II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2022-11-02

Problem 1 *Let Ω be the sample space of a random phenomenon and let $\mathcal{E}_1$, $\mathcal{E}_2$ algebras [resp. σ-algebras] of events of Ω. May we say that the family $\mathcal{E}_1 \cup \mathcal{E}_2$ of events of Ω given by*

$$\mathcal{E}_1 \cup \mathcal{E}_2 \stackrel{def}{=} \{E \in \mathcal{P}(\Omega) : E \in \mathcal{E}_1 \text{ or } E \in \mathcal{E}_2\}$$

is an algebra [resp. σ-algebra]?

Solution. Clearly, since $\mathcal{E}_1$, $\mathcal{E}_2$ algebras [resp. σ-algebras] of events of Ω, the family $\mathcal{E}_1 \cup \mathcal{E}_2$ is not empty. Now, assume that an event E is in $\mathcal{E}_1 \cup \mathcal{E}_2$, then E is in $\mathcal{E}_1$ or E is in $\mathcal{E}_2$. As a consequence, E^c is in $\mathcal{E}_1$ or E^c is in $\mathcal{E}_2$. Hence, $E^c \in \mathcal{E}_1 \cup \mathcal{E}_2$. However, assuming that E and F are in $\mathcal{E}_1 \cup \mathcal{E}_2$, unless they are both in $\mathcal{E}_1$ or $\mathcal{E}_2$, there is no reason why $E \cup F$ should be in $\mathcal{E}_1 \cup \mathcal{E}_2$. This is confirmed by the following example: with reference to the die sample space $\Omega \equiv \{\omega_1, \ldots, \omega_6\}$ choose

$$\mathcal{E}_1 \equiv \{\varnothing, \{\omega_1, \omega_3, \omega_5\}, \{\omega_2, \omega_4, \omega_6\}, \Omega\} \quad \text{and} \quad \mathcal{E}_2 \equiv \{\varnothing, \{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5, \omega_6\}, \Omega\}.$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ are algebras of events of Ω, but

$$\mathcal{E}_1 \cup \mathcal{E}_2 = \{\varnothing, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_3, \omega_5\}, \{\omega_2, \omega_4, \omega_6\} \{\omega_4, \omega_5, \omega_6\}, \Omega\}$$

is not. Note that $\mathcal{E}_1 \cup \mathcal{E}_2$ is closed with respect to the complement operator, but not with respect to the union.

Problem 2 *Let Ω be the infinite sample space of a random phenomenon. The family*

$$\mathcal{E}_{count} \equiv \{E \in \mathcal{P}(\Omega) : |E| \leq \aleph_0\}$$

of all countable events of Ω is a σ-algebra of events of Ω if and only if Ω itself is countable. In this case, we have $\mathcal{E}_{count} = \mathcal{P}(\Omega)$. On the other hand, the family $\mathcal{E}_{count\text{-}cocount}$ of all events of Ω that are countable or have countable complement, in symbols

$$\mathcal{E}_{count\text{-}cocount} \equiv \{E \in \mathcal{P}(\Omega) : |E| \leq \aleph_0 \ \vee \ |E^c| \leq \aleph_0\},$$

is a σ-algebra of events of Ω.

Solution. Given any $\omega \in \Omega$ we have

$$|\{\omega\}| = 1 \leq \aleph_0.$$

Hence, $\{\omega\} \in \mathcal{E}_{\text{count}}$. Assume that $\mathcal{E}_{\text{count}}$ is a σ-algebra, then also $\{\omega\}^c \equiv \Omega - \{\omega\}$ is in $\mathcal{E}_{\text{count}}$. By definition, it follows

$$|\Omega - \{\omega\}| \leq \aleph_0,$$

which clearly implies

$$|\Omega| \leq \aleph_0.$$

Conversely, if Ω is countable, then every $E \in \mathcal{P}(\Omega)$ is countable. This implies

$$\mathcal{P}(\Omega) \subseteq \mathcal{E}_{\text{count}},$$

that is

$$\mathcal{E}_{\text{count}} = \mathcal{P}(\Omega).$$

Trivially, $\mathcal{E}_{\text{count}}$ is σ-algebra of events of Ω. As a consequence of the above argument, if Ω is not countable, that is

$$|\Omega| > \aleph_0$$

or, according the continuum hypothesis,

$$|\Omega| \geq \aleph_1,$$

the family $\mathcal{E}_{\text{count}}$ cannot be a σ-algebra. On the other hand, the family $\mathcal{E}_{\text{count-cocount}}$ is. In fact, clearly $\mathcal{E}_{\text{count-cocount}} \neq \varnothing$. Furthermore, if $E \in \mathcal{E}_{\text{count-cocount}}$, according to the definition, we have two cases:

$$|E| \leq \aleph_0 \quad \text{or} \quad |E^c| \leq \aleph_0.$$

In the first case,

$$|(E^c)^c| = |E| \leq \aleph_0.$$

This implies that $E^c \in \mathcal{E}_{\text{count-cocount}}$. In the second case, we have $E^c \in \mathcal{E}_{\text{count-cocount}}$ straightforwardly. Hence, in either cases $E^c \in \mathcal{E}_{\text{count-cocount}}$. In the end, consider a sequence $(E_n)_{n\geq 1}$ of elements in $\mathcal{E}_{\text{count-cocount}}$. If $|E_n| \leq \aleph_0$ for every $n \in \mathbb{N}$, then

$$\left|\bigcup_{n\geq 1} E_n\right| \leq \aleph_0,$$

which implies $\bigcup_{n\geq 1} E_n \in \mathcal{E}_{\text{count-cocount}}$. Otherwise, there exists at least $n_0 \in \mathbb{N}$ such that $|E_{n_0}| > \aleph_0$. However, in this case, since $E_{n_0} \in \mathcal{E}_{\text{count-cocount}}$, we necessarily have

$$|E_{n_0}^c| \leq \aleph_0.$$

This implies

$$\left|\left(\bigcup_{n\geq 1} E_n\right)^c\right| = \left|\bigcap_{n\geq 1} E_n^c\right| \leq |E_{n_0}^c| \leq \aleph_0.$$

Thus, it still follows that $\bigcup_{n\geq 1} E_n \in \mathcal{E}_{\text{count-cocount}}$.

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$. be a probability space, where $\mathbf{P} : \mathcal{E} \to \mathbb{R}_+$ is a countably additive probability on Ω. Prove that we have*

1. *$\mathbf{P}\left(\bigcup_{k=1}^n E_k\right) = \sum_{k=1}^n \mathbf{P}(E_k)$ for any finite sequence $(E_k)_{k=1}^n$ of pairwise incompatible events in $\mathcal{E}$;*
2. *$\mathbf{P}\left(\bigcup_{n=1}^\infty E_n\right) \leq \sum_{n=1}^\infty \mathbf{P}(E_n)$ for any sequence $(E_n)_{n\geq 1}$ of events in $\mathcal{E}$.*

Solution.

1. Let $(E_k)_{k=1}^n$ be a finite sequence of pairwise incompatible events in $\mathcal{E}$. Then the sequence $(F_k)_{k\geq 1}$ given by

$$F_k \stackrel{\text{def}}{=} \begin{cases} E_k & \forall k = 1, \ldots, n \\ \varnothing & \forall k \geq n+1 \end{cases},$$

is a denumerable sequence of pairwise incompatible events in $\mathcal{E}$ such that

$$\bigcup_{k=1}^{n} E_k = \bigcup_{k=1}^{\infty} F_k$$

and

$$\mathbf{P}(F_k) = \begin{cases} \mathbf{P}(E_k) & \forall k = 1, \ldots, n \\ 0 & \forall k \geq n+1 \end{cases}.$$

As a consequence,

$$\mathbf{P}\left(\bigcup_{k=1}^{n} E_k\right) = \mathbf{P}\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mathbf{P}(F_k) = \sum_{k=1}^{n} \mathbf{P}(F_k) = \sum_{k=1}^{n} \mathbf{P}(E_k),$$

This proves that is an additive probability

2. Let $(E_n)_{n\geq 1}$ be any sequence of events in $\mathcal{E}$. Then the sequence $(F_n)_{n\geq 1}$ given by

$$F_n \stackrel{\text{def}}{=} \begin{cases} E_1 & \text{if } n = 1 \\ E_n - \bigcup\limits_{k=1}^{n-1} E_k & \text{if } n > 1 \end{cases},$$

is a sequence of pairwise incompatible events in $\mathcal{E}$ such that

$$\bigcup_{n\geq 1} E_n = \bigcup_{n\geq 1} F_n.$$

In fact, clearly $F_n \subseteq E_n$, for every $n \geq 1$. Hence,

$$\bigcup_{n\geq 1} F_n \subseteq \bigcup_{n\geq 1} E_n.$$

Conversely, if $x \in \bigcup_{n\geq 1} E_n$, then the set $N_x \equiv \{n \in \mathbb{N} : x \in E_n\} \neq \varnothing$. Write $\hat{n}_x \equiv \min(N_x)$. We have $x \in E_{\hat{n}_x}$. In case $\hat{n}_x = 1$, by definition we have $x \in F_1$. In case $\hat{n}_x > 1$, we have $x \notin E_k$ for every $k = 1, \ldots, \hat{n}_x - 1$. Hence, $x \notin \bigcup\limits_{k=1}^{\hat{n}_x - 1} E_k$ and, again by definition, $x \in F_{\hat{n}_x}$. Therefore, in any case, we obtain $x \in F_{\hat{n}_x}$. This implies that $x \in \bigcup_{n\geq 1} F_n$.
Now, given $n_1, n_2 \in \mathbb{N}$ such that $n_1 \neq n_2$, we have

$$F_{n_1} \cap F_{n_2} = \varnothing.$$

In fact, assuming for instance $n_1 < n_2$, we have

$$F_{n_1} \subseteq E_{n_1} \quad \text{and} \quad F_{n_2} \cap E_{n_1} = \left(E_{n_2} - \bigcup_{k=1}^{n_2-1} E_k\right) \cap E_{n_1} = \varnothing.$$

This implies that the events of the sequence $(F_n)_{n\geq 1}$ are pairwise incompatible. As a consequence of the above arguments, it follows

$$\mathbf{P}\left(\bigcup_{n\geq 1} E_n\right) = \mathbf{P}\left(\bigcup_{n\geq 1} F_n\right) = \sum_{n\geq 1} \mathbf{P}(F_n) \leq \sum_{n\geq 1} \mathbf{P}(E_n),$$

which is the desired result.

Problem 4 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *a probability space and let* $E, F \in \mathcal{E}$ *such that*

$$\mathbf{P}(E) + \mathbf{P}(F) \geq 1. \tag{1}$$

Prove that

$$\mathbf{P}(E) + \mathbf{P}(F) - 1 \leq \mathbf{P}(E \cap F) \leq \min\{\mathbf{P}(E), \mathbf{P}(F)\} \tag{2}$$

Determine a similar lower and upper bound for $\mathbf{P}(E \cap F)$ *under the assumption*

$$\mathbf{P}(E) + \mathbf{P}(F) < 1. \tag{3}$$

Solution. First, observe that

$$E \cap F \subseteq E \qquad \text{and} \qquad E \cap F \subseteq F. \tag{4}$$

This implies

$$\mathbf{P}(E \cap F) \subseteq \mathbf{P}(E) \qquad \text{and} \qquad \mathbf{P}(E \cap F) \subseteq \mathbf{P}(F).$$

It clearly follows

$$\mathbf{P}(E \cap F) \leq \min\{\mathbf{P}(E), \mathbf{P}(F)\}. \tag{5}$$

Second, we have

$$\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F), \tag{6}$$

which implies

$$\mathbf{P}(E \cap F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cup F). \tag{7}$$

On the other hand,

$$\mathbf{P}(E \cup F) \leq 1. \tag{8}$$

Combining (7) and (8) we then obtain

$$\mathbf{P}(E) + \mathbf{P}(F) - 1 \leq \mathbf{P}(E \cap F). \tag{9}$$

From (5) and (9) the desired (2) immedately follows. Note that in case

$$E \cap F = \varnothing \qquad \text{and} \qquad E \cup F = \Omega,$$

that is

$$F = E^c,$$

we have both

$$\mathbf{P}(E) + \mathbf{P}(F) - 1 = 0 \qquad \text{and} \qquad \mathbf{P}(E \cap F) = 0.$$

Therefore, the lower bound in (2) is achieved. Instead, in case

$$E \subseteq F \qquad \text{or} \qquad F \subseteq E,$$

we have

$$\mathbf{P}(E \cap F) = \mathbf{P}(E) \quad \text{and} \quad \min\{\mathbf{P}(E), \mathbf{P}(F)\} = \mathbf{P}(E)$$

or

$$\mathbf{P}(E \cap F) = \mathbf{P}(F) \quad \text{and} \quad \min\{\mathbf{P}(E), \mathbf{P}(F)\} = \mathbf{P}(F).$$

Hence, the upper bound in (2) is achieved. Now, under Assumption (3), combining (7) and (5) we obtain

$$0 \leq \mathbf{P}(E \cap F) \leq \min\{\mathbf{P}(E), \mathbf{P}(F)\},$$

where the lower bound is achieved when $E \cap F = \varnothing$ while the upper bound $\min\{\mathbf{P}(E), \mathbf{P}(F)\}$ is still achieved when $E \subseteq F$ or $F \subseteq E$. On the other hand, still under Assumption (3), from Equation (7) it follows

$$\mathbf{P}(E \cap F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cup F) < 1 - \mathbf{P}(E \cup F). \tag{10}$$

Therefore, one wonders whether a sharper bound for $\mathbf{P}(E \cap F)$ is given by $\min\{\mathbf{P}(E), \mathbf{P}(F)\}$ or $1 - \mathbf{P}(E \cup F)$. Considering the possibility

$$\min\{\mathbf{P}(E), \mathbf{P}(F)\} < 1 - \mathbf{P}(E \cup F), \tag{11}$$

in case $F = E^c$, we obtain

$$\mathbf{P}(E \cup F) = 1.$$

Thus, Equation (11) would imply

$$\min\{\mathbf{P}(E), \mathbf{P}(F)\} < 0.$$

which is clearly false. However, considering the possibility

$$1 - \mathbf{P}(E \cup F) < \min\{\mathbf{P}(E), \mathbf{P}(F)\}, \tag{12}$$

in case $E \subseteq F$, we obtain

$$1 - \mathbf{P}(E \cup F) = 1 - \mathbf{P}(F)$$

and

$$\min\{\mathbf{P}(E), \mathbf{P}(F)\} = \mathbf{P}(E).$$

Hence, from (12) we would obtain

$$1 - \mathbf{P}(F) < \mathbf{P}(E),$$

that is

$$1 < \mathbf{P}(E) + \mathbf{P}(F),$$

which is also clearly false. As a consequence, under Assumption (3), both $\min\{\mathbf{P}(E), \mathbf{P}(F)\}$ and $1 - \mathbf{P}(E \cup F)$ are superior bounds for $\mathbf{P}(E \cap F)$. This implies that a sharper superior bound is given by $\min\{\mathbf{P}(E), \mathbf{P}(F), 1 - \mathbf{P}(E \cup F)\}$. Summarizing, under Assumption (3), we have

$$0 \leq \mathbf{P}(E \cap F) \leq \min\{\mathbf{P}(E), \mathbf{P}(F), 1 - \mathbf{P}(E \cup F)\}.$$

Problem 5 *Five Italian players are playing poker. The deck of poker cards contains 36 cards of the usual ranks $(6, 7, 8, 9, 10, J, Q, K, A)$ and of the usual suites (hearts ♡, clubs ♠, diamonds ♢, flowers ♣).*

1. *How many hands are possible by a random deal?*
2. *How many hands give a straight flush by a random deal?*
3. *How many hands give a four of a kind by a random deal?*
4. *How many hands give a flush by a random deal?*
5. *How many hands give a full house by a random deal?*

a p.48 è fatto meglio, con qualche punto in più

6. *How many hands give a straight by a random deal?*
7. *How many hands give a three of a kind by a random deal?*
8. *How many hands give two pair by a random deal?*
9. *How many hands give one pair by a random deal?*
10. *How many hands give no pair by a random deal?*
11. *How many hands fail to give any of the above combinations by a random deal?*
12. *What about if the players are Americans? In this case the deck of poker card contains 56 cards of the usual ranks $(1, \ldots, 10, J, Q, K, A)$ and of the usual suites.*

Solution.

1. Since a poker hand is indifferent to the order in which is arranged by the deal, the number of all possible hands is just the number of all possible subsets of 5 elements that can be selected from a set of 36 elements. Hence,
$$\binom{36}{5}$$
is the number of all possible hands.

2. According to the (Italian) poker rules, there are 6 possibilities for choosing the rank of the first card of a straight. The ranks of the other cards are then consequently determined. That is the ony possible straights in a deck of 36 cards are
$$A, 6, 7, 8, 9; \quad 6, 7, 8, 9, 10; \quad ; \ldots ; \quad 10, J, Q, K, A.$$
We have a straight flush when the cards of the straight have all the same suits. We can choose the suit for the straight flush in 4 differeent ways. Hence, we have
$$4 \cdot 6$$
possible hands giving a straight flush by a random deal.

3. There are 9 possibility for choosing the rank of card for the four of a kind, once the card has been chosen there is no room for the choice of the suites. Then, there are 8 possibilities for choosing the rank of fifth card and for each rank there are $\binom{4}{1} = 4$ possibilities for choosing the suits. As a consequence, we have
$$9 \cdot 8 \cdot 4$$
possible hands giving a four of a kind by a random deal.

4. There are $\binom{4}{1} = 4$ possibilities for choosing the suits of the flush and $\binom{9}{5} - 6$ possibilities for choosing the rank of the cards in the flush discarding the 6 possible straights (which is necessary to discard to avoid straight flushes). Therefore, we have
$$4 \cdot \left(\binom{9}{5} - 6\right)$$
possible hands giving a flush by a random deal.

5. There are 9 possibilities for choosing the rank of the card for the three of a kind in a full house and, for each choice of this rank, we have $\binom{4}{3}$ possible choices for the suites. Hence, we have 8 possibilities for choosing the rank of the card for the pair in a full house and, for each choice of this rank, we have $\binom{4}{2}$ possible choices for the suites. In the end, we have

$$9 \cdot \binom{4}{3} \cdot 8 \cdot \binom{4}{2}$$

possible hands giving a full house by a random deal.

6. As discusse above, following the (Italian) poker rules, there are 6 possibilities for choosing the rank of the first card of a straight. The ranks of the other cards are then consequently determined. Then choosing for each card of the straight one of the possible suites we build all possible straights, including the straight flushes, which have to be discarded. Therefore. we have

$$6 \cdot 4^5 - 24$$

possible hands giving a straight wich is not a straight flush by a random deal.

7. There are 9 possibilities for choosing the rank of the card for a three of a kind and for each choice of this rank, we have $\binom{4}{3}$ possible choices for the suites of the cards constituting the three of a kind. Now, we need to choose two additional ranks for the two remaining cards which should not constitute a pair. This can be done in $\binom{8}{2}$ ways, and we can choose the suits of each remaining card in 4 ways. Therefore there are

$$9 \cdot \binom{4}{3} \cdot \binom{8}{2} \cdot 4^2$$

possible hands giving a three of a kind by a random deal.

8. There are $\binom{9}{2}$ possibilities for choosing the rank of the cards for the two pairs and for each choice of these ranks we have $\binom{4}{2}$ possible choices for the suites. The rank of the last card of the hand can be chosen among the remaining 7 ranks and the suits of the last card can be chosen in 4 ways. Hence, we have

$$\binom{9}{2} \cdot \binom{4}{2}^2 \cdot 7 \cdot 4$$

possible hands giving two pair by a random deal.

9. There are 9 possibilities for choosing the rank of the cards for the pair and for each choice of this rank we have $\binom{4}{2}$ possible choices for the suites. Then the rank of the remaining three cards have to be chosen to be different, which can be done in $\binom{8}{3}$ ways, and the suits of each of the remaining three cards can be chosen in 4 different. Therefore, we have

$$9 \cdot \binom{4}{2} \cdot \binom{8}{3} \cdot 4^3$$

possible hands giving a pair by a random deal.

Problem 6 *An urn contains N distinguishable balls of which M are white, with $1 \leq M < N$, and the remaining $N - M$ are black. The urn is shaken and the balls are drawn from the urn one after the other without replacement. How many of the possible drawn sequences show the first white ball at the kth draw?*

Ans.

$$\binom{N-M}{k-1}(k-1)!M(N-k)!$$

Solution. Since the balls are distinguishable it is convenient to think in terms of permutations. In fact, we can assume that the white balls are numbered from 1 to M and the white balls are numbered from $M+1$ to N. If we want the first white ball at the kth attempt, we have to start drawing $k-1$ black balls, and this can be done in $\binom{N-M}{k-1}$ ways. The chosen balls can be arranged in $(k-1)!$ different orders. The red ball at the kth place can be choosen in M different ways and the remaining $N-k$ balls can be arranged in $(n-k)!$ different orders.

Problem 7 *An urn contains N balls of which M are white, with $1 \leq M < N$, and the remaining $N - M$ are black. The urn is shaken and the balls are drawn one after the other without replacement. Suppose that both the white balls and the black ones are undistinguishable among them. How many of the possible drawn sequences show the first white ball at the kth draw?*

Ans.

$$\binom{N-k}{M-1}$$

Solution. The drawn sequences which contain the first white ball at the kth draw are distinguishable just for the position of the other $M-1$ white balls after the kth draw. So each of them identifies with the choice of $M-1$ places from the $N-k$ available ones.

Problem 8 *An urn contains N distinguishable balls of which M are white, with $1 \leq M < N$, and the remaining $N - M$ are black. The urn is shaken and k balls are drawn without replacement. If $k \leq M$, how many of the possible unordered samples contains $j \leq k$ white balls and $k - j$ black balls?*

Ans.

$$\binom{M}{j}\binom{N-M}{k-j}.$$

Solution. .

Problem 9 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $E, F \in \mathcal{E}$. Show that E and F are independent if and only if:*

1. *E^c and F^c are independent;*
2. *E and F^c are independent;*
3. *E^c and F are independent.*

Solution. We need to prove that

$$\begin{aligned}
&\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F) \\
&\quad \Leftrightarrow \mathbf{P}(E^c \cap F^c) = \mathbf{P}(E^c)\mathbf{P}(F^c) \\
&\quad \Leftrightarrow \mathbf{P}(E \cap F^c) = \mathbf{P}(E)\mathbf{P}(F^c) \\
&\quad \Leftrightarrow \mathbf{P}(E^c \cap F) = \mathbf{P}(E^c)\mathbf{P}(F).
\end{aligned}$$

First, assume that $\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$ and consider $\mathbf{P}(E^c \cap F^c)$. On account of the properties of the probability function, we can write

$$\begin{aligned}\mathbf{P}(E^c \cap F^c) &= \mathbf{P}((E \cup F)^c) \\ &= 1 - \mathbf{P}(E \cup F) \\ &= 1 - (\mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)) \\ &= 1 - (\mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E)\mathbf{P}(F)) \\ &= 1 - \mathbf{P}(E) - \mathbf{P}(F)(1 - \mathbf{P}(E)) \\ &= (1 - \mathbf{P}(E))(1 - \mathbf{P}(F)) \\ &= \mathbf{P}(E^c)\mathbf{P}(F^c).\end{aligned}$$

Second, assume that $\mathbf{P}(E^c \cap F^c) = \mathbf{P}(E^c)\mathbf{P}(F^c)$ and consider $\mathbf{P}(E \cap F^c)$. We can write

$$\mathbf{P}(F^c) = \mathbf{P}(\Omega \cap F^c) = \mathbf{P}((E \cup E^c) \cap F^c) = \mathbf{P}((E \cap F^c) \cup (E^c \cap F^c)) = \mathbf{P}(E \cap F^c) + \mathbf{P}(E^c \cap F^c).$$

Therefore,

$$\mathbf{P}(F^c) = \mathbf{P}(E \cap F^c) + \mathbf{P}(E^c)\mathbf{P}(F^c).$$

It then follows

$$\mathbf{P}(E \cap F^c) = (1 - \mathbf{P}(E^c))\mathbf{P}(F^c) = \mathbf{P}(E)\mathbf{P}(F^c).$$

Third, assume that $\mathbf{P}(E \cap F^c) = \mathbf{P}(E)\mathbf{P}(F^c)$ and consider $\mathbf{P}(E^c \cap F)$. We can write

$$\begin{aligned}\mathbf{P}(E^c \cap F) &= 1 - \mathbf{P}((E^c \cap F)^c) = 1 - \mathbf{P}(E \cup F^c) = 1 - (\mathbf{P}(E) + \mathbf{P}(F^c) - \mathbf{P}(E \cap F^c)) \\ &= 1 - (\mathbf{P}(E) + \mathbf{P}(F^c) - \mathbf{P}(E)\mathbf{P}(F^c)) = 1 - (\mathbf{P}(E)(1 - \mathbf{P}(F^c)) + \mathbf{P}(F^c)) \\ &= 1 - \mathbf{P}(E)(1 - \mathbf{P}(F^c)) - \mathbf{P}(F^c) = (1 - \mathbf{P}(E))(1 - \mathbf{P}(F^c)) = \mathbf{P}(E^c)\mathbf{P}((F^c)^c) \\ &= \mathbf{P}(E^c)\mathbf{P}(F).\end{aligned}$$

Fourth, assume that $\mathbf{P}(E^c \cap F) = \mathbf{P}(E^c)\mathbf{P}(F)$ and consider $\mathbf{P}(E \cap F)$. We can write

$$\begin{aligned}\mathbf{P}(F) &= \mathbf{P}(F \cap \Omega) = \mathbf{P}(F \cap (E \cup E^c)) = \mathbf{P}((E \cap F) \cup (E^c \cap F)) = \mathbf{P}(E \cap F) + \mathbf{P}(E^c \cap F) \\ &= \mathbf{P}(E \cap F) + \mathbf{P}(E^c)\mathbf{P}(F).\end{aligned}$$

Hence,

$$\mathbf{P}(F) = \mathbf{P}(E \cap F) + \mathbf{P}(E^c)\mathbf{P}(F).$$

As a consequence,

$$\mathbf{P}(E \cap F) = (1 - \mathbf{P}(E^c))\mathbf{P}(F) = \mathbf{P}((E^c)^c)\mathbf{P}(F) = \mathbf{P}(E)\mathbf{P}(F).$$

This case closes the chain of implications which allow us to conclude that the desired equivalences hold true. □

Problem 10 *A urn, say Urn A, contains 5 white balls and 10 black balls. Another urn, say Urn B, contains 3 white balls and 12 black balls. A fair coin is tossed. If it shows heads [resp. tails] a ball is drawn from Urn A [resp. B]. Suppose that this random experiment has been done and we know that a white ball has been drawn. What is the probability that the ball has been drawn from Urn A? What is the probability that the ball has been drawn from Urn B?*

Solution. .

H = testa, T = croce, P(W | H) = 5/15, P(B | H) = 10/15
P(W | T) = 3/15 , P(B | T) = 4/5 P(H) = P(T) = 1/2
P(W) = P(W | H)*P(H) + P(W | T)*P(T) = 4/15
allora P(H | W) = P(W | H)* P(H) / P(W) = 5/8 allora
P(T | W) = 3/8

Problem 11 *In a large town, after a robbery, a thief jumped into a cab and disappeared. An eyewitness on the crime scene told the police that the cab was yellow. Having some doubt on the reliability of the eyewitness, the police consulted a mathematician. Assuming that*

1. *20% of the cabs in the town are yellow;*
2. *from the past experience police knows that an eyewitness is 80% accurate, that is an eyewitness identifies correctly whather the colour of a taxi is yellow or not 8 out 10 times.*

   Compute the probability that the information reported by the eyewitness is true. That is the probability that the cab was yellow given that the eyewitness said so. Do you think this information is useful?

   Hint: consider the events "the cab was yellow", "the cab was not yellow", "the eyewitness says the cab was yellow", and "the eyewitness says the cab was not yellow" and formulate in terms of conditional probability the accuracy of the eyewitness.

Solution. Consider the events "the cab was yellow" and "the cab was not yellow", which we may denote by $\{C = 1\}$ and $\{C = 0\}$, respectively. We have

$$\mathbf{P}(C = 1) = 0.2 \quad \text{and} \quad \mathbf{P}(C = 0) = 0.8.$$

Hence consider the events "the eyewitness says the cab was yellow" and "the eyewitness says the cab was not yellow", which we may denote by $\{E = 1\}$ and $\{E = 0\}$, respectively. The 80% of accuracy of the eyewitness means that

$$\mathbf{P}(E = 1 \mid C = 1) = \mathbf{P}(E = 0 \mid C = 0) = 0.8.$$

Our goal is to compute

$$\mathbf{P}(C = 1 \mid E = 1).$$

By the symmetry formula, we have

$$\mathbf{P}(C = 1 \mid E = 1) = \frac{\mathbf{P}(E = 1 \mid C = 1)\mathbf{P}(C = 1)}{\mathbf{P}(E = 1)}.$$

Now, on account of the total probability form, we have

$$\begin{aligned}\mathbf{P}(E = 1) &= \mathbf{P}(E = 1 \mid C = 1)\mathbf{P}(C = 1) + \mathbf{P}(E = 1 \mid C = 0)\mathbf{P}(C = 1) \\ &= \mathbf{P}(E = 1 \mid C = 1)\mathbf{P}(C = 1) + (1 - \mathbf{P}(E = 0 \mid C = 0))\mathbf{P}(C = 0) \\ &= 0.8 \cdot 0.2 + (1 - 0.8) \cdot 0.8 \\ &= 0.32\end{aligned}$$

In the end, we obtain

$$\mathbf{P}(C = 1 \mid E = 1) = \frac{0.8 \cdot 0.2}{0.32} = 0.5.$$

This means that the information provided by the eyewitness is not that useful.

Problem 12 *The National Health Service (NHS) aims to introduce a new test for the screening of a disease. The pharmaceutical company which produces the test states that:*

- *the test yields a positive result on the 95% of people who are affected by the disease (*sensitivity *or* true positive rate *of the test);*
- *the test yields a negative result on the 99% of people who are not affected by the disease (*specificity *or* true negative rate *of the test);*

On the other hand, the NHS knows the the disease is currently affecting the 10% of the population. Compute:

1. *the probability that a randomly chosen individual of the population is affected by the disease given that the test yields a positive result;*
2. *the probability that a randomly chosen individual of the population is not affected by the disease given that the test yields a positive result;*
3. *the probability that a randomly chosen individual of the population is affected by the disease given that the test yields a negative result;*
4. *the probability that a randomly chosen individual of the population is not affected by the disease given that the test yields a negative result;*
5. *the probability that the test yields a positive result on a randomly chosen individual of the population.*
6. *the probability that the test yields a negative result on a randomly chosen individual of the population.*

Solution. Write D [resp. H] for the event "a randomly chosen individual of the population is affected [resp. not affected] by the disease". We have

$$D \cup H = \Omega \quad \text{and} \quad D \cap H = \varnothing.$$

It follows

$$\mathbf{P}(D) + \mathbf{P}(H) = 1. \tag{13}$$

Write T_+ [resp. T_-] for the event "the test yields a positive [resp. negative] result on a randomly chosen individual of the population". We have

$$T_+ \cup T_- = \Omega \quad \text{and} \quad T_+ \cap T_- = \varnothing.$$

It follows

$$\mathbf{P}(T_+) + \mathbf{P}(T_-) = 1. \tag{14}$$

In terms of the above notations, the information provided by the pharmaceutical company means

$$\mathbf{P}(T_+ \mid D) = 0.95, \qquad \text{and} \qquad \mathbf{P}(T_- \mid H) = 0.99. \tag{15}$$

The information provided by NHS is

$$\mathbf{P}(D) = 0.10, \tag{16}$$

which clearly implies

$$\mathbf{P}(H) = 0.90. \tag{17}$$

To answer Questions **??-??**, we need to compute the following probabilities

$$\mathbf{P}(D \mid T_+), \quad \mathbf{P}(H \mid T_+), \quad \mathbf{P}(D \mid T_-), \quad \mathbf{P}(H \mid T_-), \quad \mathbf{P}(T_+), \quad \mathbf{P}(T_-),$$

respectively. Now, from the symmetry formula of conditional probabilities and on account of (15)-(17), we know that

$$\mathbf{P}(D \mid T_+) = \frac{\mathbf{P}(T_+ \mid D)\,\mathbf{P}(D)}{\mathbf{P}(T_+)} = \frac{0.95 * 0.10}{\mathbf{P}(T_+)} = \frac{0.095}{\mathbf{P}(T_+)}, \tag{18}$$

$$\mathbf{P}(H \mid T_+) = \frac{\mathbf{P}(T_+ \mid H)\,\mathbf{P}(H)}{\mathbf{P}(T_+)} = \frac{0.90 * \mathbf{P}(T_+ \mid H)}{\mathbf{P}(T_+)}, \tag{19}$$

$$\mathbf{P}(D \mid T_-) = \frac{\mathbf{P}(T_- \mid D)\,\mathbf{P}(D)}{\mathbf{P}(T_-)} = \frac{0.10 * \mathbf{P}(T_- \mid D)}{\mathbf{P}(T_-)}, \tag{20}$$

$$\mathbf{P}(H \mid T_-) = \frac{\mathbf{P}(T_- \mid H)\,\mathbf{P}(H)}{\mathbf{P}(T_-)} = \frac{0.99 * 0.90}{\mathbf{P}(T_-)} = \frac{0.891}{\mathbf{P}(T_-)}, \tag{21}$$

On the other hand, we know that the conditional probability is a probability concentrated on the conditioning event. Hence, we have

$$\mathbf{P}(T_+ \mid H) = 1 - \mathbf{P}(T_- \mid H) = 1 - 0.99 = 0.01. \tag{22}$$

and

$$\mathbf{P}(T_- \mid D) = 1 - \mathbf{P}(T_+ \mid D) = 1 - 0.95 = 0.05 \tag{23}$$

Therefore, on account of (14), to answer Questions **??-??**, we are left with computing $\mathbf{P}(T_+)$. To this, by the Total Probability Formula, we can write

$$\begin{aligned}\mathbf{P}(T_+) &= \mathbf{P}(T_+ \mid H)\,\mathbf{P}(H) + \mathbf{P}(T_+ \mid D)\,\mathbf{P}(D) \\ &= 0.01 * 0.90 + 0.95 * 0.10 = 0.104.\end{aligned} \tag{24}$$

As a consequence,

$$\mathbf{P}(T_-) = 1 - 0.104 = 0.896. \tag{25}$$

In the end, replacing (22)-(25) into (18)-(21), we obtain

$$\begin{aligned}\mathbf{P}(D \mid T_+) &= \frac{0.095}{0.104} = 0.913, \\ \mathbf{P}(H \mid T_+) &= \frac{0.90 * 0.01}{0.104} = 8.654 \times 10^{-2}, \\ \mathbf{P}(D \mid T_-) &= \frac{0.10 * 0.05}{0.896} = 5.58 \times 10^{-3}, \\ \mathbf{P}(H \mid T_-) &= \frac{0.891}{0.896} = 0.994,\end{aligned}$$

which complete the answers.

Problem 13 *Machineries* 1, 2 *and* 3 *contribute to the total production of an industry with the percentages of* 50%, 30% *and* 20% *respectively. The percentages of defective items produced by the machineries are* 2%, 4% *and* 5% *respectively. Compute the naive probability that a randomly chosen item is defective. Compute also the naive probability that a randomly chosen item has been produced by the machinery* 1*, provided it is defective.*

Solution. .

Problem 14 *Guns A and B are shooting the same target. Gun A shoots on the average 9 shots during the same time that gun B shoots 10 shots. On the average, out of 10 shots from gun A only 8 hit the target, and from gun B, only 7. During the shooting the target has been hit by a bullet. Compute the naive probability that the target was hit by gun A or B. Compute also the probability of hitting the target.*

Solution. .

Problem 15 (Monty's hall) *A quiz master shows a participant three closed boxes labeled by A, B, C. One of the boxes, chosen by the quiz organisers with uniform probability, contains a prize of* $1,000. *The remaining two are empty. The quiz master asks the participant to choose a box. Once the participant makes own choice the quiz master opens one of the two rejected boxes and shows that it is empty. Thereafter, the quiz master gives the participant the opportunity either to stick to the first choice or to exchange the chosen box with the one of the rejected boxes which is still closed. Assume that the quiz master knows what box contains the prize, the quiz master never shows a box containing the prize, and the quiz master chooses an empty box between two with uniform probability. What should the participant do? To stick to her first choice, to accept the exchange or it does not matter at all because the odds are now fifty-fifty?*

Solution. Assume the quiz participant chooses box A and thereafter the quiz master shows the empty box B or C. Denote by PX the event "the prize is in box X" for $X = A, B, C$ and EY the event "the quiz master shows empty box Y", for $Y = B, C$. Under the assumptions of the problem, we have

$$\mathbf{P}(PA) = \mathbf{P}(PB) = \mathbf{P}(PC) = \frac{1}{3}.$$

Now, since the family of events $\{PA, PB, PC\}$ is a partition of the sure event Ω, by virtue of the Total Probability Formula, we have

$$\mathbf{P}(EB) = \mathbf{P}(EB \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EB \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EB \mid PC)\,\mathbf{P}(PC) \tag{26}$$

and

$$\mathbf{P}(EC) = \mathbf{P}(EC \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EC \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EC \mid PC)\,\mathbf{P}(PC) \tag{27}$$

In addition, still the assumptions of the problem yield

$$\begin{array}{lll} \mathbf{P}(EB \mid PA) = \frac{1}{2}, & \mathbf{P}(EB \mid PB) = 0, & \mathbf{P}(EB \mid PC) = 1, \\ \mathbf{P}(EC \mid PA) = \frac{1}{2}, & \mathbf{P}(EC \mid PB) = 1, & \mathbf{P}(EC \mid PC) = 0. \end{array} \tag{28}$$

Combining (26), (27) and (28), we obtain the somewhat obvious result

$$\mathbf{P}(EB) = \mathbf{P}(EC) = \frac{1}{2}. \tag{29}$$

By (**??**), it Then, follows

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}, \tag{30}$$

and

$$\mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}. \tag{31}$$

Hence, whether the quiz master shows empty box B or C the conditional probability that the prize is in the other box C or B is higher than the conditional probability that the prize is contained in the initially chosen box A. We can conclude that the quiz participant should exchange the initially chosen box A with closed box C or B.

Another argument is the following. Assume the participant chooses box A. Denote by V the event "the participant wins", by PA the event "the prize is in box A" and by PA^c the event "the prize is not in box A". We have clearly

$$\mathbf{P}(PA) = \frac{1}{3}, \quad \mathbf{P}(PA^c) = \frac{2}{3}.$$

The total probability formula yields

$$\mathbf{P}(V) = \mathbf{P}(V \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(V \mid PA^c)\,\mathbf{P}(PA^c) = \frac{1}{3}\mathbf{P}(V \mid PA) + \frac{2}{3}\mathbf{P}(V \mid PA^c). \tag{32}$$

Now, if the participant sticks to the choice of box A we have

$$\mathbf{P}(V \mid PA) = 1, \quad \mathbf{P}(V \mid PA^c) = 0,$$

which implies

$$\mathbf{P}(V) = \frac{1}{3}.$$

In contrast, if the participant exchanges the chosen box with the still-closed rejected box we have

$$\mathbf{P}(V \mid PA) = 0, \quad \mathbf{P}(V \mid PA^c) = 1,$$

which implies

$$\mathbf{P}(V) = \frac{2}{3}.$$

This confirms that the participant should exchange chosen box A with closed box B or C.

Problem 16 (Monty's Hall Strikes Back) *Consider Monty's Hall problem. Still assume that the quiz master knows what box contains the prize and the quiz master never shows a box containing the prize. However, in this case assume that after watching the game many times you notice that when a quiz participant chooses box A the quiz master shows empty box B [resp. C] the* 60% *[resp.* 40%*] of the times. May this information improve the quiz participant's strategy for winning the prize? What about if you notice that when a quiz participant chooses box A the quiz master shows empty box B [resp. C] the* 75% *[resp.* 25%*] of the time?*

Solution. .

Problem 17 (The Return of Monty's Hall) *Consider Monty's Hall problem. Still assume that the quiz master knows what box contains the prize, the quiz master never shows a box containing the prize, and the quiz master chooses an empty box between two with uniform probability. However, in this case assume that after watching the game many times you notice that the prize turns out to be in box A [resp. B] for* 45% *[resp.* 30%*] of the time and in box C the rest of the time. What is the quiz participant's best strategy?*

Solution. .

Problem 18 (Monty's Hall Awakens) *Consider Monty's Hall problem again. Still, assume that the quiz master knows what box contains the prize and that the quiz master never shows a box containing the prize. Assume also that after watching the game many times, you notice that the prize is in box A [resp. B] for* 50% *[resp.* 30%*] of the time and in box C the rest of the time. Moreover, assume that after a quiz participant chooses a box, the quiz master chooses a box to show between two empty boxes, by flipping a rigged coin with success probability p (the participant does not see the flip of the coins). What is the quiz participant's best strategy?*

tutti i monty hall sono fatti a p.52

Solution. Retaining the notation of Example 15, there are only two differences between this episode of Monty's Hall saga and Episode 17. The differences are that, from your observations, the organizers do not select the box in which to put the price by the uniform distribution, but according to the distribution

$$\mathbf{P}(PA) = 0.50 = \frac{1}{2}, \qquad \mathbf{P}(PB) = 0.30 = \frac{3}{10}, \qquad \mathbf{P}(PC) = 0.20 = \frac{1}{5},$$

and that the quiz master chooses a box between two which are empty, not by the uniform distribution, but by tossing a rigged coin. Assuming that in the favorable (heads) result of the flip, the quiz master chooses the box which comes first in the lexicografic ordere, the latter implies

$$\begin{aligned} &\mathbf{P}(EB \mid PA) = p, \quad \mathbf{P}(EC \mid PA) = 1-p, \\ &\mathbf{P}(EA \mid PB) = p, \quad \mathbf{P}(EC \mid PB) = 1-p, \\ &\mathbf{P}(EA \mid PC) = p, \quad \mathbf{P}(EB \mid PC) = 1-p, \end{aligned}$$

Clearly, we still have

$$\mathbf{P}(EA \mid PA) = \mathbf{P}(EB \mid PB) = \mathbf{P}(EC \mid PC) = 0$$

As a consequence, assuming that the quiz participant's first choice is A box. Then, we also have

$$\mathbf{P}(EB \mid PC) = \mathbf{P}(EC \mid PB) = 1$$

Thus, we end up with evaluating

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{p \cdot \frac{1}{2}}{\mathbf{P}(EB)}, \quad \mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB \mid PC)\,\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{(1-p) \cdot \frac{1}{5}}{\mathbf{P}(EB)},$$

and

$$\mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{(1-p) \cdot \frac{1}{2}}{\mathbf{P}(EC)}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC \mid PB)\,\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{(1-p) \cdot \frac{3}{10}}{\mathbf{P}(EC)}.$$

Therefore, if the quiz master shows the empty box C, the participant should stick to her first choice, A. But, if the quiz master shows the empty box B, the participant should stick to her first choice, A for $p > 2/7$, exchange her first choice with C for $p < 2/7$, and the exchange would be irrelevant for $p = 2/7$.

Note that the explicit computation of $\mathbf{P}(EB)$ and $\mathbf{P}(EC)$ is not necessary to determine the quiz participant's best strategy. However, by the total probability formula,we have

$$\mathbf{P}(EB) = \mathbf{P}(EB \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EB \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EB \mid PC)\,\mathbf{P}(PC) = p \cdot \frac{1}{2} + \frac{1}{5},$$

and

$$\mathbf{P}(EC) = \mathbf{P}(EC \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EC \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EC \mid PC)\,\mathbf{P}(PC) = (1-p) \cdot \frac{1}{2} + \frac{3}{10}.$$

It follows,

$$\mathbf{P}(PA \mid EB) = \frac{\frac{1}{2}p}{\frac{1}{2}p + \frac{1}{5}}, \qquad \mathbf{P}(PC \mid EB) = \frac{\frac{1}{5}(1-p)}{\frac{1}{2}p + \frac{1}{5}},$$

and

$$\mathbf{P}(PA \mid EC) = \frac{\frac{1}{2}(1-p)}{\frac{1}{2}(1-p) + \frac{3}{10}}, \qquad \mathbf{P}(PB \mid EC) = \frac{\frac{3}{10}(1-p)}{\frac{1}{2}(1-p) + \frac{3}{10}},$$

which quantify the probability of finding the prize in the different boxes corresponding to the quiz master's behavior.

Assume the quiz participant's first choice is B box. Then, we also have

$$\mathbf{P}(EA \mid PC) = \mathbf{P}(EC \mid PA) = 1.$$

In this case, we end up with evaluating

$$\mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA \mid PB)\,\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{p \cdot \frac{3}{10}}{\mathbf{P}(EA)}, \quad \mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA \mid PC)\,\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{p \cdot \frac{1}{5}}{\mathbf{P}(EA)},$$

and

$$\mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC \mid PB)\,\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{(1-p) \cdot \frac{3}{10}}{\mathbf{P}(EC)}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{(1-p) \cdot \frac{1}{2}}{\mathbf{P}(EC)}.$$

Therefore, if the quiz master shows the empty box A, the participant should always stick to her first choice B, but if the quiz master shows the empty box C, the participant should always exchange her first choice B with the box A.

In the end, assume the quiz participant's first choice is C box. Then, we also have

$$\mathbf{P}(EA \mid PB) = \mathbf{P}(EB \mid PA) = 1.$$

In this case, we end up with evaluating

$$\mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA \mid PC)\,\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{p \cdot \frac{1}{5}}{\mathbf{P}(EA)}, \quad \mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA \mid PB)\,\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{p \cdot \frac{3}{10}}{\mathbf{P}(EA)},$$

and

$$\mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB \mid PC)\,\mathbf{P}(PB)}{\mathbf{P}(EB)} = \frac{(1-p) \cdot \frac{3}{10}}{\mathbf{P}(EB)}, \quad \mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{p \cdot \frac{1}{2}}{\mathbf{P}(EB)}.$$

Hence, if the quiz master shows the empty box A, the participant should always exchange her first choice C with box B. But, if the quiz master shows the empty box B, then the quiz participant should stick to her first choice C for $p < 3/8$, exchange the box C with A for $p > 3/8$, and the exchange would be irrelevant for $p = 3/8$.

Problem 19 (The Last Monty's Hall) *Consider Monty's Hall problem. However, in this case assume that the quiz master does not know what box contains the prize, the quiz master chooses a box between two with uniform probability, and the quiz master shows an empty box by chance. In this episode of the Monty's Hall saga, what should the participant do? To stick to her first choice, to accept the exchange or it does not matter at all because the odds are now fifty-fifty?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Distribution Functions 2022-11-09

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $F : \mathbb{R} \to \mathbb{R}_+$ given by*

$$F(x) \stackrel{def}{=} ae^x 1_{\mathbb{R}_{--}}(x) - \left(\tfrac{1}{2}e^{-x} - b\right) 1_{\mathbb{R}_+}(x), \quad \forall x \in \mathbb{R},$$

where $a, b \in \mathbb{R}$.

1. *Can you determine $a, b \in \mathbb{R}$ such that $F : \mathbb{R} \to \mathbb{R}_+$ is a distribution function of a random variable $X : \Omega \to \mathbb{R}$.*

2. *Is it possible to determine $a, b \in \mathbb{R}$ such that $X : \Omega \to \mathbb{R}$ is absolutely continuous? In this case, compute $\mathbf{P}(-1 \le X \le 1)$.*

Solution. We have

$$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} ae^x = 0$$

and

$$\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty} b - \tfrac{1}{2}e^{-x} = b.$$

Therefore, to make $F : \mathbb{R} \to \mathbb{R}_+$ a distribution function we need

$$b = 1.$$

Moreover, we have

$$\lim_{x\to0^-} F(x) = \lim_{x\to0^-} ae^x = a \qquad \text{and} \qquad \lim_{x\to0^+} F(x) = \lim_{x\to0^+} \left(1 - \tfrac{1}{2}e^{-x}\right) = \frac{1}{2}.$$

To make $F : \mathbb{R} \to \mathbb{R}_+$ a distribution function we also need

$$a \le \frac{1}{2}.$$

Under these conditions we have

$$F(x) = ae^x 1_{\mathbb{R}_{--}}(x) + \left(1 - \tfrac{1}{2}e^{-x}\right) 1_{\mathbb{R}_+}(x),$$

for every $x \in \mathbb{R}$. The latter is an increasing function on $\mathbb{R}$ which turns out to be a distribution function of a random variable $X : \Omega \to \mathbb{R}$. To make $X : \Omega \to \mathbb{R}$ absolutely continuous, we need that $F : \mathbb{R} \to \mathbb{R}_+$ is absolutely continuous. In particular, we need that $F : \mathbb{R} \to \mathbb{R}_+$ is continuous. Hence, we need

$$a = \lim_{x\to0^-} F(x) = \lim_{x\to0^+} F(x) = \frac{1}{2}.$$

Now, the function

$$F(x) \stackrel{\text{def}}{=} \tfrac{1}{2}e^x 1_{\mathbb{R}_{--}}(x) + \left(1 - \tfrac{1}{2}e^{-x}\right) 1_{\mathbb{R}_+}(x), \quad \forall x \in \mathbb{R},$$

is a continuous function on $\mathbb{R}$, which is differentiable everywhere in $\mathbb{R} - \{0\}$ with derivative

$$F'(x) = \tfrac{1}{2}e^x 1_{\mathbb{R}_{--}}(x) + \tfrac{1}{2}e^{-x} 1_{\mathbb{R}_{++}}(x),$$

for every $x \in \mathbb{R} - \{0\}$. Now, we have

$$\lim_{x\to0^-} F'(x) = \tfrac{1}{2} = \lim_{x\to0^+} F'(x).$$

It follows that F is actually differentiable everywhere in $\mathbb{R}$ with derivative

$$F'(x) = \tfrac{1}{2}e^x 1_{\mathbb{R}_{--}}(x) + \tfrac{1}{2}e^{-x} 1_{\mathbb{R}_+}(x),$$

or every $x \in \mathbb{R}$. Such derivative is clealy bounded. It then follows that $F : \mathbb{R} \to \mathbb{R}_+$ is absolutely continuous.

is a Lebesgue integrable function and we have

$$F(x) = \int_{(-\infty,x]}^{\prime} \tilde{F}'(u)\, d\mu_L(u), \quad \forall x \in \mathbb{R}.$$

thanks to the positivity of both the functions $e^x 1_{\mathbb{R}_{--}}(x)$ and $e^{-x} 1_{\mathbb{R}_{++}}(x)$ on varying of $x \in \mathbb{R}$, we can write

$$\begin{aligned}
\int_{(-\infty,x]} \tilde{F}'(u)\, d\mu_L(u) &= \int_{(-\infty,x]} \left(\tfrac{1}{2}e^u 1_{\mathbb{R}_{--}}(u) + \tfrac{1}{2}e^{-u} 1_{\mathbb{R}_+}(u)\right) d\mu_L(u) \\
&= \tfrac{1}{2}\int_{(-\infty,x]} \left(e^u 1_{\mathbb{R}_{--}}(u) + e^{-u} 1_{\mathbb{R}_+}(u)\right) d\mu_L(u) \\
&= \tfrac{1}{2}\left(\int_{(-\infty,x]} e^u 1_{\mathbb{R}_{--}}(u)\, d\mu_L(u) + \int_{(-\infty,x]} e^{-u} 1_{\mathbb{R}_+}(u)\, d\mu_L(u)\right) \\
&= \tfrac{1}{2}\left(\int_{(-\infty,x]\cap\mathbb{R}_{--}} e^u\, d\mu_L(u) + \int_{(-\infty,x]\cap\mathbb{R}_+} e^{-u}\, d\mu_L(u)\right).
\end{aligned}$$

Hence,

$$\int_{(-\infty,x]} F'(u)\, d\mu_L(u) = \begin{cases} \tfrac{1}{2}\int_{(-\infty,x]} e^u\, d\mu_L(u) & \text{if } x < 0 \\ \tfrac{1}{2}\left(\int_{(-\infty,0]} e^u\, d\mu_L(u) + \int_{[0,x]} e^{-u}\, d\mu_L(u)\right) & \text{if } x \ge 0 \end{cases}.$$

Now, we have

$$\int_{(-\infty,x]} e^u\, d\mu_L(u) = \int_{-\infty}^x e^u\, du = e^u\big|_{-\infty}^x = e^x$$

and

$$\int_{[0,x]} e^{-u}\, d\mu_L(u) = \int_0^x e^{-u}\, du = -\,e^{-u}\big|_0^x = 1 - e^{-x}$$

These imply

$$\int_{(-\infty,x]} F'(u)\, d\mu_L(u) = \begin{cases} \tfrac{1}{2}e^x & \text{if } x < 0 \\ \tfrac{1}{2}(1 + 1 - e^{-x}) & \text{if } x \ge 0 \end{cases} = \tfrac{1}{2}e^x 1_{\mathbb{R}_{--}}(x) + \left(1 - \tfrac{1}{2}e^{-x}\right) 1_{\mathbb{R}_+}(x) = F(x)$$

for every $x \in \mathbb{R}$. It then follows that In the end, we have

$$\mathbf{P}(-1 \le X \le 1) = F(1) - F(-1) = 1 - \tfrac{1}{2}e^{-1} - \tfrac{1}{2}e^{-1} = 1 - e^{-1}.$$

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \stackrel{def}{=} \alpha + \beta x, \quad \forall x \in \mathbb{R},$$

where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*

3. *Is Y absolutely continuous?*

4. *Are the first and second order moments of Y finite?*

5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution.

1. The function $g : \mathbb{R} \to \mathbb{R}$ is a Borel function. Therefore, $Y = g \circ X$ is a random variable.

2. Recall that $X \sim Unif(-1, 1)$ is absolutely continuous with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x),$$

for every $x \in \mathbb{R}$. Hence, writing $F_X : \mathbb{R} \to \mathbb{R}$ for the distribution function of X, we have

$$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\, d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{2} 1_{[-1,1]}(u)\, d\mu_L(u) \\ &= \frac{1}{2} \int_{(-\infty,x]\cap[-1,1]} d\mu_L(u) = \frac{1}{2} \mu_L((-\infty, x] \cap [-1, 1]). \end{aligned}$$

On the other hand,

$$(-\infty, x] \cap [-1, 1] = \begin{cases} \varnothing, & \text{if } x < -1, \\ \{-1\}, & \text{if } x = -1, \\ [-1, x], & \text{if } x < -1. \end{cases}$$

Therefore,

$$F_X(x) = \begin{cases} 0, & \text{if } x < -1, \\ \frac{x+1}{2}, & \text{if } -1 \leq x < 1, \\ 1, & \text{if } 1 \leq x. \end{cases}$$

Now, since g is a continuously differentiable real function on $\mathbb{R}$, in particular a Borel function, then $Y \equiv g(X) = \alpha + \beta X$ is a real random variable. To compute the distribution function F_Y, we apply the definition

$$F_Y(y) \stackrel{\text{def}}{=} \mathbf{P}(Y \leq y), \quad \forall y \in \mathbb{R}.$$

On the other hand, considering that $\beta \neq 0$, we have

$$\begin{aligned} \mathbf{P}(Y \leq y) &= \mathbf{P}(\alpha + \beta X \leq y) = \mathbf{P}\left(X \leq \frac{y-\alpha}{\beta}\right) \\ &= F_X\left(\frac{y-\alpha}{\beta}\right) = \begin{cases} 0, & \text{if } \frac{y-\alpha}{\beta} < -1 \Leftrightarrow y < \alpha - \beta, \\ \frac{\frac{y-\alpha}{\beta}+1}{2} = \frac{y+\beta-\alpha}{2\beta}, & \text{if } -1 \leq \frac{y-\alpha}{\beta} < 1 \Leftrightarrow \alpha - \beta \leq y < \alpha + \beta, \\ 1, & \text{if } 1 \leq \frac{y-\alpha}{\beta} \Leftrightarrow \alpha + \beta \leq y. \end{cases} \end{aligned}$$

Summarizing,

$$F_Y(y) = \begin{cases} 0, & \text{if } y < \alpha - \beta, \\ \frac{y+\beta-\alpha}{2\beta}, & \text{if } \alpha - \beta \leq y \leq \alpha + \beta, \\ 1, & \text{if } \alpha + \beta < y. \end{cases}$$

Therefore, the random variable Y turns out to be a uniformly distributed random variable on the interval $[\alpha - \beta, \alpha + \beta]$. In symbols, $Y \sim Unif(\alpha - \beta, \alpha + \beta)$. It then follows that Y is absolutely continuous with density $f_Y : \mathbb{R} \to \mathbb{R}$ given by

$$f_Y(y) = \frac{1}{2\beta} 1_{[\alpha-\beta,\alpha+\beta]}(y).$$

3. Since X is in the linear space $\mathcal{L}^2(\Omega; \mathbb{R})$, the random variable $Y = \alpha + \beta X$ is also in the linear space $\mathcal{L}^2(\Omega; \mathbb{R})$. Hence, Y has finite moments of order 1 and 2.

4. Thanks to the linearity of the expectation operator, we have

$$\mathbf{E}[Y] = \mathbf{E}[\alpha + \beta X] = \alpha + \beta \mathbf{E}[X],$$

where

$$\mathbf{E}[X] = \int_{\mathbb{R}} \frac{1}{2} x 1_{[-1,1]}(x)\, d\mu_L(x) = \frac{1}{2} \int_{[-1,1]} x\, d\mu_L(x) = \frac{1}{2} \int_{-1}^{1} x\, dx = \frac{1}{4}\, x^2\big|_{-1}^{1} = 0.$$

Therefore,

$$\mathbf{E}[Y] = \alpha.$$

Moreover considering the properties of the variance operator, we have

$$\mathbf{D}^2[Y] = \mathbf{D}^2[\alpha + \beta X] = \beta^2 \mathbf{D}^2[X],$$

where

$$\mathbf{D}^2[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X] = \mathbf{E}\left[X^2\right]$$

and

$$\mathbf{E}\left[X^2\right] = \int_{\mathbb{R}} \frac{1}{2} x^2 1_{[-1,1]}(x)\, d\mu_L(x) = \frac{1}{2} \int_{[-1,1]} x^2\, d\mu_L(x) = \frac{1}{2} \int_{-1}^{1} x^2\, dx = \frac{1}{6}\, x^3\big|_{-1}^{1} = \frac{1}{3}.$$

Therefore,

$$\mathbf{D}^2[Y] = \frac{\beta^2}{3}.$$

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \stackrel{def}{=} |x|, \quad \forall x \in \mathbb{R},$$

where $|x|$ is the absolute value of x.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. Recall that, since $X \sim Unif(-1, 1)$, the random variable X is absolutely continuous with density

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x),$$

for every $x \in \mathbb{R}$. Now, we have

$$F_Y(y) \stackrel{\text{def}}{=} \mathbf{P}(Y \leq y) = \mathbf{P}(g(X) \leq y) = \mathbf{P}(|X| \leq y) = \begin{cases} 0, & \text{if } y < 0, \\ \mathbf{P}(-y \leq X \leq y), & \text{if } y \geq 0. \end{cases}$$

On the othe hand, under the assumption $y \geq 0$, we have

$$\begin{aligned}
\mathbf{P}(-y \leq X \leq y) &= \int_{[-y,y]} f_X(x)\, d\mu_X(x) \\
&= \int_{[-y,y]} \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\
&= \frac{1}{2} \int_{[-y,y]\cap[-1,1]} d\mu_X(x) \\
&= \frac{1}{2} \mu_X([-y, y] \cap [-1, 1]),
\end{aligned}$$

where

$$\mu_X([-y, y] \cap [-1, 1]) = \begin{cases} \mu_X([-y, y]) = 2y, & \text{if } y \leq 1, \\ \mu_X([-1, 1]) = 2, & \text{if } y > 1. \end{cases}$$

It follows

$$F_Y(y) = \begin{cases} 0, & \text{if } y < 0, \\ y, & \text{if } 0 \leq y \leq 1, \\ 1, & \text{if } y > 1. \end{cases}$$

We can then recognize that $Y \sim Unif(0, 1)$, which implies that Y is absolutely continuous with density given by

$$f_Y(y) = 1_{[0,1]}(y),$$

for every $y \in \mathbb{R}$, and Y has finite first and second order moments. More specifically

$$\begin{aligned}
\mathbf{E}[Y] &= \int_{\mathbb{R}} y f_Y(y)\, d\mu_X(y) = \int_{\mathbb{R}} y 1_{[0,1]}(y)\, d\mu_X(y) \\
&= \int_{[0,1]} y d\mu_X(y) = \int_0^1 y dy = \frac{1}{2} y^2 \Big|_0^1 \\
&= \frac{1}{2}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}\left[Y^2\right] &= \int_{\mathbb{R}} y^2 f_Y(y)\, d\mu_X(y) = \int_{\mathbb{R}} y^2 1_{[0,1]}(y)\, d\mu_X(y) \\
&= \int_{[0,1]} y^2 d\mu_X(y) = \int_0^1 y^2 dy = \frac{1}{3} y^3 \Big|_0^1 \\
&= \frac{1}{3}.
\end{aligned}$$

It follows

$$\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Note that, since $Y = |X|$ it would be possible to compute $\mathbf{E}[Y]$ and $\mathbf{E}\left[Y^2\right]$ by using the density of X. That is

$$\begin{aligned}
\mathbf{E}[Y] = \mathbf{E}[|X|] &= \int_{\mathbb{R}} |x| f_X(x)\, d\mu_X(x) \\
&= \int_{\mathbb{R}} |x| \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\
&= \frac{1}{2} \int_{[-1,1]} |x|\, d\mu_X(x) \\
&= \frac{1}{2} \left(\int_{[-1,0]} -x d\mu_X(x) + \int_{[0,1]} x d\mu_X(x) \right) \\
&= \frac{1}{2} \left(\int_{-1}^0 -x dx + \int_0^1 x dx \right) \\
&= \frac{1}{2} \left(-\int_{-1}^0 x dx + \int_0^1 x dx \right) \\
&= \frac{1}{2} \left(-\frac{1}{2} x^2 \Big|_{-1}^0 + \frac{1}{2} x^2 \Big|_0^1 \right) \\
&= \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2} \right) \\
&= \frac{1}{2}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[|X|^2\right] = \mathbf{E}\left[X^2\right] &= \int_{\mathbb{R}} x^2 f_X(x)\, d\mu_X(x) \\
&= \int_{\mathbb{R}} x^2 \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\
&= \frac{1}{2}\int_{[-1,1]} x^2 d\mu_X(x) \\
&= \frac{1}{2}\int_{-1}^{1} x^2 dx \\
&= \frac{1}{2}\frac{1}{3}x^3\Big|_{-1}^{1} \\
&= \frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}\right) \\
&= \frac{1}{3}.
\end{aligned}$$

.

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} x^2, \quad \forall x \in \mathbb{R}.$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 5 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} x^3, \quad \forall x \in \mathbb{R}.$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. Recall that $X \sim Unif(-1, 1)$ is absolutely continuous, with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x).$$

1. The function g is clearly continuous. In particular, g is a Borel function. Therefore, $Y = g \circ X$ is a random variable.

2. The distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of Y is given by

$$F_Y(y) = \mathbf{P}(Y \le y) = \mathbf{P}(g(X) \le y)$$

for every $y \in \mathbb{R}$. Now, due to the definition of g, we have

$$\{x \in \mathbb{R} : g(x) \le y\} = \begin{cases} \left\{x \in \mathbb{R} : x \le -\sqrt[3]{|y|}\right\}, & \text{if } y < 0, \\ \left\{x \in \mathbb{R} : x \le \sqrt{y}\right\}, & \text{if } y \ge 0. \end{cases}$$

Hence,

$$\{g(X) \le y\} = \begin{cases} \left\{X \le -\sqrt[3]{|y|}\right\}, & \text{if } y < 0, \\ \left\{X \le \sqrt{y}\right\}, & \text{if } y \ge 0. \end{cases}$$

It follows,

$$\mathbf{P}(g(X) \le y)) = \begin{cases} \mathbf{P}\left(X \le -\sqrt[3]{|y|}\right) & \text{if } y < 0, \\ \mathbf{P}\left(X \le \sqrt{y}\right), & \text{if } y \ge 0. \end{cases}$$

On the other hand, since $X \sim Unif(-1, 1)$, for every $y < 0$, we have

$$\begin{aligned}
\mathbf{P}\left(X \le -\sqrt[3]{|y|}\right) &= \int_{\left(-\infty, -\sqrt[3]{|y|}\right]} f_X(x)\, d\mu_L(x) \\
&= \int_{\left(-\infty, -\sqrt[3]{|y|}\right]} \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_L(x) \\
&= \frac{1}{2}\int_{\left(-\infty, -\sqrt[3]{|y|}\right] \cap [-1,1]} d\mu_L(x) \\
&= \frac{1}{2}\mu_L\left(\left(-\infty, -\sqrt[3]{|y|}\right] \cap [-1, 1]\right),
\end{aligned}$$

where

$$\left(-\infty, -\sqrt[3]{|y|}\right] \cap [-1, 1] = \begin{cases} \varnothing, & \text{if } y < -1, \\ \left[-1, -\sqrt[3]{|y|}\right], & \text{if } -1 \le y < 0, \end{cases}$$

Therefore,

$$\mathbf{P}\left(X \le -\sqrt[3]{|y|}\right) = \begin{cases} 0, & \text{if } y < -1, \\ \frac{1}{2}\left(-\sqrt[3]{|y|} + 1\right), & \text{if } -1 \le y < 0. \end{cases}$$

Similarly, for every $y>0$, we have

$$\begin{aligned}\mathbf{P}\left(X\le\sqrt[3]{y}\right)&=\int_{(-\infty,\sqrt[3]{y}]}f_X(x)\,d\mu_L(x)\\&=\int_{(-\infty,\sqrt[3]{y}]}\frac{1}{2}1_{[-1,1]}(x)\,d\mu_L(x)\\&=\frac{1}{2}\int_{(-\infty,\sqrt[3]{y}]\cap[-1,1]}d\mu_L(x)\\&=\frac{1}{2}\mu_L\left((-\infty,\sqrt[3]{y}]\cap[-1,1]\right),\end{aligned}$$

where

$$(-\infty,\sqrt[3]{y}]\cap[-1,1]=\begin{cases}\left[-1,\sqrt[3]{y}\right], & \text{if } 0\le y\le 1,\\ [-1,1], & \text{if } 1<y<1.\end{cases}$$

Therefore,

$$\mathbf{P}\left(X\le\sqrt[3]{y}\right)=\begin{cases}\sqrt[3]{y}+1, & \text{if } 0\le y\le 1,\\ 1, & \text{if } 1<y.\end{cases}$$

We can then write,

$$F_Y(y)=\frac{1}{2}\left(1-\sqrt[3]{|y|}\right)1_{[-1,0]}(y)+\frac{1}{2}\left(1+\sqrt[3]{y}\right)1_{[0,1]}(y)+1_{(1,+\infty)}(y).$$

3. Note that $F_Y:\mathbb{R}\to\mathbb{R}$ is continuous in $\mathbb{R}$, it is differentiable in $\mathbb{R}-\{-1,0,1\}$ and we have

$$F_Y'(y)=\begin{cases}0, & \text{if } y<-1,\\ \frac{1}{6}\frac{\sqrt[3]{|y|}}{|y|}, & \text{if } -1<y<0,\\ \frac{1}{6}\frac{\sqrt[3]{y}}{y}, & \text{if } 0<y<1,\\ 0, & \text{if } 1<y.\end{cases}$$

Therefore, $F_Y:\mathbb{R}\to\mathbb{R}$ is not differentiable in $y=-1$, $y=0$, and $y=1$. On the other hand, consider the function $f:\mathbb{R}\to\mathbb{R}$, given by

$$f(y)\overset{\text{def}}{=}\frac{1}{6}\left(\frac{\sqrt[3]{|y|}}{|y|}1_{(-1,0)}(y)+\frac{\sqrt[3]{y}}{y}1_{(0,1)}(y)\right),\quad\forall y\in\mathbb{R}.$$

we have

$$\int_{(-\infty,y)}f(v)\,d\mu_L(v)=\frac{1}{6}\left(\int_{(-\infty,y)}\frac{\sqrt[3]{|v|}}{|v|}1_{(-1,0)}(v)\,d\mu_L(v)+\int_{(-\infty,y)}\frac{\sqrt[3]{v}}{v}1_{(0,1)}(v)\,d\mu_L(v)\right),$$

where

$$\int_{(-\infty,y)}\frac{\sqrt[3]{|v|}}{|v|}1_{(-1,0)}(v)\,d\mu_L(v)=\int_{(-\infty,y)\cap(-1,0)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v)=\begin{cases}0, & \text{if } y\le-1,\\ \int_{(-1,y)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v), & \text{if } -1<y<0,\\ \int_{(-1,0)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v), & \text{if } 0\le y,\end{cases}$$

and

$$\int_{(-\infty,y)}\frac{\sqrt[3]{v}}{v}1_{(0,1)}(v)\,d\mu_L(v)=\int_{(-\infty,y)\cap(0,1)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v)=\begin{cases}0, & \text{if } y\le 0,\\ \int_{(0,y)}\frac{\sqrt[3]{v}}{v}d\mu_L(v), & \text{if } 0<y<1,\\ \int_{(0,1)}\frac{\sqrt[3]{v}}{v}d\mu_L(v), & \text{if } 1\le y.\end{cases}$$

On the other hand,

$$\int_{(-1,y)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v)=\int_{-1}^{y}\frac{\sqrt[3]{|v|}}{|v|}dv=-3\sqrt[3]{|v|}\Big|_{-1}^{y}=3\left(1-\sqrt[3]{|y|}\right),$$

for every $y\in(-1,0)$,

$$\int_{(-1,0)}\frac{\sqrt[3]{|v|}}{|v|}d\mu_L(v)=\lim_{y\to0^-}\int_{-1}^{y}\frac{\sqrt[3]{|v|}}{|v|}dv=\lim_{y\to0^-}3\left(1-\sqrt[3]{|y|}\right)=3.$$

Furthermore,

$$\int_{(0,y)}\frac{\sqrt[3]{v}}{v}d\mu_L(v)=\lim_{x\to0^+}\int_{x}^{y}\frac{\sqrt[3]{v}}{v}dv=\lim_{x\to0^+}3\sqrt[3]{v}\Big|_{x}^{y}=\lim_{x\to0^+}3\left(\sqrt[3]{y}-\sqrt[3]{x}\right)=3\sqrt[3]{y},$$

and

$$\int_{(0,1)}\frac{\sqrt[3]{v}}{v}d\mu_L(v)=\lim_{x\to0^+}\int_{x}^{1}\frac{\sqrt[3]{v}}{v}dv=\lim_{x\to0^+}3\sqrt[3]{v}\Big|_{x}^{1}=\lim_{x\to0^+}3\left(1-\sqrt[3]{x}\right)=3.$$

It then follows,

$$\int_{(-\infty,y)}f(v)\,d\mu_L(v)=\begin{cases}0, & \text{if } y\le-1\\ \frac{1}{2}\left(1-\sqrt[3]{|y|}\right), & \text{if } -1<y\le 0,\\ \frac{1}{2}\left(1+\sqrt[3]{y}\right), & \text{if } 0<y\le 1\\ 1, & \text{if } 1<y.\end{cases}$$

Hence,

$$\int_{(-\infty,y)}f(v)\,d\mu_L(v)=\frac{1}{2}\left(1-\sqrt[3]{|y|}\right)1_{(-1,0]}(y)+\frac{1}{2}\left(1+\sqrt[3]{y}\right)1_{(0,1]}(y)+1_{(0,+\infty)}(y)=F_Y(y)$$

almost everywhere in $\mathbb{R}$. Therefore, Y is absolutely continuous in $\mathbb{R}$ and a density for Y is given by $f:\mathbb{R}\to\mathbb{R}$.

4. We have

$$\int_\Omega Y^2 d\mathbf{P}=\int_\Omega g(X)^2\,d\mathbf{P}.$$

Therefore, Y has finite moment of order 2 or not according to whether

$$\int_\Omega g(X)^2\,d\mathbf{P}<\infty.$$

Now, since X is absolutely continuous, we can write

$$\begin{aligned}\int_\Omega g(X)^2\,d\mathbf{P}&=\int_{\mathbb{R}}g(x)^2 f_X(x)\,d\mu_L(x)\\&=\int_{\mathbb{R}}x^4 1_{(0,+\infty)}(x)\,1_{(-1,1)}(x)\,d\mu_L(x)\\&=\frac{1}{2}\int_{(0,1)}x^4 d\mu_L(x)\\&=\frac{1}{2}\int_0^1 x^4 dx\\&=\frac{1}{10}x^5\Big|_0^1\\&=\frac{1}{10}.\end{aligned}$$

It follows, that Y has finite moment of order 2 and

$$\mathbf{E}\left[Y^2\right] = \int_\Omega Y^2 d\mathbf{P} = \frac{1}{10}.$$

A fortiori Y has finite moment of order 1 and

$$\begin{aligned}
\mathbf{E}[Y] = \mathbf{E}[g(X)] &= \int_\mathbb{R} g(x) f_X(x) d\mu_L(x) \\
&= \int_\mathbb{R} x^2 1_{(0,+\infty)}(x) 1_{(-1,1)}(x) d\mu_L(x) \\
&= \frac{1}{2}\int_{(0,1)} x^2 d\mu_L(x) \\
&= \frac{1}{2}\int_0^1 x^2 dx \\
&= \frac{1}{6} x^3 \Big|_0^1 \\
&= \frac{1}{6}.
\end{aligned}$$

In the end,

$$\mathbf{D}^2[Y] = \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 = \frac{1}{10} - \frac{1}{36} = \frac{13}{180}.$$

.

Problem 6 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ the real state space, and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1,1]$. In symbols, $X \sim Unif(-1,1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \stackrel{def}{=} \begin{cases} 0, & \text{if } x \le 0. \\ x^2, & \text{if } x > 0. \end{cases}$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. Recall that $X \sim Unif(-1,1)$ is absolutely continuous, with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x).$$

Note also that we can write

$$g(x) = x^2 1_{(0,+\infty)}(x),$$

for every $x \in \mathbb{R}$.

1. The function g is clearly continuous. In particular, g is a Borel function. Therefore, $Y = g \circ X$ is a random variable.

2. The distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of Y is given by

$$F_Y(y) = \mathbf{P}(Y \le y) = \mathbf{P}(g(X) \le y)$$

for every $y \in \mathbb{R}$. Now, due to the definition of g, we have

$$\{x \in \mathbb{R} : g(x) \le y\} = \begin{cases} \varnothing, & \text{if } y < 0, \\ \{x \in \mathbb{R} : x \le \sqrt{y}\}, & \text{if } y \ge 0. \end{cases}$$

Hence,

$$\{g(X) \le y\} = \begin{cases} \varnothing, & \text{if } y < 0, \\ \{X \le \sqrt{y}\}, & \text{if } y \ge 0. \end{cases}$$

It follows,

$$\mathbf{P}(g(X) \le y)) = \begin{cases} 0, & \text{if } y < 0, \\ \mathbf{P}(X \le \sqrt{y}), & \text{if } y \ge 0. \end{cases}$$

On the other hand, since $X \sim Unif(-1,1)$, we have

$$\begin{aligned}
\mathbf{P}(X \le \sqrt{y}) &= \int_{(-\infty,\sqrt{y}]} f_X(x) d\mu_L(x) \\
&= \int_{(-\infty,\sqrt{y}]} \frac{1}{2} 1_{[-1,1]}(x) d\mu_L(x) \\
&= \frac{1}{2}\int_{(-\infty,\sqrt{y}]\cap[-1,1]} d\mu_L(x) \\
&= \frac{1}{2}\mu_L\left((-\infty,\sqrt{y}] \cap [-1,1]\right),
\end{aligned}$$

where

$$(-\infty,\sqrt{y}] \cap [-1,1] = \begin{cases} [-1,\sqrt{y}], & \text{if } 0 \le y < 1, \\ [-1,1], & \text{if } y \ge 1. \end{cases}$$

Therefore,

$$\mathbf{P}(X \le \sqrt{y}) = \begin{cases} \frac{1}{2}\left(\sqrt{y}+1\right), & \text{if } y < 1, \\ 1, & \text{if } y \ge 1. \end{cases}$$

We can then write,

$$F_Y(y) = \frac{1}{2}\left(\sqrt{y}+1\right) 1_{[0,1]}(y) + 1_{(1,+\infty)}(y).$$

Note that

$$\mathbf{P}(Y < 0) = F_Y(0) = 0.$$

Hence, Y is a non negative random variable.

3. Note that $F_Y : \mathbb{R} \to \mathbb{R}$ is not continuous since

$$\lim_{y \to 0-^-} F_Y(y) = 0 \quad \text{and} \quad \lim_{y \to 0^+} F_Y(y) = \frac{1}{2}.$$

A fortiori it is not absolutely continuous.

4. We have
$$\int_{\Omega} Y^2 d\mathbf{P} = \int_{\Omega} g(X)^2 \, d\mathbf{P}.$$
Therefore, Y has finite moment of order 2 or not according to whether
$$\int_{\Omega} g(X)^2 \, d\mathbf{P} < \infty.$$
Now, since X is absolutely continuous, we can write
$$\begin{aligned}\int_{\Omega} g(X)^2 \, d\mathbf{P} &= \int_{\mathbb{R}} g(x)^2 f_X(x) \, d\mu_L(x) \\ &= \int_{\mathbb{R}} x^4 1_{(0,+\infty)}(x) \, 1_{(-1,1)}(x) \, d\mu_L(x) \\ &= \frac{1}{2} \int_{(0,1)} x^4 d\mu_L(x) \\ &= \frac{1}{2} \int_0^1 x^4 dx \\ &= \frac{1}{10} x^5 \Big|_0^1 \\ &= \frac{1}{10}.\end{aligned}$$
It follows, that Y has finite moment of order 2 and
$$\mathbf{E}\left[Y^2\right] = \int_{\Omega} Y^2 d\mathbf{P} = \frac{1}{10}.$$
A fortiori Y has finite moment of order 1 and
$$\begin{aligned}\mathbf{E}[Y] = \mathbf{E}[g(X)] &= \int_{\mathbb{R}} g(x) f_X(x) \, d\mu_L(x) \\ &= \int_{\mathbb{R}} x^2 1_{(0,+\infty)}(x) \, 1_{(-1,1)}(x) \, d\mu_L(x) \\ &= \frac{1}{2} \int_{(0,1)} x^2 d\mu_L(x) \\ &= \frac{1}{2} \int_0^1 x^2 dx \\ &= \frac{1}{6} x^3 \Big|_0^1 \\ &= \frac{1}{6}.\end{aligned}$$
In the end,
$$\mathbf{D}^2[Y] = \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 = \frac{1}{10} - \frac{1}{36} = \frac{13}{180}.$$

Problem 7 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*
$$g(x) \stackrel{def}{=} x^2 - 2x, \quad \forall x \in \mathbb{R},$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*
$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$
is a random variable?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 8 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be an exponentially distributed random variable with rate parameter $\lambda = 1$. In symbols, $X \sim Exp(1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*
$$g(x) \stackrel{def}{=} 1 - \exp(-x), \quad \forall x \in \mathbb{R},$$
where $\exp : \mathbb{R} \to \mathbb{R}$ is the Neper exponential function.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*
$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$
is a random variable?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 9 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $(0, 1)$. In symbols, $X \sim Unif(0, 1)$. Consider the function $g : \mathbb{R}_{++} \to \mathbb{R}$ given by*
$$g(y) \stackrel{def}{=} -\frac{1}{\lambda} \ln(y), \quad \forall \in \mathbb{R}_{++},$$
where $\ln : \mathbb{R}_{++} \to \mathbb{R}$ is the natural logarithm function and $\lambda > 0$.

1. *Can you state that the function $Y : \Omega \to \mathbb{R}$ given by*
$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$
is a real random variable on Ω?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of $Y : \Omega \to \mathbb{R}$?*
3. *Is Y absolutely continuous?*

fatta a pagina 66 del pdf

4. *Are the first and second order moments of Y finite?*

5. *If the first and second order moments of Y are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*
Hint: recall the properties of the ***logarithm*** *and* ***exponential*** *function.*

Solution. .

Problem 10 *Consider the function* $f:\mathbb{R}\to\mathbb{R}_+$ *given by*

$$f(x) \overset{def}{=} \begin{matrix} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{if } x<0 \end{matrix}.$$

Prove that f is a density function (it could be helpful to draw the graph of f). Assume that $X:\Omega\to\mathbb{R}$ is a real random variable on some probability space with density f. Determine the distribution function $F_X:\mathbb{R}\to\mathbb{R}$ of X. Compute $\mathbf{E}[X]$ *and* $\mathbf{D}^2[X]$*. In the end, consider the function $Y:\Omega\to\mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} e^{X(\omega)}, \quad \forall\omega\in\Omega.$$

Is Y a random variable? In case it is, determine the distribution function $F_Y:\mathbb{R}\to\mathbb{R}$ of Y. Can you say that Y is absolutely continuous? In the affirmative case, can you compute the density function of Y? Does Y have finite expectation and variance? What about computing $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. .

Problem 11 *Let $(\Omega,\mathcal{E},\mathbf{P})\equiv\Omega$ be a probability space and let $X\sim Exp(1)$ be an exponentially distributed random variable with rate parameter $\lambda=1$. Consider the function $Y:\Omega\to\mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} \ln(X(\omega)), \quad \forall\omega\in\Omega.$$

Is Y a random variable? In the affirmative case, determine the distribution function $F_Y:\mathbb{R}\to\mathbb{R}$ of Y. Can you say that Y is absolutely continuous? In the affirmative case, can you compute the density function of Y? Does Y have finite expectation and variance? What about computing $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. We have

$$f_X(x) = e^{-x}1_{\mathbb{R}_+}(x),$$

for every $x\in\mathbb{R}$. Therefore,

$$\mathbf{P}(X\leq 0) = \int_{\mathbb{R}_-} f_X(x)\ d\mu_L(x) = 0.$$

Hence, the function Y, composition of the strictly positive random variable $X:\Omega\to\mathbb{R}$ and the continuous function $\ln:\mathbb{R}_+\to\mathbb{R}$ turns out to be well defined as a random variable. Now, according to the definition and considering that the exponential function is increasing, we have

$$\begin{aligned} F_Y(y) &= \mathbf{P}(Y\leq y) = \mathbf{P}(\ln(X)\leq y) \\ &= \mathbf{P}(\exp(\ln(X))\leq\exp(y)) = \mathbf{P}(X\leq e^y) \\ &= \int_{(-\infty,e^y]} f_X(u)\ d\mu_L(u), \end{aligned}$$

for every $y\in\mathbb{R}$. On the other hand, since $e^y>0$, we can write

$$\begin{aligned} \int_{(-\infty,e^y]} f_X(u)\ d\mu_L(u) &= \int_{(-\infty,e^y]} e^{-x}1_{\mathbb{R}_+}(x)\ d\mu_L(u) = \int_{(-\infty,e^y]\cap\mathbb{R}_+} e^{-x}\ d\mu_L(u) \\ &= \int_{(0,e^y]} e^{-x}\ d\mu_L(u) = \int_0^{e^y} e^{-u}\ du \\ &= 1-e^{-e^y}. \end{aligned}$$

It then follows

$$F_Y(y) = 1-e^{-e^y},$$

for every $y\in\mathbb{R}$. The distribution function $F_Y:\mathbb{R}\to\mathbb{R}$ is continuously differentiable on $\mathbb{R}$ and

$$F_Y'(y) = e^y e^{-e^y}$$

In addition, $F_Y'(y)$ is clearly bounded. In fact, we have

$$F_Y''(y) = e^y e^{-e^y}(1-e^y).$$

Hence, $F_Y'(y)$ takes a unique maximum at the point $y=0$ with value $F_Y'(0)=e^{-1}$. As a consequence, Y is absolutely continuous and has a density $f_Y:\mathbb{R}\to\mathbb{R}$ given by

$$f_Y(y) = F_Y'(y) = e^y e^{-e^y}.$$

To check whether Y has finite first order moment, we study

$$\begin{aligned} \int_\Omega |Y|\ d\mathbf{P} &= \int_{\mathbb{R}} |y|\, f_Y(y)\ d\mu_L(y) = \int_{\mathbb{R}} |y|\, e^y e^{-e^y}\ d\mu_L(y) \\ &= \int_{\mathbb{R}_-} -ye^y e^{-e^y}\ d\mu_L(y) + \int_{\mathbb{R}_+} ye^y e^{-e^y}\ d\mu_L(y) \\ &= -\int_{-\infty}^0 ye^y e^{-e^y}\ dy + \int_0^{+\infty} ye^y e^{-e^y}\ dy. \end{aligned}$$

Consider $\int_0^{+\infty} ye^y e^{-e^y}\ dy$. Setting $e^y=z$ we have

$$y=\ln(z), \quad dy=\frac{1}{z}dz$$

and

$$\int_0^{+\infty} ye^y e^{-e^y}\ dy = \int_1^{+\infty} \ln(z)\, e^{-z}\ dz,$$

where

$$\begin{aligned} \int_1^{+\infty} \ln(z)\, e^{-z}\ dz &\leq \int_1^{+\infty} ze^{-z}\ dz = -\int_1^{+\infty} z\ de^{-z} \\ &= -\left(ze^{-z}\Big|_1^{+\infty} - \int_1^{+\infty} e^{-z}\ dz \right) \\ &= -\left(ze^{-z}\Big|_1^{+\infty} + e^{-z}\Big|_1^{+\infty} \right) \\ &= 2e^{-1}. \end{aligned}$$

More precisely,

$$\int_0^{+\infty} ye^y e^{-e^y}\ dy = \int_1^{+\infty} \ln(z)\, e^{-z}\ dz = -\int_1^{+\infty} \ln(z)\ de^{-z}$$
$$= -\left[\ln(z)\, e^{-z}\big|_1^{+\infty} - \int_1^{+\infty} \frac{e^{-z}}{z}\ dz\right]$$
$$= \int_1^{+\infty} \frac{e^{-z}}{z}\ dz$$
$$= \Gamma(0,1) \simeq 0.21939.$$

Similarly,

$$\int_{-\infty}^{0} ye^y e^{-e^y}\ dy = \int_0^1 \ln(z)\, e^{-z}\ dz,$$

where

$$\int_0^1 \ln(z)\, e^{-z}\ dz \geq \int_0^1 \ln(z)\ dz$$
$$= \left(z\ln(z)|_0^1 - \int_0^1 z\frac{1}{z}\ dz\right)$$
$$= -1$$

More precisely,

$$\int_{-\infty}^{0} ye^y e^{-e^y}\ dy = \int_0^1 \ln(z)\, e^{-z}\ dz \simeq -0.7966.$$

As a consequence,

$$\int_\Omega |Y|\ d\mathbf{P} < \infty.$$

More precisely,

$$\int_\Omega |Y|\ d\mathbf{P} \simeq 0.2194 + 0.7966 = 1.016\,0.$$

It follows that Y has finite expectation given by

$$\int_\Omega Y\ d\mathbf{P} = \int_{\mathbb{R}_-} ye^y e^{-e^y}\ d\mu_L(y) + \int_{\mathbb{R}_+} ye^y e^{-e^y}\ d\mu_L(y)$$
$$= \int_{-\infty}^{0} ye^y e^{-e^y}\ dy + \int_0^{+\infty} ye^y e^{-e^y}\ dy$$
$$= -0.7966 + 0.2194$$
$$= -0.5772.$$

Now, we have

$$\int_\Omega Y^2\ d\mathbf{P} = \int_{\mathbb{R}} y^2 f_X(y)\ d\mu_L(y) = \int_{\mathbb{R}} y^{2y} e^{-e^y}\ d\mu_L(y) = \int_{-\infty}^{+\infty} y^2 e^y e^{-e^y}\ dy$$

and, with a similar argument as above, it is possible to prove that

$$\int_{-\infty}^{+\infty} y^2 e^y e^{-e^y}\ dy < \infty.$$

More precisely

$$\int_{-\infty}^{+\infty} y^2 e^y e^{-e^y}\ dy = \gamma^2 + \frac{\pi^2}{6},$$

where γ is the Euler gamma constant. □

SPERANZA CONDIZIONATA

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Conditional Expectation with Solution 2022-12-08

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$ such that*

$$\mathbf{E}[Y \mid X] = X \quad \text{and} \quad \mathbf{E}\left[Y^2 \mid X\right] = X^2.$$

Prove that $Y = X$, $\mathbf{P}$-a.s. on Ω.

Solution. We have $Y = X$, $\mathbf{P}$-a.s. on Ω if and only if there exists an event $E \in \mathcal{E}$ such that $\mathbf{P}(E) = 0$ and $Y(\omega) = X(\omega)$ for every $\omega \in \Omega - E$. By virtue of the properties of the Lebesgue integral, we have

$$Y = X,\ \mathbf{P}\text{-a.s. on } \Omega \Leftrightarrow \int_\Omega (X - Y)^2\, d\mathbf{P} = 0.$$

On the other hand,

$$\int_\Omega (X - Y)^2\, d\mathbf{P} \equiv \mathbf{E}\left[(X - Y)^2\right].$$

Hence, we evaluate

$$\mathbf{E}\left[(X - Y)^2\right] = \mathbf{E}\left[X^2 - 2XY + Y^2\right] = \mathbf{E}\left[X^2\right] - 2\mathbf{E}[XY] + \mathbf{E}\left[Y^2\right]. \tag{1}$$

Now, by virtue of the properties of the conditional expectation operator, under our assumptions, we have

$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[XY \mid X]] = \mathbf{E}[X\mathbf{E}[Y \mid X]] = \mathbf{E}\left[X^2\right] \tag{2}$$

and

$$\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[\mathbf{E}\left[Y^2 \mid X\right]\right] = \mathbf{E}\left[X^2\right]. \tag{3}$$

Combining (1)-(3) it follows

$$\mathbf{E}\left[(X - Y)^2\right] = 0,$$

which implies the desired result.

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the real Borel state space. Let $N \subseteq \mathbb{N}$, let $\{F_n\}_{n\in N}$ be a complete system of mutually exclusive events of Ω and let $\mathcal{F}$ be the σ-algebra generated by $\{F_n\}_{n\in N}$. In symbols $\mathcal{F} \equiv \sigma(\{F_n\}_{n\in N})$. We know that a map $Y : \Omega \to \mathbb{R}$ is an $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variable if and only if*

$$Y(\omega) = \sum_{n\in N} y_n 1_{F_n}(\omega), \quad \forall \omega \in \Omega,$$

where $(y_n)_{n\in N}$ is a suitable sequence of real numbers.

Consider a random variable $X \in L^2(\Omega; \mathbb{R})$ and let $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ the subspace $L^2(\Omega; \mathbb{R})$ of space of all $\mathcal{F}$-random variables. Use the above claim to prove that

$$\mathbf{E}[X \mid \mathcal{F}] = \underset{Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})}{\arg\min}\ \mathbf{E}\left[(X - Y)^2\right]$$

As a consequence, show that $\mathbf{E}[X \mid \mathcal{F}]$ is the orthogonal projection of X on $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$.

Solution. The space $L^2(\Omega_{\mathcal{F}};\mathbb{R})$ of all real $\mathcal{F}$-random variables with finite moment of order 2 is a subspace of $L^2(\Omega;\mathbb{R})$ because it fulfills the conditions for a subset of Hilbert space to be a subspace of the Hilbert space. In fact, for all $X,Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})$ and all $\alpha,\beta\in\mathbb{R}$ the linear combination $\alpha X+\beta Y$ is also in $L^2(\Omega_{\mathcal{F}};\mathbb{R})$, to say $L^2(\Omega_{\mathcal{F}};\mathbb{R})$ is closed for the linear combination. In addition, if $(X_n)_{n\geq 1}$ is a sequence belonging to $L^2(\Omega_{\mathcal{F}};\mathbb{R})$ and such that $X_n\xrightarrow{L^2}X$, where $X\in L^2(\Omega;\mathbb{R})$, we also have $X\in L^2(\Omega_{\mathcal{F}};\mathbb{R})$, to say $L^2(\Omega_{\mathcal{F}};\mathbb{R})$ is a closed subset of $L^2(\Omega;\mathbb{R})$ in the topology induced by the norm.

Now, given $X\in L^2(\Omega_{\mathcal{F}};\mathbb{R})$, consider the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R})\to\mathbb{R}_+$ given by

$$\Delta_X(Y)\overset{\text{def}}{=}\mathbf{E}\left[(X-Y)^2\right],\quad \forall Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R}).$$

Since in the case under concern

$$Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})\Leftrightarrow Y(\omega)=\sum_{n\in N}y_n 1_{F_n}(\omega),\quad \forall\omega\in\Omega,$$

we can write

$$\Delta_X(Y)=\mathbf{E}\left[\left(X-\sum_{n\in N}y_n 1_{F_n}\right)^2\right]\equiv\Delta_X(y_1,\dots,y_n,\dots).$$

Hence,

$$\begin{aligned}\Delta_X(y_1,\dots,y_n,\dots)&=\mathbf{E}\left[X^2-2\sum_{n\in N}y_nX1_{F_n}+\sum_{m,n\in N}y_my_n1_{F_m}1_{F_n}\right]\\&=\mathbf{E}\left[X^2\right]-2\sum_{n\in N}y_n\mathbf{E}\left[X1_{F_n}\right]+\sum_{m,n\in N}y_my_n\mathbf{E}\left[1_{F_m}1_{F_n}\right].\end{aligned}$$

On the other hand,

$$1_{F_m}1_{F_n}=\begin{cases}1_{F_n} & \text{if } m=n\\ 1_{\varnothing} & \text{if } m\neq n\end{cases}.$$

Moreover,

$$\mathbf{E}\left[1_E\right]=\mathbf{P}(E),\quad\forall E\in\mathcal{E}$$

and

$$\mathbf{E}\left[X1_E\right]=\int_\Omega X1_E\,d\mathbf{P}=\int_E X\,d\mathbf{P},\quad\forall E\in\mathcal{E}.$$

Therefore,

$$\Delta_X(Y)=\mathbf{E}\left[X^2\right]-2\sum_{n\in N}y_n\int_{F_n}X\,d\mathbf{P}+\sum_{n\in N}y_n^2\mathbf{P}(F_n).$$

As a consequence,

$$\partial_{y_m}\Delta_X(y_1,\dots,y_n,\dots)=-2\int_{Fm}X\,d\mathbf{P}+2y_m\mathbf{P}(F_m),\quad\forall m\in N,$$

which implies

$$\partial_{y_m}\Delta_X(y_1,\dots,y_n,\dots)=0\Leftrightarrow y_m=\frac{1}{\mathbf{P}(F_m)}\int_{F_m}X\,d\mathbf{P}=\mathbf{E}\left[X\mid F_m\right],\quad\forall m\in N.$$

Thus, a candidate minimum Y for the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R})\to\mathbb{R}_+$ takes the form

$$Y=\sum_{n\in N}\mathbf{E}\left[X\mid F_n\right]1_{F_n}=\mathbf{E}\left[X\mid\mathcal{F}\right].$$

Now, we have

$$\partial^2_{y_m}\Delta_X(y_1,\dots,y_n,\dots)=\mathbf{P}(F_m)>0$$

and the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R})\to\mathbb{R}_+$ is known to be convex[1]. It then follow that

$$\mathbf{E}\left[X\mid\mathcal{F}\right]=\underset{Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})}{\arg\min}\,\mathbf{E}\left[(X-Y)^2\right].$$

To complete the proof, it is sufficient to observe that in a Hilbert space the ortoghonal projection of a given vector onto a subspace determines the vector in the subspace of the minimum distance from the given vector.

□

Problem 3 *Let $(\Omega,\mathcal{E},\mathbf{P})\equiv\Omega$ be a probability space and let $(\mathbb{R},\mathcal{B}(\mathbb{R}))\equiv\mathbb{R}$ be the real Borel state space. Let $X,Y\in L^2(\Omega;\mathbb{R})$.*

1. *Prove in all details that $\mathbf{E}[Y\mid X]=\mathbf{E}[Y]$ a.e. on Ω implies $Cov(X,Y)=0$, but X and Y may not be independent.*

2. *Prove in all details that $Cov(X,Y)=0$ does not imply $\mathbf{E}[Y\mid X]=\mathbf{E}[Y]$.*

Hint: in the first case, to generate a suitable counterexample one may consider the random variables $X\sim Ber(p)$, $Z\sim N(0,1)$, independent of X, and $Y=XZ$. In the second case consider $X\sim N(0,1)$ and $Y=X^2$.

---

[1]To prove the convexity of the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R})\to\mathbb{R}_+$, we may observe that thanks, to the Cauchy-Schwarz inequality and the convexity of the standard quadratic function $f(u)\overset{\text{def}}{=}u^2$, we have

$$\begin{aligned}\Delta_X(\theta Y+(1-\theta)Z)&=\mathbf{E}\left[(X-(\theta Y+(1-\theta)Z))^2\right]\\&=\mathbf{E}\left[(\theta(X-Y)+(1-\theta)(X-Z))^2\right]\\&=\mathbf{E}\left[\theta^2(X-Y)^2+2\theta(1-\theta)(X-Y)(X-Z)+(1-\theta)^2(X-Z)^2\right]\\&=\theta^2\mathbf{E}\left[(X-Y)^2\right]+2\theta(1-\theta)\mathbf{E}\left[(X-Y)(X-Z)\right]+(1-\theta)^2\mathbf{E}\left[(X-Z)^2\right]\\&\leq\theta^2\mathbf{E}\left[(X-Y)^2\right]+2\theta(1-\theta)\left|\mathbf{E}\left[(X-Y)(X-Z)\right]\right|+(1-\theta)^2\mathbf{E}\left[(X-Z)^2\right]\\&\leq\theta^2\mathbf{E}\left[(X-Y)^2\right]+2\theta(1-\theta)\mathbf{E}\left[(X-Y)^2\right]^{1/2}\mathbf{E}\left[(X-Z)^2\right]^{1/2}+(1-\theta)^2\mathbf{E}\left[(X-Z)^2\right]\\&=\left(\theta\mathbf{E}\left[(X-Y)^2\right]^{1/2}+(1-\theta)\mathbf{E}\left[(X-Z)^2\right]^{1/2}\right)^2\\&\leq\theta\mathbf{E}\left[(X-Y)^2\right]+(1-\theta)\mathbf{E}\left[(X-Z)^2\right],\end{aligned}$$

for every $\theta\in[0,1]$.

To show the convexity of the standard quadratic function, $f(u)\overset{\text{def}}{=}u^2$, we may observe that the inequality

$$(u-v)^2\geq 0,$$

which holds true for every $u,v\in\mathbb{R}$, implies

$$-\theta(1-\theta)(u-v)^2\leq 0,$$

which holds true for every $u,v\in\mathbb{R}$ and $\theta\in[0,1]$. The latter can be rewritten as

$$-\theta(1-\theta)\left(u^2-2uv+v^2\right)\leq 0$$

or equivalently

$$\theta^2u^2-\theta u^2+2\theta(1-\theta)uv+(1-\theta)^2v^2-(1-\theta)v^2.$$

This implies

$$\theta^2u^2+2\theta(1-\theta)uv+(1-\theta)^2v^2\leq\theta u^2+(1-\theta)v^2.$$

Hence,

$$(\theta u+(1-\theta)v)^2\leq\theta u^2+(1-\theta)v^2,$$

which proves the desired result.

Solution.

1. Under the assumption $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω, by virtue of the properties of the conditional expectation operator, we can write

$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[XY \mid X]] = \mathbf{E}[X\mathbf{E}[Y \mid X]] = \mathbf{E}[X\mathbf{E}[Y]] = \mathbf{E}[X]\,\mathbf{E}[Y]$$

Therefore,

$$Cov(X,Y) = \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = 0.$$

Now, if we consider the random $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$, we have

$$\begin{aligned} Cov(X,Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = \mathbf{E}\left[X^2 Z\right] - \mathbf{E}[X]\,\mathbf{E}[XZ] \\ &= \mathbf{E}\left[X^2\right]\mathbf{E}[Z] - \mathbf{E}[X]^2\,\mathbf{E}[Z] \\ &= 0. \end{aligned}$$

On the other hand, we have

$$\mathbf{P}(X \le 0) = q,$$

and, on account that X and Z are independent,

$$\begin{aligned} \mathbf{P}(Y \le 0) &= \mathbf{P}(XZ \le 0) \\ &= \mathbf{P}(XZ \le 0, X = 0) + \mathbf{P}(XZ \le 0, X = 1) \\ &= \mathbf{P}(XZ \le 0 \mid X = 0)\,\mathbf{P}(X = 0) + \mathbf{P}(XZ \le 0 \mid X = 1)\,\mathbf{P}(X = 1) \\ &= \mathbf{P}(0 \le 0 \mid X = 0)\,\mathbf{P}(X = 0) + \mathbf{P}(Z \le 0 \mid X = 1)\,\mathbf{P}(X = 1) \\ &= \mathbf{P}(\Omega)\,\mathbf{P}(X = 0) + \mathbf{P}(Z \le 0)\,\mathbf{P}(X = 1) \\ &= q + \frac{1}{2}p. \end{aligned}$$

Furthermore, the same arguments as above shows that

$$\begin{aligned} \mathbf{P}(X \le 0, Y \le 0) &= \mathbf{P}(X \le 0, XZ \le 0) \\ &= \mathbf{P}(X = 0, XZ \le 0) \\ &= \mathbf{P}(XZ \le 0 \mid X = 0)\,\mathbf{P}(X = 0) \\ &= q. \end{aligned}$$

Hence, we have

$$\mathbf{P}(X \le 0)\,\mathbf{P}(Y \le 0) = q\left(q + \frac{1}{2}p\right) \ne q = \mathbf{P}(X \le 0, Y \le 0)$$

which shows that X and Y are not be independent.

2. To show that $Cov(X,Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$, we consider $X \sim N(0,1)$ and $Y = X^2$. We have

$$\begin{aligned} Cov(X,Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] \\ &= \mathbf{E}\left[X^3\right] - \mathbf{E}[X]\,\mathbf{E}\left[X^2\right] \\ &= 0, \end{aligned}$$

but

$$\mathbf{E}[Y \mid X] = \mathbf{E}\left[X^2 \mid X\right] = X^2 \ne \mathbf{E}\left[X^2\right] = \mathbf{E}[Y].$$

□

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be independent standard Gaussian distributed random variables on Ω. Set*

$$U \overset{def}{=} X + Y, \qquad V \overset{def}{=} X - Y.$$

1. *Compute the distributions of U and V.*

risolto a pagina 70

2. *Prove that U and V are independent.*

3. *Compute $\mathbf{E}[X \mid U]$, $\mathbf{E}[X \mid V]$, $\mathbf{E}[Y \mid U]$, $\mathbf{E}[Y \mid V]$.*

4. *Compute $\mathbf{E}[XY \mid U]$.*

Exercise 5 *Hint: First, concentrate your attention on the circumstance that X and Y are independent and standard Gaussian distributed. Second, it might be useful to consider $\mathbf{E}\left[X^2 \mid U\right]$ and $\mathbf{E}\left[Y^2 \mid U\right]$.*

Solution. .

Problem 6 *Let N be a geometric random variable with success probability p, which models the first occurrence of success in n independent trials, and let $(X_n)_{n\ge1}$ be a sequence of independent and normally distributed random variables with mean μ and variance σ^2, which are also independent of N. Study the conditional expectation*

$$\mathbf{E}\left[\textstyle\sum_{k=1}^{N} X_k \mid N\right].$$

Use the properties of the conditional expectation to compute the expectation and the variance of the random sum

$$S_N \overset{def}{=} \textstyle\sum_{k=1}^{N} X_k.$$

Solution. Since the random variables of the sequence $(X_n)_{n\ge1}$ are independent and are also independent of N, which is geometrically distributed, we can write

$$\begin{aligned}
\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right] &= \sum_{n=1}^{\infty} \mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N=n\right] 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega} \left(\sum_{k=1}^{N} X_k\right) 1_{\{N=n\}} \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}} \left(\sum_{k=1}^{N} X_k\right) \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}} \left(\sum_{k=1}^{n} X_k\right) \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega} \left(\sum_{k=1}^{n} X_k\right) 1_{\{N=n\}} \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right) 1_{\{N=n\}}\right] \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \left(\sum_{k=1}^{n} \mathbf{E}[X_k]\right) \mathbf{E}\left[1_{\{N=n\}}\right] \right) 1_{\{N=n_m\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \left(\sum_{k=1}^{n} \mu\right) \mathbf{P}(N=n) \right) 1_{\{N=n_m\}} \\
&= \sum_{n=1}^{\infty} n\mu 1_{\{N=n_m\}} \\
&= \mu \sum_{n=1}^{\infty} n 1_{\{N=n\}} \\
&\overset{\mathbf{P}\text{-a.s.}}{=} \mu N.
\end{aligned}$$

Now, we can write

$$\mathbf{E}\left[\sum_{k=1}^{N} X_k\right] = \mathbf{E}\left[\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right]\right] = \mathbf{E}[\mu N] = \mu \mathbf{E}[N] = \frac{\mu}{p}.$$

and

$$\mathbf{D}^2\left[\sum_{k=1}^{N} X_k\right] = \mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2\right] - \mathbf{E}\left[\sum_{k=1}^{N} X_k\right]^2 = \mathbf{E}\left[\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right]\right] - \frac{\mu^2}{p^2}.$$

Thus, we are left with computing

$$\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right].$$

A straightforward computation yields

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right] &= \sum_{n=1}^{\infty} \mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N=n\right] 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega} \left(\sum_{k=1}^{N} X_k\right)^2 1_{\{N=n\}} \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}} \left(\sum_{k=1}^{N} X_k\right)^2 \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}} \left(\sum_{k=1}^{n} X_k\right)^2 \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega} \left(\sum_{k=1}^{n} X_k\right)^2 1_{\{N=n\}} \, d\mathbf{P} \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2 1_{\{N=n\}}\right] \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] \mathbf{E}\left[1_{\{N=n\}}\right] \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] \mathbf{P}(N=n) \right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] 1_{\{N=n\}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] &= \mathbf{E}\left[\sum_{k=1}^{n} X_k^2 + \sum_{k,\ell=1}^{n} X_k X_\ell\right] \\
&= \sum_{k=1}^{n} \mathbf{E}\left[X_k^2\right] + \sum_{k,\ell=1}^{n} \mathbf{E}[X_k]\, \mathbf{E}[X_k] \\
&= \sum_{k=1}^{n} \left(\mu^2 + \sigma^2\right) + \sum_{k,\ell=1}^{n} \mu^2 \\
&= \left(\mu^2 + \sigma^2\right) n + \mu^2 (n-1)\, n \\
&= \sigma^2 n + \mu^2 n^2.
\end{aligned}$$

Therefore,

$$\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right] = \sum_{n=1}^{\infty}\left(\sigma^2 n + \mu^2 n^2\right) 1_{\{N=n\}}$$
$$= \sigma^2 \sum_{n=1}^{\infty} n 1_{\{N=n\}} + \mu^2 \sum_{n=1}^{\infty} n^2 1_{\{N=n\}}$$
$$= \sigma^2 N + \mu^2 N^2.$$

It then follows

$$\mathbf{E}\left[\sigma^2 N + \mu^2 N^2\right] = \sigma^2 \mathbf{E}[N] + \mu^2 \mathbf{E}\left[N^2\right]$$
$$= \frac{\sigma^2}{p} + \mu^2\left(\mathbf{D}^2[N] + \mathbf{E}[N]^2\right)$$
$$= \frac{\sigma^2}{p} + \mu^2\left(\frac{2-p}{p^2}\right).$$

In the end,

$$\mathbf{D}^2\left[\sum_{k=1}^{N} X_k\right] = \frac{\sigma^2}{p} + \mu^2\left(\frac{2-p}{p^2}\right) - \frac{\mu^2}{p^2} = \frac{\sigma^2}{p} + \mu^2\left(\frac{1-p}{p^2}\right).$$

Problem 7 *Let N be a Poisson random variable with rate parameters λ and let $(X_k)_{k=1}^n$ a finite sequence of independent standard Bernoulli random variables with success parameter p, which are also independent of N. Study the conditional expectation*

$$\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right].$$

Use the properties of the conditional expectation to compute the expectation and the variance of the random sum

$$S_N \overset{def}{=} \sum_{k=1}^{N} X_k.$$

Solution. .

Problem 8 *Let B be a binomial random variable with number of trials paramenter n and success probability p, which models the number of successes in n independent trials, and let $(X_k)_{k=1}^n$ be a finite sequence of independent and exponentially distributed random variables with rate parameter λ, which are also independent of B. Study the conditional expectation*

$$\mathbf{E}\left[\sum_{k=1}^{B} X_k \mid B\right].$$

Use the properties of the conditional expectation to compute the expectation (and the variance) of the random sum

$$S_B \overset{def}{=} \sum_{k=1}^{B} X_k.$$

Solution. .

Problem 9 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Rademacher random variables*[2] *on Ω. Set $Z \overset{def}{=} X + Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with standard Rademacher random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. Since X and Y be independent standard Rademacher random variables, we have

$$Z(\omega) = (X+Y)(\omega) = \begin{cases} -2, & \text{if } \omega \in \{X=-1, Y=-1\}, \\ 0, & \text{if } \omega \in \{X=-1, Y=1\} \cup \{X=1, Y=-1\}, \\ 2, & \text{if } \omega \in \{X=1, Y=1\}. \end{cases}$$

That is to say,

$$X + Y = -2 \cdot 1_{\{X=-1,Y=-1\}} + 2 \cdot 1_{\{X=1,Y=1\}} + 0 \cdot 1_{\{X=-1,Y=1\}\cup\{X=1,Y=-1\}},$$

equivalently

$$Z = -2 \cdot 1_{\{Z=-2\}} + 2 \cdot 1_{\{Z=2\}} + 0 \cdot 1_{\{Z=0\}}.$$

Furthermore,

$$\mathbf{P}(Z=-2) = \mathbf{P}(X+Y=-2) = \mathbf{P}(X=-1, Y=-1) = \mathbf{P}(X=-1)\mathbf{P}(Y=-1) = \frac{1}{4},$$

$$\mathbf{P}(Z=2) = \mathbf{P}(X+Y=2) = \mathbf{P}(X=1, Y=1) = \mathbf{P}(X=1)\mathbf{P}(Y=1) = \frac{1}{4},$$

and

$$\mathbf{P}(Z=0) = \mathbf{P}(X+Y=0) = \mathbf{P}(\{X=-1, Y=1\} \cup \{X=1, Y=-1\})$$
$$= \mathbf{P}(X=-1, Y=1) + \mathbf{P}(X=1, Y=-1)$$
$$= \mathbf{P}(X=-1)\mathbf{P}(Y=1) + \mathbf{P}(X=1)\mathbf{P}(Y=-1)$$
$$= \frac{1}{2}.$$

1. Since Z is a discrete random variable, to compute $\mathbf{E}[X \mid Z]$ we can apply the formula

$$\mathbf{E}[X \mid Z] = \mathbf{E}[X \mid Z=-2] 1_{\{Z=-2\}} + \mathbf{E}[X \mid Z=2] 1_{\{Z=2\}} + \mathbf{E}[X \mid Z=0] 1_{\{Z=0\}},$$

where

$$\mathbf{E}[X \mid Z=-2] = \frac{1}{\mathbf{P}(Z=-2)} \int_{\{Z=2\}} X d\mathbf{P} = 4 \int_{\{X=-1,Y=-1\}} X d\mathbf{P}$$
$$= -4 \int_{\{X=-1,Y=-1\}} d\mathbf{P} = -4\mathbf{P}(X=-1, Y=-1)$$
$$= -1,$$

[2] A standard Rademacher random variable R is given by

$$R \overset{\text{def}}{=} \begin{cases} 1, & \mathbf{P}(R=1) = 1/2, \\ -1, & \mathbf{P}(R=-1) = 1/2. \end{cases}$$

$$
\begin{aligned}
\mathbf{E}\left[X \mid Z=2\right] &= \frac{1}{\mathbf{P}\left(Z=2\right)} \int_{\{Z=2\}} X d\mathbf{P} = 4 \int_{\{X=1,Y=1\}} X d\mathbf{P} \\
&= 4 \int_{\{X=1,Y=1\}} d\mathbf{P} = 4\mathbf{P}\left(X=1, Y=1\right) \\
&= 1,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{E}\left[X \mid Z=0\right] &= \frac{1}{\mathbf{P}\left(Z=0\right)} \int_{\{Z=0\}} X d\mathbf{P} = 2 \left(\int_{\{X=-1,Y=1\}\cup\{X=1,Y=-1\}} X d\mathbf{P} \right) \\
&= 2 \left(\int_{\{X=-1,Y=1\}} X d\mathbf{P} + \int_{\{X=1,Y=-1\}} X d\mathbf{P} \right) \\
&= 2\left(-1 \cdot \mathbf{P}\left(X=-1, Y=1\right) + 1 \cdot \mathbf{P}\left(X=1, Y=-1\right)\right) \\
&= 2\left(-1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4}\right) \\
&= 0.
\end{aligned}
$$

It follows

$$
\mathbf{E}\left[X \mid Z\right] = -1 \cdot 1_{\{Z=-2\}} + 1 \cdot 1_{\{Z=-2\}} + 0 \cdot 1_{\{Z=0\}} = \frac{1}{2} Z.
$$

In addition, since X and Y clearly play the same role,

$$
\mathbf{E}\left[Y \mid Z\right] = \frac{1}{2} Z.
$$

Another argument, based on the properties of the conditional expectation, is the following. Observe that

$$
Z = \mathbf{E}\left[Z \mid Z\right] = \mathbf{E}\left[X + Y \mid Z\right] = \mathbf{E}\left[X \mid Z\right] + \mathbf{E}\left[Y \mid Z\right].
$$

On the other hand, we know that

$$
\mathbf{E}\left[X \mid Z\right] = g_X\left(Z\right) \quad \text{and} \quad \mathbf{E}\left[Y \mid Z\right] = g_Y\left(Z\right)
$$

where $g_X : \mathbb{R} \to \mathbb{R}$ and $g_X : \mathbb{R} \to \mathbb{R}$ are suitable Borel functions. The structure of the function $g_X : \mathbb{R} \to \mathbb{R}$ [resp. $g_Y : \mathbb{R} \to \mathbb{R}$] depends on the joint distribution of X and Z [resp. Y and Z] and on the distribution of Z. However, in our case, it is not difficult to show that

$$
F_{X,Z}\left(u, z\right) = F_{Y,Z}\left(u, z\right),
$$

for every $(u, z) \in \mathbb{R}^2$. In fact,

$$
\begin{aligned}
& F_{X,Z}\left(u, z\right) \\
&= \mathbf{P}\left(X \le u, Z \le z\right) \\
&= \mathbf{P}\left(X \le u, X + Y \le z\right) \\
&= \mathbf{P}\left(X \le u, X + Y \le z, X = 1\right) + \mathbf{P}\left(X \le u, X + Y \le z, X = -1\right) \\
&= \mathbf{P}\left(X \le u, X + Y \le z \mid X = 1\right)\mathbf{P}\left(X = 1\right) + \mathbf{P}\left(X \le u, X + Y \le z \mid X = -1\right)\mathbf{P}\left(X = -1\right) \\
&= \frac{1}{2}\left(\mathbf{P}\left(X \le u, X + Y \le z \mid X = 1\right) + \mathbf{P}\left(X \le u, X + Y \le z \mid X = -1\right)\right) \\
&= \frac{1}{2}\left(\mathbf{P}\left(1 \le u, 1 + Y \le z \mid X = 1\right) + \mathbf{P}\left(-1 \le u, -1 + Y \le z \mid X = -1\right)\right) \\
&= \frac{1}{2}\left(\mathbf{P}\left(1 \le u, Y \le z - 1 \mid X = 1\right) + \mathbf{P}\left(-1 \le u, Y \le z + 1 \mid X = -1\right)\right) \\
&= \begin{cases}
0, & \text{if } u < -1, \\
\frac{1}{2}\mathbf{P}\left(Y \le z + 1 \mid X = -1\right) = \frac{1}{2}\mathbf{P}\left(Y \le z + 1\right), & \text{if } -1 \le u < 1, \\
\frac{1}{2}\left(\mathbf{P}\left(Y \le z - 1 \mid X = 1\right) + \mathbf{P}\left(Y \le z + 1 \mid X = -1\right)\right) = \frac{1}{2}\left(\mathbf{P}\left(Y \le z - 1\right) + \mathbf{P}\left(Y \le z + 1\right)\right), & \text{if } 1 \le u.
\end{cases}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
& F_{Y,Z}\left(u, z\right) \\
&= \begin{cases}
0, & \text{if } u < -1, \\
\frac{1}{2}\mathbf{P}\left(X \le z + 1 \mid Y = -1\right) = \frac{1}{2}\mathbf{P}\left(X \le z + 1\right), & \text{if } -1 \le u < 1, \\
\frac{1}{2}\left(\mathbf{P}\left(X \le z - 1 \mid Y = 1\right) + \mathbf{P}\left(X \le z + 1 \mid Y = -1\right)\right) = \frac{1}{2}\left(\mathbf{P}\left(X \le z - 1\right) + \mathbf{P}\left(X \le z + 1\right)\right), & \text{if } 1 \le u.
\end{cases}
\end{aligned}
$$

Therefore, on account that X and Y have the same distribution, we obtain the desired result. As a consequence, we can asses that

$$
g_X = g_Y,
$$

which implies

$$
\mathbf{E}\left[X \mid Z\right] = \mathbf{E}\left[Y \mid Z\right].
$$

It then follows

$$
2\mathbf{E}\left[X \mid Z\right] = 2\mathbf{E}\left[Y \mid Z\right] = Z,
$$

which yields

$$
\mathbf{E}\left[X \mid Z\right] = \mathbf{E}\left[Y \mid Z\right] = \frac{1}{2} Z,
$$

as expected.

2. Thanks to what shown above, we have

$$
\mathbf{E}\left[X \mid Z\right]\mathbf{E}\left[Y \mid Z\right] = \frac{1}{4} Z^2 \sim Ber\left(\frac{1}{2}\right).
$$

Hence,

$$
\mathbf{E}\left[\mathbf{E}\left[X \mid Z\right]\mathbf{E}\left[Y \mid Z\right]\right] = \frac{1}{2}.
$$

On the other hand,

$$
\mathbf{E}\left[\mathbf{E}\left[X \mid Z\right]\right] = \mathbf{E}\left[\mathbf{E}\left[Y \mid Z\right]\right] = \mathbf{E}\left[\frac{1}{2} Z\right] = \frac{1}{2}\mathbf{E}\left[Z\right] = 0.
$$

Hence, $\mathbf{E}\left[X \mid Z\right]$ and $\mathbf{E}\left[Y \mid Z\right]$ are not uncorrelated.

3. Since $\mathbf{E}\left[X \mid Z\right]$ and $\mathbf{E}\left[Y \mid Z\right]$ are not not uncorrelated, they cannot be independent.

4. By virtue of the properties of the conditional expectation, we have

$$
\mathbf{E}\left[\left(X + Y\right)^2 \mid Z\right] = \mathbf{E}\left[Z^2 \mid Z\right] = Z^2.
$$

On the other hand,

$$
\begin{aligned}
\mathbf{E}\left[\left(X + Y\right)^2 \mid Z\right] &= \mathbf{E}\left[X^2 + 2XY + Y^2 \mid Z\right] \\
&= \mathbf{E}\left[X^2 \mid Z\right] + 2\mathbf{E}\left[XY \mid Z\right] + \mathbf{E}\left[Y^2 \mid Z\right].
\end{aligned}
$$

Now, since $X \sim Y \sim Rad\left(1/2\right)$, we have $X^2 \sim Y^2 \sim Dir\left(1\right)$. We then obtain

$$
Z^2 = \mathbf{E}\left[\left(X + Y\right)^2 \mid Z\right] = \mathbf{E}\left[1 \mid Z\right] + 2\mathbf{E}\left[XY \mid Z\right] + \mathbf{E}\left[1 \mid Z\right] = 2 + 2\mathbf{E}\left[XY \mid Z\right].
$$

The latter yields

$$
\mathbf{E}\left[XY \mid Z\right] = \frac{1}{2} Z^2 - 1.
$$

Problem 10 *Let Z [resp. R] be a standard Gaussian [Rademacher] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $R \sim Rad(1/2)$. Assume that X and R are independent and define $Y \equiv R \cdot X$.*

1. *Is the random variable Y Gaussian?*
2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*
3. *Are the random variables R and Y uncorrelated? Are R and Y independent?*
4. *Does the random vector $(X,Y)^{\intercal}$ have a bivariate Gaussian distribution? Hint: consider the possibility that $(X,Y)^{\intercal}$ has a bivariate Gaussian distribution; how the random variable $Z \equiv X+Y$ should be distributed?*
5. *Can you compute $\mathbf{E}[Y \mid X]$ and $\mathbf{E}[X \mid Y]$?*

Solution. . Problema 2 del 2019

Problem 11 *Let X [resp. B] be a standard Gaussian [Bernoulli] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $B \sim Ber(1/2)$. Assume that X and B are independent and define $Y \equiv B \cdot X$. **Specifying carefully the properties used**, answer the following questions:*

1. *Is the random variable Y Gaussian? Is Y absolutely continuous?*
2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*
3. *Are the random variables B and Y uncorrelated? Are B and Y independent?*
4. *Does the random vector $(X,Y)^{\intercal}$ have a bivariate Gaussian distribution?*
5. *Can you compute $\mathbf{E}[Y \mid X]$? What about $\mathbf{E}[X \mid Y]$?*

Solution. .

Problem 12 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Rademacher random variables[3] on Ω. Set $Z \stackrel{\text{def}}{=} X+Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with standard Rademacher random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. .

[3] A standard Rademacher random variable R is given by

$$R \stackrel{\text{def}}{=} \begin{cases} 1, & \mathbf{P}(R=1)=1/2, \\ -1, & \mathbf{P}(R=-1)=1/2. \end{cases}$$

Problem 13 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Bernoulli random variables on Ω. Define $Z \stackrel{\text{def}}{=} X+Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with Bernoulli random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. .

Problem 14 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let R be a standard Rademacher random variable on Ω, and let X be a real random variable on Ω symmetric about 0 with finite second order moment. Assume that X and R are independent and define $Y \stackrel{\text{def}}{=} R \cdot X$.*

1. *Has the random varable Y the same distribution of X?*
2. *Are the random variables X and Y uncorrelated?*
3. *Are the random variables X and Y independent?*
4. *Can you compute $\mathbf{E}[Y \mid X]$?*

Solution. .

Problem 15 *Let X [resp. R] be a standard Gaussian [Rademacher] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $R \sim Rad(1/2)$. Assume that X and R are independent and define $Y \equiv R \cdot X$.*

1. *Is the random varable Y Gaussian?*
2. *Are the random variables X and Y independent?*
3. *Does the random vector $(X,Y)^{\intercal}$ have a bivariate Gaussian distribution? Hint: consider the possibility that $(X,Y)^{\intercal}$ has a bivariate Gaussian distribution; how the random variable $Z \equiv X+Y$ should be distributed?*
4. *Can you compute $\mathbf{E}[Y \mid X]$ and $\mathbf{E}[X \mid Y]$?*

Solution. .

Problem 16 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{E}$, and let X, Y be real random variables on Ω with finite second order moment.*

1. *Show that*
$$\mathbf{E}\left[(X-\mathbf{E}[X \mid \mathcal{F}])^2\right] \leq \mathbf{E}\left[(X-\mathbf{E}[X])^2\right].$$
2. *Show that*
$$\mathbf{E}[XY \mid \mathcal{F}]^2 \leq \mathbf{E}\left[X^2 \mid \mathcal{F}\right] \mathbf{E}\left[Y^2 \mid \mathcal{F}\right]. \tag{4}$$

Solution. .

1. In the space $\mathcal{L}^2(\Omega_{\mathcal{F}};\mathbb{R})$ of the real $\mathcal{F}$-random variables having finite moment of the second order the conditional expectation of X given $\mathcal{F}$ is characterized as
$$\mathbf{E}[X \mid \mathcal{F}] = \underset{Y\in\mathcal{L}^2(\Omega_{\mathcal{F}};\mathbb{R})}{\arg\min}\ \mathbf{E}\left[(X-Y)^2\right].$$
This means that
$$\mathbf{E}\left[(X-\mathbf{E}[X \mid \mathcal{F}])^2\right] \leq \mathbf{E}\left[(X-Y)^2\right],$$
for every $Y \in \mathcal{L}^2(\Omega_{\mathcal{F}};\mathbb{R})$. In particular, since the deterministic random variable $\mathbf{E}[X] \equiv \mathbf{E}[X]\ 1_\Omega$ is clearly in $\mathcal{L}^2(\Omega_{\mathcal{F}};\mathbb{R})$, setting $Y \equiv \mathbf{E}[X]$ we obtain the desired inequality.

2. Note first that for all real random variables X, Y on Ω we have
$$|XY| \leq \frac{1}{2}\left(X^2+Y^2\right).$$
Therefore, the assumption that X and Y have finite second moment implies that XY has finite first order moment. Hence, both the sides of (5) are well defined. Now, given any $z \in \mathbb{R}$, the random variable $X + zY$ has finite second order moment and, by virtue of the positivity of the conditional expectation operator, we have
$$\mathbf{E}\left[(X+zY)^2 \mid \mathcal{F}\right] \geq 0.$$
On the other hand, the linearity of the conditional expectation operator implies
$$\mathbf{E}\left[(X+zY)^2 \mid \mathcal{F}\right] = \mathbf{E}\left[X^2 \mid \mathcal{F}\right] + 2z\mathbf{E}[XY \mid \mathcal{F}] + z^2\mathbf{E}\left[Y^2 \mid \mathcal{F}\right].$$
As a consequence, we can write
$$\mathbf{E}\left[X^2 \mid \mathcal{F}\right] + 2z\mathbf{E}[XY \mid \mathcal{F}] + z^2\mathbf{E}\left[Y^2 \mid \mathcal{F}\right] \geq 0$$
for every $z \in \mathbb{R}$. It follows that
$$\Delta \equiv \mathbf{E}[XY \mid \mathcal{F}]^2 - \mathbf{E}\left[X^2 \mid \mathcal{F}\right]\mathbf{E}\left[Y^2 \mid \mathcal{F}\right] \leq 0,$$
which is the desired (5).

Problem 17 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the real Borel state space. Let $X, Y \in L^2(\Omega;\mathbb{R})$.*

1. *Prove in all details that $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω implies $Cov(X,Y) = 0$, but X and Y may not be independent.*

2. *Prove in all details that $Cov(X,Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$.*

Hint: in the first case, to generate a suitable counterexample one may consider the random variables $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$. In the second case consider $X \sim N(0,1)$ and $Y = X^2$.

Solution.

1. Under the assumption $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω, by virtue of the properties of the conditional expectation operator, we can write
$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[XY \mid X]] = \mathbf{E}[X\mathbf{E}[Y \mid X]] = \mathbf{E}[X\mathbf{E}[Y]] = \mathbf{E}[X]\,\mathbf{E}[Y]$$
Therefore,
$$Cov(X,Y) = \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = 0.$$
Now, if we consider the random $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$, we have
$$\begin{aligned} Cov(X,Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = \mathbf{E}\left[X^2 Z\right] - \mathbf{E}[X]\,\mathbf{E}[XZ] \\ &= \mathbf{E}\left[X^2\right]\mathbf{E}[Z] - \mathbf{E}[X]^2\,\mathbf{E}[Z] \\ &= 0. \end{aligned}$$
On the other hand, we have
$$\mathbf{P}(X \leq 0) = q,$$
and, on account that X and Z are independent,
$$\begin{aligned} \mathbf{P}(Y \leq 0) &= \mathbf{P}(XZ \leq 0) \\ &= \mathbf{P}(XZ \leq 0, X = 0) + \mathbf{P}(XZ \leq 0, X = 1) \\ &= \mathbf{P}(XZ \leq 0 \mid X = 0)\,\mathbf{P}(X = 0) + \mathbf{P}(XZ \leq 0 \mid X = 1)\,\mathbf{P}(X = 1) \\ &= \mathbf{P}(0 \leq 0 \mid X = 0)\,\mathbf{P}(X = 0) + \mathbf{P}(Z \leq 0 \mid X = 1)\,\mathbf{P}(X = 1) \\ &= \mathbf{P}(\Omega)\,\mathbf{P}(X = 0) + \mathbf{P}(Z \leq 0)\,\mathbf{P}(X = 1) \\ &= q + \frac{1}{2}p. \end{aligned}$$
Furthermore, the same arguments as above shows that
$$\begin{aligned} \mathbf{P}(X \leq 0, Y \leq 0) &= \mathbf{P}(X \leq 0, XZ \leq 0) \\ &= \mathbf{P}(X = 0, XZ \leq 0) \\ &= \mathbf{P}(XZ \leq 0 \mid X = 0)\,\mathbf{P}(X = 0) \\ &= q. \end{aligned}$$
Hence, we have
$$\mathbf{P}(X \leq 0)\,\mathbf{P}(Y \leq 0) = q\left(q + \frac{1}{2}p\right) \neq q = \mathbf{P}(X \leq 0, Y \leq 0)$$
which shows that X and Y are not be independent.

2. To show that $Cov(X,Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$, we consider $X \sim N(0,1)$ and $Y = X^2$. We have
$$\begin{aligned} Cov(X,Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] \\ &= \mathbf{E}\left[X^3\right] - \mathbf{E}[X]\,\mathbf{E}\left[X^2\right] \\ &= 0, \end{aligned}$$
but
$$\mathbf{E}[Y \mid X] = \mathbf{E}\left[X^2 \mid X\right] = X^2 \neq \mathbf{E}\left[X^2\right] = \mathbf{E}[Y].$$
□

Problem 18 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra, and let $X, Y \in L^2(\Omega; \mathbb{R})$. Fixed any sub-$\sigma$-algebra $\mathcal{F}$ of $\mathcal{E}$, we call conditional variance of X given $\mathcal{F}$ the random variable*

$$\mathbf{D}^2[X \mid \mathcal{F}] \stackrel{def}{=} \mathbf{E}\left[(X - \mathbf{E}[X \mid \mathcal{F}])^2 \mid \mathcal{F}\right].$$

Prove that:

1. *we have*
$$\mathbf{D}^2[X \mid \mathcal{F}] = \mathbf{E}\left[X^2 \mid \mathcal{F}\right] - \mathbf{E}[X \mid \mathcal{F}]^2;$$
2. *if X is an $\mathcal{F}$-random variable, then we have*
$$\mathbf{D}^2[X \mid \mathcal{F}] = 0;$$
3. *if X is $\mathcal{F}$-independent, then we have*
$$\mathbf{D}^2[X \mid \mathcal{F}] = \mathbf{D}^2[X];$$
4. *if X is an $\mathcal{F}$-random variable and Y is $\mathcal{F}$-independent, then we have*
$$\mathbf{D}^2[X + Y \mid \mathcal{F}] = \mathbf{D}^2[Y].$$

Solution. .

Problem 19 *Let X [resp. B] be a standard Gaussian [Bernoulli] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $B \sim Ber(1/2)$. Assume that X and B are independent and define $Y \equiv B \cdot X$.* ***Specifying carefully the properties used,*** *answer the following questions:*

1. *Is the random variable Y Gaussian? Is Y absolutely continuous?*
2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*
3. *Are the random variables B and Y uncorrelated? Are B and Y independent?*
4. *Does the random vector $(X, Y)^\intercal$ have a bivariate Gaussian distribution?*
5. *Can you compute $\mathbf{E}[Y \mid X]$? What about $\mathbf{E}[X \mid Y]$?*

Solution.

1. To check whether Y is Gaussian we compute $\mathbf{P}(Y \leq y)$, for every $y \in \mathbb{R}$. On account that $\{B = 0\}$, $\{B = 1\}$ constitute a partition of Ω, the random variables B and X are independent, and $X \sim N(0,1)$, we can write
$$\begin{aligned}
\mathbf{P}(Y \leq y) &= \mathbf{P}(BX \leq y) \\
&= \mathbf{P}(BX \leq y, B = 0) + \mathbf{P}(BX \leq y, B = 1) \\
&= \mathbf{P}(BX \leq y \mid B = 0)\,\mathbf{P}(B = 0) + \mathbf{P}(BX \leq y \mid B = 1)\,\mathbf{P}(B = 1) \\
&= \frac{1}{2}\left(\mathbf{P}(0 \leq y \mid R = 0) + \mathbf{P}(X \leq y \mid R = 1)\right) \\
&= \frac{1}{2}\left(\mathbf{P}(0 \leq y) + \mathbf{P}(X \leq y)\right) \\
&= \begin{cases} \frac{1}{2}\mathbf{P}(X \leq y) = \frac{1}{2}F_X(x), & \text{if } y < 0, \\ \frac{1}{2}(1 + \mathbf{P}(X \leq y)) = \frac{1}{2}(1 + F_X(x)) & \text{if } 0 \leq y. \end{cases}
\end{aligned}$$
Now, we have
$$\lim_{y \to 0^-} \mathbf{P}(Y \leq y) = \lim_{y \to 0^-} \frac{1}{2} F_X(x) = \frac{1}{4}$$
and
$$\lim_{y \to 0^+} \mathbf{P}(Y \leq y) = \lim_{y \to 0^+} \frac{1}{2}(1 + F_X(x)) = \frac{3}{4}$$
This proves that Y is not Gaussian.

2. Since $Y \equiv B \cdot X$, the intuition is that the observation of the values taken by X transmits information on the values taken by Y. That is X and Y are not independent. In fact, thanks to the independence of X and B and on account that $X \sim N(0,1)$, we have
$$\mathbf{E}[XY] = \mathbf{E}[XBX] = \mathbf{E}\left[BX^2\right] = \mathbf{E}[B]\,\mathbf{E}\left[X^2\right] = \frac{1}{2} \cdot 1 = \frac{1}{2}.$$
On the other hand,
$$\mathbf{E}[X]\,\mathbf{E}[Y] = 0.$$
Hence,
$$\mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = \frac{1}{2}$$
This shows that X and Y are correlated, which prevents that X^2 and Y^2 are independent.

3. Still thanks to the independence of X and B and on account that $X \sim N(0,1)$ and $B^2 \sim Ber(1/2)$, we have
$$\mathbf{E}[BY] = \mathbf{E}[BBX] = \mathbf{E}\left[B^2 X\right] = \mathbf{E}\left[B^2\right]\mathbf{E}[X] = \frac{1}{2}\mathbf{E}[X] = 0.$$
On the other hand,
$$\mathbf{E}[B]\,\mathbf{E}[Y] = \mathbf{E}[B]\,\mathbf{E}[BX] = \mathbf{E}[B]\,\mathbf{E}[B]\,\mathbf{E}[X] = \frac{1}{4}\mathbf{E}[X] = 0.$$
This shows that B and Y are uncorrelated. Here, the intuition is that the observation of the values taken by B transmits information on the values taken by Y. Hence, the intuition is that B and Y are not independent. To prove this, we show that B^2 and Y^2 are correlated. In fact, thanks to the independence of X and B and on account that $B^4 \sim B^2 \sim Ber(1/2)$, we have
$$\mathbf{E}\left[B^2 Y^2\right] = \mathbf{E}\left[B^2 B^2 X^2\right] = \mathbf{E}\left[B^4 X^2\right] = \mathbf{E}[B]\,\mathbf{E}\left[X^2\right] = \frac{1}{2} \cdot 1 = \frac{1}{2}$$
and
$$\mathbf{E}\left[B^2\right]\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[B^2\right]\mathbf{E}\left[B^2 X^2\right] = \mathbf{E}\left[B^2\right]\mathbf{E}\left[B^2\right]\mathbf{E}\left[X^2\right] = \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}.$$
This shows that B^2 and Y^2 are correlated, which prevents that B^2 and Y^2 are independent. Eventually, B and Y cannot be independent.

4. Since $B \sim Ber(1/2)$ is independent of X, we have
$$\mathbf{E}[Y \mid X] = \mathbf{E}[BX \mid X] = X\mathbf{E}[B \mid X] = \mathbf{E}[B]\,X = \frac{1}{2}X$$
$$\mathbf{E}[X \mid Y] = \mathbf{E}[X \mid BX]$$

By virtue of what shown above and the properties of the conditional expectation, on account that B and X are independent, we have,

$$\mathbf{E}[Y \mid X] = \mathbf{E}[BX \mid X] = X\mathbf{E}[B \mid X] = X\mathbf{E}[B] = \frac{1}{2}X.$$

Now, to compute $\mathbf{E}[X \mid Y]$, observe preliminarly that

$$B(\omega) = 1_{\{B=1\}}(\omega)$$

for every $\omega \in \Omega$. In fact,

$$1_{\{B=1\}}(\omega) = \begin{cases} 1 & \text{if } B(\omega) = 1 \\ 0 & \text{if } B(\omega) = 0 \end{cases}$$

Hence,

$$Y = BX = X1_{\{B=1\}}.$$

As a consequence, on account that

$$1_{\{B=1\}} + 1_{\{B=0\}} = 1_\Omega,$$

we have

$$\begin{aligned}\mathbf{E}[X \mid Y] &= \mathbf{E}\left[X\left(1_{\{B=1\}} + 1_{\{B=0\}}\right) \mid Y\right] \\ &= \mathbf{E}\left[X1_{\{B=1\}} \mid Y\right] + \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] \\ &= \mathbf{E}[Y \mid Y] + \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] \\ &= Y + \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] \\ &= BX + \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right].\end{aligned}$$

Thus, we are left with computing

$$\mathbf{E}\left[X1_{\{B=0\}} \mid Y\right].$$

To this goal, observe that

$$\int_{\{Y\in C\}} \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] d\mathbf{P}_{|\sigma(Y)} = \int_{\{Y\in C\}} X1_{\{B=0\}} d\mathbf{P} = \int_{\{Y\in C\}\cap\{B=0\}} X d\mathbf{P},$$

where

$$\{Y \in C\} \cap \{B = 0\} = \{BX \in C\} \cap \{B = 0\} = \begin{cases} \{B = 0\}, & \text{if } 0 \in C, \\ \varnothing, & \text{if } 0 \notin C. \end{cases}$$

Hence,

$$\int_{\{Y\in C\}} \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] d\mathbf{P}_{|\sigma(Y)} = \begin{cases} \int_{\{B=0\}} X d\mathbf{P}, & \text{if } 0 \in C, \\ 0, & \text{if } 0 \notin C, \end{cases}$$

for every $C \in \mathcal{B}(\mathbb{R})$. On the other hand,

$$\int_{\{B=0\}} X d\mathbf{P} = \mathbf{E}\left[X1_{\{B=0\}}\right] = \mathbf{E}[X]\,\mathbf{E}\left[1_{\{B=0\}}\right] = \mathbf{E}[X]\,\mathbf{P}(B = 0) = \frac{1}{2}\mathbf{E}[X].$$

It then follows,

$$\int_{\{Y\in C\}} \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] d\mathbf{P}_{|\sigma(Y)} = \begin{cases} \frac{1}{2}\mathbf{E}[X], & \text{if } 0 \in C, \\ 0, & \text{if } 0 \notin C. \end{cases}$$

We then claim that

$$\mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] = \mathbf{E}[X]\,1_{\{Y=0\}}, \ \mathbf{P}_{|\sigma(Y)}\text{-a.s. on } \Omega.$$

In fact, we have

$$\{Y = 0\} = \{XB = 0\} = \{X \in \mathbb{R}, B = 0\} \cup \{X = 0, B = 1\},$$

where

$$\{X \in \mathbb{R}, B = 0\} \cap \{X = 0, B = 1\} = \varnothing.$$

This, on account of the independence of X and B and that X is a continuous random variable, implies

$$\begin{aligned}\mathbf{P}(Y = 0) &= \mathbf{P}(X \in \mathbb{R}, B = 0) + \mathbf{P}(X = 0, B = 1) \\ &= \mathbf{P}(X \in \mathbb{R})\,\mathbf{P}(B = 0) + \mathbf{P}(X = 0)\,\mathbf{P}(B = 1) \\ &= \mathbf{P}(B = 0) \\ &= \frac{1}{2}.\end{aligned}$$

In addition, since

$$\{Y \in C\} \cap \{Y = 0\} = \begin{cases} \{Y = 0\} & \text{if } 0 \in C \\ \varnothing & \text{if } 0 \notin C \end{cases},$$

for every $C \in \mathcal{B}(\mathbb{R})$, we have

$$\begin{aligned}\int_{\{Y\in C\}} \mathbf{E}[X]\,1_{\{Y=0\}} d\mathbf{P}_{|\sigma(Y)} &= \mathbf{E}[X] \int_{\{Y\in C\}\cap\{Y=0\}} d\mathbf{P}_{|\sigma(Y)} \\ &= \mathbf{E}[X]\,\mathbf{P}(\{Y \in C\} \cap \{Y = 0\}) \\ &= \begin{cases} \mathbf{E}[X]\,\mathbf{P}(Y = 0) & \text{if } 0 \in C \\ 0 & \text{if } 0 \notin C \end{cases} \\ &= \begin{cases} \frac{1}{2}\mathbf{E}[X] & \text{if } 0 \in C \\ 0 & \text{if } 0 \notin C \end{cases},\end{aligned}$$

for every $C \in \mathcal{B}(\mathbb{R})$. From what shown above, we can write

$$\int_{\{Y\in C\}} \mathbf{E}\left[X1_{\{B=0\}} \mid Y\right] d\mathbf{P}_{|\sigma(Y)} = \int_{\{Y\in C\}} \mathbf{E}[X]\,1_{\{Y=0\}} d\mathbf{P}_{|\sigma(Y)},$$

for every $C \in \mathcal{B}(\mathbb{R})$. This implies our claim. Summarizing we obtain

$$\mathbf{E}[X \mid Y] = BX + \mathbf{E}[X]\,1_{\{Y=0\}}, \quad \mathbf{P}_{|\sigma(Y)}\text{-a.s. on } \Omega,$$

which completes the solution of 5.

Problem 20 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let $B \sim Ber(1/2)$ [resp. $R \sim Rad(1/2)$] a standard Bernoulli [resp. Rademacher] random variable on Ω. Assume that B and R are independent and define $X \overset{def}{=} B + R$.*

1. *Compute $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$, $\mathbf{E}[B \mid X]$, and $\mathbf{E}[R \mid X]$. In addition,* ***specifying carefully the properties used,*** *answer the following questions:*
2. *Are the random variables $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$ uncorrelated? Are $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$ independent?*
3. *Are the random variables $\mathbf{E}[B \mid X]$ and $\mathbf{E}[R \mid X]$ uncorrelated? Are $\mathbf{E}[B \mid X]$ and $\mathbf{E}[R \mid X]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with a Bernoulli and a Rademacher random variable, can you compute $\mathbf{E}[BR \mid X]$?*

Solution. .

Problem 21 *Let U, V real random variables on a probability space Ω such that such that $U \sim V \sim N(0,1)$, the vector $(U,V)^{\intercal}$ is Gaussian, and $Corr(U,V) \equiv \rho < 1$. Consider the real random variables*

$$X \stackrel{def}{=} U - \rho V \quad \text{and} \quad Y \stackrel{def}{=} \sqrt{1-\rho} V.$$

1. *Can you prove that the vector $(X,Y)^{\intercal}$ Gaussian?*
2. *Are the random variables X and Y independent?*
3. *Compute the distributions of X and Y;*
4. *Compute $\mathbf{E}\left[X^2Y^2\right]$, $\mathbf{E}\left[XY^3\right]$, $\mathbf{E}\left[Y^4\right]$.*
5. *Compute $\mathbf{E}\left[U^2V^2\right]$.*

Solution. .

Problem 22 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let B_1 and B_2 be standard Bernoulli random variables on Ω. In symbols, $B_k \sim Ber(1/2)$, for $k = 1, 2$. Assume that B_1 and B_2 are independent and set*

$$X \stackrel{def}{=} B_1 + B_2, \quad Y \stackrel{def}{=} B_1 \cdot B_2$$

1. *Compute $\mathbf{E}[B_k \mid X]$ and $\mathbf{E}[B_k \mid Y]$ for $k = 1, 2$.*
2. *Are the random variables $\mathbf{E}[B_1 \mid X]$ and $\mathbf{E}[B_2 \mid X]$ uncorrelated? Are they independent?*
3. *Are the random variables $\mathbf{E}[B_1 \mid Y]$ and $\mathbf{E}[B_2 \mid Y]$ uncorrelated? Are they independent?*
4. *Compute $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$.*
5. *Are the random variables $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$ uncorrelated? Are they independent?*
6. *Compute $\mathbf{E}\left[X^2 \mid Y\right]$ and $\mathbf{E}\left[Y^2 \mid X\right]$.*

Solution. .

Problem 23 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Bernoulli random variables on Ω. Define $Z \stackrel{def}{=} X + Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with Bernoulli random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. .

Problem 24 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let R_1 and R_2 be standard Rademacher random variables on Ω. In symbols, $R_k \sim Rad(1/2)$, for $k = 1, 2$. Assume that R_1 and R_2 are independent and set*

$$X \stackrel{def}{=} R_1 - R_2, \quad Y \stackrel{def}{=} -R_1 \cdot R_2$$

1. *Compute $\mathbf{E}[R_k \mid X]$ and $\mathbf{E}[R_k \mid Y]$ for $k = 1, 2$.*
2. *Are the random variables $\mathbf{E}[R_1 \mid X]$ and $\mathbf{E}[R_2 \mid X]$ uncorrelated? Are they independent?*
3. *Are the random variables $\mathbf{E}[R_1 \mid Y]$ and $\mathbf{E}[R_2 \mid Y]$ uncorrelated? Are they independent?*
4. *Compute $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$.*
5. *Are the random variables $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$ uncorrelated? Are they independent?*
6. *Compute $\mathbf{E}\left[X^2 \mid Y\right]$ and $\mathbf{E}\left[Y^2 \mid X\right]$.*

Solution. .

Problem 25 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Rademacher random variables[4] on Ω. Set $Z \stackrel{def}{=} X + Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with standard Rademacher random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. .

Problem 26 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Bernoulli random variables on Ω. Define $Z \stackrel{def}{=} X + Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with Bernoulli random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. .

[4] A standard Rademacher random variable R is given by

$$R \stackrel{\text{def}}{=} \begin{cases} 1, & \mathbf{P}(R=1) = 1/2, \\ -1, & \mathbf{P}(R=-1) = 1/2. \end{cases}$$

Problem 27 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let R be a standard Rademacher random variable on Ω, and let X be a real random variable on Ω symmetric about 0 with finite second order moment. Assume that X and R are independent and define $Y \overset{def}{=} R \cdot X$.*

1. *Has the random varable Y the same distribution of X?*
2. *Are the random variables X and Y uncorrelated?*
3. *Are the random variables X and Y independent?*
4. *Can you compute $\mathbf{E}[Y \mid X]$?*

Solution. .

Problem 28 *Let X [resp. R] be a standard Gaussian [Rademacher] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $R \sim Rad(1/2)$. Assume that X and R are independent and define $Y \equiv R \cdot X$.*

1. *Is the random varable Y Gaussian?*
2. *Are the random variables X and Y independent?*
3. *Does the random vector $(X, Y)^{\intercal}$ have a bivariate Gaussian distribution? Hint: consider the possibility that $(X, Y)^{\intercal}$ has a bivariate Gaussian distribution; how the random variable $Z \equiv X + Y$ should be distributed?*
4. *Can you compute $\mathbf{E}[Y \mid X]$ and $\mathbf{E}[X \mid Y]$?*

Solution. .

Problem 29 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{E}$, and let X, Y be real random variables on Ω with finite second order moment.*

1. *Show that*
$$\mathbf{E}\left[(X - \mathbf{E}[X \mid \mathcal{F}])^2\right] \leq \mathbf{E}\left[(X - \mathbf{E}[X])^2\right].$$
2. *Show that*
$$\mathbf{E}[XY \mid \mathcal{F}]^2 \leq \mathbf{E}\left[X^2 \mid \mathcal{F}\right] \mathbf{E}\left[Y^2 \mid \mathcal{F}\right]. \tag{5}$$

Solution. .

Random Vectors

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Random Vectors with Solution 2022-12-08

Problem 1 *Let (X_1, X_2) a real random vector with a joint density $f_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}$ given by*
$$f_{X_1,X_2}(x_1, x_2) \overset{def}{=} 1_{[0,1]\times[0,1]}(x_1, x_2), \quad \forall (x_1, x_2) \in \mathbb{R}^2.$$
Consider the real random variables $Y \equiv \min(X_1, X_2)$ and $Z \equiv \max(X_1, X_2)$. Determine:

1. *the distribution functions of Y and Z;*
2. *the joint distribution function of Y and Z;*
3. *the marginal distributions functions of Y and Z;*
4. *the expectations of Y and Z.*

Solution.

1. We have
$$1_{[0,1]\times[0,1]}(x_1, x_2) = 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2),$$
for every $(x_1, x_2) \in \mathbb{R}^2$. As a consequence, for the marginal density $f_{X_2} : \mathbb{R} \to \mathbb{R}$ [resp. $f_{X_2} : \mathbb{R} \to \mathbb{R}$] of the entry X_1 [resp. X_2] of the random vector $(X_1, X_2)^{\intercal}$, we obtain
$$\begin{aligned} f_{X_1}(x_1) &= \int_{\mathbb{R}} 1_{[0,1]\times[0,1]}(x_1, x_2)\, d\mu_L(x_2) = \int_{\mathbb{R}} 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, d\mu_L(x_2) \\ &= 1_{[0,1]}(x_1) \int_{\mathbb{R}} 1_{[0,1]}(x_2)\, d\mu_L(x_2) = 1_{[0,1]}(x_1) \int_{[0,1]} d\mu_L(x_2) = 1_{[0,1]}(x_1)\, \mu_L([0,1]) \\ &= 1_{[0,1]}(x_1), \end{aligned}$$
for every $x_1 \in \mathbb{R}$ [resp.
$$f_{X_2}(x_2) = 1_{[0,1]}(x_2),$$
for every $x_2 \in \mathbb{R}$]. It then follows,
$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2),$$
for every $(x_1, x_2) \in \mathbb{R}^2$. Hence, the entries X_1 and X_2 of the random vector (X_1, X_2) are independent random variables, and both are standard uniformly distributed. Now, we have
$$\{Y \leq y\} = \{X_1 \leq y, X_2 \leq y\} \cup \{X_1 > y, X_2 \leq y\} \cup \{X_1 \leq y, X_2 > y\},$$
for every $y \in \mathbb{R}$, where the three events on the right hand side are pairwise incompatible, and
$$\{Z \leq z\} = \{X_1 \leq z, X_2 \leq z\},$$
for every $z \in \mathbb{R}$. By virtue of the indepencence of X_1 and X_2, it then follows,
$$\begin{aligned} &F_Y(y) \\ &= \mathbf{P}(Y \leq y) = \mathbf{P}(X_1 \leq y, X_2 \leq y) + \mathbf{P}(X_1 > y, X_2 \leq y) + \mathbf{P}(X_1 \leq y, X_2 > y) \\ &= \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 > y)\,\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 > y) \\ &= \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) + (1 - \mathbf{P}(X_1 \leq y))\,\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y)\,(1 - \mathbf{P}(X_2 \leq y)) \\ &= \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_2 \leq y) - \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y) - \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) \\ &= \mathbf{P}(X_1 \leq y) + \mathbf{P}(X_2 \leq y) - \mathbf{P}(X_1 \leq y)\,\mathbf{P}(X_2 \leq y) \\ &= F_{X_1}(y) + F_{X_2}(y) - F_{X_1}(y)\, F_{X_2}(y) \end{aligned}$$

and

$$F_Z(z) = \mathbf{P}(X_1 \le z, X_2 \le z) = \mathbf{P}(X_1 \le z)\,\mathbf{P}(X_2 \le z) = F_{X_1}(z)\,F_{X_2}(z).$$

Note that instead of the event $\{Y \le y\}$ we could have considered the event

$$\{Y > y\} = \{X_1 > y, X_2 > y\},$$

for every $y \in \mathbb{R}$, obtaining

$$\begin{aligned}
F_Y(y) &= \mathbf{P}(Y \le y) = 1 - \mathbf{P}(Y > y) = 1 - \mathbf{P}(X_1 > y, X_2 > y) \\
&= 1 - \mathbf{P}(X_1 > y)\,\mathbf{P}(X_2 > y) = 1 - (1 - \mathbf{P}(X_1 \le y))(1 - \mathbf{P}(X_2 \le y)) \\
&= 1 - (1 - F_{X_1}(y))((1 - F_{X_2}(y))) \\
&= 1 - (1 - F_{X_2}(y) - F_{X_1}(y) + F_{X_1}(y)\,F_{X_2}(y)) \\
&= F_{X_1}(y) + F_{X_2}(y) - F_{X_1}(y)\,F_{X_2}(y),
\end{aligned}$$

for every $y \in \mathbb{R}$, as above. On the other hand, both the random variables X_1 and X_2 are standard uniformly distributed on the interval $[0,1]$. Therefore,

$$F_Y(y) = F_X(y)(2 - F_X(y)) \qquad \text{and} \qquad F_Z(z) = F_X(z)^2,$$

for all $y, x \in \mathbb{R}$, where F_X is the distribution function of the random variable $X\tilde{\ }Unif(0,1)$, given by

$$F_X(x) = x \cdot 1_{[0,1]}(x) + 1_{(1,+\infty)}(x),$$

for every $x \in \mathbb{R}$. It then follows

$$\begin{aligned}
&F_Y(y) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,+\infty)}(y) - \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\right) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,0)}(y) + 2 \cdot 1_{[0,1]}(y) + 2 \cdot 1_{(1,+\infty)}(y) - \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\right) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,0)}(y) + (2 - y) \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right) \\
&= (2 - y)\,y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y),
\end{aligned}$$

for every $y \in \mathbb{R}$, and

$$F_Z(z) = \left(z \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z)\right)^2 = z^2 \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z).$$

for every $z \in \mathbb{R}$. Note that we have

$$F'_Y(y) = 2(1 - y) \cdot 1_{(0,1)}(y) \qquad \text{and} \qquad F'_Z(z) = 2z \cdot 1_{(0,1)}(z),$$

for every $y, z \in \mathbb{R} - \{0,1\}$. These imply

$$\begin{aligned}
\int_{(-\infty,y)} F'_Y(u)\,d\mu_L(u) &= \int_{(-\infty,y)} 2(1-u)\,1_{(0,1)}(u)\,d\mu_L(u) \\
&= \begin{cases} 0, & \text{if } y \le 0, \\ \int_{(0,y)} 2(1-u)\,d\mu_L(u), & \text{if } 0 < y < 1, \\ \int_{(0,1)} 2(1-u)\,d\mu_L(u), & \text{if } 1 \le y, \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
\int_{(-\infty,z)} F'_Z(v)\,d\mu_L(v) &= \int_{(-\infty,z)} 2z \cdot 1_{(0,1)}(z)\,d\mu_L(v) \\
&= \begin{cases} 0, & \text{if } z \le 0, \\ \int_{(0,z)} 2v d\mu_L(v), & \text{if } 0 < z < 1, \\ \int_{(0,1)} 2v d\mu_L(v), & \text{if } 1 \le z. \end{cases}
\end{aligned}$$

On the other hand,

$$\int_{(0,y)} 2(1-u)\,d\mu_L(u) = \int_0^y 2(1-u)\,du = 2u - u^2\Big|_0^y = y(2-y),$$

for every $0 < y \le 1$, and

$$\int_{(0,z)} 2v d\mu_L(v) = \int_0^z 2v dv = v^2\Big|_0^z = z^2,$$

for every $0 < z \le 1$. We can then write

$$\int_{(-\infty,y)} F'_Y(u)\,d\mu_L(u) = y(2-y) \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y) = F_Y(y),$$

for every $y \in \mathbb{R}$, and

$$\int_{(-\infty,z)} F'_Z(v)\,d\mu_L(v) = z^2 \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z) = F_Z(z),$$

for every $z \in \mathbb{R}$. These imply that Y and Z are absolutely continuous random variables.

2. We have

$$\begin{aligned}
&\{Y \le y, Z \le z\} \\
&= (\{X_1 \le y, X_2 \le y\} \cup \{X_1 > y, X_2 \le y\} \cup \{X_1 \le y, X_2 > y\}) \cap \{X_1 \le z, X_2 \le z\} \\
&= (\{X_1 \le y, X_2 \le y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&\quad \cup (\{X_1 > y, X_2 \le y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&\quad \cup (\{X_1 \le y, X_2 > y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&= \{X_1 \le \min(y,z), X_2 \le \min(y,z)\} \\
&\quad \cup \{y < X_1 \le z, X_2 \le \min(y,z)\} \\
&\quad \cup \{X_1 \le \min(y,z), y < X_2 \le z\}.
\end{aligned}$$

Therefore, considering the joint distribution function $F_{Y,Z} : \mathbb{R}^2 \to \mathbb{R}$ of Y and Z, on account of the independence of X_1 and X_2, we can write

$$\begin{aligned}
F_{Y,Z}(y,z) &= \mathbf{P}(Y \le y, Z \le z) \\
&= \mathbf{P}(X_1 \le \min(y,z), X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(y < X_1 \le z, X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(X_1 \le \min(y,z), y < X_2 \le z) \\
&= \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(y < X_1 \le z)\,\mathbf{P}(X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(y < X_2 \le z),
\end{aligned}$$

for every $(y,z) \in \mathbb{R}^2$. On the other hand,

$$\begin{array}{lll} \min(y,z) = y, & & \text{if } y \le z, \\ \mathbf{P}(y < X_1 \le z) = 0 & \text{and} \quad \min(y,z) = z, & \text{if } y > z. \end{array}$$

Hence, considerning that X_1 and X_2 have the same distribution, we obtain

$$F_{Y,Z}(y,z) = \begin{cases} F_X(y)(2F_X(z) - F_X(y)), & \text{if } y \le z, \\ F_X(z)^2, & \text{if } y > z. \end{cases}$$

In fact, if $y \le z$

$$\begin{aligned}
&\mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(X_2 \le \min(y,z)) + \mathbf{P}(y < X_1 \le z)\,\mathbf{P}(X_2 \le \min(y,z)) \\
&+ \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(y < X_2 \le z) \\
&= \mathbf{P}(X \le y)\,\mathbf{P}(X \le y) + 2\mathbf{P}(X \le y)\,\mathbf{P}(y < X \le z) \\
&= F_X(y)^2 + 2F_X(y)(F_X(z) - F_X(y)) \\
&= F_X(y)(2F_X(z) - F_X(y))
\end{aligned}$$

and if $y > z$

$$\begin{aligned}
&\mathbf{P}\left(X_1 \leq \min\left(y,z\right)\right)\mathbf{P}\left(X_2 \leq \min\left(y,z\right)\right) + \mathbf{P}\left(y < X_1 \leq z\right)\mathbf{P}\left(X_2 \leq \min\left(y,z\right)\right)\\
&+\mathbf{P}\left(X_1 \leq \min\left(y,z\right)\right)\mathbf{P}\left(y < X_2 \leq z\right)\\
&=\mathbf{P}\left(X \leq z\right)\mathbf{P}\left(X \leq z\right) + 2\mathbf{P}\left(X \leq z\right)\mathbf{P}\left(y < X \leq z\right)\\
&=F_X\left(z\right)^2.
\end{aligned}$$

Note that we can write

$$F_{Y,Z}\left(y,z\right) = F_X\left(y\right)\left(2F_X\left(z\right) - F_X\left(y\right)\right)1_{\{(y,z)\in\mathbb{R}^2:y\leq z\}} + F_X\left(z\right)^2 1_{\{(y,z)\in\mathbb{R}^2:y>z\}}.$$

3. To determine the marginal distribution functions $F_Y : \mathbb{R} \to \mathbb{R}$ and $F_Z : \mathbb{R} \to \mathbb{R}$ of the random vector $(Y,Z)^\intercal$, respectively, we can apply the formula

$$\begin{aligned}
F_Y\left(y\right) &= \lim_{z\to+\infty} F_{Y,Z}\left(y,z\right)\\
&= \lim_{z\to+\infty}\Big(F_X\left(y\right)\left(2F_X\left(z\right) - F_X\left(y\right)\right)1_{\{(y,z)\in\mathbb{R}^2:y\leq z\}}\left(y,z\right) + F_X\left(z\right)^2 1_{\{(y,z)\in\mathbb{R}^2:y>z\}}\left(y,z\right)\Big)
\end{aligned}$$

and

$$\begin{aligned}
F_Z\left(z\right) &= \lim_{y\to+\infty} F_{Y,Z}\left(y,z\right) =\\
&= \lim_{y\to+\infty}\Big(F_X\left(y\right)\left(2F_X\left(z\right) - F_X\left(y\right)\right)1_{\{(y,z)\in\mathbb{R}^2:y\leq z\}}\left(y,z\right) + F_X\left(z\right)^2 1_{\{(y,z)\in\mathbb{R}^2:y>z\}}\left(y,z\right)\Big).
\end{aligned}$$

as $z \to +\infty$ for every $y \in \mathbb{R}$ we have

$$1_{\{(y,z)\in\mathbb{R}^2:y\leq z\}}\left(y,z\right) = 1 \qquad \text{and} \qquad 1_{\{(y,z)\in\mathbb{R}^2:y>z\}}\left(y,z\right) = 0.$$

Conversely, as $y \to +\infty$ for every $z \in \mathbb{R}$ we have

$$1_{\{(y,z)\in\mathbb{R}^2:y\leq z\}}\left(y,z\right) = 0 \qquad \text{and} \qquad 1_{\{(y,z)\in\mathbb{R}^2:y>z\}}\left(y,z\right) = 1.$$

It then follows

$$F_Y\left(y\right) = F_X\left(y\right)\left(2F_X\left(z\right) - F_X\left(y\right)\right) \qquad \text{and} \qquad F_Z\left(z\right) = F_X\left(z\right)^2,$$

which shows that the marginal distribution functions of the random vector (Y,Z) concide with the distribution functions of the random variables X and Y. As a consequence, the random variables $Y \equiv \min\left(X_1,X_2\right)$ and $Z \equiv \max\left(X_1,X_2\right)$ are indepemdent.

4. In the end, we have

$$\begin{aligned}
\mathbf{E}\left[Y\right] &= \int_{\mathbb{R}} y f_Y\left(y\right) d\mu_L\left(y\right) = \int_{\mathbb{R}} 2y\left(1-y\right)1_{[0,1]}\left(y\right) d\mu_L\left(y\right) = \int_{[0,1]} 2y\left(1-y\right) d\mu_L\left(y\right)\\
&= \int_0^1 2\left(1-y\right) y dy = 2\left(\int_0^1 y dy - \int_0^1 y^2 dy\right) = 2\left(\frac{1}{2}y^2\Big|_0^1 - \frac{1}{3}y^3\Big|_0^1\right) = \frac{1}{3}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}\left[Z\right] &= \int_{\mathbb{R}} z f_Z\left(z\right) d\mu_L\left(z\right) = \int_{\mathbb{R}} 2z^2 \cdot 1_{[0,1]}\left(z\right) d\mu_L\left(z\right) = \int_{[0,1]} 2z^2 d\mu_L\left(z\right)\\
&= \int_0^1 2z^2 dz = 2\int_0^1 z^2 dz = 2\,\frac{1}{3}z^3\Big|_0^1 = \frac{2}{3}.
\end{aligned}$$

Problem 2 *Let $(\Omega,\mathcal{E},\mathbf{P}) \equiv \Omega$ be a probability space and let $\left(\mathbb{R}^2,\mathcal{B}\left(\mathbb{R}^2\right),\mu_L^2\right) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra and the Borel-Lebesgue measure $\mu_L^2 : \mathcal{B}\left(\mathbb{R}^2\right)$. Let $f : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f\left(x,y\right) \stackrel{def}{=} kxye^{-\left(x^2+y^2\right)}1_{\mathbb{R}_+^2}\left(x,y\right), \quad \forall\left(x,y\right)\in\mathbb{R}^2$$

where $\mathbb{R}_+^2 \equiv \left\{(x,y)\in\mathbb{R}^2 : x \geq 0, y \geq 0\right\}$. Determine $k \in \mathbb{R}$ such that $f : \mathbb{R}^2 \to \mathbb{R}_+$ is a probability density and let $Z \equiv (X,Y)$ be the random vector of density $f : \mathbb{R}^2 \to \mathbb{R}_+$.

1. *Determine the distribution function $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ of the vector Z and check that*

$$\frac{\partial F^2}{\partial x\partial y}\left(x,y\right) = f\left(x,y\right), \quad \mu_L^2 - a.e.\ on\ \mathbb{R}^2.$$

2. *Determine the marginal distribution function $F_X : \mathbb{R} \to \mathbb{R}_+$ and $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z.*

3. *Determine the densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z and check that*

$$\frac{dF_X}{dx}\left(x\right) = f_X\left(x\right) \quad and \quad \frac{dF_Y}{dy}\left(y\right) = f_Y\left(y\right), \quad \mu_L - a.e.\ on\ \mathbb{R}.$$

4. *Are X and Y independent random varables?*

5. *Compute $\mathbf{E}\left[X\right]$, $\mathbf{E}\left[Y\right]$, $\mathbf{D}^2\left[X\right]$, $\mathbf{D}^2\left[Y\right]$ and $Cov\left(X,Y\right)$.*

6. *Compute $\mathbf{E}\left[(X,Y)\right]$ and the covariance matrix of the vector (X,Y).*

Solution. . □

Problem 3 *Determine the value of the parameter k such that the function $f : \mathbb{R}^3 \to \mathbb{R}$ given by*

$$f\left(x_1,x_2,x_3\right) \stackrel{def}{=} \begin{cases} k\left(x_1 + x_2^2 + x_3^3\right), & \text{if } \left(x_1,x_2,x_3\right)\in[0,1]\times[0,1]\times[0,1],\\ 0, & \text{otherwise,}\end{cases}$$

is a probability density. Hence, consider the random vector $(X_1,X_2,X_3)^\intercal$ with density $f_{X_1,X_2,X_3} : \mathbb{R}^3 \to \mathbb{R}$ given by

$$f_{X_1,X_2,X_3}\left(x_1,x_2,x_3\right) \stackrel{def}{=} f\left(x_1,x_2,x_3\right).$$

Compute:

1. *the probability $\mathbf{P}\left(X_2 \leq 1/2, X_3 > 1/2\right)$;*

2. *the marginal densities of the random vector $(X_1,X_2)^\intercal$;*

3. *the expectation of $(X_1,X_2)^\intercal$;*

4. *the conditional density $f_{X_1,X_2|X_3=1/2}\left(x_1,x_2\right)$.*

Solution. To determine the value of the parameter k such that the function $f : \mathbb{R}^3 \to \mathbb{R}$ is a probability density we have to solve the equation

$$\int_{\mathbb{R}^3} f\left(x_1,x_2,x_3\right) d\mu_L\left(x_1,x_2,x_3\right) = 1.$$

We have

$$f\left(x_1,x_2,x_3\right) = k\left(x_1 + x_2^2 + x_3^3\right)1_{[0,1]\times[0,1]\times[0,1]}\left(x_1,x_2,x_3\right),$$

Hence,

$$\begin{aligned}
\int_{\mathbb{R}^3} f\left(x_1,x_2,x_3\right) d\mu_L\left(x_1,x_2,x_3\right) &= \int_{\mathbb{R}^3} k\left(x_1 + x_2^2 + x_3^3\right)1_{[0,1]\times[0,1]\times[0,1]}\left(x_1,x_2,x_3\right) d\mu_L\left(x_1,x_2,x_3\right)\\
&= \int_{[0,1]\times[0,1]\times[0,1]} k\left(x_1 + x_2^2 + x_3^3\right) d\mu_L\left(x_1,x_2,x_3\right)\\
&= k\int_{[0,1]\times[0,1]\times[0,1]} \left(x_1 + x_2^2 + x_3^3\right) d\mu_L\left(x_1,x_2,x_3\right)
\end{aligned}$$

Now the real function $x_1 + x_2^2 + x_3^3$ is continuous on $[0,1] \times [0,1] \times [0,1]$. Therefore, the Lebesue integral can be computed as a Riemann integral. A as consequence, on account of the additive property of the Riemann integral and the separability of the integrand function on the pluri-interval domain, we can write

$$
\begin{aligned}
&\int_{[0,1]\times[0,1]\times[0,1]} \left(x_1 + x_2^2 + x_3^3\right) d\mu_L\left(x_1, x_2, x_3\right) \\
&= \int_{x_1=0}^{1}\int_{x_2=0}^{1}\int_{x_3=0}^{1} \left(x_1 + x_2^2 + x_3^3\right) dx_1 dx_2 dx_3 \\
&= \int_{x_1=0}^{1}\int_{x_2=0}^{1}\int_{x_3=0}^{1} x_1 dx_1 dx_2 dx_3 \\
&+ \int_{x_1=0}^{1}\int_{x_2=0}^{1}\int_{x_3=0}^{1} x_2^2 dx_1 dx_2 dx_3 \\
&+ \int_{x_1=0}^{1}\int_{x_2=0}^{1}\int_{x_3=0}^{1} x_3^3 dx_1 dx_2 dx_3 \\
&= \int_{x_1=0}^{1} x_1 dx_1 \int_{x_2=0}^{1} dx_2 \int_{x_3=0}^{1} dx_3 \\
&+ \int_{x_1=0}^{1} dx_1 \int_{x_2=0}^{1} x_2^2 dx_2 \int_{x_3=0}^{1} dx_3 \\
&+ \int_{x_1=0}^{1} dx_1 \int_{x_2=0}^{1} dx_2 \int_{x_3=0}^{1} x_3^3 dx_3 \\
&= \frac{1}{2}\, x_1^2\Big|_{x_1=0}^{1} \cdot x_2\Big|_{x_2=0}^{1} \cdot x_3\Big|_{x_3=0}^{1} \\
&+ x_1\Big|_{x_1=0}^{1} \cdot \frac{1}{3}\, x_2^3\Big|_{x_2=0}^{1} \cdot x_3\Big|_{x_3=0}^{1} \\
&+ x_1\Big|_{x_1=0}^{1} \cdot x_2\Big|_{x_2=0}^{1} \frac{1}{4} \cdot x_3^4\Big|_{x_3=0}^{1} \\
&= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \\
&= \frac{13}{12}
\end{aligned}
$$

It follows

$$k = \frac{12}{13}.$$

With similar computation, we have

$$
\begin{aligned}
\mathbf{P}\left(X_2 \leq 1/2, X_3 > 1/2\right) &= \int_{x_1=0}^{1}\int_{x_2=0}^{1/2}\int_{x_3=1/2}^{1} \frac{12}{13}\left(x_1 + x_2^2 + x_3^3\right) dx_1 dx_2 dx_3 \\
&= \frac{12}{13}\left(\frac{1}{2}\, x_1^2\Big|_{x_1=0}^{1} \cdot x_2\Big|_{x_2=0}^{1/2} \cdot x_3\Big|_{x_3=1/2}^{1}\right. \\
&+ x_1\Big|_{x_1=0}^{1} \cdot \frac{1}{3}\, x_2^3\Big|_{x_2=0}^{1/2} \cdot x_3\Big|_{x_3=1/2}^{1} \\
&+ \left. x_1\Big|_{x_1=0}^{1} \cdot x_2\Big|_{x_2=0}^{1/2} \cdot \frac{1}{4}\, x_3^4\Big|_{x_3=1/2}^{1}\right) \\
&= \frac{12}{13}\left(\frac{1}{8} + \frac{1}{48} + \frac{15}{128}\right) \\
&= \frac{101}{416}.
\end{aligned}
$$

The marginal density of the random vector $\left(X_1, X_2\right)^\intercal$ is given by

$$
\begin{aligned}
f_{X_1,X_2}\left(x_1, x_2\right) &= \int_{\mathbb{R}} f\left(x_1, x_2, x_3\right) d\mu_L\left(x_3\right) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]\times[0,1]\times[0,1]}\left(x_1, x_2, x_3\right) d\mu_L\left(x_3\right) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) 1_{[0,1]}\left(x_3\right) d\mu_L\left(x_3\right) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) d\mu_L\left(x_3\right) \\
&= k 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) \int_{x_3=0}^{1} \left(x_1 + x_2^2 + x_3^3\right) dx_3 \\
&= k 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) \left(\int_{x_3=0}^{1} x_1 d\mu_L\left(x_3\right) + \int_{x_3=0}^{1} x_2^2 d\mu_L\left(x_3\right) + \int_{x_3=0}^{1} x_3^3 d\mu_L\left(x_3\right)\right) \\
&= k 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) \left(x_1 \cdot x_3\Big|_{x_3=0}^{1} + x_2^2 \cdot x_3\Big|_{x_3=0}^{1} + \frac{1}{4}\, x_3^4\Big|_{x_3=0}^{1}\right) \\
&= k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) \\
&= k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}\left(x_1, x_2\right).
\end{aligned}
$$

We have

$$\mathbf{E}\left[\left(X_1, X_2\right)^\intercal\right] = \left(\mathbf{E}\left[X_1\right], \mathbf{E}\left[X_2\right]\right)^\intercal,$$

where

$$\mathbf{E}\left[X_k\right] = \int_{\mathbb{R}} x_k f_{X_k}\left(x_k\right) d\mu_L\left(x_k\right), \quad k = 1, 2,$$

and $f_{X_k}\left(x_k\right)$ is the marginal density of the random variable X_k, for $k = 1, 2$. Now,

$$
\begin{aligned}
f_{X_1}\left(x_1\right) &= \int_{\mathbb{R}} f_{X_1,X_2}\left(x_1, x_2\right) d\mu_L\left(x_2\right) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}\left(x_1, x_2\right) d\mu_L\left(x_2\right) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}\left(x_1\right) 1_{[0,1]}\left(x_2\right) d\mu_L\left(x_2\right) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}\left(x_1\right) d\mu_L\left(x_2\right) \\
&= k 1_{[0,1]}\left(x_1\right) \int_{x_2=0}^{1} \left(x_1 + x_2^2 + \frac{1}{4}\right) dx_2 \\
&= k 1_{[0,1]}\left(x_1\right) \left(x_1 \cdot x_2\Big|_{x_2=0}^{1} + \frac{1}{3} \cdot x_2^3\Big|_{x_2=0}^{1} + \frac{1}{4} \cdot x_2\Big|_{x_2=0}^{1}\right) \\
&= k 1_{[0,1]}\left(x_1\right) \left(x_1 + \frac{1}{3} + \frac{1}{4}\right) \\
&= k\left(x_1 + \frac{7}{12}\right) 1_{[0,1]}\left(x_1\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
f_{X_2}(x_2) &= \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\, d\mu_L(x_1) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)\, d\mu_L(x_1) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, d\mu_L(x1) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_2)\, d\mu_L(x_1) \\
&= k1_{[0,1]}(x_2) \int_{x_1=0}^{1} \left(x_1 + x_2^2 + \frac{1}{4}\right) dx_1 \\
&= k1_{[0,1]}(x_2) \left(\frac{1}{3}\cdot x_1^2\Big|_{x_1=0}^{1} + x_2^2\cdot x_1\Big|_{x_1=0}^{1} + \frac{1}{4}\cdot x_1\Big|_{x_1=0}^{1}\right) \\
&= k1_{[0,1]}(x_2) \left(\frac{1}{3} + x_2^2 + \frac{1}{4}\right) \\
&= k\left(x_2^2 + \frac{7}{12}\right) 1_{[0,1]}(x_2).
\end{aligned}
$$

It follows

$$
\begin{aligned}
\mathbf{E}[X_1] &= \int_{\mathbb{R}} k\left(x_1 + \frac{7}{12}\right) 1_{[0,1]}(x_1) = k\int_{x_1=0}^{1} \left(x_1 + \frac{7}{12}\right) dx_1 \\
&= k\left(\frac{1}{2}\cdot x_1^2\Big|_{x_1=0}^{1} + \frac{7}{12}\cdot x_1\Big|_{x_1=0}^{1}\right) = \frac{13}{12}k
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{E}[X_2] &= \int_{\mathbb{R}} k\left(x_2^2 + \frac{7}{12}\right) 1_{[0,1]}(x_2) = k\int_{x_2=0}^{1} \left(x_2^2 + \frac{7}{12}\right) dx_2 \\
&= k\left(\frac{1}{3}\cdot x_1^3\Big|_{x_2=0}^{1} + \frac{7}{12}\cdot x_2\Big|_{x_2=0}^{1}\right) = \frac{11}{12}k.
\end{aligned}
$$

The conditional density $f_{X_1,X_2|X_3=1/2}(x_1,x_2)$ is simply given by

$$
f_{X_1,X_2|X_3=1/2}(x_1,x_2) = \frac{f_{X_1,X_2,X_3}(x_1,x_2,1/2)}{\int_{\mathbb{R}^2} f_{X_1,X_2,X_3}(x_1,x_2,1/2)\, d\mu_L(x_1,x_2)} = \frac{f_{X_1,X_2,X_3}(x_1,x_2,1/2)}{f_{X_3}(1/2)},
$$

for every $(x_1,x_2)\in\mathbb{R}^2$. Now, since

$$
f_{X_1,X_2,X_3}(x_1,x_2,1/2) = k\left(x_1 + x_2^2 + \frac{1}{8}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)
$$

and

$$
\begin{aligned}
&\int_{\mathbb{R}^2} f_{X_1,X_2,X_3}(x_1,x_2,1/2)\, d\mu_L(x_1,x_2) \\
&= \int_{\mathbb{R}^2} k\left(x_1 + x_2^2 + \frac{1}{8}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)\, d\mu_L(x_1,x_2) \\
&= \int_{[0,1]\times[0,1]} k\left(x_1 + x_2^2 + \frac{1}{8}\right) d\mu_L(x_1,x_2) \\
&= \int_{x_1=0}^{1}\int_{x_2=0}^{1} k\left(x_1 + x_2^2 + \frac{1}{8}\right) dx_1 dx_2 \\
&= k\left(\int_{x_1=0}^{1}\int_{x_2=0}^{1} x_1 dx_1 dx_2 + \int_{x_1=0}^{1}\int_{x_2=0}^{1} x_2^2 dx_1 dx_2 + \int_{x_1=0}^{1}\int_{x_2=0}^{1} \frac{1}{8} dx_1 dx_2\right) \\
&= k\left(\int_{x_1=0}^{1} x_1 dx_1 \int_{x_2=0}^{1} dx_2 + \int_{x_1=0}^{1} dx_1 \int_{x_2=0}^{1} x_2^2 dx_2 + \frac{1}{8}\int_{x_1=0}^{1} dx_1 \int_{x_2=0}^{1} dx_2\right) \\
&= k\left(\frac{1}{2}\cdot x_1^2\Big|_{x_1=0}^{1}\cdot x_2\Big|_{x_2=0}^{1} + x_1\Big|_{x_1=0}^{1}\cdot\frac{1}{3}\cdot x_2^3\Big|_{x_2=0}^{1} + \frac{1}{8}\cdot x_1\Big|_{x_1=0}^{1} x_2\Big|_{x_2=0}^{1}\right) \\
&= k\left(\frac{1}{2} + \frac{1}{3} + \frac{1}{8}\right) \\
&= \frac{23}{24}k,
\end{aligned}
$$

we obtain

$$
f_{X_1,X_2|X_3=1/2}(x_1,x_2) = \frac{24}{23}\left(x_1 + x_2^2 + \frac{1}{8}\right) 1_{[0,1]\times[0,1]}(x_1,x_2).
$$

Problem 4 *Determine the value of the parameter k such that the function $f:\mathbb{R}^3\to\mathbb{R}$ given by*

$$
f(x_1,x_2,x_3) \overset{def}{=} \begin{cases} k\left(x_1 + x_2^2 + x_3^3\right), & \text{if } (x_1,x_2,x_3)\in[0,1]\times[0,1]\times[0,1], \\ 0, & \text{otherwise,} \end{cases}
$$

is a probability density. Hence, consider the random vector $X\equiv(X_1,X_2,X_3)^\intercal$ with density $f_X:\mathbb{R}^3\to\mathbb{R}$ given by

$$
f_X(x_1,x_2,x_3) \overset{def}{=} f(x_1,x_2,x_3).
$$

1. *Determine the distribution function $F_X:\mathbb{R}^3\to\mathbb{R}_+$ and check that*
$$
\frac{\partial^3 F_X}{\partial x_1\partial x_2\partial x_3}(x_1,x_2,x_3) = f_X(x_1,x_2,x_3), \quad \mu_L^3\text{-a.e. on } \mathbb{R}^3.
$$
2. *Determine the marginal distribution functions $F_{X_1}:\mathbb{R}\to\mathbb{R}$, $F_{X_2}:\mathbb{R}\to\mathbb{R}$, and $F_{X_3}:\mathbb{R}\to\mathbb{R}$ of the entries X_1, X_2, and X_3 of X.*
3. *Determine the marginal densities $f_{X_1}:\mathbb{R}\to\mathbb{R}$, $f_{X_2}:\mathbb{R}\to\mathbb{R}$, and $f_{X_3}:\mathbb{R}\to\mathbb{R}$ of the entries X_1, X_2, and X_3 of X and check that*
$$
\frac{dF_{X_n}}{dx}(x) = f_{X_n}(x), \quad \text{for } n=1,2,3, \quad \mu_L\text{-a.e. on } \mathbb{R}.
$$
4. *Determine the joint distribution function $F_{X_1,X_2}:\mathbb{R}^2\to\mathbb{R}$, $F_{X_1,X_3}:\mathbb{R}\to\mathbb{R}$, and $F_{X_2,X_3}:\mathbb{R}\to\mathbb{R}$.*
5. *Determine the joint densities $f_{X_1,X_2}:\mathbb{R}^2\to\mathbb{R}$, $f_{X_1,X_3}:\mathbb{R}\to\mathbb{R}$, and $f_{X_2,X_3}:\mathbb{R}\to\mathbb{R}$. What is the relationship between the joint distribution function $F_{X_m,X_n}:\mathbb{R}^2\to\mathbb{R}$ and the joint density $f_{X_m,X_n}:\mathbb{R}^2\to\mathbb{R}$ for $m,n=1,2,3$, $m<n$.*
6. *Determine the the expectation of X.*

7. *Determine the variance-covariance matrix of X.*

Solution. .

Problem 5 *Determine the value of the parameter k such that the function $f:\mathbb{R}^3\to\mathbb{R}$ given by*

$$f(x_1,x_2,x_3)\overset{def}{=}\begin{cases} k\left(x_1^2+x_2^2+x_3^3\right), & \text{if } (x_1,x_2,x_3)\in[0,1]\times[0,1]\times[0,1],\\ 0, & \text{otherwise,}\end{cases}$$

is a probability density. Hence, consider the random vector $(X_1,X_2,X_3)^\intercal$ with density $f_{X_1,X_2,X_3}:\mathbb{R}^3\to\mathbb{R}$ given by

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3)\overset{def}{=}f(x_1,x_2,x_3).$$

Compute:

1. *the marginal density of the random vector $(X_1,X_2)^\intercal$;*
2. *the expectation of the product $X_1\cdot X_2$;*
3. *the conditional density $f_{X_1|X_2=1/2,X_3=3/4}(x_1)$;*
4. *the probability $\mathbf{P}(X_1\le 1/2, X_2<1/2, X_3<1/2)$.*

Solution. .

Problem 6 *Let $F:\mathbb{R}^2\to\mathbb{R}_+$, briefly F, given by*

$$F(x_1,x_2)\overset{def}{=}\left(1-e^{-x_1}-e^{-x_2}+e^{-(x_1+x_2)}\right)1_{\mathbb{R}_+}(x_1)\,1_{\mathbb{R}_+}(x_2),\quad\forall(x_1,x_2)\in\mathbb{R}^2.$$

Show that F is the distribution function of a real random vector (X_1,X_2) and compute the marginal distribution functions of (X_1,X_2).

1. *Is the function F absolutely continuous?*
2. *Are the entries X_1 and X_2 of the random vector (X_1,X_2) independent random variables?*
3. *Are the entries X_1 and X_2 of the random vector (X_1,X_2) absolutely continuous random variables?*
4. *What is the distribution $F_Z:\mathbb{R}^2\to\mathbb{R}_+$, briefly F_Z, of the real random variable $Z=\max\{X_1,X_2\}$.*
5. *Is the function F_Z absolutely continuous?*

Hint: it might be useful to rewrite $F(x_1,x_2)$ in a more convenient form.

Solution. .

Problem 7 *Let $(\Omega,\mathcal{E},\mathbf{P})\equiv\Omega$ be a probability space and let $\left(\mathbb{R}^2,\mathcal{B}\left(\mathbb{R}^2\right),\mu_L^2\right)\equiv\mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}\left(\mathbb{R}^2\right)$ and the Lebesgue measure $\mu_L^2:\mathcal{B}\left(\mathbb{R}^2\right)\to\mathbb{R}_+$. Let $f:\mathbb{R}^2\to\mathbb{R}$ given by*

$$f(x,y)\overset{def}{=}ke^{-(x^2-xy+y^2/2)},\quad\forall(x,y)\in\mathbb{R}^2,$$

where $k\in\mathbb{R}$ is a parameter.

1. *Determine k such that $f:\mathbb{R}^2\to\mathbb{R}$ is a probability density. Hint: can you compute $\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(y-x)^2}dy$ with no computation?*

   Let $Z\equiv(X,Y)$ be the random vector on Ω with density $f:\mathbb{R}^2\to\mathbb{R}_+$.
2. *Determine the marginal density of the entries X and Y. Are the random variables X and Y Gaussian?*
3. *Is the random vector Z Gaussian?*
4. *Compute $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$, and $Cov(X,Y)$.*
5. *Are X and Y independent random variables?*
6. *Is the random vector Z Gaussian? Hint: consider the answer you gave to 4., what you know from the theory, and try to make a simple guess.*

Solution.

1. We can write

$$\int_{\mathbb{R}^2}f(x,y)\,d\mu_L^2(x,y)=k\int_{\mathbb{R}^2}e^{-(x^2-xy+y^2/2)}d\mu_L^2(x,y).$$

On the other hand, since $e^{-(x^2-xy+y^2/2)}$ is a continuous positive function

$$\begin{aligned}\int_{\mathbb{R}^2}e^{-(x^2-xy+y^2/2)}d\mu_L^2(x,y)&=\int_{y=-\infty}^{+\infty}\int_{x=-\infty}^{+\infty}e^{-(x^2-xy+y^2/2)}dxdy\\&=\int_{y=-\infty}^{+\infty}\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}\left(y^2-2xy+x^2\right)}e^{-\frac{1}{2}x^2}dxdy\\&=\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}x^2}\left(\int_{y=-\infty}^{+\infty}e^{-\frac{1}{2}(y-x)^2}dy\right)dx.\end{aligned}$$

Now, we have

$$\int_{y=-\infty}^{+\infty}e^{-\frac{1}{2}(y-x)^2}dy=\int_{z=-\infty}^{+\infty}e^{-\frac{1}{2}z^2}dz=\sqrt{2\pi},$$

for every $x\in\mathbb{R}$. Therefore,

$$\int_{y=-\infty}^{+\infty}\int_{x=-\infty}^{+\infty}e^{-(x^2-xy+y^2/2)}dxdy=\sqrt{2\pi}\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}x^2}dx=2\pi.$$

If follows that

$$\int_{\mathbb{R}^2}f(x,y)\,d\mu_L^2(x,y)=1\Rightarrow k=\frac{1}{2\pi}.$$

2. Considering what shown above, we have

$$f_X(x)=\int_{\mathbb{R}}\frac{1}{2\pi}f(x,y)\,d\mu_L(y)=\frac{1}{2\pi}\int_{y=-\infty}^{+\infty}e^{-\frac{1}{2}\left(y^2-2xy+x^2\right)}e^{-\frac{1}{2}x^2}dy=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2},$$

for every $x\in\mathbb{R}$. Similarly, since

$$e^{-(x^2-xy+y^2/2)}=e^{-\frac{1}{2}\left(2x^2-2xy+y^2\right)}=e^{-\frac{1}{2}\left(\left(\sqrt{2}x\right)^2-2xy+\left(\frac{y}{\sqrt{2}}\right)^2\right)}e^{-\frac{1}{2}\left(\frac{y}{\sqrt{2}}\right)^2}=e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2}e^{-\frac{y^2}{4}},$$

we have

$$f_Y(y)=\int_{\mathbb{R}}\frac{1}{2\pi}f(x,y)\,d\mu_L(x)=\frac{1}{2\pi}\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2}e^{-\frac{y^2}{4}}dx=\frac{1}{2\pi}e^{-\frac{y^2}{4}}\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2}dx,$$

for every $y\in\mathbb{R}$. Furthermore,

$$\int_{x=-\infty}^{+\infty}e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2}dx=\frac{1}{\sqrt{2}}\int_{z=-\infty}^{+\infty}e^{-\frac{1}{2}z^2}dz=\sqrt{\pi}.$$

Hence,

$$f_Y(y)=\frac{1}{2\sqrt{\pi}}e^{-\frac{y^2}{4}}=\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{1}{2}\left(\frac{y}{\sigma_Y}\right)^2},\quad\sigma_Y\equiv\sqrt{2}.$$

This shows that the random variables X and Y are Gaussian.

3. We clearly have

$$\mathbf{E}[X] = \mathbf{E}[Y] = 0.$$

Moreover,

$$\mathbf{D}^2[X] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{1}{2}x^2} dx = 1, \quad \mathbf{D}^2[Y] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} y^2 e^{-\frac{1}{2}\left(\frac{y}{\sqrt{2}}\right)^2} dy = 2.$$

In addition,

$$\begin{aligned} Cov(X,Y) &= \mathbf{E}[XY] = \int_{\mathbb{R}^2} xyf(x,y)\, d\mu_L^2(x,y) = \frac{1}{2\pi}\int_{\mathbb{R}^2} xye^{-(x^2-xy+y^2/2)} d\mu_L^2(x,y) \\ &= \frac{1}{2\pi}\int_{x=-\infty}^{+\infty} xe^{-\frac{1}{2}x^2}\left(\int_{y=-\infty}^{+\infty} ye^{-\frac{1}{2}(y-x)^2} dy\right) dx. \end{aligned}$$

On the other hand,

$$\begin{aligned} \int_{y=-\infty}^{+\infty} ye^{-\frac{1}{2}(y-x)^2} dy &= \int_{y=-\infty}^{+\infty} (y-x)\, e^{-\frac{1}{2}(y-x)^2} dy + \int_{y=-\infty}^{+\infty} xe^{-\frac{1}{2}(y-x)^2} dy \\ &= \int_{z=-\infty}^{+\infty} ze^{-\frac{1}{2}z^2} dz + x\int_{z=-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz \\ &= \sqrt{2\pi}x. \end{aligned}$$

Hence,

$$Cov(X,Y) = \frac{1}{2\pi}\int_{x=-\infty}^{+\infty} \sqrt{2\pi}x^2 e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{+\infty} x^2 e^{-\frac{1}{2}x^2} = 1.$$

4. Since

$$Cov(X,Y) \neq 0,$$

the random variables X and Y are not independent.

5. Since not independent, despite X and Y are Gaussian, we cannot state at present whether the random vector $(X,Y)^\intercal$ is Gaussian or not. To solve this doubt, we can try to write

$$\begin{pmatrix} X \\ Z \end{pmatrix} = A\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

for independent standard Gaussian random variables Z_1 and Z_1 and a suitable matrix

$$A \equiv \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.$$

If this is true, we have

$$\Sigma^2_{X,Y} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = AA^\intercal.$$

Thus, we are led to find a matrix A such that

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}\begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1}^2 + a_{1,2}^2 & a_{1,1}a_{2,1} + a_{1,2}a_{2,2} \\ a_{1,1}a_{2,1} + a_{1,2}a_{2,2} & a_{2,1}^2 + a_{2,2}^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

To this goal, observe that $\Sigma^2_{X,Y}$ has eigenvalues

$$\frac{3}{2} + \frac{1}{2}\sqrt{5} \quad \text{and} \quad \frac{3}{2} - \frac{1}{2}\sqrt{5},$$

with corresponding orthogonal eigenvectors

$$\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix}.$$

In fact, we have

$$\left(\frac{3}{2} + \frac{1}{2}\sqrt{5}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\left(\frac{3}{2} - \frac{1}{2}\sqrt{5}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = 0.$$

Therefore, normalizing the eigenvectors, we have that

$$B \equiv \left\{\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \end{pmatrix}, \begin{pmatrix} -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}\right\}$$

is a basis of orthonormal eigenvectors in $\mathbb{R}^2$. We then have

$$M_E^B(id)\,\Lambda M_B^E(id) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

where

$$E \equiv \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$$

is the standard orthonormal basia in $\mathbb{R}^2$,

$$M_E^B(id) = \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}, \quad \Lambda \equiv \begin{pmatrix} \frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix},$$

and

$$M_B^E(id) = M_E^B(id)^{-1} = M_E^B(id)^\intercal = \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}\begin{pmatrix} \frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix}\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In addition, we can write

$$\begin{aligned} &\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}\begin{pmatrix} \frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix}\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}\begin{pmatrix} \sqrt{\frac{3}{2}+\frac{1}{2}\sqrt{5}} & 0 \\ 0 & \sqrt{\frac{3}{2}-\frac{1}{2}\sqrt{5}} \end{pmatrix} \\ &\cdot\begin{pmatrix} \sqrt{\frac{3}{2}+\frac{1}{2}\sqrt{5}} & 0 \\ 0 & \sqrt{\frac{3}{2}-\frac{1}{2}\sqrt{5}} \end{pmatrix}\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}\begin{pmatrix} \frac{1}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{1}{2}\sqrt{5}-\frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{1}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{1}{2}\sqrt{5}-\frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix}. \end{aligned}$$

Therefore, we obtain

$$\begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Setting

$$A = \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix},$$

it then follows

$$a_{1,1}^2 + a_{1,2}^2 = \left(\frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\right)^2 + \left(-\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\right)^2 = 1,$$

$$a_{2,1}^2 + a_{2,2}^2 = \left(\frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\right)^2 + \left(\frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\right)^2 = 2,$$

$$a_{1,1}a_{2,1} + a_{1,2}a_{2,2} = \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} - \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} = 1.$$

This proves that $(X,Y)^\intercal$ is Gaussian. Note that, from

$$\begin{pmatrix} X \\ Y \end{pmatrix} = A \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

it follows

$$\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X & Y \end{pmatrix} = A \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} A^\intercal,$$

that is to say

$$\begin{pmatrix} X^2 & XY \\ XY & Y^2 \end{pmatrix} = A \begin{pmatrix} Z_1^2 & Z_1Z_2 \\ Z_1Z_2 & Z_2^2 \end{pmatrix} A^\intercal.$$

It follows,

$$\begin{aligned} \Sigma_{X,Y}^2 &= \begin{pmatrix} \mathbf{D}^2[X] & Cov(X,Y) \\ Cov(X,Y) & \mathbf{D}^2[X] \end{pmatrix} = \begin{pmatrix} \mathbf{E}[X^2] & \mathbf{E}[XY] \\ \mathbf{E}[XY] & \mathbf{E}[Y^2] \end{pmatrix} \\ &= A \begin{pmatrix} \mathbf{E}[Z_1^2] & \mathbf{E}[Z_1Z_2] \\ \mathbf{E}[Z_1Z_2] & \mathbf{E}[Z_2^2] \end{pmatrix} A^\intercal = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^\intercal = AA^\intercal. \end{aligned}$$

Problem 8 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_L^2) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^2)$ and the Lebesgue measure $\mu_L^2 : \mathcal{B}(\mathbb{R}^2) \to \mathbb{R}_+$. Let $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x,y) \overset{def}{=} ke^{-\frac{x^2-xy+y^2}{2}}, \quad \forall (x,y) \in \mathbb{R}^2,$$

where $k \in \mathbb{R}$ is a parameter.

1. *Determine k such that $f : \mathbb{R}^2 \to \mathbb{R}$ is a probability density. Hint: It may be useful to recall that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = 1$.*
2. *Determine the marginal density functions of the entries X and Y. Are X and Y independent?*
3. *Compute $\mathbf{P}(X = Y)$ and $\mathbf{P}(X \geq Y)$.*

Solution. .

Exercise 9 (Sheldon M. Ross - 4.11) *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be real random variables on Ω such that the random vector (X,Y) is absolutely continuous with a density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f_{X,Y}(x,y) \overset{def}{=} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \cdot 1_{(0,1)\times(0,2)}(x,y), \quad \forall (x,y) \in \mathbb{R}^2.$$

1. *Check that $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ is a density function.*
2. *Are the random variables X and Y absolutely continuous? In case of affirmative answer determine the marginal densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ of X and Y, respectively.*
3. *Check whether the random variables X and Y are independent.*
4. *Compute $\mathbf{P}(X > Y)$.*

Solution.

1. We will have proven that $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ is a density function if we can show that

$$\int_{\mathbb{R}^2} f_{X,Y}(x,y)\, d\mu_L^2(x,y) = 1.$$

On the other hand, considering the properties of the Lebesgue integral, we have

$$\begin{aligned} \int_{\mathbb{R}^2} f_{X,Y}(x,y)\, d\mu_L(x,y) &= \int_{\mathbb{R}^2} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \cdot 1_{(0,1)\times(0,2)}(x,y)\ d\mu_L^2(x,y) \\ &= \int_{(0,1)\times(0,2)} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) d\mu_L^2(x,y) \\ &= \int_{(0,1)\times(0,2)} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) dxdy \\ &= \int_{y=0}^{2}\int_{x=0}^{1} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) dxdy \\ &= \frac{6}{7}\int_{y=0}^{2}\left(\int_{x=0}^{1}\left(x^2 + \frac{xy}{2}\right) dx\right) dy \\ &= \frac{6}{7}\int_{y=0}^{2}\left(\frac{x^3}{3} + \frac{x^2y}{4}\bigg|_0^1\right) dy \\ &= \frac{6}{7}\int_{y=0}^{2}\left(\frac{1}{3} + \frac{y}{4}\right) dy \\ &= \frac{6}{7}\left(\frac{y}{3} + \frac{y^2}{8}\bigg|_0^2\right) \\ &= \frac{6}{7}\left(\frac{2}{3} + \frac{1}{2}\right) \\ &= 1. \end{aligned}$$

2. Since the random vector is absolutely continuous the entries X and Y are absolutely continuous random variables with densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ given by

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\, d\mu_L(y) \qquad \text{and} \qquad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\, d\mu_L(x),$$

μ_L-a.e. on $\mathbb{R}$, respectively. Now, we have

$$\begin{aligned}
\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(y) &= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L(y) \\
&= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y)\,d\mu_L(y) \\
&= \int_{(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,d\mu_L(y) \\
&= \frac{6}{7}\left(\int_0^2\left(x^2+\frac{xy}{2}\right)dy\right)\cdot 1_{(0,1)}(x) \\
&= \frac{6}{7}\left(x^2y+\left.\frac{xy^2}{4}\right|_{y=0}^{2}\right)\cdot 1_{(0,1)}(x) \\
&= \frac{6}{7}\left(2x^2+x\right)\cdot 1_{(0,1)}(x).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(x) &= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L(x) \\
&= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y)\,d\mu_L(x) \\
&= \int_{(0,1)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,2)}(y)\,d\mu_L(y) \\
&= \frac{6}{7}\left(\int_0^1\left(x^2+\frac{xy}{2}\right)dx\right)\cdot 1_{(0,2)}(y) \\
&= \frac{6}{7}\left(\frac{x^3}{3}+\left.\frac{x^2y}{4}\right|_{x=0}^{1}\right)\cdot 1_{(0,2)}(y) \\
&= \frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y).
\end{aligned}$$

Therefore, we can write

$$f_X(x)=\frac{6}{7}\left(x+2x^2\right)\cdot 1_{(0,1)}(x) \qquad \text{and} \qquad f_Y(y)=\frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y),$$

μ_L-a.e. on $\mathbb{R}$, respectively.

3. The random variables X and Y are independent if and only if

$$f_X(x)\,f_Y(y)=f_{X,Y}(x,y),$$

μ_L^2-a.e. on $\mathbb{R}^2$. On the other hand,

$$\begin{aligned}
f_X(x)\,f_Y(x) &= \left(\frac{6}{7}\left(x+2x^2\right)\cdot 1_{(0,1)}(x)\right)\left(\frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y)\right) \\
&= \frac{36}{49}\left(\frac{x}{3}+\frac{xy}{4}+\frac{2x^2}{3}+\frac{x^2y}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y) \\
&= \frac{36}{49}\left(\frac{x}{3}+\frac{xy}{4}+\frac{2x^2}{3}+\frac{x^2y}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y) \\
&\neq \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)
\end{aligned}$$

for almost all points $(x,y)\in(0,1)\times(0,2)$. Therefore, X and Y are not independent.

4. To compute $\mathbf{P}(X>Y)$ we apply the formula

$$\mathbf{P}((X,Y)\in B)=\int_B f_{X,Y}(x,y)\;d\mu_L^2(x,y),$$

which holds true for every $B\in\mathcal{B}(\mathbb{R}^2)$, by suitably choosing B to represent the event $\{X>Y\}$ in terms of the event $\{(X,Y)\in B\}$. Eventually, setting

$$B\equiv\left\{(x,y)\in\mathbb{R}^2 : x>y\right\},$$

it turns out that we can write

$$\{X>Y\}=\{(X,Y)\in B\}.$$

In fact, assume that $\omega\in\{X>Y\}\equiv\{\omega\in\Omega : X(\omega)>Y(\omega)\}$, then we have $X(\omega)>Y(\omega)$ so that $(X(\omega),Y(\omega))\in B$ and $\omega\in\{(X,Y)\in B\}\equiv\{\omega\in\Omega : (X(\omega),Y(\omega))\in B\}$. Conversely, assume that $\omega\in\{(X,Y)\in B\}$, then $(X(\omega),Y(\omega))\in B$, which implies $X(\omega)>Y(\omega)$ and consequently $\omega\in\{X>Y\}$. As a consequence, we have

$$\begin{aligned}
\mathbf{P}(X>Y) &= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}} f_{X,Y}(x,y)\,d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}\cap(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}\cap(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)dxdy \\
&= \frac{6}{7}\int_{x=0}^{1}\left(\int_{y=0}^{x}\left(x^2+\frac{xy}{2}\right)dy\right)dx \\
&= \frac{6}{7}\int_{x=0}^{1}\left(x^2y+\left.\frac{xy^2}{4}\right|_0^x\right)dx \\
&= \frac{6}{7}\int_{x=0}^{1}\frac{5x^3}{4}dx \\
&= \frac{6}{7}\left.\frac{5x^4}{16}\right|_0^1 \\
&= \frac{6}{7}\frac{5}{16} \\
&= \frac{15}{56}\approx 0.26786
\end{aligned}$$

.

Problem 10 *Let $f:\mathbb{R}^2\to\mathbb{R}_+$ given by*

$$f(x,y)\overset{def}{=}\frac{4x+2y}{3}1_{[0,1]}(x)\,1_{[0,1]}(y),\quad \forall (x,y)\in\mathbb{R}^2.$$

1. *Show that $f:\mathbb{R}^2\to\mathbb{R}_+$ is the density function of a real random vector (X,Y).*
2. *Compute the marginal densities of (X,Y) and check that the computed marginal densities are actually probability densities.*
3. *May we say that the entries X and Y of the random vector (X,Y) are independent random variables?*
4. *Compute the conditional density function $f_{X|Y}(x,y)$ of X given that $Y=y$ and check the computed density is actually a probability density.*
5. *Compute the function $\mathbf{E}[X\mid Y=y]$ and the conditional expectation $\mathbf{E}[X\mid Y]$.*

Solution. .

Problem 11 (Sheldon M. Ross - 4.17) *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be absolutely continuous real random variables on Ω with densities $f_X : \mathbb{R} \to \mathbb{R}$ and $f_Y : \mathbb{R} \to \mathbb{R}$, respectively. Assume that the densities of X and Y have at most a finite number of discontinuity points and that X and Y are independent.*

1. *Prove that we have*
$$\mathbf{P}(X + Y < a) = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\ dy,$$
for every $a \in \mathbb{R}$, and
$$\mathbf{P}(X \leq Y) = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\ dy,$$
where $F_X : \mathbb{R} \to \mathbb{R}_+$ is the distribution function of X.

2. *Show that the latter equation does not hold true if we drop the assumption of independence.*

Solution.

1. Since the random variables X and Y are absolutely continuous and independent, the random vector (X, Y) is absolutely continuous with a density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ given by
$$f_{X,Y}(x, y) = f_X(x) f_Y(y),$$
for every $(x, y) \in \mathbb{R}^2$. To compute $\mathbf{P}(X + Y < a)$ we apply the formula
$$\mathbf{P}((X, Y) \in B) = \int_B f_{X,Y}(x, y)\ d\mu_L^2(x, y),$$
which holds true for every $B \in \mathcal{B}(\mathbb{R}^2)$, by suitably choosing B to represent the event $\{X + Y < a\}$ in terms of the event $\{(X, Y) \in B\}$. Eventually, setting
$$B \equiv \left\{(x, y) \in \mathbb{R}^2 : x + y < a\right\},$$
it turns out that we can write
$$\{X + Y < a\} = \{(X, Y) \in B\}.$$
Hence, on account of the continuity property of the densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$, we have
$$\begin{aligned}
\mathbf{P}(X + Y < a) &= \int_{\{(x,y)\in\mathbb{R}^2 : x+y<a\}} f_{X,Y}(x, y)\ d\mu_L^2(x, y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2 : x+y<a\}} f_X(x) f_Y(y)\ d\mu_L^2(x, y) \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{a-y} f_X(x) f_Y(y)\ dxdy \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{a-y} f_X(x) f_Y(y)\ dxdy \\
&= \int_{y=-\infty}^{\infty} f_Y(y) \left(\int_{x=-\infty}^{a-y} f_X(x)\ dx\right) dy \\
&= \int_{-\infty}^{\infty} f_Y(y) F_X(a - y)\ dy.
\end{aligned}$$

Similarly,
$$\begin{aligned}
\mathbf{P}(X \leq Y) &= \int_{\{(x,y)\in\mathbb{R}^2 : x\leq y\}} f_{X,Y}(x, y)\ d\mu_L^2(x, y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2 : x\leq y\}} f_X(x) f_Y(y)\ d\mu_L^2(x, y) \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{y} f_X(x) f_Y(y)\ dxdy \\
&= \int_{y=-\infty}^{\infty} f_Y(y) \left(\int_{x=-\infty}^{y} f_X(x)\ dx\right) dy \\
&= \int_{-\infty}^{\infty} f_Y(y) F_X(y)\ dy.
\end{aligned}$$

2. To show that
$$\mathbf{P}(X \leq Y) = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\ dy,$$
does not hold true if we drop the assumption of independence, consider the random variables X and Y with densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ given by
$$f_X(x) \stackrel{\text{def}}{=} \frac{6}{7}\left(x + 2x^2\right) \cdot 1_{(0,1)}(x), \quad \forall x \in \mathbb{R} \qquad \text{and} \qquad f_Y(y) \stackrel{\text{def}}{=} \frac{6}{7}\left(\frac{1}{3} + \frac{y}{4}\right) \cdot 1_{(0,2)}(y), \quad \forall y \in \mathbb{R}$$
and a joint density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ given by
$$f_{X,Y}(x, y) \stackrel{\text{def}}{=} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \cdot 1_{(0,1)\times(0,2)}(x, y), \quad \forall (x, y) \in \mathbb{R}^2.$$
We know that
$$\mathbf{P}(X > Y) = \frac{15}{56}.$$
Therefore,
$$\mathbf{P}(X \leq Y) = 1 - \mathbf{P}(X > Y) = 1 - \frac{15}{56} = \frac{41}{56} \approx 0.73214$$
On the other hand, the distribution function $F_X : \mathbb{R} \to \mathbb{R}_+$ is given by
$$F_X(x) = \frac{6}{7}\left(\frac{1}{2}x^2 + \frac{2}{3}x^3\right) \cdot 1_{(0,1]}(x) + 1_{(1,+\infty)}(x)$$
for every $x \in \mathbb{R}$. In fact, we have
$$\begin{aligned}
F_X(x) &= \int_{(-\infty,x)} f_X(u)\ d\mu_L(u) \\
&= \int_{(-\infty,x)} \frac{6}{7}\left(u + 2u^2\right) \cdot 1_{(0,1)}(u)\ d\mu_L(u) \\
&= \int_{(-\infty,x)\cap(0,1)} \frac{6}{7}\left(u + 2u^2\right)\ d\mu_L(u) \\
&= \begin{cases} 0 & \text{if } x \leq 0 \\ \int_{(0,x)} \frac{6}{7}\left(u + 2u^2\right)\ d\mu_L(u) & \text{if } 0 < x < 1 \\ \int_{(0,1)} \frac{6}{7}\left(u + 2u^2\right)\ d\mu_L(u) & \text{if } 1 \leq x \end{cases},
\end{aligned}$$
where
$$\int_{(0,x)} \frac{6}{7}\left(u + 2u^2\right)\ d\mu_L(u) = \frac{6}{7}\int_0^x \left(u + 2u^2\right)\ du = \frac{6}{7}\left.\frac{1}{2}u^2 + \frac{2}{3}u^3\right|_0^x = \frac{6}{7}\left(\frac{1}{2}x^2 + \frac{2}{3}x^3\right),$$

for every $x \in (0,1]$. As a consequence,

$$
\begin{aligned}
&\int_{-\infty}^{\infty} f_Y(y) F_X(y)\ dy \\
&\int_{-\infty}^{\infty} \left(\frac{6}{7}\left(\frac{1}{2}y^2+\frac{2}{3}y^3\right)\cdot 1_{(0,1]}(y)+1_{(1,+\infty)}(y)\right)\left(\frac{6}{7}\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(0,2)}(y)\right)\ dy \\
&=\int_{-\infty}^{\infty} \left(\frac{36}{49}\left(\frac{1}{2}y^2+\frac{2}{3}y^3\right)\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(0,1]}(y)+\frac{6}{7}\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(1,2)}(y)\right)\ dy \\
&=\frac{36}{49}\int_0^1\left(\frac{1}{6}y^2+\frac{25}{72}y^3+\frac{1}{6}y^4\right)\ dy+\frac{6}{7}\int_1^2\left(\frac{1}{3}+\frac{1}{4}y\right)\ dy \\
&=\frac{36}{49}\left(\frac{1}{18}y^3+\frac{25}{288}y^4+\frac{1}{30}y^5\Big|_0^1\right)+\frac{6}{7}\left(\frac{1}{3}y+\frac{1}{8}y^2\Big|_1^2\right) \\
&=\frac{36}{49}\left(\frac{1}{18}+\frac{25}{288}+\frac{1}{30}\right)+\frac{6}{7}\left(\left(\frac{2}{3}+\frac{4}{8}\right)-\left(\frac{1}{3}+\frac{1}{8}\right)\right) \\
&=\frac{1443}{1960}\approx 0.73622.
\end{aligned}
$$

It the follows

$$\mathbf{P}(X\leq Y)\neq\int_{-\infty}^{\infty} f_Y(y) F_X(y)\ dy.$$

It may be interesting to observe that if we assume that X and Y are independent, then a joint density $f_{X,Y}:\mathbb{R}^2\to\mathbb{R}_+$ is given by

$$
\begin{aligned}
f_{X,Y}(x,y)&=f_X(x)\,f_Y(y) \\
&=\left(\frac{6}{7}\left(x+2x^2\right)\cdot 1_{(0,1)}(x)\right)\left(\frac{6}{7}\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(0,2)}(y)\right) \\
&=\frac{36}{49}\left(x+2x^2\right)\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(0,1)\times(0,2)}(x,y),
\end{aligned}
$$

for every $(x,y)\in\mathbb{R}^2$. In this case,

$$
\begin{aligned}
\mathbf{P}(X\leq Y)&=\int_{\{(x,y)\in\mathbb{R}^2:x\leq y\}}\frac{36}{49}\left(x+2x^2\right)\left(\frac{1}{3}+\frac{1}{4}y\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\ d\mu_L^2(x,y) \\
&=\int_{\{(x,y)\in\mathbb{R}^2:x\leq y\}\cap(0,1)\times(0,2)}\frac{36}{49}\left(x+2x^2\right)\left(\frac{1}{3}+\frac{1}{4}y\right)\ d\mu_L^2(x,y) \\
&=\frac{36}{49}\int_{x=0}^1\left(x+2x^2\right)\left(\int_{y=x}^2\left(\frac{1}{3}+\frac{1}{4}y\right)\ dy\right)\ dx \\
&=\frac{36}{49}\int_{x=0}^1\left(x+2x^2\right)\left(\frac{1}{3}y+\frac{1}{8}y^2\Big|_x^2\right)\ dx \\
&=\frac{36}{49}\int_{x=0}^1\left(x+2x^2\right)\left(\frac{7}{6}-\left(\frac{1}{3}x+\frac{1}{8}x^2\right)\right)\ dx \\
&=\frac{36}{49}\int_{x=0}^1\left(\frac{7}{6}x+2x^2-\frac{19}{24}x^3-\frac{1}{4}x^4\right)\ dx \\
&=\frac{36}{49}\ \frac{7}{12}x^2+\frac{2}{3}x^3-\frac{19}{96}x^4-\frac{1}{20}x^5\Big|_0^1 \\
&=\frac{36}{49}\left(\frac{7}{12}+\frac{2}{3}-\frac{19}{96}-\frac{1}{20}\right) \\
&=\frac{1443}{1960} \\
&=\int_{-\infty}^{\infty} f_Y(y)\,F_X(y)\ dy
\end{aligned}
$$

Exercise 12 (Sheldon M. Ross - 4.18) *Let $(\Omega,\mathcal{E},\mathbf{P})\equiv\Omega$ be a probability space and let X and Z be absolutely continuous real random variables on Ω with densities $f_X:\mathbb{R}\to\mathbb{R}_+$ and $f_Z:\mathbb{R}\to\mathbb{R}_+$ given by*

$$f_X(x)\overset{def}{=}6x(1-x)\cdot 1_{(0,1)}(x),\quad\forall x\in\mathbb{R}\qquad\textit{and}\qquad f_Z(y)\overset{def}{=}2z\cdot 1_{(0,1)}(z),\quad\forall z\in\mathbb{R}.$$

Assume that X and Z are independent and show that the random variable $W=X^2Z$ is absolutely continuous.

Solution. .

Problem 13 *Let U,V real random variables on a probability space Ω such that such that $U\sim V\sim N(0,1)$, the vector (U,V) is Gaussian, and $Corr(U,V)\equiv\rho<1$. Consider the real random variables*

$$X\overset{def}{=}U-\rho V\quad\textit{and}\quad Y\overset{def}{=}\sqrt{1-\rho}V.$$

1. *Can you prove that the vector (X,Y) Gaussian?*
2. *Are the random variables X and Y independent?*
3. *Compute the distributions of X and Y;*
4. *Compute $\mathbf{E}\left[X^2Y^2\right]$, $\mathbf{E}\left[XY^3\right]$, $\mathbf{E}\left[Y^4\right]$.*
5. *Compute $\mathbf{E}\left[U^2V^2\right]$.*

Solution. .

Problem 14 *Show that the function $f:\mathbb{R}^2\to\mathbb{R}$ given by*

$$f(x,y)\overset{def}{=}\begin{cases}\frac{y-x}{2}, & \textit{if } (x,y)\in[-1,0]\times[0,1],\\ \frac{x-y}{2}, & \textit{if } (x,y)\in[0,1]\times[-1,0],\\ 0, & \textit{otherwise,}\end{cases}$$

is a probability density. Hence, consider the random vector $(X,Y)^{\intercal}$ with density $f_{X,Y}:\mathbb{R}^2\to\mathbb{R}$ given by

$$f_{X,Y}(x,y)\overset{def}{=}f(x,y).$$

Determine the marginal densities of entries X and Y of $(X,Y)^{\intercal}$. Are X and Y correlated? Are X and Y independent? Compute

$$\mathbf{P}(X+Y\geq 0).$$

Solution. .

Problem 15 *Determine the value of the parameter k such that the function $f:\mathbb{R}^2\to\mathbb{R}$ given by*

$$f(x,y)\overset{def}{=}\begin{cases}ke^{-(x+y)} & \textit{if } 0<x<y\\ 0 & \textit{otherwise}\end{cases}$$

is a probability density. Hence, consider the random vector $(X,Y)^{\intercal}$ with density $f_{X,Y}:\mathbb{R}^2\to\mathbb{R}$ given by

$$f_{X,Y}(x,y)\overset{def}{=}f(x,y).$$

Determine the marginal densities of the entries of $(X,Y)^{\intercal}$. Are X and Y correlated? Are X and Y independent? What is the distribution of Y?

Solution. .

Problem 16 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2)) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra. Let $f : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f(x,y) \overset{def}{=} kxe^{-(x+y)} 1_{\mathbb{R}^2_+}(x,y), \quad \forall (x,y) \in \mathbb{R}^2$$

where $\mathbb{R}^2_+ \equiv \{(x,y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$. Determine $k \in \mathbb{R}$ such that $f : \mathbb{R}^2 \to \mathbb{R}_+$ is a probability density. Let $Z \equiv (X, Y)$ be the random vector of density $f : \mathbb{R}^2 \to \mathbb{R}_+$.

1. *Determine the distribution function $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ of the vector Z and check that*

$$\frac{\partial F^2}{\partial x \partial y}(x,y) = f(x,y), \quad \mu_L^2\text{-a.e. on } \mathbb{R}^2.$$

2. *Determine the marginal distribution function $F_X : \mathbb{R} \to \mathbb{R}_+$ and $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z.*

3. *Determine the densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z and check that*

$$\frac{dF_X}{dx}(x) = f_X(x) \quad \text{and} \quad \frac{dF_Y}{dy}(y) = f_Y(y), \quad \mu_L\text{-a.e. on } \mathbb{R}.$$

4. *Are X and Y independent random varables?*

5. *Compute $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$ and $Cov(X,Y)$.*

6. *Compute $\mathbf{E}[(X,Y)]$ and the covariance matrix of the vector (X,Y).*

Solution. . □

Exercise 17 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X, Z be real random variables on Ω such that $X \sim N(0,1)$, $Z \sim Rad\left(\frac{1}{2}\right)$ and X and Z are independent. Set*

$$Y \overset{def}{=} XZ.$$

Prove that $Y \sim N(0,1)$, but the random vector (X,Y) is not Gaussian.

Solution. Since X and Z are independent and X is symmetric about zero, we have

$$\begin{aligned}
\mathbf{P}(Y \leq y) &= \mathbf{P}(XZ \leq y) = \mathbf{P}(XZ \leq y, Z = -1) + \mathbf{P}(XZ \leq y, Z = 1) \\
&= \mathbf{P}(XZ \leq y \mid Z = -1)\,\mathbf{P}(Z = -1) + \mathbf{P}(XZ \leq y \mid Z = 1)\,\mathbf{P}(Z = 1) \\
&= \frac{1}{2}\left(\mathbf{P}(-X \leq y \mid Z = -1) + \mathbf{P}(X \leq y \mid Z = 1)\right) \\
&= \frac{1}{2}\left(\mathbf{P}(-X \leq y) + \mathbf{P}(X \leq y)\right) \\
&= \frac{1}{2}\left(\mathbf{P}(X \geq -y) + \mathbf{P}(X \leq y)\right) \\
&= \mathbf{P}(X \leq y),
\end{aligned}$$

for every $y \in \mathbb{R}$. This proves that $Y \sim N(0,1)$. Now, consider $Cov(X,Y)$. We have $\mathbf{E}[X] = \mathbf{E}[Y] = 0$ and, thanks again to the independence of X and Z,

$$Cov(X,Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] = \mathbf{E}[XY] = \mathbf{E}\left[X^2 Z\right] = \mathbf{E}\left[X^2\right]\mathbf{E}[Z] = 0.$$

That is to say that the random variables X and Y are uncorrelated. Therefore, if random vector (X,Y) were Gaussian, the random variables X and Y would be independent. In particular, we would have

$$\mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(X \leq x)\,\mathbf{P}(Y \leq y),$$

for all $x, y \in \mathbb{R}$. On the other hand, still on account of the independence of X and Z, we have

$$\begin{aligned}
\mathbf{P}(X \leq x, Y \leq y) &= \mathbf{P}(X \leq x, XZ \leq y) = \mathbf{P}(X \leq x, XZ \leq y, Z = -1) + \mathbf{P}(X \leq x, XZ \leq y, Z = 1) \\
&= \mathbf{P}(X \leq x, XZ \leq y \mid Z = -1)\,\mathbf{P}(Z = -1) + \mathbf{P}(X \leq x, XZ \leq y \mid Z = 1)\,\mathbf{P}(Z = 1) \\
&= \mathbf{P}(X \leq x, -X \leq y \mid Z = -1)\,\mathbf{P}(Z = -1) + \mathbf{P}(X \leq x, X \leq y \mid Z = 1)\,\mathbf{P}(Z = 1) \\
&= \mathbf{P}(X \leq x, -X \leq y)\,\mathbf{P}(Z = -1) + \mathbf{P}(X \leq x, X \leq y)\,\mathbf{P}(Z = 1) \\
&= \frac{1}{2}\left(\mathbf{P}(X \leq x, X \geq -y) + \mathbf{P}(X \leq x, X \leq y)\right).
\end{aligned}$$

for all $x, y \in \mathbb{R}$. As a consequence, we would obtain

$$\mathbf{P}(X \leq x)\,\mathbf{P}(Y \leq y) = \frac{1}{2}\left(\mathbf{P}(X \leq x, X \geq -y) + \mathbf{P}(X \leq x, X \leq y)\right)$$

for all $x, y \in \mathbb{R}$, which is clearly false if we consider, for instance, $x = y = 0$. □

Problem 18 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_L^2) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^2)$ and the Lebesgue measure $\mu_L^2 : \mathcal{B}(\mathbb{R}^2) \to \mathbb{R}_+$. Let*

$$\mathbb{R}^2_+(x > y) \equiv \{(x,y) \in \mathbb{R}^2_+ : x > y\}, \qquad \mathbb{R}^2_+(x \leq y) \equiv \{(x,y) \in \mathbb{R}^2_+ : x \leq y\},$$

and let $F : \mathbb{R}^2 \to \mathbb{R}_+$ given by

$$F(x,y) \overset{def}{=} \left(1 - e^{-y} - \frac{1}{2}ye^{-x}\right) 1_{\mathbb{R}^2_+(x>y)}(x,y) + \left(1 - e^{-x} - \frac{1}{2}xe^{-y}\right) 1_{\mathbb{R}^2_+(x \leq y)}(x,y), \quad \forall (x,y) \in \mathbb{R}^2.$$

1. *Can you show that the function $F : \mathbb{R}^2 \to \mathbb{R}_+$ is a distribution function? Hint: consider carefully the sets $\mathbb{R}^2_+(x > y)$ and $\mathbb{R}^2_+(x \leq y)$ (draw a graph).*
   Let $Z \equiv (X,Y)^\intercal$ be the random vector on Ω with distribution function $F : \mathbb{R}^2 \to \mathbb{R}$.

2. *Can you determine the marginal distribution of the entries X and Y?*

3. *Is the random vector Z absolutely continuous? Can you determine a density $f_Z : \mathbb{R}^2 \to \mathbb{R}_+$ for Z? Hint: it may be useful to rewrite the indicator fnctions $1_{\mathbb{R}^2_+(x>y)}(x,y)$ and $1_{\mathbb{R}^2_+(x \leq y)}(x,y)$ in terms of product of other indicator functions.*

4. *If Z is not absolutely continuous, can you determine the marginal densities of the entries X and Y?*

Solution.

1. We clearly have

$$\lim_{x \to -\infty} F(x,y) = \lim_{y \to -\infty} F(x,y) = 0$$

and

$$\lim_{y \to +\infty} \lim_{x \to +\infty} F(x,y) = \lim_{x \to +\infty} \lim_{y \to +\infty} F(x,y) = 1.$$

We also have

$$\frac{\partial F}{\partial x}(x,y) = \frac{1}{2}ye^{-x} 1_{\mathbb{R}^2_+(x>y)}(x,y) + \left(e^{-x} - \frac{1}{2}e^{-y}\right) 1_{\mathbb{R}^2_+(x<y)}(x,y)$$

and

$$\frac{\partial F}{\partial y}(x,y) = \frac{1}{2}\left(e^{-y} - \frac{1}{2}e^{-x}\right) 1_{\mathbb{R}^2_+(x>y)}(x,y) + \frac{1}{2}xe^{-y} 1_{\mathbb{R}^2_+(x<y)}(x,y).$$

Therefore,

$$\frac{\partial F}{\partial x}(x,y) \geq 0 \quad \text{and} \quad \frac{\partial F}{\partial y}(x,y) \geq 0,$$

for every $(x,y) \in \mathbb{R}^2_+ - \{(x,y) \in \mathbb{R}^2_+ : x = y\}$. Note also that

$$F(0,y) = 0,$$

for every $y \in \mathbb{R}^2_+$. In fact, this is clearly true if $y < 0$ and it is also true for $y \geq 0$ since

$$1_{\mathbb{R}^2_+(x>y)}(0,y) = 0 \quad \text{and} \quad \left(1 - e^{-x} - \frac{1}{2}xe^{-y}\right)_{x=0} = 0.$$

Similarly

$$F(x,0) = 0,$$

for every $x \in \mathbb{R}^2_+$. In the end, for every $(x,y) \in \mathbb{R}^2_+ (x \leq y)$ such that $x = y$ we have

$$F(x,x) = 0,$$

if $x \leq 0$ and

$$F(x,x) = 1 - e^{-x} - \frac{1}{2}xe^{-x}$$

Hence, for every $(x,y) \in \mathbb{R}^2_+ (x \leq y)$ such that $x = y$ we have

$$\lim_{u\to x^\cdot} F(u,y) = \lim_{u\to x^\cdot} F(u,x) = \lim_{u\to x^\cdot}\left(1 - e^{-u} - \frac{1}{2}ue^{-x}\right) = 1 - e^{-x} - \frac{1}{2}xe^{-x} = F(x,x)$$

and

$$\lim_{v\to y^\cdot} F(x,v) = \lim_{v\to x^\cdot}\left(1 - e^{-v} - \frac{1}{2}ve^{-x}\right) = 1 - e^{-x} - \frac{1}{2}xe^{-x} = F(x,x).$$

This is enough to show that F is a distribution function.

2. With regard to the marginal distributions, we have

$$F_X(x) = \lim_{y\to+\infty} F(x,y) = \begin{cases} 0, & \text{if } x \leq 0, \\ \lim_{y\to+\infty}\left(1 - e^{-x} - \frac{1}{2}xe^{-y}\right) = 1 - e^{-x}, & \text{if } x > 0. \end{cases}$$

That is,

$$F_X(x) = \left(1 - e^{-x}\right)1_{\mathbb{R}_+}(x).$$

Similarly,

$$F_Y(y) = \lim_{x\to+\infty} F(x,y) = \begin{cases} 0, & \text{if } y \leq 0, \\ \lim_{x\to+\infty}\left(1 - e^{-y} - \frac{1}{2}ye^{-x}\right) = 1 - e^{-y}, & \text{if } y > 0. \end{cases}$$

That is

$$F_Y(y) = \left(1 - e^{-y}\right)1_{\mathbb{R}_+}(y).$$

Note that X and Y are exponential random variables with rate parameter $\lambda = 1$.

3. We check whether Z is absolutely continuous. To this goal, observe that we have

$$\frac{\partial^2 F_Z}{\partial y \partial x}(x,y) = \frac{1}{2}\left(e^{-x}1_{\mathbb{R}^2_+(x>y)}(x,y) + e^{-y}1_{\mathbb{R}^2_+(x<y)}(x,y)\right) = \frac{\partial^2 F_Z}{\partial x \partial y}(x,y)$$

for every $(x,y) \in \mathbb{R}^2 - \left\{(x,y) \in \mathbb{R}^2_+ : x = y\right\}$. However,

$$\mu^2_L\left(\left\{(x,y) \in \mathbb{R}^2_+ : x = y\right\}\right) = 0.$$

Hence, we check whether

$$\begin{aligned}
F_Z(x,y) &= \int_{(-\infty,x]\times(-\infty,y]} \frac{\partial^2 F_Z}{\partial y \partial x}(u,v)\, d\mu^2_L(u,v) \\
&= \frac{1}{2}\int_{(-\infty,x]\times(-\infty,y]} \left(e^{-u}1_{\mathbb{R}^2_+(x>y)}(u,v) + e^{-v}1_{\mathbb{R}^2_+(x<y)}(u,v)\right) d\mu^2_L(u,v) \\
&= \frac{1}{2}\left(\int_{(-\infty,x]\times(-\infty,y]} e^{-u}1_{\mathbb{R}^2_+(x>y)}(u,v)\, d\mu^2_L(u,v) + \int_{(-\infty,x]\times(-\infty,y]} e^{-v}1_{\mathbb{R}^2_+(x<y)}(u,v)\, d\mu^2_L(u,v)\right).
\end{aligned}$$

Note that the above equality is trivially true if $x \leq 0$ or $y \leq 0$. In fact, in this case, we have

$$\frac{\partial^2 F_Z}{\partial y \partial x}(u,v) = 0,$$

for every $(u,v) \in$ is identically $(-\infty,x] \times (-\infty,y]$. Therefore, we consider only the case $x, y \in \mathbb{R}_{++}$ and distinguish two subcases $x > y$ and $x \leq y$. Observe that we can write

$$1_{\mathbb{R}^2_+(x>y)}(x,y) = 1_{[0,x)}(y)\,1_{\mathbb{R}_+}(x) = 1_{(y,+\infty)}(x)\,1_{\mathbb{R}_+}(y) \quad \text{and} \quad 1_{\mathbb{R}^2_+(x\leq y)}(x,y) = 1_{\mathbb{R}_+}(x)\,1_{[x,+\infty)}(y) = 1_{[0,y)}(x)\,1_{\mathbb{R}_+}(y),$$

for every $(x,y) \in \mathbb{R}^2$. Therefore, for every $x, y \in \mathbb{R}_{++}$ such that $x > y$, applying the Fubini therorem, we can write

$$\begin{aligned}
\int_{(-\infty,x]\times(-\infty,y]} e^{-u}1_{\mathbb{R}^2_+(x>y)}(u,v)\, d\mu^2_L(u,v) &= \int_{(-\infty,x]\times(-\infty,y]} e^{-u}1_{[0,u)}(v)\,1_{\mathbb{R}_+}(u)\, d\mu^2_L(u,v) \\
&= \int_{(-\infty,x]} e^{-u}1_{\mathbb{R}_+}(u)\, d\mu_L(u) \int_{(-\infty,y]} 1_{[0,u)}(v)\, d\mu_L(v) \\
&= \int_{(-\infty,x]\cap\mathbb{R}_+} e^{-u} d\mu_L(u) \int_{(-\infty,y]\cap[0,u)} d\mu_L(v).
\end{aligned}$$

Now,

$$\int_{(-\infty,y]\cap[0,u)} d\mu_L(v) = \mu_L((-\infty,y]\cap[0,u)) = \begin{cases} \mu_L([0,y)) = y, & \text{if } y < u, \\ \mu_L([0,u)) = u, & \text{if } y \geq u, \end{cases} = u1_{(-\infty,y]}(u) + y1_{(y,+\infty)}(u).$$

It follows,

$$\begin{aligned}
\int_{(-\infty,x]\times(-\infty,y]} e^{-u}1_{\mathbb{R}^2_+(x>y)}(u,v)\, d\mu^2_L(u,v) &= \int_{[0,x]} e^{-u}\left(u1_{(-\infty,y]}(u) + y1_{(y,+\infty)}(u)\right) d\mu_L(u) \\
&= \int_{[0,x]} ue^{-u}1_{(-\infty,y]}(u)\, d\mu_L(u) + \int_{[0,x]} ye^{-u}1_{(y,+\infty)}(u)\, d\mu_L(u) \\
&= \int_{[0,x]\cap(-\infty,y]} ue^{-u} d\mu_L(u) + y\int_{[0,x]\cap(y,+\infty)} e^{-u} d\mu_L(u)
\end{aligned}$$

where, since $x > y$,

$$\begin{aligned}
\int_{[0,x]\cap(-\infty,y]} ue^{-u} d\mu_L(u) + y\int_{[0,x]\cap(y,+\infty)} e^{-u} d\mu_L(u) &= \int_{[0,y]} ue^{-u} d\mu_L(u) + y\int_{(y,x]} e^{-u} d\mu_L(u) \\
&= \int_0^y ue^{-u} du + y\int_y^x e^{-u} du \\
&= 1 - ye^{-y} - e^{-y} + y\left(e^{-y} - e^{-x}\right) \\
&= 1 - e^{-y} - ye^{-x}.
\end{aligned}$$

On the other hand, if $y \geq x$, we have

$$\begin{aligned}
\int_{[0,x]\cap(-\infty,y]} ue^{-u} d\mu_L(u) + y\int_{[0,x]\cap(y,+\infty)} e^{-u} d\mu_L(u) &= \int_{[0,x]} ue^{-u} d\mu_L(u) + y\int_{\varnothing} e^{-u} d\mu_L(u) \\
&= \int_0^x ue^{-u} du \\
&= 1 - xe^{-x} - e^{-x}.
\end{aligned}$$

We can then write

$$\frac{1}{2}\int_{(-\infty,x]\times(-\infty,y]} e^{-u}1_{\mathbb{R}^2_+(x>y)}(u,v)\, d\mu^2_L(u,v) = \frac{1}{2}\left(1 - e^{-y} - ye^{-x}\right)1_{\mathbb{R}^2_+(x>y)}(x,y) + \frac{1}{2}\left(1 - e^{-x} - xe^{-x}\right)1_{\mathbb{R}^2_+(x\leq y)}(x,y).$$

Similarly,

$$\begin{aligned}\int_{(-\infty,x]\times(-\infty,y]} e^{-v}1_{\mathbb{R}^2_+(x<y)}(u,v)\,d\mu_L^2(u,v) &= \int_{(-\infty,x]\times(-\infty,y]} e^{-v}1_{[0,v)}(u)\,1_{\mathbb{R}_+}(v)\,d\mu_L^2(u,v)\\ &= \int_{(-\infty,y]} e^{-v}1_{\mathbb{R}_+}(v)\,d\mu_L(v)\int_{(-\infty,x]} 1_{[0,v)}(u)\,d\mu_L(v)\\ &= \int_{(-\infty,y]\cap\mathbb{R}_+} e^{-v}d\mu_L(v)\int_{(-\infty,x]\cap[0,v)} d\mu_L(v)\,.\end{aligned}$$

Now,

$$\int_{(-\infty,x]\cap[0,v)} d\mu_L(v) = \mu_L((-\infty,x]\cap[0,v)) = \begin{cases}\mu_L([0,x)) = x, & \text{if } x<v,\\ \mu_L([0,v)) = v, & \text{if } x\geq v,\end{cases} = v1_{(-\infty,x]}(u)+x1_{(x,+\infty)}(u)\,.$$

It follows,

$$\begin{aligned}\int_{(-\infty,x]\times(-\infty,y]} e^{-v}1_{\mathbb{R}^2_+(x<y)}(u,v)\,d\mu_L^2(u,v) &= \int_{[0,y]} e^{-v}\left(v1_{(-\infty,x]}(u)+x1_{(x,+\infty)}(u)\right)d\mu_L(v)\\ &\int_{[0,y]} ve^{-v}1_{(-\infty,x]}(u)\,d\mu_L(v)+\int_{[0,y]} xe^{-v}1_{(x,+\infty)}(u)\,d\mu_L(v)\\ &\int_{[0,y]\cap(-\infty,x]} ve^{-v}d\mu_L(v)+x\int_{[0,y]\cap(x,+\infty)} e^{-v}d\mu_L(v)\end{aligned}$$

where, in case $x>y$,

$$\begin{aligned}\int_{[0,y]\cap(-\infty,x]} ve^{-v}d\mu_L(v)+x\int_{[0,y]\cap(x,+\infty)} e^{-u}d\mu_L(u) &= \int_{[0,y]} ve^{-v}d\mu_L(v)+x\int_{\varnothing} e^{-u}d\mu_L(u)\\ &= \int_0^y ve^{-v}dv\\ &= 1-e^{-y}-ye^{-y}\end{aligned}$$

and, in case $x\leq y$,

$$\begin{aligned}\int_{[0,y]\cap(-\infty,x]} ve^{-v}d\mu_L(v)+x\int_{[0,y]\cap(x,+\infty)} e^{-u}d\mu_L(u) &= \int_{[0,x]} ve^{-v}d\mu_L(v)+x\int_{[x,y]} e^{-u}d\mu_L(u)\\ &= \int_0^x ve^{-v}dv+x\int_x^y e^{-u}du\\ &= 1-e^{-x}-xe^{-x}+x\left(e^{-x}-e^{-y}\right)\\ &= 1-e^{-x}-xe^{-y}.\end{aligned}$$

We then have

$$\frac{1}{2}\int_{(-\infty,x]\times(-\infty,y]} e^{-v}1_{\mathbb{R}^2_+(x<y)}(u,v)\,d\mu_L^2(u,v) = \frac{1}{2}\left(1-e^{-y}-ye^{-y}\right)1_{\mathbb{R}^2_+(x>y)}(x,y)+\frac{1}{2}\left(1-e^{-x}-xe^{-y}\right)1_{\mathbb{R}^2_+(x\leq y)}(x,y)\,.$$

Summarizing,

$$\int_{(-\infty,x]\times(-\infty,y]}\frac{\partial^2 F_Z}{\partial y\partial x}(u,v)\,d\mu_L^2(u,v) = \left(1-e^{-y}-\frac{1}{2}ye^{-x}-\frac{1}{2}ye^{-y}\right)1_{\mathbb{R}^2_+(x>y)}(x,y)+\left(1-e^{-x}-\frac{1}{2}xe^{-y}-\frac{1}{2}xe^{-x}\right)1_{\mathbb{R}^2_+}$$

As a consequence,

$$F(x,y)\neq\int_{(-\infty,x]\times(-\infty,y]}\frac{\partial^2 F_Z}{\partial y\partial x}(u,v)\,d\mu_L^2(u,v)\,,$$

which implies that Z is not absolutely continuous.

4. Despite Z is not absolutely continuous, X and Y, which are exponential random variables with rate parameter $\lambda=1$, are absolutely continuous with the same density $f:\mathbb{R}\to\mathbb{R}$, given by

$$f(z)=e^{-z}1_{\mathbb{R}_+}(z)\,,$$

for every $z\in\mathbb{R}$.

Problem 19 *Let Z_1,Z_2,Z_3 independent random variables on a probability space Ω such that such that $X_k\sim N(0,1)$, fo $k=1,2,3$. Consider the real random variables*

$$X_1\overset{def}{=}Z_1+Z_2+Z_3,\quad X_2\overset{def}{=}Z_1-Z_2+Z_3,\quad X_3\overset{def}{=}Z_1-Z_3.$$

1. *What is the distribution of the vector $X\equiv(X_1,X_2,X_3)^\intercal$?*
2. *Can you compute the distribution function of X?*
3. *Among the pairs (X_1,X_2), (X_1,X_3), and (X_2,X_3) of entries of X what are made by independent random variables?*
4. *Compute the distributions of X_1, X_2, and X_3;*
5. *Think on a quick and smart way to compute $\mathbf{E}\left[X_1X_2^2\right]$, $\mathbf{E}\left[X_1^2X_2^2\right]$, $\mathbf{E}\left[X_2X_3^2\right]$, $\mathbf{E}\left[X_2^2X_3^2\right]$.*

Solution. .

Variabili Random

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Random Variables 2022-12-08

Problem 1 *Let X be a geometrically distributed random variable with success probability p.*

1. *Are the moments of order 1 and 2 of X finite?*
2. *If the moments of order 1 and 2 are finite, can you compute $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*
3. *Assume that $p = 0.1$. Cand you apply the Tchebychev inequality to estimate $\mathbf{P}(-2 < X < 23)$.*

Solution. . □

Problem 2 *Let X be a uniform continuous real random variable with states in the interval $(0,1)$. In symbols, $X \sim Unif(0,1)$. Let $a, b \in \mathbb{R}$ such that $a < b$. Prove that the random variable $Y \overset{def}{=} a + (b-a)X$ is a uniform continuous real random variable with states in the interval (a,b). In symbols, $Y \sim Unif(a,b)$.*

Solution. .

Problem 3 *Let $X_1, \dots, X_n$ (totally) independent random variables uniformly distributed in the interval $(0,1)$. Hence, $X_1, \dots, X_n$ are absolutely continuous with density $f_{X_k} : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_{X_k}(x) \overset{def}{=} 1_{(0,1)}(x) \quad \forall x \in \mathbb{R},$$

on varying of $k = 1, \dots, n$. Consider the random variables

$$\check{X}_n \overset{def}{=} \max(X_1, \dots, X_n) \quad \text{and} \quad \hat{X}_n \overset{def}{=} \min(X_1, \dots, X_n).$$

Compute $\mathbf{E}[\check{X}_n]$ and $\mathbf{E}\left[\hat{X}_n\right]$.

Exercise 4 *Hint: it might be useful to determine the distribution functions $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ and $F_{\hat{X}_n} : \mathbb{R} \to \mathbb{R}_+$.*

Solution. .

Problem 5 *The decoration of a Christmas tree in a mall is made by 1.000 small light bulbs. The lifetime of each light bulb is exponentially distributed with an average life time of 20 days (rather cheap bulbs indeed!). The mall manager decides to turn on the lights of the Christmas tree on the midnight of the 15-th of November. Estimate the probability that at least 800 bulbs are still working on the midnight of the 25-th of December.*

Hints: compute the probability that each bulb will last until the midnight of the 25-th of December; write X for the random variable counting the number of light bulbs out of 1.000 which are still on at the midnight of the 25-th of December and guess how it is distributed; use the Markov inequality to make the estimate.

Solution. Let $T : \Omega \to \mathbb{R}_+$ be the random variable representing the life time of each bulb. Under our assumptions T is exponentially distributed with rate λ. That is T is absolutely continuous with density

$$f_T(x) \overset{\text{def}}{=} \lambda e^{-\lambda x} 1_{\mathbb{R}_+}(x), \quad \forall x \in \mathbb{R}.$$

Recall that

$$\begin{aligned}\int_{\mathbb{R}} f_T(x)\, d\mu_L(x) &= \int_{\mathbb{R}} \lambda e^{-\lambda x} 1_{\mathbb{R}_+}(x)\, d\mu_L(x) = \int_{\mathbb{R}_+} \lambda e^{-\lambda x} d\mu_L(x) = \int_0^{+\infty} \lambda e^{-\lambda x} dx \\ &= \lim_{x\to+\infty} \int_0^x \lambda e^{-\lambda u} du = -\lim_{x\to+\infty} \int_0^x e^{-\lambda u} d(-\lambda u) = -\lim_{x\to+\infty} \int_0^{-\lambda x} e^v dv \\ &= \lim_{x\to+\infty} e^v\big|_{-\lambda x}^0 = \lim_{x\to+\infty} \left(1 - e^{-\lambda x}\right) = 1.\end{aligned}$$

Recall also that

$$\mathbf{E}[X] = \int_{\mathbb{R}} x f_T(x)\, d\mu_L(x) = \int_{\mathbb{R}_+} \lambda x e^{-\lambda x} d\mu_L(x) = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \lim_{x\to+\infty} \int_0^x \lambda u e^{-\lambda u} du.$$

Now, integrating by parts,

$$\begin{aligned}\int_0^x \lambda u e^{-\lambda u} du &= -\int_0^x u\, d\left(e^{-\lambda u}\right) = -\, u e^{-\lambda u}\Big|_0^x + \int_0^x e^{-\lambda u} du = -\, u e^{-\lambda u}\Big|_0^x - \frac{1}{\lambda}\int_0^x e^{-\lambda u} d(-\lambda u) \\ &= -\, u e^{-\lambda u}\Big|_0^x - \frac{1}{\lambda}\int_0^{-\lambda x} e^v dv = -\, u e^{-\lambda u}\Big|_0^x + \frac{1}{\lambda} e^v\big|_{-\lambda x}^0 \\ &= -x e^{-\lambda x} + \frac{1}{\lambda}\left(1 - e^{-\lambda x}\right)\end{aligned}$$

It follows,

$$\mathbf{E}[X] = \frac{1}{\lambda}.$$

Hence, since the average life time of each bulb is 20 days, we set

$$\frac{1}{\lambda} = 20 \Leftrightarrow \lambda = \frac{1}{20}.$$

Then, we then have

$$\begin{aligned}\mathbf{P}(T \le t) &= \int_{(-\infty,t]} f_T(x)\, d\mu_L(x) = \int_{(-\infty,t]} \frac{1}{20} e^{-\frac{1}{20}x} 1_{\mathbb{R}_+}(x)\, d\mu_L(x) \\ &= \int_{(0,t]} \frac{1}{20} e^{-\frac{1}{20}x} d\mu_L(x) = \int_0^t \frac{1}{20} e^{-\frac{1}{20}x} dx = -\int_0^t e^{-\frac{1}{20}x} d\left(-\frac{1}{20}x\right) \\ &= -\int_0^{-\frac{t}{20}} e^u du = e^u\big|_{-\frac{t}{20}}^0 = 1 - e^{-\frac{t}{20}},\end{aligned}$$

for any $t \ge 0$. We are interested to compute the probability that a bulb is still working after 40 days it has been turned on. Hence, we have to set $t = 40$ and compute

$$\mathbf{P}(T > 40) = 1 - \mathbf{P}(T \le 40) = 1 - \left(1 - e^{-2}\right) = e^{-2} \simeq 0.13534.$$

Now, let $X : \Omega \to \mathbb{N}_0$ be the random variable representing the number of bulbs which are still working after 40 days. Since the bulbs in the circuit are parallel connected, we can assume that they are independent from each other. Therefore, X is a binomial random variable with parameters $n = 1000$ and $p = e^{-2}$. As a consequence, thanks to the Markov inequality, we can write

$$\mathbf{P}(X \geq 800) \leq \frac{\mathbf{E}[X]}{800} = \frac{1000e^{-2}}{800} = \frac{5}{4}e^{-2} \simeq 0.16917.$$

However, we have also

$$\mathbf{P}(X \geq 800) = 1-\mathbf{P}(X < 800) = 1-\sum_{k=0}^{799}\mathbf{P}(X = k) = 1-\sum_{k=0}^{799}\binom{1000}{k}e^{-2k}\left(1-e^{-2}\right)^{1000-k} \simeq 2.64233\cdot 10^{-14},$$

which shows that the bound obtained via the Markov inequality may be rather loose. □

Problem 6 *Let X be a geometrically distributed random variable with success probability $p = 0.1$. Compute $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$. Use the Tchebychev inequality to estimate $\mathbf{P}(-2 < X < 23)$.*

Solution. . □

Exercise 7 *An empathic professor aims to help his students to pass the hard final exam of his course in Probability and Statistics. To this goal, he splits the course program in two parts and gives his students an intermediate written test on the first part of the course. Assume that*

- *the 18% of the students attending the course who pass the final exam on their first try got a mark not lower than 25 in the intermediate test;*
- *the 24% of the students attending the course who pass the final exam on their first try got a mark in the range 20 − 24 in the intermediate test;*
- *the 30% of the students attending the course who pass the final exam on their first try got a mark not higher than 19 in the intermediate test;*
- *the 4% of the students attending the course who pass the final exam on their first try did not take the intermediate test.*

Assume also that

- *the 20% of the students attending the course get a mark not lower than 25 in the intermediate test;*
- *the 30% of the students attending the course get a mark in the range 20 − 24 in the intermediate test;*
- *the 10% of the students attending the course do not take the intermediate test.*

Compute:

1. *the probability that a student attending the course passes the final exam on her first try, given that she got a mark not lower than 25 in the intermediate test;*
2. *the probability that a student attending the course passes the final exam on her first try, given that she got a mark not higer than 19 in the intermediate test;*
3. *the probability that a student attending the course passes the final exam on her first try, given that she did not take the intermediate test;*
4. *the probability that a student attending the course passes the final exam on her first try;*
5. *the probability that a student attending the course got a mark not lower than 25 in the intermediate test, given that she passes the final exam on her first try;*
6. *the probability that a student attending the course got a mark not lower than 25 in the intermediate test, given that she does not pass the final exam on her first try;*
7. *the probability that a student attending the course does not pass the final exam on her first try given that she did not take the intermediate test.*

   Hint: write $T_{\geq 25}$ [resp. T_{20-24}, $T_{\leq 19}$] for the event "a randomly choosen student attending the course gets a mark not lower than 25 [resp. in the range 20 − 24, not higher than 19] in the intermediate test". Write also T_0 for the event "a randomly choosen student attendin

Solution. .

Problem 8 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f(a,b) \stackrel{def}{=} \mathbf{E}\left[(X-(a+bY))^2\right].$$

Prove that there exists

$$(a^*, b^*) \equiv \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\{f(a,b)\}$$

and compute it.

Solution. By virtue of the properties of the expectation operator, a straightforward computation yields

$$f(a,b) = a^2 + 2ab\mathbf{E}[Y] + b^2\mathbf{E}\left[Y^2\right] - 2a\mathbf{E}[X] - 2b\mathbf{E}[XY] + \mathbf{E}\left[X^2\right].$$

Therefore, $f(a,b)$ is a second order polynomial in the variables a and b. Consider the zeroes of the partial derivatives of $f(a,b)$, we have

$$\begin{aligned} \partial f_a(a,b) = 0 &\Leftrightarrow a + b\mathbf{E}[Y] = \mathbf{E}[X], \\ \partial f_b(a,b) = 0 &\Leftrightarrow a\mathbf{E}[Y] + b\mathbf{E}\left[Y^2\right] = \mathbf{E}[XY]. \end{aligned}$$

Thus, a point (a,b) is a candidate local minimum only if

$$\begin{aligned} a &= \frac{\begin{vmatrix} \mathbf{E}[X] & \mathbf{E}[Y] \\ \mathbf{E}[XY] & \mathbf{E}\left[Y^2\right] \end{vmatrix}}{\begin{vmatrix} 1 & \mathbf{E}[Y] \\ \mathbf{E}[Y] & \mathbf{E}\left[Y^2\right] \end{vmatrix}} = \frac{\mathbf{E}[X]\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]\mathbf{E}[XY]}{\mathbf{D}^2[Y]} \\ &= \frac{\mathbf{E}[X]\mathbf{E}\left[Y^2\right] - \mathbf{E}[X]\mathbf{E}[Y]^2 + \mathbf{E}[X]\mathbf{E}[Y]^2 - \mathbf{E}[Y]\mathbf{E}[XY]}{\mathbf{D}^2[Y]} \\ &= \frac{\mathbf{E}[X]\left(\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2\right) + \mathbf{E}[Y](\mathbf{E}[X]\mathbf{E}[Y] - \mathbf{E}[XY])}{\mathbf{D}^2[Y]} \\ &= \mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y], \\ b &= \frac{\begin{vmatrix} 1 & \mathbf{E}[X] \\ \mathbf{E}[Y] & \mathbf{E}[XY] \end{vmatrix}}{\begin{vmatrix} 1 & \mathbf{E}[Y] \\ \mathbf{E}[Y] & \mathbf{E}\left[Y^2\right] \end{vmatrix}} = \frac{\mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}{\mathbf{D}^2[Y]} = \frac{cov(X,Y)}{\mathbf{D}^2[Y]}. \end{aligned}$$

Moreover, we have

$$\partial^2 f_{a,a}(a,b) = 2 > 0, \qquad \partial^2 f_{a,b}(a,b) = 2\mathbf{E}[Y], \qquad \partial^2 f_{b,b}(a,b) = 2\mathbf{E}\left[Y^2\right] > 0.$$

Hence, the Hessian matrix Hf at (a,b) is given by

$$(Hf)(a,b) = \begin{pmatrix} 2 & 2\mathbf{E}[Y] \\ 2\mathbf{E}[Y] & 2\mathbf{E}\left[Y^2\right] \end{pmatrix}$$

and has determinant

$$\det\left((Hf)(a,b)\right) = 4\left(\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]\right) = 4\mathbf{D}^2[Y] > 0.$$

It follows that the point

$$(a^*, b^*) = \left(\mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y], \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\right)$$

is actually a local minimum. On the other hand, it is not difficult to show that $f(a,b)$ is a convex function. In fact, since

$$0 < \mathbf{D}^2[Y] = \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2,$$

we have

$$4\mathbf{E}[Y]^2 \leq 4\mathbf{E}\left[Y^2\right]$$

and

$$\left(\mathbf{E}\left[Y^2\right] + 1\right)^2 \geq 4\mathbf{E}[Y]^2 + \left(\mathbf{E}\left[Y^2\right] - 1\right)^2.$$

The latter implies

$$0 \leq \mathbf{E}\left[Y^2\right] + 1 - \sqrt{4\mathbf{E}[Y]^2 + (\mathbf{E}[Y^2] - 1)^2} \leq \mathbf{E}\left[Y^2\right] + 1 + \sqrt{4\mathbf{E}[Y]^2 + (\mathbf{E}[Y^2] - 1)^2},$$

that is to say, the eigenvalues the Hessian matrix $(Hf)(a,b)$ are positive. As a consequence,

$$(a^*, b^*) = \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\{f(a,b)\}.$$

Note that the first order Y-polynomial

$$\phi(Y) \stackrel{\text{def}}{=} \mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y] + \frac{cov(X,Y)}{\mathbf{D}^2[Y]}Y$$

turns out to be the best $(\sigma(Y), \mathcal{B}(\mathbb{R}))$-random first order Y-polynomial which approximates the $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$-random variable X w.r.t. the square norm.

Problem 9 *Students are allowed to take a test twice in an examination session. Assume that 8 students over 10 pass the test the first time. For those who fail or do not show up, only 6 students over 10 pass the test the second time.*

1. *Find the probability that a randomly selected student (who needs to pass the test) passes the test.*
2. *Assuming that a student passed the test what is the probability she passed on the first try?*
3. *Consider that a part of the first test presented the following problem: two dice are rolled and the number of the upper faces are observed. Is the event "the sum of the observed numbers is 7" independent of the event "the number observed on the upper face of a die is 5"? Could you give a solution to this problem?*

Solution. . □

Problem 10 *There is a group of n persons who checked their hat at a theatre. When they went to take their hats back the hatter went mad and started giving random persons random hats. What is the expected amount of persons who get their hat back? What is the probability that everyone gets their hat back?*

Hint: consider the k-th person of the group, for $k = 1, \dots, n$, and write C_k for the random variable expressing the circumstance that the k-th person gets her hat back.

Solution. . □

Problem 11 (Sheldon M. Ross - 4.2 - 4.3) *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{n\geq 1}$ be a sequence of independent Bernoulli random variables. Recall that for a Bernoulli random variable X we have*

$$X = \begin{cases} 1 & \mathbf{P}(X = 1) = p \\ 0 & \mathbf{P}(X = 0) = q \end{cases},$$

where $p \in (0,1)$ and $q \equiv 1 - p$. Consider the sequence $(Z_n)_{n\geq 1}$ of random variables given by

$$Z_n \stackrel{def}{=} \sum_{k=1}^n X_k, \quad \forall n \geq 1$$

and let $(H_n)_{n\geq 1}$ be the sequence of random variables given by

$$H_n \stackrel{def}{=} 2Z_n - n, \quad \forall n \geq 1.$$

1. *Assume that X_n is the random variable which represents the toss of a coin with the convention that "success" [resp. "failure"] is for the outcome "heads" [resp. "tails"] represented, in turn, by the outcome 1 [resp. 0]. Give an interpretation of the random variables Z_n and H_n and compute their mean and variance.*
2. *Assume that $p = 1/2$. Prove that we have*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{H_n}{\sqrt{n}} < z\right) = \Phi(z),$$

for every $z \in \mathbb{R}$, where Φ is the standard notation for the distribution function of the standard normal random variable.

Solution.

1. By virtue of the meaning of the random variable X_n it is clearly seen that the random variable Z_n represents the number of successes $k = 0, 1, \dots, n$ that we may obtain in n tosses and

$$H_n \stackrel{\text{def}}{=} 2Z_n - n = Z_n - (n - Z_n)$$

represents the number of heads minus the number of tails that we obtain in n tosses, namely the number of successes minus the number of failures that we may obtain in n tosses. Note that Z_n is binomially distributed with number of trials parameter n and success parameter p. That is

$$\mathbf{P}(Z_n = k) = \binom{n}{k} p^k q^{n-k}$$

for every $k = 0, 1, \ldots, n$. Now, we have

$$\mathbf{E}[X_n] = 1 \cdot \mathbf{P}(X=1) + 0 \cdot \mathbf{P}(X=0) = p$$

and

$$\mathbf{E}\left[X_n^2\right] = 1^2 \cdot \mathbf{P}(X=1) + 0^2 \cdot \mathbf{P}(X=0) = p.$$

These imply

$$\mathbf{D}^2[X_n] = \mathbf{E}\left[X_n^2\right] - \mathbf{E}[X_n]^2 = p - p^2 = p(1-p) = pq.$$

As a consequence,

$$\mathbf{E}[Z_n] = \mathbf{E}\left[\textstyle\sum_{k=1}^n X_k\right] = \textstyle\sum_{k=1}^n \mathbf{E}[X_k] = \sum_{k=1}^n p = np.$$

Furthermore, thanks to the independence of the random variables of the sequence $(X_n)_{n\geq 1}$,

$$\mathbf{D}^2[Z_n] = \mathbf{D}^2\left[\textstyle\sum_{k=1}^n X_k\right] = \textstyle\sum_{k=1}^n \mathbf{D}^2[X_k] = \sum_{k=1}^n pq = npq.$$

, since sum of independent random variables which are Bernoulli distributed with success parameter p We then have

$$\mathbf{E}[H_n] = \mathbf{E}[2Z_n - n] = 2\mathbf{E}[Z_n] - n = 2np - n = n(2p-1)$$

and

$$\mathbf{D}^2[H_n] = \mathbf{D}^2[2Z_n - n] = 4\mathbf{D}^2[Z_n] = 4npq.$$

2. Under the additional assumption $p = 1/2$, we have

$$\mathbf{E}[Z_n] = \frac{1}{2}n, \quad \mathbf{D}^2[Z_n] = \frac{1}{4}n, \quad \mathbf{E}[H_n] = 0, \quad \mathbf{D}^2[H_n] = n.$$

It follows

$$\mathbf{P}\left(\frac{H_n}{\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{2Z_n - n}{\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{Z_n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{Z_n - \mathbf{E}[Z_n]}{\mathbf{D}[Z_n]} < z\right) \tag{1}$$

On the other hand, by the Central limit Theorem, we have

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z_n - \mathbf{E}[Z_n]}{\mathbf{D}[Z_n]} < z\right) = \Phi(z), \tag{2}$$

for every $z \in \mathbb{R}$. Combining (1) and (2), the desired claim immediately follows.

Problem 12 *Suppose that we roll a standard fair die* 100 *times. Let X be the sum of the numbers that appear over the* 100 *rolls. Use the Tchebychev inequality to bound* $\mathbf{P}[(|X - 350| \geq 50)$.

Solution. . □

Problem 13 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let X and Y discrete real random variables on Ω. Write* $X(\Omega) \equiv \{x_m\}_{m\in M}$ *and* $Y(\Omega) \equiv \{y_n\}_{n\in N}$, *where* $M, N \subseteq \mathbb{N}$. *Prove that X and Y are independent if and only if*

$$\mathbf{P}(X = x_m, Y = y_n) = \mathbf{P}(X = x_m)\,\mathbf{P}(Y = y_n), \quad \forall (m,n) \in M \times N. \tag{3}$$

Solution. The condition is clearly necessary. To prove that it is also sufficient, consider any couple $B, C \in \mathcal{B}(\mathbb{R})$. We have

$$\{X \in B\} = \bigcup_{m|x_m\in B} \{X = x_m\} \quad \text{and} \quad \{Y \in C\} = \bigcup_{n|y_n\in C} \{Y = y_n\}.$$

Therefore,

$$\begin{aligned}
\{X \in B, Y \in C\} &= \{X \in B\} \cap \{Y \in C\} \\
&= \left(\bigcup_{m|x_m\in B} \{X = x_m\}\right) \cap \left(\bigcup_{n|y_n\in C} \{Y = y_n\}\right) \\
&= \bigcup_{m|x_m\in B,\ n|y_n\in C} \{X = x_m\} \cap \{Y = y_n\} \\
&\equiv \bigcup_{m|x_m\in B,\ n|y_n\in C} \{X = x_m, Y = y_n\},
\end{aligned}$$

where the events of the family $\{\{X = x_m, Y = y_n\}\}_{m|x_m\in B,\ n|y_n\in C}$ are pairwise incompatible. As a consequence, under Assumption (3), we have

$$\begin{aligned}
\mathbf{P}(X \in B, Y \in C) &= \mathbf{P}\left(\bigcup_{m|x_m\in B,\ n|y_n\in C} \{X = x_m, Y = y_n\}\right) \\
&= \sum_{m|x_m\in B\ ,n|y_n\in C} \mathbf{P}(X = x_m, Y = y_n) \\
&= \sum_{m|x_m\in B\ ,n|y_n\in C} \mathbf{P}(X = x_m)\,\mathbf{P}(Y = y_n) \\
&= \left(\sum_{m|x_m\in B} \mathbf{P}(X = x_m)\right)\left(\sum_{n|y_n\in C} \mathbf{P}(Y = y_n)\right) \\
&= \mathbf{P}\left(\bigcup_{m|x_m\in B} \{X = x_m\}\right)\mathbf{P}\left(\bigcup_{n|y_n\in C} \{Y = y_n\}\right) \\
&= \mathbf{P}(X \in B)\,\mathbf{P}(Y \in C).
\end{aligned}$$

This, thanks to the arbitrariness of $B, C \in \mathcal{B}(\mathbb{R})$, yields the independence of X and Y.

.

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2021-10-12

Problem 1 *Let Ω be the sample space of a random phenomenon and let $\mathcal{E}_1$, $\mathcal{E}_2$ algebras [resp. σ-algebras] of events of Ω. May we say that the family $\mathcal{E}_1 \cup \mathcal{E}_2$ of events of Ω given by*

$$\mathcal{E}_1 \cup \mathcal{E}_2 \overset{def}{=} \{E \in \mathcal{P}(\Omega) : E \in \mathcal{E}_1 \text{ or } E \in \mathcal{E}_2\}$$

is an algebra [resp. σ-algebras]?

Solution. Clearly, since $\mathcal{E}_1$, $\mathcal{E}_2$ algebras [resp. σ-algebras] of events of Ω, the family $\mathcal{E}_1 \cup \mathcal{E}_2$ is not empty. Now, assume that an event E is in $\mathcal{E}_1 \cup \mathcal{E}_2$, then E is in $\mathcal{E}_1$ or E is in $\mathcal{E}_2$. As a consequence, E^c is in $\mathcal{E}_1$ or E^c is in $\mathcal{E}_2$. Hence, $E^c \in \mathcal{E}_1 \cup \mathcal{E}_2$. However, assuming that E and F are in $\mathcal{E}_1 \cup \mathcal{E}_2$, unless they are both in $\mathcal{E}_1$ or $\mathcal{E}_2$, there is no reason why $E \cup F$ should be in $\mathcal{E}_1 \cup \mathcal{E}_2$. This is confirmed by the following example: with reference to the die sample space $\Omega \equiv \{\omega_1, \ldots, \omega_6\}$ choose

$$\mathcal{E}_1 \equiv \{\varnothing, \{\omega_1, \omega_3, \omega_5\}, \{\omega_2, \omega_4, \omega_6\}, \Omega\} \quad \text{and} \quad \mathcal{E}_2 \equiv \{\varnothing, \{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5, \omega_6\}, \Omega\}.$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ are algebras of events of Ω, but

$$\mathcal{E}_1 \cup \mathcal{E}_2 = \{\varnothing, \{\omega_1, \omega_2, \omega_3\}, \{\omega_1, \omega_3, \omega_5\}, \{\omega_2, \omega_4, \omega_6\} \{\omega_4, \omega_5, \omega_6\}, \Omega\}$$

is not. Note that $\mathcal{E}_1 \cup \mathcal{E}_2$ is closed with respect to the complement operator, but not with respect to the union.

Problem 2 *Let Ω be the infinite sample space of a random phenomenon. The family*

$$\mathcal{E}_{count} \equiv \{E \in \mathcal{P}(\Omega) : |E| \leq \aleph_0\}$$

of all countable events of Ω is a σ-algebra of events of Ω if and only if Ω itself is countable. In this case, we have $\mathcal{E}_{count} = \mathcal{P}(\Omega)$. On the other hand, the family $\mathcal{E}_{count\text{-}cocount}$ of all events of Ω that are countable or have countable complement, in symbols

$$\mathcal{E}_{count\text{-}cocount} \equiv \{E \in \mathcal{P}(\Omega) : |E| \leq \aleph_0 \ \vee \ |E^c| \leq \aleph_0\},$$

is a σ-algebra of events of Ω.

Solution. Given any $\omega \in \Omega$ we have

$$|\{\omega\}| = 1 \leq \aleph_0.$$

Hence, $\{\omega\} \in \mathcal{E}_{\text{count}}$. Assume that $\mathcal{E}_{\text{count}}$ is a σ-algebra, then also $\{\omega\}^c \equiv \Omega - \{\omega\}$ is in $\mathcal{E}_{\text{count}}$. By definition, it follows

$$|\Omega - \{\omega\}| \leq \aleph_0,$$

which clearly implies

$$|\Omega| \leq \aleph_0.$$

Conversely, if Ω is countable, then every $E \in \mathcal{P}(\Omega)$ is countable. This implies

$$\mathcal{P}(\Omega) \subseteq \mathcal{E}_{\text{count}},$$

that is

$$\mathcal{E}_{\text{count}} = \mathcal{P}(\Omega).$$

As a trivial consequence, $\mathcal{E}_{\text{count}}$ is σ-algebra of events of Ω. As a consequence of the above argument, if Ω is not countable, that is

$$|\Omega| > \aleph_0$$

or, according the continuum hypothesis,

$$|\Omega| \geq \aleph_1,$$

the family $\mathcal{E}_{\text{count}}$ cannot be a σ-algebra. On the other hand, the family $\mathcal{E}_{\text{count-cocount}}$ is. In fact, clearly $\mathcal{E}_{\text{count-cocount}} \neq \varnothing$. Furthermore, if $E \in \mathcal{E}_{\text{count-cocount}}$, according to the definition, we have two cases:

$$|E| \leq \aleph_0 \quad \text{or} \quad |E^c| \leq \aleph_0.$$

In the first case,

$$|(E^c)^c| = |E| \leq \aleph_0.$$

This implies that $E^c \in \mathcal{E}_{\text{count-cocount}}$. In the second case, we have $E^c \in \mathcal{E}_{\text{count-cocount}}$ straightforwardly. Hence, in either cases $E^c \in \mathcal{E}_{\text{count-cocount}}$. In the end, consider a sequence $(E_n)_{n \geq 1}$ of elements in $\mathcal{E}_{\text{count-cocount}}$. If $|E_n| \leq \aleph_0$ for every $n \in \mathbb{N}$, then

$$\left| \bigcup_{n \geq 1} E_n \right| \leq \aleph_0,$$

which implies $\bigcup_{n \geq 1} E_n \in \mathcal{E}_{\text{count-cocount}}$. Otherwise, there exists at least $n_0 \in \mathbb{N}$ such that $|E_{n_0}| > \aleph_0$. However, in this case, since $E_{n_0} \in \mathcal{E}_{\text{count-cocount}}$, we necessarily have

$$|E^c_{n_0}| \leq \aleph_0.$$

This implies

$$\left| \left(\bigcup_{n \geq 1} E_n \right)^c \right| = \left| \bigcap_{n \geq 1} E^c_n \right| \leq |E^c_{n_0}| \leq \aleph_0.$$

Thus, it still follows that $\bigcup_{n \geq 1} E_n \in \mathcal{E}_{\text{count-cocount}}$.

Problem 3 *Let Ω be the sample space of a random phenomenon or experiment, let $\mathcal{E}$ be an algebra of events of Ω and let $\mathbf{P} : \mathcal{E} \to \mathbb{R}_+$ be an additive probability on Ω. Prove that we have*

1. $\mathbf{P}(\varnothing) = 0$;
2. $\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$ *for any* $E \in \mathcal{E}$;
3. $\mathbf{P}(F - E) = \mathbf{P}(F) - \mathbf{P}(E \cap F)$ *for all* $E, F \in \mathcal{E}$, *in particular* $\mathbf{P}(F - E) = \mathbf{P}(F) - \mathbf{P}(E)$ *when* $E \subseteq F$;
4. $\mathbf{P}(E) \leq \mathbf{P}(F)$ *for all* $E, F \in \mathcal{E}$ *such that* $E \subseteq F$;
5. $\mathbf{P}(E) \leq 1$ *for any* $E \in \mathcal{E}$;
6. $\mathbf{P}(\bigcup_{k=1}^n E_k) = \sum_{k=1}^n \mathbf{P}(E_k)$ *for any finite sequence* $(E_k)_{k=1}^n$ *in* $\mathcal{E}$ *of pairwise exclusive events;*

7. $\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)$ *for all* $E, F \in \mathcal{E}$, *in particular* $\mathbf{P}(E \cup F) \leq \mathbf{P}(E) + \mathbf{P}(F)$;
8. $\mathbf{P}\left(\bigcup_{k=1}^{n} E_k\right) \leq \sum_{k=1}^{n} \mathbf{P}(E_k)$ *for any finite sequence* $(E_k)_{k=1}^{n}$ *in* $\mathcal{E}$.

Solution. See Notes on Probability and Statistics, Proposition 146.

Problem 4 *Let* Ω *be the sample space of a random phenomenon or experiment, let* $\mathcal{E}$ *be a* σ*-algebra of events of* Ω *and let* $\mathbf{P} : \mathcal{E} \to \mathbb{R}_+$ *be a countably additive probability on* Ω. *Prove that we have*

1. $\mathbf{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbf{P}(E_k)$ *for any finite sequence* $(E_k)_{k=1}^{n}$ *of pairwise incompatible events in* $\mathcal{E}$;
2. $\mathbf{P}\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mathbf{P}(E_n)$ *for any sequence* $(E_n)_{n\geq 1}$ *in* $\mathcal{E}$.

Solution.

1. Let $(E_k)_{k=1}^{n}$ be a finite sequence of pairwise incompatible events in $\mathcal{E}$. Then the sequence $(F_k)_{k\geq 1}$ given by
$$F_k \stackrel{\text{def}}{=} \begin{cases} E_k & \forall k = 1, \ldots, n \\ \varnothing & \forall k \geq n+1 \end{cases},$$
is a denumerable sequence of pairwise incompatible events in $\mathcal{E}$ such that
$$\bigcup_{k=1}^{n} E_k = \bigcup_{k=1}^{\infty} F_k$$
and
$$\mathbf{P}(F_k) = \begin{cases} \mathbf{P}(E_k) & \forall k = 1, \ldots, n \\ 0 & \forall k \geq n+1 \end{cases}.$$
As a consequence,
$$\mathbf{P}\left(\bigcup_{k=1}^{n} E_k\right) = \mathbf{P}\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mathbf{P}(F_k) = \sum_{k=1}^{n} \mathbf{P}(F_k) = \sum_{k=1}^{n} \mathbf{P}(E_k),$$
This proves that is an additive probability
2. Let $(E_n)_{n\geq 1}$ be any sequence of events in $\mathcal{E}$. Then the sequence $(F_n)_{n\geq 1}$ given by
$$F_n \stackrel{\text{def}}{=} \begin{cases} E_1 & \text{if } n = 1 \\ E_n - \bigcup_{k=1}^{n-1} E_k & \text{if } n > 1 \end{cases},$$
is a sequence of pairwise incompatible events in $\mathcal{E}$ such that
$$\bigcup_{n\geq 1} E_n = \bigcup_{n\geq 1} F_n.$$
In fact, clearly $F_n \subseteq E_n$, for every $n \geq 1$. Hence,
$$\bigcup_{n\geq 1} F_n \subseteq \bigcup_{n\geq 1} E_n.$$
Conversely, if $x \in \bigcup_{n\geq 1} E_n$, then the set $N_x \equiv \{n \in \mathbb{N} : x \in E_n\} \neq \varnothing$. Write $\hat{n}_x \equiv \min(N_x)$. We have $x \in E_{\hat{n}_x}$. In case $\hat{n}_x = 1$, by definition we have $x \in F_1$. In case $\hat{n}_x > 1$, we have $x \notin E_k$ for every $k = 1, \ldots, \hat{n}_x - 1$. Hence, $x \notin \bigcup_{k=1}^{\hat{n}_x - 1} E_k$ and, again by definition, $x \in F_{\hat{n}_x}$. Therefore, in any case, we obtain $x \in F_{\hat{n}_x}$. This implies that $x \in \bigcup_{n\geq 1} F_n$.
Now, given $n_1, n_2 \in \mathbb{N}$ such that $n_1 \neq n_2$, we have
$$F_{n_1} \cap F_{n_2} = \varnothing.$$
In fact, assuming for instance $n_1 < n_2$, we have
$$F_{n_1} \subseteq E_{n_1} \quad \text{and} \quad F_{n_2} \cap E_{n_1} = \left(E_{n_2} - \bigcup_{k=1}^{n_2-1} E_k\right) \cap E_{n_1} = \varnothing.$$
This implies that the events of the sequence $(F_n)_{n\geq 1}$ are pairwise incompatible. As a consequence of the above arguments, it follows
$$\mathbf{P}\left(\bigcup_{n\geq 1} E_n\right) = \mathbf{P}\left(\bigcup_{n\geq 1} F_n\right) = \sum_{n\geq 1} \mathbf{P}(F_n) \leq \sum_{n\geq 1} \mathbf{P}(E_n),$$
which is the desired result.

Problem 5 *Let* Ω *be the sample space of a random phenomenon or experiment, let* $\mathcal{E}$ *be a* σ*-algebra of events of* Ω *and let* $\mathbf{P} : \mathcal{E} \to \mathbb{R}_+$ *be a countably additive probability on* Ω. *A sequence* $(E_n)_{n\geq 1}$ *of events in* $\mathcal{E}$ *is said to be increasing [resp. decreasing] if*
$$E_n \subseteq E_{n+1} \quad [\textit{resp. } E_n \supseteq E_{n+1}] \quad \forall n \in \mathbb{N}.$$
Let $(E_n)_{n\geq 1}$ *be an increasing sequence of events belonging to* $\mathcal{E}$. *Prove that*
$$\mathbf{P}\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mathbf{P}(E_n).$$
Use this result to show that for a decreasing sequence $(E_n)_{n\geq 1}$ *of events belonging to* $\mathcal{E}$ *we have*
$$\mathbf{P}\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mathbf{P}(E_n).$$

Solution. .

Problem 6 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *a probability space and let* $E, F \in \mathcal{E}$ *such that*
$$\mathbf{P}(E) + \mathbf{P}(F) \geq 1. \tag{1}$$
Prove that
$$\mathbf{P}(E) + \mathbf{P}(F) - 1 \leq \mathbf{P}(E \cap F) \leq \min\{\mathbf{P}(E), \mathbf{P}(F)\} \tag{2}$$
Determine a similar lower and upper bound for $\mathbf{P}(E \cap F)$ *under the assumption*
$$\mathbf{P}(E) + \mathbf{P}(F) < 1. \tag{3}$$

Solution. .

Problem 7 *Five Italian players are playing poker. The deck of poker cards contains 36 cards of the usual ranks ($6, 7, 8, 9, 10, J, Q, K, A$) and of the usual suites (hearts $\heartsuit$, clubs $\spadesuit$, diamonds $\diamondsuit$, flowers $\clubsuit$).*

1. *How many hands are possible by a random deal?*
2. *How many hands give a straight flush by a random deal?*
3. *How many hands give a four of a kind by a random deal?*
4. *How many hands give a flush by a random deal?*
5. *How many hands give a full house by a random deal?*
6. *How many hands give a straight by a random deal?*
7. *How many hands give a three of a kind by a random deal?*
8. *How many hands give two pair by a random deal?*
9. *How many hands give one pair by a random deal?*
10. *How many hands give no pair by a random deal?*
11. *How many hands fail to give any of the above combinations by a random deal?*
12. *What about if the players are Americans? In this case the deck of poker card contains 56 cards of the usual ranks ($1, \ldots, 10, J, Q, K, A$) and of the usual suites.*

Solution.

1. Since a poker hand is indifferent to the order in which is arranged by the deal, the number of all possible hands is just the number of all possible subsets of 5 elements that can be selected from a set of 36 elements. Hence,
$$\binom{36}{5}$$
is the number of all possible hands.

2. According to the (Italian) poker rules, there are 6 possibilities for choosing the rank of the first card of a straight. The ranks of the other cards are then consequently determined. That is the ony possible straights in a deck of 36 cards are
$$A, 6, 7, 8, 9; \quad 6, 7, 8, 9, 10; \quad ; \ldots ; \quad 10, J, Q, K, A.$$
We have a straight flush when the cards of the straight have all the same suits. We can choose the suit for the straight flush in 4 differeent ways. Hence, we have
$$4 \cdot 6$$
possible hands giving a straight flush by a random deal.

3. There are 9 possibility for choosing the rank of card for the four of a kind, once the card has been chosen there is no room for the choice of the suites. Then, there are 8 possibilities for choosing the rank of fifth card and for each rank there are $\binom{4}{1} = 4$ possibilities for choosing the suits. As a consequence, we have
$$9 \cdot 8 \cdot 4$$
possible hands giving a four of a kind by a random deal.

4. To be continued.

Problem 8 *An urn contains n distinguishable balls of which r are red, with $1 \le r < n$, and $n - r$ are white.The urn is shaken and the balls are drawn from the urn one after the other without replacement. How many of the possible drawn sequences show the first red ball at the kth draw?*

Ans.
$$\binom{n-r}{k-1}(k-1)!r\,(n-k)!$$

Solution. .

Problem 9 *An urn contains n balls of which r are red, with $1 \le r < n$, and $n - r$ are white. The urn is shaken and the balls are drawn one after the other without replacement. Suppose that both the red balls and the white ones are undistinguishable among them. How many of the possible drawn sequences show the first red ball at the kth draw?*

Ans.
$$\binom{n-k}{r-1}$$

Solution. .

Problem 10 *An urn contains n distinguishable balls of which r are red, with $1 \le r < n$, and $n - r$ are white. The urn is shaken and k balls are drawn without replacement. If $k \le r$, how many of the possible unordered samples contains $s \le k$ red balls and $k - s$ white ones*

Ans.
$$C_{r,s} \cdot C_{n-r,k-s}.$$

Solution. .

Problem 11 *The key that opens a room is in a box containing $n \ge 1$ keys. How many ordered samples of keys containing the right key at the kth place ($1 \le k \le n$) we can draw from?*

Ans.
$$(n-1)!$$

Solution. .

Problem 12 *An urn contains N different balls numbered from 1 to N. A sample of $n \leq N$ balls is drawn from the urn without replacement. Show that the probability to obtain a given combination of $k \leq n$ elements is*

$$p_{N,n,k} = \frac{(n-k+1) \cdot \ldots \cdot (n-1)\, n}{(N-k+1) \cdot \ldots \cdot (N-1)\, N}.$$

Note that when $n = k$ we obtain

$$p_{N,n,n} = \frac{n!}{N\,(N-1) \cdot \ldots \cdot (N-n+1)}.$$

Solution. .

Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac statistics

Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac Statistics consider the number of ways in which m objects can be placed into n cells, according to whether the objects are distinguishable from each other and the cells can contain more than one or only one object.

Problem 13 (Maxwell-Boltzman Statistics) *Suppose we have m distinguishable balls and n boxes. How many ways are there to distribute the balls in the boxes if each box can contain more than one ball? How many ways are there to distribute the balls in the boxes if a specific box has to contain $\ell \leq m$ balls?*

Solution. .

Problem 14 (Bose-Einstein Statistics) *Suppose we have m indistinguishable balls and n boxes. How many ways are there in which the balls can be distributed in the boxes if every box can contain more than one ball?*

Solution. .

Problem 15 (Fermi-Dirac Statistics) *Suppose we have m indistinguishable balls and n boxes. How many ways are there in which the balls can be distributed in the boxes if every box cannot contain more than one ball?*

Solution. .

non risolti ma spesso gli esercizi il prof li prende da questo gruppetto qui

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2021-10-13

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $E, F \in \mathcal{E}$. Show that E and F are independent if and only if:*

1. *E^c and F^c are independent;*
2. *E and F^c are independent;*
3. *E^c and F are independent.*

Solution. .

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(E_k)_{k=1}^n$ a finite sequence of independent events of Ω.*

1. *Show that the probability that none of the events of the sequence $(E_k)_{k=1}^n$ occurs is less or equal to $\exp\left(-\sum_{k=1}^n \mathbf{P}(E_k)\right)$.*
   Hint: consider the elementary properties of the exponential and logarithmic functions.
2. *Let E, F independent events of Ω such that $\mathbf{P}(F) > 0$. Prove that*
$$\mathbf{P}(E \cap F \mid E \cup F) \leq \mathbf{P}(E \cap F \mid F).$$
3. *Let E, F, G independent events of Ω such that $\mathbf{P}(E \cap F) > 0$. Prove that*
$$\mathbf{P}(G \mid E \cap F) = \mathbf{P}(G).$$

Solution. .

Problem 3 *A urn, say Urn A, contains 5 white balls and 10 black balls. Another urn, say Urn B, contains 3 white balls and 12 black balls. A fair coin is tossed. If it shows heads [resp. tails] a ball is drawn from Urn A [resp. B]. Suppose that this random experiment has been done and we know that a white ball has been drawn. What is the probability that the ball has been drawn from Urn A? What is the probability that the ball has been drawn from Urn B?*

Solution. .

Problem 4 *An urn contains N balls of which M are white and the remaining $N - M$ are black. From the urn the balls are drawn* with replacement. *We know that the naive probability of drawing the white ball at the nth drawn is M/N. Show that the naive probability of drawing the white ball at the n-th draw, given that all previous draws have failed is again M/N. Assume then that the balls are drawn* without replacement. *Show that the naive probability of drawing a white ball at the n-th draw is still M/N and the naive probability of drawing the white ball at the n-th draw, given that all previous draws have failed is $M/(N - n + 1)$, where $n \leq N - M + 1$.*

Solution. .

Problem 5 *Students are allowed to take a test twice in an examination session. Assume that 8 students over 10 pass the test the first time. For those who fail or do not show up, only 6 students over 10 pass the test the second time.*

1. *Find the probability that a randomly selected student (who needs to pass the test) passes the test.*
2. *Assuming that a student passed the test what is the probability she passed on the first try?*
3. *Consider that a part of the first test presented the following problem: two dice are rolled and the number on the upper faces are observed. Is the event "the sum of the observed numbers is 7" independent of the event "the number observed on the upper face of a die is 5"? Could you give a solution to this problem?*

Solution. .

Problem 6 *The National Health Service (NHS) aims to introduce a new test for the screening of a disease. The pharmaceutical company which produces the test states that:*

- *the test yields a positive result on the 95% of people who are affected by the disease (*sensitivity *or* true positive rate *of the test);*
- *the test yields a negative result on the 99% of people who are not affected by the disease (*specificity *or* true negative rate *of the test);*

On the other hand, the NHS knows the the disease is currently affecting the 10% of the population. Compute:

1. *the probability that a randomly chosen individual of the population is affected by the disease given that the test yields a positive result;*
2. *the probability that a randomly chosen individual of the population is not affected by the disease given that the test yields a positive result;*
3. *the probability that a randomly chosen individual of the population is affected by the disease given that the test yields a negative result;*
4. *the probability that a randomly chosen individual of the population is not affected by the disease given that the test yields a negative result;*
5. *the probability that the test yields a positive result on a randomly chosen individual of the population.*
6. *the probability that the test yields a negative result on a randomly chosen individual of the population.*

Solution. .

Problem 7 *The scrutiny of group of* $100,000$ *randomly chosen male people in the age* $40-79$ *in UK during* $2013-2015$ *reveals the following table of average lung cancer incidence*

	smoker	*not smoker*	*total*
lung cancer	$10,395$	$7,407$	$17,802$
not lung cancer	$50,078$	$32,120$	$82,198$
	$60,473$	$39,527$	$100,000$

Write Ω *for the sample space consisting of these* $100,000$ *people and write* S *[resp.* C*] for the subsets of* Ω *consisting of the smokers [the people affected by lung canger]. Let* $1_S : \Omega \to \{0,1\}$ *and* $1_C : \Omega \to \{0,1\}$ *the indicator functions of the events* S *and* C *respectively.*

1. *Determine the joint distribution of the random vector* $(1_S, 1_C)$ *and the marginal distributions of the random variables* 1_S *and* 1_C*.*
2. *Are the random variables* 1_S *and* 1_C *independent?*
3. *What is the probability that a randomly chosen person in* Ω *is affected by lung cancer, given that he is a smoker?*
4. *What is the probability that a randomly chosen person in* Ω *is a smoker, given that he is affected by lung cancer?*
5. *Check the validity of the total probability formula for* $\mathbf{P}(S)$ *and the Bayes Formula for* $\mathbf{P}(C \mid S)$*.*

Solution. .

Problem 8 *Consider an e-mail server equipped with a spam filter. The spam filter scrutinizes each incoming e-mail for the appearance of some key sentences.which make it likely that the message is spam. Assume that a key sentence is "check this out" in the the subject line of the e-mail. Assume also that*

1. 40% *of incoming e-mails are spam;*
2. 1% *of emails contain the sentence "check this out" in the subject line, given that they are spam;*
3. 0.4% *of emails contain the sentence "check this out" in the subject line, given that they are not spam.*

   Compute the probability that an e-mail is spam given that it contains the sentence "check this out" in the subject line.

Solution. .

Problem 9 *In a large town, after a robbery, a thief jumped into a cab and disappeared. An eyewitness on the crime scene told the police that the cab is yellow. Having some doubt on the reliability of the eyewitness, the police consulted a mathematician. Assuming that*

1. 20% *of the cabs in the town are yellow;*
2. *from the past experience police knows that an eyewitness is %80 accurate, that an eyewitness identifies correctly the colour of a taxi yellow or not yellow* 8 *out* 10 *times.*

   Compute the probability that the information reported by the eyewitness is true. That is the probability that the cab was yellow given that the eyewitness said so. Do you think this information is useful?

   Hint: consider the events "the cab is yellow", "the cab is not yellow", "the eyewitness says the cab is yellow", and "the eyewitness says the cab is not yellow" and formulate in terms of conditional probability the accuracy of the eyewitness.

Solution. .

Problem 10 *A box contains five coins labeled by* $C_1, \dots, C_5$*. Assume that both the faces of the coin* C_1 *[resp.* C_5*] show head [resp. tail]; assume that both the coins* C_2 *and* C_4 *are rigged in such a way that the probability of getting head [resp. tail] when tossing the coin* C_2 *[resp.* C_4*] is double than the probability of getting tail [resp. head]; in the end assume that* C_3 *is a fair coin. A coin is randomly drawn from the box and is flipped* 5 *times. It happens that* 5 *heads show up in a row, what is the probability that we have drawn the coin* C_k*, for* $k = 1, \dots, 5$*.*

Hint: given a probability space $(\Omega, \mathcal{E}, \mathbf{P})$ *and three events* $E, F, G \in \mathcal{E}$ *it may occur that*

$$\mathbf{P}(E \cap F \mid G) = \mathbf{P}(E \mid G)\,\mathbf{P}(F \mid G).$$

In this case, the events E *and* F *are said to be* conditionally independent *given* G*.*

Solution. .

Problem 11 *An empathic professor aims to help his students to pass the hard final exam of his course in Probability and Statistics. To this goal he splits the course propram in two parts and gives his students an intermediate written test on the first part of the course. Assume that*

- *the* 18% *of the students who pass the final exam on their first try got a mark not lower than* 25 *in the intermediate test;*
- *the* 24% *of the students who pass the final exam on their first try got a mark in the range* $20-24$ *in the intermediate test;*
- *the* 30% *of the students who pass the final exam on their first try got a mark not higher than* 19 *in the intermediate test;*
- *the* 4% *of the students who pass the final exam on their first try did not take the intermediate test.*

  Assume also that

- *the* 20% *of the students attending the course get a mark not lower than* 25 *in the intermediate test;*
- *the* 30% *of the students attending the course get a mark in the range* $20-24$ *in the intermediate test;*
- *the* 10% *of the students attending the course do not take the intermediate test.*

Exercise 12 *Compute:*

1. *the probability that a student passes the final exam on her first try, given that she gets a mark not lower than* 25 *in the intermediate test;*

2. *the probability that a student passes the final exam on her first try, given that she gets a mark not higer than* 19 *in the intermediate test;*

3. *the probability that a student passes the final exam on her first try, given that she does not take the intermediate test;*

4. *the probability that a student passes the final exam on her first try;*

5. *the probability that a student gets a mark not lower than* 25 *in the intermediate test, given that she passes the final exam on her first try;*

6. *the probability that a student gets a mark not lower than* 25 *in the intermediate test, given that she does not pass the final exam on her first try;*

7. *the probability that a student does not pass the final exam on her first try, given that she does not take the intermediate test.*

Solution. Write $T_{\geq 25}$ [resp. T_{20-24}, $T_{\leq 19}$] for the event "a randomly choosen student attending the course got a mark not lower than 25 [resp. in the range $20-24$, not higher than 19] in the intermediate test". Write also T_0 for the event "a randomly choosen student did not take the intermediate test". In the end, write S for the event "a randomly choosen student attending the course passed the final exam on her first try".

We clearly have that $\{T_0, T_{\leq 19}, T_{20-24}, T_{\geq 25}\}$ is a partition of the sure event Ω. In addition, we know that

$$\mathbf{P}(T_{\geq 25}) = 0.20, \quad \mathbf{P}(T_{20-24}) = 0.30, \quad \mathbf{P}(T_0) = 0.10.$$

It follows

$$\mathbf{P}(T_{\leq 19}) = 1 - (\mathbf{P}(T_{\geq 25}) + \mathbf{P}(T_{20-24}) + \mathbf{P}(T_0)) = 1 - (0.20 + 0.30 + 0.10) = 0.40.$$

We also know the probabilities

$$\mathbf{P}(S \cap T_{\geq 25}) = 0.18, \quad \mathbf{P}(S \cap T_{20-24}) = 0.24, \quad \mathbf{P}(S \cap T_{\leq 19}) = 0.30, \quad \mathbf{P}(S \cap T_0) = 0.04.$$

.

Example 13 (Monty's Hall Strikes Back) *Consider Monty's Hall problem. Still assume that the quiz master knows what box contains the prize and the quiz master never shows a box containing the prize. However, in this case assume that after watching the game many times you notice that when a quiz participant chooses box A the quiz master shows empty box B [resp. C] the* 60% *[resp.* 40%*] of the time. May this information improve the quiz participant's strategy for winning the prize? What about if you notice that when a quiz participant chooses box A the quiz master shows empty box B [resp. C] the* 75% *[resp.* 25%*] of the time?*

Solution. Retaining the notation of Example **??**, assume the quiz participant chooses box A and set

$$\mathbf{P}(EB) = x, \qquad \mathbf{P}(EC) = 1 - x \tag{1}$$

for the probability that the quiz master shows empty box B and C, respectively. Applying the Total Probability Formula we have

$$\mathbf{P}(EB) = \mathbf{P}(EB \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EB \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EB \mid PC)\,\mathbf{P}(PC) \tag{2}$$

and

$$\mathbf{P}(EC) = \mathbf{P}(EC \mid PA)\,\mathbf{P}(PA) + \mathbf{P}(EC \mid PB)\,\mathbf{P}(PB) + \mathbf{P}(EC \mid PC)\,\mathbf{P}(PC). \tag{3}$$

On the other hand, under the assumptions of this version of Monty's Hall problem, we still have

$$\mathbf{P}(PA) = \mathbf{P}(PB) = \mathbf{P}(PC) = \frac{1}{3}. \tag{4}$$

and

$$\begin{aligned} &\mathbf{P}(EB \mid PB) = 0, \quad \mathbf{P}(EB \mid PC) = 1, \\ &\mathbf{P}(EC \mid PB) = 1, \quad \mathbf{P}(EC \mid PC) = 0. \end{aligned} \tag{5}$$

Combining (1)-(5) it then follows

$$\mathbf{P}(EB \mid PA) = 3x - 1 \qquad \text{and} \qquad \mathbf{P}(EC \mid PA) = 2 - 3x.$$

Now, thanks to symmetry formula (**??**) we obtain

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{1}{3}\frac{3x-1}{x}, \qquad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC \mid PA)\,\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{1}{3}\frac{2-3x}{1-x}, \tag{6}$$

and

$$\mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB \mid PC)\,\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{1}{3}\frac{1}{x}, \qquad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC \mid PB)\,\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{1}{3}\frac{1}{1-x}. \tag{7}$$

In the end, note that since we are dealing with probabilities the choice of x in the interval $(0,1)$ cannot be free but subject to the further constraints

$$0 < 3x - 1 < 1 \qquad \text{and} \qquad 0 < 2 - 3x < 1.$$

These yields

$$\frac{1}{3} < x < \frac{2}{3}, \tag{8}$$

which excludes the possibility that we notice that when a quiz participant chooses box A the quiz master shows empty box B [resp. C] the 75% [resp. 25%] of the time. However, when $x = 0.6 = \frac{3}{5}$ which fulfills the constraint (8) we obtain

$$\mathbf{P}(PA \mid EB) = \frac{4}{9} = 0.4\bar{4}, \qquad \mathbf{P}(PA \mid EC) = \frac{1}{6} = 0.167$$

and

$$\mathbf{P}(PB \mid EC) = \frac{5}{6} = 0.8\bar{3}, \qquad \mathbf{P}(PC \mid EB) = \frac{5}{9} = 0.5\bar{5}.$$

These imply that the participant should again exchange chosen box A with closed box B or C. Note that comparing Equations (6) and (7) it turns out that the condition who could suggest the quiz participant to stick to her first choice would be

$$3x - 1 > 1 \qquad \text{or} \qquad 2 - 3x > 1.$$

Bothe these conditions do not respect the constraint (8). At most, an indifference condition can be realized for

$$x = \frac{1}{3} = 0.3\bar{3} \qquad \text{or} \qquad x = 0.6\bar{6}.$$

in thsi cases we have

$$\begin{aligned} &\mathbf{P}(PA \mid EB) = 0, \quad \mathbf{P}(PA \mid EC) = \tfrac{1}{2}, \\ &\mathbf{P}(PC \mid EB) = 1, \quad \mathbf{P}(PB \mid EC) = \tfrac{1}{2}, \end{aligned}$$

or

$$\begin{aligned} &\mathbf{P}(PA \mid EB) = \tfrac{1}{2}, \quad \mathbf{P}(PA \mid EC) = 0, \\ &\mathbf{P}(PC \mid EB) = \tfrac{1}{2}, \quad \mathbf{P}(PB \mid EC) = 1. \end{aligned}$$

Hence, in the first [resp. second] case the quiz participant could either stick to is first choice or exchange with the rejected box, the odds being fifty-fifty, if the quiz master shows the empty box C [resp. B].

Example 14 (The Return of Monty's Hall) *Consider Monty's Hall problem. Still assume that the quiz master knows what box contains the prize, the quiz master never shows a box containing the prize, and the quiz master chooses an empty box between two with uniform probability. However, in this case assume that after watching the game many times you notice that the prize turns out to be in box A [resp. B] for 45% [resp. 30%] of the time and in box C the rest of the time. What is the quiz participant's best strategy?*

Solution. Retaining the notation of Example **??**, in this third episode of Monty's Hall saga ve have

$$\mathbf{P}(PA) = 0.50 = \frac{1}{2}, \qquad \mathbf{P}(PB) = 0.30 = \frac{3}{10}, \qquad \mathbf{P}(PC) = 0.20 = \frac{1}{5}.$$

However, in this case, we have to determine the quiz participant's best strategy which is made by a first and a second choice. To this, assume the quiz participant's first choice is A box. We can clearly replicate the argument of Example **??** and end up with evaluating

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}} = \frac{1}{2},$$

and

$$\mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{3}{10}}{\frac{1}{2}} = \frac{3}{5}.$$

Now, assume the quiz participant first choice is B box. In this case, with a similar argument, we end up with evaluating

$$\mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA|PB)\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{\frac{1}{2}} = \frac{3}{10}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{\frac{1}{2}} = \frac{3}{10},$$

and

$$\mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA|PC)\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{1\cdot\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{1}{2}}{\frac{1}{2}} = 1.$$

In the end, assume the quiz participant first choice is C box. We end up with evaluating

$$\mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA|PC)\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{\frac{1}{2}\cdot\frac{1}{5}}{\frac{1}{2}} = \frac{1}{5}, \quad \mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PB)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{\frac{1}{2}} = \frac{3}{10},$$

and

$$\mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA|PB)\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{1\cdot\frac{3}{10}}{\frac{1}{2}} = \frac{3}{5}, \quad \mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{2}}{\frac{1}{2}} = 1.$$

In light of what shown above, the best strategy is to choose box C and change with the other rejected box as the quiz master gives the opportunity.

Example 15 (Monty's Hall Awakens) *Consider Monty's Hall problem. Still assume that the quiz master knows what box contains the prize, the quiz master never shows a box containing the prize. Assume also that after watching the game many times you notice that the prize turns out to be in box A [resp. B] for 45% [resp. 30%] of the time and in box C the rest of the time. Moreover, assume that after a quiz participant chooses a box the quiz master chooses an empty box to show between two by fliping a rigged coin with success probability p. What is the quiz participant's best strategy?*

Solution. Retaining the notation of Example **??**, the only difference between this episode of Monty's Hall saga and Episode 14 is that we have

$$\begin{aligned} &\mathbf{P}(EB) = p, \quad \mathbf{P}(EC) = 1-p, \\ &\mathbf{P}(EA) = p, \quad \mathbf{P}(EC) = 1-p, \\ &\mathbf{P}(EA) = p, \quad \mathbf{P}(EB) = 1-p, \end{aligned}$$

according to whether the quiz participant's first choice is A or B or C box In particular, we still have

$$\mathbf{P}(PA) = 0.50 = \frac{1}{2}, \qquad \mathbf{P}(PB) = 0.30 = \frac{3}{10}, \qquad \mathbf{P}(PC) = 0.20 = \frac{1}{5}.$$

As a consequence, assume the quiz participant's first choice is A box. We end up with evaluating

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{p} = \frac{1}{4p}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{1-p} = \frac{1}{4(1-p)},$$

and

$$\mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{5}}{p} = \frac{1}{5p}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{3}{10}}{1-p} = \frac{3}{10(1-p)}.$$

Assume the quiz participant's first choice is B box. We obtain

$$\mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA|PB)\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{p} = \frac{3}{20p}, \quad \mathbf{P}(PB \mid EC) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{1-p} = \frac{3}{20(1-p)},$$

and

$$\mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA|PC)\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{1\cdot\frac{1}{5}}{p} = \frac{1}{5p}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{1}{2}}{1-p} = \frac{1}{2(1-p)}.$$

In the end, assume the quiz participant's first choice is C box. We obtain

$$\mathbf{P}(PC \mid EA) = \frac{\mathbf{P}(EA|PC)\mathbf{P}(PC)}{\mathbf{P}(EA)} = \frac{\frac{1}{2}\cdot\frac{1}{5}}{p} = \frac{1}{10p}, \quad \mathbf{P}(PC \mid EB) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PB)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{3}{10}}{1-p} = \frac{3}{20(1-p)},$$

and

$$\mathbf{P}(PB \mid EA) = \frac{\mathbf{P}(EA|PB)\mathbf{P}(PB)}{\mathbf{P}(EA)} = \frac{1\cdot\frac{3}{10}}{p} = \frac{3}{10p}, \quad \mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{2}}{1-p} = \frac{1}{2(1-p)}.$$

Hence, also in this case, no matter of the value of the success probability p, the best strategy is to choose box C and change with the other rejected box as the quiz master gives the opportunity.

Example 16 (The Last Monty's Hall) *Consider Monty's Hall problem. However, in this case assume that the quiz master does not know what box contains the prize, the quiz master chooses a box between two with uniform probability, and the quiz master shows an empty box by chance. In this episode of the Monty's Hall saga, what should the participant do? To stick to her first choice, to accept the exchange or it does not matter at all because the odds are now fifty-fifty?*

Solution. Retaining the notation of Example **??**, still Equations (**??**) and (**??**) hold true, but since the quiz master does not know what box contains the prize and he shows an empty box by chance then the is a substantial change in Equation (**??**) which becomes

$$\begin{aligned} &\mathbf{P}(EB \mid PA) = \tfrac{1}{2}, \quad \mathbf{P}(EB \mid PB) = 0, \quad \mathbf{P}(EB \mid PC) = \tfrac{1}{2}, \\ &\mathbf{P}(EC \mid PA) = \tfrac{1}{2}, \quad \mathbf{P}(EC \mid PB) = \tfrac{1}{2}, \quad \mathbf{P}(EC \mid PC) = 0. \end{aligned} \tag{9}$$

Combining (**??**), (**??**) and (9), we obtain

$$\mathbf{P}(EB) = \mathbf{P}(EC) = \frac{1}{3}. \tag{10}$$

It then follows

$$\mathbf{P}(PA \mid EB) = \frac{\mathbf{P}(EB|PA)\mathbf{P}(PA)}{\mathbf{P}(EB)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{3}} = \frac{1}{2}, \quad \mathbf{P}(PA \mid EC) = \frac{\mathbf{P}(EC|PA)\mathbf{P}(PA)}{\mathbf{P}(EC)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{3}} = \frac{1}{2}, \tag{11}$$

and

$$\mathbf{P}\left(PC \mid EB\right) = \frac{\mathbf{P}(EB|PC)\mathbf{P}(PC)}{\mathbf{P}(EB)} = \frac{1\cdot\frac{1}{3}}{\frac{1}{3}} = \frac{1}{3}, \quad \mathbf{P}\left(PB \mid EC\right) = \frac{\mathbf{P}(EC|PB)\mathbf{P}(PB)}{\mathbf{P}(EC)} = \frac{1\cdot\frac{1}{3}}{\frac{1}{3}} = \frac{1}{3}. \tag{12}$$

Hence, whether the quiz master shows empty box B or C the conditional probability that the prize is in the other box C or B is the same of the conditional probability that the prize is contained in the firstly chosen box A. We can conclude that the quiz participant should be indifferent between sticking to her first choice or accepting the exchange. Note that in this episode of the Monty's Hall saga, the assumption that the quiz master ignores what box contains the prize, clearly makes the of couples of events $\{PA, EB\}$, $\{PA, EB\}$, $\{PC, EB\}$, and $\{PB, EC\}$ independent. Therefore, we could write straightforwardly

$$\mathbf{P}\left(PA \mid EB\right) = \mathbf{P}\left(PA\right) = \tfrac{1}{3}, \quad \mathbf{P}\left(PA \mid EC\right) = \mathbf{P}\left(PA\right) = \tfrac{1}{3}, \tag{13}$$

and

$$\mathbf{P}\left(PC \mid EB\right) = \mathbf{P}\left(PC\right) = \tfrac{1}{3}, \quad \mathbf{P}\left(PB \mid EC\right) = \mathbf{P}\left(PB\right) = \tfrac{1}{3}. \tag{14}$$

It is also interesting to note that the second argument applied in the solution of Example **??** for a not very evident reason. What a reasony?

Theorem 17 (Bayes Formula) *Let $N \subseteq \mathbb{N}$ and let $(F_n)_{n\in N}$ be a countable partition of Ω. Then for any $n \in N$, we have*

$$\mathbf{P}\left(F_n \mid E\right) = \frac{\mathbf{P}\left(E \mid F_n\right)\mathbf{P}\left(F_n\right)}{\sum_{m\in N}\mathbf{P}\left(E \mid F_m\right)\mathbf{P}\left(F_m\right)}, \quad \forall E \in \mathcal{E} : \mathbf{P}\left(E\right) > 0.$$

Proof. From the symmetry formula (**??**), we can write

$$\mathbf{P}\left(F_n \mid E\right) = \frac{\mathbf{P}\left(E \mid F_n\right)\mathbf{P}\left(F_n\right)}{\mathbf{P}\left(E\right)},$$

for any $n \in N$. Hence, applying Equation (**??**) to evaluate $\mathbf{P}\left(E\right)$, we obtain the desired result. □

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Distribution Functions 2021-10-28

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} \alpha + \beta x, \quad \forall x \in \mathbb{R},$$

where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*
$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$
is a random variable?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} |x|, \quad \forall x \in \mathbb{R},$$

where $|x|$ is the absolute value of x.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*
$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$
is a random variable?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 3 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $X : \Omega \to \mathbb{R}$ *be a uniformly distributed random variable with states in the interval* $[-1, 1]$*. In symbols,* $X \sim Unif(-1, 1)$*. Consider the function* $g : \mathbb{R} \to \mathbb{R}$ *given by*

$$g(x) \overset{def}{=} x^2, \quad \forall x \in \mathbb{R}.$$

1. *Can you show that the function* $Y : \Omega \to \mathbb{R}$ *given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

   is a random variable?
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}$ *of the random variable* Y*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments of* Y *are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. .

Problem 4 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $X : \Omega \to \mathbb{R}$ *be a uniformly distributed random variable with states in the interval* $[-1, 1]$*. In symbols,* $X \sim Unif(-1, 1)$*. Consider the function* $g : \mathbb{R} \to \mathbb{R}$ *given by*

$$g(x) \overset{def}{=} \begin{cases} 0, & \text{if } x \leq 0. \\ x^2, & \text{if } x > 0. \end{cases}$$

1. *Can you show that the function* $Y : \Omega \to \mathbb{R}$ *given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

   is a random variable?
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}_+$ *of the random variable* Y*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. .

Problem 5
1. *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $X : \Omega \to \mathbb{R}$ *be a uniformly distributed random variable with states in the interval* $[-1, 1]$*. In symbols,* $X \sim Unif(-1, 1)$*. Consider the function* $g : \mathbb{R} \to \mathbb{R}$ *given by*

   $$g(x) \overset{def}{=} x^2 - 2x, \quad \forall x \in \mathbb{R},$$

   Can you show that the function $Y : \Omega \to \mathbb{R}$ *given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

   is a random variable?
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}_+$ *of the random variable* Y*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. .

Problem 6 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space, and let* $X : \Omega \to \mathbb{R}$ *be an exponentially distributed random variable with rate parameter* $\lambda = 1$*. In symbols,* $X \sim Exp(1)$*. Consider the function* $g : \mathbb{R} \to \mathbb{R}$ *given by*

$$g(x) \overset{def}{=} 1 - \exp(-x), \quad \forall x \in \mathbb{R},$$

where $\exp : \mathbb{R} \to \mathbb{R}$ *is the Neper exponential function.*

1. *Can you show that the function* $Y : \Omega \to \mathbb{R}$ *given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

   is a random variable?
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}_+$ *of the random variable* Y*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. .

Problem 7 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $X : \Omega \to \mathbb{R}$ *be a uniformly distributed random variable with states in the interval* $(-1, 1)$*. In symbols,* $X \sim Unif(-1, 1)$*. Consider the function* $g : \mathbb{R}_{++} \to \mathbb{R}$ *given by*

$$g(y) \overset{def}{=} -\frac{1}{\lambda} \ln(y), \quad \forall \in \mathbb{R}_{++},$$

where $\ln : \mathbb{R}_{++} \to \mathbb{R}$ *is the natural logarithm function and* $\lambda > 0$*.*

1. *Can you state that the function* $Y : \Omega \to \mathbb{R}$ *given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$

   is a real random variable on Ω*?*
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}$ *of* $Y : \Omega \to \mathbb{R}$*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments of* Y *are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*
   Hint: recall the properties of the ***logarithm*** *and* ***exponential*** *function.*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Random Variables 2021-11-10

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra. Let $F : \mathbb{R} \to \mathbb{R}_+$ given by*

$$F(x) \stackrel{def}{=} ae^x 1_{\mathbb{R}_{--}}(x) - \left(\tfrac{1}{2}e^{-x} - b\right) 1_{\mathbb{R}_+}(x), \quad \forall x \in \mathbb{R},$$

where $a, b \in \mathbb{R}$.

1. *Determine $a, b \in \mathbb{R}$ such that $F : \mathbb{R} \to \mathbb{R}_+$ is a distribution function of a random variable $X : \Omega \to \mathbb{R}$.*
2. *Is it possible to determine $a, b \in \mathbb{R}$ such that $X : \Omega \to \mathbb{R}$ is absolutely continuous? In this case, compute $\mathbf{P}(-1 \leq X \leq 1)$.*

Solution. .

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(a, b) \stackrel{def}{=} \mathbf{E}\left[(X - (a + bY))^2\right].$$

Prove that there exists

$$(a^*, b^*) \equiv \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\{f(a, b)\}$$

and compute it.

Solution. .

Problem 3 *The decoration of a Christmas tree in a mall is made by 1000 small light bulbs. The life-time of each light bulb is exponentially distributed with an average life time of 20 days (rather cheap bulbs indeed!). The mall manager decides to turn on the lights of the Christmas tree on the midnight of the 15-th of November. Estimate the probability that at least 800 bulbs are still working on the midnight of the 25-th of December.*

Hints: write X for the random variable compute representing the life time of each bulb and compute the probability that each bulb will last until the midnight of the 25-th of December; write Y for the random variable counting the number of light bulbs out of 1.000 which are still on at the midnight of the 25-th of December and guess how it is distributed; use the Markov inequality to make the estimate.

Solution. .

Problem 4 *Let X be a geometrically distributed random variable with success probability $p = 0.1$. Compute $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$. Use the Tchebychev inequality to estimate $\mathbf{P}(0 < X < 23)$.*

Solution. .

Problem 5 *Suppose that we roll a standard fair die 100 times. Let X be the sum of the numbers that appear over the 100 rolls. Use the Tchebychev inequality to bound $\mathbf{P}(200 < X < 400)$.*

Solution. .

Random Vectors

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Random Vectors with Some Solutions 2021-11-23

Problem 1 *Let (X_1, X_2) a real random vector with a joint density $f_{X_1,X_2} : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f_{X_1,X_2}(x_1, x_2) \stackrel{def}{=} 1_{[0,1]\times[0,1]}(x_1, x_2), \quad \forall (x_1, x_2) \in \mathbb{R}^2.$$

Consider the real random variables $Y \equiv \min(X_1, X_2)$ and $Z \equiv \max(X_1, X_2)$. Determine:

1. *the distribution functions of Y and Z;*
2. *the joint distribution function of Y and Z;*
3. *the marginal distributions functions of Y and Z;*
4. *the expectations of Y and Z.*

Solution.

1. We clearly have
$$1_{[0,1]\times[0,1]}(x_1, x_2) = 1_{[0,1]}(x_1) 1_{[0,1]}(x_2),$$
for every $(x_1, x_2) \in \mathbb{R}^2$. As a consequence, considering the marginal densities of the entries X_1 and X_2 of the random vector (X_1, X_2), we obtain
$$\begin{aligned} f_{X_1}(x_1) &= \int_{\mathbb{R}} 1_{[0,1]\times[0,1]}(x_1, x_2)\, d\mu_L(x_2) = \int_{\mathbb{R}} 1_{[0,1]}(x_1) 1_{[0,1]}(x_2)\, d\mu_L(x_2) \\ &= 1_{[0,1]}(x_1) \int_{\mathbb{R}} 1_{[0,1]}(x_2)\, d\mu_L(x_2) = 1_{[0,1]}(x_1) \int_{[0,1]} d\mu_L(x_2) = 1_{[0,1]}(x_1)\, \mu_L([0,1]) \\ &= 1_{[0,1]}(x_1). \end{aligned}$$
Similarly,
$$f_{X_2}(x_2) = 1_{[0,1]}(x_2).$$
Hence, the entries X_1 and X_2 of the random vector (X_1, X_2) are independent random variables and both standard uniformly distributed. We have
$$\{Y \leq y\} = \{X_1 \leq y, X_2 \leq y\} \cup \{X_1 > y, X_2 \leq y\} \cup \{X_1 \leq y, X_2 > y\},$$
$F_{X_1}(y) F_{X_2}(y)$where the three events on the right hand side are pairwise incompatible, and
$$\{Z \leq z\} = \{X_1 \leq z, X_2 \leq z\},$$
for every $z \in \mathbb{R}$. By virtue of the indepencence of X_1 and X_2, it follows,
$$\begin{aligned} &F_Y(y) \\ &= \mathbf{P}(Y \leq y) = \mathbf{P}(X_1 \leq y, X_2 \leq y) + \mathbf{P}(X_1 > y, X_2 \leq y) + \mathbf{P}(X_1 \leq y, X_2 > y) \\ &= \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 > y)\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 > y) \\ &= \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) + (1 - \mathbf{P}(X_1 \leq y))\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y)(1 - \mathbf{P}(X_2 \leq y)) \\ &= \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_2 \leq y) - \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) + \mathbf{P}(X_1 \leq y) - \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) \\ &= \mathbf{P}(X_1 \leq y) + \mathbf{P}(X_2 \leq y) - \mathbf{P}(X_1 \leq y)\mathbf{P}(X_2 \leq y) \\ &= F_{X_1}(y) + F_{X_2}(y) - F_{X_1}(y) F_{X_2}(y) \end{aligned}$$
and
$$F_Z(z) = \mathbf{P}(X_1 \leq z, X_2 \leq z) = \mathbf{P}(X_1 \leq z)\mathbf{P}(X_2 \leq z) = F_{X_1}(z) F_{X_2}(z),$$

Note that instead of the event $\{Y \le y\}$ we could have considered the event

$$\{Y > y\} = \{X_1 > y, X_2 > y\}$$

for every $y \in \mathbb{R}$, obtaining

$$\begin{aligned}
F_Y(y) &= \mathbf{P}(Y \le y) = 1 - \mathbf{P}(Y > y) = 1 - \mathbf{P}(X_1 > y, X_2 > y) \\
&= 1 - \mathbf{P}(X_1 > y)\,\mathbf{P}(X_2 > y) = 1 - (1 - \mathbf{P}(X_1 \le y))(1 - \mathbf{P}(X_2 \le y)) \\
&= 1 - (1 - F_{X_1}(y))((1 - F_{X_2}(y))) \\
&= 1 - (1 - F_{X_2}(y) - F_{X_1}(y) + F_{X_1}(y)\,F_{X_2}(y)) \\
&= F_{X_1}(y) + F_{X_2}(y) - F_{X_1}(y)\,F_{X_2}(y)
\end{aligned}$$

as above. On the other hand, both the random variables X_1 and X_2 are standard uniformly distributed on the interval $[0,1]$. Therefore,

$$F_Y(y) = F_X(y)(2 - F_X(y)) \qquad \text{and} \qquad F_Z(z) = F_X(z)^2,$$

for all $y, x \in \mathbb{R}$, where F_X is the distribution function of the standard uniformly distributed random variable X, given by

$$F_X(x) = x \cdot 1_{[0,1]}(x) + 1_{(1,+\infty)}(x),$$

for every $x \in \mathbb{R}$. It then follows

$$\begin{aligned}
&F_Y(y) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,+\infty)}(y) - \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\right) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,0)}(y) + 2 \cdot 1_{[0,1]}(y) + 2 \cdot 1_{(1,+\infty)}(y) - \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\right) \\
&= \left(y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right)\left(2 \cdot 1_{(-\infty,0)}(y) + (2 - y) \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)\right) \\
&= (2 - y)\,y \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y)
\end{aligned}$$

and

$$F_Z(z) = \left(z \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z)\right)^2 = z^2 \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z).$$

Note that we have

$$F_Y'(y) = 2(1 - y) \cdot 1_{(0,1)}(y) \qquad \text{and} \qquad F_Z'(z) = 2z \cdot 1_{(0,1)}(z),$$

for every $y, z \in \mathbb{R} - \{0,1\}$. These imply

$$\begin{aligned}
\int_{(-\infty,y)} F_Y'(u)\,d\mu_L(u) &= \int_{(-\infty,y)} 2(1-u)\,1_{(0,1)}(u)\,d\mu_L(u) \\
&= \begin{cases} 0, & \text{if } y \le 0, \\ \int_{(0,y)} 2(1-u)\,d\mu_L(u), & \text{if } 0 < y < 1, \\ \int_{(0,1)} 2(1-u)\,d\mu_L(u), & \text{if } 1 \le y, \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
\int_{(-\infty,z)} F_Z'(v)\,d\mu_L(v) &= \int_{(-\infty,z)} 2z \cdot 1_{(0,1)}(z)\,d\mu_L(v) \\
&= \begin{cases} 0, & \text{if } z \le 0, \\ \int_{(0,z)} 2v\,d\mu_L(v), & \text{if } 0 < z < 1, \\ \int_{(0,1)} 2v\,d\mu_L(v), & \text{if } 1 \le z. \end{cases}
\end{aligned}$$

On the other hand,

$$\int_{(0,y)} 2(1-u)\,d\mu_L(u) = \int_0^y 2(1-u)\,du = 2u - u^2\big|_0^y = y(2-y),$$

for every $0 < y \le 1$, and

$$\int_{(0,z)} 2v\,d\mu_L(v) = \int_0^z 2v\,dv = v^2\big|_0^z = z^2,$$

for every $0 < z \le 1$. We can then write

$$\int_{(-\infty,y)} F_Y'(u)\,d\mu_L(u) = y(2-y) \cdot 1_{[0,1]}(y) + 1_{(1,+\infty)}(y) = F_Y(y),$$

for every $y \in \mathbb{R}$, and

$$\int_{(-\infty,z)} F_Z'(v)\,d\mu_L(v) = z^2 \cdot 1_{[0,1]}(z) + 1_{(1,+\infty)}(z) = F_Z(z),$$

for every $z \in \mathbb{R}$. These imply that Y and Z are absolutely continuous random variables.

2. We have

$$\begin{aligned}
&\{Y \le y, Z \le z\} \\
&= (\{X_1 \le y, X_2 \le y\} \cup \{X_1 > y, X_2 \le y\} \cup \{X_1 \le y, X_2 > y\}) \cap \{X_1 \le z, X_2 \le z\} \\
&= (\{X_1 \le y, X_2 \le y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&\quad \cup (\{X_1 > y, X_2 \le y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&\quad \cup (\{X_1 \le y, X_2 > y\} \cap \{X_1 \le z, X_2 \le z\}) \\
&= \{X_1 \le \min(y,z), X_2 \le \min(y,z)\} \\
&\quad \cup \{y < X_1 \le z, X_2 \le \min(y,z)\} \\
&\quad \cup \{X_1 \le \min(y,z), y < X_2 \le z\}.
\end{aligned}$$

Therefore, considering the joint distribution function $F_{Y,Z} : \mathbb{R}^2 \to \mathbb{R}_+$ of Y and Z, on account of the independence of X_1 and X_2, we can write

$$\begin{aligned}
F_{Y,Z}(y,z) &= \mathbf{P}(Y \le y, Z \le z) \\
&= \mathbf{P}(X_1 \le \min(y,z), X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(y < X_1 \le z, X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(X_1 \le \min(y,z), y < X_2 \le z) \\
&= \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(y < X_1 \le z)\,\mathbf{P}(X_2 \le \min(y,z)) \\
&\quad + \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(y < X_2 \le z),
\end{aligned}$$

for every $(y,z) \in \mathbb{R}^2$. On the other hand,

$$\begin{array}{ll} \min(y,z) = y, & \text{if } y \le z, \\ \mathbf{P}(y < X_1 \le z) = 0 \quad \text{and} \quad \min(y,z) = z, & \text{if } y > z. \end{array}$$

Hence, considerning that X_1 and X_2 have the same distribution, we obtain

$$F_{Y,Z}(y,z) = \begin{cases} F_X(y)(2F_X(z) - F_X(y)), & \text{if } y \le z, \\ F_X(z)^2, & \text{if } y > z. \end{cases}$$

In fact, if $y \le z$

$$\begin{aligned}
&\mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(X_2 \le \min(y,z)) + \mathbf{P}(y < X_1 \le z)\,\mathbf{P}(X_2 \le \min(y,z)) \\
&+ \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(y < X_2 \le z) \\
&= \mathbf{P}(X \le y)\,\mathbf{P}(X \le y) + 2\mathbf{P}(X \le y)\,\mathbf{P}(y < X \le z) \\
&= F_X(y)^2 + 2F_X(y)(F_X(z) - F_X(y)) \\
&= F_X(y)(2F_X(z) - F_X(y)).
\end{aligned}$$

and if $y > z$

$$\begin{aligned}
&\mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(X_2 \le \min(y,z)) + \mathbf{P}(y < X_1 \le z)\,\mathbf{P}(X_2 \le \min(y,z)) \\
&+ \mathbf{P}(X_1 \le \min(y,z))\,\mathbf{P}(y < X_2 \le z) \\
&= \mathbf{P}(X \le z)\,\mathbf{P}(X \le z) + 2\mathbf{P}(X \le z)\,\mathbf{P}(y < X \le z) \\
&= F_X(z)^2
\end{aligned}$$

Note that we can write

$$F_{Y,Z}(y,z) = F_X(y)\left(2F_X(z) - F_X(y)\right) 1_{\{(y,z)\in\mathbb{R}^2 : y\le z\}} + F_X(z)^2\, 1_{\{(y,z)\in\mathbb{R}^2 : y> z\}}.$$

3. To determine the marginal distribution functions $F_Y : \mathbb{R} \to \mathbb{R}_+$ and $F_Z : \mathbb{R} \to \mathbb{R}_+$ of the random vector (Y, Z), respectively, we can apply the formula

$$\begin{aligned} F_Y(y) &= \lim_{z\to+\infty} F_{Y,Z}(y,z) \\ &= \lim_{z\to+\infty} \Big(F_X(y)\left(2F_X(z) - F_X(y)\right) 1_{\{(y,z)\in\mathbb{R}^2 : y\le z\}}(y,z) + F_X(z)^2\, 1_{\{(y,z)\in\mathbb{R}^2 : y> z\}}(y,z) \Big) \end{aligned}$$

and

$$\begin{aligned} F_Z(z) &= \lim_{y\to+\infty} F_{Y,Z}(y,z) = \\ &= \lim_{y\to+\infty} \Big(F_X(y)\left(2F_X(z) - F_X(y)\right) 1_{\{(y,z)\in\mathbb{R}^2 : y\le z\}}(y,z) + F_X(z)^2\, 1_{\{(y,z)\in\mathbb{R}^2 : y> z\}}(y,z) \Big). \end{aligned}$$

as $z \to +\infty$ for every $y \in \mathbb{R}$ we have

$$1_{\{(y,z)\in\mathbb{R}^2 : y\le z\}}(y,z) = 1 \qquad \text{and} \qquad 1_{\{(y,z)\in\mathbb{R}^2 : y> z\}}(y,z) = 0.$$

Conversely, as $y \to +\infty$ for every $z \in \mathbb{R}$ we have

$$1_{\{(y,z)\in\mathbb{R}^2 : y\le z\}}(y,z) = 0 \qquad \text{and} \qquad 1_{\{(y,z)\in\mathbb{R}^2 : y> z\}}(y,z) = 1.$$

It then follows

$$F_Y(y) = F_X(y)\left(2F_X(z) - F_X(y)\right) \qquad \text{and} \qquad F_Z(z) = F_X(z)^2,$$

which shows that the marginal distribution functions of the random vector (Y, Z) concide with the distribution functions of the random variables X and Y. As a consequence, the random variables $Y \equiv \min(X_1, X_2)$ and $Z \equiv \max(X_1, X_2)$ are indepemdent.

4. In the end, we have

$$\begin{aligned} \mathbf{E}[Y] &= \int_{\mathbb{R}} y f_Y(y)\, d\mu_L(y) = \int_{\mathbb{R}} 2y(1-y)\, 1_{[0,1]}(y)\, d\mu_L(y) = \int_{[0,1]} 2y(1-y)\, d\mu_L(y) \\ &= \int_0^1 2(1-y)\, y\, dy = 2\left(\int_0^1 y\,dy - \int_0^1 y^2 dy \right) = 2\left(\frac{1}{2} y^2 \Big|_0^1 - \frac{1}{3} y^3 \Big|_0^1 \right) = \frac{1}{3} \end{aligned}$$

and

$$\begin{aligned} \mathbf{E}[Z] &= \int_{\mathbb{R}} z f_Z(z)\, d\mu_L(z) = \int_{\mathbb{R}} 2z^2 \cdot 1_{[0,1]}(z)\, d\mu_L(z) = \int_{[0,1]} 2z^2 d\mu_L(z) \\ &= \int_0^1 2z^2 dz = 2\int_0^1 z^2 dz = 2\, \frac{1}{3} z^3 \Big|_0^1 = \frac{2}{3}. \end{aligned}$$

Problem 2 *Let $F : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$F(x_1, x_2) \overset{def}{=} \left(1 - e^{-x_1} - e^{-x_2} + e^{-(x_1+x_2)}\right) 1_{\mathbb{R}_+}(x_1)\, 1_{\mathbb{R}_+}(x_2), \quad \forall (x_1, x_2) \in \mathbb{R}^2.$$

Show that $F : \mathbb{R}^2 \to \mathbb{R}_+$ is the distribution function of a real random vector (X_1, X_2) and compute the marginal distribution functions of (X_1, X_2).

1. *Is the function $F : \mathbb{R}^2 \to \mathbb{R}_+$ absolutely continuous?*
2. *Are the entries X_1 and X_2 of the random vector (X_1, X_2) independent random variables?*
3. *Are the entries X_1 and X_2 of the random vector (X_1, X_2) absolutely continuous random variables?*
4. *What is the distribution $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ of the real random variable $Z = \max\{X_1, X_2\}$.*
5. *Is the function $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ absolutely continuous?*

Hint: it might be useful to rewrite $F(x_1, x_2)$ in a more convenient form.

Solution. .

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $\left(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_L^2\right) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^2)$ and the Lebesgue measure $\mu_L^2 : \mathcal{B}(\mathbb{R}^2) \to \mathbb{R}_+$. Let $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x,y) \overset{def}{=} k e^{-(x^2 - xy + y^2/2)}, \quad \forall (x,y) \in \mathbb{R}^2,$$

where $k \in \mathbb{R}$ is a parameter.

1. *Determine k such that $f : \mathbb{R}^2 \to \mathbb{R}$ is a probability density. Hint: can you compute $\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(y-x)^2} dy$ with no computation?*

   Let $Z \equiv (X, Y)$ be the random vector on Ω with density $f : \mathbb{R}^2 \to \mathbb{R}_+$.
2. *Determine the marginal density of the entries X and Y. Are the random variables X and Y Gaussian?*
3. *Is the random vector Z Gaussian?*
4. *Compute $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$, and $Cov(X, Y)$.*
5. *Are X and Y independent random variables?*
6. *Is the random vector Z Gaussian? Hint: consider the answer you gave to 4., what you know from the theory, and try to make a simple guess.*

Solution.

1. We can write

$$\int_{\mathbb{R}^2} f(x,y)\, d\mu_L^2(x,y) = k \int_{\mathbb{R}^2} e^{-(x^2 - xy + y^2/2)} d\mu_L^2(x,y).$$

On the other hand, since $e^{-(x^2 - xy + y^2/2)}$ is a continuous positive function

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-(x^2 - xy + y^2/2)} d\mu_L^2(x,y) &= \int_{y=-\infty}^{+\infty} \int_{x=-\infty}^{+\infty} e^{-(x^2 - xy + y^2/2)} dx dy \\ &= \int_{y=-\infty}^{+\infty} \int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}\left(y^2 - 2xy + x^2\right)} e^{-\frac{1}{2}x^2} dx dy \\ &= \int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}x^2} \left(\int_{y=-\infty}^{+\infty} e^{-\frac{1}{2}(y-x)^2} dy \right) dx. \end{aligned}$$

Now, we have

$$\int_{y=-\infty}^{+\infty} e^{-\frac{1}{2}(y-x)^2} dy = \int_{z=-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi},$$

for every $x \in \mathbb{R}$. Therefore,

$$\int_{y=-\infty}^{+\infty} \int_{x=-\infty}^{+\infty} e^{-(x^2 - xy + y^2/2)} dx dy = \sqrt{2\pi} \int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx = 2\pi.$$

If follows that

$$\int_{\mathbb{R}^2} f(x,y)\, d\mu_L^2(x,y) = 1 \Rightarrow k = \frac{1}{2\pi}.$$

2. Considering what shown above, we have

$$f_X(x) = \int_{\mathbb{R}} \frac{1}{2\pi} f(x,y)\, d\mu_L(y) = \frac{1}{2\pi}\int_{y=-\infty}^{+\infty} e^{-\frac{1}{2}\left(y^2-2xy+x^2\right)} e^{-\frac{1}{2}x^2}\, dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

for every $x \in \mathbb{R}$. Similarly, since

$$e^{-(x^2-xy+y^2/2)} = e^{-\frac{1}{2}(2x^2-2xy+y^2)} = e^{-\frac{1}{2}\left(\left(\sqrt{2}x\right)^2-2xy+\left(\frac{y}{\sqrt{2}}\right)^2\right)} e^{-\frac{1}{2}\left(\frac{y}{\sqrt{2}}\right)^2} = e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2} e^{-\frac{y^2}{4}},$$

we have

$$f_Y(y) = \int_{\mathbb{R}} \frac{1}{2\pi} f(x,y)\, d\mu_L(x) = \frac{1}{2\pi}\int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2} e^{-\frac{y^2}{4}}\, dx = \frac{1}{2\pi} e^{-\frac{y^2}{4}} \int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2} dx,$$

for every $y \in \mathbb{R}$. Furthermore,

$$\int_{x=-\infty}^{+\infty} e^{-\frac{1}{2}\left(\sqrt{2}x-\frac{y}{\sqrt{2}}\right)^2} dx = \frac{1}{\sqrt{2}}\int_{z=-\infty}^{+\infty} e^{-\frac{1}{2}z^2}\, dz = \sqrt{\pi}.$$

Hence,

$$f_Y(y) = \frac{1}{2\sqrt{\pi}} e^{-\frac{y^2}{4}} = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{1}{2}\left(\frac{y}{\sigma_Y}\right)^2}, \quad \sigma_Y \equiv \sqrt{2}.$$

This shows that the random variables X and Y are Gaussian.

3. We clearly have

$$\mathbf{E}[X] = \mathbf{E}[Y] = 0.$$

Moreover,

$$\mathbf{D}^2[X] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{1}{2}x^2}\, dx = 1, \quad \mathbf{D}^2[Y] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} y^2 e^{-\frac{1}{2}\left(\frac{y}{\sqrt{2}}\right)^2} dy = 2.$$

In addition,

$$\begin{aligned} Cov(X,Y) = \mathbf{E}[XY] &= \int_{\mathbb{R}^2} xy f(x,y)\, d\mu_L^2(x,y) = \frac{1}{2\pi}\int_{\mathbb{R}^2} xy e^{-(x^2-xy+y^2/2)} d\mu_L^2(x,y) \\ &= \frac{1}{2\pi}\int_{x=-\infty}^{+\infty} x e^{-\frac{1}{2}x^2}\left(\int_{y=-\infty}^{+\infty} y e^{-\frac{1}{2}(y-x)^2} dy\right) dx. \end{aligned}$$

On the other hand,

$$\begin{aligned} \int_{y=-\infty}^{+\infty} y e^{-\frac{1}{2}(y-x)^2} dy &= \int_{y=-\infty}^{+\infty} (y-x)\, e^{-\frac{1}{2}(y-x)^2} dy + \int_{y=-\infty}^{+\infty} x e^{-\frac{1}{2}(y-x)^2} dy \\ &= \int_{z=-\infty}^{+\infty} z e^{-\frac{1}{2}z^2} dz + x\int_{z=-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz \\ &= \sqrt{2\pi}x. \end{aligned}$$

Hence,

$$Cov(X,Y) = \frac{1}{2\pi}\int_{x=-\infty}^{+\infty} \sqrt{2\pi}x^2 e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{+\infty} x^2 e^{-\frac{1}{2}x^2} = 1.$$

4. Since

$$Cov(X,Y) \neq 0,$$

the random variables X and Y are not independent.

5. Since not independent, despite X and Y are Gaussian, we cannot state at present whether the random vector $(X,Y)^{\intercal}$ is Gaussian or not. To solve this doubt, we can try to write

$$\begin{pmatrix} X \\ Z \end{pmatrix} = A \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

for independent standard Gaussian random variables Z_1 and Z_1 and a suitable matrix

$$A \equiv \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.$$

If this is true, we have

$$\Sigma^2_{X,Y} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = AA^{\intercal}.$$

Thus, we are led to find a matrix A such that

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}\begin{pmatrix} a_{1,1} & a_{2,1} \\ a_{1,2} & a_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1}^2 + a_{1,2}^2 & a_{1,1}a_{2,1} + a_{1,2}a_{2,2} \\ a_{1,1}a_{2,1} + a_{1,2}a_{2,2} & a_{2,1}^2 + a_{2,2}^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

To this goal, observe that $\Sigma^2_{X,Y}$ has eigenvalues

$$\frac{3}{2} + \frac{1}{2}\sqrt{5} \quad \text{and} \quad \frac{3}{2} - \frac{1}{2}\sqrt{5},$$

with corresponding orthogonal eigenvectors

$$\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix}.$$

In fact, we have

$$\left(\frac{3}{2} + \frac{1}{2}\sqrt{5}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\left(\frac{3}{2} - \frac{1}{2}\sqrt{5}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$\begin{pmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{pmatrix} = 0.$$

Therefore, normalizing the eigenvectors, we have that

$$B \equiv \left\{ \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \end{pmatrix}, \begin{pmatrix} -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix} \right\}$$

is a basis of orthonormal eigenvectors in $\mathbb{R}^2$. We then have

$$M_E^B(id)\, \Lambda M_B^E(id) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

where

$$E \equiv \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

is the standard orthonormal basia in $\mathbb{R}^2$,

$$M_E^B(id) = \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix}, \quad \Lambda \equiv \begin{pmatrix} \frac{3}{2} + \frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2} - \frac{1}{2}\sqrt{5} \end{pmatrix},$$

and

$$M_B^E(id) = M_E^B(id)^{-1} = M_E^B(id)^{\intercal} = \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix} \begin{pmatrix} \frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In addition, we can write

$$\begin{aligned} &\begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix} \begin{pmatrix} \frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix} \begin{pmatrix} \sqrt{\frac{3}{2}+\frac{1}{2}\sqrt{5}} & 0 \\ 0 & \sqrt{\frac{3}{2}-\frac{1}{2}\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{\frac{3}{2}+\frac{1}{2}\sqrt{5}} & 0 \\ 0 & \sqrt{\frac{3}{2}-\frac{1}{2}\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \\ \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}+\frac{1}{2}\sqrt{5}}} \end{pmatrix} \begin{pmatrix} \frac{1}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{1}{2}\sqrt{5}-\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & \frac{1}{2}\sqrt{5}-\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\frac{1}{2}\sqrt{5}-\frac{1}{2}}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{\frac{1}{2}\sqrt{5}+\frac{1}{2}}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix}. \end{aligned}$$

Therefore, we obtain

$$\begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} \\ -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Setting

$$A = \begin{pmatrix} \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & -\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \\ \frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} & \frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} \end{pmatrix},$$

it then follows

$$a_{1,1}^2 + a_{1,2}^2 = \left(\frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\right)^2 + \left(-\frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\right)^2 = 1,$$

$$a_{2,1}^2 + a_{2,2}^2 = \left(\frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\right)^2 + \left(\frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\right)^2 = 2,$$

$$a_{1,1}a_{2,1} + a_{1,2}a_{2,2} = \frac{1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}}\frac{1}{2}\frac{\sqrt{5}+1}{\sqrt{\frac{5}{2}-\frac{1}{2}\sqrt{5}}} - \frac{1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}}\frac{1}{2}\frac{\sqrt{5}-1}{\sqrt{\frac{1}{2}\sqrt{5}+\frac{5}{2}}} = 1.$$

This proves that $(X,Y)^\intercal$ is Gaussian. Note that, from

$$\begin{pmatrix} X \\ Y \end{pmatrix} = A \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

it follows

$$\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X & Y \end{pmatrix} = A \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} A^\intercal,$$

that is to say

$$\begin{pmatrix} X^2 & XY \\ XY & Y^2 \end{pmatrix} = A \begin{pmatrix} Z_1^2 & Z_1Z_2 \\ Z_1Z_2 & Z_2^2 \end{pmatrix} A^\intercal.$$

It follows,

$$\begin{aligned} \Sigma_{X,Y}^2 &= \begin{pmatrix} \mathbf{D}^2[X] & Cov(X,Y) \\ Cov(X,Y) & \mathbf{D}^2[X] \end{pmatrix} = \begin{pmatrix} \mathbf{E}[X^2] & \mathbf{E}[XY] \\ \mathbf{E}[XY] & \mathbf{E}[Y^2] \end{pmatrix} \\ &= A \begin{pmatrix} \mathbf{E}[Z_1^2] & \mathbf{E}[Z_1Z_2] \\ \mathbf{E}[Z_1Z_2] & \mathbf{E}[Z_2^2] \end{pmatrix} A^\intercal = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^\intercal = AA^\intercal. \end{aligned}$$

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_L^2) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^2)$ and the Lebesgue measure $\mu_L^2 : \mathcal{B}(\mathbb{R}^2) \to \mathbb{R}_+$. Let $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x,y) \overset{def}{=} ke^{-\frac{x^2-xy+y^2}{2}}, \quad \forall (x,y) \in \mathbb{R}^2,$$

where $k \in \mathbb{R}$ is a parameter.

1. *Determine k such that $f : \mathbb{R}^2 \to \mathbb{R}$ is a probability density. Hint: It may be useful to recall that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx = 1$.*

2. *Determine the marginal density functions of the entries X and Y. Are X and Y independent?*

3. *Compute $\mathbf{P}(X = Y)$ and $\mathbf{P}(X \geq Y)$.*

Solution. .

Exercise 5 (Sheldon M. Ross - 4.11) *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be real random variables on Ω such that the random vector (X,Y) is absolutely continuous with a density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f_{X,Y}(x,y) \overset{def}{=} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \cdot 1_{(0,1)\times(0,2)}(x,y), \quad \forall (x,y) \in \mathbb{R}^2.$$

1. *Check that $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ is a density function.*

2. *Are the random variables X and Y absolutely continuous? In case of affirmative answer determine the marginal densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ of X and Y, respectively.*

3. *Check whether the random variables X and Y are independent.*

4. *Compute $\mathbf{P}(X > Y)$.*

Solution.

1. We will have proven that $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$ is a density function if we can show that

$$\int_{\mathbb{R}^2} f_{X,Y}(x,y)\, d\mu_L^2(x,y) = 1.$$

On the other hand, considering the properties of the Lebesgue integral, we have

$$\begin{aligned}
\int_{\mathbb{R}^2} f_{X,Y}(x,y)\,d\mu_L(x,y) &= \int_{\mathbb{R}^2} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\ \ d\mu_L^2(x,y)\\
&= \int_{(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right) d\mu_L^2(x,y)\\
&= \int_{(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right) dxdy\\
&= \int_{y=0}^{2}\int_{x=0}^{1} \frac{6}{7}\left(x^2+\frac{xy}{2}\right) dxdy\\
&= \frac{6}{7}\int_{y=0}^{2}\left(\int_{x=0}^{1}\left(x^2+\frac{xy}{2}\right)dx\right)dy\\
&= \frac{6}{7}\int_{y=0}^{2}\left(\frac{x^3}{3}+\frac{x^2y}{4}\bigg|_0^1\right)dy\\
&= \frac{6}{7}\int_{y=0}^{2}\left(\frac{1}{3}+\frac{y}{4}\right)dy\\
&= \frac{6}{7}\left(\frac{y}{3}+\frac{y^2}{8}\bigg|_0^2\right)\\
&= \frac{6}{7}\left(\frac{2}{3}+\frac{1}{2}\right)\\
&= 1.
\end{aligned}$$

2. Since the random vector is absolutely continuous the entries X and Y are absolutely continuous random variables with densities $f_X:\mathbb{R}\to\mathbb{R}_+$ and $f_Y:\mathbb{R}\to\mathbb{R}_+$ given by

$$f_X(x)=\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(y) \qquad \text{and} \qquad f_Y(y)=\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(x),$$

μ_L-a.e. on $\mathbb{R}$, respectively. Now, we have

$$\begin{aligned}
\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(y) &= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L(y)\\
&= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y)\,d\mu_L(y)\\
&= \int_{(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,d\mu_L(y)\\
&= \frac{6}{7}\left(\int_0^2\left(x^2+\frac{xy}{2}\right)dy\right)\cdot 1_{(0,1)}(x)\\
&= \frac{6}{7}\left(x^2y+\frac{xy^2}{4}\bigg|_{y=0}^2\right)\cdot 1_{(0,1)}(x)\\
&= \frac{6}{7}\left(2x^2+x\right)\cdot 1_{(0,1)}(x).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_{\mathbb{R}} f_{X,Y}(x,y)\,d\mu_L(x) &= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L(x)\\
&= \int_{\mathbb{R}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y)\,d\mu_L(x)\\
&= \int_{(0,1)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,2)}(y)\,d\mu_L(y)\\
&= \frac{6}{7}\left(\int_0^1\left(x^2+\frac{xy}{2}\right)dx\right)\cdot 1_{(0,2)}(y)\\
&= \frac{6}{7}\left(\frac{x^3}{3}+\frac{x^2y}{4}\bigg|_{x=0}^1\right)\cdot 1_{(0,2)}(y)\\
&= \frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y).
\end{aligned}$$

Therefore, we can write

$$f_X(x)=\frac{6}{7}\left(x+2x^2\right)\cdot 1_{(0,1)}(x) \qquad \text{and} \qquad f_Y(y)=\frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y),$$

μ_L-a.e. on $\mathbb{R}$, respectively.

3. The random variables X and Y are independent if and only if

$$f_X(x)\,f_Y(y)=f_{X,Y}(x,y),$$

μ_L^2-a.e. on $\mathbb{R}^2$. On the other hand,

$$\begin{aligned}
f_X(x)\,f_Y(x) &= \left(\frac{6}{7}\left(x+2x^2\right)\cdot 1_{(0,1)}(x)\right)\left(\frac{6}{7}\left(\frac{1}{3}+\frac{y}{4}\right)\cdot 1_{(0,2)}(y)\right)\\
&= \frac{36}{49}\left(\frac{x}{3}+\frac{xy}{4}+\frac{2x^2}{3}+\frac{x^2y}{2}\right)\cdot 1_{(0,1)}(x)\,1_{(0,2)}(y)\\
&= \frac{36}{49}\left(\frac{x}{3}+\frac{xy}{4}+\frac{2x^2}{3}+\frac{x^2y}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\\
&\neq \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)
\end{aligned}$$

for almost all points $(x,y)\in(0,1)\times(0,2)$. Therefore, X and Y are not independent.

4. To compute $\mathbf{P}(X>Y)$ we apply the formula

$$\mathbf{P}((X,Y)\in B)=\int_B f_{X,Y}(x,y)\ \ d\mu_L^2(x,y),$$

which holds true for every $B\in\mathcal{B}(\mathbb{R}^2)$, by suitably choosing B to represent the event $\{X>Y\}$ in terms of the event $\{(X,Y)\in B\}$. Eventually, setting

$$B\equiv\left\{(x,y)\in\mathbb{R}^2 : x>y\right\},$$

it turns out that we can write

$$\{X>Y\}=\{(X,Y)\in B\}.$$

In fact, assume that $\omega\in\{X>Y\}\equiv\{\omega\in\Omega : X(\omega)>Y(\omega)\}$, then we have $X(\omega)>Y(\omega)$ so that $(X(\omega),Y(\omega))\in B$ and $\omega\in\{(X,Y)\in B\}\equiv\{\omega\in\Omega:(X(\omega),Y(\omega))\in B\}$. Conversely, assume that $\omega\in\{(X,Y)\in B\}$, then $(X(\omega),Y(\omega))\in B$, which implies $X(\omega)>Y(\omega)$ and consequently $\omega\in\{X>Y\}$.

As a consequence, we have

$$
\begin{aligned}
\mathbf{P}(X>Y) &= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}} f_{X,Y}(x,y)\,d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)\cdot 1_{(0,1)\times(0,2)}(x,y)\,d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}\cap(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)d\mu_L^2(x,y) \\
&= \int_{\{(x,y)\in\mathbb{R}^2:x>y\}\cap(0,1)\times(0,2)} \frac{6}{7}\left(x^2+\frac{xy}{2}\right)dxdy \\
&= \frac{6}{7}\int_{x=0}^{1}\left(\int_{y=0}^{x}\left(x^2+\frac{xy}{2}\right)dy\right)dx \\
&= \frac{6}{7}\int_{x=0}^{1}\left(x^2y+\frac{xy^2}{4}\Big|_0^x\right)dx \\
&= \frac{6}{7}\int_{x=0}^{1}\frac{5x^3}{4}dx \\
&= \frac{6}{7}\frac{5x^4}{16}\Big|_0^1 \\
&= \frac{6}{7}\frac{5}{16} \\
&= \frac{15}{56}\approx 0.26786
\end{aligned}
$$

.

Problem 6 *Let $f:\mathbb{R}^2\to\mathbb{R}_+$ given by*

$$f(x,y)\overset{def}{=}\frac{4x+2y}{3}1_{[0,1]}(x)\,1_{[0,1]}(y),\quad \forall(x,y)\in\mathbb{R}^2.$$

1. *Show that $f:\mathbb{R}^2\to\mathbb{R}_+$ is the density function of a real random vector (X,Y).*
2. *Compute the marginal densities of (X,Y) and check that the computed marginal densities are actually probability densities.*
3. *May we say that the entries X and Y of the random vector (X,Y) are independent random variables?*
4. *Compute the conditional density function $f_{X|Y}(x,y)$ of X given that $Y=y$ and check the computed density is actually a probability density.*
5. *Compute the function $\mathbf{E}[X\mid Y=y]$ and the conditional expectation $\mathbf{E}[X\mid Y]$.*

Solution. .

Problem 7 *Determine the value of the parameter k such that the function $f:\mathbb{R}^3\to\mathbb{R}$ given by*

$$f(x_1,x_2,x_3)\overset{def}{=}\begin{cases} k\left(x_1+x_2^2+x_3^3\right) & \text{if } (x_1,x_2,x_3)\in[0,1]\times[0,1]\times[0,1] \\ 0 & \text{otherwise}\end{cases}$$

is a probability density. Hence, consider the random vector $(X_1,X_2,X_3)^\intercal$ with density $f_{X_1,X_2,X_3}:\mathbb{R}^3\to\mathbb{R}$ given by

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3)\overset{def}{=}f(x_1,x_2,x_3).$$

Compute:

1. *the probability $\mathbf{P}(X_2\le 1/2, X_3>1/2)$;*
2. *the marginal density of the random vector $(X_1,X_2)^\intercal$;*
3. *the expectation of $(X_1,X_2)^\intercal$;*
4. *the conditional density $f_{X_1,X_2|X_3=1/2}(x_1,x_2)$.*

Solution. .

Problem 8 *Determine the value of the parameter k such that the function $f:\mathbb{R}^3\to\mathbb{R}$ given by*

$$f(x_1,x_2,x_3)\overset{def}{=}\begin{cases} k\left(x_1^2+x_2^2+x_3^3\right) & \text{if } (x_1,x_2,x_3)\in[0,1]\times[0,1]\times[0,1] \\ 0 & \text{otherwise}\end{cases}$$

is a probability density. Hence, consider the random vector $(X_1,X_2,X_3)^\intercal$ with density $f_{X_1,X_2,X_3}:\mathbb{R}^3\to\mathbb{R}$ given by

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3)\overset{def}{=}f(x_1,x_2,x_3).$$

Compute:

1. *the marginal density of the random vector $(X_1,X_2)^\intercal$;*
2. *the expectation of the product $X_1\cdot X_2$;*
3. *the conditional density $f_{X_1|X_2=1/2,X_3=3/4}(x_1)$;*
4. *the probability $\mathbf{P}(X_1\le 1/2, X_2<1/2, X_3<1/2)$.*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Distribution Functions 2021-11-17 with Some Solutions

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1, 1]$. In symbols, $X \sim Unif(-1, 1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \stackrel{def}{=} \alpha + \beta x, \quad \forall x \in \mathbb{R},$$

where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution.

1. The function $g : \mathbb{R} \to \mathbb{R}$ is a Borel function. Therefore, $Y = g \circ X$ is a random variable.
2. Recall that $X \sim Unif(-1, 1)$ is absolutely continuous with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x),$$

for every $x \in \mathbb{R}$. Hence, writing $F_X : \mathbb{R} \to \mathbb{R}$ for the distribution function of X, we have

$$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\, d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{2} 1_{[-1,1]}(u)\, d\mu_L(u) \\ &= \frac{1}{2} \int_{(-\infty,x] \cap [-1,1]} d\mu_L(u) = \frac{1}{2} \mu_L((-\infty, x] \cap [-1, 1]). \end{aligned}$$

On the other hand,

$$(-\infty, x] \cap [-1, 1] = \begin{cases} \varnothing, & \text{if } x < -1, \\ \{-1\}, & \text{if } x = -1, \\ [-1, x], & \text{if } x < -1. \end{cases}$$

Therefore,

$$F_X(x) = \begin{cases} 0, & \text{if } x < -1, \\ \frac{x+1}{2}, & \text{if } -1 \le x < 1, \\ 1, & \text{if } 1 \le x. \end{cases}$$

Now, since g is a continuously differentiable real function on $\mathbb{R}$, in particular a Borel function, then $Y \equiv g(X) = \alpha + \beta X$ is a real random variable. To compute the distribution function F_Y, we apply the definition

$$F_Y(y) \stackrel{\text{def}}{=} \mathbf{P}(Y \le y), \quad \forall y \in \mathbb{R}.$$

On the other hand, considering that $\beta \neq 0$, we have

$$\begin{aligned} \mathbf{P}(Y \le y) &= \mathbf{P}(\alpha + \beta X \le y) = \mathbf{P}\left(X \le \frac{y-\alpha}{\beta}\right) \\ &= F_X\left(\frac{y-\alpha}{\beta}\right) = \begin{cases} 0, & \text{if } \frac{y-\alpha}{\beta} < -1 \Leftrightarrow y < \alpha - \beta, \\ \frac{\frac{y-\alpha}{\beta}+1}{2} = \frac{y+\beta-\alpha}{2\beta}, & \text{if } -1 \le \frac{y-\alpha}{\beta} < 1 \Leftrightarrow \alpha - \beta \le y < \alpha + \beta, \\ 1, & \text{if } 1 \le \frac{y-\alpha}{\beta} \Leftrightarrow \alpha + \beta \le y. \end{cases} \end{aligned}$$

Summarizing,

$$F_Y(y) = \begin{cases} 0, & \text{if } y < \alpha - \beta, \\ \frac{y+\beta-\alpha}{2\beta}, & \text{if } \alpha - \beta \le y \le \alpha + \beta, \\ 1, & \text{if } \alpha + \beta < y. \end{cases}$$

Therefore, the random variable Y turns out to be a uniformly distributed random variable on the interval $[\alpha - \beta, \alpha + \beta]$. In symbols, $Y \sim Unif(\alpha - \beta, \alpha + \beta)$. It then follows that Y is absolutely continuous with density $f_Y : \mathbb{R} \to \mathbb{R}$ given by

$$f_Y(y) = \frac{1}{2\beta} 1_{[\alpha-\beta,\alpha+\beta]}(y).$$

3. Since X is in the linear space $\mathcal{L}^2(\Omega; \mathbb{R})$, the random variable $Y = \alpha + \beta X$ is also in the linear space $\mathcal{L}^2(\Omega; \mathbb{R})$. Hence, Y has finite moments of order 1 and 2.
4. Thanks to the linearity of the expectation operator, we have

$$\mathbf{E}[Y] = \mathbf{E}[\alpha + \beta X] = \alpha + \beta \mathbf{E}[X],$$

where

$$\mathbf{E}[X] = \int_{\mathbb{R}} \frac{1}{2} x 1_{[-1,1]}(x)\, d\mu_L(x) = \frac{1}{2} \int_{[-1,1]} x\, d\mu_L(x) = \frac{1}{2} \int_{-1}^{1} x\, dx = \frac{1}{4} x^2 \Big|_{-1}^{1} = 0.$$

Therefore,

$$\mathbf{E}[Y] = \alpha.$$

Moreover considering the properties of the variance operator, we have

$$\mathbf{D}^2[Y] = \mathbf{D}^2[\alpha + \beta X] = \beta^2 \mathbf{D}^2[X],$$

where

$$\mathbf{D}^2[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X] = \mathbf{E}\left[X^2\right]$$

and

$$\mathbf{E}\left[X^2\right] = \int_{\mathbb{R}} \frac{1}{2} x^2 1_{[-1,1]}(x)\, d\mu_L(x) = \frac{1}{2} \int_{[-1,1]} x^2 d\mu_L(x) = \frac{1}{2} \int_{-1}^{1} x^2 dx = \frac{1}{6} x^3 \Big|_{-1}^{1} = \frac{1}{3}.$$

Therefore,

$$\mathbf{D}^2[Y] = \frac{\beta^2}{3}.$$

Problem 2 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $X : \Omega \to \mathbb{R}$ *be a uniformly distributed random variable with states in the interval* $[-1, 1]$*. In symbols,* $X \sim Unif(-1, 1)$*. Consider the function* $g : \mathbb{R} \to \mathbb{R}$ *given by*

$$g(x) \overset{def}{=} |x|, \quad \forall x \in \mathbb{R},$$

where $|x|$ *is the absolute value of* x*.*

1. *Can you show that the function* $Y : \Omega \to \mathbb{R}$ *given by*
$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$
is a random variable?
2. *Can you compute the distribution function* $F_Y : \mathbb{R} \to \mathbb{R}_+$ *of the random variable* Y*?*
3. *Is* Y *absolutely continuous?*
4. *Are the first and second order moments of* Y *finite?*
5. *If the first and second order moments of* Y *are finite, can you compute* $\mathbf{E}[Y]$ *and* $\mathbf{D}^2[Y]$*?*

Solution. Recall that, since $X \sim Unif(-1, 1)$, the random variable X is absolutely continuous with density

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x),$$

for every $x \in \mathbb{R}$. Now, we have

$$F_Y(y) \overset{\text{def}}{=} \mathbf{P}(Y \le y) = \mathbf{P}(g(X) \le y) = \mathbf{P}(|X| \le y) = \begin{cases} 0, & \text{if } y < 0, \\ \mathbf{P}(-y \le X \le y), & \text{if } y \ge 0. \end{cases}$$

On the othe hand, under the assumption $y \ge 0$, we have

$$\begin{aligned} \mathbf{P}(-y \le X \le y) &= \int_{[-y,y]} f_X(x)\, d\mu_X(x) \\ &= \int_{[-y,y]} \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\ &= \frac{1}{2} \int_{[-y,y]\cap[-1,1]} d\mu_X(x) \\ &= \frac{1}{2} \mu_X([-y, y] \cap [-1, 1]), \end{aligned}$$

where

$$\mu_X([-y, y] \cap [-1, 1]) = \begin{cases} \mu_X([-y, y]) = 2y, & \text{if } y \le 1, \\ \mu_X([-1, 1]) = 2, & \text{if } y > 1. \end{cases}$$

It follows

$$F_Y(y) = \begin{cases} 0, & \text{if } y < 0, \\ y, & \text{if } 0 \le y \le 1, \\ 1, & \text{if } y > 1. \end{cases}$$

We can then recognize that $Y \sim Unif(0, 1)$, which implies that Y is absolutely continuous with density given by

$$f_Y(y) = 1_{[0,1]}(y),$$

for every $y \in \mathbb{R}$, and Y has finite first and second order moments. More specifically

$$\begin{aligned} \mathbf{E}[Y] &= \int_{\mathbb{R}} y f_Y(y)\, d\mu_X(y) = \int_{\mathbb{R}} y 1_{[0,1]}(y)\, d\mu_X(y) \\ &= \int_{[0,1]} y d\mu_X(y) = \int_0^1 y dy = \frac{1}{2} y^2 \Big|_0^1 \\ &= \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} \mathbf{E}\left[Y^2\right] &= \int_{\mathbb{R}} y^2 f_Y(y)\, d\mu_X(y) = \int_{\mathbb{R}} y^2 1_{[0,1]}(y)\, d\mu_X(y) \\ &= \int_{[0,1]} y^2 d\mu_X(y) = \int_0^1 y^2 dy = \frac{1}{3} y^3 \Big|_0^1 \\ &= \frac{1}{3}. \end{aligned}$$

It follows

$$\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Note that, since $Y = |X|$ it would be possible to compute $\mathbf{E}[Y]$ and $\mathbf{E}\left[Y^2\right]$ by using the density of X. That is

$$\begin{aligned} \mathbf{E}[Y] = \mathbf{E}[|X|] &= \int_{\mathbb{R}} |x| f_X(x)\, d\mu_X(x) \\ &= \int_{\mathbb{R}} |x| \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\ &= \frac{1}{2} \int_{[-1,1]} |x|\, d\mu_X(x) \\ &= \frac{1}{2} \left(\int_{[-1,0]} -x d\mu_X(x) + \int_{[0,1]} x d\mu_X(x) \right) \\ &= \frac{1}{2} \left(\int_{-1}^0 -x dx + \int_0^1 x dx \right) \\ &= \frac{1}{2} \left(-\int_{-1}^0 x dx + \int_0^1 x dx \right) \\ &= \frac{1}{2} \left(-\frac{1}{2} x^2 \Big|_{-1}^0 + \frac{1}{2} x^2 \Big|_0^1 \right) \\ &= \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2} \right) \\ &= \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[|X|^2\right] = \mathbf{E}\left[X^2\right] &= \int_{\mathbb{R}} x^2 f_X(x)\, d\mu_X(x) \\
&= \int_{\mathbb{R}} x^2 \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_X(x) \\
&= \frac{1}{2}\int_{[-1,1]} x^2 d\mu_X(x) \\
&= \frac{1}{2}\int_{-1}^{1} x^2 dx \\
&= \frac{1}{2}\frac{1}{3}x^3\Big|_{-1}^{1} \\
&= \frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}\right) \\
&= \frac{1}{3}.
\end{aligned}$$

.

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1,1]$. In symbols, $X \sim Unif(-1,1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} x^2, \quad \forall x \in \mathbb{R}.$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of the random variable Y?*

3. *Is Y absolutely continuous?*

4. *Are the first and second order moments of Y finite?*

5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1,1]$. In symbols, $X \sim Unif(-1,1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} \begin{cases} 0, & \text{if } x \le 0. \\ x^2, & \text{if } x > 0. \end{cases}$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

già fatta intorno a p.30

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*

3. *Is Y absolutely continuous?*

4. *Are the first and second order moments of Y finite?*

5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. Recall that $X \sim Unif(-1,1)$ is absolutely continuous, with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) = \frac{1}{2} 1_{[-1,1]}(x).$$

Note also that we can write

$$g(x) = x^2 1_{(0,+\infty)}(x),$$

for every $x \in \mathbb{R}$.

1. The function g is clearly continuous. In particular, g is a Borel function. Therefore, $Y = g \circ X$ is a random variable.

2. The distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of Y is given by

$$F_Y(y) = \mathbf{P}(Y \le y) = \mathbf{P}(g(X) \le y)$$

for every $y \in \mathbb{R}$. Now, due to the definition of g, we have

$$\{x \in \mathbb{R} : g(x) \le y\} = \begin{cases} \varnothing, & \text{if } y < 0, \\ \{x \in \mathbb{R} : x \le \sqrt{y}\}, & \text{if } y \ge 0. \end{cases}$$

Hence,

$$\{g(X) \le y\} = \begin{cases} \varnothing, & \text{if } y < 0, \\ \{X \le \sqrt{y}\}, & \text{if } y \ge 0. \end{cases}$$

It follows,

$$\mathbf{P}(g(X) \le y)) = \begin{cases} 0, & \text{if } y < 0, \\ \mathbf{P}(X \le \sqrt{y}), & \text{if } y \ge 0. \end{cases}$$

On the other hand, since $X \sim Unif(-1,1)$, we have

$$\begin{aligned}
\mathbf{P}(X \le \sqrt{y}) &= \int_{(-\infty,\sqrt{y}]} f_X(x)\, d\mu_L(x) \\
&= \int_{(-\infty,\sqrt{y}]} \frac{1}{2} 1_{[-1,1]}(x)\, d\mu_L(x) \\
&= \frac{1}{2}\int_{(-\infty,\sqrt{y}]\cap[-1,1]} d\mu_L(x) \\
&= \frac{1}{2}\mu_L\left((-\infty,\sqrt{y}]\cap[-1,1]\right),
\end{aligned}$$

where

$$(-\infty,\sqrt{y}]\cap[-1,1] = \begin{cases} [-1,\sqrt{y}], & \text{if } 0 \le y < 1, \\ [-1,1], & \text{if } y \ge 1. \end{cases}$$

Therefore,

$$\mathbf{P}(X \le \sqrt{y}) = \begin{cases} \frac{1}{2}\left(\sqrt{y}+1\right), & \text{if } y < 1, \\ 1, & \text{if } y \ge 1. \end{cases}$$

We can then write,

$$F_Y(y) = \frac{1}{2}(\sqrt{y}+1)\,1_{[0,1]}(y) + 1_{(1,+\infty)}(y).$$

Note that

$$\mathbf{P}(Y<0) = F_Y(0) = 0.$$

Hence, Y is a non negative random variable.

3. Note that $F_Y : \mathbb{R} \to \mathbb{R}$ is not continuous since

$$\lim_{x\to 0^-} F_Y(x) = 0 \quad \text{and} \quad \lim_{x\to 0^+} F_Y(x) = \frac{1}{2}.$$

A fortiori it is not absolutely continuous.

4. We have

$$\int_\Omega Y^2 d\mathbf{P} = \int_\Omega g(X)^2 \, d\mathbf{P}.$$

Therefore, Y has finite moment of order 2 or not according to whether

$$\int_\Omega g(X)^2 \, d\mathbf{P} < \infty.$$

Now, since X is absolutely continuous, we can write

$$\begin{aligned}\int_\Omega g(X)^2 \, d\mathbf{P} &= \int_\mathbb{R} g(x)^2 f_X(x)\, d\mu_L(x) \\ &= \int_\mathbb{R} x^4 1_{(0,+\infty)}(x)\, 1_{(-1,1)}(x)\, d\mu_L(x) \\ &= \frac{1}{2}\int_{(0,1)} x^4 d\mu_L(x) \\ &= \frac{1}{2}\int_0^1 x^4 dx \\ &= \frac{1}{10}\, x^5\Big|_0^1 \\ &= \frac{1}{10}.\end{aligned}$$

It follows, that Y has finite moment of order 2 and

$$\mathbf{E}\left[Y^2\right] = \int_\Omega Y^2 d\mathbf{P} = \frac{1}{10}.$$

A fortiori Y has finite moment of order 1 and

$$\begin{aligned}\mathbf{E}[Y] = \mathbf{E}[g(X)] &= \int_\mathbb{R} g(x) f_X(x) \; d\mu_L(x) \\ &= \int_\mathbb{R} x^2 1_{(0,+\infty)}(x)\, 1_{(-1,1)}(x) \; d\mu_L(x) \\ &= \frac{1}{2}\int_{(0,1)} x^2 \; d\mu_L(x) \\ &= \frac{1}{2}\int_0^1 x^2 \; dx \\ &= \frac{1}{6}\, x^3\Big|_0^1 \\ &= \frac{1}{6}.\end{aligned}$$

In the end,

$$\mathbf{D}^2[Y] = \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 = \frac{1}{10} - \frac{1}{36} = \frac{13}{180}.$$

Problem 5 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $[-1,1]$. In symbols, $X \sim Unif(-1,1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} x^2 - 2x, \quad \forall x \in \mathbb{R},$$

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 6 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be an exponentially distributed random variable with rate parameter $\lambda = 1$. In symbols, $X \sim Exp(1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} 1 - \exp(-x), \quad \forall x \in \mathbb{R},$$

where $\exp : \mathbb{R} \to \mathbb{R}$ is the Neper exponential function.

1. *Can you show that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega,$$

is a random variable?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the random variable Y?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 7 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a uniformly distributed random variable with states in the interval $(0,1)$. In symbols, $X \sim Unif(0,1)$. Consider the function $g : \mathbb{R}_{++} \to \mathbb{R}$ given by*

$$g(y) \overset{def}{=} -\frac{1}{\lambda}\ln(y), \quad \forall \in \mathbb{R}_{++},$$

where $\ln : \mathbb{R}_{++} \to \mathbb{R}$ is the natural logarithm function and $\lambda > 0$.

1. *Can you state that the function $Y : \Omega \to \mathbb{R}$ given by*

$$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$

is a real random variable on Ω?

2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}$ of $Y : \Omega \to \mathbb{R}$?*

3. *Is Y absolutely continuous?*

4. *Are the first and second order moments of Y finite?*

5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

*Hint: recall the properties of the **logarithm** and **exponential** function.*

Solution.

1. Note that, since $X \sim Unif(0,1)$, that is X has density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) \overset{\text{def}}{=} 1_{[0,1]}(x), \quad \forall x \in \mathbb{R},$$

we have

$$\mathbf{P}(X \leq 0) = \int_{(-\infty,0]} f_X(x)\,d\mu_L(x) = \int_{(-\infty,0]} 1_{[0,1]}(x)\,d\mu_L(x) = \int_{(-\infty,0]\cap[0,1]} d\mu_L(x) = \int_{\{0\}} d\mu_L(x) = \mu_L(0) =$$

Therefore, since $\ln : \mathbb{R}_{++} \to \mathbb{R}$ is a Borel function on $\mathbb{R}_{++}$, the function $Y : \Omega \to \mathbb{R}$ is well defined and it is a real random variable on Ω.

2. Considering that $\lambda > 0$ and the exponential function is the inverse of the logarithm function, we have

$$\{Y \leq y\} = \left\{-\frac{1}{\lambda}\ln(X) \leq y\right\} = \{\ln(X) \geq -\lambda y\} = \left\{X \geq e^{-\lambda y}\right\}$$

for every $y \in \mathbb{R}$. As a consequence,

$$\mathbf{P}(Y \leq y) = \mathbf{P}\left(X \geq e^{-\lambda y}\right) = \int_{\left[e^{-\lambda y},+\infty\right)} 1_{[0,1]}(x)\,d\mu_L(x) = \int_{\left[e^{-\lambda y},+\infty\right)\cap[0,1]} d\mu_L(x).$$

On the other hand,

$$\left[e^{-\lambda y},+\infty\right) \cap [0,1] = \begin{cases} \left[e^{-\lambda y},1\right], & \text{if } y \geq 0, \\ \varnothing, & \text{if } y < 0. \end{cases}$$

Therefore,

$$\mathbf{P}(Y \leq y) = \begin{cases} \mu_L\left(e^{-\lambda y},1\right) = 1 - e^{-\lambda y}, & \text{if } y \geq 0, \\ \mu_L(\varnothing) = 0, & \text{if } y < 0. \end{cases}$$

That is

$$F_Y(y) = \left(1 - e^{-\lambda y}\right) 1_{\mathbb{R}_+}(y).$$

It then follows that Y is an exponentially distributed random variable with rate parameter λ, in symbols $Y \sim Exp(\lambda)$.

3. Since $Y \sim Exp(\lambda)$ it is well known that Y is absolutely continuous with density $f_Y : \mathbb{R} \to \mathbb{R}$ given by

$$f_Y(y) \overset{\text{def}}{=} \lambda e^{-\lambda y} 1_{\mathbb{R}_+}(y), \quad \forall y \in \mathbb{R}.$$

If we are not aware of this, we can observe that

$$F_Y'(y) = \begin{cases} 0, & \text{if } y < 0, \\ \lambda e^{-\lambda y}, & \text{if } y > 0. \end{cases}$$

That is

$$F_Y'(y) = f_Y(y),$$

for every $y \in \mathbb{R} - \{0\}$. On the other hand, F_Y is not differentiable at $y = 0$. Nevertheless, we have

$$\int_{(-\infty,y]} f_Y(v)\,d\mu_L(v) = \int_{(-\infty,y]} \lambda e^{-\lambda v} 1_{\mathbb{R}_+}(v)\,d\mu_L(v) = \int_{(-\infty,y]\cap\mathbb{R}_+} \lambda e^{-\lambda v}\,d\mu_L(v),$$

where

$$(-\infty,y] \cap \mathbb{R}_+ = \begin{cases} \varnothing, & \text{if } y < 0, \\ 0 & \text{if } y = 0, \\ [0,y], & \text{if } 0 < y. \end{cases}$$

Hence,

$$\int_{(-\infty,y]\cap\mathbb{R}_+} \lambda e^{-\lambda y}\,d\mu_L(v) = \begin{cases} 0, & \text{if } y \leq 0, \\ \int_{[0,y]} \lambda e^{-\lambda v}\,d\mu_L(v), & \text{if } y > 0. \end{cases}$$

Now, we have

$$\int_{[0,y]} \lambda e^{-\lambda v}\,d\mu_L(v) = \int_0^y \lambda e^{-\lambda v}\,dv = -\int_0^y de^{-\lambda v} = -\,e^{-\lambda v}\Big|_0^y = 1 - e^{-\lambda y}.$$

It then follows

$$\int_{(-\infty,y]} f_Y(v)\,d\mu_L(v) = \left(1 - e^{-\lambda y}\right) 1_{\mathbb{R}_+}(y) = F_Y(y),$$

which shows that Y is absolutely continuous with density $f_Y : \mathbb{R} \to \mathbb{R}$.

4. Since $Y \sim Exp(\lambda)$ it is well known that Y ha finite moments of order 1 and 2. If we are not aware of this, we can observe that

$$\int_\Omega Y^2\,d\mathbf{P} = \int_{\mathbb{R}} y^2 f_Y(y)\,d\mu_L(y) = \int_{\mathbb{R}} y^2 \lambda e^{-\lambda y} 1_{\mathbb{R}_+}(y)\,d\mu_L(y) = \int_{\mathbb{R}_+} \lambda y^2 e^{-\lambda y}\,d\mu_L(y) = \int_0^{+\infty} \lambda y^2 e^{-\lambda y}\,dy.$$

On the other hand,

$$\int_0^{+\infty} y^2 \lambda e^{-\lambda y}\,dy = \lim_{y\to+\infty} \int_0^y \lambda v^2 e^{-\lambda v}\,dv,$$

where integrating by parts

$$\begin{aligned}\int_0^y \lambda v^2 e^{-\lambda v} dv &= -\int_0^y v^2 de^{-\lambda v} \\ &= -\left. v^2 e^{-\lambda v}\right|_0^y + 2\int_0^y ve^{-\lambda v} dv \\ &= -y^2 e^{-\lambda y} + \frac{2}{\lambda}\int_0^y \lambda ve^{-\lambda v} dv \\ &= -y^2 e^{-\lambda y} - \frac{2}{\lambda}\int_0^y vde^{-\lambda v} \\ &= -y^2 e^{-\lambda y} - \frac{2}{\lambda}\left. ve^{-\lambda v}\right|_0^y + \frac{2}{\lambda}\int_0^y e^{-\lambda v} dv \\ &= -y^2 e^{-\lambda y} - \frac{2}{\lambda} ye^{-\lambda y} - \frac{2}{\lambda^2}\int_0^y de^{-\lambda v} \\ &= -y^2 e^{-\lambda y} - \frac{2}{\lambda} ye^{-\lambda y} - \frac{2}{\lambda^2}\left. e^{-\lambda v}\right|_0^y \\ &= -y^2 e^{-\lambda y} - \frac{2}{\lambda} ye^{-\lambda y} - \frac{2}{\lambda^2} e^{-\lambda v} + \frac{2}{\lambda^2}.\end{aligned}$$

It follows

$$\lim_{y\to+\infty}\int_0^y \lambda v^2 e^{-\lambda v} dv = \lim_{y\to+\infty}\left(-y^2 e^{-\lambda y} - \frac{2}{\lambda} ye^{-\lambda y} - \frac{2}{\lambda^2} e^{-\lambda v} + \frac{2}{\lambda^2}\right) = \frac{2}{\lambda^2}.$$

Therefore, Y ha finite moments of order 2 and we have

$$\mathbf{E}\left[Y^2\right] = \frac{2}{\lambda^2}.$$

This implies also that Y has finite moment of order 1.

5. We have

$$\mathbf{E}\left[Y\right] = \int_\Omega Y d\mathbf{P} = \int_\mathbb{R} yf_Y(y)\, d\mu_L(y) = \int_\mathbb{R} y\lambda e^{-\lambda y} 1_{\mathbb{R}_+}(y)\, d\mu_L(y) = \int_{\mathbb{R}_+} \lambda ye^{-\lambda y} d\mu_L(y) = \int_0^{+\infty} \lambda ye^{-\lambda y} dy.$$

On the other hand,

$$\int_0^{+\infty} y\lambda e^{-\lambda y} dy = \lim_{y\to+\infty}\int_0^y \lambda ve^{-\lambda v} dv,$$

where integrating by parts

$$\begin{aligned}\int_0^y \lambda ve^{-\lambda v} dv &= -\int_0^y vde^{-\lambda v} \\ &= -\left. ve^{-\lambda v}\right|_0^y + \int_0^y e^{-\lambda v} dv \\ &= -ye^{-\lambda y} + \frac{1}{\lambda}\int_0^y \lambda e^{-\lambda v} dv \\ &= -ye^{-\lambda y} - \frac{1}{\lambda}\int_0^y de^{-\lambda v} \\ &= -ye^{-\lambda y} - \frac{1}{\lambda}\left. e^{-\lambda v}\right|_0^y \\ &= -ye^{-\lambda y} - \frac{1}{\lambda} e^{-\lambda y} + \frac{1}{\lambda}\end{aligned}$$

It follows

$$\lim_{y\to+\infty}\int_0^y \lambda ve^{-\lambda v} dv = \lim_{y\to+\infty}\left(-ye^{-\lambda y} - \frac{1}{\lambda} e^{-\lambda y} + \frac{1}{\lambda}\right) = \frac{1}{\lambda}.$$

Thus,

$$\mathbf{E}\left[Y\right] = \frac{1}{\lambda}.$$

In the end,

$$\mathbf{D}^2\left[Y\right] = \mathbf{E}\left[Y^2\right] - \mathbf{E}\left[Y\right]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2},$$

as it is well known.

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Conditional Expectation with Solution 2021-11-25

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$. Assume that $\mathbf{E}\left[Y^2 \mid X\right] = X^2$ and $\mathbf{E}[Y \mid X] = X$. Show that $X = Y$ a.s. on Ω. Hint: is it true that $X = Y$ a.s. on Ω if and only if $\mathbf{E}\left[(X-Y)^2\right] = 0$?*

Solution. Recall that, since $(X-Y)^2 \geq 0$, thanks to the properties of the Lebesgue integral, we have

$$\mathbf{E}\left[(X-Y)^2\right] = 0 \Leftrightarrow X = Y \text{ a.s.on } \Omega.$$

Now, by virtues of the properties of the conditional expectation operator, we can write

$$\mathbf{E}\left[(X-Y)^2\right] = \mathbf{E}\left[\mathbf{E}\left[(X-Y)^2 \mid X\right]\right].$$

On the other hand, since the random variable X is clearly measurable with respect to the σ-algebra $\sigma(X)$, we have

$$\begin{aligned}\mathbf{E}\left[(X-Y)^2 \mid X\right] &= \mathbf{E}\left[X^2 - 2XY + Y^2 \mid X\right] \\ &= \mathbf{E}\left[X^2 \mid X\right] - 2\mathbf{E}[XY \mid X] + \mathbf{E}\left[Y^2 \mid X\right] \\ &= X^2 - 2X\mathbf{E}[Y \mid X] + \mathbf{E}\left[Y^2 \mid X\right].\end{aligned}$$

Therefore, the assumptions on $\mathbf{E}\left[Y^2 \mid X\right]$ and $\mathbf{E}[Y \mid X]$, allow us to conclude that

$$\mathbf{E}\left[(X-Y)^2 \mid X\right] = 0.$$

It follows

$$\mathbf{E}\left[(X-Y)^2\right] = 0,$$

which implies the desired result.

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra. Let $N \subseteq \mathbb{N}$, let $\{F_n\}_{n\in N}$ be a complete system of mutually exclusive events of Ω and let $\mathcal{F}$ be the σ-algebra generated by $\{F_n\}_{n\in N}$. In symbols $\mathcal{F} \equiv \sigma(\{F_n\}_{n\in N})$. We know that a map $Y : \Omega \to \mathbb{R}$ is an $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variable if and only if*

$$Y(\omega) = \sum_{n\in N} y_n 1_{F_n}(\omega), \quad \forall \omega \in \Omega,$$

già fatta

where $(y_n)_{n\in N}$ is a suitable sequence of real numbers.

Exercise 3 *Consider a random variable $X \in L^2(\Omega; \mathbb{R})$ and let $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ the space of all $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variables with finite second-order moment. Use the above claim to prove that*

$$\mathbf{E}[X \mid \mathcal{F}] = \underset{Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})}{\arg\min}\ \mathbf{E}\left[(X-Y)^2\right]$$

As a consequence, after proving that $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ is a subspace of $L^2(\Omega; \mathbb{R})$, show that $\mathbf{E}[X \mid \mathcal{F}]$ is the orthogonal projection of X on $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$.

Solution. The space $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ of all $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variables with finite second-order moment is a subspace of $L^2(\Omega; \mathbb{R})$ because fulfills the conditions for a subset of a Hilbert space to be a subspace of the Hilbert space. In fact, for all $X, Y \in L^2(\Omega_{\mathcal{F}}; \mathbb{R})$, $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variables with finite second-order moment, and all $\alpha, \beta \in \mathbb{R}$ the linear combination $\alpha X + \beta Y$ is also an $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$-random variable with finite second order moment, that is to say $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ is closed with respect to the linear combination. In addition, if $(X_n)_{n\geq 1}$ is a sequence belonging to $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ and such that $X_n \xrightarrow{L^2} X$, where $X \in L^2(\Omega; \mathbb{R})$, we have also $X \in L^2(\Omega_{\mathcal{F}}; \mathbb{R})$, that is to say $L^2(\Omega_{\mathcal{F}}; \mathbb{R})$ is a closed subset of $L^2(\Omega; \mathbb{R})$ in the topology induced by the norm.

Now, given $X \in L^2(\Omega_{\mathcal{F}}; \mathbb{R})$, consider the functional $\Delta_X : L^2(\Omega_{\mathcal{F}}; \mathbb{R}) \to \mathbb{R}_+$ given by

$$\Delta_X(Y) \stackrel{\text{def}}{=} \mathbf{E}\left[(X-Y)^2\right], \quad \forall Y \in L^2(\Omega_{\mathcal{F}}; \mathbb{R}).$$

Since in the case under consideration

$$Y \in L^2(\Omega_{\mathcal{F}}; \mathbb{R}) \Leftrightarrow Y(\omega) = \sum_{n\in N} y_n 1_{F_n}(\omega), \quad \forall \omega \in \Omega,$$

we can write

$$\Delta_X(Y) = \mathbf{E}\left[\left(X - \sum_{n\in N} y_n 1_{F_n}\right)^2\right] \equiv \Delta_X(y_1, \dots, y_n, \dots).$$

Hence,

$$\begin{aligned}\Delta_X(y_1, \dots, y_n, \dots) &= \mathbf{E}\left[X^2 - 2\sum_{n\in N} y_n X 1_{F_n} + \sum_{m,n\in N} y_m y_n 1_{F_m} 1_{F_n}\right] \\ &= \mathbf{E}\left[X^2\right] - 2\sum_{n\in N} y_n \mathbf{E}[X 1_{F_n}] + \sum_{m,n\in N} y_m y_n \mathbf{E}[1_{F_m} 1_{F_n}].\end{aligned}$$

On the other hand,

$$1_{F_m} 1_{F_n} = \begin{cases} 1_{F_n} & \text{if } m = n \\ 1_{\varnothing} & \text{if } m \neq n \end{cases}.$$

Moreover,

$$\mathbf{E}[1_E] = \mathbf{P}(E), \quad \forall E \in \mathcal{E}$$

and

$$\mathbf{E}[X 1_E] = \int_\Omega X 1_E \, d\mathbf{P} = \int_E X \, d\mathbf{P}, \quad \forall E \in \mathcal{E}.$$

Therefore,

$$\Delta_X(Y) = \mathbf{E}\left[X^2\right] - 2\sum_{n\in N} y_n \int_{F_n} X d\mathbf{P} + \sum_{n\in N} y_n^2 \mathbf{P}(F_n).$$

As a consequence,

$$\partial_{y_m} \Delta_X(y_1, \dots, y_n, \dots) = -2\int_{Fm} X d\mathbf{P} + 2y_m \mathbf{P}(F_m), \quad \forall m \in N,$$

which implies

$$\partial_{y_m} \Delta_X(y_1, \dots, y_n, \dots) = 0 \Leftrightarrow y_m = \frac{1}{\mathbf{P}(F_m)} \int_{F_m} X d\mathbf{P} = \mathbf{E}[X \mid F_m], \quad \forall m \in N.$$

Thus, a candidate minimum Y for the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R}) \to \mathbb{R}_+$ takes the form

$$Y = \sum_{n\in N} \mathbf{E}[X \mid F_n] 1_{F_n} = \mathbf{E}[X \mid \mathcal{F}].$$

Now, we have

$$\partial^2_{y_m} \Delta_X (y_1, \ldots, y_n, \ldots) = \mathbf{P}(F_m) > 0$$

and the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R}) \to \mathbb{R}_+$ is known to be convex[1]. It then follow that

$$\mathbf{E}[X \mid \mathcal{F}] = \underset{Y\in L^2(\Omega_{\mathcal{F}};\mathbb{R})}{\arg\min}\ \mathbf{E}\left[(X-Y)^2\right].$$

To complete the proof, it is sufficient to observe that in a Hilbert space the ortoghonal projection of a given vector onto a subspace determines the vector in the subspace of the minimum distance from the given vector.

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra. Let $X, Y \in L^2(\Omega;\mathbb{R})$.*

già fatta p.18/19

1. *Prove in all details that $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω implies $Cov(X,Y) = 0$, but X and Y may not be independent.*

[1] To prove the convexity of the functional $\Delta_X : L^2(\Omega_{\mathcal{F}};\mathbb{R}) \to \mathbb{R}_+$, we may observe that thanks, to the Cauchy-Schwarz inequality and the convexity of the standard quadratic function $f(u) \overset{\text{def}}{=} u^2$, we have

$$\begin{aligned}
\Delta_X(\theta Y + (1-\theta) Z) &= \mathbf{E}\left[(X - (\theta Y + (1-\theta) Z))^2\right] \\
&= \mathbf{E}\left[(\theta(X-Y) + (1-\theta)(X-Z))^2\right] \\
&= \mathbf{E}\left[\theta^2 (X-Y)^2 + 2\theta(1-\theta)(X-Y)(X-Z) + (1-\theta)^2 (X-Z)^2\right] \\
&= \theta^2 \mathbf{E}\left[(X-Y)^2\right] + 2\theta(1-\theta)\mathbf{E}[(X-Y)(X-Z)] + (1-\theta)^2 \mathbf{E}\left[(X-Z)^2\right] \\
&\le \theta^2 \mathbf{E}\left[(X-Y)^2\right] + 2\theta(1-\theta)|\mathbf{E}[(X-Y)(X-Z)]| + (1-\theta)^2 \mathbf{E}\left[(X-Z)^2\right] \\
&\le \theta^2 \mathbf{E}\left[(X-Y)^2\right] + 2\theta(1-\theta)\mathbf{E}\left[(X-Y)^2\right]^{1/2}\mathbf{E}\left[(X-Z)^2\right]^{1/2} + (1-\theta)^2 \mathbf{E}\left[(X-Z)^2\right] \\
&= \left(\theta \mathbf{E}\left[(X-Y)^2\right]^{1/2} + (1-\theta)\mathbf{E}\left[(X-Z)^2\right]^{1/2}\right)^2 \\
&\le \theta \mathbf{E}\left[(X-Y)^2\right] + (1-\theta)\mathbf{E}\left[(X-Z)^2\right],
\end{aligned}$$

for every $\theta \in [0,1]$.

To show the convexity of the standard quadratic function, $f(u) \overset{\text{def}}{=} u^2$, we may observe that the inequality

$$(u-v)^2 \ge 0,$$

which holds true for every $u, v \in \mathbb{R}$, implies

$$-\theta(1-\theta)(u-v)^2 \le 0,$$

which holds true for every $u, v \in \mathbb{R}$ and $\theta \in [0,1]$. The latter can be rewritten as

$$-\theta(1-\theta)\left(u^2 - 2uv + v^2\right) \le 0$$

or equivalently

$$\theta^2 u^2 - \theta u^2 + 2\theta(1-\theta) uv + (1-\theta)^2 v^2 - (1-\theta) v^2.$$

This implies

$$\theta^2 u^2 + 2\theta(1-\theta) uv + (1-\theta)^2 v^2 \le \theta u^2 + (1-\theta) v^2.$$

Hence,

$$(\theta u + (1-\theta) v)^2 \le \theta u^2 + (1-\theta) v^2,$$

which proves the desired result.

2. *Prove in all details that $Cov(X,Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$.*

   Hint: in the first case, to generate a suitable counterexample one may consider the random variables $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$. In the second case consider $X \sim N(0,1)$ and $Y = X^2$.

Solution. .

Problem 5 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be independent standard Gaussian distributed random variables on Ω. Set*

$$U \overset{def}{=} X + Y, \qquad V \overset{def}{=} X - Y.$$

1. *Compute the distributions of U and V.*
2. *Prove that U and V are independent.*
3. *Compute $\mathbf{E}[X \mid U]$, $\mathbf{E}[X \mid V]$, $\mathbf{E}[Y \mid U]$, $\mathbf{E}[Y \mid V]$.*
4. *Compute $\mathbf{E}[XY \mid U]$.*

   Hint: It might be useful to consider $\mathbf{E}\left[X^2 \mid U\right]$ and $\mathbf{E}\left[Y^2 \mid U\right]$.

Solution.

1. Since X and Y are independent Gaussian random variables, X and Y are also jointly Gaussian, that is the random vector $(X,Y)^\intercal$ is Gaussian. By virtue of the equation

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix},$$

   it then follows that the vector $(U,V)^\intercal$ is Gaussian. Hence, its entries U and V are Gaussian. Now,

$$\mathbf{E}[U] = \mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y] = 0$$

   and

$$\mathbf{E}[V] = \mathbf{E}[X-Y] = \mathbf{E}[X] - \mathbf{E}[Y] = 0.$$

   Furthermore,

$$\mathbf{D}^2[U] = \mathbf{D}^2[X+Y] = \mathbf{D}^2[X] + \mathbf{D}^2[Y] = 2$$

   and

$$\mathbf{D}^2[V] = \mathbf{D}^2[X-Y] = \mathbf{D}^2[X] + \mathbf{D}^2[Y] = 2.$$

   We then have

$$U \sim V \sim N(0,2).$$

2. We clearly have,

$$\mathbf{E}[U]\,\mathbf{E}[V] = 0$$

   Moreover,

$$\mathbf{E}[UV] = \mathbf{E}[(X+Y)(X-Y)] = \mathbf{E}\left[X^2 - Y^2\right] = \mathbf{E}\left[X^2\right] - \mathbf{E}\left[Y^2\right] = \mathbf{D}^2[X] - \mathbf{D}^2[Y] = 0$$

   As a consequence,

$$Cov(U,V) = \mathbf{E}[UV] - \mathbf{E}[U]\,\mathbf{E}[V] = 0.$$

   On the other hand, the vector $(U,V)^\intercal$ is Gaussian. Thus, the zero correlation of U and Y implies the independence of U and V.

3. Note that we can write
$$\begin{pmatrix} X \\ U \end{pmatrix} = \begin{pmatrix} X \\ X+Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$
Similarly
$$\begin{pmatrix} X \\ V \end{pmatrix} = \begin{pmatrix} X \\ X-Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$
Hence, the vectors $(X,U)^\intercal$ and $(X,V)^\intercal$ are Gaussian. As a consequence, thanks to the conditional expectation formula for the entries of a Gaussian vector, we can write
$$\mathbf{E}[X \mid U] = \mathbf{E}[X] + \frac{Cov(X,U)}{\mathbf{D}^2[U]}(U - \mathbf{E}[U]) = \frac{Cov(X, X+Y)}{2}U = \frac{\mathbf{D}^2[X] + Cov(X,Y)}{2}U = \frac{1}{2}U.$$
Similarly
$$\mathbf{E}[X \mid V] = \mathbf{E}[X] + \frac{Cov(X,V)}{\mathbf{D}^2[U]}(V - \mathbf{E}[V]) = \frac{Cov(X, X-Y)}{2}V = \frac{\mathbf{D}^2[X] - Cov(X,Y)}{2}V = \frac{1}{2}V.$$
The same argument implies
$$\mathbf{E}[Y \mid U] = -\frac{1}{2}U \quad \text{and} \quad \mathbf{E}[Y \mid V] = -\frac{1}{2}V.$$
Alternatively, thanks to the properties of the conditional expectation and independence of U and V we can write
$$\mathbf{E}[X \mid U] + \mathbf{E}[Y \mid U] = \mathbf{E}[X+Y \mid U] = \mathbf{E}[U \mid U] = U$$
and
$$\mathbf{E}[X \mid U] - \mathbf{E}[Y \mid U] = \mathbf{E}[X-Y \mid U] = \mathbf{E}[V \mid U] = \mathbf{E}[V] = 0$$
Solving for $\mathbf{E}[X \mid U] + \mathbf{E}[Y \mid U]$, we obtain
$$\mathbf{E}[X \mid U] = \frac{1}{2}U \quad \text{and} \quad \mathbf{E}[Y \mid U] = -\frac{1}{2}U$$
We can also write
$$\mathbf{E}[X \mid V] - \mathbf{E}[Y \mid V] = \mathbf{E}[X-Y \mid V] = \mathbf{E}[V \mid V] = V.$$
and
$$\mathbf{E}[X \mid V] + \mathbf{E}[Y \mid V] = \mathbf{E}[X+Y \mid V] = \mathbf{E}[U \mid V] = \mathbf{E}[U] = 0.$$
It follows
$$\mathbf{E}[X \mid V] = \frac{1}{2}V \quad \text{and} \quad \mathbf{E}[Y \mid V] = -\frac{1}{2}V.$$
Note that $\mathbf{E}[X \mid U]$ and $\mathbf{E}[X \mid V]$ are the linear regressions of X against U and V, respectively while $\mathbf{E}[Y \mid U]$ and $\mathbf{E}[Y \mid V]$ are the linear regressions of Y against U and V, respectively.

4. In the end, observe that we have
$$X = \frac{1}{2}(U+V) \quad \text{and} \quad Y = \frac{1}{2}(U-V).$$
Hence,
$$XY = \frac{1}{4}\left(U^2 - V^2\right).$$
It follows
$$\begin{aligned}\mathbf{E}[XY \mid U] &= \frac{1}{4}\mathbf{E}\left[U^2 - V^2 \mid U\right] \\ &= \frac{1}{4}\left(\mathbf{E}\left[U^2 \mid U\right] - \mathbf{E}\left[V^2 \mid U\right]\right) \\ &= \frac{1}{4}\left(U^2 - \mathbf{E}\left[V^2\right]\right) \\ &= \frac{1}{4}\left(U^2 - \mathbf{D}^2[V]\right) \\ &= \frac{1}{4}\left(U^2 - 2\right).\end{aligned}$$

Problem 6 *Let N be a geometric random variable with success probability p, which models the first occurrence of success in n independent trials, and let $(X_n)_{n\geq 1}$ be a sequence of independent and normally distributed random variables with mean μ and variance σ^2, which are also independent of N. Study the conditional expectation*
$$\mathbf{E}\left[\textstyle\sum_{k=1}^N X_k \mid N\right].$$
già fatta

Use the properties of the conditional expectation to compute the expectation and the variance of the random sum
$$S_N \overset{def}{=} \textstyle\sum_{k=1}^N X_k.$$

Solution. .

Problem 7 *Let Z [resp. R] be a standard Gaussian [Rademacher] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $R \sim Rad(1/2)$. Assume that X and R are independent and define $Y \equiv R \cdot X$.*

1. *Is the random variable Y Gaussian?*
2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*
3. *Are the random variables R and Y uncorrelated? Are R and Y independent?*
4. *Does the random vector $(X,Y)^\intercal$ have a bivariate Gaussian distribution? Hint: consider the possibility that $(X,Y)^\intercal$ has a bivariate Gaussian distribution; how the random variable $Z \equiv X+Y$ should be distributed?*
5. *Can you compute $\mathbf{E}[Y \mid X]$ and $\mathbf{E}[X \mid Y]$?*

Solution.

1. To prove that Y is Gaussian we show that
$$\mathbf{P}(Y \leq y) = \mathbf{P}(X \leq y), \tag{1}$$

for every $y \in \mathbb{R}$. To this, on account that $\{R=1\}$, $\{R=-1\}$ constitute a partition of Ω, the random variables R and X are independent, and X is symmetric about 0, we can write

$$\begin{aligned}\mathbf{P}(Y \le y) &= \mathbf{P}(RX \le y)\\ &= \mathbf{P}(RX \le y, R=1) + \mathbf{P}(RX \le y, R=-1)\\ &= \mathbf{P}(RX \le y \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(RX \le y \mid R=-1)\,\mathbf{P}(R=-1)\\ &= \frac{1}{2}\left(\mathbf{P}(X \le y \mid R=1) + \mathbf{P}(X \ge -y \mid R=-1)\right)\\ &= \frac{1}{2}\left(\mathbf{P}(X \le y) + \mathbf{P}(X \ge -y)\right)\\ &= \mathbf{P}(X \le y),\end{aligned}$$

for every $y \in \mathbb{R}$. This proves that $Y \sim X \sim N(0,1)$.

2. Since $Y \equiv R \cdot X$, the intuition is that the observation of the values taken by X transmits information on the values taken by Y. That is X and Y are not independent. However, on account that $\mathbf{E}[R] = 0$ and thanks to the independence of X and R, which implies the independence of X^2 and R, we have

$$\mathbf{E}[XY] = \mathbf{E}[XRX] = \mathbf{E}\left[RX^2\right] = \mathbf{E}[R]\,\mathbf{E}\left[X^2\right] = 0 = \mathbf{E}[X]\,\mathbf{E}[R].$$

This shows that X and Y are uncorrelated. On the other hand, since $X \sim N(0,1)$, we have

$$\mathbf{E}\left[X^2Y^2\right] = \mathbf{E}\left[X^2R^2X^2\right] = \mathbf{E}\left[X^4\right] = 3$$

and

$$\mathbf{E}\left[X^2\right]\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[X^2\right]\mathbf{E}\left[R^2X^2\right] = \mathbf{E}\left[X^2\right]\mathbf{E}\left[X^2\right] = \mathbf{E}\left[X^2\right]^2 = 1.$$

This shows that X^2 and Y^2 are not uncorrelated, which prevents that X^2 and Y^2 are not independent. Eventually, X and Y cannot be independent.

3. On account that $R^2 \sim Dirac(1)$, we have

$$\mathbf{E}[RY] = \mathbf{E}[RRX] = \mathbf{E}\left[R^2X\right] = \mathbf{E}[X] = 0 = \mathbf{E}[X]\,\mathbf{E}[R].$$

This shows that R and Y are uncorrelated. On the other hand, since $Y \equiv R \cdot X \sim N(0,1)$ the intuition is that the observation of the values taken by R transmits no information on the values taken by Y. Hence, the intuition is that R and Y are independent. To prove this, we show that

$$\mathbf{P}(R \le r, Y \le y) = \mathbf{P}(R \le r)\,\mathbf{P}(Y \le y), \tag{2}$$

for all $r, y \in \mathbb{R}$. In fact, still on account that $\{R=1\}$, $\{R=-1\}$ constitute a partition of Ω, the random variables R and X are independent and X is symmetric about 0, we have

$$\begin{aligned}&\mathbf{P}(R \le r, Y \le y,)\\ &= \mathbf{P}(R \le r, Y \le y, R=1) + \mathbf{P}(R \le r, Y \le y, R=-1)\\ &= \mathbf{P}(R \le r, XR \le y, R=1) + \mathbf{P}(R \le r, XR \le y, R=-1)\\ &= \mathbf{P}(R \le r, XR \le y \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(R \le r, XR \le y \mid R=-1)\,\mathbf{P}(R=-1)\\ &= \tfrac{1}{2}\left(\mathbf{P}(1 \le r, X \le y \mid R=1) + \mathbf{P}(-1 \le r, X \ge -y \mid R=-1)\right)\\ &= \begin{cases} 0 & \text{if } r < -1\\ \frac{1}{2}\mathbf{P}(X \ge -y \mid R=-1) = \frac{1}{2}\mathbf{P}(X \ge -y) = \frac{1}{2}\mathbf{P}(X \le y) & \text{if } -1 \le r < 1\\ \frac{1}{2}\left(\mathbf{P}(X \le y \mid R=1) + \mathbf{P}(X \ge -y \mid R=-1)\right) = \frac{1}{2}\left(\mathbf{P}(X \le y) + \mathbf{P}(X \ge -y)\right) = \mathbf{P}(X \le y) & \text{if } 1 \le r \end{cases}\end{aligned}$$

On the other hand

$$\mathbf{P}(R \le r)\,\mathbf{P}(Y \le y) = \begin{cases} 0 & \text{if } r < -1\\ \frac{1}{2}\mathbf{P}(Y \le y) = \frac{1}{2}\mathbf{P}(X \le y) & \text{if } -1 \le r < 1\\ \mathbf{P}(Y \le y) = \mathbf{P}(X \le y) & \text{if } 1 \le r \end{cases}$$

Therefore, the random variables R and Y are independent.

4. If the random vector $(X,Y)^\intercal$ had a bivariate Gaussian distribution, the random variable $Z \equiv X + Y$ should be Gaussian distributed. On the other hand,

$$Z = X + Y = X + RX = (R+1)X.$$

Hence,

$$\begin{aligned}F_Z(x) &= \mathbf{P}(Z \le z)\\ &= \mathbf{P}(Z \le z, R=1) + \mathbf{P}(Z \le z, R=-1)\\ &= \mathbf{P}(Z \le z \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(Z \le z \mid R=-1)\,\mathbf{P}(R=-1)\\ &= \tfrac{1}{2}\left(\mathbf{P}((R+1)X \le z \mid R=1) + \mathbf{P}((R+1)X \le z \mid R=-1)\right)\\ &= \tfrac{1}{2}\left(\mathbf{P}(2X \le z \mid R=1) + \mathbf{P}(0 \le z \mid R=-1)\right).\end{aligned}$$

Now, we have that the events

$$\{2X \le z\} \quad \text{and} \quad \{R=1\}$$

are independent. Moreover,

$$\begin{aligned}\{0 \le z\} &= \Omega \quad \text{if } z \ge 0\\ \{0 \le z\} &= \varnothing \quad \text{if } z < 0\end{aligned}$$

Hence,

$$F_Z(x) = \begin{cases} \frac{1}{2}\mathbf{P}(2X \le z) & \text{if } z < 0\\ \frac{1}{2}\left(\mathbf{P}(2X \le z) + 1\right) & \text{if } z \ge 0 \end{cases}$$

in particular, if $z < 0$, we have

$$F_Z(x) \le \frac{1}{2}\mathbf{P}(2X \le 0) = \frac{1}{4}$$

and if $z \ge 0$

$$F_Z(x) \ge \frac{1}{2}\left(\mathbf{P}(2X \le 0) + 1\right) = \frac{1}{2}\left(\frac{1}{2} + 1\right) = \frac{3}{4},$$

Hence, F_Z cannot not continuous at $z = 0$. This prevents Z to be Gaussian.

5. By virtue of what shown above and the properties of the conditional expectation, we have,

$$\mathbf{E}[Y \mid X] = \mathbf{E}[RX \mid X] = X\mathbf{E}[R \mid X] = X\mathbf{E}[R] = 0$$

and

$$\mathbf{E}[X \mid Y] = \mathbf{E}\left[XR^2 \mid Y\right] = \mathbf{E}[XRR \mid Y] = \mathbf{E}[YR \mid Y] = Y\mathbf{E}[R \mid Y] = Y\mathbf{E}[R] = 0.$$

Problem 8 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Rademacher random variables on Ω. Define $Z \stackrel{\text{def}}{=} X + Y$.*

1. Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.

2. *Are the random variables* $\mathbf{E}[X \mid Z]$ *and* $\mathbf{E}[Y \mid Z]$ *uncorrelated?*

3. *Are the random variables* $\mathbf{E}[X \mid Z]$ *and* $\mathbf{E}[Y \mid Z]$ *independent?*

4. *By using the properties of the conditional expectation, on account that you are dealing with standard Rademacher random variables, can you compute* $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ *and* $\mathbf{E}[XY \mid Z]$*?*

Solution. .

Problem 9 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *a probability space, let* X *and* Y *be independent standard Bernoulli random variables on* Ω*. Define* $Z \stackrel{def}{=} X + Y$.

1. *Compute* $\mathbf{E}[X \mid Z]$ *and* $\mathbf{E}[Y \mid Z]$.

2. *Are the random variables* $\mathbf{E}[X \mid Z]$ *and* $\mathbf{E}[Y \mid Z]$ *uncorrelated?*

3. *Are the random variables* $\mathbf{E}[X \mid Z]$ *and* $\mathbf{E}[Y \mid Z]$ *independent?*

4. *By using the properties of the conditional expectation, on account that you are dealing with standard Bernoulli random variables, can you compute* $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ *and* $\mathbf{E}[XY \mid Z]$*?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Sequences of Random Variables with Solutions 2021-11-23

Problem 1 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* X *be a uniformly distributed real random variable on the interval* $[0,1]$*. In symbols,* $X \sim U(0,1)$*. Consider the sequence* $(Y_n)_{n\geq 1}$ *of real random variables given by*

$$Y_n \stackrel{def}{=} \begin{cases} n, & \text{if } 0 \leq X < \frac{1}{n}, \\ 0, & \text{if } 1/n \leq X \leq 1, \end{cases} \qquad \forall n \geq 1.$$

Check whether the sequence $(Y_n)_{n\geq 1}$ *converges in distribution, converges in probability, converges in mean, converges almost surely, in the assigned order.*

Exercise 2 *Hint: to deal with the almost sure convergence consider the event* $E_0 \equiv \{\omega \in \Omega : X(\omega) = 0\}$ *and the complement* E_0^c.

Solution. . □

Problem 3 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ *be the Euclidean real line endowed with the Borel* σ*-algebra. Prove that the function* $f : \mathbb{R} \to \mathbb{R}_+$ *given by*

$$f(x) \stackrel{def}{=} \frac{\alpha - 1}{x^\alpha} 1_{[1,+\infty)}, \quad \forall x \in \mathbb{R},$$

where $\alpha > 1$*, is a density. Then, consider a random variable* X *with density* $f_X = f$ *and the sequence* $(Y_n)_{n\geq 1}$ *of random variables given by*

$$Y_n \stackrel{def}{=} \frac{X}{n}, \quad \forall n \in \mathbb{N}.$$

Exercise 4 *Study the convergence in distribution, in probability and in p-th mean of the sequence* $(Y_n)_{n\geq 1}$ *on varying of* $\alpha > 1$.

Solution. .

Problem 5 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a complete probability space and let* $(X_n)_{n\geq 1}$ *be a sequence of independent real random variables such that* $X_n \sim Ber(1/n^\alpha)$ *for some* $\alpha > 0$*. Consider the sequence* $(Y_n)_{n\geq 1}$ *of real random variables on* Ω *given by*

$$Y_n \stackrel{def}{=} \min\{X_1, \dots, X_n\}.$$

1. *study the convergence in distribution, in probability and in* $L^p(\Omega; \mathbb{R})$ *of* $(X_n)_{n\geq 1}$ *and* $(Y_n)_{n\geq 1}$ *on varying of* $\alpha > 0$*;*

2. *study the almost sure convergence of* $(X_n)_{n\geq 1}$ *and* $(Y_n)_{n\geq 1}$ *on varying of* $\alpha > 0$.

Solution. .

Exercise 6 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{\geq n}$ be a sequence of real random variables on Ω. Assume that $(X_n)_{\geq n}$ are identically distributed and let $f_X : \mathbb{R} \to \mathbb{R}_+$ their common density function given by*

$$f_X(x) \stackrel{def}{=} \frac{2}{x^3} 1_{(1,+\infty)}(x), \quad \forall x \in \mathbb{R}.$$

Set

$$Y_n \equiv \frac{X_n}{n^\alpha}, \quad \forall n \geq 1,$$

where $\alpha > 0$.

1. *Study the convergence in distribution, probability, and L^p of the sequence $(Y_n)_{n\geq 1}$ on varying of $\alpha > 0$.*

2. *Under the additional assumption of independence of the random variables of the sequence $(X_n)_{\geq n}$, compute $\limsup_{n\to\infty} Y_n$ and $\liminf_{n\to\infty} Y_n$ on varying of $\alpha > 0$. Does the sequence $(Y_n)_{n\geq 1}$ converge almost surely?*

Solution. . □

Sequence di VA random

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Problems on Sequences of Random Variables with Some Solutions 2021-12-07

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X be a uniformly distributed real random variable on the interval $[0,1]$. In symbols, $X \sim U(0,1)$. Consider the sequence $(Y_n)_{n\geq 1}$ of real random variables given by*

$$Y_n \stackrel{def}{=} \begin{cases} n, & \text{if } 0 \leq X < \frac{1}{n}, \\ 0, & \text{if } 1/n \leq X \leq 1, \end{cases} \qquad \forall n \geq 1.$$

Check whether the sequence $(Y_n)_{n\geq 1}$ converges in distribution, converges in probability, converges in mean, converges almost surely, in the order indicated.

Hint: to deal with the almost sure convergence consider the event $E_0 \equiv \{\omega \in \Omega : X(\omega) = 0\}$ and the complement E_0^c.

Solution. Write $F_{Y_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of Y_n. We have

$$F_{Y_n}(y) = \mathbf{P}(Y_n \leq y) = \begin{cases} 0, & \text{if } y < 0, \\ \mathbf{P}(1/n \leq X \leq 1) = 1 - 1/n, & \text{if } 0 \leq y < n, \\ 1, & \text{if } n \leq y. \end{cases}$$

On the other hand, for every $y \geq 0$ there exists $n_y \in \mathbb{N}$, (e.g. $n_y = \lceil y \rceil$, where $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{R}$, is the ceiling function), such that $y < n$ for every $n > n_y$. Therefore, definitively,

$$\mathbf{P}(Y_n \leq y) = 1 - 1/n.$$

It then follows

$$\lim_{n\to\infty} F_{Y_n}(y) = \begin{cases} 0, & \text{if } y < 0, \\ 1 - \lim_{n\to\infty} 1/n = 1, & \text{if } 0 \leq y. \end{cases}$$

Considering the Heavside function $H : \mathbb{R} \to \mathbb{R}$ given by

$$H(y) = \begin{cases} 0, & \text{if } y < 0, \\ 1, & \text{if } 0 \leq y, \end{cases}$$

we clearly have

$$\lim_{n\to\infty} F_{Y_n}(y) = H(y),$$

at any point $y \in \mathbb{R}$. Hence, the sequence $(Y_n)_{n\geq 1}$ converges in distribution to the standard Dirac real random variable $Dir(0)$. With regard to the convergence in probability, we know that the convergence in distribution to a Dirac random variables $Dir(y_0)$, concentrated at some $y_0 \in \mathbb{R}$, implies also the convergence in probability to $Dir(y_0)$. By direct approach, according to the definition, we have

$$\mathbf{P}(Y_n = n) = \mathbf{P}\left(0 \leq X < \frac{1}{n}\right) = \frac{1}{n} \quad \text{and} \quad \mathbf{P}(Y_n = 0) = \mathbf{P}\left(\frac{1}{n} \leq X \leq 1\right) = 1 - \frac{1}{n}.$$

Therefore, definitively,

$$\mathbf{P}(|Y_n| \leq \varepsilon) \geq \mathbf{P}(Y_n = 0) = 1 - \frac{1}{n},$$

for every $\varepsilon > 0$. It follows

$$\lim_{n\to\infty} \mathbf{P}(|Y_n| \leq \varepsilon) \geq 1 - \lim_{n\to\infty} \frac{1}{n} = 1,$$

for every $\varepsilon > 0$, which is the convergence in probability of $(Y_n)_{n\geq 1}$ to $Dir(0)$. Now, to check the convergence in mean, since $Y_n(\Omega) = \{0, 1\}$, we consider

$$\mathbf{E}[Y_n] = n\mathbf{P}(Y_n = n) = n\frac{1}{n} = 1.$$

It follows that

$$\lim_{n\to\infty} \mathbf{E}[|Y_n|] = \lim_{n\to\infty} \mathbf{E}[Y_n] = 1 \neq 0.$$

Hence, $(Y_n)_{n\geq 1}$ does not converge in mean to $Dir(0)$, which implies that $(Y_n)_{n\geq 1}$ does not converge in mean at all (recall that convergence in mean at some random variable implies convergence in probability at the same random variable). In the end, consider the event

$$E_0 \equiv \{\omega \in \Omega : X(\omega) = 0\}.$$

Since $X \sim U(0,1)$, we have $\mathbf{P}(E_0) = 0$. In addition, for every $\omega \in E_0^c$ we have $X(\omega) > 0$ and it is possible to find n_ω such that

$$\frac{1}{n} < X(\omega),$$

for every $n > n_\omega$. It then follows that

$$Y_n(\omega) = 0,$$

for every $n > n_\omega$. This implies

$$\lim_{n\to\infty} Y_n(\omega) = 0,$$

for every $\omega \in E_0^c$, which is the almost sure convergence of the sequence $(Y_n)_{n\geq 1}$ to $Dir(0)$. □

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. Prove that the function $f : \mathbb{R} \to \mathbb{R}$ given by*

$$f(x) \stackrel{def}{=} \frac{\alpha - 1}{x^\alpha} 1_{[1,+\infty)}, \quad \forall x \in \mathbb{R},$$

where $\alpha > 1$, is a density. Then, consider a random variable X with density $f_X = f$ and the sequence $(Y_n)_{n\geq 1}$ of random variables given by

$$Y_n \stackrel{def}{=} \frac{X}{n}, \quad \forall n \in \mathbb{N}.$$

Study the convergence in distribution, in probability, and in p-th mean of the sequence $(Y_n)_{n\geq 1}$, on varying of $\alpha > 1$, in the order indicated.

Solution. Since $\alpha > 1$, that is $\alpha - 1 > 0$ and $1 - \alpha < 0$, we have clearly

$$f(x) \geq 0,$$

for every $x \in \mathbb{R}$, and

$$\begin{aligned}\int_{\mathbb{R}} f(x)\, d\mu_L(x) &= \int_{\mathbb{R}} \frac{\alpha - 1}{x^\alpha} 1_{[1,+\infty)}(x)\, d\mu_L(x) = \int_{[1,+\infty)} \frac{\alpha - 1}{x^\alpha} d\mu_L(x) \\ &= \int_1^{+\infty} \frac{\alpha - 1}{x^\alpha} dx = \lim_{x\to+\infty} \int_1^x \frac{\alpha - 1}{u^\alpha} du = \lim_{x\to+\infty} -\int_1^x du^{1-\alpha} \\ &= \lim_{x\to+\infty} -u^{1-\alpha}\Big|_1^x = 1 - \lim_{x\to+\infty} x^{1-\alpha} = 1.\end{aligned}$$

This proves that $f : \mathbb{R} \to \mathbb{R}$ is a density.

Write $F_{Y_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of Y_n, for every $n \geq 1$. We have

$$F_{Y_n}(y) = \mathbf{P}(Y_n \leq y) = \mathbf{P}(X/n \leq y) = \mathbf{P}(X \leq ny) = \int_{(-\infty,ny]} f(x)\, d\mu_L(x).$$

On the other hand,

$$\begin{aligned}\int_{(-\infty,ny]} f(x)\, d\mu_L(x) &= \int_{(-\infty,ny]} \frac{\alpha - 1}{x^\alpha} 1_{[1,+\infty)}(x)\, d\mu_L(x) \\ &= \int_{(-\infty,ny]\cap[1,+\infty)} \frac{\alpha - 1}{x^\alpha} d\mu_L(x) \\ &= \begin{cases} \int_\varnothing \frac{\alpha-1}{x^\alpha} d\mu_L(x), & \text{if } ny < 1, \\ \int_{\{ny\}} \frac{\alpha-1}{x^\alpha} d\mu_L(x), & \text{if } ny = 1, \\ \int_{[1,ny]} \frac{\alpha-1}{x^\alpha} d\mu_L(x), & \text{if } 1 < ny, \end{cases}\end{aligned}$$

where

$$\int_\varnothing \frac{\alpha - 1}{x^\alpha} d\mu_L(x) = \int_{\{ny\}} \frac{\alpha - 1}{x^\alpha} d\mu_L(x) = 0$$

and

$$\int_{[1,ny]} \frac{\alpha - 1}{x^\alpha} d\mu_L(x) = \int_1^{ny} \frac{\alpha - 1}{x^\alpha} dx = -\int_1^{ny} dx^{1-\alpha} = -x^{1-\alpha}\Big|_1^{ny} = 1 - \frac{1}{n^{\alpha-1}y^{\alpha-1}}.$$

Therefore,

$$F_{Y_n}(y) = \begin{cases} 0, & \text{if } y \leq \frac{1}{n}, \\ 1 - \frac{1}{n^{\alpha-1}y^{\alpha-1}}, & \text{if } \frac{1}{n} < y. \end{cases}$$

As a consequence,

$$\lim_{n\to\infty} F_{Y_n}(y) = \begin{cases} 0, & \text{if } y \leq 0, \\ 1 - \lim_{n\to\infty} \frac{1}{n^{\alpha-1}y^{\alpha-1}} = 1, & \text{if } 0 < y. \end{cases}$$

Considering the Heavside function $H : \mathbb{R} \to \mathbb{R}$ given by

$$H(y) = \begin{cases} 0, & \text{if } y < 0, \\ 1, & \text{if } 0 \leq y, \end{cases}$$

we then have

$$\lim_{n\to\infty} F_{Y_n}(y) = H(y),$$

at any point $y \in \mathbb{R} - \{0\}$, where the Heavside function is continuous. Hence, the sequence $(Y_n)_{n\geq 1}$ converges in distribution to the standard Dirac real random variable $Dir(0)$. With regard to the convergence in probability, we know that the convergence in distribution to a Dirac random variables $Dir(y_0)$, concentrated at some $y_0 \in \mathbb{R}$, implies also the convergence in probability to $Dir(y_0)$. By direct approach, since

$$F_{Y_n}(y) = \int_{(-\infty,y]} \frac{1 - \alpha}{n^{\alpha-1}u^\alpha} 1_{(1/n,+\infty)}(u)\, d\mu_L(u)$$

for every $y \in \mathbb{R}$, we have that the random variables of the sequence $(Y_n)_{n\geq 1}$ are absolutely continuous with density $f_{Y_n} : \mathbb{R} \to \mathbb{R}$ given by

$$f_{Y_n}(y) = \frac{1 - \alpha}{n^{\alpha-1}y^\alpha} 1_{(1/n,+\infty)}(y), \quad \forall y \in \mathbb{R}.$$

Note that

$$F'_{Y_n}(y) = f_{Y_n}(y),$$

for every $y \neq 1/n$. As a consequence, provided n is sufficiently large,

$$
\begin{aligned}
\mathbf{P}(|Y_n| > \varepsilon) = \mathbf{P}(Y_n > \varepsilon) &= \int_{(\varepsilon,+\infty)} f_{Y_n}(y)\, d\mu_L(y) \\
&= \int_{(\varepsilon,+\infty)} \frac{1-\alpha}{n^{\alpha-1}y^{\alpha}} 1_{(1/n,+\infty)}(y)\, d\mu_L(y) \\
&= \int_{(\varepsilon,+\infty)\cap(1/n,+\infty)} \frac{1-\alpha}{n^{\alpha-1}y^{\alpha}} d\mu_L(y) \\
&= \int_{(\varepsilon,+\infty)} \frac{1-\alpha}{n^{\alpha-1}y^{\alpha}} d\mu_L(y) \\
&= \int_{\varepsilon}^{+\infty} \frac{1-\alpha}{n^{\alpha-1}y^{\alpha}} dy \\
&= \frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} \int_{\varepsilon}^{y} \frac{\alpha-1}{u^{\alpha}} du \\
&= -\frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} \int_{\varepsilon}^{y} du^{\alpha-1} \\
&= -\frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} u^{1-\alpha}\big|_{\varepsilon}^{y} \\
&= -\frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} \left(\frac{1}{y^{\alpha-1}} - \frac{1}{\varepsilon^{\alpha-1}}\right) \\
&= \frac{1}{n^{\alpha-1}\varepsilon^{\alpha-1}}.
\end{aligned}
$$

It follows

$$
\lim_{n\to\infty} \mathbf{P}(|Y_n| > \varepsilon) = \lim_{n\to\infty} \frac{1}{n^{\alpha-1}\varepsilon^{\alpha-1}} = 0,
$$

for every $\varepsilon > 0$. This proves directly that $(Y_n)_{n\geq1}$ converges in probability to $Dir(0)$.

By virtue of what shown above, to study the convergence in p-th mean of the sequence $(Y_n)_{n\geq1}$ it is sufficient to consider

$$
\begin{aligned}
\mathbf{E}[Y_n^p] &= \int_{\mathbb{R}} y^p f_{Y_n}(y)\, d\mu_L(u) = \int_{\mathbb{R}} \frac{1-\alpha}{n^{\alpha-1}y^{\alpha-p}} 1_{(1/n,+\infty)}(y)\, d\mu_L(u) = \frac{1-\alpha}{n^{\alpha-1}} \int_{(1/n,+\infty)} \frac{1}{y^{\alpha-p}} d\mu_L(u) \\
&= \frac{1-\alpha}{n^{\alpha-1}} \int_{1/n}^{+\infty} \frac{1}{y^{\alpha-p}} dy = \frac{\alpha-1}{n^{\alpha-1}} \lim_{y\to+\infty} \int_{1/n}^{y} \frac{1}{u^{\alpha-p}} du \\
&= \begin{cases} \frac{\alpha-1}{p-\alpha+1}\frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} \int_{1/n}^{y} du^{p-\alpha+1} = \frac{\alpha-1}{p-\alpha+1}\frac{1}{n^{\alpha-1}} \lim_{y\to+\infty} u^{p-\alpha+1}\big|_{1/n}^{y}, & \text{if } p \neq \alpha-1, \\ \frac{\alpha-1}{n^{\alpha-1}} \lim_{y\to+\infty} \int_{1/n}^{y} d\ln(u) = \frac{\alpha-1}{n^{\alpha-1}} \lim_{y\to+\infty} \ln(u)\big|_{1/n}^{y}, & \text{if } p = \alpha-1. \end{cases}
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
\mathbf{E}[Y_n^p] = \mathbf{E}\left[\left(\frac{X}{n}\right)^p\right] &= \int_{\mathbb{R}} \frac{x^p}{n^p} f_X(x)\, d\mu_L(x) = \int_{\mathbb{R}} \frac{x^p}{n^p}\frac{\alpha-1}{x^{\alpha}} 1_{[1,+\infty)}(x)\, d\mu_L(x) \\
&= \frac{\alpha-1}{n^p} \int_{[1,+\infty)} \frac{1}{x^{\alpha-p}} d\mu_L(x) = \frac{\alpha-1}{n^p} \int_{1}^{+\infty} \frac{1}{x^{\alpha-p}} dx = \frac{\alpha-1}{n^p} \lim_{x\to+\infty} \int_{1}^{x} \frac{1}{u^{\alpha-p}} du \\
&= \begin{cases} \frac{\alpha-1}{p-\alpha+1}\frac{1}{n^p} \lim_{x\to+\infty} \int_{1}^{x} \frac{1}{u^{\alpha-p}} u^{p-\alpha+1} = \frac{\alpha-1}{p-\alpha+1}\frac{1}{n^p} \lim_{x\to+\infty} u^{p-\alpha+1}\big|_{1}^{x}, & \text{if } p \neq \alpha-1, \\ \frac{\alpha-1}{n^p} \lim_{x\to+\infty} \int_{1}^{x} d\ln(u) = \frac{\alpha-1}{n^p} \lim_{x\to+\infty} \ln(u)\big|_{1}^{x}, & \text{if } p = \alpha-1. \end{cases}
\end{aligned}
$$

Now, if $p \geq \alpha - 1$ we have that $\mathbf{E}[Y_n^p]$ is not finite. The sequence $(Y_n)_{n\geq1}$ cannot converge in p-th mean. If $1 \leq p < \alpha - 1$, we have

$$
\mathbf{E}[Y_n^p] = -\frac{\alpha-1}{p-\alpha+1}\frac{1}{n^{\alpha-1}}\frac{1}{n^{p-\alpha+1}} = -\frac{\alpha-1}{p-\alpha+1}\frac{1}{n^p}.
$$

Hence,

$$
\lim_{n\to\infty} \mathbf{E}[Y_n^p] = \lim_{n\to\infty} -\frac{\alpha-1}{p-\alpha+1}\frac{1}{n^p} = 0.
$$

The sequence converges in p-th mean to the standard Dirac random variable.

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{\geq n}$ be a sequence of real random variables on Ω. Assume that $(X_n)_{\geq n}$ are identically distributed and let $f_X : \mathbb{R} \to \mathbb{R}$ their common density function given by*

$$
f_X(x) \overset{def}{=} \frac{2}{x^3} 1_{(1,+\infty)}(x), \quad \forall x \in \mathbb{R}.
$$

Set

$$
Y_n \equiv \frac{X_n}{n^{\alpha}}, \quad \forall n \geq 1,
$$

where $\alpha > 0$.

1. *Study the convergence in distribution, probability, and L^p of the sequence $(Y_n)_{n\geq1}$, on varying of $\alpha > 0$, in the order indicated.*

2. *Under the additional assumption of independence of the random variables of the sequence $(X_n)_{\geq n}$, compute $\limsup_{n\to\infty} Y_n$ and $\liminf_{n\to\infty} Y_n$ on varying of $\alpha > 0$. Does the sequence $(Y_n)_{n\geq1}$ converge almost surely?*

Solution. . □ fatta a p.103

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a complete probability space and let $(X_n)_{n\geq1}$ be a sequence of independent real random variables such that $X_n \sim Ber(1/n^{\alpha})$ for some $\alpha > 0$. Consider the sequence $(Y_n)_{n\geq1}$ of real random variables on Ω given by*

$$
Y_n \overset{def}{=} \min\{X_1, \ldots, X_n\}.
$$

1. *study the convergence in distribution, in probability and in $L^p(\Omega; \mathbb{R})$ of $(X_n)_{n\geq1}$ and $(Y_n)_{n\geq1}$ on varying of $\alpha > 0$;*

2. *study the almost sure convergence of $(X_n)_{n\geq1}$ and $(Y_n)_{n\geq1}$ on varying of $\alpha > 0$.*

Solution.

1. We clearly have

$$
Y_n(\omega) = \begin{cases} 1 & \Leftrightarrow X_1(\omega) = \cdots = X_n(\omega) = 1 \\ 0 & \text{otherwise} \end{cases}
$$

Hence, by virtue of the independence of the random variables of the sequence $(X_n)_{n\geq1}$, we have

$$
\mathbf{P}(Y_n = 1) = \mathbf{P}(X_1 = 1, \ldots, X_n = 1) = \mathbf{P}(X_1 = 1)\cdots\mathbf{P}(X_n = 1) = \textstyle\prod_{k=1}^{n} \frac{1}{k^{\alpha}} = \frac{1}{n!^{\alpha}}
$$

and

$$
\mathbf{P}(Y_n = 0) = 1 - \mathbf{P}(Y_n = 1) = 1 - \frac{1}{n!^{\alpha}}.
$$

In other words, $(Y_n)_{n\geq 1}$ is a sequence of standard Bernoulli random variables with succes probability $\frac{1}{n!^\alpha}$. Considering the distribution functions $F_{X_n}:\mathbb{R}\to\mathbb{R}_+$ and $F_{Y_n}:\mathbb{R}\to\mathbb{R}_+$ of X_n and Y_n, respectively, we have

$$F_{X_n}(x) \stackrel{\text{def}}{=} \mathbf{P}(X_n \leq x) = \begin{cases} 0, & \text{if } x<0,\\ 1-\frac{1}{n^\alpha}, & \text{if } 0\leq x<1,\\ 1, & \text{if } 1\leq x,\end{cases}$$

and

$$F_{Y_n}(x) \stackrel{\text{def}}{=} \mathbf{P}(Y_n \leq x) = \begin{cases} 0, & \text{if } x<0,\\ 1-\frac{1}{n!^\alpha}, & \text{if } 0\leq x<1,\\ 1, & \text{if } 1\leq x.\end{cases}$$

Therefore, considering the Heaviside function $H:\mathbb{R}\to\mathbb{R}_+$ given by

$$H(x) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } x<0,\\ 1, & \text{if } 0\leq x,\end{cases}$$

we have

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty} F_{Y_n}(x) = H(x),$$

for every $x\in\mathbb{R}$. Thus, both the sequences $(F_{X_n})_{n\geq 1}$ and $(F_{Y_n})_{n\geq 1}$ converge pointwise to H. It follows that both the sequences $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ converge to the standard Dirac real random variable $Dir(0)$. With regard to the convergence in probability, we know that the convergence in distribution to a Dirac random variables $Dir(y_0)$, concentrated at some $y_0\in\mathbb{R}$, implies also the convergence in probability to $Dir(y_0)$. However, according to the definition, we have definitively

$$\mathbf{P}(|X_n - Dir(0)|<\varepsilon) = \mathbf{P}(X_n<\varepsilon) = \mathbf{P}(X_n=0) = 1-\frac{1}{n^\alpha}$$

and

$$\mathbf{P}(|Y_n - Dir(0)|<\varepsilon) = \mathbf{P}(Y_n<\varepsilon) = \mathbf{P}(Y_n=0) = 1-\frac{1}{n!^\alpha},$$

for every $\varepsilon>0$. Therefore,

$$\lim_{n\to\infty}\mathbf{P}(|X_n - Dir(0)|<\varepsilon) = \lim_{n\to\infty}\left(1-\frac{1}{n^\alpha}\right) = 1$$

and

$$\lim_{n\to\infty}\mathbf{P}(|Y_n - Dir(0)|<\varepsilon) = \lim_{n\to\infty}\left(1-\frac{1}{n!^\alpha}\right) = 1,$$

which is the convergence in probabiliyty of both $(X_n)_{n\geq 1}$ and $(X_n)_{n\geq 1}$ to $Dir(0)$. To check the convergence in $L^p(\Omega;\mathbb{R})$, we need to consider $\|X_n - Dir(0)\|_p$ and $\|Y_n - Dir(0)\|_p$, because in case of convergence the limit has to be $Dir(0)$. We then have

$$\|X_n - Dir(0)\|_p = \left(\int_\Omega |X_n - Dir(0)|^p \, d\mathbf{P}\right)^{1/p} = \left(\int_\Omega X_n^p \, d\mathbf{P}\right)^{1/p} = \mathbf{P}(X_n=1)^{1/p} = \frac{1}{n^{\frac{\alpha}{p}}}$$

and

$$\|Y_n - Dir(0)\|_p = \left(\int_\Omega |Y_n - Dir(0)|^p \, d\mathbf{P}\right)^{1/p} = \left(\int_\Omega Y_n^p \, d\mathbf{P}\right)^{1/p} = \mathbf{P}(Y_n=1)^{1/p} = \frac{1}{n!^{\frac{\alpha}{p}}},$$

for every $\alpha>0$. Hence, we obtain

$$\lim_{n\to\infty}\|X_n - X\|_p = \lim_{n\to\infty}\|Y_n - X\|_p = 0,$$

which proves the convergence of both $(X_n)_{n\geq 1}$ and $(X_n)_{n\geq 1}$ to $Dir(0)$ in $L^p(\Omega;\mathbb{R})$.

2. With regard to the almost sure convergence, note that also in this case, if the sequences $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ converge almost surely, the limit has to be $Dir(0)$. Hence, for any fixed $\varepsilon>0$ consider the events $E_0\equiv\{|X_n - Dir(0)|\geq\varepsilon\}$ and $F_0\equiv\{|Y_n - Dir(0)|\geq\varepsilon\}$ we have

$$\{|X_n - Dir(0)|\geq\varepsilon\} = \{X_n\geq\varepsilon\} \quad\text{and}\quad \{|Y_n - Dir(0)|\geq\varepsilon\} = \{Y_n\geq\varepsilon\}.$$

Hence,

$$\mathbf{P}(|X_n - Dir(0)|\geq\varepsilon) = \begin{cases} \mathbf{P}(X_n=1)=\frac{1}{n^\alpha}, & \text{if } 0<\varepsilon\leq 1,\\ 0, & \text{if } \varepsilon>1,\end{cases}$$

and

$$\mathbf{P}(|Y_n - Dir(0)|\geq\varepsilon) = \begin{cases} \mathbf{P}(Y_n=1)=\frac{1}{n!^\alpha}, & \text{if } 0<\varepsilon\leq 1,\\ 0, & \text{if } \varepsilon>1.\end{cases}$$

As a consequence,

$$\textstyle\sum_{n=1}^\infty \mathbf{P}(|X_n - Dir(0)|\geq\varepsilon) = \begin{cases} \sum_{n=1}^\infty\frac{1}{n^\alpha}, & \text{if } 0<\varepsilon\leq 1,\\ 0, & \text{if } \varepsilon>1,\end{cases}$$

and

$$\textstyle\sum_{n=1}^\infty \mathbf{P}(|Y_n - Dir(0)|\geq\varepsilon) = \begin{cases} \sum_{n=1}^\infty\frac{1}{n!^\alpha}, & \text{if } 0<\varepsilon\leq 1,\\ 0, & \text{if } \varepsilon>1.\end{cases}$$

It then follows that $\sum_{n=1}^\infty \mathbf{P}(|X_n - Dir(0)|\geq\varepsilon)$ converges for every $\alpha>1$ and $\sum_{n=1}^\infty \mathbf{P}(|X_n - Dir(0)|\geq\varepsilon)$ converges for every $\alpha>0$. This yields the almost sure convergence of $(X_n)_{n\geq 1}$ to $Dir(0)$ for every $\alpha>1$ and the almost sure convergence of $(Y_n)_{n\geq 1}$ to $Dir(0)$ for every $\alpha>0$. In fact, the convergence of the series implies that

$$\lim_{m\to\infty}\mathbf{P}\left(\bigcup_{n\geq m}\{|Z_n - Z|\geq\varepsilon\}\right) \leq \textstyle\sum_{n=m}^\infty \mathbf{P}(|Z_n - Z|\geq\varepsilon) = 0.$$

To check the almost sure convergence of the sequence $(X_n)_{n\geq 1}$ to $Dir(0)$ when $0<\alpha\leq 1$, let us start by considering the case $\alpha=1$. Choosing any $\varepsilon<1$, on account of the independence of the random variables of the sequence $(X_n)_{n\geq 1}$, we estimate

$$\begin{aligned}\mathbf{P}\left(\textstyle\bigcap_{n\geq m}\{|X_n|\leq\varepsilon\}\right) &\leq \mathbf{P}\left(\textstyle\bigcap_{n=m}^{2m}\{|X_n|\leq\varepsilon\}\right) = \textstyle\prod_{n=m}^{2m}\mathbf{P}(|X_n|\leq\varepsilon)\\ &= \textstyle\prod_{n=m}^{2m}\mathbf{P}(X_n=0) = \prod_{n=m}^{2m}\left(1-\frac{1}{n}\right)\\ &\leq \textstyle\prod_{n=m}^{2m}\left(1-\frac{1}{2m}\right) = \left(1-\frac{1}{2m}\right)^m.\end{aligned}$$

As a consequence,

$$\lim_{m\to\infty}\mathbf{P}\left(\textstyle\bigcap_{n\geq m}\{|X_n|\leq\varepsilon\}\right) \leq \lim_{m\to\infty}\left(1-\frac{1}{2m}\right)^m = e^{-1/2}<1.$$

This prevents that

$$\lim_{m\to\infty}\mathbf{P}\left(\textstyle\bigcap_{n\geq m}\{|X_n|\leq\varepsilon\}\right) = 1,$$

so that $X_n \stackrel{\text{a.s.}}{\not\to} 0$. With regard to the case $0<\alpha<1$, it is then sufficient to observe that we have

$$\left(1-\frac{1}{n^\alpha}\right) < \left(1-\frac{1}{n}\right),$$

for every $n \in \mathbb{N}$. By an analogous computation as above, it then follows

$$\lim_{m\to\infty} \mathbf{P}\left(\bigcap_{n\geq m}\{|X_n| \leq \varepsilon\}\right) \leq \lim_{n\to\infty}\left(1-\frac{1}{2m^{\alpha}}\right)^m \leq \lim_{n\to\infty}\left(1-\frac{1}{2m}\right)^m = e^{-1/2} < 1.$$

Alternatively, considering that

$$\log(1-x) = -x + o(x),$$

we can directly compute

$$\lim_{n\to\infty}\left(1-\frac{1}{2m^{\alpha}}\right)^m = \lim_{n\to\infty}\exp\left(m\log\left(1-\frac{1}{2m^{\alpha}}\right)\right) = \lim_{n\to\infty}\exp\left(-\frac{m}{2m^{\alpha}}\right) = \lim_{n\to\infty}\exp\left(-\frac{m^{1-\alpha}}{2}\right) = 0.$$

As a consequence, we still have $X_n \overset{\text{a.s.}}{\nrightarrow} 0$.

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Point Estimators 2021-12-17

Problem 1 *A real random variable X on a probability space $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$, which represents the reaction time at some stimulus, has a uniform distribution on an interval $[0,\theta]$, where $\theta > 0$ is a parameter. An investigator wants to estimate θ on the basis of a simple random sample $X_1, \dots, X_n$ of reaction times. Since θ is the largest possible time in the entire population of reaction times, the investigator considers as a first estimator for the parameter θ the largest sample reaction time, that is the statistic*

$$\hat{\theta}_1 \equiv \check{X}_n \equiv \max(X_1, \dots, X_n).$$

1. *Is $\check{X}_n$ unbiased? In case $\check{X}_n$ is not unbiased, is it possible to derive from $\check{X}_n$ an unbiased estimator of θ?*
2. *As a second estimator, the investigator considers the statistic*

   $$\hat{\theta}_2 \equiv \bar{X}_n \equiv \frac{1}{n}\sum_{k=1}^n X_k.$$

   Is $\bar{X}_n$ unbiased? In case $\bar{X}_n$ is not unbiased, is it possible to derive from $\bar{X}_n$ an unbiased estimator of θ?
3. *In the investigator's shoes, what estimator would you prefer among those considered?*
4. *Is $\check{X}_n$ consistent in probability? Is $\check{X}_n$ consistent in mean square?*

Solution.

1. Writing $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of the statistic $\check{X}_n$, we have

   $$\begin{aligned} F_{\check{X}_n}(x) &= \mathbf{P}\left(\check{X}_n \leq x\right) = \mathbf{P}(X_1 \leq x, \dots, X_n \leq x) = \textstyle\prod_{k=1}^n \mathbf{P}(X_k \leq x) \\ &= \textstyle\prod_{k=1}^n \mathbf{P}(X \leq x) = \mathbf{P}(X \leq x)^n = F_X(x)^n, \end{aligned}$$

   for every $x \in \mathbb{R}$, where $F_X : \mathbb{R} \to \mathbb{R}$ is the distribution function of X. On the other hand, since X is uniformly distributed on $[0,\theta]$, we know that X is absolutely continuous with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

   $$f_X(x) \stackrel{\text{def}}{=} \frac{1}{\theta}1_{[0,\theta]}(x), \quad \forall x \in \mathbb{R}.$$

   Hence,

   $$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\, d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{\theta}1_{[0,\theta]}(u)\, d\mu_L(u) = \frac{1}{\theta}\int_{(-\infty,x]\cap[0,\theta]} d\mu_L(u) \\ &= \begin{cases} \frac{1}{\theta}\int_{\varnothing} d\mu_L(u) = 0, & \text{if } x < 0, \\ \frac{1}{\theta}\int_{[0,x]} d\mu_L(u) = \frac{x}{\theta}, & \text{if } 0 \leq x \leq \theta, \\ \frac{1}{\theta}\int_{[0,\theta]} d\mu_L(u) = 1, & \text{if } \theta < x. \end{cases} \end{aligned}$$

   Briefly,

   $$F_X(x) = \frac{x}{\theta}1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

   for every $x \in \mathbb{R}$. It then follows,

   $$F_{\check{X}_n}(x) = F_X(x)^n = \frac{x^n}{\theta^n}1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

   for every $x \in \mathbb{R}$. Now, we have

   $$F'_{\check{X}_n}(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{nx^{n-1}}{\theta^n}, & \text{if } 0 < x < \theta, \\ 0, & \text{if } \theta < x, \end{cases}$$

but $F_{\check{X}_n}$ is not everywhere differentable. Eventually, is not differentiable at the point $x = \theta$. However, considering the function $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ given by

$$f_{\check{X}_n}(x) \stackrel{\text{def}}{=} \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x), \quad \forall x \in \mathbb{R},$$

a straightforward computation shows that

$$F_{\check{X}_n}(x) = \int_{(-\infty,x]} f_{\check{X}_n}(u)\ d\mu_L(u),$$

for every $x \in \mathbb{R}$. This implies that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n}$. As a consequence,

$$\begin{aligned}\mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} x \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\ d\mu_L(x) = \frac{n}{\theta^n}\int_{(0,\theta)} x^n\ d\mu_L(x) \\ &= \frac{n}{\theta^n}\int_0^\theta x^n\ dx = \frac{n}{\theta^n}\left.\frac{x^{n+1}}{n+1}\right|_0^\theta = \frac{n}{n+1}\theta.\end{aligned}$$

We conclude that $\check{X}_n$ is not a unbiased estimator of θ but $\frac{n+1}{n}\check{X}_n$ is an unbiased estimator of θ.

2. We have

$$\begin{aligned}\mathbf{E}\left[\bar{X}_n\right] &= \mathbf{E}[X] = \int_{\mathbb{R}} x f_X(x)\ d\mu_L(x) = \int_{\mathbb{R}} \frac{x}{\theta} 1_{[0,\theta]}(x)\ d\mu_L(x) \\ &= \frac{1}{\theta}\int_{[0,\theta]} x\ d\mu_L(x) = \frac{1}{\theta}\int_0^\theta x\ dx = \frac{1}{\theta}\left.\frac{x^2}{2}\right|_0^\theta = \frac{\theta}{2}.\end{aligned}$$

Hence, $\bar{X}_n$ is not a unbiased estimator of θ but $2\bar{X}_n$ is an unbiased estimator of θ.

3. From 1. and 2. we know that

$$\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right] = \theta \quad \text{and} \quad \mathbf{E}\left[2\bar{X}_n\right] = \theta.$$

Hence, both $\frac{n+1}{n}\check{X}_n$ and $2\bar{X}_n$ are unbiased estimators of the parameter θ. To choose which is preferable between them, we consider

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] \quad \text{and} \quad \mathbf{D}^2\left[2\bar{X}_n\right].$$

We have

$$\begin{aligned}\mathbf{E}\left[\check{X}_n^2\right] &= \int_{\mathbb{R}} x^2 f_{\check{X}_n}(x)\, d\mu_L(x) = \int_{\mathbb{R}} x^2 \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\, d\mu_L(x) = \frac{n}{\theta^n}\int_{(0,\theta)} x^{n+1} d\mu_L(x) \\ &= \frac{n}{\theta^n}\int_0^\theta x^{n+1}\ dx = \frac{n}{\theta^n}\left.\frac{x^{n+2}}{n+2}\right|_0^\theta = \frac{n}{n+2}\theta^2.\end{aligned}$$

Therefore,

$$\mathbf{D}^2\left[\check{X}_n\right] = \mathbf{E}\left[\check{X}_n^2\right] - \mathbf{E}\left[\check{X}_n\right]^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n}{(n+1)^2(n+2)}\theta^2.$$

As a consequence,

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \mathbf{D}^2\left[\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \frac{n}{(n+1)^2(n+2)}\theta^2 = \frac{\theta^2}{n(n+2)}.$$

On the other hand,

$$\mathbf{D}^2\left[2\bar{X}_n\right] = 4\mathbf{D}^2\left[\bar{X}_n\right] = \frac{4}{n}\mathbf{D}^2[X] = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

Now, we clearly have

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] < \mathbf{D}^2\left[2\bar{X}_n\right],$$

for every $n > 1$. It follows that the estimator $\frac{n+1}{n}\check{X}_n$ is preferable to $2\bar{X}_n$.

4. We have

$$\mathbf{P}\left(\left|\frac{n+1}{n}\check{X}_n - \theta\right| \geq \varepsilon\right) = \mathbf{P}\left(\left|\frac{n+1}{n}\check{X}_n - \mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right]\right| \geq \varepsilon\right) \leq \frac{\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right]}{\varepsilon^2} = \frac{\theta^2}{n(n+2)\varepsilon^2}.$$

It clearly follows that

$$\frac{n+1}{n}\check{X}_n \xrightarrow{\mathbf{P}} \theta.$$

On the other hand, trivially

$$\frac{n}{n+1} \xrightarrow{\mathbf{P}} 1.$$

As a consequence, we have

$$\check{X}_n = \frac{n}{n+1}\cdot\frac{n+1}{n}\check{X}_n \xrightarrow{\mathbf{P}} 1\cdot\theta = \theta.$$

Hence, both the estimators $\frac{n+1}{n}\check{X}_n$ and $\check{X}_n$ are consistent in probability. In addition, considering that

$$\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right] = \theta \quad \text{and} \quad \mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] = \frac{\theta^2}{n(n+2)},$$

we obtain

$$\begin{aligned}\mathbf{E}\left[\left(\check{X}_n - \theta\right)^2\right] &= \left(\frac{n}{n+1}\right)^2 \mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n - \frac{n+1}{n}\theta\right)^2\right] \\ &= \left(\frac{n}{n+1}\right)^2 \mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n - \theta - \frac{\theta}{n}\right)^2\right] \\ &= \left(\frac{n}{n+1}\right)^2 \mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n - \theta\right)^2 - 2\left(\frac{n+1}{n}\check{X}_n - \theta\right)\frac{\theta}{n} + \frac{\theta^2}{n^2}\right] \\ &= \left(\frac{n}{n+1}\right)^2 \left(\mathbf{E}\left[\left(\frac{n+1}{n}\check{X}_n - \theta\right)^2\right] - \frac{2\theta}{n}\mathbf{E}\left[\frac{n+1}{n}\check{X}_n - \theta\right] + \mathbf{E}\left[\frac{\theta^2}{n^2}\right]\right) \\ &= \left(\frac{n}{n+1}\right)^2 \left(\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] + \frac{\theta^2}{n^2}\right) \\ &= \left(\frac{n}{n+1}\right)^2 \left(\frac{\theta^2}{n(n+2)} + \frac{\theta^2}{n^2}\right) \\ &= \frac{2\theta^2}{(n+1)(n+2)}.\end{aligned}$$

It follows that both the estimators $\frac{n+1}{n}\check{X}_n$ and $\check{X}_n$ are consistent in mean square.

Problem 2 *Let X be a binomially distributed random variable on a probability space $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ with known number of trials parameter m and unknown success parameter p. An investigator wants to estimate p on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the method of moments. What is the estimator $\hat{p}_n^M$?*
2. *Is $\hat{p}_n^M$ biased? Is $\hat{p}_n^M$ consistent?*
3. *Assume the investigator applies the likelihood method. What is the estimator $\hat{p}_n^{ML}$?*
4. *Given that $m = 10$ and we observe a realization $4, 4, 3, 5, 6$ of a sample $X_1, \dots, X_5$ of size 5 drawn from X what is the estimate of p by the estimators $\hat{p}_n^M$?*
5. *Can you give an estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3?*

Solution.

1. Write $\mu_1' : (0,1) \to \mathbb{R}$ for the first order population raw moment. We have

$$\mu_1'(p) = \mathbf{E}[X] = mp$$

Replacing p with the estimator $\hat{p}_n^M$, and equating the first order population raw moment to the first order sample moment $\bar{X}_n$, we can write

$$m\hat{p}_n^M = \bar{X}_n.$$

It follows

$$\hat{p}_n^M = \frac{1}{m}\bar{X}_n.$$

2. We have

$$\mathbf{E}\left[\hat{p}_n^M\right] = \mathbf{E}\left[\frac{1}{m}\bar{X}_n\right] = \frac{1}{m}\mathbf{E}\left[\bar{X}_n\right] = \frac{1}{m}\mathbf{E}[X] = p.$$

Hence, the estimator $\hat{p}_n^M$ is unbiased. In addition, we have

$$\begin{aligned}\mathbf{E}\left[\left(\hat{p}_n^M - p\right)^2\right] &= \mathbf{E}\left[\left(\hat{p}_n^M - \mathbf{E}\left[\hat{p}_n^M\right]\right)^2\right] = \mathbf{D}^2\left[\hat{p}_n^M\right] \\ &= \mathbf{D}^2\left[\frac{1}{m}\bar{X}_n\right] = \frac{1}{m^2}\mathbf{D}^2\left[\bar{X}_n\right] = \frac{1}{m^2}\frac{1}{n}\mathbf{D}^2[X] \\ &= \frac{1}{nm^2}mp(1-p) = \frac{p(1-p)}{nm}.\end{aligned}$$

It follows,

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{p}_n^M - p\right)^2\right] = \lim_{n\to\infty}\frac{p(1-p)}{nm} = 0$$

foe every $p \in (0,1)$. This means that

$$\hat{p}_n^M \xrightarrow{\mathbf{L}^2} p.$$

That is the estimator $\hat{p}_n^M$ is mean square consistent. A fortiori $\hat{p}_n^M$ is probability consistent.

3. The density function $f_X : \mathbb{N}_0 \times (0,1) \to \mathbb{R}_+$ of a binomial random variable with known number of trials parameter m and unknown success parameter p can be written as

$$f_X(x;p) = \frac{m!}{(m-x)!x!}p^x(1-p)^{m-x} \cdot 1_{\{0,1,\ldots,m\}}(x),$$

for every $x \in \mathbb{N}_0$ and $p \in (0,1)$. Let $X_1,\ldots,X_n$ be a simple random sample of size n drawn from X. Then the likelihood function $\mathcal{L}_{X_1,\ldots,X_n} : (0,1) \times \mathbb{N}_0^n \to \mathbb{R}$ of the sample $X_1,\ldots,X_n$ is given by

$$\begin{aligned}&\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \\ &= \prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}p^{x_k}(1-p)^{m-x_k} \cdot 1_{\{0,1,\ldots,m\}}(x_k) \\ &= \left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k} 1_{\{0,1,\ldots,m\}^n}(x_1,\ldots,x_n)\end{aligned}$$

for every $p \in (0,1)$ and every realization $(x_1,\ldots,x_n) \in \mathbb{N}_0^n$ of the sample $X_1,\ldots,X_n$. Note that

$$\begin{aligned}&\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \\ &= \begin{cases}\left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k} > 0, & \text{if } (x_1,\ldots,x_n) \in \{0,1,\ldots,m\}^n, \\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Therefore,

$$\underset{p\in\mathbb{R}_{++}}{\arg\max}\,\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) = \underset{p\in\mathbb{R}_{++}}{\arg\max}\left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k}$$

Hence, we can consider as the log-likelihood function of the sample $X_1,\ldots,X_n$ the function $\log\mathcal{L}_{X_1,\ldots,X_n} : (0,1) \times \{0,1,\ldots,m\}^n \to \mathbb{R}$ given by

$$\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \overset{\text{def}}{=} \ln\left(\left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right)p^{\sum_{k=1}^n x_k}(1-p)^{n\cdot m-\sum_{k=1}^n x_k}\right),$$
$$\forall (p;x_1,\ldots,x_n) \in (0,1) \times \{0,1,\ldots,m\}^n.$$

That is

$$\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) = \sum_{k=1}^n \ln\left(\frac{m!}{(m-x_k)!x_k!}\right) + \left(\sum_{k=1}^n x_k\right)\ln(p) + \left(n\cdot m - \sum_{k=1}^n x_k\right)\ln(1-p).$$

To determine $\underset{p\in(0,1)}{\arg\max}\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n)$, we consider the first order condition

$$\frac{d}{dp}\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) = 0,$$

which yields

$$\left(\sum_{k=1}^n x_k\right)\frac{1}{p} - \left(n\cdot m - \sum_{k=1}^n x_k\right)\frac{1}{1-p} = 0.$$

On account that $p \in (0,1)$, the latter becomes

$$\left(\sum_{k=1}^n x_k\right)(1-p) - \left(n\cdot m - \sum_{k=1}^n x_k\right)p = 0.$$

That is

$$\sum_{k=1}^n x_k - n\cdot m\cdot p = 0,$$

which implies

$$p = \frac{\sum_{k=1}^n x_k}{n\cdot m} = \frac{\bar{x}_n}{m}.$$

In addition,

$$\begin{aligned}\frac{d^2}{dp^2}\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) &= -\left(\sum_{k=1}^n x_k\right)\frac{1}{p^2} - \left(n\cdot m - \sum_{k=1}^n x_k\right)\frac{1}{(1-p)^2} \\ &= \frac{-\left(\sum_{k=1}^n x_k\right)(1-p)^2 - \left(n\cdot m - \sum_{k=1}^n x_k\right)p^2}{p^2(1-p)^2} \\ &= \frac{-\sum_{k=1}^n x_k + 2\left(\sum_{k=1}^n x_k\right)p - n\cdot m\cdot p^2}{p^2(1-p)^2} \\ &= \frac{-n\bar{x}_n + 2n\bar{x}_n p - n\cdot m\cdot p^2}{p^2(1-p)^2} \\ &= -\frac{n}{p^2(1-p)^2}\left(\bar{x}_n - 2\bar{x}_n p + m\cdot p^2\right).\end{aligned}$$

Now, we have

$$\left(\bar{x}_n - 2\bar{x}_n p + m\cdot p^2\right)_{p=\frac{\bar{x}_n}{m}} = \left(\bar{x}_n - \frac{2}{m}\bar{x}_n^2 + m\cdot\frac{\bar{x}_n^2}{m^2}\right) = \bar{x}_n\left(1 - \frac{1}{m}\bar{x}_n\right).$$

On the other hand, we clearly have

$$\bar{x}_n \leq m,$$

for every $(x_1,\ldots,x_n) \in \{0,1,\ldots,m\}^n$. It follows

$$\frac{d^2}{dp^2}\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n) \leq 0$$

which implies that

$$\frac{\bar{x}_n}{m} = \underset{p\in(0,1)}{\arg\max}\log\mathcal{L}_{X_1,\ldots,X_n}(p;x_1,\ldots,x_n).$$

As a consequence, we obtain that the maximum likelihood estimator fo p is given by

$$\hat{p}_n^{ML} = \frac{\bar{X}_n}{m}.$$

4. Given that $m = 10$ and we observe a realization $4, 4, 3, 5, 6$ of a sample $X_1, \dots, X_5$ of size 5 drawn from X, we obtain
$$\hat{p}_n^{M}(\omega) = \frac{\bar{X}_5(\omega)}{10} = \frac{\frac{1}{5}(4+4+3+5+6)}{10} = 0.44.$$

5. We know that
$$\mathbf{E}[X] = m \cdot p \qquad \text{and} \qquad \mathbf{D}^2[X] = m \cdot p\,(1-p)\,,$$
where p is the true value of the success parameter, Hence, an estimator $\hat{\mu}_n^M$ [resp. $\hat{\sigma}_n^{2^M}$] of the expectation [resp. variance] of X build from $\hat{p}_n^M$ is given by
$$\hat{\mu}_n^M = m \cdot \hat{p}_n^M \quad \text{and} \quad \hat{\sigma}_n^{2^M} = m \cdot \hat{p}_n^M \left(1 - \hat{p}_n^M\right).$$
An estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3. is the given by
$$\hat{\mu}_X^M(\omega) = m \cdot \hat{p}_n^M(\omega) = 10 \cdot 0.44 = 4.4$$
and
$$\hat{\sigma}_n^{2^M}(\omega) = m \cdot \hat{p}_n^M(\omega)\left(1 - \hat{p}_n^M(\omega)\right) = 10 \cdot 0.44 \cdot (1 - 0.44) = 2.464.$$
This completes the solution.

Problem 3 *Let X be a normally distributed random variable on a probability space $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ with unknown mean μ_X and variance σ_X^2. An investigator wants to estimate μ and σ^2 on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the likelihood methods. What are the estimator $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$?*

2. *Assume the investigator applies the method of moments. What are the estimators $\hat{\mu}_n^M$ and $\hat{\sigma}_n^{2^M}$? Hint: guess what $\hat{\sigma}_n^{2^M}$ could be and get it!*

3. *Are the estimators $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$ unbiased? Are the estimators $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$ consistent in probability? Are they consistent in mean square?*

Solution.

1. We know that the joint density function $f_{X_1,\dots,X_n} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ of the sample $X_1, \dots, X_n$ is given by
$$f_{X_1,\dots,X_n}(x_1, \dots, x_n; \mu, \sigma) \stackrel{\text{def}}{=} \prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_k-\mu)^2}{2\sigma^2}}, \quad \forall (x_1, \dots, x_n; \mu, \sigma) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++}.$$
That is
$$f_{X_1,\dots,X_n}(x_1, \dots, x_n; \mu, \sigma) = \frac{1}{\sqrt{2^n \pi^n}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k-\mu)^2}.$$
Hence, the likelihood function $\mathcal{L}_{X_1,\dots,X_n} : \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$ of the sample is given by
$$\mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = \frac{1}{\sqrt{2^n \pi^n}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k-\mu)^2}, \quad \forall (\mu, \sigma; x_1, \dots, x_n) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n.$$
Thanks to the stucture of $\mathcal{L}_{X_1,\dots,X_n}$ it is convenient to consider the log-likelihood function $\log \mathcal{L}_{X_1,\dots,X_n} : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ of the sample which is given by
$$\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) \stackrel{\text{def}}{=} (\log \circ \mathcal{L}_{X_1,\dots,X_n})(\mu, \sigma; x_1, \dots, x_n), \quad \forall (\mu, \sigma; x_1, \dots, x_n) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n.$$

That is
$$\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = -n\left(\frac{1}{2}\ln(2\pi) + \ln(\sigma)\right) - \frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \mu)^2.$$
Now, to determine $\underset{(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\dots,X_n}$ we consider the first order conditions
$$\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = 0 \quad \text{and} \quad \frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = 0.$$
We have
$$\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = \frac{1}{\sigma^2}\sum_{k=1}^n (x_k - \mu)$$
and
$$\frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = \frac{1}{\sigma}\left(\frac{1}{\sigma^2}\sum_{k=1}^n (x_k - \mu)^2 - n\right).$$
Therefore,
$$\frac{\partial}{\partial\mu}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = 0 \Rightarrow \sum_{k=1}^n (x_k - \mu) = 0,$$
$$\Rightarrow \mu = \frac{1}{n}\sum_{k=1}^n x_k \equiv \bar{x}_n$$
and
$$\frac{\partial}{\partial\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = 0 \Rightarrow \frac{1}{\sigma^2}\sum_{k=1}^n (x_k - \mu)^2 - n,$$
$$\Rightarrow \sigma^2 = \frac{1}{n}\sum_{k=1}^n (x_k - \mu)^2 \equiv \tilde{s}_{X,n}^2.$$
In addition,
$$\frac{\partial^2}{\partial\mu^2}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = -\frac{n}{\sigma^2},$$
$$\frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = -2\frac{1}{\sigma^3}\sum_{k=1}^n (x_k - \mu),$$
$$\frac{\partial^2}{\partial\sigma^2}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) = \frac{1}{\sigma^2}\left(-\frac{3}{\sigma^2}\sum_{k=1}^n (x_k - \mu)^2 + n\right).$$
Hence,
$$\begin{aligned}
&J\left(\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} \\
&= \begin{pmatrix} \frac{\partial^2}{\partial\mu^2}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) & \frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) \\ \frac{\partial^2}{\partial\mu\sigma}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) & \frac{\partial^2}{\partial\sigma^2}\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n) \end{pmatrix}_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} \\
&= \begin{pmatrix} -\frac{n}{\sigma^2} & -2\frac{1}{\sigma^3}\sum_{k=1}^n (x_k - \mu) \\ -2\frac{1}{\sigma^3}\sum_{k=1}^n (x_k - \mu) & \frac{1}{\sigma^2}\left(-\frac{3}{\sigma^2}\sum_{k=1}^n (x_k - \mu)^2 + n\right) \end{pmatrix}_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} \\
&= \frac{1}{\tilde{s}_{X,n}^2}\begin{pmatrix} -n & 0 \\ 0 & -2n \end{pmatrix}
\end{aligned}$$
Because,
$$\sum_{k=1}^n (x_k - \mu)\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} = \sum_{k=1}^n x_k - n\mu\Big|_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} = 0$$
and
$$-\frac{3}{\sigma^2}\sum_{k=1}^n (x_k - \mu)^2\Bigg|_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} = -\frac{3n}{\sigma^2}\frac{1}{n}\sum_{k=1}^n (x_k - \mu)^2\Bigg|_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} = -\frac{3n}{\tilde{s}_{X,n}^2}\tilde{s}_{X,n}^2 = -3n.$$
We then have
$$\det J\left(\log \mathcal{L}_{X_1,\dots,X_n}(\mu, \sigma; x_1, \dots, x_n)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n, \tilde{s}_{X,n}^2\right)} = \frac{3n^2}{\tilde{s}_{X,n}^4}$$

and

$$\operatorname{tr} J\left(\log \mathcal{L}_{X_1,\ldots,X_n}\left(\mu,\sigma;x_1,\ldots,x_n\right)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)} = -\frac{3n}{\tilde{s}^4_{X,n}}$$

It follows that the eigenvalues of the Jacobian matrix $J\left(\log \mathcal{L}_{X_1,\ldots,X_n}\left(\mu,\sigma;x_1,\ldots,x_n\right)\right)_{(\mu,\sigma^2)=\left(\bar{x}_n,\tilde{s}^2_{X,n}\right)}$ are stictly negative. This imples that

$$\left(\bar{x}_n,\tilde{s}^2_{X,n}\right) = \underset{(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\ldots,X_n}.$$

As a consequence, we obtain the maximum likelihood estimators

$$\hat{\mu}_n^{LM} = \bar{X}_n \quad \text{and} \quad \hat{\sigma}_n^{2^{LM}} = \tilde{S}^2_{X,n},$$

where $\bar{X}_n$ [resp. $\tilde{S}^2_n(X)$] is the sample mean [resp. unbiased sample variance] of $X_1,\ldots,X_n$.

2. We know that

$$\mathbf{E}[X] = \mu \quad \text{and} \quad \mathbf{E}\left[X^2\right] = \mu^2 + \sigma^2.$$

Hence, applying the method of moments, the investigator writes

$$\frac{1}{n}\sum_{k=1}^n X_k = \hat{\mu}_n^M \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^n X_k^2 = \left(\hat{\mu}_n^M\right)^2 + \hat{\sigma}_n^{2^M}.$$

The first of the two equations clearly yields

$$\hat{\mu}_n^M = \bar{X}_n.$$

The second equation, on account of the first, yields

$$\hat{\sigma}_n^{2^M} = \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2 = \tilde{S}^2_{X,n}.$$

Recall that

$$\begin{aligned}
\tilde{S}^2_{X,n} &= \frac{1}{n}\sum_{k=1}^n\left(X_k - \bar{X}_n\right)^2 = \frac{1}{n}\sum_{k=1}^n\left(X_k^2 - 2X_k\bar{X}_n + \bar{X}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2\bar{X}_n\sum_{k=1}^n X_k + \sum_{k=1}^n \bar{X}_n^2\right) = \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - n\bar{X}_n^2\right) = \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2.
\end{aligned}$$

3. It is well known that the estimator $\hat{\mu}_n^{ML} = \bar{X}_n$ [resp. $\hat{\sigma}_n^{2^{ML}} = \tilde{S}^2_{X,n}$] is unbiased [resp. biased]. In addition, since

$$\mathbf{D}^2\left[\hat{\mu}_n^{ML}\right] = \mathbf{D}^2\left[\bar{X}_n\right] = \frac{1}{n}\mathbf{D}^2[X] = \frac{1}{n}\sigma^2$$

we clearly have

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\mu}_n^{ML} - \mu\right)^2\right] = \lim_{n\to\infty}\mathbf{D}^2\left[\hat{\mu}_n^{ML}\right] = \lim_{n\to\infty}\frac{1}{n}\sigma^2 = 0$$

which means that $\hat{\mu}_n^{ML}$ is consistent in mean square. With regard to the biased estimator $\hat{\sigma}_n^{2^{ML}}$, observe that we can write

$$\begin{aligned}
\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \sigma\right)^2\right] &= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \left(\frac{n-1}{n}\sigma^2 + \frac{1}{n}\sigma^2\right)\right)^2\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2 - \frac{1}{n}\sigma^2\right)^2\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right)^2 - \frac{1}{n}\sigma^2\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right) + \frac{1}{n}\sigma^4\right] \\
&= \mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right)^2\right] - \frac{1}{n}\sigma^2\mathbf{E}\left[\hat{\sigma}_n^{2^{ML}} - \frac{n-1}{n}\sigma^2\right] + \frac{1}{n}\sigma^2 \\
&= \mathbf{E}\left[\left(\tilde{S}^2_{X,n} - \mathbf{E}\left[\tilde{S}^2_{X,n}\right]\right)^2\right] - \frac{1}{n}\sigma^2\mathbf{E}\left[\tilde{S}^2_{X,n} - \mathbf{E}\left[\tilde{S}^2_{X,n}\right]\right] + \frac{1}{n}\sigma^2 \\
&= \mathbf{D}^2\left[\tilde{S}^2_{X,n}\right] + \frac{1}{n}\sigma^2.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbf{D}^2\left[\tilde{S}^2_{X,n}\right] &= \mathbf{D}^2\left[\frac{n}{n+1}S^2_{X,n}\right] = \frac{n^2}{(n+1)^2}\mathbf{D}^2\left[S^2_{X,n}\right] \\
&= \frac{n^2}{(n+1)^2}\frac{\sigma^4}{n}\left(3 - \frac{n-3}{n-1}\right) = \frac{n^2}{(n+1)^2}\frac{\sigma^4}{n}\frac{2n}{n+1} \\
&= \frac{2n^2}{(n+1)^3}\sigma^4.
\end{aligned}$$

Therefore,

$$\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \sigma\right)^2\right] = \frac{2n^2}{(n+1)^3}\sigma^4 + \frac{1}{n}\sigma^2$$

It folows

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\sigma}_n^{2^{ML}} - \sigma\right)^2\right] = \lim_{n\to\infty}\left(\frac{2n^2}{(n+1)^3}\sigma^4 + \frac{1}{n}\sigma^2\right) = 0,$$

which means that also $\hat{\sigma}_n^{2^{ML}}$ is consistent in mean square. A fortiori, both $\hat{\mu}_n^{ML}$ and $\hat{\sigma}_n^{2^{ML}}$ are consistent in probability.

Problem 4 *Let $\theta > 0$ and let X be an uniformly distributed real random variable on the interval $[0,\theta]$. In symbols $X \sim Unif(0,\theta)$.*

1. *Write the joint likelihood function of a simple random sample $X_1,\ldots,X_n$ of size n drawn from X and determine $\hat{\theta}_n^{ML}$.*

2. *Check whether the MLE is unbiased or biased.*

3. *Determine $\hat{\theta}_n^M$, check that $\hat{\theta}_n^M$ is unbiased and consistent.*

Solution.

1. The density function $f_X : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$, depending of the parameter θ, is given by

$$f_X(x;\theta) \overset{\text{def}}{=} \frac{1}{\theta}1_{[0,\theta]}(x), \quad \forall x\in\mathbb{R}, \quad \forall\theta\in\mathbb{R}_+.$$

Therefore, the sample likelihood $\mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R}_+\times\mathbb{R}^n \to \mathbb{R}$ can be written as

$$\mathcal{L}_{X_1,\ldots,X_n}(\theta;x_1,\ldots,x_n) \overset{\text{def}}{=} \prod_{k=1}^n \frac{1}{\theta}1_{[0,\theta]}(x_k), \quad \forall\theta\in\mathbb{R}_+, \quad \forall x_1,\ldots,x_n\in\mathbb{R}.$$

That is to say

$$\mathcal{L}_{X_1,\ldots,X_n}(\theta;x_1,\ldots,x_n) = \frac{1}{\theta^n}\prod_{k=1}^n 1_{[0,\theta]}(x_k).$$

Now, to enhance the role of θ as a variable, note that

$$1_{[0,\theta]}(x_k) = 1_{\mathbb{R}_+}(x_k)\,1_{[x_n,+\infty)}(\theta)$$

for all $\theta\in\mathbb{R}_+$ and all $x_1,\ldots,x_n\in\mathbb{R}$. Hence, it may be convenient to write the joint likelihood in the form

$$\mathcal{L}_{X_1,\ldots,X_n}(\theta;x_1,\ldots,x_n) = \frac{1}{\theta^n}\prod_{k=1}^n 1_{\mathbb{R}_+}(x_k)\,1_{[x_k,+\infty)}(\theta) = \frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_k,+\infty)}(\theta)\prod_{k=1}^n 1_{\mathbb{R}_+}(x_k).$$

Given any realization $x_1,\ldots,x_n$ of the random sample $X_1,\ldots,X_n$, it follows that

$$\mathcal{L}_{X_1,\ldots,X_n}(\theta;x_1,\ldots,x_n) = \begin{cases} 0 & \text{if } x_k < 0, \quad \exists k\in\{1,\ldots,n\} \\ \frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_k,+\infty)}(\theta) & \text{if } x_k \geq 0, \quad \forall k\in\{1,\ldots,n\} \end{cases}.$$

Therefore, under the condition $x_k \geq 0$ for every $k \in \{1, \dots, n\}$, we have

$$\underset{\theta\in\mathbb{R}_+}{\arg\max}\mathcal{L}_{X_1,\dots,X_n}(\theta; x_1, \dots, x_n) = \underset{\theta\in\mathbb{R}_+}{\arg\max}\frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_k,+\infty)}(\theta) = \max\{x_1, \dots, x_n\}.$$

In fact,

$$\prod_{k=1}^n 1_{[x_n,+\infty)}(\theta) = \begin{cases} 0 & \text{if } \theta < x_k, \quad \exists k \in \{1, \dots, n\} \\ 1 & \text{if } \theta \geq x_k, \quad \forall k \in \{1, \dots, n\} \end{cases}.$$

Hence, $\frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_n,+\infty)}(\theta)$ attains its maximum for $\theta \geq x_k$, for every $k = 1, \dots, n$. That is

$$\max\{x_1, \dots, x_n\} \leq \underset{\theta\in\mathbb{R}_+}{\arg\max}\frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_k,+\infty)}(\theta).$$

In addition,

$$\frac{1}{\theta^n}\prod_{k=1}^n 1_{[x_k,+\infty)}(\theta) = \begin{cases} \frac{1}{\max\{x_1,\dots,x_n\}^n} & \text{if } \theta = \max\{x_1, \dots, x_n\} \\ \frac{1}{\theta^n} < \frac{1}{\max\{x_1,\dots,x_n\}^n} & \text{if } \theta > \max\{x_1, \dots, x_n\} \end{cases}.$$

In the end, since $\mathbf{P}(X < 0) = 0$ implies $\mathbf{P}(X_k \geq 0) = 1$ for every $k \in \{1, \dots, n\}$, we obtain

$$\hat{\theta}_n^{ML} = \max\{X_1, \dots, X_n\}.$$

2. To check whether $\hat{\theta}_{MLE}$ is unbiased or biased we need check whether

$$\mathbf{E}\left[\hat{\theta}_n^{ML}\right] = \theta$$

or not. Write $\check{X}_n \equiv \max\{X_1, \dots, X_n\} \equiv \hat{\theta}_n^{ML}$. We will be able to compute $\mathbf{E}\left[\check{X}_n\right]$ if we determine the distribution function $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ of $\check{X}_n$. We have

$$F_{\check{X}_n}(x) = \mathbf{P}\left(\check{X}_n \leq x\right) = \mathbf{P}(X_1 \leq x, \dots, X_n \leq x) = \prod_{k=1}^n \mathbf{P}(X_k \leq x) = \mathbf{P}(X \leq x)^n,$$

for every $x \in \mathbb{R}$. On the other hand,

$$\mathbf{P}(X \leq x) = F_X(x) = \frac{x}{\theta}1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

Therefore,

$$F_{\check{X}_n}(x) = \frac{x^n}{\theta^n}1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

It follows that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ given by

$$f_{\check{X}_n}(x) = n\frac{x^{n-1}}{\theta^n}1_{[0,\theta]}(x),$$

for every $x \in \mathbb{R}$. As a consequence, we can write

$$\begin{aligned}\mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} n\frac{x^n}{\theta^n}1_{[0,\theta]}(x)\ d\mu_L(x) \\ &= \frac{n}{\theta^n}\int_{[0,\theta]} x^n\ d\mu_L(x) = \frac{n}{\theta^n}\int_0^\theta x^n\ dx = \frac{n}{\theta^n}\frac{1}{n+1}\ x^{n+1}\Big|_0^\theta \\ &= \frac{n}{n+1}\theta.\end{aligned}$$

This proves that $\hat{\theta}_n^{ML}$ is a biased estimator of θ.

3. The first population moment and the first sample moment are given by

$$\mathbf{E}[X] = \frac{\theta}{2} \quad \text{and} \quad \bar{X}_n \equiv \frac{1}{n}\sum_{k=1}^n X_k,$$

respectively. Equating

$$\frac{\theta}{2} = \bar{X}_n,$$

it follows that

$$\hat{\theta}_n^M = 2\bar{X}_n.$$

A straightforward computation yields

$$\mathbf{E}\left[\hat{\theta}_n^M\right] = \mathbf{E}\left[2\bar{X}_n\right] = 2\mathbf{E}\left[\bar{X}_n\right] = 2\mathbf{E}\left[\frac{1}{n}\sum_{k=1}^n X_k\right] = \frac{2}{n}\sum_{k=1}^n \mathbf{E}[X_k] = \frac{2}{n}\sum_{k=1}^n \mathbf{E}[X] = 2\mathbf{E}[X] = \theta,$$

which shows that $\hat{\theta}_n^M$ is unbiased. In addition, from Remark **??** we know that

$$\bar{X}_n \xrightarrow{\mathbf{P}} \mathbf{E}[X].$$

This implies (see Theorem **??**)

$$2\bar{X}_n \xrightarrow{\mathbf{P}} 2\mathbf{E}[X] = \theta,$$

which shows that $\hat{\theta}_n^M$ is consistent.

Problem 5 *Let $\theta > 0$ and let X be an absolutely continuous real random variable with density function $f_X : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_X(x) \stackrel{def}{=} \theta x^{\theta-1}1_{[0,1]}(x), \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to determine the estimator $\hat{\theta}_n^M$ for θ.*
2. *Check whether $\hat{\theta}_n^M$ is unbiased, consistent in probability, and consistent in mean square.*
3. *Apply the method of maximum likelihood to determine the estimator $\hat{\theta}_n^{ML}$ for θ.*
4. *Check whether $\hat{\theta}_n^{ML}$ is unbiased, consistent in probability, and consistent in mean square.*
5. *Use the estimators obtained to build estimators of the mean μ_X and the variance σ_X^2 of the random variable X.*

Solution.

1. Note that

$$\begin{aligned}\mathbf{P}(0 < X < 1) &= \int_{(0,1)} f_X(x;\theta)\,d\mu_L(x) = \int_{(0,1)} \theta x^{\theta-1}1_{[0,1]}(x)\,d\mu_L(x) \\ &= \int_{(0,1)} \theta x^{\theta-1} d\mu_L(x) = \int_0^1 \theta x^{\theta-1}dx = \theta\int_0^1 x^{\theta-1}dx \\ &= \theta\ \frac{x^\theta}{\theta}\Big|_0^1 \\ &= 1.\end{aligned}$$

Hence, considering a simple random sample $X_1, \dots, X_n$ of size n drawn from X, we have

$$\mathbf{P}(0 < X_k < 1) = 1$$

for every $k = 1, \dots, n$. Moreover, we clearly have

$$\bigcap_{k=1}^n \{0 < X_k < 1\} \subseteq \left\{0 < \bar{X}_n < 1\right\},$$

which implies

$$\mathbf{P}\left(\bigcap_{k=1}^n \{0 < X_k < 1\}\right) \leq \mathbf{P}\left(0 < \bar{X}_n < 1\right).$$

On the other hand,

$$\mathbf{P}\left(\bigcap_{k=1}^{n}\{0<X_k<1\}\right)=\prod_{k=1}^{n}\mathbf{P}\left(0<X_k<1\right)=1$$

It follows

$$\mathbf{P}\left(0<\bar{X}_n<1\right)=1.$$

Now, we have

$$\begin{aligned}\mathbf{E}[X]&=\int_{\mathbb{R}}xf_X(x;\theta)\,d\mu_L(x)=\int_{\mathbb{R}}\theta x^{\theta}1_{[0,1]}(x)\,d\mu_L(x)\\&=\int_{[0,1]}\theta x^{\theta}d\mu_L(x)=\theta\int_0^1x^{\theta}dx=\frac{\theta}{\theta+1}x^{\theta+1}\Big|_0^1\\&=\frac{\theta}{1+\theta}\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\left[X^2\right]&=\int_{\mathbb{R}}x^2f_X(x;\theta)\,d\mu_L(x)=\int_{\mathbb{R}}\theta x^{\theta+1}1_{[0,1]}(x)\,d\mu_L(x)\,d\\&=\int_{[0,1]}\theta x^{\theta+1}d\mu_L(x)=\theta\int_0^1x^{\theta+1}dx=\frac{\theta}{\theta+2}x^{\theta+2}\Big|_0^1\\&=\frac{\theta}{2+\theta}\end{aligned}$$

It follows,

$$\mathbf{D}^2[X]=\mathbf{E}\left[X^2\right]-\mathbf{E}[X]^2=\frac{\theta}{2+\theta}-\frac{\theta^2}{(1+\theta)^2}=\frac{\theta}{(1+\theta)^2(2+\theta)}$$

As a consequence,

$$\mathbf{E}\left[\bar{X}_n\right]=\mathbf{E}[X]=\frac{\theta}{1+\theta}\quad\text{and}\quad\mathbf{D}^2\left[\bar{X}_n\right]=\frac{1}{n}\mathbf{D}^2[X]=\frac{\theta}{n(1+\theta)^2(2+\theta)}.$$

The estimator $\hat{\theta}_n^M$ for θ is then obtained by solving the equation

$$\frac{\hat{\theta}_n^M}{1+\hat{\theta}_n^M}=\bar{X}_n,$$

which yields

$$\hat{\theta}_n^M=\frac{\bar{X}_n}{1-\bar{X}_n}.$$

Note that, since $\mathbf{P}\left(0<\bar{X}_n<1\right)=1$, the estimator $\hat{\theta}_n^M$ is well defined.

2. To check the properties of the estimator $\hat{\theta}_n^M$, we apply the so called delta method. Considering the function $f:(0,1)\to\mathbb{R}$ given by

$$f(x)\overset{\text{def}}{=}\frac{x}{1-x},\quad\forall x\in(0,1)$$

by virtue of the Taylor formula, fixed any $x_0\in(0,1)$, we can write

$$f(x)\approx f(x_0)+f'(x_0)(x-x_0),$$

for any $x\in(0,1)$, where

$$f'(x_0)=\frac{1}{(1-x_0)^2}.$$

On the other hand, we have

$$\hat{\theta}_n^M=\frac{\bar{X}_n}{1-\bar{X}_n}\equiv f\left(\bar{X}_n\right).$$

Hence, setting

$$\mu_{\bar{X}_n}\equiv\mathbf{E}\left[\bar{X}_n\right]=\mathbf{E}[X]=\frac{\theta}{1+\theta},$$

the Taylor formula yields

$$f\left(\bar{X}_n\right)\approx f\left(\mu_{\bar{X}_n}\right)+f'\left(\mu_{\bar{X}_n}\right)\left(\bar{X}_n-\mu_{\bar{X}_n}\right),$$

where

$$f\left(\mu_{\bar{X}_n}\right)=\frac{\mu_{\bar{X}_n}}{1-\mu_{\bar{X}_n}}=\frac{\frac{\theta}{1+\theta}}{1-\frac{\theta}{1+\theta}}=\theta$$

and

$$f'\left(\mu_{\bar{X}_n}\right)=\frac{1}{\left(1-\mu_{\bar{X}_n}\right)^2}=\frac{1}{\left(1-\frac{\theta}{1+\theta}\right)^2}=(1+\theta)^2.$$

We then obtain

$$\hat{\theta}_n^M\approx\theta+(1+\theta)^2\left(\bar{X}_n-\frac{\theta}{1+\theta}\right).$$

It follows

$$\mathbf{E}\left[\hat{\theta}_n^M\right]\approx\theta+(1+\theta)^2\left(\mathbf{E}\left[\bar{X}_n\right]-\frac{\theta}{1+\theta}\right)=\theta$$

and

$$\mathbf{D}^2\left[\hat{\theta}_n^M\right]\approx\mathbf{D}^2\left[f\left(\mu_{\bar{X}_n}\right)+f'\left(\mu_{\bar{X}_n}\right)\left(\bar{X}_n-\mu_{\bar{X}_n}\right)\right]=f'\left(\mu_{\bar{X}_n}\right)^2\mathbf{D}^2\left[\bar{X}_n\right]=\frac{\theta(\theta+1)^2}{n(\theta+2)}.$$

As a consequence, we have that the estimator $\hat{\theta}_n^M$ is approximatively unbiased. In addition, since

$$\bar{X}_n\xrightarrow{\mathbf{P}}\mathbf{E}\left[\bar{X}_n\right]=\frac{\theta}{1+\theta},\quad\text{and}\quad\mathbf{P}\left(0<\bar{X}_n<1\right)=1$$

we have that

$$1-\bar{X}_n\xrightarrow{\mathbf{P}}\frac{1}{1+\theta}\quad\text{and}\quad\frac{\bar{X}_n}{1-\bar{X}_n}\xrightarrow{\mathbf{P}}\frac{\frac{\theta}{1+\theta}}{\frac{1}{1+\theta}}=\theta.$$

Thus, the estimator $\hat{\theta}_n^M$ is consistent in probability. We also have

$$\mathbf{E}\left[\left(\hat{\theta}_n^M-\theta\right)^2\right]\approx\mathbf{E}\left[\left(\hat{\theta}_n^M-\mathbf{E}\left[\hat{\theta}_n^M\right]\right)^2\right]=\mathbf{D}^2\left[\hat{\theta}_n^M\right]\approx\frac{\theta(\theta+1)^2}{n(\theta+2)}.$$

It follows

$$\lim_{n\to\infty}\mathbf{E}\left[\left(\hat{\theta}_n^M-\theta\right)^2\right]=0,$$

that is the estimator $\hat{\theta}_n^M$ is consistent in mean square.

3. To determine the estimator $\hat{\theta}_n^{ML}$ we start by building the likelihood function of the sample $X_1,\dots,X_n$. Writing $f_{X_1,\dots,X_1}:\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$ for the density function of the sample $X_1,\dots,X_n$, we clearly have

$$\begin{aligned}f_{X_1,\dots,X_1}(x_1,\dots,x_n;\theta)&=\prod_{k=1}^{n}\theta x_k^{\theta-1}1_{[0,1]}(x_k)=\theta^n\prod_{k=1}^{n}x_k^{\theta-1}\prod_{k=1}^{n}1_{[0,1]}(x_k)\\&=\theta^n\left(\prod_{k=1}^{n}x_k^{\theta-1}\right)1_{[0,1]\times\cdots\times[0,1]}(x_1,\dots,x_n),\end{aligned}$$

for every $(x_1,\dots,x_n;\theta)\in\mathbb{R}^n\times\mathbb{R}_+$. Hence, the likelihood function $\mathcal{L}_{X_1,\dots,X_n}:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}$ is given by

$$\mathcal{L}_{X_1,\dots,X_n}(\theta,x_1,\dots,x_n)=\theta^n\left(\prod_{k=1}^{n}x_k^{\theta-1}\right)1_{[0,1]\times\cdots\times[0,1]}(x_1,\dots,x_n),$$

for every $(\theta;x_1,\dots,x_n)\in\mathbb{R}^n\times\mathbb{R}_+$. Observing that we have

$$\mathcal{L}_{X_1,\dots,X_n}(\theta,x_1,\dots,x_n)>0\qquad[\text{resp. }\mathcal{L}_{X_1,\dots,X_n}(\theta,x_1,\dots,x_n)=0]$$

for every $(x_1, \ldots, x_n) \in (0,1] \times \cdots \times (0,1]$ [resp. $(x_1, \ldots, x_n) \notin (0,1] \times \cdots \times (0,1]$] we have

$$\underset{\theta \in \mathbb{R}_+}{\arg\max} \mathcal{L}_{X_1,\ldots,X_n}(\theta, x_1, \ldots, x_n) = \underset{\theta \in \mathbb{R}_+}{\arg\max} \theta^n \left(\prod_{k=1}^n x_k^{\theta-1} \right).$$

Therefore, we rather consider

$$\underset{\theta \in \mathbb{R}_+}{\arg\max} \ln \left(\theta^n \left(\prod_{k=1}^n x_k^{\theta-1} \right) \right) = \underset{\theta \in \mathbb{R}_+}{\arg\max} n \ln(\theta) + (\theta - 1) \sum_{k=1}^n \log(x_k)$$

where $(x_1, \ldots, x_n) \in (0,1] \times \cdots \times (0,1]$. Applying the first order condition to the function to be maximized, we obtain

$$\frac{n}{\theta} + \sum_{k=1}^n \log(x_k) = 0$$

which yields

$$\theta = -\frac{n}{\sum_{k=1}^n \log(x_k)}.$$

In addition the second order derivative of the function to be maximized is

$$-\frac{n}{\theta^2} < 0.$$

As a consequence, we can write

$$\underset{\theta \in \mathbb{R}_+}{\arg\max} \mathcal{L}_{X_1,\ldots,X_n}(\theta, x_1, \ldots, x_n) = -\frac{n}{\sum_{k=1}^n \log(x_k)}.$$

It follows

$$\hat{\theta}_n^{ML} = -\frac{n}{\sum_{k=1}^n \log(X_k)}.$$

4. We can also write

$$\hat{\theta}_n^{ML} = -\frac{n}{\log\left(\prod_{k=1}^n X_k \right)}$$

Therefore, considering the function $f : (0,1) \to \mathbb{R}$ given by

$$f(x) \overset{\text{def}}{=} -\frac{n}{\log(x)}, \quad \forall x \in (0,1),$$

by virtue of the Taylor formula, fixed any $x_0 \in (0,1)$, we can write

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0),$$

for any $x \in (0,1)$, where

$$f'(x_0) = -\frac{n}{x_0 \log^2(x_0)}.$$

On the other hand, we have

$$\hat{\theta}_n^{ML} = -\frac{n}{\log\left(\prod_{k=1}^n X_k \right)} \equiv f\left(\prod_{k=1}^n X_k \right)$$

Hence, setting

$$\mu_{\Pi_{k=1}^n X_k} \equiv \mathbf{E}\left[\prod_{k=1}^n X_k \right] = \prod_{k=1}^n \mathbf{E}[X_k] = \prod_{k=1}^n \mathbf{E}[X] = \frac{\theta^n}{(1+\theta)^n},$$

the Taylor formula yields

$$f\left(\prod_{k=1}^n X_k \right) \approx f\left(\mu_{\Pi_{k=1}^n X_k} \right) + f'\left(\mu_{\Pi_{k=1}^n X_k} \right) \left(\prod_{k=1}^n X_k - \mu_{\Pi_{k=1}^n X_k} \right),$$

where

$$f\left(\mu_{\Pi_{k=1}^n X_k} \right) = -\frac{n}{\log\left(\frac{\theta^n}{(1+\theta)^n} \right)} = -\frac{n}{n\left(\log(\theta) - \log(1+\theta) \right)} = \frac{1}{\log(1+\theta) - \log(\theta)}$$

and

$$f'\left(\mu_{\Pi_{k=1}^n X_k} \right) = -\frac{n}{\frac{\theta^n}{(1+\theta)^n} \log^2\left(\frac{\theta^n}{(1+\theta)^n} \right)} = -\frac{n(1+\theta)^n}{\theta^n n^2 \left(\log(1+\theta) - \log(\theta) \right)^2} = -\frac{(1+\theta)^n}{\theta^n n \left(\log(1+\theta) - \log(\theta) \right)^2}.$$

It follows

$$\begin{aligned} \mathbf{E}\left[\hat{\theta}_n^{ML} \right] &\approx f\left(\mu_{\Pi_{k=1}^n X_k} \right) + f'\left(\mu_{\Pi_{k=1}^n X_k} \right) \left(\mathbf{E}\left[\prod_{k=1}^n X_k \right] - \mu_{\Pi_{k=1}^n X_k} \right) \\ &= \frac{1}{\log(1+\theta) - \log(\theta)} = \frac{1}{\log\left(\frac{1+\theta}{\theta} \right)} = \frac{1}{\log\left(1 + \frac{1}{\theta} \right)} \approx \frac{1}{\frac{1}{\theta}} \\ &= \theta, \end{aligned}$$

for large θ. The stimator $\hat{\theta}_n^{ML}$ is approximatively unbiased for large θ

$$\begin{aligned} \mathbf{D}^2\left[\hat{\theta}_n^{ML} \right] &\approx f'\left(\mu_{\Pi_{k=1}^n X_k} \right)^2 \mathbf{D}^2\left[\prod_{k=1}^n X_k \right] \\ &= \frac{(1+\theta)^{2n}}{\theta^{2n} n^2 \left(\log(1+\theta) - \log(\theta) \right)^4} \theta^n \frac{(\theta+1)^{4n} - \theta^{3n} (\theta+2)^n}{(\theta+2)^n (\theta+1)^{4n}} \\ &= \frac{1}{n^2} \frac{(1+\theta)^{4n} - \theta^{3n} (2+\theta)^n}{\theta^n (1+\theta)^{2n} (2+\theta)^n \left(\log(1+\theta) - \log(\theta) \right)^2} \end{aligned}$$

In fact,

$$\begin{aligned} \mathbf{D}^2\left[\prod_{k=1}^n X_k \right] &= \mathbf{E}\left[\left(\prod_{k=1}^n X_k \right)^2 \right] - \mathbf{E}\left[\prod_{k=1}^n X_k \right]^2 \\ &= \mathbf{E}\left[X_k^2 \right] - \left(\prod_{k=1}^n \mathbf{E}[X_k] \right)^2 \\ &= \prod_{k=1}^n \mathbf{E}\left[X_k^2 \right] - \left(\prod_{k=1}^n \mathbf{E}[X_k] \right)^2 \\ &= \prod_{k=1}^n \mathbf{E}\left[X^2 \right] - \left(\prod_{k=1}^n \mathbf{E}[X] \right)^2 \\ &= \mathbf{E}\left[X^2 \right]^n - \mathbf{E}[X]^{2n} \\ &= \frac{\theta^n}{(2+\theta)^n} - \frac{\theta^{4n}}{(1+\theta)^{4n}} \\ &= \theta^n \frac{(1+\theta)^{4n} - \theta^{3n} (2+\theta)^n}{(2+\theta)^n (1+\theta)^{4n}} \end{aligned}$$

As a consequence, for large θ

$$\begin{aligned} \mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \theta \right)^2 \right] &\approx \mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \mathbf{E}\left[\hat{\theta}_n^{ML} \right] \right)^2 \right] = \mathbf{D}^2\left[\hat{\theta}_n^{M} \right] \\ &\approx \frac{1}{n^2} \frac{(1+\theta)^{4n} - \theta^{3n} (2+\theta)^n}{\theta^n (1+\theta)^{2n} (2+\theta)^n \left(\log(1+\theta) - \log(\theta) \right)^2}. \end{aligned}$$

It follows

$$\lim_{n \to \infty} \mathbf{E}\left[\left(\hat{\theta}_n^{ML} - \theta \right)^2 \right] = 0,$$

that is the estimator $\hat{\theta}_n^M$ is consistent in mean square. In particular, the estimator $\hat{\theta}_n^M$ is consistent in probability.

5. We have

$$\mathbf{E}[X] = \frac{\theta}{1+\theta} \quad \text{and} \quad \mathbf{D}^2[X] = \frac{\theta}{(1+\theta)^2(2+\theta)}.$$

Hence, writing $\hat{\mu}_{X,n}^M$ [resp. $\hat{\mu}_{X,n}^{ML}$] for an estimator of size n of $\mathbf{E}[X]$ built from $\hat{\theta}_n^M$ [resp. $\hat{\theta}_n^{ML}$], we have

$$\hat{\mu}_{X,n}^M = \frac{\hat{\theta}_n^M}{1+\hat{\theta}_n^M} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{1+\frac{\bar{X}_n}{1-\bar{X}_n}} = \bar{X}_n$$

and

$$\hat{\mu}_{X,n}^{ML} = \frac{\hat{\theta}_n^{ML}}{1+\hat{\theta}_n^{ML}} = \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{1-\frac{n}{\sum_{k=1}^n \log(X_k)}} = \frac{n}{n-\sum_{k=1}^n \log(X_k)}.$$

Similarly, writing $\hat{\sigma}_{X,n}^{2,M}$ [resp. $\hat{\sigma}_{X,n}^{2,ML}$] for an estimator of size n of $\mathbf{D}^2[X]$ built from $\hat{\theta}_n^M$ [resp. $\hat{\theta}_n^{ML}$], we have

$$\hat{\sigma}_{X,n}^{2,M} = \frac{\hat{\theta}_n^M}{\left(1+\hat{\theta}_n^M\right)^2\left(2+\hat{\theta}_n^M\right)} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{\left(1+\frac{\bar{X}_n}{1-\bar{X}_n}\right)^2\left(2+\frac{\bar{X}_n}{1-\bar{X}_n}\right)} = \frac{\frac{\bar{X}_n}{1-\bar{X}_n}}{\frac{1}{(1-\bar{X}_n)^2}\frac{2-\bar{X}_n}{1-\bar{X}_n}} = \frac{\bar{X}_n\left(1-\bar{X}_n\right)^2}{2-\bar{X}_n}$$

and

$$\begin{aligned}\hat{\sigma}_{X,n}^{2,ML} &= \frac{\hat{\theta}_n^{ML}}{\left(1+\hat{\theta}_n^{ML}\right)^2\left(2+\hat{\theta}_n^{ML}\right)} = \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{\left(1-\frac{n}{\sum_{k=1}^n \log(X_k)}\right)^2\left(2-\frac{n}{\sum_{k=1}^n \log(X_k)}\right)} \\ &= \frac{-\frac{n}{\sum_{k=1}^n \log(X_k)}}{\left(\frac{\sum_{k=1}^n \log(X_k)-n}{\sum_{k=1}^n \log(X_k)}\right)^2\left(\frac{2\sum_{k=1}^n \log(X_k)-n}{\sum_{k=1}^n \log(X_k)}\right)} = \frac{n\left(\sum_{k=1}^n \log(X_k)\right)^2}{\left(n-\sum_{k=1}^n \log(X_k)\right)\left(n-2\sum_{k=1}^n \log(X_k)\right)}.\end{aligned}$$

Problem 6 *Let X a random variable representing a characteristic of a certain population. Assume that X has a density $f_X : \mathbb{R} \to \mathbb{R}$ given by*

$$f_X(x) \overset{def}{=} \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x), \quad \forall x \in \mathbb{R},$$

where θ is a positive parameter.

1. *Apply the method of moments to find the estimator $\hat{\theta}_M$ of the parameter θ.*
2. *Apply the maximum likelihood method to find the estimator $\hat{\theta}_{ML}$ of the parameter θ.*
3. *Use the estimators $\hat{\theta}_M$ and $\hat{\theta}_{ML}$ to build estimators for $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*

Solution. .

Problem 7 *Let $\theta > 0$ and let X be an absolutely continuous real random variable with density function $f_X : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_X(x) \overset{def}{=} \frac{1}{2} e^{-|x-\theta|}, \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to determine the estimator $\hat{\theta}_n^M$ for θ.*
2. *Check whether $\hat{\theta}_n^M$ is unbiased, consistent in probability, and consistent in mean square.*
3. *Can you "guess" the result of the method of maximum likelihood to determine the estimator $\hat{\theta}_n^{ML}$ for θ?*

Hint: recall that an estimator $\hat{\theta}_n$ for the true value of a parameter θ is said to be consistent in probability [resp. in mean square] if

$$\hat{\theta}_n \xrightarrow{\mathbf{P}} \theta \quad \text{[resp. } \hat{\theta}_n \xrightarrow{\mathbf{L}^2} \theta\text{]},$$

as $n \to \infty$.

Solution. .

Problem 8 *Let $\lambda > 0$ and let X be a Poisson real random variable with rate parameter λ, in symbols $X \sim Poiss(\lambda)$. Consider a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Let Z_n be the sample sum $X_1, \dots, X_n$, namely $Z_n \equiv \sum_{k=1}^n X_k$. Write the distribution of Z_n and compute $\mathbf{E}[Z_n]$ and $\mathbf{D}^2[Z_n]$.*
2. *Consider the sample mean $\bar{X}_n \equiv \frac{1}{n}Z_n$ and the unbiased sample variance $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^n \left(X_k - \bar{X}_n\right)^2$. Might they both be used to estimate λ? Which would perform better?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Confidence and Prediction Intervals 2021-12-17

Problem 1 *Assume that the log-returns of a stock in a financial market are Gaussian distributed with unknown mean μ and variance σ^2. Let X be the normal random variable representing the realization of the log-returns and let $X_1, \dots, X_n$ be a simple random sample of size n drawn from X. Assume that $n = 5$ and the realizations of the sample are*

$$x_1 \equiv -1.5, \quad x_2 \equiv -0.5, \quad x_3 \equiv 1.5, \quad x_4 \equiv 2.0, \quad x_5 \equiv 2.5$$

1. *Determine a 99% confidence interval for the mean μ.*
2. *Find the confidence for an interval of width 0.1.*
3. *Determine a 90% confidence interval for the standard deviation σ.*

Solution.

1. From data we obtain
$$\bar{x}_5 \equiv \frac{1}{5}\sum_{k=1}^{5} x_k = 0.8$$
and
$$s_{X,5}^2 = \frac{1}{4}\sum_{k=1}^{5} (x_k - \bar{x}_5)^2 = 2.95 \Rightarrow s_{X,5} = 1.72$$
Now, since X is Gaussian distributed with unknown variance, to determine a $100\,(1-\alpha)\,\%$ confidence interval for the mean μ the statistic to be considered is
$$\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} \sim t_{n-1}.$$
The achievement of a $100\,(1-\alpha)\,\%$ confidence interval requires to use the $\alpha/2$ upper and lower critical values $t^{+}_{n-1,\alpha/2} \equiv t_{n-1,1-\alpha/2}$ and $t^{-}_{n-1,\alpha/2} \equiv t_{n-1,\alpha/2} = -t_{n-1,1-\alpha/2}$ of t_{n-1} for $\alpha = 0.01$, where $t_{n-1,1-\alpha/2}$ [resp. $t_{n-1,\alpha/2}$] denotes the $(1-\alpha/2)$ [resp. $\alpha/2$]-quantile. In fact, we have
$$-t_{n-1,1-\alpha/2} < \frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < t_{n-1,1-\alpha/2} \Leftrightarrow -\left(\bar{X}_n + t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right) < -\mu < -\left(\bar{X}_n - t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right)$$
$$\Leftrightarrow \bar{X}_n - t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}},$$
which implies
$$\begin{aligned} 1-\alpha &= \mathbf{P}\left(-t_{n-1,1-\alpha/2} < \frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < t_{n-1,1-\alpha/2}\right) \\ &= \mathbf{P}\left(\bar{X}_n - t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right). \end{aligned}$$
It follows that the desired confidence interval for μ is given by the random interval
$$\left(\bar{X}_n - t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}},\ \bar{X}_n + t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right)$$
A realization of such a confidence interval is then given by
$$\left(\bar{x}_n - t_{n-1,1-\alpha/2}\frac{s_{X,n}}{\sqrt{n}},\ \bar{x}_n + t_{n-1,1-\alpha/2}\frac{s_{X,n}}{\sqrt{n}}\right).$$
In the case considered, since $t_{n-1,1-\alpha/2} \equiv t_{4,0.995} = 4.60$, $\bar{x}_n \equiv \bar{x}_5 = 0.80$, $s_{X,n} \equiv s_{X,5} = 1.72$, the realization of the confidence interval becomes
$$(-3.16,\ 4.76)\,.$$

2. From 1. it is clearly seen the width w of a $100\,(1-\alpha)\,\%$ confidence interval is given by
$$w = 2t_{n-1,1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}.$$
As a consequence, the size n of the sample which gives a $100\,(1-\alpha)\,\%$ confidence interval of a given widht w is given by the solution of the equation
$$\frac{n}{t^2_{n-1,1-\alpha/2}} = \left[4\frac{S^2_{X,n}}{w^2}\right] + 1.$$
To determine n we need to consider as many realizations $x_1, \dots, x_n$ of the simple random sample $X_1, \dots, X_n$ such that
$$\frac{n}{t^2_{n-1,1-\alpha/2}} = \left[4\frac{s^2_{X,n}}{0.01}\right] + 1.$$

3. Again, since X is Gaussian distributed, to determine a $100\,(1-\alpha)\,\%$ confidence interval for the standard deviation σ the statistic to be considered is
$$\frac{(n-1)\,S^2_{X,n}}{\sigma^2} \sim \chi^2_{n-1}.$$
Since χ^2_{n-1} is not symmetric, the achievement of a $100\,(1-\alpha)\,\%$ confidence interval requires to use the $\alpha/2$ and the $1-\alpha/2$ critical value $\chi^{2,-}_{n-1,\alpha/2} \equiv \chi^2_{n-1,\alpha/2}$ and $\chi^{2,+}_{n-1,\alpha/2} = \chi^2_{n-1,1-\alpha/2}$ of χ^2_{n-1} for $\alpha = 0.1$, where $\chi^2_{n-1,\alpha/2}$ [resp. $\chi^2_{n-1,1-\alpha/2}$] is the $\alpha/2$-quantile [$1-\alpha/2$-quantile] of the χ^2_{n-1} distribution. In fact, we have
$$\chi^{2,-}_{n-1,\alpha/2} < \frac{(n-1)\,S^2_{X,n}}{\sigma^2} < \chi^{2,+}_{n-1,\alpha/2} \Leftrightarrow \frac{1}{\chi^{2,-}_{n-1,\alpha/2}} > \frac{\sigma^2}{(n-1)\,S^2_{X,n}} > \frac{1}{\chi^{2,+}_{n-1,\alpha/2}}$$
$$\Leftrightarrow \frac{(n-1)\,S^2_{X,n}}{\chi^{2,+}_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)\,S^2_{X,n}}{\chi^{2,-}_{n-1,\alpha/2}},$$
which implies
$$1-\alpha = \mathbf{P}\left(\chi^{2,-}_{n-1,\alpha/2} < \frac{(n-1)\,S^2_{X,n}}{\sigma^2} < \chi^{2,+}_{n-1,\alpha/2}\right) = \mathbf{P}\left(\frac{(n-1)\,S^2_{X,n}}{\chi^{2,+}_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)\,S^2_{X,n}}{\chi^{2,-}_{n-1,\alpha/2}}\right).$$
It follows that the desired confidence interval for the variance σ^2 is given by
$$\left(\frac{(n-1)\,S^2_{X,n}}{\chi^{2,+}_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)\,S^2_{X,n}}{\chi^{2,-}_{n-1,\alpha/2}}\right) = \left(\frac{(n-1)\,S^2_{X,n}}{\chi^{2}_{n-1,1-\alpha/2}},\ \frac{(n-1)\,S^2_{X,n}}{\chi^{2}_{n-1,\alpha/2}}\right).$$

In the case considered, since $\chi^2_{n-1,\alpha/2} \equiv \chi^2_{4,0.5} = 0.71$, $\chi^2_{n-1,1-\alpha/2} \equiv \chi^2_{4,0.95} = 9.49$, $\bar{x}_n \equiv \bar{x}_5 = 0.80$, $s^2_{X,n} \equiv s^2_{X,5} = 2.95$, a realization of the confidence interval is given by

$$\left(\frac{4s^2_{X,n}}{\chi^2_{4,0.95}},\ \frac{4s^2_{X,n}}{\chi^2_{4,0.05}}\right) = \left(\frac{4\cdot 2.95}{9.49},\ \frac{4\cdot 2.95}{0.71}\right) = (1.24,\ 16.62)\,.$$

As a consequence the $100\,(1-0.1)\,\%$ confidence interval for the standard deviation σ is

$$(1.11,\ 4.08)\,.$$

This completes the solution.

Problem 2 *Assume that a library master believes that the mean duration in days of the borrowing period is* $20d$. *However, the library master selects randomly a simple random sample of* 100 *books in the library and discovers that the sample mean and variance of the borrowing days are* $18d$ *and* $8d^2$, *respectively. Determine a* 99% *confidence interval for the mean duration of the borrowing days to check whether library master's initial guess is correct.*

Solution. Note that the distribution of the random variable X representing the duration in days of the borrowing period is unknown. However, are known the sample mean and variance realizations referred to a simple sample of size $n = 100$, which may be considered a large sample. In this case the statistic to be considered is given by

$$\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}},$$

which is approximatively distributed as a standard Gaussian random variable $Z \sim N(0,1)$. The achievement of a $100\,(1-\alpha)\,\%$ confidence interval requires to use the $\alpha/2$ upper and lower critical values $z^+_{\alpha/2} \equiv z_{1-\alpha/2}$ and $z^-_{\alpha/2} \equiv z_{\alpha/2} = -z_{1-\alpha/2}$ of Z for $\alpha = 0.01$, where $z_{1-\alpha/2}$ [resp. $z_{\alpha/2}$] denotes the $(1-\alpha/2)$ [resp. $\alpha/2$]-quantile. In fact, we have

$$\begin{aligned}-z_{1-\alpha/2} < \frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < z_{1-\alpha/2} &\Leftrightarrow -\left(\bar{X}_n + z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right) < -\mu < -\left(\bar{X}_n - z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right)\\ &\Leftrightarrow \bar{X}_n - z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}} < \mu < \bar{X}_n + z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}},\end{aligned}$$

whuihc implies

$$\begin{aligned}1-\alpha &\approx \mathbf{P}\left(-z_{1-\alpha/2} < \frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < z_{1-\alpha/2}\right)\\ &= \mathbf{P}\left(\bar{X}_n - z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}} < \mu < \bar{X}_n + z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right)\end{aligned}$$

It follows that the desired confidence interval for μ is given by

$$\left(\bar{X}_n - z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}},\ \bar{X}_n + z_{1-\alpha/2}\frac{S_{X,n}}{\sqrt{n}}\right).$$

In the case considered,

$$n = 100, \quad \bar{x}_n \equiv \bar{x}_{100} = 18, \quad s_{X,n} \equiv s_{X,100} = \sqrt{8}, \quad z_{1-\alpha/2} \equiv 2.58.$$

Therefore, a realization of the confidence interval is given by

$$\left(\bar{x}_n - z_{1-\alpha/2}\frac{s_{X,n}}{\sqrt{n}},\ \bar{x}_n + z_{1-\alpha/2}\frac{s_{X,n}}{\sqrt{n}}\right) = \left(18 - 2.58\cdot\frac{\sqrt{8}}{\sqrt{100}}, 18 + 2.58\cdot\frac{\sqrt{8}}{\sqrt{100}}\right) = (17.27,\ 18.73)\,.$$

It follows that library master's initial guess is not supported by data. Note that this problem can be tackled also exploiting the hypothesis test method. In fact, assume as the null hypothesis that library master's assumption is correct, that is $H_0 : \mu = \mu_0$, and as the alternative hypothesis that library master's assumption is wrong, that is $H_0 : \mu \neq \mu_0$. The same consideration as above on the available information on the random variable X led to consider the rejection region

$$R = \left\{\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < z_{\alpha/2}\right\} \cup \left\{\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} > z_{1-\alpha/2}\right\} = \left\{\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} < -2.58\right\} \cup \left\{\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}} > 2.58\right\},$$

where $\mu = 20$. Computing the statistic $\frac{\bar{X}_n - \mu}{S_{X,n}/\sqrt{n}}$ for the available realization, we obtain

$$\frac{\bar{x}_n - \mu}{s_{X,n}/\sqrt{n}} = \frac{18 - 20}{\sqrt{8}/10} = -7.07 \in R.$$

Hence, the library master's assumption has to be rejected.

Problem 3 *A sample of* 60 *cars of the same model are tested for gasoline consumption, expressed as litres per* 100 *kilometers (L/100km). The result of the test yield a mean consumption* μ *of* 9.4 *L/100km and a standard deviation* σ *of* 1.5 *L/100km.*

1. *Determine a* 99% *confidence interval for the mean consumption. Is it necessary to make any assumption on the consumption distribution?.*
2. *Assume that the consumption is normally distributed with variance* $\sigma^2 = 2$ *L/100km. How large has to be the sample if, with the same confidence, we want the maximum error to be* 0.25 *L/100km? Apply the method of the confidence interval and the Chebychev inequality.*

Solution. .

Problem 4 *Let* X *[resp.* Y*] a Gaussian distributed random variables with (unknown) mean* $\mu_X \in \mathbb{R}$ *[resp.* $\mu_Y \in \mathbb{R}$*] and variance* $\sigma^2_X > 0$ *[resp.* $\sigma^2_Y > 0$*]. Assume that* X *describes a population before a treatment and* Y *describes the same population after a treatment. Let* $X_1, \ldots, X_n$ *be a simple random sample drawn by* X *and let* $Y_1, \ldots, Y_n$ *be the corresponding sample drawn from* Y. *Note that we can still assume that* $Y_1, \ldots, Y_n$ *is a simple random sample but we cannot assume that the samples* $X_1, \ldots, X_n$ *and* $Y_1, \ldots, Y_n$ *are independent. Actually, there is no reason at all to think that the random variables* X *and* Y *are independent. However, it is still reasonable to assume that the random variable* $D \equiv Y - X$ *is Gaussian distributed and that* $D_1 \equiv Y_1 - X_1, \ldots, D_n \equiv Y_n - X_n$ *is a simple random sample drawn from* D.

1. *Given* $\alpha > 0$, *can you build a* $100\,(1-\alpha)\,\%$ *confidence interval for the difference* $\mu_Y - \mu_X$*?*
2. *Assume to have measured*

$$\begin{array}{lllllll} x_1 = 3.85 & x_2 = 2.82 & x_3 = 3.44 & x_4 = 3.48 & x_5 = 1.92 & x_6 = 4.39 & x_7 = 3.12\\ y_1 = 5.73 & y_2 = 3.84 & y_3 = 4.78 & y_4 = 4.40 & y_5 = 1.91 & y_6 = 4.98 & y_7 = 4.94\end{array}.$$

Determine a realization of the 95% *confidence interval built above.*

Solution. We clearly have

$$\mu_D \equiv \mathbf{E}[D] \equiv \mathbf{E}[Y - X] = \mathbf{E}[Y] - \mathbf{E}[X] = \mu_Y - \mu_X.$$

Therefore, introducing the sample mean of size n drawn from D, that is

$$\bar{D}_n = \frac{1}{n}\sum_{k=1}^{n} D_k = \frac{1}{n}\sum_{k=1}^{n}(Y_k - X_k) = \frac{1}{n}\sum_{k=1}^{n} Y_k - \frac{1}{n}\sum_{k=1}^{n} X_k = \bar{Y}_n - \bar{X}_n$$

and the unbiased sample variance of size n drawn from D, that is

$$S_n^2(D) = \frac{1}{n-1}\sum_{k=1}^{n}\left(D_k - \bar{D}_n\right)^2,$$

under the assumptions considered, we have that the statistic

$$\frac{\bar{D}_n - \mu_D}{S_n(D)/\sqrt{n}},$$

has the Student distribution with $n-1$ degrees of freedom. As a consequence, given $\alpha > 0$ a $100(1-\alpha)\%$ confidence interval for μ_D is given by

$$\left(\bar{D}_n - t^-_{\frac{\alpha}{2},n-1}\frac{S_n(D)}{\sqrt{n}}, \bar{D}_n + t^+_{\frac{\alpha}{2},n-1}\frac{S_n(D)}{\sqrt{n}}\right).$$

The realization of such a confidence interval are of the form

$$\left(\bar{d}_n - t_{\frac{\alpha}{2},n-1}\frac{s_n(D)}{\sqrt{n}}, \bar{d}_n + t_{\frac{\alpha}{2},n-1}\frac{s_n(D)}{\sqrt{n}}\right),$$

where $\bar{d}_n$ [resp. $s_n(D)$] is the value taken by the sample mean estimator $\bar{D}_n$ [resp. unbiased sample standard deviation estimator $S_n(D)$] on the available realizations $d_1, \ldots, d_n$ of the sample $D_1, \ldots, D_n$. Since in our case $n = 7$ and $\alpha \equiv 0.05$, we have

$$t_{\frac{\alpha}{2},n-1} \equiv t_{0.025,6} = 2.45$$

Furthermore,

$$\bar{d}_n \equiv \bar{d}_7 = 1.08 \quad \text{and} \quad s_n(D) \equiv s_7(D) = 0.45$$

Therefore,

$$\bar{d}_n - t_{\frac{\alpha}{2},n-1}\frac{s_n(D)}{\sqrt{n}} = 1.08 - 2.45\frac{0.45}{\sqrt{7}} = 0.66$$

and

$$\bar{d}_n + t_{\frac{\alpha}{2},n-1}\frac{s_n(D)}{\sqrt{n}} = 1.08 + 2.45\frac{0.45}{\sqrt{7}} = 1.50.$$

Thus the realization of the confidence interval is

$$(0.66, 1.50).$$

Problem 5 *Two independent groups of* 50 *people, say group A and group B, are affected by a disease. A drug is given to the individuals of group A (*treatment group*) but not to the individuals of group B (*control group*). After a week since the administration of the drug a medical check shows that* 45 *[resp.* 30*] people of group A [resp. B] got recovered from the disease. Let p_A [resp. p_B] the proportion of inividuals of group A [resp. B] who got recovered.*

1. *Given $\alpha > 0$, can you build a $100(1-\alpha)\%$ confidence interval for the true value of difference $p_A - p_B$?*

2. *Determine the realization of the* 95% *confidence interval built above and use it to comment on the efficacy of the drug.*

Solution. .

Problem 6 *Let X be a standard Bernoulli random variable with unknown success parameter p. Let $X_1, \ldots, X_n$ be a simple random sample of size n drawn from X and let $Z_n \equiv \sum_{k=1}^{n} X_k$ be the sample sum. It is well known that $Z_n \sim Bin(n,p)$. In addition, when n is large ($np \geq 10$ and $n(1-p) \geq 10$) the sample sum has approximately a normal distribution.*

1. *Determine a confidence interval for the parameter p with confidence level approximately $100(1-\alpha)\%$.*

2. *Determine the size n of the sample $X_1, \ldots, X_n$ which allows a confidence interval for the parameter p with confidence level approximately $100(1-\alpha)\%$ and width w, where both α and w are given in advance.*

Solution. .

Problem 7 *Let $X_1, \ldots, X_n, X_{n+1}$ be a simple random sample of size $n+1$ drawn from a Gaussian distributed random variable X with unknown mean μ and variance σ^2. Assume that we have observed $X_1, \ldots, X_n$ and we want use the observed values $x_1, \ldots, x_n$ to determine a confidence interval for the prediction of X_{n+1}. To this goal give detailed answers to the following questions:*

1. *what is the distribution of the statistic $\bar{X}_n$?*

2. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/\sigma\sqrt{1+1/n}$?*

3. *what is the distribution of the statistic $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)^2$?*

4. *are the statistics $X_{n+1} - \bar{X}_n$ and $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)^2$ independent? Why?*

5. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/S_n\sqrt{1+1/n}$?*

6. *After answering the above questions, build an interval in which the random variable X_{n+1} takes its values with probability α and determine the corresponding confidence interval for the prediction of X_{n+1}. In the end, assume that $n = 7$ and we have*

   $$x_1 = 7005, \quad x_2 = 7432, \quad x_3 = 7420, \quad x_4 = 6822, \quad x_5 = 6752, \quad x_6 = 5333, \quad x_7 = 6552.$$

   compute the 95% *confidence interval for the prediction of X_8.*

Solution. .

Problem 8 *A test on the reaction time measured in seconds to a sudden emergency has lead to the following results in* 10 *people:*

$$t_1 = 0.77, \quad t_2 = 0.75, \quad t_3 = 0.70, \quad t_4 = 0.72, \quad t_5 = 0.70,$$
$$t_6 = 0.69, \quad t_7 = 0.67, \quad t_8 = 0.79, \quad t_9 = 0.64, \quad t_{10} = 0.72.$$

Assume that the reaction time can be modelled by a normal random variable $T \sim N\left(\mu, \sigma^2\right)$ *with unknown* μ *and* σ^2.

1. *Compute the sample mean and variance referred to the above sample.*
2. *Find the confidence interval for* μ *[resp.* σ^2*] at the confidence level* 95%.
3. *What would the confidence interval for* μ *be if the variance were known and we had* $\sigma^2 = 0.0025$*?*
4. *What should the size of the sample be to achieve a precision of* 10*?*

Solution. .

Problem 9 *Assume we need to measure the same trait* X *in two different population and the results of our measurement for a sample of size* $n_1 = 15$ *[resp.* $n_2 = 20$*] of the first [resp. second] population gives a value* $\bar{x}_{n_1} = 24.0$ *[resp.* $\bar{x}_{n_2} = 26.0$*] of the sample mean* μ_X *[resp.* μ_Y*] of the characteristic under investigation, with a sample variance* $s^2_{n_1} = 4.5$ *[resp.* $s^2_{n_2} = 5.0$*]. Assume that the trait is normally distributed with the same unknown variance* σ^2_X. *Computed a confidence interval for the value of the difference* $\mu_X - \mu_Y$ *with a* 95% *confidence level.*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2019-10-31

Problem 1 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $\left(\mathbb{R}^2, \mathcal{B}\left(\mathbb{R}^2\right), \mu_L^2\right) \equiv \mathbb{R}^2$ *be the Euclidean real plane endowed with the Borel* σ*-algebra and the Borel-Lebesgue measure* $\mu_L^2 : \mathcal{B}\left(\mathbb{R}^2\right)$. *Let* $f : \mathbb{R}^2 \to \mathbb{R}_+$ *given by*

$$f(x,y) \stackrel{def}{=} kxye^{-(x^2+y^2)} 1_{\mathbb{R}^2_+}(x,y), \quad \forall (x,y) \in \mathbb{R}^2$$

where $\mathbb{R}^2_+ \equiv \left\{(x,y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\right\}$. *Determine* $k \in \mathbb{R}$ *such that* $f : \mathbb{R}^2 \to \mathbb{R}_+$ *is a probability density and let* $Z \equiv (X,Y)$ *be the random vector of density* $f : \mathbb{R}^2 \to \mathbb{R}_+$.

1. *Determine the distribution function* $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ *of the vector* Z *and check that*
$$\frac{\partial F^2}{\partial x \partial y}(x,y) = f(x,y), \quad \mu_L^2 - a.e. \text{ on } \mathbb{R}^2.$$
2. *Determine the marginal distribution function* $F_X : \mathbb{R} \to \mathbb{R}_+$ *and* $F_Y : \mathbb{R} \to \mathbb{R}_+$ *of the entries* X *and* Y *of* Z.
3. *Determine the densities* $f_X : \mathbb{R} \to \mathbb{R}_+$ *and* $f_Y : \mathbb{R} \to \mathbb{R}_+$ *of the entries* X *and* Y *of* Z *and check that*
$$\frac{dF_X}{dx}(x) = f_X(x) \quad \text{and} \quad \frac{dF_Y}{dy}(y) = f_Y(y), \quad \mu_L - a.e. \text{ on } \mathbb{R}.$$
4. *Are* X *and* Y *independent random varables?*
5. *Compute* $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$ *and* $Cov(X,Y)$.
6. *Compute* $\mathbf{E}[(X,Y)]$ *and the covariance matrix of the vector* (X,Y).

Solution. . □

Problem 2 *Determine the value of the parameter* k *such that the function* $f : \mathbb{R}^3 \to \mathbb{R}$ *given by*

$$f(x_1,x_2,x_3) \stackrel{def}{=} \begin{cases} k\left(x_1 + x_2^2 + x_3^3\right) & \text{if } (x_1,x_2,x_3) \in [0,1] \times [0,1] \times [0,1] \\ 0 & \text{otherwise} \end{cases}$$

is a probability density. Hence, consider the random vector $X \equiv (X_1, X_2, X_3)^{\intercal}$ *with density* $f_X : \mathbb{R}^3 \to \mathbb{R}$ *given by*

$$f_X(x_1,x_2,x_3) \stackrel{def}{=} f(x_1,x_2,x_3).$$

1. *Determine the distribution function* $F_X : \mathbb{R}^3 \to \mathbb{R}_+$ *and check that*
$$\frac{\partial F_X^2}{\partial x_1 \partial x_2 \partial x_3}(x_1,x_2,x_3) = f_X(x_1,x_2,x_3), \quad \mu_L^3\text{-a.e. on } \mathbb{R}^3.$$
2. *Determine the marginal distribution function* $F_{X_1} : \mathbb{R} \to \mathbb{R}_+$, $F_{X_2} : \mathbb{R} \to \mathbb{R}_+$, *and* $F_{X_3} : \mathbb{R} \to \mathbb{R}_+$ *of the entries* X_1, X_2, *and* X_3 *of* X.

3. *Determine the marginal densities* $f_{X_1}:\mathbb{R}\to\mathbb{R}_+$, $f_{X_2}:\mathbb{R}\to\mathbb{R}_+$, *and* $f_{X_3}:\mathbb{R}\to\mathbb{R}_+$ *of the entries* X_1, X_2, *and* X_3 *of* X *and check that*

$$\frac{dF_{X_n}}{dx}(x)=f_{X_n}(x),\ \text{for } n=1,2,3,\quad \mu_L\text{-a.e. on } \mathbb{R}.$$

4. *Determine the joint distribution function* $F_{X_1,X_2}:\mathbb{R}^2\to\mathbb{R}_+$, $F_{X_1,X_3}:\mathbb{R}\to\mathbb{R}_+$, *and* $F_{X_2,X_3}:\mathbb{R}\to\mathbb{R}_+$. *Is it useful to compute the joint distribution function* $F_{X_2,X_1}:\mathbb{R}^2\to\mathbb{R}_+$, $F_{X_3,X_1}:\mathbb{R}\to\mathbb{R}_+$, *and* $F_{X_3,X_2}:\mathbb{R}\to\mathbb{R}_+$*?*

5. *Determine the joint densities* $f_{X_1,X_2}:\mathbb{R}^2\to\mathbb{R}_+$, $f_{X_1,X_3}:\mathbb{R}\to\mathbb{R}_+$, *and* $f_{X_2,X_3}:\mathbb{R}\to\mathbb{R}_+$. *What is the relationship between the joint distribution function* $F_{X_m,X_n}:\mathbb{R}^2\to\mathbb{R}_+$ *and the joint density* $f_{X_m,X_n}:\mathbb{R}^2\to\mathbb{R}_+$ *for* $m,n=1,2,3$, $m<n$.

6. *Determine the the expectation of* X.

7. *Determine the variance-covariance matrix of* X.

Solution. To determine the value of the parameter k such that the function $f:\mathbb{R}^3\to\mathbb{R}$ is a probability density we have to solve the equation

$$\int_{\mathbb{R}^3} f(x_1,x_2,x_3)\,d\mu_L(x_1,x_2,x_3)=1.$$

We have

$$f(x_1,x_2,x_3)=k\left(x_1+x_2^2+x_3^3\right)1_{[0,1]\times[0,1]\times[0,1]}(x_1,x_2,x_3),$$

Hence,

$$\begin{aligned}\int_{\mathbb{R}^3} f(x_1,x_2,x_3)\,d\mu_L(x_1,x_2,x_3)&=\int_{\mathbb{R}^3} k\left(x_1+x_2^2+x_3^3\right)1_{[0,1]\times[0,1]\times[0,1]}(x_1,x_2,x_3)\,d\mu_L(x_1,x_2,x_3)\\&=\int_{[0,1]\times[0,1]\times[0,1]} k\left(x_1+x_2^2+x_3^3\right)d\mu_L(x_1,x_2,x_3)\\&=k\int_{[0,1]\times[0,1]\times[0,1]}\left(x_1+x_2^2+x_3^3\right)d\mu_L(x_1,x_2,x_3)\end{aligned}$$

Now the real function $x_1+x_2^2+x_3^3$ is continuous on $[0,1]\times[0,1]\times[0,1]$. Therefore, the Lebesue integral can be computed as a Riemann integral. A as consequence, on account of the additive property of the Riemann integral and the separability of the integrand function on the pluri-interval domain, we can write

$$\begin{aligned}&\int_{[0,1]\times[0,1]\times[0,1]}\left(x_1+x_2^2+x_3^3\right)d\mu_L(x_1,x_2,x_3)\\&=\int_{x_1=0}^1\int_{x_2=0}^1\int_{x_3=0}^1\left(x_1+x_2^2+x_3^3\right)dx_1dx_2dx_3\\&=\int_{x_1=0}^1\int_{x_2=0}^1\int_{x_3=0}^1 x_1dx_1dx_2dx_3\\&+\int_{x_1=0}^1\int_{x_2=0}^1\int_{x_3=0}^1 x_2^2dx_1dx_2dx_3\\&+\int_{x_1=0}^1\int_{x_2=0}^1\int_{x_3=0}^1 x_3^3dx_1dx_2dx_3\\&=\int_{x_1=0}^1 x_1dx_1\int_{x_2=0}^1 dx_2\int_{x_3=0}^1 dx_3\\&+\int_{x_1=0}^1 dx_1\int_{x_2=0}^1 x_2^2dx_2\int_{x_3=0}^1 dx_3\\&+\int_{x_1=0}^1 dx_1\int_{x_2=0}^1 dx_2\int_{x_3=0}^1 x_3^3dx_3\\&=\frac{1}{2}\,x_1^2\Big|_{x_1=0}^1\cdot x_2\Big|_{x_2=0}^1\cdot x_3\Big|_{x_3=0}^1\\&+x_1\Big|_{x_1=0}^1\cdot\frac{1}{3}\,x_2^3\Big|_{x_2=0}^1\cdot x_3\Big|_{x_3=0}^1\\&+x_1\Big|_{x_1=0}^1\cdot x_2\Big|_{x_2=0}^1\frac{1}{4}\cdot x_3^4\Big|_{x_3=0}^1\\&=\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\\&=\frac{13}{12}\end{aligned}$$

It follows

$$k=\frac{12}{13}.$$

With similar computation, we have

$$\begin{aligned}\mathbf{P}(X_2\le 1/2,X_3>1/2)&=\int_{x_1=0}^1\int_{x_2=0}^{1/2}\int_{x_3=1/2}^1\frac{12}{13}\left(x_1+x_2^2+x_3^3\right)dx_1dx_2dx_3\\&=\frac{12}{13}\left(\frac{1}{2}\,x_1^2\Big|_{x_1=0}^1\cdot x_2\Big|_{x_2=0}^{1/2}\cdot x_3\Big|_{x_3=1/2}^1\right.\\&+x_1\Big|_{x_1=0}^1\cdot\frac{1}{3}\,x_2^3\Big|_{x_2=0}^{1/2}\cdot x_3\Big|_{x_3=1/2}^1\\&\left.+x_1\Big|_{x_1=0}^1\cdot x_2\Big|_{x_2=0}^{1/2}\cdot\frac{1}{4}\,x_3^4\Big|_{x_3=1/2}^1\right)\\&=\frac{12}{13}\left(\frac{1}{8}+\frac{1}{48}+\frac{15}{128}\right)\\&=\frac{101}{416}.\end{aligned}$$

The marginal density of the random vector $(X_1, X_2)^\intercal$ is given by

$$\begin{aligned}
f_{X_1,X_2}(x_1,x_2) &= \int_{\mathbb{R}} f(x_1,x_2,x_3)\, d\mu_L(x_3) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]\times[0,1]\times[0,1]}(x_1,x_2,x_3)\, d\mu_L(x_3) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, 1_{[0,1]}(x_3)\, d\mu_L(x_3) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + x_3^3\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, d\mu_L(x_3) \\
&= k1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2) \int_{x_3=0}^{1} \left(x_1 + x_2^2 + x_3^3\right) dx_3 \\
&= k1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2) \left(\int_{x_3=0}^{1} x_1 d\mu_L(x_3) + \int_{x_3=0}^{1} x_2^2 d\mu_L(x_3) + \int_{x_3=0}^{1} x_3^3 d\mu_L(x_3)\right) \\
&= k1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2) \left(x_1 \cdot x_3\big|_{x_3=0}^{1} + x_2^2 \cdot x_3\big|_{x_3=0}^{1} + \frac{1}{4}\, x_3^4\big|_{x_3=0}^{1}\right) \\
&= k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2) \\
&= k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}(x_1,x_2).
\end{aligned}$$

We have

$$\mathbf{E}\left[(X_1, X_2)^\intercal\right] = \left(\mathbf{E}[X_1], \mathbf{E}[X_2]\right)^\intercal,$$

where

$$\mathbf{E}[X_k] = \int_{\mathbb{R}} x_k f_{X_k}(x_k)\, d\mu_L(x_k), \quad k = 1, 2,$$

and $f_{X_k}(x_k)$ is the marginal density of the random variable X_k, for $k = 1, 2$. Now,

$$\begin{aligned}
f_{X_1}(x_1) &= \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\, d\mu_L(x_2) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)\, d\mu_L(x_2) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, d\mu_L(x_2) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_1)\, d\mu_L(x_2) \\
&= k1_{[0,1]}(x_1) \int_{x_2=0}^{1} \left(x_1 + x_2^2 + \frac{1}{4}\right) dx_2 \\
&= k1_{[0,1]}(x_1) \left(x_1 \cdot x_2\big|_{x_2=0}^{1} + \frac{1}{3} \cdot x_2^3\big|_{x_2=0}^{1} + \frac{1}{4} \cdot x_2\big|_{x_2=0}^{1}\right) \\
&= k1_{[0,1]}(x_1) \left(x_1 + \frac{1}{3} + \frac{1}{4}\right) \\
&= k\left(x_1 + \frac{7}{12}\right) 1_{[0,1]}(x_1).
\end{aligned}$$

Similarly,

$$\begin{aligned}
f_{X_2}(x_2) &= \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\, d\mu_L(x_1) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)\, d\mu_L(x_1) \\
&= \int_{\mathbb{R}} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_1)\, 1_{[0,1]}(x_2)\, d\mu_L(x1) \\
&= \int_{[0,1]} k\left(x_1 + x_2^2 + \frac{1}{4}\right) 1_{[0,1]}(x_2)\, d\mu_L(x_1) \\
&= k1_{[0,1]}(x_2) \int_{x_1=0}^{1} \left(x_1 + x_2^2 + \frac{1}{4}\right) dx_1 \\
&= k1_{[0,1]}(x_2) \left(\frac{1}{3} \cdot x_1^2\big|_{x_1=0}^{1} + x_2^2 \cdot x_1\big|_{x_1=0}^{1} + \frac{1}{4} \cdot x_1\big|_{x_1=0}^{1}\right) \\
&= k1_{[0,1]}(x_2) \left(\frac{1}{3} + x_2^2 + \frac{1}{4}\right) \\
&= k\left(x_2^2 + \frac{7}{12}\right) 1_{[0,1]}(x_2).
\end{aligned}$$

It follows

$$\begin{aligned}
\mathbf{E}[X_1] &= \int_{\mathbb{R}} k\left(x_1 + \frac{7}{12}\right) 1_{[0,1]}(x_1) = k\int_{x_1=0}^{1} \left(x_1 + \frac{7}{12}\right) dx_1 \\
&= k\left(\frac{1}{2} \cdot x_1^2\big|_{x_1=0}^{1} + \frac{7}{12} \cdot x_1\big|_{x_1=0}^{1}\right) = \frac{13}{12}k
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}[X_2] &= \int_{\mathbb{R}} k\left(x_2^2 + \frac{7}{12}\right) 1_{[0,1]}(x_2) = k\int_{x_2=0}^{1} \left(x_2^2 + \frac{7}{12}\right) dx_2 \\
&= k\left(\frac{1}{3} \cdot x_1^3\big|_{x_2=0}^{1} + \frac{7}{12} \cdot x_2\big|_{x_2=0}^{1}\right) = \frac{11}{12}k.
\end{aligned}$$

The conditional density $f_{X_1,X_2|X_3=1/2}(x_1,x_2)$ is simply given by

$$f_{X_1,X_2|X_3=1/2}(x_1,x_2) = \frac{f_{X_1,X_2,X_3}(x_1,x_2,1/2)}{\int_{\mathbb{R}^2} f_{X_1,X_2,X_3}(x_1,x_2,1/2)\, d\mu_L(x_1,x_2)} = \frac{f_{X_1,X_2,X_3}(x_1,x_2,1/2)}{f_{X_3}(1/2)},$$

for every $(x_1,x_2) \in \mathbb{R}^2$. Now, since

$$f_{X_1,X_2,X_3}(x_1,x_2,1/2) = k\left(x_1 + x_2^2 + \frac{1}{8}\right) 1_{[0,1]\times[0,1]}(x_1,x_2)$$

and

$$
\begin{aligned}
&\int_{\mathbb{R}^2} f_{X_1,X_2,X_3}(x_1,x_2,1/2)\,d\mu_L(x_1,x_2)\\
&= \int_{\mathbb{R}^2} k\left(x_1+x_2^2+\frac{1}{8}\right)1_{[0,1]\times[0,1]}(x_1,x_2)\,d\mu_L(x_1,x_2)\\
&= \int_{[0,1]\times[0,1]} k\left(x_1+x_2^2+\frac{1}{8}\right)d\mu_L(x_1,x_2)\\
&= \int_{x_1=0}^{1}\int_{x_2=0}^{1} k\left(x_1+x_2^2+\frac{1}{8}\right)dx_1dx_2\\
&= k\left(\int_{x_1=0}^{1}\int_{x_2=0}^{1} x_1dx_1dx_2+\int_{x_1=0}^{1}\int_{x_2=0}^{1} x_2^2dx_1dx_2+\int_{x_1=0}^{1}\int_{x_2=0}^{1}\frac{1}{8}dx_1dx_2\right)\\
&= k\left(\int_{x_1=0}^{1} x_1dx_1\int_{x_2=0}^{1} dx_2+\int_{x_1=0}^{1} dx_1\int_{x_2=0}^{1} x_2^2dx_2+\frac{1}{8}\int_{x_1=0}^{1} dx_1\int_{x_2=0}^{1} dx_2\right)\\
&= k\left(\frac{1}{2}\cdot x_1^2\Big|_{x_1=0}^{1}\cdot x_2\Big|_{x_2=0}^{1}+x_1\Big|_{x_1=0}^{1}\cdot\frac{1}{3}\cdot x_2^3\Big|_{x_2=0}^{1}+\frac{1}{8}\cdot x_1\Big|_{x_1=0}^{1}\,x_2\Big|_{x_2=0}^{1}\right)\\
&= k\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{8}\right)\\
&= \frac{23}{24}k,
\end{aligned}
$$

we obtain

$$f_{X_1,X_2|X_3=1/2}(x_1,x_2)=\frac{24}{23}\left(x_1+x_2^2+\frac{1}{8}\right)1_{[0,1]\times[0,1]}(x_1,x_2)\,.$$

Problem 3 *Let $F:\mathbb{R}^2\to\mathbb{R}_+$ given by*

$$F(x_1,x_2)\overset{def}{=}\left(1-e^{-x_1}-e^{-x_2}+e^{-(x_1+x_2)}\right)1_{\mathbb{R}_+}(x_1)\,1_{\mathbb{R}_+}(x_2)\,,\quad\forall (x_1,x_2)\in\mathbb{R}^2.$$

Show that $F:\mathbb{R}^2\to\mathbb{R}_+$ is the distribution function of a real random vector (X_1,X_2) and compute the marginal distribution functions of the entries X_1 and X_2 of (X_1,X_2).

Exercise 4 *1. May we say that $F:\mathbb{R}^2\to\mathbb{R}_+$ is absolutely continuous?*

2. May we say that the entries X_1 and X_2 of the random vector (X_1,X_2) are independent random variables?

3. May we say that the entries X_1 and X_2 of the random vector (X_1,X_2) are absolutely continuous random variables?

4. Consider the real random variable $Z=\max\{X_1,X_2\}$. Determine the distribution function $F_Z:\mathbb{R}^2\to\mathbb{R}_+$ of Z is $F_Z:\mathbb{R}^2\to\mathbb{R}_+$ absolutely continuous?

Solution. We have

$$
\begin{aligned}
\left(1-e^{-x_1}-e^{-x_2}+e^{-(x_1+x_2)}\right)1_{\mathbb{R}_+}(x_1)\,1_{\mathbb{R}_+}(x_2) &= \left(\left(1-e^{-x_1}\right)-\left(1-e^{-x_1}\right)e^{-x_2}\right)1_{\mathbb{R}_+}(x_1)\,1_{\mathbb{R}_+}(x_2)\\
&= \left(1-e^{-x_1}\right)1_{\mathbb{R}_+}(x_1)\left(1-e^{-x_2}\right)1_{\mathbb{R}_+}(x_2)\,,
\end{aligned}
$$

for every $(x_1,x_2)\in\mathbb{R}^2$. Therefore,

$$\lim_{x_1\to-\infty}\lim_{x_2\to-\infty}F(x_1,x_2)=0\quad\text{and}\quad\lim_{x_1\to+\infty}\lim_{x_2\to+\infty}F(x_1,x_2)=1.$$

In addition, $F(x_1,x_2)$ is non decreasing and right-hand continuous in each variable. These properties imply that $F:\mathbb{R}^2\to\mathbb{R}_+$ the distribution function of a real random vector (X_1,X_2). The marginal distributions $F_{X_1}:\mathbb{R}\to\mathbb{R}_+$ and $F_{X_2}:\mathbb{R}\to\mathbb{R}_+$ are given by

$$F_{X_1}(x_1)=\lim_{x_2\to-\infty}F(x_1,x_2)\,,\quad\forall x_1\in\mathbb{R}\quad\text{and}\quad F_{X_2}(x_2)=\lim_{x_1\to+\infty}F(x_1,x_2)\,,\quad\forall x_2\in\mathbb{R},$$

respectively. Hence,

$$F_{X_1}(x_1)=\lim_{x_2\to-\infty}\left(1-e^{-x_1}\right)1_{\mathbb{R}_+}(x_1)\left(1-e^{-x_2}\right)1_{\mathbb{R}_+}(x_2)=\left(1-e^{-x_1}\right)1_{\mathbb{R}_+}(x_1)$$

and

$$F_{X_2}(x_2)=\lim_{x_1\to-\infty}\left(1-e^{-x_1}\right)1_{\mathbb{R}_+}(x_1)\left(1-e^{-x_2}\right)1_{\mathbb{R}_+}(x_2)=\left(1-e^{-x_2}\right)1_{\mathbb{R}_+}(x_2)\,.$$

Note that both $F_{X_1}:\mathbb{R}\to\mathbb{R}_+$ and $F_{X_2}:\mathbb{R}\to\mathbb{R}_+$ are absolutely continuous functions with density $f_{X_1}:\mathbb{R}\to\mathbb{R}_+$ and $f_{X_2}:\mathbb{R}\to\mathbb{R}_+$ given by

$$f_{X_1}(x_1)=e^{-x_1}1_{\mathbb{R}_+}(x_1)\quad\text{and}\quad f_{X_2}(x_2)=e^{-x_2}1_{\mathbb{R}_+}(x_2)\,,$$

respectively. This proves that the entries X_1 and X_2 of the random vector (X_1,X_2) are absolutely continuous random variables. As a consequence $F:\mathbb{R}^2\to\mathbb{R}_+$ is itself absolutely continuous. In fact

$$
\begin{aligned}
F(x_1,x_2) &= \int_{(-\infty,x_1]\times(-\infty,x_2]} f_{X_1}(u_1)\,f_{X_2}(u_2)\;d\mu_L(u_1,u_2)\\
&= \int_{(-\infty,x_1]} f_{X_1}(u_1)\;d\mu_L(u_1)\int_{(-\infty,x_2]} f_{X_2}(u_2)\;d\mu_L(u_2)\\
&= \int_{(-\infty,x_1]} e^{-u_1}1_{\mathbb{R}_+}(u_1)\;d\mu_L(u_1)\int_{(-\infty,x_2]} e^{-u_2}1_{\mathbb{R}_+}(u_2)\;d\mu_L(u_2)\\
&= \int_{(-\infty,x_1]\cap\mathbb{R}_+} e^{-u_1}\;d\mu_L(u_1)\int_{(-\infty,x_2]\cap\mathbb{R}_+} e^{-u_2}\;d\mu_L(u_2)\,.
\end{aligned}
$$

Hence,

$$F(x_1,x_2)=\begin{cases}\int_{[0,x_1]}e^{-u_1}\;d\mu_L(u_1)\int_{[0,x_2]}e^{-u_2}\;d\mu_L(u_2)=\int_1^{x_1}e^{-u_1}\;du_1\int_1^{x_1}e^{-x_2}\;du_2 & \text{if } x_1,x_2>0\\ 0 & \text{otherwise}\end{cases}.$$

It follows

$$F(x_1,x_2)=\left(1-e^{x_1}\right)\left(1-e^{x_2}\right)1_{\mathbb{R}_{++}}(x_1)\,1_{\mathbb{R}_{++}}(x_2)=\left(1-e^{x_1}\right)\left(1-e^{x_2}\right)1_{\mathbb{R}_+}(x_1)\,1_{\mathbb{R}_+}(x_2)$$

almost everywhere in $\mathbb{R}^2$. Moreover, we have

$$F(x_1,x_2)=\int_{(-\infty,x_1]} f_{X_1}(u_1)\;d\mu_L(u_1)\int_{(-\infty,x_2]} f_{X_2}(u_2)\;d\mu_L(u_2)=F_{X_1}(x_1)\,F_{X_2}(x_2)\,,$$

almost everywhere in $\mathbb{R}^2$. This proves that the entries X_1 and X_2 of the random vector (X_1,X_2) are independent random variables.

In the end, to determine the distribution function $F_Z:\mathbb{R}^2\to\mathbb{R}_+$ of Z, we consider the event $\{Z\le z\}$. We have

$$\{Z\le z\}=\{\max\{X_1,X_2\}\le z\}=\{X_1\le z,X_2\le z\}.$$

Hence, on accont of the independence of X_1 and X_2, we can write

$$
\begin{aligned}
F_Z(z) &= \mathbf{P}(Z\le z)=\mathbf{P}(X_1\le z,X_2\le z)=\mathbf{P}(X_1\le z)\,\mathbf{P}(X_2\le z)\\
&= F_{X_1}(z)\,F_{X_2}(z)=\left(1-e^{-z}\right)1_{\mathbb{R}_+}(z)\left(1-e^{-z}\right)1_{\mathbb{R}_+}(z)\\
&= \left(1-e^{-z}\right)^2 1_{\mathbb{R}_+}(z)=\left(1-2e^{-z}+e^{-2z}\right)1_{\mathbb{R}_+}(z)\,.
\end{aligned}
$$

Now, since

$$\int_0^z e^{-v}\, dv = 1 - e^{-z} \quad \text{and} \quad \int_0^z e^{-2v}\, dv = \frac{1}{2}\left(1 - e^{-2z}\right)$$

for every $z > 0$, we have

$$2\int_0^z \left(e^{-v} - e^{-2v}\right)\, dv = 2\left(1 - e^{-z}\right) - 2\frac{1}{2}\left(1 - e^{-2z}\right) = 1 - 2e^{-z} + e^{-2z},$$

for every $z > 0$. It clearly follows

$$\begin{aligned} F_Z(z) &= \int_{(-\infty,z]} 2\left(e^{-v} - e^{-2v}\right) 1_{\mathbb{R}_+}(v)\, d\mu_L(v) \\ &= \int_{(-\infty,z]\cap\mathbb{R}_+} 2\left(e^{-v} - e^{-2v}\right)\, d\mu_L(v) \\ &= \begin{cases} \int_{[0,z]} 2\left(e^{-v} - e^{-2v}\right)\, d\mu_L(v) = \int_0^z 2\left(e^{-v} - e^{-2v}\right)\, dv & \text{if } z > 0 \\ 0 & \text{if } z \leq 0 \end{cases}, \end{aligned}$$

that is

$$F_Z(z) = \left(1 - 2e^{-z} + e^{-2z}\right) 1_{\mathbb{R}_+}(z)$$

for every $z \in \mathbb{R}$. This proves that $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ is absolutely continuous. □

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2019-11-16

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space, let X and Y be independent standard Rademacher random variables*[1] *on Ω. Set $Z \stackrel{def}{=} X + Y$.*

1. *Compute $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$.*
2. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ uncorrelated?*
3. *Are the random variables $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with standard Rademacher random variables, can you compute $\mathbf{E}\left[(X+Y)^2 \mid Z\right]$ and $\mathbf{E}[XY \mid Z]$?*

Solution. Since X and Y be independent standard Rademacher random variables, we have

$$(X+Y)(\omega) = \begin{cases} -2, & \text{if } \omega \in \{X = -1, Y = -1\}, \\ 0, & \text{if } \omega \in \{X = -1, Y = 1\} \cup \{X = 1, Y = -1\}, \\ 2, & \text{if } \omega \in \{X = 1, Y = 1\}. \end{cases}$$

That is to say,

$$X + Y = -2 \cdot 1_{\{X=-1,Y=-1\}} + 2 \cdot 1_{\{X=1,Y=1\}}$$

Furthermore,

$$\mathbf{P}(X+Y=-2) = \mathbf{P}(X=-1, Y=-1) = \mathbf{P}(X=-1)\,\mathbf{P}(Y=-1) = \frac{1}{4},$$

$$\mathbf{P}(X+Y=2) = \mathbf{P}(X=1, Y=1) = \mathbf{P}(X=1)\,\mathbf{P}(Y=1) = \frac{1}{4},$$

and

$$\begin{aligned} \mathbf{P}(X+Y=0) &= \mathbf{P}(\{X=-1, Y=1\} \cup \{X=1, Y=-1\}) \\ &= \mathbf{P}(X=-1, Y=1) + \mathbf{P}(X=1, Y=-1) \\ &= \mathbf{P}(X=-1)\,\mathbf{P}(Y=1) + \mathbf{P}(X=1)\,\mathbf{P}(Y=-1) \\ &= \frac{1}{2}. \end{aligned}$$

1. Since Z is a discrete random variable, to compute $\mathbf{E}[X \mid Z]$ we can apply the formula

$$\mathbf{E}[X \mid Z] = \mathbf{E}[X \mid Z=-2]\, 1_{\{Z=-2\}} + \mathbf{E}[X \mid Z=2]\, 1_{\{Z=2\}},$$

[1] A standard Rademacher random variable R is given by

$$R \stackrel{\text{def}}{=} \begin{cases} 1, & \mathbf{P}(R=1) = 1/2, \\ -1, & \mathbf{P}(R=-1) = 1/2. \end{cases}$$

where

$$\begin{aligned}\mathbf{E}[X \mid Z=-2] &= \frac{1}{\mathbf{P}(Z=-2)}\int_{\{Z=2\}} X\,d\mathbf{P} = 4\int_{\{X=-1,Y=-1\}} X\,d\mathbf{P} \\ &= -4\int_{\{X=-1,Y=-1\}} d\mathbf{P} = -4\mathbf{P}(X=-1,Y=-1) = -1.\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}[X \mid Z=2] &= \frac{1}{\mathbf{P}(Z=2)}\int_{\{Z=2\}} X\,d\mathbf{P} = 4\int_{\{X=1,Y=1\}} X\,d\mathbf{P} \\ &= 4\int_{\{X=1,Y=1\}} d\mathbf{P} = 4\mathbf{P}(X=1,Y=1) = 1.\end{aligned}$$

It follows

$$\mathbf{E}[X \mid Z] = -1_{\{Z=-2\}} + 1_{\{Z=2\}} = \frac{1}{2}Z.$$

Similarly,

$$\mathbf{E}[Y \mid Z] = -1_{\{Z=-2\}} + 1_{\{Z=2\}} = \frac{1}{2}Z.$$

Another argument, based on the properties of the conditional expectation, is the following. Observe that

$$Z = \mathbf{E}[Z \mid Z] = \mathbf{E}[X+Y \mid Z] = \mathbf{E}[X \mid Z] + \mathbf{E}[Y \mid Z].$$

On the other hand, we know that

$$\mathbf{E}[X \mid Z] = g_X(Z) \quad \text{and} \quad \mathbf{E}[Y \mid Z] = g_Y(Z)$$

where $g_X : \mathbb{R} \to \mathbb{R}$ and $g_X : \mathbb{R} \to \mathbb{R}$ are suitable Borel functions. The structure of the function $g_X : \mathbb{R} \to \mathbb{R}$ [resp. $g_Y : \mathbb{R} \to \mathbb{R}$] depends on the joint distribution of X and Z [resp. Y and Z] and on the distribution of Z. However, in our case, it is not difficult to prove that

$$F_{X,Z}(u,z) = F_{Y,Z}(u,z),$$

for every $(u,z) \in \mathbb{R}^2$. In fact,

$$\begin{aligned}&F_{X,Z}(u,z) \\ &= \mathbf{P}(X \le u, Z \le z) \\ &= \mathbf{P}(X \le u, X+Y \le z) \\ &= \mathbf{P}(X \le u, X+Y \le z, X=1) + \mathbf{P}(X \le u, X+Y \le z, X=-1) \\ &= \mathbf{P}(X \le u, X+Y \le z \mid X=1)\mathbf{P}(X=1) + \mathbf{P}(X \le u, X+Y \le z \mid X=-1)\mathbf{P}(X=-1) \\ &= \frac{1}{2}(\mathbf{P}(X \le u, X+Y \le z \mid X=1) + \mathbf{P}(X \le u, X+Y \le z \mid X=-1)) \\ &= \frac{1}{2}(\mathbf{P}(1 \le u, 1+Y \le z \mid X=1) + \mathbf{P}(-1 \le u, -1+Y \le z \mid X=-1)) \\ &= \frac{1}{2}(\mathbf{P}(1 \le u, Y \le z-1 \mid X=1) + \mathbf{P}(-1 \le u, Y \le z+1 \mid X=-1)) \\ &= \begin{cases} 0, & \text{if } u < -1, \\ \frac{1}{2}\mathbf{P}(Y \le z+1 \mid X=-1) = \frac{1}{2}\mathbf{P}(Y \le z+1), & \text{if } -1 \le u < 1, \\ \frac{1}{2}(\mathbf{P}(Y \le z-1 \mid X=1) + \mathbf{P}(Y \le z+1 \mid X=-1)) = \frac{1}{2}(\mathbf{P}(Y \le z-1) + \mathbf{P}(Y \le z+1)), & \text{if } 1 \le u. \end{cases}\end{aligned}$$

Similarly,

$$\begin{aligned}&F_{Y,Z}(u,z) \\ &= \begin{cases} 0, & \text{if } u < -1, \\ \frac{1}{2}\mathbf{P}(X \le z+1 \mid Y=-1) = \frac{1}{2}\mathbf{P}(X \le z+1), & \text{if } -1 \le u < 1, \\ \frac{1}{2}(\mathbf{P}(X \le z-1 \mid Y=1) + \mathbf{P}(X \le z+1 \mid Y=-1)) = \frac{1}{2}(\mathbf{P}(X \le z-1) + \mathbf{P}(X \le z+1)), & \text{if } 1 \le u. \end{cases}\end{aligned}$$

Therefore, on account that X and Y have the same distribution, we obtain the desired result. As a consequence, we can asses that

$$g_X = g_Y,$$

which implies

$$\mathbf{E}[X \mid Z] = \mathbf{E}[Y \mid Z].$$

It then follows

$$2\mathbf{E}[X \mid Z] = 2\mathbf{E}[Y \mid Z] = Z,$$

which yields

$$\mathbf{E}[X \mid Z] = \mathbf{E}[Y \mid Z] = \frac{1}{2}Z,$$

as expected.

2. Thanks to what shown above, we have

$$\mathbf{E}[X \mid Z]\,\mathbf{E}[Y \mid Z] = \frac{1}{4}Z^2 \sim Dir\left(\frac{1}{4}\right).$$

Hence,

$$\mathbf{E}[\mathbf{E}[X \mid Z]\,\mathbf{E}[Y \mid Z]] = \frac{1}{4}.$$

On the other hand,

$$\mathbf{E}[\mathbf{E}[X \mid Z]] = \mathbf{E}[\mathbf{E}[Y \mid Z]] = \mathbf{E}\left[\frac{1}{2}Z\right] = \frac{1}{2}\mathbf{E}[Z] = 0.$$

Hence, $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ are not uncorrelated.

3. Since $\mathbf{E}[X \mid Z]$ and $\mathbf{E}[Y \mid Z]$ are not not uncorrelated, they cannot be independent.

4. By virtue of the properties of the conditional expectation, we have

$$\mathbf{E}\left[(X+Y)^2 \mid Z\right] = \mathbf{E}\left[Z^2 \mid Z\right] = Z^2.$$

On the other hand,

$$\begin{aligned}\mathbf{E}\left[(X+Y)^2 \mid Z\right] &= \mathbf{E}\left[X^2 + 2XY + Y^2 \mid Z\right] \\ &= \mathbf{E}\left[X^2 \mid Z\right] + 2\mathbf{E}[XY \mid Z] + \mathbf{E}\left[Y^2 \mid Z\right].\end{aligned}$$

Now, since $X \sim Y \sim Rad(1/2)$, we have $X^2 \sim Y^2 \sim Dir(1)$. We then obtain

$$Z^2 = \mathbf{E}\left[(X+Y)^2 \mid Z\right] = \mathbf{E}[1 \mid Z] + 2\mathbf{E}[XY \mid Z] + \mathbf{E}[1 \mid Z] = 2 + 2\mathbf{E}[XY \mid Z].$$

The latter yields

$$\mathbf{E}[XY \mid Z] = \frac{1}{2}Z^2 - 1.$$

Problem 2 *Let X [resp. R] be a standard Gaussian [Rademacher] random variable on a probability space Ω. In symbols, $X \sim N(0,1)$ and $R \sim Rad(1/2)$. Assume that X and R are independent and define $Y \equiv R \cdot X$.*

1. *Is the random variable Y Gaussian?*

2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*

3. *Are the random variables R and Y uncorrelated? Are R and Y independent?*

4. *Does the random vector $(X,Y)^\intercal$ have a bivariate Gaussian distribution? Hint: consider the possibility that $(X,Y)^\intercal$ has a bivariate Gaussian distribution; how the random variable $Z \equiv X+Y$ should be distributed?*

5. *Can you compute $\mathbf{E}[Y \mid X]$ and $\mathbf{E}[X \mid Y]$?*

Solution.

1. To prove that Y is Gaussian we show that

$$\mathbf{P}(Y \leq y) = \mathbf{P}(X \leq y), \tag{1}$$

for every $y \in \mathbb{R}$. To this, on account that $\{R=1\}$, $\{R=-1\}$ constitute a partition of Ω, the random variables R and X are independent and X is symmetric about 0, we can write

$$\begin{aligned}
\mathbf{P}(Y \leq y) &= \mathbf{P}(RX \leq y) \\
&= \mathbf{P}(RX \leq y, R=1) + \mathbf{P}(RX \leq y, R=-1) \\
&= \mathbf{P}(RX \leq y \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(RX \leq y \mid R=-1)\,\mathbf{P}(R=-1) \\
&= \frac{1}{2}\left(\mathbf{P}(X \leq y, \mid R=1) + \mathbf{P}(X \geq -y \mid R=-1)\right) \\
&= \frac{1}{2}\left(\mathbf{P}(X \leq y,) + \mathbf{P}(X \geq -y)\right) \\
&= \mathbf{P}(X \leq y),
\end{aligned}$$

for every $y \in \mathbb{R}$. This proves that $Y \sim X \sim N(0,1)$.

2. Since $Y \equiv R \cdot X$, the intuition is that the observation of the values taken by X transmits information on the values taken by Y. That is X and Y are not independent. However, on account that $R^2 \sim Dirac(1)$, thanks to the independence of X and R, we have

$$\mathbf{E}[XY] = \mathbf{E}[XRX] = \mathbf{E}\left[R^2X\right] = \mathbf{E}[X] = 0 = \mathbf{E}[X]\,\mathbf{E}[R].$$

This shows that X and Y are uncorrelated. On the other hand, since $X \sim N(0,1)$, we have

$$\mathbf{E}\left[X^2Y^2\right] = \mathbf{E}\left[X^2R^2X^2\right] = \mathbf{E}\left[X^4\right] = 3$$

and

$$\mathbf{E}\left[X^2\right]\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[X^2\right]\mathbf{E}\left[R^2X^2\right] = \mathbf{E}\left[X^2\right]\mathbf{E}\left[X^2\right] = \mathbf{E}\left[X^2\right]^2 = 1.$$

This shows that X^2 and Y^2 are not uncorrelated, which prevents that X^2 and Y^2 are not independent. Eventually, X and Y cannot be independent.

3. Still on account that $R^2 \sim Dirac(1)$, we have

$$\mathbf{E}[RY] = \mathbf{E}[RRX] = \mathbf{E}\left[R^2X\right] = \mathbf{E}[X] = 0 = \mathbf{E}[X]\,\mathbf{E}[R].$$

This shows that R and Y are uncorrelated. On the other hand, since $Y \equiv R \cdot X \sim N(0,1)$ the intuition is that the observation of the values taken by R transmits no information on the values taken by Y. Hence, the intuition is that R and Y are independent. To prove this, we show that

$$\mathbf{P}(R \leq r, Y \leq y) = \mathbf{P}(R \leq r)\,\mathbf{P}(Y \leq y), \tag{2}$$

for all $r, y \in \mathbb{R}$. In fact, still on account that $\{R=1\}$, $\{R=-1\}$ constitute a partition of Ω, the random variables R and X are independent and X is symmetric about 0, we have

$$\begin{aligned}
&\mathbf{P}(R \leq r, Y \leq y,) \\
&= \mathbf{P}(R \leq r, Y \leq y, R=1) + \mathbf{P}(R \leq r, Y \leq y, R=-1) \\
&= \mathbf{P}(R \leq r, XR \leq y, R=1) + \mathbf{P}(R \leq r, XR \leq y, R=-1) \\
&= \mathbf{P}(R \leq r, XR \leq y \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(R \leq r, XR \leq y \mid R=-1)\,\mathbf{P}(R=-1) \\
&= \tfrac{1}{2}\left(\mathbf{P}(1 \leq r, X \leq y \mid R=1) + \mathbf{P}(-1 \leq r, X \geq -y \mid R=-1)\right) \\
&= \begin{cases} 0 & \text{if } r < -1 \\ \frac{1}{2}\mathbf{P}(X \geq -y \mid R=-1) = \frac{1}{2}\mathbf{P}(X \geq -y) = \frac{1}{2}\mathbf{P}(X \leq y) & \text{if } -1 \leq r < 1 \\ \frac{1}{2}\left(\mathbf{P}(X \leq y \mid R=1) + \mathbf{P}(X \geq -y \mid R=-1)\right) = \frac{1}{2}\left(\mathbf{P}(X \leq y) + \mathbf{P}(X \geq -y)\right) = \mathbf{P}(X \leq y) & \text{if } 1 \leq r \end{cases}
\end{aligned}$$

On the other hand

$$\mathbf{P}(R \leq r)\,\mathbf{P}(Y \leq y) = \begin{cases} 0 & \text{if } r < -1 \\ \frac{1}{2}\mathbf{P}(Y \leq y) = \frac{1}{2}\mathbf{P}(X \leq y) & \text{if } -1 \leq r < 1 \\ \mathbf{P}(Y \leq y) = \mathbf{P}(X \leq y) & \text{if } 1 \leq r \end{cases}$$

Therefore, the random variables R and Y are independent.

4. If the random vector $(X,Y)^\intercal$ had a bivariate Gaussian distribution, the random variable $Z \equiv X+Y$ should be Gaussian distributed. On the other hand,

$$Z = X + Y = X + RX = (R+1)X.$$

Hence,

$$\begin{aligned}
F_Z(x) &= \mathbf{P}(Z \leq z) \\
&= \mathbf{P}(Z \leq z, R=1) + \mathbf{P}(Z \leq z, R=-1) \\
&= \mathbf{P}(Z \leq z \mid R=1)\,\mathbf{P}(R=1) + \mathbf{P}(Z \leq z \mid R=-1)\,\mathbf{P}(R=-1) \\
&= \tfrac{1}{2}\left(\mathbf{P}((R+1)X \leq z \mid R=1) + \mathbf{P}((R+1)X \leq z \mid R=-1)\right) \\
&= \tfrac{1}{2}\left(\mathbf{P}(2X \leq z \mid R=1) + \mathbf{P}(0 \leq z \mid R=-1)\right).
\end{aligned}$$

Now, we have that the events

$$\{2X \leq z\} \quad \text{and} \quad \{R=1\}$$

are independent. Moreover,

$$\begin{aligned} \{0 \leq z\} = \Omega \quad &\text{if } z \geq 0 \\ \{0 \leq z\} = \varnothing \quad &\text{if } z < 0 \end{aligned}$$

Hence,

$$F_Z(x) = \begin{cases} \frac{1}{2}\mathbf{P}(2X \leq z) & \text{if } z < 0 \\ \frac{1}{2}\left(\mathbf{P}(2X \leq z) + 1\right) & \text{if } z \geq 0 \end{cases}$$

in particular, if $z < 0$, we have

$$F_Z(x) \leq \frac{1}{2}\mathbf{P}(2X \leq 0) = \frac{1}{4}$$

and if $z \geq 0$

$$F_Z(x) \geq \frac{1}{2}\left(\mathbf{P}(2X \leq 0) + 1\right) = \frac{1}{2}\left(\frac{1}{2} + 1\right) = \frac{3}{4},$$

Hence, F_Z cannot not continuous at $z=0$. This prevents Z to be Gaussian.

5. By virtue of what shown above and the properties of the conditional expectation, we have,

$$\mathbf{E}[Y \mid X] = \mathbf{E}[RX \mid X] = X\mathbf{E}[R \mid X] = X\mathbf{E}[R] = 0$$

and

$$\mathbf{E}[X \mid Y] = \mathbf{E}\left[XR^2 \mid Y\right] = \mathbf{E}[XRR \mid Y] = \mathbf{E}[YR \mid Y] = Y\mathbf{E}[R \mid Y] = Y\mathbf{E}[R] = 0.$$

Problem 3 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra. Let $X, Y \in L^2(\Omega; \mathbb{R})$.*

1. *Prove in all details that $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω implies $Cov(X, Y) = 0$, but X and Y may not be independent.*

2. *Prove in all details that $Cov(X, Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$.*

Exercise 4 *Hint: in the first case, to generate a suitable counterexample one may consider the random variables $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$. In the second case consider $X \sim N(0,1)$ and $Y = X^2$.*

Solution.

1. Under the assumption $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$ a.e. on Ω, by virtue of the properties of the conditional expectation operator, we can write

$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[XY \mid X]] = \mathbf{E}[X\mathbf{E}[Y \mid X]] = \mathbf{E}[X\mathbf{E}[Y]] = \mathbf{E}[X]\mathbf{E}[Y]$$

Therefore,

$$Cov(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = 0.$$

Now, if we consider the random $X \sim Ber(p)$, $Z \sim N(0,1)$, independent of X, and $Y = XZ$, we have

$$\begin{aligned} Cov(X, Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \mathbf{E}\left[X^2 Z\right] - \mathbf{E}[X]\mathbf{E}[XZ] \\ &= \mathbf{E}\left[X^2\right]\mathbf{E}[Z] - \mathbf{E}[X]^2\mathbf{E}[Z] \\ &= 0. \end{aligned}$$

On the other hand, we have

$$\mathbf{P}(X \leq 0) = q,$$

and, on account that X and Z are independent,

$$\begin{aligned} \mathbf{P}(Y \leq 0) &= \mathbf{P}(XZ \leq 0) \\ &= \mathbf{P}(XZ \leq 0, X = 0) + \mathbf{P}(XZ \leq 0, X = 1) \\ &= \mathbf{P}(XZ \leq 0 \mid X = 0)\mathbf{P}(X = 0) + \mathbf{P}(XZ \leq 0 \mid X = 1)\mathbf{P}(X = 1) \\ &= \mathbf{P}(0 \leq 0 \mid X = 0)\mathbf{P}(X = 0) + \mathbf{P}(Z \leq 0 \mid X = 1)\mathbf{P}(X = 1) \\ &= \mathbf{P}(\Omega)\mathbf{P}(X = 0) + \mathbf{P}(Z \leq 0)\mathbf{P}(X = 1) \\ &= q + \frac{1}{2}p. \end{aligned}$$

Furthermore, the same arguments as above shows that

$$\begin{aligned} \mathbf{P}(X \leq 0, Y \leq 0) &= \mathbf{P}(X \leq 0, XZ \leq 0) \\ &= \mathbf{P}(X = 0, XZ \leq 0) \\ &= \mathbf{P}(XZ \leq 0 \mid X = 0)\mathbf{P}(X = 0) \\ &= q. \end{aligned}$$

Hence, we have

$$\mathbf{P}(X \leq 0)\mathbf{P}(Y \leq 0) = q\left(q + \frac{1}{2}p\right) \neq q = \mathbf{P}(X \leq 0, Y \leq 0)$$

which shows that X and Y are not be independent.

2. To show that $Cov(X, Y) = 0$ does not imply $\mathbf{E}[Y \mid X] = \mathbf{E}[Y]$, we consider $X \sim N(0,1)$ and $Y = X^2$. We have

$$\begin{aligned} Cov(X, Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] \\ &= \mathbf{E}\left[X^3\right] - \mathbf{E}[X]\mathbf{E}\left[X^2\right] \\ &= 0, \end{aligned}$$

but

$$\mathbf{E}[Y \mid X] = \mathbf{E}\left[X^2 \mid X\right] = X^2 \neq \mathbf{E}\left[X^2\right] = \mathbf{E}[Y].$$

Problem 5 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{E}$, and let X, Y be real random variables on Ω with finite second order moment.*

1. *Show that*

$$\mathbf{E}\left[(X - \mathbf{E}[X \mid \mathcal{F}])^2\right] \leq \mathbf{E}\left[(X - \mathbf{E}[X])^2\right].$$

2. *Show that*

$$\mathbf{E}[XY \mid \mathcal{F}]^2 \leq \mathbf{E}\left[X^2 \mid \mathcal{F}\right]\mathbf{E}\left[Y^2 \mid \mathcal{F}\right]. \tag{3}$$

Solution.

1. In the space $\mathcal{L}^2(\Omega_{\mathcal{F}}; \mathbb{R})$ of the real $\mathcal{F}$-random variables having fnite moment of the second order the conditional expectation of X given $\mathcal{F}$ is characterized as

$$\mathbf{E}[X \mid \mathcal{F}] = \underset{Y \in \mathcal{L}^2(\Omega_{\mathcal{F}}; \mathbb{R})}{\arg\min} \mathbf{E}\left[(X - Y)^2\right].$$

This means that

$$\mathbf{E}\left[(X - \mathbf{E}[X \mid \mathcal{F}])^2\right] \leq \mathbf{E}\left[(X - Y)^2\right],$$

for every $Y \in \mathcal{L}^2(\Omega_{\mathcal{F}}; \mathbb{R})$. In particular, since the deterministic random variable $\mathbf{E}[X] \equiv \mathbf{E}[X]\ 1_{\Omega}$ is clearly in $\mathcal{L}^2(\Omega_{\mathcal{F}}; \mathbb{R})$, setting $Y \equiv \mathbf{E}[X]$ we obtain the desired inequality.

2. Note first that for all real random variables X, Y on Ω we have

$$|XY| \leq \frac{1}{2}\left(X^2 + Y^2\right).$$

Therefore, the assumption that X and Y have finite second moment implies that XY has finite first order moment. Hence, both the sides of (3) are well defined. Now, given any $z \in \mathbb{R}$, the random variable $X + zY$ has finite second order moment and, by virtue of the positivity of the conditional expectation operator, we have

$$\mathbf{E}\left[(X+zY)^2 \mid \mathcal{F}\right] \geq 0.$$

On the other hand, the linearity of the conditional expectation operator implies

$$\mathbf{E}\left[(X+zY)^2 \mid \mathcal{F}\right] = \mathbf{E}\left[X^2 \mid \mathcal{F}\right] + 2z\mathbf{E}\left[XY \mid \mathcal{F}\right] + z^2\mathbf{E}\left[Y^2 \mid \mathcal{F}\right].$$

As a consequence, we can write

$$\mathbf{E}\left[X^2 \mid \mathcal{F}\right] + 2z\mathbf{E}\left[XY \mid \mathcal{F}\right] + z^2\mathbf{E}\left[Y^2 \mid \mathcal{F}\right] \geq 0$$

for every $z \in \mathbb{R}$. It follows that

$$\Delta \equiv \mathbf{E}\left[XY \mid \mathcal{F}\right]^2 - \mathbf{E}\left[X^2 \mid \mathcal{F}\right]\mathbf{E}\left[Y^2 \mid \mathcal{F}\right] \leq 0,$$

which is the desired (3).

Problem 6 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{n\geq 1}$ be a sequence of independent and identically distributed real random variables. Consider the sample sum*

$$Z_n \stackrel{def}{=} \textstyle\sum_{k=1}^n X_k.$$

1. *Determine $\mathbf{E}[X_n \mid X_m]$ for every $m, n \geq 1$.*
2. *Determine $\mathbf{E}[Z_n \mid X_m]$ for every $n \geq 1$ and $m \leq n$.*
3. *Determine $\mathbf{E}[Z_n \mid Z_m]$ for every $n \geq 1$ and $m \leq n$.*
4. *Compute $\mathbf{E}[Z_n]$ and $\mathbf{D}[Z_n]$, for every $n \geq 1$.*
5. *Assume $X_n \sim Ber(p)$, for some $p \in (0,1)$, determine the distribution of Z_n.*
6. *Assume $X_n \sim N(\mu, \sigma^2)$, for some $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{++}$, determine the distribution of Z_n.*
7. *Assume $X \sim Exp(\lambda)$, for some $\lambda > 0$, determine the distribution of Z_n.*

Solution.

1. Since the random variables of the sequence $(X_n)_{n\geq 1}$ are independent, we have
$$\mathbf{E}[X_n \mid X_m] = \begin{cases} \mu & \text{if } n \neq m \\ X_n & \text{if } n = m \end{cases},$$
where μ is the value of the mean of the identically distributed random variables of the sequence $(X_n)_{n\geq 1}$.

2. By virtue of the linearity of the conditional expectation operator, we have
$$\begin{aligned}\mathbf{E}[Z_n \mid X_m] &= \mathbf{E}\left[\textstyle\sum_{k=1}^n X_k \mid X_m\right] = \textstyle\sum_{k=1}^n \mathbf{E}[X_k \mid X_m] \\ &= \textstyle\sum_{k=1,k\neq m}^n \mathbf{E}[X_k \mid X_m] + \mathbf{E}[X_m \mid X_m] \\ &= \textstyle\sum_{k=1,k\neq m}^n \mathbf{E}[X_k] + X_m \\ &= \textstyle\sum_{k=1,k\neq m}^n \mu + X_m \\ &= \mu\textstyle\sum_{k=1,k\neq m}^n 1 + X_m \\ &= (n-1)\mu + X_m,\end{aligned}$$
where μ is the value of the mean of the identically distributed random variables of the sequence $(X_n)_{n\geq 1}$.

3. In case $m = n$, we have trivially
$$\mathbf{E}[Z_n \mid Z_m] = \mathbf{E}[Z_n \mid Z_n] = Z_n.$$
In case $m < n$, by virtue of the linearity of the conditional expectation operator, we can write
$$\begin{aligned}\mathbf{E}[Z_n \mid Z_m] &= \mathbf{E}\left[\textstyle\sum_{k=1}^n X_k \mid \sum_{k=1}^m X_k\right] \\ &= \mathbf{E}\left[\textstyle\sum_{k=1}^m X_k + \sum_{k=m+1}^n X_k \mid \sum_{k=1}^m X_k\right] \\ &= \mathbf{E}\left[\textstyle\sum_{k=1}^m X_k \mid \sum_{k=1}^m X_k\right] + \mathbf{E}\left[\sum_{k=m+1}^n X_k \mid \sum_{k=1}^m X_k\right] \\ &= \mathbf{E}[Z_m \mid Z_m] \\ &= Z_m + \textstyle\sum_{k=m+1}^n \mathbf{E}\left[X_k \mid \sum_{k=1}^m X_k\right].\end{aligned}$$
On the other hand the independence of the random variables of the sequence $(X_n)_{n\geq 1}$ implies that each X_n is independent of $f(X_1, \ldots, X_{n-1})$ for every $n \geq 1$, where $f : \mathbb{R} \to \mathbb{R}_+$ is any Borel function. It follows that
$$\textstyle\sum_{k=m+1}^n \mathbf{E}\left[X_k \mid \sum_{k=1}^m X_k\right] = \sum_{k=m+1}^n \mathbf{E}[X_k] = \sum_{k=m+1}^n \mu = \mu\sum_{k=m+1}^n 1 = (n-m)\mu.$$
As a consequence,
$$\mathbf{E}[Z_n \mid Z_m] = Z_m + (n-m)\mu.$$

4. Writing σ^2 for the alue of the variance of the identically distributed random variables of the sequence $(X_n)_{n\geq 1}$, thanks to the linearity of the operator expectation and the linearity of the operator variance on independent random variables, we have
$$\mathbf{E}[Z_n] = \mathbf{E}\left[\textstyle\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \mathbf{E}[X_k] = \sum_{k=1}^n \mu = \mu\sum_{k=1}^n 1 = n\mu$$
and
$$\mathbf{D}^2[Z_n] = \mathbf{D}^2\left[\textstyle\sum_{k=1}^n X_k\right] = \sum_{k=1}^n \mathbf{D}^2[X_k] = \sum_{k=1}^n \sigma^2 = \sigma^2\sum_{k=1}^n 1 = n\sigma^2.$$

5. Under the assumption $X_n \sim Ber(p)$, for some $p \in (0,1)$, it is well known (see sum of independent and Bernoulli distributed random variables e.g. sec. Statistics on Simple Random Samples of Notes) that $Z_n \sim Bin(n, p)$.

6. Under the assumption $X_n \sim N(\mu, \sigma^2)$, for some $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{++}$, it is well known (see sum of independent and Gaussian distributed random variables e.g. sec. Statistics on Simple Random Samples of Notes) that $Z_n \sim N(n\mu, n\sigma^2)$.

7. Under the assumption $X \sim Exp(\lambda)$, for some $\lambda > 0$, it is well known (see sum of independent and exponentially distributed random variables e.g. sec. Statistics on Simple Random Samples of Notes) that $Z_n \sim \Gamma(n, \lambda)$.

Problem 7 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y be independent standard normally distributed random variables on Ω. Set*

$$U \overset{def}{=} X + Y, \qquad V \overset{def}{=} X - Y.$$

1. *Compute the distributions of U and V.*
2. *Prove that U and V are independent.*
3. *Compute $\mathbf{E}[X \mid U]$, $\mathbf{E}[X \mid V]$, $\mathbf{E}[Y \mid U]$, $\mathbf{E}[Y \mid V]$.*
4. *Compute $\mathbf{E}[XY \mid U]$.*

Exercise 8 *Hint: First, concentrate your attention on the circumstance that X and Y are independent and standard normally distributed. Second, it might be useful to consider $\mathbf{E}\left[X^2 \mid U\right]$ and $\mathbf{E}\left[Y^2 \mid U\right]$.*

Solution.

1. Since X and Y are independent standard normally distributed random variables, by virtue of Proposition 631 of Notes U and V are normally distributed with mean

$$\mu_U = \mu_V = \mu_X + \mu_Y = 0$$

and variance

$$\sigma_U^2 = \sigma_V^2 = \sigma_X^2 + \sigma_Y^2 = 1.$$

2. We can write

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Therefore we can apply Theorem 642 of Notes to obtain that the random variables U and V are jointly Gaussian distributed. Now, since we have

$$\begin{aligned} Cov\,(U, V) &= Cov\,(X + Y, X - Y) = Cov\,(X, X) - Cov\,(X, Y) + Cov\,(Y, X) - Cov\,(Y, Y) \\ &= \mathbf{D}^2[X] - Cov\,(X, Y) + Cov\,(X, Y) - \mathbf{D}^2[Y] \\ &= 0 \end{aligned}$$

We can apply Proposition 652 of Notes to obtain that the random variables U and V are independent.

3. Observe that, by virtue of the properties of the conditional expectation operator, we have

$$U = \mathbf{E}[U \mid U] = \mathbf{E}[X + Y \mid U] = \mathbf{E}[X \mid U] + \mathbf{E}[Y \mid U]$$

and

$$V = \mathbf{E}[U \mid U] = \mathbf{E}[X - Y \mid U] = \mathbf{E}[X \mid V] - \mathbf{E}[Y \mid V]$$

Therefore, to compute the desired conditional expectations it is clearly sufficient to compute $\mathbf{E}[X \mid U]$ and $\mathbf{E}[X \mid V]$. On the other hand, we have

$$\begin{pmatrix} X \\ U \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} X \\ V \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

This, still on account of Theorem 642 of Notes, proves that both the vectors (X, U) and (X, V) are Gaussian. It is then possible to apply Corollary 704 of Notes to obtain that

$$\mathbf{E}[X \mid U] = \mathbf{E}[X] + \frac{Cov(X, U)}{\mathbf{D}^2[U]}(U - \mathbf{E}[U]) = \frac{1}{2}U$$

and

$$\mathbf{E}[X \mid V] = \mathbf{E}[X] + \frac{Cov(X, V)}{\mathbf{D}^2[V]}(V - \mathbf{E}[V]) = \frac{1}{2}V$$

since

$$Cov(X, U) = Cov(X, X + Y) = Cov(X, X) + Cov(X, Y) = \mathbf{D}^2[X] = 1$$

and

$$Cov(X, V) = Cov(X, X - Y) = Cov(X, X) - Cov(X, Y) = \mathbf{D}^2[X] = 1$$

It follows that

$$\mathbf{E}[Y \mid U] = U - \mathbf{E}[X \mid U] = \frac{1}{2}U$$

and

$$\mathbf{E}[Y \mid V] = \mathbf{E}[X \mid V] - V = -\frac{1}{2}V.$$

4. To compute $\mathbf{E}[XY \mid U]$, observe that, by virtue of the properties of the conditional expectation operator, we can write

$$U^2 = \mathbf{E}\left[U^2 \mid U\right] = \mathbf{E}\left[(X + Y)^2 \mid U\right] = \mathbf{E}\left[X^2 \mid U\right] + 2\mathbf{E}[XY \mid U] + \mathbf{E}\left[Y^2 \mid U\right].$$

It follows

$$\mathbf{E}[XY \mid U] = U^2 - \frac{1}{2}\left(\mathbf{E}\left[X^2 \mid U\right] + \mathbf{E}\left[Y^2 \mid U\right]\right).$$

We already know that the vector (X, U) is Gaussian and likewise also the vector (Y, U) is Gaussian. Therefore, we can compute $\mathbf{E}\left[X^2 \mid U\right]$ and $\mathbf{E}\left[Y^2 \mid U\right]$ applying again Corollary 704. We obtain

$$\mathbf{E}\left[X^2 \mid U\right] = \mathbf{D}^2[X] - \frac{Cov(X, U)^2}{\mathbf{D}^2[U]} + \left(\mathbf{E}[X] + \frac{Cov(X, U)}{\mathbf{D}^2[U]}(U - \mathbf{E}[U])\right)^2 = \frac{1}{2}\left(1 + \frac{1}{2}U^2\right)$$

and

$$\mathbf{E}\left[Y^2 \mid U\right] = \mathbf{D}^2[Y] - \frac{Cov(Y, U)^2}{\mathbf{D}^2[U]} + \left(\mathbf{E}[Y] + \frac{Cov(Y, U)}{\mathbf{D}^2[U]}(U - \mathbf{E}[U])\right)^2 = \frac{1}{2}\left(1 + \frac{1}{2}U^2\right)$$

since

$$Cov(Y, U) = Cov(Y, X + Y) = Cov(Y, X) + Cov(Y, Y) = \mathbf{D}^2[Y] = 1.$$

In the end

$$\begin{aligned} \mathbf{E}[XY \mid U] &= U^2 - \frac{1}{2}\left(\mathbf{E}\left[X^2 \mid U\right] + \mathbf{E}\left[Y^2 \mid U\right]\right) \\ &= U^2 - \frac{1}{2}\left(1 + \frac{1}{2}U^2\right) \\ &= \frac{1}{2}\left(\frac{3}{2}U^2 - 1\right). \end{aligned}$$

Problem 9 *Let N be a geometric random variable with success probability p, which models the first occurrence of success in n independent trials, and let $(X_n)_{n\geq 1}$ be a sequence of independent and normally distributed random variables with mean μ and variance σ^2, which are also independent of N. Study the conditional expectation*

$$\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right].$$

Use the properties of the conditional expectation to compute the expectation and the variance of the random sum

$$S_N \overset{def}{=} \sum_{k=1}^{N} X_k.$$

Solution. Since the random variables of the sequence $(X_n)_{n\geq 1}$ are independent and are also independent of N, which is geometrically distributed, we can write

$$\begin{aligned}
\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right] &= \sum_{n=1}^{\infty} \mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N=n\right] 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega}\left(\sum_{k=1}^{N} X_k\right) 1_{\{N=n\}}\, d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}}\left(\sum_{k=1}^{N} X_k\right) d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}}\left(\sum_{k=1}^{n} X_k\right) d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega}\left(\sum_{k=1}^{n} X_k\right) 1_{\{N=n\}}\, d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right) 1_{\{N=n\}}\right]\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \left(\sum_{k=1}^{n} \mathbf{E}[X_k]\right) \mathbf{E}\left[1_{\{N=n\}}\right]\right) 1_{\{N=n_m\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \left(\sum_{k=1}^{n} \mu\right) \mathbf{P}(N=n)\right) 1_{\{N=n_m\}} \\
&= \sum_{n=1}^{\infty} n\mu 1_{\{N=n_m\}} \\
&= \mu \sum_{n=1}^{\infty} n 1_{\{N=n\}} \\
&\overset{\mathbf{P}\text{-a.s.}}{=} \mu N.
\end{aligned}$$

Now, we can write

$$\mathbf{E}\left[\sum_{k=1}^{N} X_k\right] = \mathbf{E}\left[\mathbf{E}\left[\sum_{k=1}^{N} X_k \mid N\right]\right] = \mathbf{E}[\mu N] = \mu \mathbf{E}[N] = \frac{\mu}{p}.$$

and

$$\mathbf{D}^2\left[\sum_{k=1}^{N} X_k\right] = \mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2\right] - \mathbf{E}\left[\sum_{k=1}^{N} X_k\right]^2 = \mathbf{E}\left[\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right]\right] - \frac{\mu^2}{p^2}.$$

Thus, we are left with computing

$$\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right].$$

A straightforward computation yields

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right] &= \sum_{n=1}^{\infty} \mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N=n\right] 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega}\left(\sum_{k=1}^{N} X_k\right)^2 1_{\{N=n\}}\, d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}}\left(\sum_{k=1}^{N} X_k\right)^2 d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\{N=n\}}\left(\sum_{k=1}^{n} X_k\right)^2 d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \int_{\Omega}\left(\sum_{k=1}^{n} X_k\right)^2 1_{\{N=n\}}\, d\mathbf{P}\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2 1_{\{N=n\}}\right]\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] \mathbf{E}\left[1_{\{N=n\}}\right]\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty}\left(\frac{1}{\mathbf{P}(N=n)} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] \mathbf{P}(N=n)\right) 1_{\{N=n\}} \\
&= \sum_{n=1}^{\infty} \mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] 1_{\{N=n\}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right] &= \mathbf{E}\left[\sum_{k=1}^{n} X_k^2 + \sum_{k,\ell=1}^{n} X_k X_\ell\right] \\
&= \sum_{k=1}^{n} \mathbf{E}\left[X_k^2\right] + \sum_{k,\ell=1}^{n} \mathbf{E}\left[X_k\right]\mathbf{E}\left[X_k\right] \\
&= \sum_{k=1}^{n} \left(\mu^2+\sigma^2\right) + \sum_{k,\ell=1}^{n} \mu^2 \\
&= \left(\mu^2+\sigma^2\right) n + \mu^2 (n-1) n \\
&= \sigma^2 n + \mu^2 n^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2 \mid N\right] &= \sum_{n=1}^{\infty} \left(\sigma^2 n + \mu^2 n^2\right) 1_{\{N=n\}} \\
&= \sigma^2 \sum_{n=1}^{\infty} n 1_{\{N=n\}} + \mu^2 \sum_{n=1}^{\infty} n^2 1_{\{N=n\}} \\
&= \sigma^2 N + \mu^2 N^2.
\end{aligned}$$

It then follows

$$\begin{aligned}
\mathbf{E}\left[\sigma^2 N + \mu^2 N^2\right] &= \sigma^2 \mathbf{E}[N] + \mu^2 \mathbf{E}\left[N^2\right] \\
&= \frac{\sigma^2}{p} + \mu^2 \left(\mathbf{D}^2[N] + \mathbf{E}[N]^2\right) \\
&= \frac{\sigma^2}{p} + \mu^2 \left(\frac{2-p}{p^2}\right).
\end{aligned}$$

In the end,

$$\mathbf{D}^2\left[\sum_{k=1}^{N} X_k\right] = \frac{\sigma^2}{p} + \mu^2\left(\frac{2-p}{p^2}\right) - \frac{\mu^2}{p^2} = \frac{\sigma^2}{p} + \mu^2\left(\frac{1-p}{p^2}\right).$$

.

Problem 10 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$ such that*

$$\mathbf{E}[Y \mid X] = X \quad \textit{and} \quad \mathbf{E}\left[Y^2 \mid X\right] = X^2.$$

Prove that $Y = X$, $\mathbf{P}$-a.s. on Ω.

Solution. We have $Y = X$, $\mathbf{P}$-a.s. on Ω if and only if there exists an event $E \in \mathcal{E}$ such that $\mathbf{P}(E) = 0$ and $Y(\omega) = X(\omega)$ for every $\omega \in \Omega - E$. By virtue of the properties of the Lebesgue integral, we have

$$Y = X,\ \mathbf{P}\text{-a.s. on } \Omega \Leftrightarrow \int_\Omega (X-Y)^2 \, d\mathbf{P} = 0.$$

On the other hand,

$$\int_\Omega (X-Y)^2 \, d\mathbf{P} \equiv \mathbf{E}\left[(X-Y)^2\right].$$

Hence, we evaluate

$$\mathbf{E}\left[(X-Y)^2\right] = \mathbf{E}\left[X^2 - 2XY + Y^2\right] = \mathbf{E}\left[X^2\right] - 2\mathbf{E}[XY] + \mathbf{E}\left[Y^2\right]. \tag{4}$$

Now, by virtue of the properties of the conditional expectation operator, under our assumptions, we have

$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[XY \mid X]] = \mathbf{E}[X\mathbf{E}[Y \mid X]] = \mathbf{E}\left[X^2\right] \tag{5}$$

and

$$\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[\mathbf{E}\left[Y^2 \mid X\right]\right] = \mathbf{E}\left[X^2\right]. \tag{6}$$

Combining (4)-(6) it follows

$$\mathbf{E}\left[(X-Y)^2\right] = 0,$$

which yields the deired result.

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2019-11-16

Problem 1 *Let $(X_n)_{n\geq 1}$ a sequence of independent real random variables on a probability space Ω such that*

$$\mathbf{P}(X_n = x) = \begin{cases} 1-\frac{1}{n} & \text{if } x=0 \\ \frac{1}{n} & \text{if } x=\sqrt{n} \\ 0 & \text{otherwise} \end{cases}$$

In the assigned order, check whether the sequence $(X_n)_{n\geq 1}$ converges in distribution, converges in probability, converges almost surely, converges in mean, and converges in square mean.

Solution. According to the definition of the probability distribution, X_n is a Bernoulli random variable on Ω with states 0, $\sqrt{n}$ and success probability $\frac{1}{n}$, for every $n \geq 1$. In symbols,

$$X_n = \begin{cases} 0 & \mathbf{P}(X_n = 0) = 1-\frac{1}{n} \\ \sqrt{n} & \mathbf{P}(X_n = \sqrt{n}) = \frac{1}{n} \end{cases}.$$

Then, considering the distribution function $F_{X_n} : \mathbb{R} \to \mathbb{R}_+$ of X_n, we have

$$F_{X_n}(x) = \begin{cases} 0 & \text{if } x<0 \\ 1-\frac{1}{n} & \text{if } 0 \leq x < \sqrt{n} \\ 1 & \text{if } \sqrt{n} \leq x \end{cases}.$$

As a consequence,

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 & \text{if } x<0 \\ 1 & \text{if } x \geq 0 \end{cases}.$$

In fact, for any $x<0$ we have $F_{X_n}(x) = 0$, for every $n \geq 1$, which clearly implies $\lim_{n\to\infty} F_{X_n}(x) = 0$. On the other hand, for any $x \geq 0$ there clearly exist $n_x \geq 1$ such that $x \leq \sqrt{n}$ for every $n \geq n_x$. It follows, $F_{X_n}(x) = 1-\frac{1}{n}$ for any $n \geq n_x$ and $\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}\left(1-\frac{1}{n}\right) = 1$.

Therefore, the sequence $(F_{X_n})_{n\geq 0}$ of distribution functions converges to the Heaviside function

$$F(x) = \begin{cases} 0 & \text{if } x<0 \\ 1 & \text{if } x \geq 0 \end{cases},$$

for every $x \in \mathbb{R}$. In particular for all $x \in \mathbb{R}$ where F is continuous. This means that the sequence $(X_n)_{n\geq 1}$ converges in distribution to the Dirac random variable concentrated at 0. Write $X \equiv Dir(0)$. Now, for any $0 < \varepsilon < 1$ consider the event $\{|X_n - X| \geq \varepsilon\}$ on varying of $n \geq 1$. We have

$$|X_n - X| = \begin{cases} 0 & \text{if } X_n = 0 \\ \sqrt{n} & \text{if } X_n = \sqrt{n} \end{cases}.$$

Hence,

$$|X_n - X| = X_n.$$

The latter implies

$$\{|X_n - X| \geq \varepsilon\} = \{X_n \geq \varepsilon\} = \{X_n = \sqrt{n}\}.$$

It follows

$$\mathbf{P}(|X_n - X| \geq \varepsilon) = \mathbf{P}(X_n = \sqrt{n}) = \frac{1}{n},$$

for every $n \geq 1$. As a consequence,

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X| \geq \varepsilon) = \lim_{n\to\infty} \frac{1}{n} = 0,$$

which yields the convergence in probability of the sequence $(X_n)_{n\geq 1}$ to X. Note that this result is implied by the theorem stating that any sequence $(X_n)_{n\geq 1}$ which converges in distribution to a Dirac random variable converges also in probability to this Dirac random variable.

To check the almost sure convergence of the sequence $(X_n)_{n\geq 1}$ to X, choosing any $\varepsilon < 1$, on account of the independence of the random variables of the sequence $(X_n)_{n\geq 1}$, we estimate

$$\begin{aligned} \mathbf{P}\left(\bigcap_{n\geq m} \{|X_n| \leq \varepsilon\}\right) &\leq \mathbf{P}\left(\bigcap_{n=m}^{2m} \{|X_n| \leq \varepsilon\}\right) = \prod_{n=m}^{2m} \mathbf{P}(|X_n| \leq \varepsilon) \\ &= \prod_{n=m}^{2m} \mathbf{P}(X_n = 0) = \prod_{n=m}^{2m}\left(1-\frac{1}{n}\right) \\ &\leq \prod_{n=m}^{2m}\left(1-\frac{1}{2m}\right) = \left(1-\frac{1}{2m}\right)^m. \end{aligned}$$

As a consequence,

$$\lim_{m\to\infty} \mathbf{P}\left(\bigcap_{n\geq m} \{|X_n| \leq \varepsilon\}\right) \leq \lim_{m\to\infty}\left(1-\frac{1}{2m}\right)^m = e^{-1/2} < 1.$$

This prevents that

$$\lim_{m\to\infty} \mathbf{P}\left(\bigcap_{n\geq m} \{|X_n| \leq \varepsilon\}\right) = 1,$$

so that $X_n \overset{\mathbf{a.s.}}{\nrightarrow} 0$.

To check the convergence in mean of the sequence $(X_n)_{n\geq 1}$, recall first that if a sequence $(X_n)_{n\geq 1}$ of real random variables converges in probability to a real random variable X and $(X_n)_{n\geq 1}$ converges also in mean, then $(X_n)_{n\geq 1}$ converges in mean to X. In light of this, we check the convergence in mean of the sequence $(X_n)_{n\geq 1}$ to X. By virtue of (**??**), we have

$$\mathbf{E}[|X_n - X|] = \mathbf{E}[X_n].$$

On the other hand,

$$\mathbf{E}[X_n] = 0 \cdot \mathbf{P}(X_n = 0) + \sqrt{n} \cdot \mathbf{P}(X_n = \sqrt{n}) = 0\left(1-\frac{1}{n}\right) + \sqrt{n}\frac{1}{n} = \frac{1}{\sqrt{n}}.$$

Therefore,

$$\lim_{n\to\infty} \mathbf{E}[|X_n - X|] = \lim_{n\to\infty} \mathbf{E}[X_n] = \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0,$$

which yields the convergence in mean of the sequence $(X_n)_{n\geq 1}$ to X.

To check the convergence in square mean of the sequence $(X_n)_{n\geq 1}$, recall first that if a sequence $(X_n)_{n\geq 1}$ of real random variables converges in probability to a real random variable X and $(X_n)_{n\geq 1}$ converges also in square mean, then $(X_n)_{n\geq 1}$ converges in square mean to X. In light of this, we check convergence in square mean of the sequence $(X_n)_{n\geq 1}$ to X. By virtue of (**??**), we have

$$\mathbf{E}\left[|X_n - X|^2\right] = \mathbf{E}\left[X_n^2\right].$$

On the other hand,

$$\mathbf{E}\left[X_n^2\right] = 0 \cdot \mathbf{P}(X_n = 0) + n \cdot \mathbf{P}(X_n = \sqrt{n}) = 0\left(1-\frac{1}{n}\right) + n\frac{1}{n} = 1.$$

Therefore,

$$\lim_{n\to\infty} \mathbf{E}\left[|X_n - X|^2\right] = \lim_{n\to\infty} \mathbf{E}\left[X_n^2\right] = \lim_{n\to\infty} 1 = 1.$$

This prevents the convergence in square mean of the sequence $(X_n)_{n\geq 1}$ to X.

Problem 2 *Let $X \sim U(0,1)$ and let $(Y_n)_{n\geq 1}$ be the sequence of real random variables given by*

$$Y_n \overset{def}{=} \begin{cases} n & \text{if } 0 \leq X < \frac{1}{n}, \\ 0 & \text{if } 1/n \leq X \leq 1. \end{cases}, \quad \forall n \geq 1$$

già fatta

Check whether the sequence $(Y_n)_{n\geq 1}$ converges in probability, converges in mean, converges almost surely, in the assigned order.

Exercise 3 *Hint: to deal with the almost sure convergence consider the event $E_0 \equiv \{\omega \in \Omega : X(\omega) = 0\}$ and the complement E_0^c.*

Solution. Note that, according to the definition

$$\mathbf{P}(Y_n = n) = \mathbf{P}\left(0 \leq X < \frac{1}{n}\right) = \frac{1}{n} \quad \text{and} \quad \mathbf{P}(Y_n = 0) = \mathbf{P}\left(\frac{1}{n} \leq X \leq 1\right) = 1 - \frac{1}{n}.$$

Therefore,

$$\mathbf{P}(|Y_n| \leq \varepsilon) \geq \mathbf{P}(Y_n = 0) = 1 - \frac{1}{n}.$$

It follows

$$\lim_{n\to\infty} \mathbf{P}(|Y_n| \leq \varepsilon) \geq \lim_{n\to\infty} \left(1 - \frac{1}{n}\right) = 1,$$

which implies the convergence in probability to 0. Now, to check the convergence in mean, we consider

$$\mathbf{E}[Y_n] = n\mathbf{P}(Y_n = n) + 0\mathbf{P}(Y_n = 0) = 1$$

It follows that

$$\lim_{n\to\infty} \mathbf{E}[|Y_n - 0|] = \lim_{n\to\infty} \mathbf{E}[Y_n] = 1 \neq 0.$$

Hence, $(Y_n)_{n\geq 1}$ does not converge in mean. In the end, consider the event

$$E_0 \equiv \{\omega \in \Omega : X(\omega) = 0\}.$$

Since $X \sim U(0,1)$ we have $\mathbf{P}(E_0) = 0$. In addition, for every $\omega \in E_0^c$ we have $X(\omega) > 0$ and it is possible to find n_ω such that for every $n > n_\omega$

$$\frac{1}{n} < X(\omega).$$

It then follows that

$$Y_n(\omega) = 0$$

for every $n > n_\omega$. This implies

$$\lim_{n\to\infty} Y_n(\omega) = 0, \quad \forall \omega \in E_0^c$$

which yields the almost sure convergence to 0 of the sequence $(Y_n)_{n\geq 1}$.

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{\geq n}$ be a sequence of real random variables on Ω. Assume that $(X_n)_{\geq n}$ are identically distributed and let $f_X : \mathbb{R} \to \mathbb{R}_+$ their common density function given by*

$$f_X(x) \overset{def}{=} \frac{2}{x^3} 1_{(1,+\infty)}(x), \quad \forall x \in \mathbb{R}.$$

Set

$$Y_n \equiv \frac{X_n}{n^\alpha}, \quad \forall n \geq 1,$$

where $\alpha > 0$.

1. *Study the convergence in distribution, probability, and L^p of the sequence $(Y_n)_{n\geq 1}$ on varying of $\alpha > 0$.*
2. *Under the additional assumption of independence of the random variables of the sequence $(X_n)_{\geq n}$, does the sequence $(Y_n)_{n\geq 1}$ converge almost surely?*

Solution. Note that the random variables of the sequences $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ are almost surely positive. Hence, we have

$$\begin{aligned} \mathbf{E}[|Y_n|^p] = \mathbf{E}[Y_n^p] &= \int_\Omega Y_n^p \, d\mathbf{P} = \frac{1}{n^{\alpha p}} \int_\Omega X_n^p \, d\mathbf{P} = \frac{1}{n^{\alpha p}} \int_\mathbb{R} x^p f_X(x) \, dx \\ &= \frac{1}{n^{\alpha p}} \int_\mathbb{R} 2x^{p-3} 1_{(1,+\infty)}(x) \, dx = \frac{2}{n^{\alpha p}} \int_1^{+\infty} x^{p-3} \, dx, \end{aligned}$$

for every $\alpha > 0$, $p \geq 1$. On the other hand,

$$\int_1^{+\infty} x^{p-3} \, dx = \begin{cases} \frac{1}{p-2} x^{p-2}\Big|_1^{+\infty} = \frac{1}{2-p} & \text{if } 1 \leq p < 2 \\ +\infty & \text{if } p \geq 2 \end{cases}.$$

It then follows

$$\mathbf{E}[|Y_n|^p] = \begin{cases} \frac{1}{2-p} \frac{2}{n^{\alpha p}} & \text{if } 1 \leq p < 2 \\ +\infty & \text{if } p \geq 2 \end{cases}.$$

for every $\alpha > 0$, As a consequence, $Y_n \overset{\mathbf{L}^p}{\to} 0$ for every $\alpha > 0$ if and only if $1 \leq p < 2$. In particular, $Y_n \overset{\mathbf{P}}{\to} 0$ and $Y_n \overset{\mathbf{w}}{\to} 0$ for every $\alpha > 0$. Now, we can write

$$\mathbf{P}(|Y_n| > \varepsilon) = \mathbf{P}(Y_n > \varepsilon) = \mathbf{P}(X_n > n^\alpha \varepsilon) = \int_{n^\alpha \varepsilon}^{+\infty} 2x^{-3} 1_{(1,+\infty)}(x) \, dx = -x^{-2}\Big|_{n^\alpha \varepsilon}^{+\infty} = \frac{1}{\varepsilon^2 n^{2\alpha}}$$

for every $\varepsilon > 0$ and every $n > n_\varepsilon$, where $n_\varepsilon \in \mathbb{N}$ is such that $n^\alpha \varepsilon > 1$. As a consequence,

$$\sum_{n=1}^\infty \mathbf{P}(|Y_n| > \varepsilon) < \infty \Leftrightarrow \frac{1}{\varepsilon^2} \sum_{n=1}^\infty \frac{1}{n^{2\alpha}} < \infty \Leftrightarrow \alpha > \frac{1}{2}.$$

Therefore, we have

$$Y_n \overset{\text{a.s.}}{\not\to} 0, \quad \forall \alpha \leq \frac{1}{2} \quad \text{and} \quad Y_n \overset{\text{a.s.}}{\to} 0, \quad \forall \alpha > \frac{1}{2}.$$

Exercise 5 *Consider the interval $[0,1]$ of the real Euclidean line. Let $\mathcal{B}([0,1])$ the Borel σ-algebra on $[0,1]$ and let $\mu_L : \mathcal{B}([0,1]) \to \mathbb{R}_+$ be the Lebesgue measure on $[0,1]$. Hence, consider the probability space $(\Omega, \mathcal{E}, \mathbf{P})$ where $\Omega \equiv [0,1]$, $\mathcal{E} = \mathcal{B}([0,1])$, and $\mathbf{P} \equiv \mu_L$. Prove,* ***following the prescribed order whitout independence assumption****, that the sequence $(X_n)_{n\geq 1}$ of random variables given by*

$$X_n \overset{def}{=} \sqrt{n} 1_{[0,1/n]}, \quad \forall n \geq 1 \tag{1}$$

converges in distribution, in probability, almost surely, and in mean to the Dirac random variable concentrated at 0. Prove also that $(X_n)_{n\geq 1}$ does not converge in square mean.

Solution. According to Definition 1 we have

$$X_n(\omega) = \begin{cases} \sqrt{n} & \text{if } \omega \in [0, 1/n] \\ 0 & \text{if } \omega \in (1/n, 1] \end{cases},$$

for every $n \geq 1$. Hence,

$$\mathbf{P}\left(X_n = \sqrt{n}\right) = \frac{1}{n} \quad \text{and} \quad \mathbf{P}(X_n = 0) = \frac{n-1}{n}.$$

That is X_n is a Bernoulli random variable with states $0, \sqrt{n}$ and success probability $\frac{1}{n}$. As a consequence, considering the distribution function $F_{X_n} : \mathbb{R} \to \mathbb{R}_+$ of X_n, we have

$$F_{X_n}(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{n-1}{n} & \text{if } 0 \leq x < \sqrt{n} \\ 1 & \text{if } \sqrt{n} \leq x \end{cases},$$

for every $n \geq 1$. It follows

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \leq x \end{cases}.$$

Therefore,

$$\lim_{n\to\infty} F_{X_n}(x) = H(x),$$

for every $x \in \mathbb{R}$, where $H(x)$ is the Heaviside function which is the distribution function of the Dirac random variable concentrated at 0. This shows that

$$X_n \overset{\mathbf{d}}{\to} X_0,$$

where X_0 is the Dirac random variable concentrated at 0

$$X_0(\omega) \overset{\text{def}}{=} 0, \quad \mathbf{P}(X_0 = 0) = 1.$$

To prove the convergence in probability of X_n to X_0, we have to show that

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X_0| < \varepsilon) = 1,$$

for every $\varepsilon > 0$. We have

$$|X_n - X_0| = X_n \quad \text{a.s. on } \Omega.$$

Hence,

$$\mathbf{P}(|X_n - X_0| < \varepsilon) = \mathbf{P}(X_n < \varepsilon).$$

Now, for any $\varepsilon \leq 1$ we have

$$\mathbf{P}(X_n < \varepsilon) = \mathbf{P}(X_n = 0) = \mu_L((1/n, 1]) = 1 - \frac{1}{n}.$$

It follows,

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X_0| < \varepsilon) = \lim_{n\to\infty} 1 - \frac{1}{n} = 1.$$

this is sufficient to prove that

$$X_n \overset{\mathbf{P}}{\to} X_0.$$

To prove the almost sure convergence of X_n to X_0, we need to show that there exists an event $E \in \mathcal{E}$ such that $\mathbf{P}(E) = 1$ and

$$\lim_{n\to\infty} X_n(\omega) = X_0(\omega),$$

for every $\omega \in E$. To this goal consider the event $E = (0, 1]$. We have

$$\mathbf{P}(E) = \mu_L(E) = 1.$$

In addition for every $\omega \in E$ there exists n_ω such that

$$\frac{1}{n} < \omega$$

for every $n > n_\omega$. This implies that

$$X_n(\omega) = 0$$

for every $n > n_\omega$ and it follows

$$\lim_{n\to\infty} X_n(\omega) = 0 = X_0(\omega).$$

That is

$$X_n \overset{\mathbf{a.s.}}{\to} X_0.$$

To prove the convergence in mean of X_n to X_0, we have to show that

$$\lim_{n\to\infty} \mathbf{E}[|X_n - X_0|] = 0.$$

We have

$$\mathbf{E}[|X_n - X_0|] = \mathbf{E}[X_n] = \sqrt{n}\mathbf{P}\left(X_n = \sqrt{n}\right) = \sqrt{n}\mathbf{P}([0, 1/n]) = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}.$$

It follows

$$\lim_{n\to\infty} \mathbf{E}[|X_n - X_0|] = \lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0,$$

which proves

$$X_n \overset{\mathbf{L}^1}{\to} X_0.$$

To prove that X_n does not converge in square mean to X_0, we compute

$$\mathbf{E}\left[|X_n - X_0|^2\right] = \mathbf{E}\left[X_n^2\right] = n\mathbf{P}\left(X_n = \sqrt{n}\right) = n\mathbf{P}([0, 1/n]) = \frac{n}{n} = 1.$$

It follows

$$\lim_{n\to\infty} \mathbf{E}\left[|X_n - X_0|^2\right] = 1,$$

which prevents

$$\lim_{n\to\infty} \mathbf{E}\left[|X_n - X_0|^2\right] = 0,$$

as it would be necessary to have

$$X_n \overset{\mathbf{L}^2}{\to} X_0.$$

Problem 6 *Show that the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by*

$$f(x, y) \overset{def}{=} \begin{cases} \frac{y-x}{2} & \textit{if } (x, y) \in [-1, 0] \times [0, 1] \\ \frac{x-y}{2} & \textit{if } (x, y) \in [0, 1] \times [-1, 0] \\ 0 & \textit{otherwise} \end{cases}$$

is a probability density. Hence, consider the random vector $(X, Y)^\intercal$ with density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f_{X,Y}(x, y) \overset{def}{=} f(x, y).$$

Determine the marginal densities of entries X and Y of $(X, Y)^\intercal$. Are X and Y correlated? Are X and Y independent? Compute

$$\mathbf{P}(X + Y \geq 0).$$

Solution. To prove that $f:\mathbb{R}^2\to\mathbb{R}$ is a probability density, we need to show that

$$f(x,y)\geq 0,$$

for almost every $(x,y)\in\mathbb{R}^2$ and

$$\int_{\mathbb{R}^2} f(x,y)\ d\mu_L(x,y)=1.$$

Since in $[-1,0]\times[0,1]$ [resp. $[0,1]\times[-1,0]$] we have $x\leq 0$ and $y\geq 0$ [resp. $x\geq 0$ and $y\leq 0$] it follows $y-x\geq 0$ [resp. $x-y\geq 0$]. This proves the positivity of $f(x,y)$ for every $(x,y)\in\mathbb{R}^2$. Since we can write

$$f(x,y)=\frac{y-x}{2}1_{[-1,0]\times[0,1]}(x,y)+\frac{x-y}{2}1_{[0,1]\times[-1,0]}(x,y),$$

for every $(x,y)\in\mathbb{R}^2$, thanks to the properties of the Lebesgue integral for positive functions, we obtain

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\ d\mu_L(x,y)&=\int_{\mathbb{R}^2}\left(\frac{y-x}{2}1_{[-1,0]\times[0,1]}(x,y)+\frac{x-y}{2}1_{[0,1]\times[-1,0]}(x,y)\right)\ d\mu_L(x,y)\\&=\int_{\mathbb{R}^2}\frac{y-x}{2}1_{[-1,0]\times[0,1]}(x,y)\ d\mu_L(x,y)+\int_{\mathbb{R}^2}\frac{x-y}{2}1_{[0,1]\times[-1,0]}(x,y)\ d\mu_L(x,y)\\&=\int_{[-1,0]\times[0,1]}\frac{y-x}{2}\ d\mu_L(x,y)+\int_{[0,1]\times[-1,0]}\frac{x-y}{2}\ d\mu_L(x,y).\end{aligned}$$

On the other hand, the function $y-x$ [resp. $x-y$] is continuous on $[-1,0]\times[0,1]$ [resp. $[0,1]\times[-1,0]$]. It follows

$$\int_{[-1,0]\times[0,1]}\frac{y-x}{2}\ d\mu_L(x,y)=\int_{x=-1}^{0}\int_{y=0}^{1}\frac{y-x}{2}dxdy\quad\text{[resp. }\int_{[0,1]\times[-1,0]}\frac{x-y}{2}\ d\mu_L(x,y)=\int_{x=1}^{0}\int_{y=-1}^{0}\frac{y-x}{2}dxdy\text{]}.$$

Now,

$$\begin{aligned}\int_{x=-1}^{0}\int_{y=0}^{1}\frac{y-x}{2}dxdy&=\frac{1}{2}\int_{x=-1}^{0}\left(\int_{y=0}^{1}(y-x)\,dy\right)dx\\&=\frac{1}{2}\int_{x=-1}^{0}\left(\left.\frac{(y-x)^2}{2}\right|_{y=0}^{1}\right)dx\\&=\frac{1}{2}\int_{x=-1}^{0}\left(\frac{(1-x)^2}{2}-\frac{x^2}{2}\right)dx\\&=\frac{1}{2}\int_{x=-1}^{0}\left(\frac{1-2x+x^2}{2}-\frac{x^2}{2}\right)dx\\&=\frac{1}{2}\int_{x=-1}^{0}\left(\frac{1}{2}-x\right)dx\\&=\frac{1}{2}\left.\frac{1}{2}x-\frac{x^2}{2}\right|_{x=-1}^{0}\\&=\frac{1}{4}\left.x-x^2\right|_{x=-1}^{0}\\&=\frac{1}{2}.\end{aligned}$$

Moreover, we clearly have

$$\int_{x=1}^{0}\int_{y=-1}^{0}\frac{y-x}{2}dxdy=\int_{x=-1}^{0}\int_{y=0}^{1}\frac{y-x}{2}dxdy.$$

It then follows that $f:\mathbb{R}^2\to\mathbb{R}$ is a probability density. The marginal densities $f_X:\mathbb{R}\to\mathbb{R}$ and $f_Y:\mathbb{R}\to\mathbb{R}$ are given by

$$f_X(x)\overset{\text{def}}{=}\int_{\mathbb{R}} f_{X,Y}(x,y)\ d\mu_L(y)\quad\text{and}\quad f_Y(y)\overset{\text{def}}{=}\int_{\mathbb{R}} f_{X,Y}(x,y)\ d\mu_L(x).$$

Since we can write

$$f(x,y)=\frac{y-x}{2}1_{[-1,0]}(x)\,1_{[0,1]}(y)+\frac{x-y}{2}1_{[0,1]}(x)\,1_{[-1,0]}(y),$$

we have

$$\begin{aligned}f_X(x)&=\int_{\mathbb{R}}\frac{y-x}{2}1_{[-1,0]}(x)\,1_{[0,1]}(y)\ d\mu_L(y)+\int_{\mathbb{R}}\frac{x-y}{2}1_{[0,1]}(x)\,1_{[-1,0]}(y)\ d\mu_L(y)\\&=1_{[-1,0]}(x)\int_{\mathbb{R}}\frac{y-x}{2}1_{[0,1]}(y)\ d\mu_L(y)+1_{[0,1]}(x)\int_{\mathbb{R}}\frac{x-y}{2}1_{[-1,0]}(y)\ d\mu_L(y)\\&=1_{[-1,0]}(x)\int_{[0,1]}\frac{y-x}{2}\ d\mu_L(y)+1_{[0,1]}(x)\int_{[-1,0]}\frac{x-y}{2}\ d\mu_L(y)\\&=\frac{1}{2}1_{[-1,0]}(x)\int_{0}^{1}(y-x)\ dy+\frac{1}{2}1_{[0,1]}(x)\int_{-1}^{0}(x-y)\ dy\\&=\frac{1}{2}1_{[-1,0]}(x)\left.\frac{(y-x)^2}{2}\right|_{y=0}^{1}-\frac{1}{2}1_{[0,1]}(x)\left.\frac{(x-y)^2}{2}\right|_{y=-1}^{0}\\&=\frac{1}{2}1_{[-1,0]}(x)\left(\frac{(1-x)^2}{2}-\frac{x^2}{2}\right)-\frac{1}{2}1_{[0,1]}(x)\left(\frac{x^2}{2}-\frac{(x+1)^2}{2}\right)\\&=\frac{1}{2}\left(\frac{1-2x}{2}\right)1_{[-1,0]}(x)+\frac{1}{2}\left(\frac{1+2x}{2}\right)1_{[0,1]}(x)\\&=\frac{1}{4}\left((1-2x)\,1_{[-1,0]}(x)+(1+2x)\,1_{[0,1]}(x)\right).\end{aligned}$$

Similarly,

$$\begin{aligned}f_Y(y)&=\int_{\mathbb{R}}\frac{y-x}{2}1_{[-1,0]}(x)\,1_{[0,1]}(y)\ d\mu_L(x)+\int_{\mathbb{R}}\frac{x-y}{2}1_{[0,1]}(x)\,1_{[-1,0]}(y)\ d\mu_L(x)\\&=1_{[0,1]}(y)\int_{\mathbb{R}}\frac{y-x}{2}1_{[-1,0]}(x)\ d\mu_L(x)+1_{[-1,0]}(y)\int_{\mathbb{R}}\frac{x-y}{2}1_{[0,1]}(x)\ d\mu_L(x)\\&=1_{[0,1]}(y)\int_{[-1,0]}\frac{y-x}{2}\ d\mu_L(x)+1_{[-1,0]}(y)\int_{[0,1]}\frac{x-y}{2}\ d\mu_L(x)\\&=\frac{1}{2}1_{[0,1]}(y)\int_{-1}^{0}(y-x)\ dx+\frac{1}{2}1_{[-1,0]}(y)\int_{0}^{1}(x-y)\ dx\\&=-\frac{1}{2}1_{[0,1]}(y)\left.\frac{(y-x)^2}{2}\right|_{x=-1}^{0}+\frac{1}{2}1_{[-1,0]}(y)\left.\frac{(x-y)^2}{2}\right|_{x=0}^{1}\\&=-\frac{1}{2}1_{[0,1]}(y)\left(\frac{y^2}{2}-\frac{(y+1)^2}{2}\right)+\frac{1}{2}1_{[-1,0]}(y)\left(\frac{(1-y)^2}{2}-\frac{y^2}{2}\right)\\&=\frac{1}{2}\left(\frac{1+2y}{2}\right)1_{[0,1]}(y)+\frac{1}{2}\left(\frac{1-2y}{2}\right)1_{[-1,0]}(y)\\&=\frac{1}{4}\left((1-2y)\,1_{[-1,0]}(y)+(1+2y)\,1_{[0,1]}(y)\right).\end{aligned}$$

Note that we have[1]

$$
\begin{aligned}
f_X(-x) &= \frac{1}{4}\left((1+2x)\,1_{[-1,0]}(-x) + (1-2x)\,1_{[0,1]}(-x)\right) \\
&= \frac{1}{4}\left((1+2x)\,1_{[0,1]}(x) + (1-2x)\,1_{[-1,0]}(x)\right) \\
&= f_X(x).
\end{aligned}
$$

That is $f_X : \mathbb{R} \to \mathbb{R}_+$ is an even function. Moreover,

$$f_Y(y) = f_X(x)\,|_{x=y}.$$

As a consequence,

$$\mathbf{E}[X] = \int_{\mathbb{R}} x f_X(x)\ d\mu_L(x) = 0 = \mathbf{E}[Y].$$

Otherwise, in terms of explicit computations, we can write,

$$
\begin{aligned}
\mathbf{E}[X] &= \int_{\mathbb{R}} x f_X(x)\ d\mu_L(x) \\
&= \int_{\mathbb{R}} \frac{1}{4}\left(x(1-2x)\,1_{[-1,0]}(x) + x(1+2x)\,1_{[0,1]}(x)\right)\ d\mu_L(x) \\
&= \frac{1}{4}\left(\int_{[-1,0]} x(1-2x)\ d\mu_L(x) + \int_{[0,1]} x(1+2x)\ d\mu_L(x)\right) \\
&= \frac{1}{4}\left(\int_{-1}^{0} \left(x-2x^2\right)\ dx + \int_{0}^{1} \left(x+2x^2\right)\ dx\right) \\
&= \frac{1}{4}\left(\frac{x^2}{2} - \frac{2}{3}x^3\Big|_{x=-1}^{0} + \frac{x^2}{2} + \frac{2}{3}x^3\Big|_{x=0}^{1}\right) \\
&= \frac{1}{4}\left(-\frac{1}{2} - \frac{2}{3} + \frac{1}{2} + \frac{2}{3}\right) \\
&= 0.
\end{aligned}
$$

The same computation yields

$$
\begin{aligned}
\mathbf{E}[Y] &= \int_{\mathbb{R}} y f_X(y)\ d\mu_L(y) \\
&= \int_{\mathbb{R}} \frac{1}{4}\left(1_{[-1,0]}(y)\,y(1-2y) + 1_{[0,1]}(y)\,y(1+2y)\right)\ d\mu_L(y) \\
&= \frac{1}{0}.
\end{aligned}
$$

[1] Thanks to Tiziana Mannucci

On the other hand,

$$
\begin{aligned}
\mathbf{E}[XY] &= \int_{\mathbb{R}^2} xy f_{X,Y}(x,y)\ d\mu_L(x,y) \\
&= \int_{\mathbb{R}^2} \left(\frac{xy(y-x)}{2} 1_{[-1,0]\times[0,1]}(x,y) + \frac{xy(x-y)}{2} 1_{[0,1]\times[-1,0]}(x,y)\right)\ d\mu_L(x,y) \\
&= \int_{\mathbb{R}^2} \frac{xy(y-x)}{2} 1_{[-1,0]\times[0,1]}(x,y)\ d\mu_L(x,y) + \int_{\mathbb{R}^2} \frac{xy(x-y)}{2} 1_{[0,1]\times[-1,0]}(x,y)\ d\mu_L(x,y) \\
&= \int_{[-1,0]\times[0,1]} \frac{xy(y-x)}{2}\ d\mu_L(x,y) + \int_{[0,1]\times[-1,0]} \frac{xy(x-y)}{2}\ d\mu_L(x,y) \\
&= \frac{1}{2}\left(\int_{y=0}^{1}\left(\int_{x=-1}^{0} \left(xy^2 - x^2y\right)\ dx\right) dy + \int_{y=-1}^{0}\left(\int_{x=0}^{1} \left(x^2y - xy^2\right)\ dx\right) dy\right) \\
&= \frac{1}{2}\left(\int_{y=0}^{1} \frac{1}{2}x^2y^2 - \frac{1}{3}x^3y\Big|_{x=-1}^{0}\ dy + \int_{y=-1}^{0} \frac{1}{3}x^3y - \frac{1}{2}x^2y^2\Big|_{x=0}^{1}\ dy\right) \\
&= \frac{1}{2}\left(-\int_{y=0}^{1}\left(\frac{1}{2}y^2 + \frac{1}{3}y\right) dy + \int_{y=-1}^{0} \frac{1}{3}y - \frac{1}{2}y^2 dy\right) \\
&= \frac{1}{2}\left(-\frac{1}{6}y^3 + \frac{1}{6}y^2\Big|_{y=0}^{1} + \frac{1}{6}y^2 - \frac{1}{6}y^3\Big|_{y=-1}^{0}\right) \\
&= \frac{1}{12}(-2-2) \\
&= -\frac{1}{3}.
\end{aligned}
$$

It follows

$$\mathbf{E}[XY] \neq \mathbf{E}[X]\,\mathbf{E}[Y],$$

which shows that X and Y are correlated. Therefore, X and Y are not independent. In the end, setting

$$H = \left\{(x,y) \in \mathbb{R}^2 : x + y \geq 0\right\},$$

which is the half-plane on the right of the line of equation

$$x + y = 0,$$

we have

$$\mathbf{P}(X+Y \geq 0) = \int_H f_{X,Y}(x,y)\, d\mu_L^2(x,y).$$

Now, since ve can write

$$
\begin{aligned}
f(x,y) &= \frac{y-x}{2} 1_{[-1,0]\times[0,1]}(x,y) + \frac{x-y}{2} 1_{[0,1]\times[-1,0]}(x,y) \\
&= \frac{y-x}{2} 1_{[-1,0]}(x)\, 1_{[0,1]}(y) + \frac{x-y}{2} 1_{[0,1]}(x)\, 1_{[-1,0]}(y)
\end{aligned}
$$

we have

$$
\begin{aligned}
f(-x,-y) &= \frac{-y+x}{2} 1_{[-1,0]}(-x)\, 1_{[0,1]}(-y) + \frac{-x+y}{2} 1_{[0,1]}(-x)\, 1_{[-1,0]}(-y) \\
&= \frac{x-y}{2} 1_{[0,1]}(x)\, 1_{[-1,0]}(y) + \frac{y-x}{2} 1_{[-1,0]}(x)\, 1_{[0,1]}(y) \\
&= f(x,y).
\end{aligned}
$$

and

$$
\begin{aligned}
f(y,x) &= \frac{x-y}{2} 1_{[-1,0]}(y)\, 1_{[0,1]}(x) + \frac{y-x}{2} 1_{[0,1]}(y)\, 1_{[-1,0]}(x) \\
&= \frac{y-x}{2} 1_{[-1,0]}(x)\, 1_{[0,1]}(y) + \frac{x-y}{2} 1_{[0,1]}(x)\, 1_{[-1,0]}(y) \\
&= f(x,y)
\end{aligned}
$$

This means that the function $f : \mathbb{R}^2 \to \mathbb{R}$ is symmetric with respect to the point $(0,0)$ and the line $x+y=0$. As a consequence,

$$\int_H f_{X,Y}(x,y)\, d\mu_L^2(x,y) = \frac{1}{2}.$$

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2019-12-06

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{n\geq 1}$ be a sequence of independent identically distributed Bernoulli random variables with success probability p. Set*

$$Z_n \stackrel{def}{=} \sum_{k=1}^n X_n, \qquad \bar{X}_n = \frac{1}{n}\sum_{n=1}^n X_k.$$

esonero 2022

Assume that n is large. What you can say about the distributions, expectation, and variance of Z_n and $\bar{X}_n$? Consider the case $n = 100,000$ and $p = 1/2$. Use both the Central Limit Theorem and the Tchebychev inequality to estimate the probability that Z_n lies between $49,500$ and $50,500$. What you can say about the distributions, expectation, and variance of Z_n and $\bar{X}_n$ if $(X_n)_{n\geq 1}$ is a sequence of independent and Poisson distributed random variables with the same rate parameter λ?

Solution. Under the assumption that $(X_n)_{n\geq 1}$ is a sequence of independent and identically distributed Bernoulli random variables with success probability p, the random variable Z_n has the binomial distribution with parameters n and p, for every $n \in \mathbb{N}$. In symbols

$$Z_n \sim Bin(n,p), \quad \forall n \in \mathbb{N}.$$

Hence,

$$\mathbf{E}[Z_n] = np \quad \text{and} \quad \mathbf{D}^2[Z_n] = np(1-p),$$

for every $n \in \mathbb{N}$. Therefore, the random variable

$$Z_n^* \equiv \frac{Z_n - np}{\sqrt{np(1-p)}}$$

is standardized, that is

$$\mathbf{E}[Z_n^*] = 0 \quad \text{and} \quad \mathbf{D}^2[Z_n^*] = 1.$$

for every $n \in \mathbb{N}$. By virtue of the Central Limit Theorem, we know that

$$Z_n^* \stackrel{\mathbf{w}}{\to} N(0,1),$$

as $n \to \infty$. As a consequence, as n is large, the random variable Z_n^* is approximately standard Gaussian distributed. On the other hand, we can write

$$\sqrt{\frac{p(1-p)}{n}} Z_n^* = \sqrt{\frac{p(1-p)}{n}} \frac{Z_n - np}{\sqrt{np(1-p)}} = \frac{Z_n}{n} - p = \bar{X}_n - p,$$

that is

$$\bar{X}_n = \sqrt{\frac{p(1-p)}{n}} Z_n^* + p$$

This implies that, as n is large, the distribution of $\bar{X}_n$ is approximately Gaussian with

$$\mathbf{E}\left[\bar{X}_n\right] = p \quad \text{and} \quad \mathbf{D}^2\left[\bar{X}_n\right] = \frac{p(1-p)}{n}.$$

In the case $n = 100,000$ and $p = 1/2$, thanks to the Central Limit Theorem, we can then write

$$\begin{aligned}
\mathbf{P}(49,500 \le Z_n \le 50,500) &= \mathbf{P}(49,500 - np \le Z_n - np \le 50,500 - np) \\
&= \mathbf{P}\left(\frac{49,500 - np}{\sqrt{np(1-p)}} \le \frac{Z_n - np}{\sqrt{np(1-p)}} \le \frac{50,500 - np}{\sqrt{np(1-p)}}\right) \\
&= \mathbf{P}\left(\frac{49,500 - np}{\sqrt{np(1-p)}} \le Z_n^* \le \frac{50,500 - np}{\sqrt{np(1-p)}}\right) \\
&\simeq \Phi\left(\frac{50,500 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{49,500 - np}{\sqrt{np(1-p)}}\right) \\
&= \Phi\left(\frac{500}{\sqrt{25,000}}\right) - \Phi\left(\frac{-500}{\sqrt{25,000}}\right) \\
&= 2\Phi\left(\sqrt{10}\right) - 1 \\
&= 2 \cdot 0.9992 - 1 \\
&= 0.9984.
\end{aligned}$$

where Φ is the distribution function of the standard normal. Instead, with the goal of applying the Tchebychev inequality, we can write

$$\begin{aligned}
\mathbf{P}(49,500 \le Z_n \le 50,500) &= \mathbf{P}\left(\frac{49,500 - np}{\sqrt{np(1-p)}} \le \frac{Z_n - np}{\sqrt{np(1-p)}} \le \frac{50,500 - np}{\sqrt{np(1-p)}}\right) \\
&= \mathbf{P}\left(-\sqrt{10} \le Z_n^* \le \sqrt{10}\right) \\
&= \mathbf{P}\left(|Z_n^*| \le \sqrt{10}\right) \\
&= 1 - \mathbf{P}\left(|Z_n^*| > \sqrt{10}\right) \\
&= 1 - \mathbf{P}\left(|Z_n^*| \ge \sqrt{10}\right).
\end{aligned}$$

On the other hand, by the Tchebychev inequality we have

$$\mathbf{P}\left(|Z_n^*| \ge \sqrt{10}\right) \le \frac{\mathbf{D}^2[Z_n^*]}{10} = \frac{1}{10}.$$

Therefore,

$$\mathbf{P}(49,500 \le Z_n \le 50,500) \ge 1 - \frac{1}{10} = \frac{9}{10} = 0.9.$$

This shows that the central limit aproach provides a sharper bound for the desired probability than the Tchebychev inequality approach. Now, if $(X_n)_{n\ge1}$ is a sequence of independent and Poisson distributed random variables with the same rate parameter λ, the random variable Z_n has the exponential distribution with rate parameter $n\lambda$. In symbols

$$Z_n \sim Poiss(n\lambda), \quad \forall n \in \mathbb{N}.$$

Hence,

$$\mathbf{E}[Z_n] = n\lambda \quad \text{and} \quad \mathbf{D}^2[Z_n] = n\lambda.$$

for every $n \in \mathbb{N}$. Therefore, the random variable

$$Z_n^* \equiv \frac{Z_n - n\lambda}{\sqrt{n\lambda}}$$

is standardized, that is

$$\mathbf{E}[Z_n^*] = 0 \quad \text{and} \quad \mathbf{D}^2[Z_n^*] = 1,$$

for every $n \in \mathbb{N}$. Again, by virtue of the Central Limit Theorem, we know that

$$Z_n^* \xrightarrow{\mathbf{w}} N(0,1),$$

as $n \to \infty$. As a consequence, as n is large, the random variable Z_n^* is approximately Gaussian distributed. On the other hand, we can write

$$\sqrt{\frac{\lambda}{n}} Z_n^* = \sqrt{\frac{\lambda}{n}} \frac{Z_n - n\lambda}{\sqrt{n\lambda}} = \frac{Z_n}{n} - \lambda = \bar{X}_n - \lambda,$$

that is

$$\bar{X}_n = \sqrt{\frac{\lambda}{n}} Z_n^* + \lambda.$$

This implies that, as n is large, the distribution of $\bar{X}_n$ is approximately Gaussian with

$$\mathbf{E}\left[\bar{X}_n\right] = \lambda \quad \text{and} \quad \mathbf{D}^2\left[\bar{X}_n\right] = \frac{\lambda}{n}.$$

The solution is complete.

Problem 2 *Suppose that a random variable X, which represents the reaction time at some stimulus, has a uniform distribution on an interval $[0, \theta]$, where the parameter $\theta > 0$ is unknown. An investigator wants to estimate θ on the basis of a simple random sample $X_1, \dots, X_n$ of reaction times. Since θ is the largest possible time in the entire population of reaction times, the investigator consider as a first estimator for the parameter θ the largest sample reaction time. That is to say, the investigator consider as a first estimator the statistic*

$$\hat{\theta}_1 \equiv \check{X}_n \equiv \max(X_1, \dots, X_n).$$

1. *Is $\check{X}_n$ unbiased? In case $\check{X}_n$ is not unbiased, is it possible to derive from $\check{X}_n$ an unbiased estimator of θ?*

2. *As a second estimator, the investigator consider the statistic*

$$\hat{\theta}_2 \equiv \bar{X}_n \equiv \frac{1}{n}\textstyle\sum_{k=1}^n X_k.$$

   Is $\bar{X}_n$ unbiased? In case $\bar{X}_n$ is not unbiased, is it possible to derive from $\bar{X}_n$ an unbiased estimator of θ?

3. *In the investigator's shoes, what estimator would you prefer among those considered?*

Solution.

1. Writing $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of the statistic $\check{X}_n$, we have

$$\begin{aligned}
F_{\check{X}_n}(x) &= \mathbf{P}\left(\check{X}_n \le x\right) = \mathbf{P}(X_1 \le x, \dots, X_n \le x) = \textstyle\prod_{k=1}^n \mathbf{P}(X_k \le x) \\
&= \textstyle\prod_{k=1}^n \mathbf{P}(X \le x) = \mathbf{P}(X \le x)^n = F_X(x)^n.
\end{aligned}$$

for every $x \in \mathbb{R}$, where $F_X : \mathbb{R} \to \mathbb{R}$ is the distribution function of the random variable X. On the other hand, since X is uniformly distributed on $[0, \theta]$, X is absolutely continuous with density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) \stackrel{\text{def}}{=} \frac{1}{\theta} 1_{[0,\theta]}(x), \quad \forall x \in \mathbb{R}.$$

Hence,

$$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\ d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{\theta} 1_{[0,\theta]}(u)\ d\mu_L(u) = \frac{1}{\theta} \int_{(-\infty,x] \cap [0,\theta]} d\mu_L(u) \\ &= \begin{cases} \frac{1}{\theta} \int_{\varnothing} d\mu_L(u) = 0, & \text{if } x < 0, \\ \frac{1}{\theta} \int_{[0,x]} d\mu_L(u) = \frac{x}{\theta}, & \text{if } 0 \leq x \leq \theta, \\ \frac{1}{\theta} \int_{[0,\theta]} d\mu_L(u) = 1, & \text{if } \theta < x. \end{cases} \end{aligned}$$

More briefly

$$F_X(x) = \frac{x}{\theta} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

for every $x \in \mathbb{R}$. It then follows,

$$F_{\check{X}_n}(x) = F_X(x)^n = \frac{x^n}{\theta^n} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

for every $x \in \mathbb{R}$. Now, we have

$$F'_{\check{X}_n}(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{n x^{n-1}}{\theta^n}, & \text{if } 0 < x < \theta, \\ 0, & \text{if } \theta < x, \end{cases}$$

but $F_{\check{X}_n}$ is not everywhere differentable. Eventually, is not differentiable at the point $x = \theta$. However, considering the function $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ given by

$$f_{\check{X}_n}(x) \stackrel{\text{def}}{=} \frac{n x^{n-1}}{\theta^n} 1_{(0,\theta)}(x), \quad \forall x \in \mathbb{R},$$

a straightforward computation shows that

$$F_{\check{X}_n}(x) = \int_{(-\infty,x]} f_{\check{X}_n}(u)\ d\mu_L(u),$$

for every $x \in \mathbb{R}$. This implies that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n}$. As a consequence,

$$\begin{aligned} \mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} x \frac{n x^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\ d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^n\ d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^\theta x^n\ dx = \frac{n}{\theta^n} \left. \frac{x^{n+1}}{n+1} \right|_0^\theta = \frac{n}{n+1} \theta. \end{aligned}$$

We can conclude that $\check{X}_n$ is not a unbiased estimator of θ but $\frac{n+1}{n} \check{X}_n$ is an unbiased estimator of θ.

2. We have

$$\begin{aligned} \mathbf{E}\left[\bar{X}_n\right] = \mathbf{E}[X] &= \int_{\mathbb{R}} x f_X(x)\ d\mu_L(x) = \int_{\mathbb{R}} \frac{x}{\theta} 1_{[0,\theta]}(x)\ d\mu_L(x) \\ &= \frac{1}{\theta} \int_{[0,\theta]} x\ d\mu_L(x) = \frac{1}{\theta} \int_0^\theta x\ dx = \frac{1}{\theta} \left. \frac{x^2}{2} \right|_0^\theta = \frac{\theta}{2}. \end{aligned}$$

Hence, $\bar{X}_n$ is not a unbiased estimator of θ but $2\bar{X}_n$ is an unbiased estimator of θ.

3. From 1. and 2. we know that

$$\mathbf{E}\left[\frac{n+1}{n} \check{X}_n\right] = \theta \qquad \text{and} \qquad \mathbf{E}\left[2\bar{X}_n\right] = \theta.$$

Hence, both $\frac{n+1}{n} \check{X}_n$ and $2\bar{X}_n$ are unbiased estimators of the parameter θ. To choose which is preferable between them, we consider

$$\mathbf{D}^2\left[\frac{n+1}{n} \check{X}_n\right] \qquad \text{and} \qquad \mathbf{D}^2\left[2\bar{X}_n\right].$$

We have

$$\begin{aligned} \mathbf{E}\left[\check{X}_n^2\right] &= \int_{\mathbb{R}} x^2 f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} x^2 \frac{n x^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\ d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^{n+1}\ d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^\theta x^{n+1}\ dx = \frac{n}{\theta^n} \left. \frac{x^{n+2}}{n+2} \right|_0^\theta = \frac{n}{n+2} \theta^2. \end{aligned}$$

Therefore,

$$\mathbf{D}^2\left[\check{X}_n\right] = \mathbf{E}\left[\check{X}_n^2\right] - \mathbf{E}\left[\check{X}_n\right]^2 = \frac{n}{n+2} \theta^2 - \frac{n^2}{(n+1)^2} \theta^2 = \frac{n}{(n+1)^2 (n+2)} \theta^2.$$

As a consequence,

$$\mathbf{D}^2\left[\frac{n+1}{n} \check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \mathbf{D}^2\left[\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \frac{n}{(n+1)^2 (n+2)} \theta^2 = \frac{\theta^2}{n(n+2)}.$$

On the other hand,

$$\mathbf{D}^2\left[2\bar{X}_n\right] = 4\mathbf{D}^2\left[\bar{X}_n\right] = \frac{4}{n} \mathbf{D}^2[X] = \frac{4}{n} \frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

Now, for any $n > 1$ we clearly have

$$\mathbf{D}^2\left[\frac{n+1}{n} \check{X}_n\right] < \mathbf{D}^2\left[2\bar{X}_n\right].$$

It follows that the estimator $\frac{n+1}{n} \check{X}_n$ is preferable to $2\bar{X}_n$.

Problem 3 *Let X be a binomially distributed real random variable with known number of trials parameter m and unknown success parameter p. An investigator wants to estimate p on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the method of moments. What is the estimator $\hat{p}_n^M$?*
2. *Assume the investigator applies the likelihood method. What is the estimator $\hat{p}_n^{ML}$?*
3. *Given that $m = 10$ and we observe a realization $4, 4, 3, 5, 6$ of a sample $X_1, \dots, X_5$ of size 5 drawn from X what is the estimate of p by the estimators $\hat{p}_n^M$?*
4. *Can you give an estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3.?*

Solution.

1. Under the assumption considered, we know that the first population moment is given by

$$\mathbf{E}[X] = m \cdot p$$

Equating the first population moment to the first sample moment $\bar{X}_n$ and replacing $\hat{p}_n^M$ to p we obtain

$$m \cdot \hat{p}_n^M = \bar{X}_n.$$

Therefore,

$$\hat{p}_n^M = \frac{\bar{X}_n}{m}.$$

2. The density function $f_X : \mathbb{N}_0 \times (0,1) \to \mathbb{R}_+$ of a binomial random variable with fixed number of trials parameter m and variable success parameter p can be written as

$$f_X(x;p) = \frac{m!}{(m-x)!x!} p^x (1-p)^{m-x} \cdot 1_{\{0,1,\dots,m\}}(x),$$

for every $x \in \mathbb{N}_0$ and every $p \in (0,1)$. Let $X_1, \dots, X_n$ be a simple random sample of size n drawn from X. Then the likelihood function $\mathcal{L}_{X_1,\dots,X_n} : (0,1) \times \mathbb{N}_0^n \to \mathbb{R}$ of the sample $X_1, \dots, X_n$ is given by

$$\begin{aligned}
&\mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) \\
&= \textstyle\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!} p^{x_k} (1-p)^{m-x_k} \cdot 1_{\{0,1,\dots,m\}}(x_k) \\
&= \left(\textstyle\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right) p^{\sum_{k=1}^n x_k} (1-p)^{n\cdot m - \sum_{k=1}^n x_k} 1_{\{0,1,\dots,m\}^n}(x_1, \dots, x_n)
\end{aligned}$$

for every $p \in (0,1)$ and every realizations $x_1, \dots, x_n$ of the sample $X_1, \dots, X_n$. Note that

$$\begin{aligned}
&\mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) \\
&= \begin{cases} \left(\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right) p^{\sum_{k=1}^n x_k} (1-p)^{n\cdot m - \sum_{k=1}^n x_k} > 0, & \text{if } (x_1, \dots, x_n) \in \{0,1,\dots,m\}^n, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Therefore,

$$\underset{\theta \in \mathbb{R}_{++}}{\arg\max} \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) = \underset{\theta \in \mathbb{R}_{++}}{\arg\max} \left(\textstyle\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right) p^{\sum_{k=1}^n x_k} (1-p)^{n\cdot m - \sum_{k=1}^n x_k}$$

Hence, we can consider as the log-likelihood function of the sample $X_1, \dots, X_n$ the function $\log \mathcal{L}_{X_1,\dots,X_n} : (0,1) \times \{0,1,\dots,m\}^n \to \mathbb{R}$ given by

$$\begin{aligned}
&\log \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) \stackrel{\text{def}}{=} \ln \left(\left(\textstyle\prod_{k=1}^n \frac{m!}{(m-x_k)!x_k!}\right) p^{\sum_{k=1}^n x_k} (1-p)^{n\cdot m - \sum_{k=1}^n x_k} \right) \\
&= \sum_{k=1}^n \ln\left(\frac{m!}{(m-x_k)!x_k!}\right) + \left(\textstyle\sum_{k=1}^n x_k\right) \ln(p) + \left(n \cdot m - \textstyle\sum_{k=1}^n x_k\right) \ln(1-p).
\end{aligned}$$

To determine $\hat{p}_n^{ML}$ we consider the first order condition

$$\frac{d}{dp} \log \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) = 0,$$

which yields

$$\left(\textstyle\sum_{k=1}^n x_k\right) \frac{1}{p} - \left(n \cdot m - \textstyle\sum_{k=1}^n x_k\right) \frac{1}{1-p} = 0.$$

On account that $p \in (0,1)$, the latter becomes

$$\left(\textstyle\sum_{k=1}^n x_k\right)(1-p) - \left(n \cdot m - \textstyle\sum_{k=1}^n x_k\right) p = 0.$$

That is

$$\textstyle\sum_{k=1}^n x_k - n \cdot m \cdot p = 0,$$

which implies

$$p = \frac{\sum_{k=1}^n x_k}{n \cdot m} = \frac{\bar{x}_n}{m}.$$

In addition,

$$\begin{aligned}
\frac{d^2}{dp^2} \log \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) &= -\left(\textstyle\sum_{k=1}^n x_k\right) \frac{1}{p^2} - \left(n \cdot m - \textstyle\sum_{k=1}^n x_k\right) \frac{1}{(1-p)^2} \\
&= \frac{-\left(\sum_{k=1}^n x_k\right)(1-p)^2 - \left(n \cdot m - \sum_{k=1}^n x_k\right) p^2}{p^2 (1-p)^2} \\
&= \frac{-\sum_{k=1}^n x_k + 2\left(\sum_{k=1}^n x_k\right) p - n \cdot m \cdot p^2}{p^2 (1-p)^2} \\
&= \frac{-n\bar{x}_n + 2n\bar{x}_n p - n \cdot m \cdot p^2}{p^2 (1-p)^2} \\
&= -\frac{n}{p^2 (1-p)^2} \left(\bar{x}_n - 2\bar{x}_n p + m \cdot p^2\right).
\end{aligned}$$

Now, we have

$$\left(\bar{x}_n - 2\bar{x}_n p + m \cdot p^2\right)_{p = \frac{\bar{x}_n}{m}} = \left(\bar{x}_n - \frac{2}{m}\bar{x}_n^2 + m \cdot \frac{\bar{x}_n^2}{m^2}\right) = \bar{x}_n \left(1 - \frac{1}{m}\bar{x}_n\right).$$

On the other hand, we clearly have

$$\bar{x}_n \leq m,$$

for every $(x_1, \dots, x_n) \in \{0,1,\dots,m\}^n$. It follows

$$\frac{d^2}{dp^2} \log \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n) \leq 0$$

which implies that

$$p = \frac{\bar{x}_n}{m}$$

is a maximum for $\log \mathcal{L}_{X_1,\dots,X_n}(p; x_1, \dots, x_n)$. As a consequence, we obtain that the maximum likelihood estimator fo p is given by

$$\hat{p}_n^{ML} = \frac{\bar{X}_n}{m}.$$

3. Given that $m = 10$ and we observe a realization $4, 4, 3, 5, 6$ of a sample $X_1, \dots, X_5$ of size 5 drawn from X, we obtain

$$\hat{p}_n^M(\omega) = \frac{\bar{X}_5(\omega)}{10} = \frac{\frac{1}{5}(4+4+3+5+6)}{10} = 0.44.$$

4. We know that
$$\mathbf{E}[X] = m \cdot p \quad \text{and} \quad \mathbf{D}^2[X] = m \cdot p(1-p),$$
where p is the true value of the success parameter, As a consequence given the estimators $\hat{p}_n^M$ an estimator $\hat{\mu}_X$ [resp. $\hat{\sigma}_X^2$] of the expectation [resp. variance] of X is given by
$$\hat{\mu}_X = m \cdot \hat{p}_n^M \quad \text{and} \quad \hat{\sigma}_X^2 = m \cdot \hat{p}_n^M \left(1 - \hat{p}_n^M\right).$$
An estimate of $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$ by means of the estimator $\hat{p}_n^M$ and the information provided at 3. is the given by
$$\hat{\mu}_X(\omega) = m \cdot \hat{p}_n^M(\omega) = 10 \cdot 0.44 = 4.4$$
and
$$\hat{\sigma}_X^2(\omega) = m \cdot \hat{p}_n^M(\omega)\left(1 - \hat{p}_n^M(\omega)\right) = 10 \cdot 0.44 \cdot (1 - 0.44) = 2.464.$$
This completes the solution.

Problem 4 *Let X be a normally distributed random variable with unknown mean μ_X and variance σ_X^2. An investigator wants to estimate μ and σ^2 on the basis of a simple random sample $X_1, \ldots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the likelihood methods. What are the estimator $\hat{\mu}_n^{LM}$ and $\hat{\sigma}_n^{2^{LM}}$?*
2. *Assume the investigator applies the method of moments. What are the estimators $\hat{\mu}_n^M$ and $\hat{\sigma}_n^{2^M}$? Hint: guess what $\hat{\sigma}_n^{2^M}$ could be and get it!*

Solution.

1. We know that the joint density function $f_{X_1,\ldots,X_n} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ of the sample $X_1, \ldots, X_n$ is given by
$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\mu,\sigma) \stackrel{\text{def}}{=} \prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_k-\mu)^2}{2\sigma^2}}, \quad \forall (x_1,\ldots,x_n) \in \mathbb{R}^n, \quad \forall (\mu,\sigma) \in \mathbb{R} \times \mathbb{R}_{++},$$
that is
$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\mu,\sigma) = \frac{1}{\sqrt{2^n\pi^n}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k-\mu)^2}.$$
Hence, the likelihood function $\mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$ of the sample is given by
$$\mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = \frac{1}{\sqrt{2^n\pi^n}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k-\mu)^2}, \quad \forall (\mu,\sigma) \in \mathbb{R}\times\mathbb{R}_{++}, \quad \forall (x_1,\ldots,x_n) \in \mathbb{R}^n.$$
Thanks to the stucture of $\mathcal{L}_{X_1,\ldots,X_n}$ it is convenient to consider the log-likelihood function $\log \mathcal{L}_{X_1,\ldots,X_n} : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ of the sample which is given by
$$\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) \stackrel{\text{def}}{=} (\log \circ \mathcal{L}_{X_1,\ldots,X_n})(\mu,\sigma;x_1,\ldots,x_n), \quad \forall (\mu,\sigma) \in \mathbb{R}\times\mathbb{R}_{++}, \quad \forall (x_1,\ldots,x_n) \in \mathbb{R}^n$$
that is
$$\log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = -n\left(\frac{1}{2}\ln(2\pi) + \ln(\sigma)\right) - \frac{1}{2\sigma^2}\sum_{k=1}^n (x_k-\mu)^2.$$
Now, to determine $\underset{(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\ldots,X_n}$ we can consider the first order conditions
$$\frac{\partial}{\partial\mu} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 \quad \text{and} \quad \frac{\partial}{\partial\sigma} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0.$$
We have
$$\frac{\partial}{\partial\mu} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = \frac{1}{\sigma^2}\sum_{k=1}^n (x_k-\mu)$$
and
$$\frac{\partial}{\partial\sigma} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = \frac{1}{\sigma}\left(\frac{1}{\sigma^2}\sum_{k=1}^n (x_k-\mu)^2 - n\right).$$
Therefore,
$$\begin{aligned}\frac{\partial}{\partial\mu} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 &\Rightarrow \sum_{k=1}^n (x_k-\mu) = 0,\\ &\Rightarrow \mu = \frac{1}{n}\sum_{k=1}^n x_k,\\ &\Rightarrow \mu = \bar{x}_n\end{aligned}$$
and
$$\begin{aligned}\frac{\partial}{\partial\sigma} \log \mathcal{L}_{X_1,\ldots,X_n}(\mu,\sigma;x_1,\ldots,x_n) = 0 &\Rightarrow \frac{1}{\sigma^2}\sum_{k=1}^n (x_k-\mu)^2 - n,\\ &\Rightarrow \sigma^2 = \frac{1}{n}\sum_{k=1}^n (x_k-\mu)^2,\\ &\Rightarrow \sigma^2 = \frac{1}{n}\sum_{k=1}^n (x_k-\bar{x}_n)^2.\end{aligned}$$
The above equations imply that we have
$$\hat{\mu}_n^{LM} = \bar{X}_n \quad \text{and} \quad \hat{\sigma}_n^{2^{LM}} = \tilde{S}_n^2(X),$$
where $\bar{X}_n$ [resp. $\tilde{S}_n^2(X)$] is the sample mean [resp. unbiased sample variance] of $X_1, \ldots, X_n$.

2. We know that
$$\mathbf{E}[X] = \mu \quad \text{and} \quad \mathbf{E}\left[X^2\right] = \mu^2 + \sigma^2.$$
Hence, applying the method of moments, the investigator writes
$$\frac{1}{n}\sum_{k=1}^n X_k = \hat{\mu}_n^M \quad \text{and} \quad \frac{1}{n}\sum_{k=1}^n X_k^2 = \left(\hat{\mu}_n^M\right)^2 + \hat{\sigma}_n^{2^M}.$$
The first of the two equations clearly yields
$$\hat{\mu}_n^M = \bar{X}_n.$$
The second equation, on account of the first, yields
$$\hat{\sigma}_n^{2^M} = \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2.$$
On the other hand,
$$\begin{aligned}\tilde{S}_n^2(X) &= \frac{1}{n}\sum_{k=1}^n \left(X_k - \bar{X}_n\right)^2\\ &= \frac{1}{n}\sum_{k=1}^n \left(X_k^2 - 2X_k\bar{X}_n + \bar{X}_n^2\right)\\ &= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2\bar{X}_n\sum_{k=1}^n X_k + \sum_{k=1}^n \bar{X}_n^2\right)\\ &= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2\right)\\ &= \frac{1}{n}\left(\sum_{k=1}^n X_k^2 - n\bar{X}_n^2\right)\\ &= \frac{1}{n}\sum_{k=1}^n X_k^2 - \bar{X}_n^2\end{aligned}$$

It follows that

$$\hat{\sigma}_n^{2^M} = \tilde{S}_n^2(X).$$

This completes the solution.

Problem 5 *Let X a random variable representing a characteristic of a certain population. Assume that X has a density $f_X : \mathbb{R} \to \mathbb{R}$ given by*

$$f_X(x) \overset{def}{=} \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x), \quad \forall x \in \mathbb{R},$$

where θ is a positive parameter. esonero 2022

1. *Apply the method of moments to find the estimator $\hat{\theta}_M$ of the parameter θ.*
2. *Apply the maximum likelihood method to find the estimator $\hat{\theta}_{ML}$ of the parameter θ.*
3. *Use the estimators $\hat{\theta}_M$ and $\hat{\theta}_{ML}$ to build estimators for $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*

Solution.

1. We have

$$\begin{aligned}
\mathbf{E}[X] &= \int_{\mathbb{R}} x f_X(x)\, d\mu_L(x) = \int_{\mathbb{R}} x \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x)\, d\mu_L(x) = \int_{[3,+\infty)} x \frac{1}{\theta} e^{-\frac{x-3}{\theta}}\, d\mu_L(x) \\
&= \int_3^{+\infty} x \frac{1}{\theta} e^{-\frac{x-3}{\theta}}\, dx = \lim_{x\to+\infty} \int_3^x u \frac{1}{\theta} e^{-\frac{u-3}{\theta}}\, du = \lim_{x\to+\infty} \int_0^{\frac{x-3}{\theta}} (\theta v + 3) e^{-v}\, dv \\
&= \lim_{x\to+\infty} \left(\theta \int_0^{\frac{x-3}{\theta}} v e^{-v}\, dv + 3 \int_0^{\frac{x-3}{\theta}} e^{-v}\, dv \right) = \theta \lim_{x\to+\infty} \int_0^{\frac{x-3}{\theta}} v e^{-v}\, dv + 3 \lim_{x\to+\infty} \int_0^{\frac{x-3}{\theta}} e^{-v}\, dv.
\end{aligned}$$

Now,

$$\int_0^{\frac{x-3}{\theta}} e^{-v}\, dv = -\int_0^{-\frac{x-3}{\theta}} e^{v}\, dv = \int_{-\frac{x-3}{\theta}}^0 e^{v}\, dv = e^v\big|_{-\frac{x-3}{\theta}}^0 = 1 - e^{-\frac{x-3}{\theta}}$$

and

$$\begin{aligned}
\int_0^{\frac{x-3}{\theta}} v e^{-v}\, dv &= -\int_0^{\frac{x-3}{\theta}} v\, de^{-v} = -\left(v e^{-v}\big|_0^{\frac{x-3}{\theta}} - \int_0^{\frac{x-3}{\theta}} e^{-v}\, dv \right) \\
&= -\left(v e^{-v}\big|_0^{\frac{x-3}{\theta}} - \left(1 - e^{-\frac{x-3}{\theta}}\right)\right) = -\left(\frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} + e^{-\frac{x-3}{\theta}} - 1 \right) \\
&= 1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}}.
\end{aligned}$$

It follows

$$\mathbf{E}[X] = \theta \lim_{x\to+\infty} \left(1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}} \right) + 3 \lim_{x\to+\infty} \left(1 - e^{-\frac{x-3}{\theta}} \right) = \theta + 3.$$

As a consequence, setting

$$\frac{1}{n} \textstyle\sum_{k=1}^n X_k = \hat{\theta}_n^M + 3$$

we obtain

$$\hat{\theta}_n^M = \bar{X}_n - 3.$$

2. We know that the joint density function $f_{X_1,\dots,X_n} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ of a simple random sample $X_1, \dots, X_n$ drawn from X is given by

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n;\theta) \overset{\text{def}}{=} \textstyle\prod_{k=1}^n \frac{1}{\theta} e^{-\frac{x_k-3}{\theta}} 1_{[3,+\infty)}(x_k), \quad \forall (x_1,\dots,x_n) \in \mathbb{R}^n, \quad \forall \theta \in \mathbb{R}_{++},$$

that is

$$\begin{aligned}
f_{X_1,\dots,X_n}(x_1,\dots,x_n;\theta) &= \left(\textstyle\prod_{k=1}^n \frac{1}{\theta} \right) \left(\textstyle\prod_{k=1}^n e^{-\frac{x_k-3}{\theta}} \right) \textstyle\prod_{k=1}^n 1_{[3,+\infty)}(x_k) \\
&= \frac{1}{\theta^n} e^{-\frac{1}{\theta}\left(\sum_{k=1}^n x_k - 3n\right)} \textstyle\prod_{k=1}^n 1_{[3,+\infty)}(x_k).
\end{aligned}$$

Hence, the likelihood function $\mathcal{L}_{X_1,\dots,X_n} : \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$ of the sample is given by

$$\mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\left(\sum_{k=1}^n x_k - 3n\right)} \textstyle\prod_{k=1}^n 1_{[3,+\infty)}(x_k), \quad \forall \theta \in \mathbb{R}_{++}, \quad \forall (x_1,\dots,x_n) \in \mathbb{R}^n.$$

Note that

$$\mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = \begin{cases} \frac{1}{\theta^n} e^{-\frac{1}{\theta}\left(\sum_{k=1}^n x_k - 3n\right)} > 0, & \text{if } (x_1,\dots,x_n) \in \times_{k=1}^n [3,+\infty), \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$\underset{\theta\in\mathbb{R}_{++}}{\arg\max} \mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = \underset{\theta\in\mathbb{R}_{++}}{\arg\max} \frac{1}{\theta^n} e^{-\frac{1}{\theta}\left(\sum_{k=1}^n x_k - 3n\right)}$$

Hence, we can consider as the log-likelihood function of the sample $X_1,\dots,X_n$ the function $\log \mathcal{L}_{X_1,\dots,X_n} : \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$ given by

$$\log \mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) \overset{\text{def}}{=} \log \left(\frac{1}{\theta^n} e^{-\frac{1}{\theta}\left(\sum_{k=1}^n x_k - 3n\right)} \right) = -n \ln(\theta) - \frac{1}{\theta} \left(\textstyle\sum_{k=1}^n x_k - 3n \right)$$

Now, to determine $\underset{\theta\in\mathbb{R}_{++}}{\arg\max} \log \mathcal{L}_{X_1,\dots,X_n}$ we can consider the first order condition

$$\frac{\partial}{\partial\theta} \log \mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = 0.$$

We have

$$\frac{\partial}{\partial\theta} \log \mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = -\frac{n}{\theta} + \left(\textstyle\sum_{k=1}^n x_k - 3n \right) \frac{1}{\theta^2}$$

Therefore,

$$\begin{aligned}
\frac{\partial}{\partial\theta} \log \mathcal{L}_{X_1,\dots,X_n}(\theta; x_1,\dots,x_n) = 0 &\Rightarrow \frac{1}{\theta} \left(\left(\textstyle\sum_{k=1}^n x_k - 3n \right) \frac{1}{\theta} - n \right) = 0 \\
&\Rightarrow \theta = \frac{1}{n} \left(\textstyle\sum_{k=1}^n x_k - 3n \right) \\
&\Rightarrow \theta = \frac{1}{n} \textstyle\sum_{k=1}^n x_k - 3.
\end{aligned}$$

The above equation imply that we have

$$\hat{\theta}_n^{LM} = \bar{X}_n - 3.$$

3. Note that
$$\hat{\theta}_n^M = \hat{\theta}_n^{LM} \equiv \bar{X}_n - 3.$$
Since
$$\mathbf{E}[X] = \theta + 3,$$
it clearly follows that the desired estimator for $\mathbf{E}[X]$ is $\bar{X}_n$.

To build an estimator for $\mathbf{D}^2[X]$ we have first to determine the second order moment $\mathbf{E}[X^2]$ of X. We have
$$\begin{aligned}\mathbf{E}\left[X^2\right] &= \int_{\mathbb{R}} x^2 f_X(x)\ d\mu_L(x) = \int_{\mathbb{R}} x^2 \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x)\ d\mu_L(x) = \int_{[3,+\infty)} x^2 \frac{1}{\theta} e^{-\frac{x-3}{\theta}}\ d\mu_L(x) \\ &= \int_3^{+\infty} x^2 \frac{1}{\theta} e^{-\frac{x-3}{\theta}}\ dx = \lim_{x\to+\infty} \int_3^x u^2 \frac{1}{\theta} e^{-\frac{u-3}{\theta}}\ dx = \lim_{x\to+\infty} \int_0^{\frac{x-3}{\theta}} (\theta v + 3)^2 e^{-v}\ dv \\ &= \lim_{x\to+\infty} \int_0^{\frac{x-3}{\theta}} \left(\theta^2 v^2 + 6\theta v + 9\right) e^{-v}\ dv \\ &= \lim_{x\to+\infty} \left(\theta^2 \int_0^{\frac{x-3}{\theta}} v^2 e^{-v}\ dv + 6\theta \int_0^{\frac{x-3}{\theta}} v e^{-v}\ du + 9 \int_0^{\frac{x-3}{\theta}} e^{-v}\ dv\right)\end{aligned}$$
We already know that
$$\int_0^{\frac{x-3}{\theta}} e^{-v}\ dv = 1 - e^{-\frac{x-3}{\theta}} \quad \text{and} \quad \int_0^{\frac{x-3}{\theta}} v e^{-v}\ dv = 1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}}.$$
Hence, we compute
$$\begin{aligned}\int_0^{\frac{x-3}{\theta}} v^2 e^{-v}\ dv &= -\int_0^{\frac{x-3}{\theta}} v^2\ de^{-v} = -\left(v^2 e^{-v}\Big|_0^{\frac{x-3}{\theta}} - 2\int_0^{\frac{x-3}{\theta}} v e^{-v}\ dv\right) \\ &= 2\left(1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}}\right) - \left(\frac{x-3}{\theta}\right)^2 e^{-\frac{x-3}{\theta}}.\end{aligned}$$
We then have
$$\begin{aligned}\mathbf{E}\left[X^2\right] &= \lim_{x\to+\infty} \theta^2 \left(2\left(1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}}\right) - \left(\frac{x-3}{\theta}\right)^2 e^{-\frac{x-3}{\theta}}\right) \\ &+ \lim_{x\to+\infty} 6\theta \left(1 - \frac{x-3}{\theta} e^{-\frac{x-3}{\theta}} - e^{-\frac{x-3}{\theta}}\right) + 9 \lim_{x\to+\infty} \left(1 - e^{-\frac{x-3}{\theta}}\right) \\ &= 2\theta^2 + 6\theta + 9.\end{aligned}$$
It follows
$$\mathbf{D}^2[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2 = 2\theta^2 + 6\theta + 9 - (\theta+3)^2 = \theta^2.$$
Therefore, the desired estimators for $\mathbf{D}^2[X]$ is
$$\left(\hat{\theta}_n^M\right)^2 = \left(\bar{X}_n - 3\right)^2 = \bar{X}_n^2 - 6\bar{X}_n + 9.$$
This completes the solution.

Problem 6 *Assume that the returns of a stock in a financial market are Gaussian distributed with unknown mean μ and variance σ^2. Let X be the normal random variable representing the realization of the returns and let $X_1, \ldots, X_n$ be a simple random sample of size n drawn from X. Assume that $n = 5$ and the realizations of the sample are*
$$x_1 \equiv -1.5, \quad x_2 \equiv -0.5, \quad x_3 \equiv 1.5, \quad x_4 \equiv 2.0, \quad x_5 \equiv 2.5$$

1. *Determine a 99% confidence interval for the mean μ.*
2. *Find the confidence for an interval of width 0.1.*
3. *Determine a 90% confidence interval for the standard deviation σ.*

Solution.

1. From data we obtain
$$\bar{x}_5 \equiv \frac{1}{5}\textstyle\sum_{k=1}^5 x_k = 0.8$$
and
$$s_5^2(X) = \frac{1}{4}\textstyle\sum_{k=1}^5 (x_k - \bar{x}_5)^2 = 2.95 \Rightarrow s_5(X) = 1.72$$
Now, since X is Gaussian distributed with unknown variance and the size of the sample is small, to determine a $100(1-\alpha)\%$ confidence interval for the mean μ the statistic to be considered is
$$\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} \sim t_{n-1}.$$
The achievement of a $100(1-\alpha)\%$ confidence interval requires to use the $\alpha/2$ critical value $t_{n-1,\alpha/2}$ of t_{n-1} for $\alpha = 0.01$. In fact, we have
$$\begin{aligned}1-\alpha &= \mathbf{P}\left(-t_{n-1,\alpha/2} < \frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} < t_{n-1,\alpha/2}\right) \\ &= \mathbf{P}\left(-\left(\bar{X}_n + t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right) < -\mu < -\left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right)\right) \\ &= \mathbf{P}\left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right)\end{aligned}$$
It follows that the desired confidence interval for μ is given by the random interval
$$\left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}},\ \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right)$$
A realization of such a confidence interval is then given by
$$\left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n(X)}{\sqrt{n}},\ \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n(X)}{\sqrt{n}}\right).$$
In the case under scrutiny, since $t_{n-1,\alpha/2} \equiv t_{4,0.005} = 4.60$, $\bar{x}_n \equiv \bar{x}_5 = 0.80$, $s_n \equiv s_5 = 1.72$, the realization of the confidence interval becomes
$$(-3.16,\ 4.76).$$

2. From 1. it is clearly seen the width w of a $100(1-\alpha)\%$ confidence interval is given by

$$w = 2t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}.$$

As a consequence, the size n of the sample which gives a $100(1-\alpha)\%$ confidence interval of a given widht w is given by the solution of the equation

$$\frac{n}{t^2_{n-1,\alpha/2}} = \left[4\frac{S^2_n(X)}{w^2}\right] + 1.$$

To determine n we need to consider as many realizations $x_1, \ldots, x_n$ of the simple random sample $X_1, \ldots, X_n$ such that

$$\frac{n}{t^2_{n-1,\alpha/2}} = \left[4\frac{s^2_n(X)}{0.01}\right] + 1.$$

3. Again, since X is Gaussian distributed and the size of the sample is small, to determine a $100(1-\alpha)\%$ confidence interval for the standard deviation σ the statistic to be considered is

$$\frac{(n-1)S^2_n(X)}{\sigma^2} \sim \chi^2_{n-1}.$$

Since χ^2_{n-1} is not symmetric, the achievement of a $100(1-\alpha)\%$ confidence interval requires to exploit the $\alpha/2$ and the $1-\alpha/2$ critical value $\chi^2_{n-1,\alpha/2,-} \equiv \chi^2_{n-1,\alpha/2}$ and $\chi^2_{n-1,\alpha/2,+} = \chi^2_{n-1,1-\alpha/2}$ of χ^2_{n-1} for $\alpha = 0.1$, where $\chi^2_{n-1,\alpha/2}$ [resp. $\chi^2_{n-1,1-\alpha/2}$] is the $\alpha/2$-quantile [$1-\alpha/2$-quantile] of the χ^2_{n-1} distribution. In fact, we have

$$1-\alpha = \mathbf{P}\left(\chi^2_{n-1,\alpha/2,-} < \frac{(n-1)S^2_n(X)}{\sigma^2} < \chi^2_{n-1,\alpha/2,+}\right) = \mathbf{P}\left(\frac{(n-1)S^2_n(X)}{\chi^2_{n-1,\alpha/2,+}} < \sigma^2 < \frac{(n-1)S^2_n(X)}{\chi^2_{n-1,\alpha/2,-}}\right).$$

It follows that the desired confidence interval for the variance σ^2 is given by

$$\left(\frac{(n-1)S^2_n(X)}{\chi^2_{n-1,\alpha/2,+}}, \frac{(n-1)S^2_n(X)}{\chi^2_{n-1,\alpha/2,-}}\right) = \left(\frac{(n-1)S^2_n(X)}{\chi^2_{n-1,1-\alpha/2}}, \frac{(n-1)S^2_n(X)}{\chi^2_{n-1,\alpha/2}}\right).$$

In the case under scrutiny, since $\chi^2_{n-1,\alpha/2} \equiv \chi^2_{4,0.5} = 0.71$, $\chi^2_{n-1,1-\alpha/2} \equiv \chi^2_{4,0.95} = 9.49$, $\bar{x}_n \equiv \bar{x}_5 = 0.80$, $s^2_n(X) \equiv s^2_5(X) = 2.95$, a realization of the confidence interval is given by

$$\left(\frac{4s^2_n(X)}{\chi^2_{4,0.95}}, \frac{4s^2_n(X)}{\chi^2_{4,0.05}}\right) = \left(\frac{4\cdot 2.95}{9.49}, \frac{4\cdot 2.95}{0.71}\right) = (1.24,\ 16.62).$$

As a consequence the $100(1-0.1)\%$ confidence interval for the standard deviation σ is

$$(1.11,\ 4.08).$$

This completes the solution.

Problem 7 *Assume that a library master believes that the mean duration in days of the borrowing period is* $20d$. *However, the library master selects a simple random sample of* 100 *books in the library and discovers that the sample mean and variance of the borrowing days are* $18d$ *and* $8d^2$, *respectively. Determine a* 99% *confidence interval for the mean duration of the borrowing days to check whether library master's initial guess is correct.*

Solution. Note that the distribution of the random variable X representing the duration in days of the borrowing period is unknown. However, are known the sample mean and variance realizations referred to a simple sample of size $n = 100$, which may be considered a large sample. In this case the statistic to be considered is given by

$$\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} \to Z,$$

where $Z \sim N(0,1)$. The achievement of a $100(1-\alpha)\%$ confidence interval requires to exploit the $\alpha/2$ critical value $z_{\alpha/2}$ of Z for $\alpha = 0.01$. In fact, we have

$$\begin{aligned}1-\alpha &\approx \mathbf{P}\left(-z_{\alpha/2} < \frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} < z_{\alpha/2}\right)\\ &= \mathbf{P}\left(-\left(\bar{X}_n + z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right) < -\mu < -\left(\bar{X}_n - z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right)\right)\\ &= \mathbf{P}\left(\bar{X}_n - z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}} < \mu < \bar{X}_n + z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right)\end{aligned}$$

It follows that the desired confidence interval for μ is given by

$$\left(\bar{X}_n - z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}},\ \bar{X}_n + z_{\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right).$$

In the case under scrutiny,

$$n = 100, \quad \bar{x}_n \equiv \bar{x}_{100} = 18, \quad s_n(X) \equiv s_{100}(X) = \sqrt{8}, \quad z_{\alpha/2} \equiv 2.58.$$

Therefore, a realization of the confidence interval is given by

$$\left(\bar{x}_n - z_{\alpha/2}\frac{s_n(X)}{\sqrt{n}},\ \bar{x}_n + z_{\alpha/2}\frac{s_n(X)}{\sqrt{n}}\right) = \left(18 - 2.58\cdot\frac{\sqrt{8}}{\sqrt{100}}, 18 + 2.58\cdot\frac{\sqrt{8}}{\sqrt{100}}\right) = (17.27,\ 18.73).$$

It follows that library master's initial guess is not supported by data. Note that this problem can be tackled also exploiting the hypothesis test method. In fact, assume as the null hypothesis that library master's assumption is correct, that is $H_0 : \mu = \mu_0$, and as the alternative hypothesis that library master's assumption is wrong, that is $H_0 : \mu \neq \mu_0$. The same consideration as above on the available information on the random variable X led to consider the rejection region

$$R = \left\{\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} < -z_{\alpha/2}\right\} \cup \left\{\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} > z_{\alpha/2}\right\} = \left\{\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} < -2.58\right\} \cup \left\{\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}} > 2.58\right\},$$

where $\mu = 20$. Computing the statistic $\frac{\bar{X}_n - \mu}{S_n(X)/\sqrt{n}}$ for the available realization, we obtain

$$\frac{\bar{x}_n - \mu}{s_n(X)/\sqrt{n}} = \frac{18-20}{\sqrt{8}/10} = -7.07 \in R.$$

Hence, the library master's assumption has to be rejected.

Problem 8 *The mark of a infamous exam of Probability and Statistics are normally distributed with standard deviation* $\sigma = 2$. *A simple random sample of nine students is selected end the following evaluations are computed*

$$\sum_{k=1}^{9} x_k = 237 \quad \text{and} \quad \sum_{k=1}^{9} x_k^2 = 6295.$$

1. *Find a 90% confidence interval for the mean mark.*
2. *Discuss, without computation, whether the lenght of a 95% confidence interval would be smaller, greater or equal than the lenght of the interval previously determined.*
3. *How large the minimum sample size should be to obtain a 90% confidence interval for the mean mark with width equal to 3? Besides the confidence interval method is it possible to apply the Tchebychev inequality?*

Solution.

1. We have

$$\bar{x}_n = \frac{1}{n}\sum_{\ell=1}^{n} x_\ell$$

and

$$\begin{aligned} s_n^2(X) &= \frac{1}{n-1}\sum_{k=1}^{n}(x_k-\bar{x}_n)^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n}\left(x_k^2-2x_k\bar{x}_n+\bar{x}_n^2\right)\right) \\ &= \frac{1}{n-1}\left(\sum_{k=1}^{n}x_k^2-2\bar{x}_n\sum_{k=1}^{n}x_k+\sum_{k=1}^{n}\bar{x}_n^2\right) = \frac{1}{n-1}\left(\sum_{k=1}^{n}x_k^2-2n\bar{x}_n^2+n\bar{x}_n^2\right) \\ &= \frac{1}{n-1}\left(\sum_{k=1}^{n}x_k^2-n\bar{x}_n^2\right). \end{aligned}$$

Hence, in our case

$$n=9, \quad \bar{x}_n \equiv \bar{x}_9 = \frac{1}{9}237 = 26.33, \quad s_n^2(X) \equiv s_9^2(X) = \frac{1}{8}\left(6295-\frac{1}{9}237^2\right) = 6.75.$$

Now, the random variable X representing the mark of the exam is normally distributed and the size of the sample considered is small. On the other hand, we know the variance of X and we also know the realization of the sample variance. Therefore, we can considered two different approaches.

(i) Exploit the statistic

$$\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}} \sim Z$$

where $Z \sim N(0,1)$.

(ii) Exploit the statistic

$$\frac{\bar{X}_n-\mu}{S_n(X)/\sqrt{n}} \sim t_{n-1}.$$

where t_{n-1} is Student distributed with $n-1$ degrees of freedom.

In the first case, a $100(1-\alpha)\%$ confidence interval requires to consider the $\alpha/2$ critical value $z_{\alpha/2}$ of Z for $\alpha = 0.01$. It follows that the desired confidence interval for μ is given by

$$\left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right),$$

On account that $z_{\alpha/2} \equiv z_{0.05} \equiv 1.64$ and $\sigma = 2$, a realization of the confidence interval, is then given by

$$I_\alpha = \left(\bar{x}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = \left(26.33-1.64\frac{2}{3},\ 26.33+1.64\frac{2}{3}\right) = (25.24,\ 27.42)$$

In the second case, a $100(1-\alpha)\%$ confidence interval requires to consider the $\alpha/2$ critical value $t_{n-1,\alpha/2}$ of T for $\alpha = 0.1$. Hence, the desired confidence interval for μ is given by

$$\left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}},\ \bar{X}_n - t_{n-1,\alpha/2}\frac{S_n(X)}{\sqrt{n}}\right),$$

On account that $t_{n-1,\alpha/2} \equiv t_{8,0.05} = 1.86$, a realization of the confidence interval, is then given by

$$J_\alpha = \left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n(X)}{\sqrt{n}},\ \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n(X)}{\sqrt{n}}\right) = \left(26.33-1.86\frac{\sqrt{6.75}}{3},\ 26.33+1.86\frac{\sqrt{6.75}}{3}\right) = (24.72,\ 2$$

Note that

$$I_\alpha \subset J_\alpha.$$

2. The lenght of a 95% confidence interval will be greater then the lenght of the interval previously determined. There is a theorem stating that the higher the confidence is, the larger the confidence interval is.

3. Considering Case (i), the width w of a 90% confidence interval for the mean mark is given by

$$w = 2\ z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Solving for n we obtain

$$n = \frac{4\ z_{\alpha/2}^2\sigma^2}{w^2}$$

Setting $w = 3$, and keeping all other parameters unchanged, it follows

$$n = 12.$$

With this value of n, keeping all other parameters unchanged, the confidence interval becomes

$$(24.84,\ 27.82),$$

which is narrower than I_α.

FROM NOW ON THE SOLUTIONS TO THE PROBLEMS HAVE NOT BEEN REVIEWED YET. HOWEVER, THEY SHOULD BE ESSENTIALLY CORRECT. PLEASE, LET ME KNOW WHETHER YOU FIND INCONGRUENCES.

Problem 9 *Let $X_1,\dots,X_n,X_{n+1}$ be a simple random sample of size $n+1$ drawn from a Gaussian distributed random variable X with unknown mean μ and variance σ^2. Assume that we have observed $X_1,\dots,X_n$ and we want use the observed values $x_1,\dots,x_n$ to determine a confidence interval for the prediction of X_{n+1}. To this goal give detailed answers to the following questions:*

1. *what is the distribution of the statistic $\bar{X}_n$?*
2. *what is the distribution of the statistic $\left(X_{n+1}-\bar{X}_n\right)/\sigma\sqrt{1+1/n}$?*

3. *are the statistics* $X_{n+1}-\bar{X}_n$ *and* $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}_n\right)^2$ *independent?*

4. *what is the distribution of the statistic* $\left(X_{n+1}-\bar{X}_n\right)/S_n\sqrt{1+1/n}$*?*

Exercise 10 *After answering the above questions, build an interval in which the random variable* X_{n+1} *takes its values with probability* α *and determine the corresponding confidence interval for the prediction of* X_{n+1}*. In the end, assume that* $n=7$ *and we have*

$$x_1=7005,\quad x_2=7432,\quad x_3=7420,\quad x_4=6822,\quad x_5=6752,\quad x_6=5333,\quad x_7=6552.$$

compute the 95% *confidence interval for the prediction of* X_8*.*

Problem 11 *Let* X *be a Gaussian random variable with unknown mean* μ_X *and variance* σ_X^2 *representing a certain characteristic of a population. Assume that testing the sample mean* $\bar{X}_n$ *and the sample standard deviation* S_n *of a simple random sample* $X_1,\dots,X_n$ *of size* $n\equiv 9$ *drawn from* X *we obtain the value* $\bar{X}_n(\omega)\equiv\bar{x}_n=251.50cm$ *and* $S_n(\omega)\equiv s_n=2.30cm$*.*

1. *Considering both the rejection region method and the p-value method, should the null hypothesis* $H_0:\mu_X=250cm$ *be rejected against the alternative* $H_a:\mu_X\neq 250cm$ *at the significance level* $\alpha=0.1$*?*

2. *Considering both the rejection region method and the p-value method, should the null hypothesis* $H_0:\sigma_X^2=4$ *be rejected against the alternative* $H_a:\sigma_X^2>4$ *at the significance level* $\alpha=0.05$*? Calculate the probability* $\beta(5)$ *of a II type error.*

Solution.

1. Since X is Gaussian distributed with unknown mean and variance and the size of the sample is small, the statistic to be used is
$$\frac{\bar{X}_n-\mu_X}{S_n/\sqrt{n}}. \tag{1}$$
Consider testing the null hypothesis $H_0:\mu_X=\mu_0$, where $\mu_0\equiv 250cm$, against the alternative $H_1:\mu_X\neq\mu_0$ at the significance level $\alpha=0.1$. Under the assumption that the null hypothesis is true the statistic (1) with $\mu_X=\mu_0$ has the Student distribution with $n-1$ degree of freedom, that is
$$\frac{\bar{X}_n-\mu_0}{S_n/\sqrt{n}}\sim T_{n-1}.$$
Moreover, the stucture of the alternative hypothesis calls for a rejection region of the form
$$R=\left\{T_{n-1}<t_{n-1,1-\alpha/2}\right\}\cup\left\{T_{n-1}>t_{n-1,\alpha/2}\right\}.$$
where,
$$t_{n-1,\alpha/2}=t_{8,0.05}=1.860\quad\text{and}\quad t_{n-1,1-\alpha/2}=-t_{n-1,\alpha/2}=-t_{8,0.05}=-1.860.$$
Hence,
$$R=(-\infty,-1.860)\cup(1.860,+\infty)$$
Computing the realization of the statistic, we have
$$\frac{\bar{X}_n(\omega)-\mu_X}{S_n(\omega)/\sqrt{n}}=\frac{\bar{x}_n-\mu_0}{s_n/\sqrt{9}}=\frac{251.50-250}{2.30/3}=1.96\in R.$$
This implies a rejection of the null hypothesis in favor of the alternative. Adopting the p-value method, we recall that, on account of the alternative hypothesis, the p-value is the probability that the absolute value of the test statistic under the null assumption yields a value not less than the realization of the statistic. We will reject the null hypothesis when the computed p-value is smaller that the given significance level α. In symbols,
$$p=\mathbf{P}\left(|T_{n-1}|\geq\frac{\bar{x}_n-\mu_0}{s_n/\sqrt{9}}\mid H_0=T\right)=\mathbf{P}\left(T_8\leq-1.96\right)+\mathbf{P}\left(T_8\geq 1.96\right)=0.087<0.1,$$
which confirms the rejection of the null hypothesis.

2. Since ve are interested in testing a hypothesis on the variance of X, which is normally distributed with unknown variance and the size of the sample is small, the statistic to be used is
$$\frac{(n-1)S_n^2}{\sigma_X^2}. \tag{2}$$
Consider testing the null hypothesis $H_0:\sigma_X^2=\sigma_0^2$, where $\sigma_0^2\equiv 4$ against the alternative $H_1:\sigma_X^2>\sigma_0^2$ at the significance level $\alpha\equiv 0.05$. Under the assumption that the null hypothesis is true the statistic (2) with $\sigma_X^2=\sigma_0^2$ has the chi-square distribution with $n-1$ degrees of freedom, that is
$$\frac{(n-1)S_n^2}{\sigma_0^2}\sim\chi_{n-1}^2.$$
Hence, the upper tail rejection region is given by
$$R=\left\{\chi_{n-1}^2>\chi_{n-1,\alpha}^2\right\}$$
where, the upper $\alpha=0.05$ critical value $\chi_{n-1,\alpha}^2$ of the chi-square distribution with $n-1=8$ degrees of freedom is given by
$$\chi_{8,0.05}^2\simeq 15.51.$$
Hence,
$$R=(15.51,+\infty)$$
Computing the realization of the statistic, we have
$$\frac{(n-1)S_n^2(\omega)}{\sigma_X^2}=\frac{(n-1)s_n^2}{\sigma_0^2}=\frac{8\cdot 2.30^2}{4}=10.58\notin R$$
Hence, the realization of the statistic does not belong to the rejection region. This implies that H_0 cannot be rejected.

In terms of p-value we have to compute
$$\mathbf{P}\left(\chi_{n-1}^2\geq\frac{(n-1)s_n^2}{\sigma_0^2}\mid H_0\text{ true}\right)=\mathbf{P}\left(\chi_8^2\geq 10.58\right)=1-\mathbf{P}\left(\chi_8^2\leq 10.58\right)=0.227>0.05.$$
The p-value method confirms that H_0 cannot be rejected.

With regard to the evaluation of $\beta(5)$, setting $\sigma_1^2 \equiv 5$, we have

$$
\begin{aligned}
\beta(5) &= \mathbf{P}\left(\text{accept } H_0 \mid \sigma_X^2 = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_0^2} \notin R \mid \sigma_X = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_0^2}\frac{\sigma_1^2}{\sigma_1^2} \notin R \mid \sigma_X = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_1^2}\frac{\sigma_1^2}{\sigma_0^2} \notin R \mid \sigma_X = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_X^2}\frac{\sigma_1^2}{\sigma_0^2} \notin R\right) \\
&= \mathbf{P}\left(\chi_{n-1}^2\frac{\sigma_1^2}{\sigma_0^2} \notin R\right) \\
&= \mathbf{P}\left(\chi_8^2\frac{\sigma_1^2}{\sigma_0^2} \leq 15.51\right) \\
&= \mathbf{P}\left(\chi_8^2 \leq 15.51 \cdot \frac{\sigma_0^2}{\sigma_1^2}\right) \\
&= \mathbf{P}\left(\chi_8^2 \leq 8.46\right) \\
&= 0.61.
\end{aligned}
$$

.

Problem 12 *Let X be a Gaussian random variable with unknown mean μ and variance σ^2 representing a certain characteristic of a population and let $X_1, \ldots, X_n$ be a simple random sample of size n drawn from X. Assume that $n = 25$ and that the realizations $x_1, \ldots, x_{25}$ of the sample give an information summarized by*

$$\textstyle\sum_{k=1}^{25} x_k = 100 \quad \text{and} \quad \sum_{k=1}^{25} x_k^2 = 560$$

1. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \sigma^2 = 4$ be rejected against of the alternative $H_1 : \sigma^2 > 4$ with a significance level $\alpha = 0.05$? Calculate the probability $\beta(5)$ of a II type error.*

2. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \sigma^2 = 4$ be rejected against of the alternative $H_1 : \sigma^2 \neq 4$ with a significance level $\alpha = 0.05$? Calculate the probability $\beta(5)$ of a II type error.*

Solution. Note that the given information allows the knowledge of the realization of sample variance, which is given by

$$
\begin{aligned}
s_n^2 &= \frac{1}{n-1}\textstyle\sum_{k=1}^n\left(x_k - \frac{1}{n}\sum_{\ell=1}^n x_\ell\right)^2 \\
&= \frac{1}{n-1}\textstyle\sum_{k=1}^n\left(x_k^2 - \frac{2}{n}x_k\sum_{\ell=1}^n x_\ell + \frac{1}{n^2}\left(\sum_{\ell=1}^n x_\ell\right)^2\right) \\
&= \frac{1}{n-1}\left(\textstyle\sum_{k=1}^n x_k^2 - \frac{2}{n}\left(\sum_{k=1}^n x_k\right)\left(\sum_{\ell=1}^n x_\ell\right) + \frac{1}{n^2}\sum_{k=1}^n\left(\sum_{\ell=1}^n x_\ell\right)^2\right) \\
&= \frac{1}{n-1}\left(\textstyle\sum_{k=1}^n x_k^2 - \frac{2}{n}\left(\sum_{k=1}^n x_k\right)^2 + \frac{1}{n^2}n\left(\sum_{\ell=1}^n x_\ell\right)^2\right) \\
&= \frac{1}{n-1}\left(\textstyle\sum_{k=1}^n x_k^2 - \frac{1}{n}\left(\sum_{k=1}^n x_k\right)^2\right).
\end{aligned}
$$

Hence, in our case

$$s_n^2 = \frac{1}{24}\left(560 - \frac{1}{25}100^2\right) = \frac{20}{3} = 6.67.$$

Now, since ve are interested in testing a hypothesis on the variance of X, which is normally distributed, the statistic to be used is

$$\frac{(n-1)S_n^2}{\sigma_X^2}. \tag{3}$$

Consider testing the null hypothesis $H_0 : \sigma_X^2 = \sigma_0^2$, where $\sigma_0^2 \equiv 4$ against the alternative $H_1 : \sigma_X^2 > \sigma_0^2$ at the significance level $\alpha \equiv 0.05$. Under the assumption that the null hypothesis is true the statistic (3) with $\sigma_X^2 = \sigma_0^2$ has the chi-square distribution with $n-1$ degrees of freedom, that is

$$\frac{(n-1)S_n^2}{\sigma_0^2} \sim \chi_{n-1}^2.$$

1. By virtue of the above considerations, the upper tail rejection region is given by

$$R = \left\{\chi_{n-1}^2 > \chi_{n-1,\alpha}^2\right\}$$

where, the upper $\alpha = 0.05$ critical value of the chi-square distribution with $n-1 = 24$ degrees of freedom is given by

$$\chi_{24,0.05}^2 = 36.415.$$

Hence,

$$R = (36.415, +\infty)$$

Now, we have

$$\frac{(n-1)S_n^2(\omega)}{\sigma_X^2} = \frac{(n-1)s_n^2}{\sigma_0^2} = \frac{24}{4}\cdot\frac{20}{3} = 40.00 \in R.$$

Thus. the realization of the statistic belongs to the rejection region. This implies that H_0 is rejected.

In terms of p-value we have to compute

$$\mathbf{P}\left(\chi_{n-1}^2 \geq \frac{(n-1)s_n^2}{\sigma_0^2} \mid H_0 \text{ true}\right) = \mathbf{P}\left(\chi_{24}^2 \geq 40.00\right) = 1 - \mathbf{P}\left(\chi_{24}^2 \leq 40.00\right) = 0.0213 < 0.05.$$

The p-value method confirms the rejection of H_0.

With regard to the evaluation of $\beta(5)$, $\sigma_0^2 \equiv 4$, $\sigma_1^2 \equiv 5$, and $n = 25$, we have that

$$\begin{aligned}
\beta(5) &= \mathbf{P}(\text{II type error}) = \mathbf{P}(\text{accept } H_0 \mid H_0 \text{ false}) = \mathbf{P}\left(\text{accept } H_0 \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_0^2} \notin R \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_0^2}\frac{\sigma_1^2}{\sigma_1^2} \notin R \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_1^2}\frac{\sigma_1^2}{\sigma_0^2} \notin R \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(\chi_{n-1}^2\frac{\sigma_1^2}{\sigma_0^2} \notin R\right) \\
&= \mathbf{P}\left(\chi_{24}^2 \leq 36.415\frac{\sigma_0^2}{\sigma_1^2}\right) \\
&= \mathbf{P}\left(\chi_{24}^2 \leq 36.415\frac{4}{5}\right) \\
&= \mathbf{P}\left(\chi_{24}^2 \leq 29.132\right) \\
&= 0.7848.
\end{aligned}$$

2. In this case the rejection region is given by

$$R = \left\{\frac{(n-1)S_n^2}{\sigma^2} < \chi_{n-1,1-\alpha/2}^2\right\} \cup \left\{\frac{(n-1)S_n^2}{\sigma^2} > \chi_{n-1,\alpha/2}^2\right\}$$

where, lower $1-\alpha/2 = 0.975$ and the upper $\alpha/2 = 0.025$ critical values of the chi-square distribution with $n-1 = 24$ degrees of freedom are given by

$$\chi_{24,0.975}^2 = 12.40 \quad \text{and} \quad \chi_{24,0.025}^2 = 39.36.$$

In this case

$$\frac{(n-1)s_n^2}{\sigma^2} = \frac{24}{4}\cdot\frac{20}{3} = 40.00 \in R$$

belongs to the rejection region. Therefore, H_0 can be rejected against H_1. In terms of p-value we have to compute

$$2\mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma^2} \geq \frac{(n-1)s_n^2}{\sigma^2} \mid H_0 \text{ true}\right) = 2\mathbf{P}\left(\chi_{24}^2 \geq 40.00\right) = 2\left(1-\mathbf{P}\left(\chi_{24}^2 \leq 40.00\right)\right) = 0.0428 < 0.050.$$

Hence, the p-value method confirms that H_0 should not be rejected.

With regard to the evaluation of $\beta(5)$, setting $\sigma_0^2 \equiv 4$, $\sigma_1^2 \equiv 5$, and $n = 25$, we have that

$$\begin{aligned}
\beta(5) &= \mathbf{P}(\text{II type error}) = \mathbf{P}(\text{accept } H_0 \mid H_0 \text{ false}) = \mathbf{P}\left(\text{accept } H_0 \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(\frac{(n-1)S_n^2}{\sigma_0^2} \notin R \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(12.42 \leq \frac{(n-1)S_n^2}{\sigma_0^2}\frac{\sigma_1^2}{\sigma_1^2} \leq 39.36 \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(12.42\frac{\sigma_0^2}{\sigma_1^2} \leq \frac{(n-1)S_n^2}{\sigma_1^2} \leq 39.36\frac{\sigma_0^2}{\sigma_1^2} \mid \sigma = \sigma_1^2\right) \\
&= \mathbf{P}\left(12.42\frac{4}{5} \leq \frac{(n-1)S_n^2}{\sigma^2} \leq 39.36\frac{4}{5}\right) \\
&= \mathbf{P}\left(9.94 \leq \chi_{n-1}^2 \leq 31.49\right) \\
&= \mathbf{P}\left(\chi_{n-1}^2 \leq 31.49\right) - \mathbf{P}\left(\chi_{n-1}^2 \leq 9.94\right) \\
&= 0.860 - 0.005 \\
&= 0.855.
\end{aligned}$$

l'ho ricopiato per sbaglio, è quello già visto nelle slide di prima

NO

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica
Homework - 2019-12-06

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{n\geq 1}$ be a sequence of independent identically distributed Bernoulli random variables with success probability p. Set*

$$Z_n \stackrel{def}{=} \sum_{k=1}^n X_n, \qquad \bar{X}_n = \frac{1}{n}\sum_{n=1}^n X_k.$$

Assume that n is large. What you can say about the distributions, expectation, and variance of Z_n and $\bar{X}_n$? Consider the case $n = 100{,}000$ and $p = 1/2$. Use both the Central Limit Theorem and the Tchebychev inequality to estimate the probability that Z_n lies between $49{,}500$ and $50{,}500$. What you can say about the distributions, expectation, and variance of Z_n and $\bar{X}_n$ if $(X_n)_{n\geq 1}$ is a sequence of independent and Poisson distributed random variables with the same rate parameter λ?

Solution. Independently of n, the random variable Z_n has the binomial distribution with parameters n and p. In symbols

$$Z_n \sim Bin\,(n,p)\,.$$

As a consequence,

$$\mathbf{E}\,[Z_n] = np \quad \text{and} \quad \mathbf{D}^2\,[Z_n] = np\,(1-p)\,.$$

By virtue of the Central Limit Theorem, we know that

$$\frac{Z_n - np}{\sqrt{np\,(1-p)}} \xrightarrow{\mathbf{w}} N\,(0,1)\,.$$

Otherwise saying: as n is large, the distribution of $(Z_n - np)\,/\sqrt{np\,(1-p)}$ is approximately the standard normal distribution. On the other hand, we can write

$$\sqrt{\frac{p\,(1-p)}{n}}\frac{Z_n - np}{\sqrt{np\,(1-p)}} = \frac{Z_n}{n} - p = \bar{X}_n - p,$$

that is

$$\bar{X}_n = \sqrt{\frac{p\,(1-p)}{n}}\frac{Z_n - np}{\sqrt{np\,(1-p)}} + p$$

This implies that, as n is large, the distribution of $\bar{X}_n$ is approximately normal with

$$\mathbf{E}\left[\bar{X}_n\right] = p \quad \text{and} \quad \mathbf{D}^2\left[\bar{X}_n\right] = \frac{p\,(1-p)}{n}.$$

In the case $n = 100{,}000$ and $p = 1/2$, thanks to the Central Limit Theorem, we can write

$$\begin{aligned}
\mathbf{P}\,(49{,}500 \leq Z_n \leq 50{,}500) &= \mathbf{P}\,(49{,}500 - np \leq Z_n - np \leq 50{,}500 - np) \\
&= \mathbf{P}\left(\frac{49{,}500 - np}{\sqrt{np\,(1-p)}} \leq \frac{Z_n - np}{\sqrt{np\,(1-p)}} \leq \frac{50{,}500 - np}{\sqrt{np\,(1-p)}}\right) \\
&\simeq \Phi\left(\frac{50{,}500 - np}{\sqrt{np\,(1-p)}}\right) - \Phi\left(\frac{49{,}500 - np}{\sqrt{np\,(1-p)}}\right) \\
&= \Phi\left(\frac{500}{\sqrt{25{,}000}}\right) - \Phi\left(\frac{-500}{\sqrt{25{,}000}}\right) \\
&= 2\Phi\left(\sqrt{10}\right) - 1 \\
&= 2 \cdot 0.9992 - 1 \\
&= 0.9984.
\end{aligned}$$

where Φ is the distribution function of the standard normal. Instead, with the goal of applying the Tchebychev inequality, we can write

$$\begin{aligned}
\mathbf{P}\,(49{,}500 \leq Z_n \leq 50{,}500) &= \mathbf{P}\left(\frac{49{,}500 - np}{\sqrt{np\,(1-p)}} \leq \frac{Z_n - np}{\sqrt{np\,(1-p)}} \leq \frac{50{,}500 - np}{\sqrt{np\,(1-p)}}\right) \\
&= \mathbf{P}\left(-\sqrt{10} \leq \frac{Z_n - 50{,}000}{50\sqrt{10}} \leq \sqrt{10}\right) \\
&= \mathbf{P}\left(\left|\frac{Z_n - 50{,}000}{50\sqrt{10}}\right| \leq \sqrt{10}\right) \\
&= 1 - \mathbf{P}\left(\left|\frac{Z_n - 50{,}000}{50\sqrt{10}}\right| > \sqrt{10}\right) \\
&= 1 - \mathbf{P}\left(\left|\frac{Z_n - 50{,}000}{50\sqrt{10}}\right| \geq \sqrt{10}\right).
\end{aligned}$$

On the other hand, by the Tchebychev inequality we have

$$\mathbf{P}\left(\left|\frac{Z_n - 50{,}000}{50\sqrt{10}}\right| \geq \sqrt{10}\right) \leq \frac{\mathbf{D}^2\left[\frac{Z_n - 50{,}000}{50\sqrt{10}}\right]}{10} = \frac{1}{10}.$$

Therefore,

$$\mathbf{P}\,(49{,}500 \leq Z_n \leq 50{,}500) \geq 1 - \frac{1}{10} = \frac{9}{10} = 0.9.$$

This shows that the central limit aproach provides a sharper bound for the desired probability than the Tchebychev inequality approach.

Problem 2 *Suppose that a random variable X, which represents the reaction time at some stimulus, has a uniform distribution on an interval $[0, \theta]$, where the parameter $\theta > 0$ is unknown. An investigator wants to estimate θ on the basis of a simple random sample $X_1, \ldots, X_n$ of reaction times. Since θ is the largest possible time in the entire population of reaction times, the investigator consider as a first estimator for the parameter θ the largest sample reaction time. That is to say, the investigator consider as a first estimator the statistic*

$$\hat{\theta}_1 \equiv \check{X}_n \equiv \max\,(X_1, \ldots, X_n)\,.$$

NO

1. *Is $\check{X}_n$ unbiased? In case $\check{X}_n$ is not unbiased, is it possible to derive from $\check{X}_n$ an unbiased estimator of θ?*

2. *As a second estimator, the investigator consider the statistic*

$$\hat{\theta}_2 \equiv \bar{X}_n \equiv \frac{1}{n}\textstyle\sum_{k=1}^{n} X_k.$$

Is $\bar{X}_n$ unbiased? In case $\bar{X}_n$ is not unbiased, is it possible to derive from $\bar{X}_n$ an unbiased estimator of θ?

3. *In the investigator's shoes, what estimator would you prefer among those considered?*

Solution.

1. Writing $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ for the distribution function of the statistic $\check{X}_n$, we have

$$\begin{aligned} F_{\check{X}_n}(x) &= \mathbf{P}\left(\check{X}_n \leq x\right) = \mathbf{P}\left(X_1 \leq x, \ldots, X_n \leq x\right) = \textstyle\prod_{k=1}^{n} \mathbf{P}\left(X_k \leq x\right) \\ &= \textstyle\prod_{k=1}^{n} \mathbf{P}\left(X \leq x\right) = \mathbf{P}\left(X \leq x\right)^n = F_X(x)^n. \end{aligned}$$

for every $x \in \mathbb{R}$, where $F_X : \mathbb{R} \to \mathbb{R}$ is the distribution function of the random variable X. On the other hand, since X is uniformly distributed on $[0, \theta]$, it has density $f_X : \mathbb{R} \to \mathbb{R}$ given by

$$f_X(x) \stackrel{\text{def}}{=} \frac{1}{\theta} 1_{[0,\theta]}(x), \quad \forall x \in \mathbb{R}.$$

Hence,

$$\begin{aligned} F_X(x) &= \int_{(-\infty,x]} f_X(u)\ d\mu_L(u) = \int_{(-\infty,x]} \frac{1}{\theta} 1_{[0,\theta]}(u)\ d\mu_L(u) = \frac{1}{\theta} \int_{(-\infty,x]\cap[0,\theta]} d\mu_L(u) \\ &= \begin{cases} \frac{1}{\theta}\int_{\varnothing} d\mu_L(u) = 0 & \text{if } x < 0 \\ \frac{1}{\theta}\int_{[0,x]} d\mu_L(u) = \frac{x}{\theta} & \text{if } 0 \leq x \leq \theta \\ \frac{1}{\theta}\int_{[0,\theta]} d\mu_L(u) = 1 & \text{if } \theta < x \end{cases} \\ &= \frac{x}{\theta} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x) \end{aligned}$$

It then follows,

$$F_{\check{X}_n}(x) = F_X(x)^n = \frac{x^n}{\theta^n} 1_{[0,\theta]}(x) + 1_{(\theta,+\infty)}(x),$$

for every $x \in \mathbb{R}$. Now, we have

$$F'_{\check{X}_n}(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{nx^{n-1}}{\theta^n} & \text{if } 0 < x < \theta \\ 0 & \text{if } \theta < x \end{cases}.$$

Note that $F_{\check{X}_n}$ is not everywhere differentable. Eventually, is not differentiable in the point $x = \theta$. However, considering the function $f_{\check{X}_n} : \mathbb{R} \to \mathbb{R}$ given by

$$f_{\check{X}_n}(x) \stackrel{\text{def}}{=} \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x), \quad \forall x \in \mathbb{R},$$

a straightforward computation shows that

$$F_{\check{X}_n}(x) = \int_{(-\infty,x]} f_{\check{X}_n}(u)\ d\mu_L(u),$$

for every $x \in \mathbb{R}$. This implies that $\check{X}_n$ is absolutely continuous with density $f_{\check{X}_n}$. As a consequence,

$$\begin{aligned} \mathbf{E}\left[\check{X}_n\right] &= \int_{\mathbb{R}} x f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} x \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\ d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^n\ d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^{\theta} x^n\ dx = \frac{n}{\theta^n} \left.\frac{x^{n+1}}{n+1}\right|_0^{\theta} = \frac{n}{n+1}\theta. \end{aligned}$$

We can conclude that $\check{X}_n$ is not a unbiased estimator of θ but $\frac{n+1}{n}\check{X}_n$ is an unbiased estimator of θ.

2. We have

$$\begin{aligned} \mathbf{E}\left[\bar{X}_n\right] &= \mathbf{E}[X] = \int_{\mathbb{R}} x f_X(x)\ d\mu_L(x) = \int_{\mathbb{R}} \frac{x}{\theta} 1_{[0,\theta]}(x)\ d\mu_L(x) \\ &= \frac{1}{\theta} \int_{[0,\theta]} x\ d\mu_L(x) = \frac{1}{\theta} \int_0^{\theta} x\ dx = \frac{1}{\theta} \left.\frac{x^2}{2}\right|_0^{\theta} = \frac{\theta}{2}. \end{aligned}$$

Hence, $\bar{X}_n$ is not a unbiased estimator of θ but $2\bar{X}_n$ is an unbiased estimator of θ.

3. From 1. and 2. we know that

$$\mathbf{E}\left[\frac{n+1}{n}\check{X}_n\right] = \theta \qquad \text{and} \qquad \mathbf{E}\left[2\bar{X}_n\right] = \theta.$$

Hence, both $\frac{n+1}{n}\check{X}_n$ and $2\bar{X}_n$ are unbiased estimators of the parameter θ. To choose which is preferable between them, we consider

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] \qquad \text{and} \qquad \mathbf{D}^2\left[2\bar{X}_n\right].$$

We have

$$\begin{aligned} \mathbf{E}\left[\check{X}_n^2\right] &= \int_{\mathbb{R}} x^2 f_{\check{X}_n}(x)\ d\mu_L(x) = \int_{\mathbb{R}} x^2 \frac{nx^{n-1}}{\theta^n} 1_{(0,\theta)}(x)\ d\mu_L(x) = \frac{n}{\theta^n} \int_{(0,\theta)} x^{n+1}\ d\mu_L(x) \\ &= \frac{n}{\theta^n} \int_0^{\theta} x^{n+1}\ dx = \frac{n}{\theta^n} \left.\frac{x^{n+2}}{n+2}\right|_0^{\theta} = \frac{n}{n+2}\theta^2. \end{aligned}$$

Therefore,

$$\mathbf{D}^2\left[\check{X}_n\right] = \mathbf{E}\left[\check{X}_n^2\right] - \mathbf{E}\left[\check{X}_n\right]^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n}{(n+1)^2(n+2)}\theta^2.$$

As a consequence,

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \mathbf{D}^2\left[\check{X}_n\right] = \left(\frac{n+1}{n}\right)^2 \frac{n}{(n+1)^2(n+2)}\theta^2 = \frac{\theta^2}{n(n+2)}.$$

On the other hand,

$$\mathbf{D}^2\left[2\bar{X}_n\right] = 4\mathbf{D}^2\left[\bar{X}_n\right] = \frac{4}{n}\mathbf{D}^2[X] = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

Now, for any $n > 1$ we clearly have

$$\mathbf{D}^2\left[\frac{n+1}{n}\check{X}_n\right] < \mathbf{D}^2\left[2\bar{X}_n\right].$$

It follows that the estimator $\frac{n+1}{n}\check{X}_n$ is preferable to $2\bar{X}_n$.

Exercise 3 *Let X be a binomially distributed real random variable with number of trials parameter m and unknown success parameter p. An investigator wants to estimate p on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the method of moments. What is the estimator $\hat{p}_{M,n}$?*
2. *Assume the investigator applies the likelihood method. What is the estimator $\hat{p}_{ML,n}$?*

Solution. .

Exercise 4 *Let X be a normally distributed random variable with unknown mean μ and variance σ^2. An investigator wants to estimate μ and σ^2 on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the likelihood methods. What are the estimator $\hat{\mu}_{LM}$ and $\hat{\sigma}^2_{LM}$?*
2. *Assume the investigator applies the method of moments. What are the estimators $\hat{\mu}_{MM}$ and $\hat{\sigma}^2_{MM}$? Hint: guess what $\hat{\sigma}^2_{MM}$ could be and get it!*

Solution. .

Exercise 5 *Let X a random variable representing a characteristic of a certain population. Assume that X has a density $f_X : \mathbb{R} \to \mathbb{R}$ given by*

$$f_X(x) \stackrel{def}{=} \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x), \quad \forall x \in \mathbb{R},$$

where θ is a positive parameter.

1. *Apply the method of moments to find the estimator $\hat{\theta}_M$ of the parameter θ.*
2. *Apply the maximum likelihood method to find the estimator $\hat{\theta}_{ML}$ of the parameter θ.*
3. *Use the estimators $\hat{\theta}_M$ and $\hat{\theta}_{ML}$ to build estimators for $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*

Solution. .

Exercise 6 *Assume that the returns of a stock in a financial market are normally distributed with unknown mean μ and variance σ^2. Let X be the normal random variable representing the realization of the returns and let $X_1, \dots, X_n$ be a simple random sample of size n drawn from X. Assume that $n = 5$ and the realizations of the sample are*

$$x_1 \equiv -1.5, \quad x_2 \equiv -0.5, \quad x_3 \equiv 1.5, \quad x_4 \equiv 2.0, \quad x_5 \equiv 2.5$$

1. *Determine a 99% confidence interval for the mean μ.*
2. *Find the confidence for an interval of width 0.1.*

no

3. *Determine a 90% confidence interval for the standard deviation σ.*

Solution. .

Exercise 7 *Assume that a library master believes that the mean duration in days of the borrowing period is $20d$. However, the library master selects a simple random sample of 100 books in the library and discovers that the sample mean and variance of the borrowing days are $18d$ and $8d^2$, respectively. Determine a 99% confidence interval for the mean duration of the borrowing days to check whether library master's initial guess is correct.*

Solution. .

Exercise 8 *The mark of a infamous exam of Probability and Statistics are normally distributed with standard deviation $\sigma = 2$. A simple random sample of nine students is selected end the following evaluations are computed*

$$\sum_{k=1}^{9} x_k = 237 \quad \text{and} \quad \sum_{k=1}^{9} x_k^2 = 6295.$$

1. *Find a 90% confidence interval for the mean mark.*
2. *Discuss, without computation, whether the lenght of a 95% confidence interval would be smaller, greater or equal than the lenght of the interval previously determined.*
3. *How large the minimum sample size should be to obtain a 90% confidence interval for the mean mark with width equal to 3? Besides the confidence interval method is it possible to apply the Tchebychev inequality?*

Solution. .

Exercise 9 *Let $X_1, \dots, X_n, X_{n+1}$ be a simple random sample of size $n+1$ drawn from a Gaussian distributed random variable X with unknown mean μ and variance σ^2. Assume that we have observed $X_1, \dots, X_n$ and we want use the observed values $x_1, \dots, x_n$ to determine a confidence interval for the prediction of X_{n+1}. To this goal give detailed answers to the following questions:*

1. *what is the distribution of the statistic $\bar{X}_n$?*
2. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/\sigma\sqrt{1+1/n}$?*
3. *are the statistics $X_{n+1} - \bar{X}_n$ and $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)^2$ independent?*
4. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/S_n\sqrt{1+1/n}$?*
   After answering the above questions, build an interval in which the random variable X_{n+1} takes its values with probability α and determine the corresponding confidence interval for the prediction of X_{n+1}. In the end, assume that $n = 7$ and we have

   $$x_1 = 7005, \quad x_2 = 7432, \quad x_3 = 7420, \quad x_4 = 6822, \quad x_5 = 6752, \quad x_6 = 5333, \quad x_7 = 6552.$$

   compute the 95% confidence interval for the prediction of X_8.

DA QUI FINO A P.128 TUTTA MONNEZZA

Exercise 10 *Let X be a Gaussian random variable with unknown mean μ_X and variance σ_X^2 representing a certain characteristic of a population. Assume that testing the sample mean $\bar{X}_n$ and the sample standard deviation S_n of a simple random sample $X_1, \dots, X_n$ of size $n \equiv 9$ drawn from X we obtain the value $\bar{X}_n(\omega) \equiv \bar{x}_n = 251.50cm$ and $S_n(\omega) \equiv s_n = 2.30cm$.*

1. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \mu_X = 250cm$ be rejected against the alternative $H_a : \mu_X \neq 250cm$ at the significance level $\alpha = 0.1$?*
2. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \sigma_X^2 = 4$ be rejected against of the alternative $H_a : \sigma_X^2 > 4$ at the significance level $\alpha = 0.05$? Calculate the probability $\beta(5)$ of a II type error.*

Solution. .

Exercise 11 *Let X be a Gaussian random variable with unknown mean μ and variance σ^2 representing a certain characteristic of a population and let $X_1, \dots, X_n$ be a simple random sample of size n drawn from X. Assume that $n = 25$ and that the realizations $x_1, \dots, x_{25}$ of the sample give an information summarized by*

$$\sum_{k=1}^{25} x_k = 100 \quad \text{and} \quad \sum_{k=1}^{25} x_k^2 = 560$$

1. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \sigma^2 = 4$ be rejected against of the alternative $H_1 : \sigma^2 > 4$ with a significance level $\alpha = 0.05$? Calculate the probability $\beta(5)$ of a II type error.*
2. *Considering both the rejection region method and the p-value method, should the null hypothesis $H_0 : \sigma^2 = 4$ be rejected against of the alternative $H_1 : \sigma^2 \neq 4$ with a significance level $\alpha = 0.05$? Calculate the probability $\beta(5)$ of a II type error.*

Solution. .

Exercise 12 *In order to measure the depencence between two random variables X and Y a simple sample of size 10 is drawn from the random vector (X, Y) and the following quantities are computed*

$$\bar{x}_{10} = 49.6670, \quad \bar{y}_n = -0.4333, \quad s_x^2 = 236.1390, \quad s_y^2 = 0,1750, \quad \gamma_{x,y} = 5.8072.$$

1. *May you find the equation of the regression live of Y against X?*
2. *May you find the estimated mean square error?*
3. *What value of y would you predict for a corresponding value of x?*

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Final Test - 2020-02-12 - Probability

Problem 1 *The scrutiny of group of 100,000 randomly chosen male people in the age 40 − 79 in UK during 2013 − 2015 reveals the following table of average lung cancer incidence*

	smoker	*not smoker*	*total*
lung cancer	10,395	7,407	17,802
not lung cancer	50,078	32,120	82,198
	60,473	39,527	100,000

Write Ω for the sample space consisting of these 100,000 people and write S [resp. C] for the subsets of Ω consisting of the smokers [the people affected by lung canger]. Let $1_S : \Omega \to \{0,1\}$ and $1_C : \Omega \to \{0,1\}$ the indicator functions of the events S and C respectively.

1. *Determine the joint distribution of the random vector $(1_S, 1_C)$ and the marginal distributions of the random variables 1_S and 1_C.*
2. *Are the random variables 1_S and 1_C independent?*
3. *What is the probability that a randomly chosen person in Ω is affected by lung cancer, given that he is a smoker?*
4. *What is the probability that a randomly chosen person in Ω is a smoker, given that he is affected by lung cancer?*
5. *Check the validity of the total probability formula for $\mathbf{P}(S)$ and the Bayes Formula for $\mathbf{P}(C \mid S)$.*

Solution. .

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra, and let X be a uniformly distributed random variable on Ω with states in the interval $(-1,1)$. In symbols $X \sim Unif(-1,1)$. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by*

$$g(x) \stackrel{def}{=} x^2, \quad \forall x \in \mathbb{R}.$$

1. *Can you state that the function $Y : \Omega \to \mathbb{R}$ given by*
   $$Y(\omega) \stackrel{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega.$$
   is a real random variable on Ω?
2. *Can you compute the distribution function $F_Y : \mathbb{R} \to \mathbb{R}_+$ of $Y : \Omega \to \mathbb{R}$?*
3. *Is Y absolutely continuous?*
4. *Are the first and second order moments of Y finite?*
5. *If the first and second order moments of Y are finite, can you compute $\mathbf{E}[Y]$ and $\mathbf{D}^2[Y]$?*

Solution. .

Problem 3 *Let U, V real random variables on a probability space Ω such that such that $U \sim V \sim N(0,1)$, the vector (U,V) is Gaussian, and $Corr(U,V) \equiv \rho < 1$. Consider the real random variables*

$$X \stackrel{def}{=} U - \rho V \quad \text{and} \quad Y \stackrel{def}{=} \sqrt{1-\rho} V.$$

1. *Can you prove that the vector (X,Y) Gaussian?*
2. *Are the random variables X and Y independent?*
3. *Compute the distributions of X and Y;*
4. *Compute $\mathbf{E}\left[X^2Y^2\right]$, $\mathbf{E}\left[XY^3\right]$, $\mathbf{E}\left[Y^4\right]$.*
5. *Compute $\mathbf{E}\left[U^2V^2\right]$.*

Solution. .

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let B_1 and B_2 be standard Bernoulli random variables on Ω. In symbols, $B_k \sim Ber(1/2)$, for $k = 1, 2$. Assume that B_1 and B_2 are independent and set*

$$X \stackrel{def}{=} B_1 + B_2, \quad Y \stackrel{def}{=} B_1 \cdot B_2$$

1. *Compute $\mathbf{E}[B_k \mid X]$ and $\mathbf{E}[B_k \mid Y]$ for $k = 1, 2$.*
2. *Are the random variables $\mathbf{E}[B_1 \mid X]$ and $\mathbf{E}[B_2 \mid X]$ uncorrelated? Are they independent?*
3. *Are the random variables $\mathbf{E}[B_1 \mid Y]$ and $\mathbf{E}[B_2 \mid Y]$ uncorrelated? Are they independent?*
4. *Compute $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$.*
5. *Are the random variables $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$ uncorrelated? Are they independent?*
6. *Compute $\mathbf{E}\left[X^2 \mid Y\right]$ and $\mathbf{E}\left[Y^2 \mid X\right]$.*

Solution. .

Problem 5 *Let $\theta > 0$ and let X be a real random variable which is uniformly distributed in the interval $[0, \theta]$, in symbols $X \sim U(0, \theta)$. Let $X_1, \dots, X_n$ a simple random sample of size $n \geq 1$ drawn from X and let $\left(\check{X}_n\right)_{n\geq 1}$ be the sequence of real random variables given by*

$$\check{X}_n \stackrel{def}{=} \max\{X_1, \dots, X_n\}, \quad \forall n \geq 1.$$

1. *Prove that the sequence $\left(\check{X}_n\right)_{n\geq 1}$ converges in distribution to the Dirac random variable concentrated in θ.*
2. *Prove directly that the sequence $\left(\check{X}_n\right)_{n\geq 1}$ converges in probability to $Y \sim Dir(\theta)$.*
3. *Prove directly that the sequence $\left(\check{X}_n\right)_{n\geq 1}$ converges in mean to $Y \sim Dir(\theta)$.*
4. *Prove that the sequence $\left(\check{X}_n\right)_{n\geq 1}$ converges in square mean to $Y \sim Dir(\theta)$.*
5. *Does $\left(\check{X}_n\right)_{n\geq 1}$ converge almost surely to $Y \sim Dir(\theta)$?*

   Hint: Determine the distribution function of $\check{X}_n$ and exploit it.

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Final Test - 2020-02-12 - Statistics

Problem 1 *Let X and Y two real random variables each of which represents a certain trait of a population. Assume that X and Y are independent and such that*

$$X \sim N\left(\mu_X, \sigma_X^2\right), \qquad Y \sim N\left(\mu_Y, \sigma_Y^2\right),$$

where $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ are known. Consider a simple random sample $X_1, \dots, X_m$ [resp. $Y_1, \dots, Y_n$] of size m [resp. n] drawn from X [resp. Y].

1. *Compute $\mathbf{P}\left(x \leq S_n^2(Y) \leq y\right)$ in terms of given $x, y \in \mathbb{R}_+$.*
2. *Determine x such that $\mathbf{P}\left(\bar{X}_m > x\right) = 0.25$.*
3. *Compute $\mathbf{P}\left(\bar{X}_m - x > \bar{Y}_n + y\right)$ in terms of given $x, y \in \mathbb{R}_+$.*

Solution. .

Problem 2 *Let $\theta > 0$ and let X be an absolutely continuous real random variable with density function $f_X : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_X(x) \stackrel{def}{=} \frac{1}{2} e^{-|x-\theta|}, \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to determine the estimator $\hat{\theta}_n^M$ for θ.*
2. *Check whether $\hat{\theta}_n^M$ is unbiased, consistent in probability, and consistent in square mean.*
3. *Can you "guess" the result of the method of maximum likelihood to determine the estimator $\hat{\theta}_n^{ML}$ for θ?*

   Hint: take for granted that the random variable X has finite moment of the first order; recall that an estimator $\hat{\theta}_n$ for the true value of a parameter θ is said to be consistent in probability [resp. in square mean] if

   $$\hat{\theta}_n \xrightarrow{\mathbf{P}} \theta \quad [\text{resp. } \hat{\theta}_n \xrightarrow{\mathbf{L}^2} \theta],$$

   as $n \to \infty$.

Solution. .

Problem 3 *Let X be a standard Bernoulli random variable with unknown success parameter p. Let $X_1, \dots, X_n$ be a simple random sample of size n drawn from X and let $Z_n \equiv \sum_{k=1}^n X_k$ be the sample sum. It is well known that $Z_n \sim Bin(n,p)$. In addition, when n is large ($np \geq 10$ and $n(1-p) \geq 10$) the sample sum has approximately a normal distribution.*

1. *Determine a confidence interval for the parameter p with confidence level approximately $100(1-\alpha)\%$.*

2. *Determine the size n of the sample $X_1, \dots, X_n$ which allows a confidence interval for the parameter p with confidence level approximately $100(1-\alpha)\%$ and width w, where both α and w are given in advance.*

Solution. .

Problem 4 *Let $X_1, \dots, X_n, X_{n+1}$ be a simple random sample of size $n+1$ drawn from a Gaussian distributed random variable X with unknown mean μ and variance σ^2. Assume that we have observed $X_1, \dots, X_n$ and we want use the observed values $x_1, \dots, x_n$ to determine a confidence interval for the prediction of X_{n+1}. To this goal give detailed answers to the following questions:*

1. *what is the distribution of the statistic $\bar{X}_n$?*
2. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/\sigma\sqrt{1+1/n}$?*
3. *what is the distribution of the statistic $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)^2$?*
4. *are the statistics $X_{n+1} - \bar{X}_n$ and $S_n^2 \equiv \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)^2$ independent? Why?*
5. *what is the distribution of the statistic $\left(X_{n+1} - \bar{X}_n\right)/S_n\sqrt{1+1/n}$?*
6. *After answering the above questions, build an interval in which the random variable X_{n+1} takes its values with probability α and determine the corresponding confidence interval for the prediction of X_{n+1}. In the end, assume that $n = 7$ and we have*

$$x_1 = 7005, \quad x_2 = 7432, \quad x_3 = 7420, \quad x_4 = 6822, \quad x_5 = 6752, \quad x_6 = 5333, \quad x_7 = 6552.$$

compute the 95% confidence interval for the prediction of X_8.

Solution. .

Problem 5 *Let X [resp. Y] be a Gaussian distributed random variables with (unknown) mean $\mu_X \in \mathbb{R}$ [resp. $\mu_Y \in \mathbb{R}$] and variance $\sigma_X^2 > 0$ [resp. $\sigma_Y^2 > 0$]. Assume that X describes a trait of some population before a treatment and Y describes the same trait of the same population after a treatment (for instance a power training period). Let $X_1, \dots, X_n$ be a simple random sample drawn by X and let $Y_1, \dots, Y_n$ be the corresponding sample drawn from Y. Note that we can still assume that $Y_1, \dots, Y_n$ is a simple random sample but we cannot assume that the samples $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ are independent. Actually, there is no reason at all to think that the random variables X and Y are independent. However, it is still reasonable to assume that the random variable $D \equiv Y - X$ is Gaussian distributed and that $D_1 \equiv Y_1 - X_1, \dots, D_n \equiv Y_n - X_n$ is a simple random sample drawn from D. Assume to have measured*

$$\begin{array}{ccccccc} x_1 = 73.80 & x_2 = 62.80 & x_3 = 73.40 & x_4 = 63.50 & x_5 = 71.90 & x_6 = 74.30 & x_7 = 63.10 \\ y_1 = 75.70 & y_2 = 63.70 & y_3 = 74.70 & y_4 = 64.40 & y_5 = 70.50 & y_6 = 74.90 & y_7 = 64.90 \end{array}.$$

1. *Should we reject the null hypothesis $H_0 = \mu_Y = \mu_X$ against the alternatives $H_1 = \mu_Y \neq \mu_X$ and $H_1 = \mu_Y > \mu_X$ at the significance level $\alpha = 0.05$? Consider both the rejection region method and the p-value method.*
2. *Assume that $\sigma_X = 5.51$ and that $\rho_{X,Y} = 0.98$. Should we reject the null hypothesis $H_0 = \sigma_Y^2 = \sigma_X^2$ against the alternatives $H_1 = \sigma_Y^2 \neq \sigma_X^2$ and $H_1 = \sigma_Y^2 < \sigma_X^2$ at the significance level $\alpha = 0.05$?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Final Test - 2020-02-25 - Probability

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space, let $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \equiv \mathbb{R}$ be the Euclidean real line endowed with the Borel σ-algebra, and let X be a real random variable on Ω. Consider the functions $g : \mathbb{R}_+ \to \mathbb{R}$ and $h : \mathbb{R}_{++} \to \mathbb{R}$ given by*

$$g(x) \overset{def}{=} \sqrt{x}, \quad \forall x \in \mathbb{R}_+ \qquad \text{and} \qquad h(x) \overset{def}{=} \ln(x), \quad \forall x \in \mathbb{R}_{++}.$$

1. *Can you always state that the functions $Y : \Omega \to \mathbb{R}$ given by*

   $$Y(\omega) \overset{def}{=} g(X(\omega)), \quad \forall \omega \in \Omega \qquad \text{and} \qquad Z(\omega) \overset{def}{=} h(X(\omega)) \quad \forall \omega \in \Omega$$

   are real random variables on Ω?
2. *Considering the answer you gave to the above question, can you compute the distribution function, $F_Y : \mathbb{R} \to \mathbb{R}_+$ of $Y : \Omega \to \mathbb{R}$ and $F_Z : \mathbb{R} \to \mathbb{R}_+$ of $Z : \Omega \to \mathbb{R}$?*
3. *Can you show that fixed any $\lambda \in (0,1)$, the function $\lambda F_Y + (1-\lambda) F_Z$ is a distribution function?*
4. *What about the functions F_Y^2, F_X^2, and $F_Y F_Z$?*

Solution. .

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $\left(\mathbb{R}^2, \mathcal{B}\left(\mathbb{R}^2\right), \mu_L^2\right) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra $\mathcal{B}\left(\mathbb{R}^2\right)$ and the Lebesgue measure $\mu_L^2 : \mathcal{B}\left(\mathbb{R}^2\right) \to \mathbb{R}_+$. Let*

$$\mathbb{R}_+^2(x > y) \equiv \left\{(x,y) \in \mathbb{R}_+^2 : x > y\right\}, \qquad \mathbb{R}_+^2(x \leq y) \equiv \left\{(x,y) \in \mathbb{R}_+^2 : x \leq y\right\},$$

and let $F : \mathbb{R}^2 \to \mathbb{R}_+$ given by

$$F(x,y) \overset{def}{=} \left(1 - e^{-y} - \frac{1}{2} y e^{-x}\right) 1_{\mathbb{R}_+^2(x>y)}(x,y) + \left(1 - e^{-x} - \frac{1}{2} x e^{-y}\right) 1_{\mathbb{R}_+^2(x \leq y)}(x,y), \quad \forall (x,y) \in \mathbb{R}^2.$$

1. *Can you show that the function $F : \mathbb{R}^2 \to \mathbb{R}_+$ is a distribution function? Hint: consider carefully the sets $\mathbb{R}_+^2(x > y)$ and $\mathbb{R}_+^2(x \leq y)$ (draw a graph).*

   Let $Z \equiv (X, Y)$ be the random vector on Ω with distribution function $F : \mathbb{R}^2 \to \mathbb{R}_+$.
2. *Can you determine the marginal distribution of the entries X and Y?*
3. *Is the random vector Z absolutely continuous? Can you determine a density $f_Z : \mathbb{R}^2 \to \mathbb{R}_+$ for Z?*
4. *If Z is absolutely continuous, can you determine the marginal densities of the entries X and Y? Hint: it may be useful to rewrite the indicator fnctions $1_{\mathbb{R}_+^2(x>y)}(x,y)$ and $1_{\mathbb{R}_+^2(x \leq y)}(x,y)$ in terms of product of other indicator functions.*

Solution. .

Problem 3 *Let Z_1, Z_2, Z_3 independent random variables on a probability space Ω such that such that $X_k \sim N(0,1)$, fo $k = 1,2,3$. Consider the real random variables*

$$X_1 \stackrel{def}{=} Z_1 + Z_2 + Z_3, \quad X_2 \stackrel{def}{=} Z_1 - Z_2 + Z_3, \quad X_3 \stackrel{def}{=} Z_1 - Z_3.$$

1. *What is the distribution of the vector $X \equiv (X_1, X_2, X_3)^{\intercal}$?*
2. *Can you compute the distribution function of X?*
3. *Among the pairs (X_1, X_2), (X_1, X_3), and (X_2, X_3) of entries of X what are made by independent random variables?*
4. *Compute the distributions of X_1, X_2, and X_3;*
5. *Think on a quick and smart way to compute $\mathbf{E}\left[X_1 X_2^2\right]$, $\mathbf{E}\left[X_1^2 X_2^2\right]$, $\mathbf{E}\left[X_2 X_3^2\right]$, $\mathbf{E}\left[X_2^2 X_3^2\right]$.*

Solution. .

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let R_1 and R_2 be standard Rademacher random variables on Ω. In symbols, $R_k \sim Rad(1/2)$, for $k = 1,2$. Assume that R_1 and R_2 are independent and set*

$$X \stackrel{def}{=} R_1 - R_2, \quad Y \stackrel{def}{=} -R_1 \cdot R_2$$

1. *Compute $\mathbf{E}[R_k \mid X]$ and $\mathbf{E}[R_k \mid Y]$ for $k = 1,2$.*
2. *Are the random variables $\mathbf{E}[R_1 \mid X]$ and $\mathbf{E}[R_2 \mid X]$ uncorrelated? Are they independent?*
3. *Are the random variables $\mathbf{E}[R_1 \mid Y]$ and $\mathbf{E}[R_2 \mid Y]$ uncorrelated? Are they independent?*
4. *Compute $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$.*
5. *Are the random variables $\mathbf{E}[X \mid Y]$ and $\mathbf{E}[Y \mid X]$ uncorrelated? Are they independent?*
6. *Compute $\mathbf{E}\left[X^2 \mid Y\right]$ and $\mathbf{E}\left[Y^2 \mid X\right]$.*

Solution. .

Problem 5 *Consider the probability space $([0,1], \mathcal{B}([0,1]), \mu_L) \equiv \Omega$, where $\mathcal{B}([0,1])$ is the Borel σ-algebra on the interval $[0,1] \subseteq \mathbb{R}$ and $\mu_L : \mathcal{B}([0,1]) \to \mathbb{R}_+$ is the Borel-Lebesgue measure on $[0,1]$. Consider the sequence $(X_n)_{n\geq 1}$ given by*

$$X_n(\omega) = \begin{cases} 1 & \text{if } 0 \leq \omega \leq \frac{n+1}{2n} \\ 0 & \text{otherwise} \end{cases}.$$

1. *Can you show that $(X_n)_{n\geq 1}$ is a sequence of random variables on Ω?*
2. *Can you prove that the sequence $(X_n)_{n\geq 1}$ converges in distribution to a random variable X on Ω?*
3. *Can you prove that the sequence $(X_n)_{n\geq 1}$ converges in probability to X.*
4. *Does the sequence $(X_n)_{n\geq 1}$ converges in mean to X.*
5. *Does the sequence $(X_n)_{n\geq 1}$ converges in square mean to X?*
6. *Can you prove that the sequence $(X_n)_{n\geq 1}$ converges almost surely to X?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Final Test - 2020-02-25 - Statistics

Problem 1 *Let X be a geometrically distributed real random variable with unknown success probability $p \in (0,1)$ representing the occurrence of the first success at the nth trial, where $n \in \mathbb{N}$, in binary random experiment. An investigator wants to estimate p on the basis of a simple random sample $X_1, \dots, X_n$ of size n drawn from X.*

1. *Assume the investigator applies the method of moments. What is the estimator $\hat{p}_n^M$? Hint: it may be useful to recall that*

$$\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}$$

and that, since $q \in (0,1)$, the formula

$$\frac{d}{dq}\sum_{n=1}^{\infty} q^n = \sum_{n=1}^{\infty} \frac{d}{dq} q^n$$

holds true.

2. *Is $\hat{p}_n^M$ unbiased? Is $\hat{p}_n^M$ consistent in probability?*
3. *Assume the investigator applies the likelihood methods. What is the estimator $\hat{p}_n^{LM}$? Hint: pay attention to writing the density function of X.*
4. *Given a realization $x_1, \dots, x_{50}$ of a sample $X_1, \dots, X_{50}$ of size 50 drawn from X such that $\sum_{k=1}^{50} x_k = 150$ apply the formulas obtained to get the realizations $\hat{p}_n^M(\omega)$ and $\hat{p}_n^{LM}(\omega)$ of the estimators $\hat{p}_n^M$ and $\hat{p}_n^{LM}$.*
5. *Give an estimate of the mean and the variance of X.*

Solution. .

Problem 2 *Let X a random variable representing a trait of a population. Assume that X has a density $f_X : \mathbb{R} \to \mathbb{R}$ given by*

$$f_X(x) \stackrel{def}{=} \frac{1}{\theta} e^{-\frac{x-3}{\theta}} 1_{[3,+\infty)}(x), \quad \forall x \in \mathbb{R}.$$

1. *Apply the method of moments to find the estimator $\hat{\theta}_{MM}$ of the parameter θ.*
2. *Apply the maximum likelihood method to find the estimator $\hat{\theta}_{ML}$ of the parameter θ.*
3. *Check whether $\hat{\theta}_n^M$ is unbiased and consistent in probability. Can you also check whether $\hat{\theta}_n^M$ is consistent in square mean?*
4. *Use the estimators $\hat{\theta}_{MM}$ and $\hat{\theta}_{ML}$ to build estimators for $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*

Solution. .

Problem 3 *Assume that a library master believes that the mean duration in days of the borrowing period is* 20d. *However, the library master selects a simple random sample of* 100 *books in the library and discovers that the sample mean and standard deviation of the borrowing days are* 18d *and* 4d, *respectively.*

1. *Determine a* 95% *confidence interval for the mean duration of the borrowing days to check if library master's initial guess is correct.*

2. *Test the null hypothesis* $H_0 : \mu = 20$ *against the alternative* $H_1 : \mu > 20$ *at the significance level* $\alpha = 0.05$.

3. *Can you also compute the probability* $\beta(23)$ *of a II type error?*

Solution. .

Problem 4 *A random sample of* 25 *chickens from a hen-house are selected and two diets are given to them to increase their weight. A sub-sample of* 13 *chickens are fed with diet* X *and the remaining* 12 *chickens are fed with diet* Y. *After one month the chickens of the two samples are weighted and the weights in kg are the following*

Diet X	1.60,	1.70,	1.55,	1.70,	1.75,	1.35,	1.65,	1.75,	1.70,	1.65,	1.56,	1.60,	1.80.
Diet Y	1.70	1.35	1.50	1.65	1.50	1.60	1.70	1.75	1.60,	1.65,	1.55,	1.45.	

The weight of a chicken is assumed to be modeled by a normal random variable.

1. *Find the mean* μ_X *and* μ_Y *and the standard deviation* σ_X *and* σ_Y *for the weight of each sample.*

2. *Find an estimate of the pooled variance* s_p^2 *which estimates the variance of the whole sample.*

3. *Find a confidence interval for* $\mu_X - \mu_Y$ *at the* 95% *confidence level.*

4. *Formulate a null and an alternative hypothesis from the farmer's point of view who aims to test whether diet* X *is better than* Y. *Test such a null hypothesis against the alternative at the* 5% *significance level by using the rejection method and the p-value method.*

Solution. .

Problem 5 *Let* X *[resp.* Y*] be a Gaussian distributed random variables with (unknown) mean* $\mu_X \in \mathbb{R}$ *[resp.* $\mu_Y \in \mathbb{R}$*] and variance* $\sigma_X^2 > 0$ *[resp.* $\sigma_Y^2 > 0$*]. Assume that* X *and* Y *describe the same trait of different populations, for instance the cholesterol concentration in the blood of* 25+ *years aged males in Iceland and Democratic Republic of Congo in* 2008. *Then it is rather natural to assume that* X *and* Y *are idependent. Let* $X_1, \ldots, X_m$ *[resp.* $Y_1, \ldots, Y_n$*] be a simple random sample drawn by* X *[resp. by* Y*]. Assume to have measured*

$$\begin{array}{lllllllll} x_1 = 5.80 & x_2 = 5.70 & x_3 = 5.30 & x_4 = 5.60 & x_5 = 5.90 & x_6 = 6.20 & x_7 = 6.10 & & \\ y_1 = 3.20 & y_2 = 3.40 & y_3 = 4.70 & y_4 = 3.30 & y_5 = 4.50 & y_6 = 3.60 & y_7 = 4.30 & y_8 = 4.20 & y_9 = 4.10 \end{array}.$$

1. *Should we reject the null hypothesis* $H_0 = \mu_Y = \mu_X$ *against the alternatives* $H_1 = \mu_Y \neq \mu_X$ *and* $H_1 = \mu_Y > \mu_X$ *at the significance level* $\alpha = 0.05$*? Consider both the rejection region method and the p-value method.*

2. *Under the assumption that two random variables* U *and* V *are independent and such that* $U \sim \chi_m^2$ *and* $V \sim \chi_n^2$, *we know that the random variable* $(U/m)/(V/n)$ *has the Fisher-Snedecor distribution with* m *numerator degrees of freedom and* n *denominator degrees of freedom. In symbols,*

$$U \sim \chi_m^2,\ V \sim \chi_n^2,\ U \perp\!\!\!\perp V \Rightarrow \frac{U/m}{Y/n} \sim F(m,n).$$

We also know that the density of any Fisher-Snedecor distribution is concentrated on the positive real axis and is not symmetric (the graph of the density is similar to the graph of a χ^2 *density). For some standard values of* α *(for instance* $\alpha = 0.1, \alpha = 0.05, \alpha = 0.25, \alpha = 0.01, \ldots$*) the lower [resp. upper] critical value* $f_{m,n,\alpha}^-$ *[resp.* $f_{m,n,\alpha}^+$*] of* $F(m,n)$ *can be found in statistical tables or computed by almost all statistical softwares, for several values of* m *and* n. *Therefore, we assume to know* $f_{m,n,\alpha}^-$ *and* $f_{m,n,\alpha}^+$. *Now, in light of what we know, can you introduce a statistic to test the null hypothesis* $H_0 : \sigma_X^2 = \sigma_Y^2$ *against the alternative* $H_1 : \sigma_X^2 > \sigma_Y^2$ *and* $H_1 : \sigma_X^2 < \sigma_Y^2$ *and describe for this statistic both the rejection method and the p-value method?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Intermediate Test - 2019-11-22

Problem 1 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let E, F, G events of Ω such that*

1. *E and G are independent;*
2. *F and G are independent;*
3. *E and F are exclusive;*
4. *$\mathbf{P}(E \cup G) \equiv a$, $\mathbf{P}(F^c \cap G^c) \equiv b$, $\mathbf{P}(E \cup F \cup G) \equiv c$, such that $a - b - c \neq 0$.*

Determine $\mathbf{P}(E)$, $\mathbf{P}(F)$, and $\mathbf{P}(G)$ in terms of a, b, c. Hint: think on a "smart" way to solve the system of equation yielding $\mathbf{P}(E)$, $\mathbf{P}(F)$, and $\mathbf{P}(G)$.

Solution. . □

Problem 2 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_L^2) \equiv \mathbb{R}^2$ be the Euclidean real plane endowed with the Borel σ-algebra, $\mathcal{B}(\mathbb{R}^2)$, and the Lebesgue measure, $\mu_L^2 : \mathcal{B}(\mathbb{R}^2) \to \mathbb{R}_+$. Let $f : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f(x, y) \stackrel{def}{=} kxe^{-(x+y)} 1_{\mathbb{R}_+^2}(x, y), \quad \forall (x, y) \in \mathbb{R}^2$$

where $\mathbb{R}_+^2 \equiv \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$. Determine $k \in \mathbb{R}$ such that $f : \mathbb{R}^2 \to \mathbb{R}_+$ is a probability density. Let $Z \equiv (X, Y)$ be the random vector of density $f : \mathbb{R}^2 \to \mathbb{R}_+$.

1. *Determine the distribution function $F_Z : \mathbb{R}^2 \to \mathbb{R}_+$ of the vector Z and check that*
$$\frac{\partial F^2}{\partial x \partial y}(x, y) = f(x, y), \quad \mu_L^2\text{-a.e. on } \mathbb{R}^2.$$
2. *Determine the marginal distribution function $F_X : \mathbb{R} \to \mathbb{R}_+$ and $F_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z.*
3. *Determine the densities $f_X : \mathbb{R} \to \mathbb{R}_+$ and $f_Y : \mathbb{R} \to \mathbb{R}_+$ of the entries X and Y of Z and check that*
$$\frac{dF_X}{dx}(x) = f_X(x) \quad \text{and} \quad \frac{dF_Y}{dy}(y) = f_Y(y), \quad \mu_L\text{-a.e. on } \mathbb{R}.$$
4. *Are X and Y independent random varables?*
5. *Compute $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$ and $Cov(X, Y)$.*
6. *Compute $\mathbf{E}[(X, Y)]$ and the covariance matrix of the vector (X, Y).*

Solution. . □

Problem 3 *Let X [resp. B] be a standard Gaussian [Bernoulli] random variable on a probability space Ω. In symbols, $X \sim N(0, 1)$ and $B \sim Ber(1/2)$. Assume that X and B are independent and define $Y \equiv B \cdot X$.* ***Specifying carefully the properties used****, answer the following questions:*

1. *Is the random variable Y Gaussian? Is Y absolutely continuous?*
2. *Are the random variables X and Y uncorrelated? Are X and Y independent?*
3. *Are the random variables B and Y uncorrelated? Are B and Y independent?*
4. *Does the random vector $(X, Y)^\intercal$ have a bivariate Gaussian distribution?*
5. *Can you compute $\mathbf{E}[Y \mid X]$? What about $\mathbf{E}[X \mid Y]$?*

Solution. .

Problem 4 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ a probability space and let $B \sim Ber(1/2)$ [resp. $R \sim Rad(1/2)$] a standard Bernoulli [resp. Rademacher] random variable on Ω. Assume that B and R are independent and define $X \stackrel{def}{=} B + R$.*

1. *Compute $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$, $\mathbf{E}[B \mid X]$, and $\mathbf{E}[R \mid X]$. In addition,* ***specifying carefully the properties used****, answer the following questions:*
2. *Are the random variables $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$ uncorrelated? Are $\mathbf{E}[X \mid B]$, $\mathbf{E}[X \mid R]$ independent?*
3. *Are the random variables $\mathbf{E}[B \mid X]$ and $\mathbf{E}[R \mid X]$ uncorrelated? Are $\mathbf{E}[B \mid X]$ and $\mathbf{E}[R \mid X]$ independent?*
4. *By using the properties of the conditional expectation, on account that you are dealing with a Bernoulli and a Rademacher random variable, can you compute $\mathbf{E}[BR \mid X]$?*

Solution. .

Problem 5 *Let $X \sim Exp(1)$ an exponential random variable of rate parameter $\lambda = 1$ and let $(Y_n)_{n \geq 1}$ be the sequence of independent real random variables such that*

$$Y_n \stackrel{def}{=} \begin{cases} n & \text{if } 0 \leq X < \frac{1}{n}, \\ 0 & \text{if } 1/n \leq X. \end{cases}, \quad \forall n \geq 1$$

1. *Does $(Y_n)_{n \geq 1}$ converges in distribution?*
2. *Does $(Y_n)_{n \geq 1}$ converges in probability?*
3. *Does $(Y_n)_{n \geq 1}$ converges almost surely?*
4. *Does $(Y_n)_{n \geq 1}$ converges in mean?*
5. *Does $(Y_n)_{n \geq 1}$ converges in quadratic mean?*

Solution. .

II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Intermediate Test - 2020-01-27 - Probability

Problem 1 *Students are allowed to take a test twice in an examination session. Assume that* 7 *students over* 10 *pass the test on the first try. For those who fail, only* 4 *students over* 10 *pass the test on the second try.*

1. *Find the probability that a randomly selected student passes the test.*
2. *Assuming that a student passed the test what is the probability she passed on the first try?*
3. *The first test presents the following problem: two dice are rolled and the number on the upper faces are observed. Is the event "the sum of the observed numbers is* 8*" independent of the event "the number observed on the upper face of a die is* 4*"? Can you give a solution to this problem?*

   Hint. it might be useful to consider the event E_k *"a randomly selected student passes the test on the* k*th try" for* $k = 1, 2$.

Solution. .

Problem 2 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $\left(\mathbb{R}^2, \mathcal{B}\left(\mathbb{R}^2\right), \mu_L^2\right) \equiv \mathbb{R}^2$ *be the Euclidean real plane endowed with the Borel* σ*-algebra* $\mathcal{B}\left(\mathbb{R}^2\right)$ *and the Lebesgue measure* $\mu_L^2 : \mathcal{B}\left(\mathbb{R}^2\right) \to \mathbb{R}_+$. *Let* $f : \mathbb{R}^2 \to \mathbb{R}_+$ *given by*

$$f(x, y) \stackrel{def}{=} k e^{-(x^2 - xy + y^2/2)}, \quad \forall (x, y) \in \mathbb{R}^2.$$

1. *Determine* $k \in \mathbb{R}$ *such that* $f : \mathbb{R}^2 \to \mathbb{R}_+$ *is a probability density. Hint: can you compute* $\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(y-x)^2}\, dy$ *with no computation?*

   Let $Z \equiv (X, Y)$ *be the random vector on* Ω *with density* $f : \mathbb{R}^2 \to \mathbb{R}_+$.
2. *Determine the marginal density of the entries* X *and* Y. *Are the random variables* X *and* Y *Gaussian?*
3. *Compute* $\mathbf{E}[X]$, $\mathbf{E}[Y]$, $\mathbf{D}^2[X]$, $\mathbf{D}^2[Y]$, *and* $Cov(X, Y)$.
4. *Are* X *and* Y *independent random varables?*
5. *Is the random vector* Z *Gaussian? Hint: consider the answer you gave to 4., what you know from the theory, and try to make a simple guess.*

Solution. .

Problem 3 *Let* B *[resp.* R*] be a standard Bernoulli [resp. Rademacher] random variable on a probability space* Ω. *In symbols,* $B \sim Ber(1/2)$ *[resp.* $R \sim Rad(1/2)$. *Assume that* B *and* R *are independent and define* $X \equiv B \cdot R$.

1. *Compute* $\mathbf{E}[X \mid B]$ *and* $\mathbf{E}[X \mid R]$.
2. *Are the random variables* $\mathbf{E}[X \mid B]$ *and* $\mathbf{E}[X \mid R]$ *uncorrelated? Are they independent?*
3. *Compute* $\mathbf{E}[B \mid X]$ *and* $\mathbf{E}[R \mid X]$.
4. *Are the random variables* $\mathbf{E}[B \mid X]$ *and* $\mathbf{E}[R \mid X]$ *uncorrelated? Are they independent?*
5. *Are the random variables* B *and* X *uncorrelated? Are* B *and* X *independent?*
6. *Are the random variables* R *and* X *uncorrelated? Are* R *and* X *independent?*
7. *Compute* $\mathbf{E}[B + R \mid X]$ *and* $\mathbf{E}\left[(B + R)^2 \mid X\right]$.

Solution. .

Problem 4 *Let* B *be a binomial random variable with number of trials paramenter* n *and success probability* p, *which models the number of successes in* n *independent trials, and let* $(X_k)_{k=1}^n$ *be a finite sequence of independent and exponentially distributed random variables with rate parameter* λ, *which are also independent of* B. *Study the conditional expectation*

$$\mathbf{E}\left[\sum_{k=1}^{B} X_k \mid B\right].$$

Use the properties of the conditional expectation to compute the expectation (and the variance) of the random sum

$$S_B \stackrel{def}{=} \sum_{k=1}^{B} X_k.$$

Solution. .

.

Problem 5 *Let* $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ *be a probability space and let* $(X_n)_{n\geq 1}$ *be a sequence of independent real random variables on* Ω *such that*

$$X_n \sim Ber(1/n^\alpha),$$

for some $\alpha > 0$. *Consider the sequence* $(Y_n)_{n\geq 1}$ *of real random variables on* Ω *given by*

$$Y_n \stackrel{def}{=} \min\{X_1, \ldots, X_n\}.$$

1. *After characterizing* Y_n, *study the convergence in distribution, in probability, in mean, and in quadratic mean of* $(X_n)_{n\geq 1}$ *and* $(Y_n)_{n\geq 1}$ *on varying of* $\alpha > 0$.
2. *Study the almost sure convergence of* $(X_n)_{n\geq 1}$ *on varying of* $\alpha > 0$. *Hint: it might be useful to recall when the series* $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ *or equivalently the integral* $\int_1^{+\infty} \frac{1}{x^\alpha} dx$ *converges.*
3. *Can you study the almost sure convergence of* $(Y_n)_{n\geq 1}$ *on varying of* $\alpha > 0$?

Solution. .