Probability Theory

Tutorial 1

1. If there are 30 people in a room (assume that no one is born on February 29 and that any day has the same chance of being anyone's birthday)

   (a) What is the probability of at least two of them to have the same birthday? .

   (b) What is the probability that exactly two people in the room have the same birthday?

2. Compute the probability that 10 married couples are seated at random at a round table, then no wife sits next to her husband.

3. A coin of radius r is tossed onto an infinite chessboard, where each square has a side length of 1. What is the probability that the coin lands entirely within the boundaries of a single square?

Problem 1. *If there are 30 people in a room (assume that no one is born on February 29 and that any day has the same chance of being anyone's birthday)>*

1. *What is the probability of at least two of them to have the same birthday? .*
2. *What is the probability that exactly two people in the room have the same birthday?*

Solution: *We need to compute the cardinality of the set of all possible birthdays where at least two are the same. It is easier and equivalent (why?) to compute the cardinality of the set where no two birthdays are the same. As there cannot be any repetition, this is an example of an 'ordering of length 30 of elements of* $\{1, \ldots, 365\}$

First, we compute the probability that no two people have the same birthday, corresponding to event B*. As discussed above,* B *is an ordering of length* 30 *(*$k = 30$*) from a set of cardinality* 365 *(*$n = 365$*), so*

$$P(B) = \frac{|B|}{|\Omega|} = \frac{365 \times \cdots \times (365 - 30 + 1)}{365^{30}} \approx 0.29$$

The event that at least two people have the same birthday is the complement of event B*, so the probability that at least two out of the 30 people have the same birthday will be close to 71%.*

Now we consider a slightly different question: what is the probability that exactly two people in the room have the same birthday? To construct such an example, we would need to

1. *Choose the two people that have the same birthday.*
2. *Choose a day for their birthday.*
3. *Choose a day for everyone else's birthday, so that no other birthdays are the same.*

In how many ways can we choose the two people that have the same birthday? We need to choose two numbers from $C = \{1, \ldots, 30\}$ *– this will be a sequence of length 2 with no repetition, so their cardinality is* $\binom{30}{2}$*. Now, back to our question: we have computed the number of ways we can pick the two people with the same birthday. There are 365 ways to choose their birthday. For the remaining 28 people, there will be* $364 \times \cdots \times (365 - 28)$ *ways of picking their birthdays since they all need to be different. So, the total number of ways of selecting an outcome in the event 'exactly two people have the same birthday' is*

$$\binom{30}{2} 365 \times 364 \times \cdots \times (365 - 28).$$

Finally, to compute the probability of exactly two people having the same birthday, we need to divide by the cardinality of all possible birthday combinations given by 365^{30}*, which gives approximately 0.38 or 38%.*

Problem 2. *Compute the probability that 10 married couples are seated at random at a round table, then no wife sits next to her husband.*

Solution: *If we let* E_i*,* $i = 1, 2, \ldots, 10$ *denote the event that the i-th couple sits next to each other, it follows that the desired probability is* $1 - P\left(\bigcup_{i=1}^{10} E_i\right)$*. Now, from Theorem 2,*

$$P\left(\bigcup_{i=1}^{10} E_i\right) = \sum_{i=1}^{10} P(E_i) - \sum_{i_1<i_2} P(E_{i_1}E_{i_2}) + \cdots + (-1)^{n+1} \sum_{i_1<i_2<\cdots<i_n} P(E_{i_1}E_{i_2}\cdots E_{i_n}) + \cdots - P(E_1E_2\cdots E_{10})$$

To compute $P(E_{i_1}E_{i_2}\cdots E_{i_n})$*, we first note that there are* $19!$ *ways of arranging 20 people around a round table. (Why?) The number of arrangements that result in a specified set of* n *men sitting next to their wives can most easily be obtained by first thinking of each of the* n *married couples as being single entities. If this were the case, then we would need to arrange* $20 - n$ *entities around a round table, and there are clearly* $(20 - n - 1)!$ *such arrangements. Finally, since each of the* n *married couples can be arranged next to each other in one of two possible ways, it follows that there* *are $2^n(20-n-1)!$ arrangements that result in a specified set of n men each sitting next to their wives. Therefore,*

$$P(E_{i_1}E_{i_2}\cdots E_{i_n}) = \frac{2^n(20-n-1)!}{19!}$$

To compute the full union probability, we must consider all possible combinations of couples at each step of the Inclusion-Exclusion formula. The term $\binom{10}{n}$ accounts for the number of ways to choose n couples out of the 10. For example:

- *The term $\binom{10}{1}$ represents the number of ways to choose exactly one couple to sit together. We choose one couple out of ten, which is $\binom{10}{1} = 10$ ways. Each of these single-couple events has a probability of $\frac{2^1(20-1-1)!}{19!} = \frac{2\cdot 18!}{19!}$.*
- *The term $\binom{10}{2}$ represents the number of ways to choose exactly two couples to sit together. We choose two couples out of ten, which is $\binom{10}{2} = 45$ ways.*
- *This pattern continues for all n couples.*

Thus, the probability that at least one married couple sits together is:

$$\binom{10}{1}\frac{2^1(18)!}{19!} - \binom{10}{2}\frac{2^2(17)!}{19!} + \binom{10}{3}\frac{2^3(16)!}{19!} - \cdots - \binom{10}{10}\frac{2^{10}9!}{19!} \approx .6605$$

Finally, the desired probability that no wife sits next to her husband is the complement:

$$1 - P\left(\bigcup_{i=1}^{10} E_i\right) \approx 1 - .6605 = .3395$$

Problem 3. *The Coin on the Infinite Chessboard : A coin of radius r is tossed onto an infinite chessboard, where each square has a side length of 1. What is the probability that the coin lands entirely within the boundaries of a single square?*

First, we define the Sample Space (Ω)
This problem can be solved by considering a single, representative 1×1 square. The position of the coin is determined by the coordinates of its center, (X,Y). Assuming the center is uniformly distributed over the square:

$$\Omega = \{(X,Y) \mid 0 \le X \le 1, \quad 0 \le Y \le 1\}$$

The measure (area) of the sample space is:

$$|\Omega| = 1 \times 1 = 1$$

Now, we define the Favorable Event (A)
The event A that the coin lands entirely within the square occurs if and only if the center of the coin, (X,Y), is at a distance of at least r from all four boundaries of the 1×1 square.
This condition restricts the center's coordinates to the following range:

- *The center must be at least r units away from the left edge ($X=0$) and the right edge ($X=1$):*

$$r \le X \le 1-r$$

- *The center must be at least r units away from the bottom edge ($Y=0$) and the top edge ($Y=1$):*

$$r \le Y \le 1-r$$

The region of the favorable event A is a central square with a reduced side length $(1-2r)$. The area of the favorable event $|A|$ is:

$$|A| = ((1-r)-r) \times ((1-r)-r) = (1-2r)^2$$

The probability $P(A)$ is the ratio of the favorable area to the total sample space area:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{(1-2r)^2}{1} = (1-2r)^2$$

Assumes a specific radius, $r = \frac{1}{3}$. Substituting this value into the general formula:

$$|A| = \left(1 - 2 \cdot \frac{1}{3}\right)^2 = \left(\frac{3-2}{3}\right)^2 = \left(\frac{1}{3}\right)^2 = \frac{1}{9}$$

Thus, the final probability is:

$$P(A) = \frac{1}{9}$$

Tutorial 2

1. A student buys 2 apples, 3 bananas and 5 coconuts. Every day (fir three days) the student chooses a fruit uniformly at random and eats it.
   (a) What is the probability that the student eats a coconut in day 1 and a banana in day 2?
   (b) What is the probability that on the third day the student will eat the last apple?
   (c) What is the probability that the student eats a coconut on day 2?
2. Assume there are two boxes: Box I and Box II. Box I contains w_1 white balls and b_1 black balls. Box II contains w_2 white balls and b_2 black balls. In the experiment, we first choose a box and then a ball from the chosen box.
   (a) What is the probability that a chosen ball is white?
   (b) Assume that a chosen ball is white. What is the probability that it was taken from Box I.
3. If an aircraft is present in a certain area, a radar correctly registers its presence with probability 0.99. If it is not present, the radar falsely registers an aircraft presence with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of a false alarm (a false indication of aircraft presence), and the probability of a missed detection (nothing registers, even though an aircraft is present)?

Problem 4. *A student buys 2 apples, 3 bananas and 5 coconuts. Every day (fir three days) the student chooses a fruit uniformly at random and eats it.*

1. *What is the probability that the student eats a coconut in day 1 and a banana in day 2?*
2. *What is the probability that on the third day the student will eat the last apple?*
3. *What is the probability that the student eats a coconut on day 2?*

Solution: *The sample space is the set of all triplets that can be constructed with the available fruits, with each outcome corresponding to the fruit eaten on each day. Since by the end of the three days we have full information, so the event space is the power set of the sample space. We define the events*

$A_i = \{\textit{the student eats an apple on day } i\}$,
$B_i = \{\textit{the student eats a banana on day } i\}$
$C_i = \{\textit{the student eats a coconut on day } i\}$.

1. The event 'the student eats a coconut in day 1 and a banana in day 2' corresponds to the event $C_1 \cap B_2$. *Note that the way information about the probability is encoded is through conditional probabilities: the statement 'every day the student chooses a fruit uniformly at random and eats it' can be interpreted as the conditional probability of choosing any of the remaining fruits uniformly at random, so we know that* $P(B_2|C_1) = \frac{3}{9}$.

It follows from the definition of conditional probability that $P(C_1 \cap B_2) = P(B_2|C_1)P(C_1) = \frac{3}{9} \cdot \frac{5}{10} = \frac{1}{6}$.

Writing the probability of an intersection of two events as a product of a conditional probability and a probability is called the 'multiplication rule' and can be extended to intersections of more than two events. For example, let us consider the following question.

2. Since there are exactly two apples, that means that the student will eat the first apple on either day 1 or day 2. So, if A is the event 'student eats last apple on the third day', we can write $A = (A_1 \cap A_2^c \cap A_3) \cup (A_1^c \cap A_2 \cap A_3)$.

Notice that the events $A_1 \cap A_2^c \cap A_3$ *and* $A_1^c \cap A_2 \cap A_3$ *are disjoint, therefore* $P(A) = P(A_1 \cap A_2^c \cap A_3) + P(A_1^c \cap A_2 \cap A_3) = P(A_1)P(A_2^c|A_1)P(A_3|A_1 \cap A_2^c) + P(A_1^c)P(A_2|A_1^c)P(A_3|A_1^c \cap A_2)$ $= \frac{2}{10} \cdot \frac{8}{9} \cdot \frac{1}{8} + \frac{8}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} = \frac{1}{45} + \frac{1}{45} = \frac{2}{45}$, *by using the multiplication rule twice.*

3. To compute the probability, we need to condition on what happened in day 1, but going through all possible options. In this case, there are two options that affect the computation of the conditional probability: whether the student also had a coconut on day 1 (event C_1*) or not (event* C_1^c*).*

Where is this formula coming from? We write $C_2 = (C_2 \cap C_1) \cup (C_2 \cap C_1^c)$. *So, from finite additivity, it follows that* $P(C_2) = P(C_2 \cap C_1) + P(C_2 \cap C_1^c)$.

By applying the multiplication rule to the conditional probabilities above, we get the formula which is a specific example of the law of total probabilities.

So
$P(C_2) = P(C_2|C_1) \cdot P(C_1) + P(C_2|C_1^c) \cdot P(C_1^c) = \frac{4}{9} \cdot \frac{5}{10} + \frac{5}{9} \cdot \frac{5}{10} = \frac{1}{2}$.

Problem 5. *Assume there are two boxes: Box I and Box II. Box I contains* w_1 *white balls and* b_1 *black balls. Box II contains* w_2 *white balls and* b_2 *black balls. In the experiment, we first choose a box and then a ball from the chosen box.*

1. *What is the probability that a chosen ball is white?*
2. *Assume that a chosen ball is white. What is the probability that it was taken from Box I.*

Solution:

A - an event that a chosen ball is white. B_1 *- an event that Box I was chosen.* B_2 *- an event that Box II was chosen.* $B_1 \cap B_2 = \emptyset, B_1 \cup B_2 = \Omega$ *hence* $\{B_1, B_2\}$ *is partition of* Ω. $P(B_1) = P(B_2) = \frac{1}{2}$, $P(A|B_1) = \frac{w_1}{w_1+b_1}$, $P(A|B_2) = \frac{w_2}{w_2+b_2}$.

1. $P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2)P(B_2) = \frac{1}{2}\frac{w_1}{w_1+b_1} + \frac{1}{2}\frac{w_2}{w_2+b_2}$.

2. $P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{\frac{w_1}{w_1+b_1}\frac{1}{2}}{\frac{w_1}{w_1+b_1}\frac{1}{2}+\frac{w_2}{w_2+b_2}\frac{1}{2}}$.

Problem 6. *If an aircraft is present in a certain area, a radar correctly registers its presence with probability 0.99. If it is not present, the radar falsely registers an aircraft presence with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of a false alarm (a false indication of aircraft presence), and the probability of a missed detection (nothing registers, even though an aircraft is present)?*

Solution: *A sequential representation of the sample space is appropriate here, as shown in the figure below.*

Let A and B be the events:

- $A = \{$*an aircraft is present*$\}$
- $B = \{$*the radar registers an aircraft presence*$\}$

and consider also their complements:

- $A^c = \{$*an aircraft is not present*$\}$
- $B^c = \{$*the radar does not register an aircraft presence*$\}$

The given probabilities are recorded along the corresponding branches of the tree describing the sample space. Each event of interest corresponds to a leaf of the tree, and its probability is equal to the product of the probabilities associated with the branches in a path from the root to the corresponding leaf.

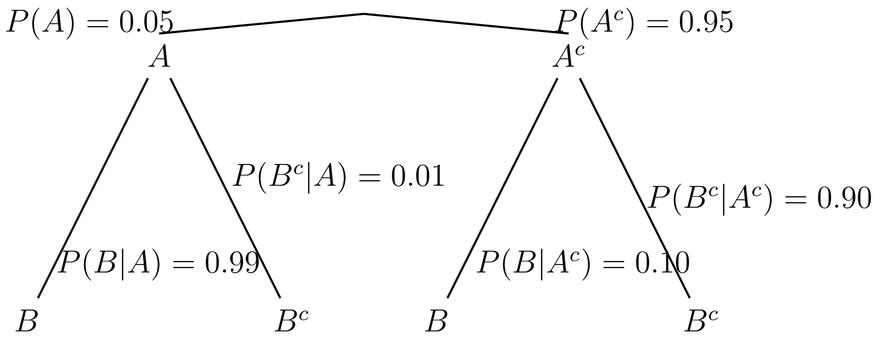

Figure 1: Sequential representation of the sample space for the radar problem.

The desired probabilities of false alarm and missed detection are:

- *The probability of a false alarm is the probability of the radar registering an aircraft when none is present. This corresponds to the event $A^c \cap B$.*

$$P(\textit{false alarm}) = P(A^c \cap B) = P(A^c)P(B|A^c) = 0.95 \cdot 0.10 = 0.095$$

- *The probability of a missed detection is the probability of the radar not registering an aircraft when one is present. This corresponds to the event $A \cap B^c$.*

$$P(\textit{missed detection}) = P(A \cap B^c) = P(A)P(B^c|A) = 0.05 \cdot 0.01 = 0.0005$$

Tutorial 3

1. Let X be a random variable with $P(X > 0) > 0$. Prove that there is a $\delta > 0$ such that $P(X \geq \delta) > 0$.

Problem 7. *Let X be a random variable with $P(X > 0) > 0$. Prove that there is a $\delta > 0$ such that $P(X \geq \delta) > 0$.*

Solution.

Method 1: Put $A = \{X > 0\}$ and $A_n = \{X \geq 1/n\}$ for all $n \in \mathbb{N}$. Then, $A_n \uparrow A$ (or A_n converges to A from below), and using Proposition 3.3.1 (Continuity of Probabilities), we have $\lim_{n\to\infty} P(A_n) = P(A)$. Given that $P(A) > 0$, it follows that there must exist some $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geq N$, we have $P(A_n) > 0$. In particular, $P(A_N) > 0$. We can therefore choose $\delta = 1/N > 0$. This completes the proof.

Method 2: Put $A = \{X > 0\}$ and $A_n = \{X \geq 1/n\}$ for all $n \in \mathbb{N}$. Then, $A = \bigcup_{n=1}^{\infty} A_n$. If for every $n \in \mathbb{N}$, we had $P(A_n) = 0$, then using a fundamental property of probability (countable subadditivity), we would have $P(A) = P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} 0 = 0$. This result, $P(A) = 0$, is a contradiction to our initial assumption that $P(X > 0) > 0$. Therefore, there must exist at least one $N \in \mathbb{N}$ such that $P(A_N) > 0$. We can then choose $\delta = 1/N > 0$. This concludes the proof.

Problem 8. *ver ejemplo 28 y 29 de convergencias. Problema 1 y 2*

Tutorial 1 Adam

Problema 3: Rompimiento de un Palo (The Stick Problem)

Un palo de 1 metro de longitud se rompió en dos puntos. ¿Cuál es la probabilidad de que con las tres piezas resultantes se pueda construir un triángulo?
Sea L_1, L_2, L_3 las longitudes de las tres piezas, tal que $L_1 + L_2 + L_3 = 1$. El espacio de parámetros inicial es $\mathbf{L} = (L_1, L_2) \in [0, 1] \times [0, 1]$.
El espacio muestral Ω relevante (donde $L_3 \geq 0$) es:

$$\Omega = \{(L_1, L_2) \mid L_1 \geq 0, L_2 \geq 0, L_1 + L_2 \leq 1\}$$

El área del espacio muestral es $|\Omega|[cite_start] = \frac{1}{2}$.
Las **condiciones de triángulo** son:

1. $L_1 + L_2 > L_3$
2. $L_1 + L_3 > L_2$
3. $L_2 + L_3 > L_1$

Sustituyendo $L_3 = 1 - L_1 - L_2$[cite: 5]:

1. $L_1 + L_2 > 1 - L_1 - L_2 \implies 2L_1 + 2L_2 > 1 \implies L_1 + L_2 > \frac{1}{2}$ [cite: 16]
2. $L_1 + (1 - L_1 - L_2) > L_2 \implies 1 - L_2 > L_2 \implies L_2 < \frac{1}{2}$ [cite: 17]
3. $L_2 + (1 - L_1 - L_2) > L_1 \implies 1 - L_1 > L_1 \implies L_1 < \frac{1}{2}$ [cite: 17]

La región favorable A está definida por las tres condiciones y $L_1, L_2 > 0$. La región A es un triángulo dentro de Ω con vértices $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$, y $(\frac{1}{2}, \frac{1}{2})$ que se obtiene del cuadrado $(\frac{1}{2}, \frac{1}{2})$.

$$|A| = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

La probabilidad $P(A)$ es[cite: 18]:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{1/8}{1/2} = \frac{1}{4}$$

Problema 4: La Moneda en el Tablero de Ajedrez (Coin on Chessboard)

Una moneda de radio r es lanzada sobre un tablero de ajedrez infinito cuyos lados tienen longitud 1. ¿Cuál es la probabilidad de que la moneda quede completamente dentro de uno de los cuadrados?
El evento A ocurre si el centro de la moneda cae en una región interior del cuadrado 1×1 que esté a una distancia de al meno r de cualquier borde. El cuadrado 1×1 se reduce a un cuadrado central de $(1-2r) \times (1-2r)$. Asumiendo $r = \frac{1}{3}$ (el valor se infiere del diagrama , el área favorable es[cite: 27, 28]:

$$|A| = \left(1 - 2 \cdot \frac{1}{3}\right)^2 = \left(\frac{1}{3}\right)^2 = \frac{1}{9}$$

La probabilidad $P(A)$, dado que el área total es $1 \times 1 = 1$, es[cite: 29]:

$$P(A) = \frac{|A|}{1} = \frac{1}{9}$$

Problema 1: El Desorden de los Impermeables (The Raincoat Problem)

Suponga n estudiantes con n impermeables. Después de clase, cada estudiante toma un impermeable al azar[cite: 30, 31]. Calcule la probabilidad de que al menos un impermeable termine con su dueño original.
El espacio muestral Ω es el conjunto de todas las permutaciones S_n[cite: 34, 38].

$$|\Omega|[cite_start] = n![cite : 40]$$

Sea A_i el evento donde el estudiante i obtiene el abrigo i[cite: 41, 43]. El evento de interés es $A = \bigcup_{i=1}^{n} A_i$ (al menos un dueño original recupera su abrigo)[cite: 45].
Usando la fórmula de Inclusión-Exclusión[cite: 45]:

$$P(A) = \sum_{j=1}^{n} (-1)^{j+1} \sum_{k_1 < \cdots < k_j} P(A_{k_1} \cap \cdots \cap A_{k_j})$$

La intersección de j eventos, $|A_{k_1} \cap \cdots \cap A_{k_j}|$, es el número de permutaciones que fijan j elementos, que es $(n-j)!$[cite: 47].

$$P(A_{k_1} \cap \cdots \cap A_{k_j}) = \frac{|A_{k_1} \cap \cdots \cap A_{k_j}|}{|\Omega|} = \frac{(n-j)!}{n!}[cite : 47]$$

El número de formas de elegir j elementos de n es $\binom{n}{j}$. La suma de las probabilidades de todas las intersecciones de j eventos es:

$$\sum_{k_1 < \cdots < k_j} P(A_{k_1} \cap \cdots \cap A_{k_j}) = \binom{n}{j} \frac{(n-j)!}{n!} = \frac{n!}{j!(n-j)!} \frac{(n-j)!}{n!} = \frac{1}{j!}$$

Sustituyendo en la fórmula de Inclusión-Exclusión[cite: 50]:

$$P(A) = \sum_{j=1}^{n} (-1)^{j+1} \frac{1}{j!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{n!}$$

Cuando n es grande $(n \to \infty)$[cite: 51, 52]:

$$P(A) \xrightarrow{n \to \infty} 1 - e^{-1} \approx 0.632[cite : 52]$$

Problema: La Fila de la Taquilla (The Ticket Office Line)

$2n$ clientes, n con billetes de \$5 y n con billetes de \$10. El boleto cuesta 5 y la taquilla comienza sin dinero Cuál es la probabilidad de que nadie tenga que esperar el cambio (evento A)?
Modelamos a los clientes como una secuencia de ± 1 de longitud $2n$, con n unos $(+\$5)$ y n menos unos $(-\$5)$.

$$|\Omega|[cite_start] = \binom{2n}{n}$$

El evento A ocurre si la suma parcial de la secuencia nunca es negativa (el dinero de la taquilla nunca cae por debajo de cero)[cite: 57]. Esto es un problema clásico de **Caminos de Lattice** (Lattice Paths).
Número de caminos que cumplen la condición (Teorema del Voto)[cite: 41, 42]:

$$|A|[cite_start] = \binom{2n}{n} - \binom{2n}{n+1}[cite : 104]$$

El término $\binom{2n}{n+1}$ es el número de caminos con $(n-1)$ subidas y $(n+1)$ bajadas, que corresponden al evento de que al menos una persona tiene que esperar el cambio[cite: 103, 96].
Probabilidad $P(A)$[cite: 104]:

$$P(A) = \frac{\binom{2n}{n} - \binom{2n}{n+1}}{\binom{2n}{n}} = 1 - \frac{\binom{2n}{n+1}}{\binom{2n}{n}}$$

$$P(A) = 1 - \frac{n}{n+1} = \frac{1}{n+1}[cite : 104]$$

Probability Theory

Tutorial 2

1. A student buys 2 apples, 3 bananas and 5 coconuts. Every day (for three days) the student chooses a fruit uniformly at random and eats it.

   (a) What is the probability that the student eats a coconut in day 1 and a banana in day 2?

   (b) What is the probability that on the third day the student will eat the last apple?

   (c) What is the probability that the student eats a coconut on day 2?

2. Assume there are two boxes: Box I and Box II. Box I contains w_1 white balls and b_1 black balls. Box II contains w_2 white balls and b_2 black balls. In the experiment, we first choose a box and then a ball from the chosen box.

   (a) What is the probability that a chosen ball is white?

   (b) Assume that a chosen ball is white. What is the probability that it was taken from Box I.

3. If an aircraft is present in a certain area, a radar correctly registers its presence with probability 0.99. If it is not present, the radar falsely registers an aircraft presence with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of a false alarm (a false indication of aircraft presence), and the probability of a missed detection (nothing registers, even though an aircraft is present)?

Problem 1. *A student buys 2 apples, 3 bananas and 5 coconuts. Every day (fir three days) the student chooses a fruit uniformly at random and eats it.*

1. *What is the probability that the student eats a coconut in day 1 and a banana in day 2?*

2. *What is the probability that on the third day the student will eat the last apple?*

3. *What is the probability that the student eats a coconut on day 2?*

Solution: *The sample space is the set of all triplets that can be constructed with the available fruits, with each outcome corresponding to the fruit eaten on each day. Since by the end of the three days we have full information, so the event space is the power set of the sample space. We define the events*

$A_i = \{\text{the student eats an apple on day } i\}$,

$B_i = \{\text{the student eats a banana on day } i\}$

$C_i = \{\text{the student eats a coconut on day } i\}$.

1. The event 'the student eats a coconut in day 1 and a banana in day 2' corresponds to the event $C_1 \cap B_2$. *Note that the way information about the probability is encoded is through conditional probabilities: the statement 'every day the student chooses a fruit uniformly at random and eats it' can be interpreted as the conditional probability of choosing any of the remaining fruits uniformly at random, so we know that* $P(B_2|C_1) = \frac{3}{9}$.

It follows from the definition of conditional probability that $P(C_1 \cap B_2) = P(B_2|C_1)P(C_1) = \frac{3}{9} \cdot \frac{5}{10} = \frac{1}{6}$.

Writing the probability of an intersection of two events as a product of a conditional probability and a probability is called the 'multiplication rule' and can be extended to intersections of more than two events. For example, let us consider the following question.

2. Since there are exactly two apples, that means that the student will eat the first apple on either day 1 or day 2. So, if A *is the event 'student eats last apple on the third day', we can write* $A = (A_1 \cap A_2^c \cap A_3) \cup (A_1^c \cap A_2 \cap A_3)$.

Notice that the events $A_1 \cap A_2^c \cap A_3$ *and* $A_1^c \cap A_2 \cap A_3$ *are disjoint, therefore* $P(A) = P(A_1 \cap A_2^c \cap A_3) + P(A_1^c \cap A_2 \cap A_3) = P(A_1)P(A_2^c|A_1)P(A_3|A_1 \cap A_2^c) + P(A_1^c)P(A_2|A_1^c)P(A_3|A_1^c \cap A_2) = \frac{2}{10} \cdot \frac{8}{9} \cdot \frac{1}{8} + \frac{8}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} = \frac{1}{45} + \frac{1}{45} = \frac{2}{45}$, *by using the multiplication rule twice.*

3. To compute the probability, we need to condition on what happened in day 1, but going through all possible options. In this case, there are two options that affect the computation of the conditional probability: whether the student also had a coconut on day 1 (event C_1*) or not (event* C_1^c*).*

Where is this formula coming from? We write $C_2 = (C_2 \cap C_1) \cup (C_2 \cap C_1^c)$. *So, from finite additivity, it follows that* $P(C_2) = P(C_2 \cap C_1) + P(C_2 \cap C_1^c)$.

By applying the multiplication rule to the conditional probabilities above, we get the formula which is a specific example of the law of total probabilities.

So

$P(C_2) = P(C_2|C_1) \cdot P(C_1) + P(C_2|C_1^c) \cdot P(C_1^c) = \frac{4}{9} \cdot \frac{5}{10} + \frac{5}{9} \cdot \frac{5}{10} = \frac{1}{2}$.

Problem 2. *Assume there are two boxes: Box I and Box II. Box I contains* w_1 *white balls and* b_1 *black balls. Box II contains* w_2 *white balls and* b_2 *black balls. In the experiment, we first choose a box and then a ball from the chosen box.*

1. *What is the probability that a chosen ball is white?*

2. *Assume that a chosen ball is white. What is the probability that it was taken from Box I.*

Solution:

A *- an event that a chosen ball is white.* B_1 *- an event that Box I was chosen.* B_2 *- an event that Box II was chosen.* $B_1 \cap B_2 = \emptyset, B_1 \cup B_2 = \Omega$ *hence* $\{B_1, B_2\}$ *is partition of* Ω. $P(B_1) = P(B_2) = \frac{1}{2}$, $P(A|B_1) = \frac{w_1}{w_1+b_1}$, $P(A|B_2) = \frac{w_2}{w_2+b_2}$.

1. $P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2)P(B_2) = \frac{1}{2}\frac{w_1}{w_1+b_1} + \frac{1}{2}\frac{w_2}{w_2+b_2}$.

2. $P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{\frac{w_1}{w_1+b_1}\frac{1}{2}}{\frac{w_1}{w_1+b_1}\frac{1}{2}+\frac{w_2}{w_2+b_2}\frac{1}{2}}$.

Problem 3. *If an aircraft is present in a certain area, a radar correctly registers its presence with probability 0.99. If it is not present, the radar falsely registers an aircraft presence with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of a false alarm (a false indication of aircraft presence), and the probability of a missed detection (nothing registers, even though an aircraft is present)?*

Solution: *A sequential representation of the sample space is appropriate here, as shown in the figure below.*

Let A and B be the events:

- $A = \{$*an aircraft is present*$\}$
- $B = \{$*the radar registers an aircraft presence*$\}$

and consider also their complements:

- $A^c = \{$*an aircraft is not present*$\}$
- $B^c = \{$*the radar does not register an aircraft presence*$\}$

The given probabilities are recorded along the corresponding branches of the tree describing the sample space. Each event of interest corresponds to a leaf of the tree, and its probability is equal to the product of the probabilities associated with the branches in a path from the root to the corresponding leaf.

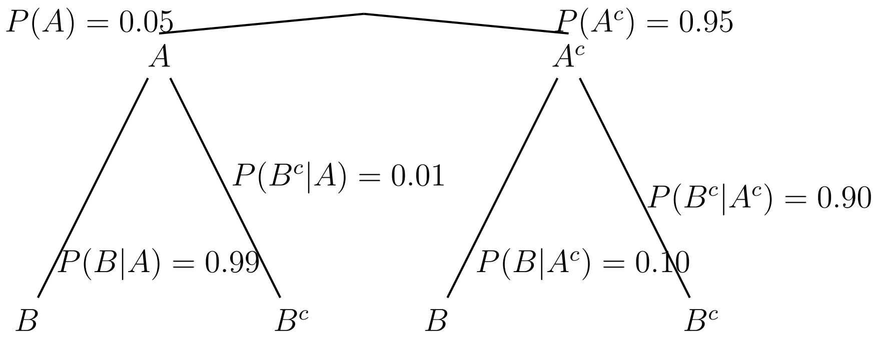

Figure 1: Sequential representation of the sample space for the radar problem.

The desired probabilities of false alarm and missed detection are:

- *The probability of a false alarm is the probability of the radar registering an aircraft when none is present. This corresponds to the event $A^c \cap B$.*

$$P(\textit{false alarm}) = P(A^c \cap B) = P(A^c)P(B|A^c) = 0.95 \cdot 0.10 = 0.095$$

- *The probability of a missed detection is the probability of the radar not registering an aircraft when one is present. This corresponds to the event $A \cap B^c$.*

$$P(\textit{missed detection}) = P(A \cap B^c) = P(A)P(B^c|A) = 0.05 \cdot 0.01 = 0.0005$$

Probability Theory

Tutorial 3

1. We throw a coin an infinite number of times. In each throw, the probability of getting a head is equal to $p \in (0, 1]$. Prove that with probability 1, a pattern of 100 heads in a row will appear infinitely many times.

2. We throw a coin an infinite number of times. In each throw, the probability of getting a head is $p \neq \frac{1}{2}$. For $n = 2, 4, 6, \ldots$, consider the events A_n: in the first n throws, the number of heads and tails is the same. Prove that with probability 1, only finitely many of the events $A_2, A_4, \ldots, A_{2k}, \ldots$ occur.

3. Suppose $P\left(\left[0, \frac{8}{4+n}\right]\right) = \frac{2+e^{-n}}{6}$ for all $n = 1, 2, 3, \ldots$. What must $P(\{0\})$ be?

4. Suppose $P([0, 1]) = 1$, but $P\left(\left[\frac{1}{n}, 1\right]\right) = 0$ for all $n = 1, 2, 3, \ldots$. What must $P(\{0\})$ be?

Problem 1

We throw a coin an infinite number of times. In each throw, the probability of getting a head is equal to $p \in (0, 1]$. Prove that with probability 1, a pattern of 100 heads in a row will appear infinitely many times.

Proof

This proof relies fundamentally on the Second Borel-Cantelli Lemma for independent events.

Second Borel-Cantelli Lemma (for Independent Events): Let $A_1, A_2, \dots$ be a sequence of independent events. Then, the probability that infinitely many of the events occur, denoted $P(A_k \text{ i.o.})$, is 1 if and only if the sum of their probabilities diverges:

$$P(A_k \text{ i.o.}) = 1 \iff \sum_{k=1}^{\infty} P(A_k) = \infty$$

The event of interest is the appearance of 100 consecutive heads. To apply the Borel-Cantelli Lemma, we must construct an infinite sequence of mutually independent events.

]We partition the infinite sequence of coin throws into non-overlapping blocks, each of length $m = 100$. For $k \in \{0, 1, 2, \dots\}$, let A_k be the event that the k-th block consists entirely of heads. That is, A_k occurs if the throws from position $100k + 1$ up to position $100k + 100$ are all heads. Since all individual coin throws are mutually independent, and the blocks A_k and A_j for $j \neq k$ occupy completely distinct sets of throws, the sequence of events $A_0, A_1, A_2, \dots$ is mutually independent.

The probability of getting a head in a single throw is $P(\text{Head}) = p$. Since A_k is the event of 100 independent heads in a row:

$$P(A_k) = p \cdot p \cdot \ldots \cdot p \quad \text{(100 times)} = p^{100}$$

We now compute the sum of the probabilities of these independent events:

$$\sum_{k=0}^{\infty} P(A_k) = \sum_{k=0}^{\infty} p^{100}$$

Since $p \in (0, 1]$, the smallest possible value for p^{100} is $\lim_{p \to 0^+} p^{100} = 0$, but since p is strictly greater than 0, $p^{100} > 0$. The sum is an infinite series where every term is the same positive constant $c = p^{100}$:

$$\sum_{k=0}^{\infty} p^{100} = p^{100} + p^{100} + p^{100} + \ldots = \infty$$

The sum diverges.

Since the events A_k are independent and $\sum_{k=0}^{\infty} P(A_k) = \infty$, the Second Borel-Cantelli Lemma guarantees that the probability of the sequence of events A_k occurring infinitely often is 1:

$$P(\{A_k \text{ i.o.}\}) = 1$$

Let A be the event that "a pattern of 100 heads in a row appears infinitely many times." The event $\{A_k \text{ i.o.}\}$ means that the pattern appears infinitely many times only within the predetermined disjoint blocks. Since every occurrence of the pattern within a disjoint block is also an occurrence of the general pattern A, we have the set containment:

$$\{A_k \text{ i.o.}\} \subseteq A$$

Because $\{A_k \text{ i.o.}\}$ is a subset of A, its probability provides a lower bound for $P(A)$:

$$P(A) \geq P(\{A_k \text{ i.o.}\}) = 1$$

Since probability cannot exceed 1, we must have $P(A) = 1$.

Therefore, with probability 1, a pattern of 100 heads in a row will appear infinitely many times.

Problem 2

We throw a coin an infinite number of times. In each throw, the probability of getting a head is $p \neq \frac{1}{2}$. For $n = 2, 4, 6, \dots$, consider the events A_n: in the first n throws, the number of heads and tails is the same. Prove that with probability 1, only finitely many of the events $A_2, A_4, \dots, A_{2k}, \dots$ occur.

Solution

This problem is solved using the First Borel-Cantelli Lemma. This lemma provides a simple condition for an event to occur only a finite number of times (finitely often, or f.o.).

First Borel-Cantelli Lemma: Let $A_1, A_2, \dots$ be any sequence of events. If the sum of their probabilities converges, then the probability that infinitely many of the events occur is zero:

$$\sum_{k=1}^{\infty} P(A_k) < \infty \implies P(\{A_k \text{ occurs infinitely often (i.o.)}\}) = 0$$

The conclusion $P(\{A_k \text{ i.o.}\}) = 0$ is equivalent to $P(\{A_k \text{ occurs finitely often (f.o.)}\}) = 1$.

Let A_{2k} be the event that in the first $2k$ throws, the number of heads equals the number of tails. This requires exactly k heads and k tails. Since each throw is an independent Bernoulli trial with success probability p (Head) and failure probability $(1-p)$ (Tail), the number of ways to get k heads in $2k$ trials is $\binom{2k}{k}$.

$$P(A_{2k}) = \binom{2k}{k} p^k (1-p)^k$$

We must check the convergence of the series $\sum_{k=1}^{\infty} P(A_{2k})$. We use the Ratio Test.

Let $a_k = P(A_{2k})$. We compute the limit of the ratio $\frac{a_{k+1}}{a_k}$ as $k \to \infty$.

$$\frac{P(A_{2(k+1)})}{P(A_{2k})} = \frac{\binom{2k+2}{k+1} p^{k+1} (1-p)^{k+1}}{\binom{2k}{k} p^k (1-p)^k}$$

We simplify the combinatorial term:

$$\frac{\binom{2k+2}{k+1}}{\binom{2k}{k}} = \frac{\frac{(2k+2)!}{(k+1)!(k+1)!}}{\frac{(2k)!}{k!k!}} = \frac{(2k+2)!}{(2k)!} \cdot \frac{(k!)^2}{((k+1)!)^2}$$

$$= (2k+2)(2k+1) \cdot \frac{1}{(k+1)^2} = \frac{2(k+1)(2k+1)}{(k+1)^2} = \frac{2(2k+1)}{k+1}$$

Now, we compute the limit L:

$$L = \lim_{k \to \infty} \left[p(1-p) \cdot \frac{2(2k+1)}{k+1} \right]$$

$$L = p(1-p) \cdot \lim_{k \to \infty} \left[\frac{4k+2}{k+1} \right] = p(1-p) \cdot 4$$

$$L = 4p(1-p)$$

The Ratio Test states that the series converges if $L < 1$. We analyze the condition $4p(1-p) < 1$:

$$4p - 4p^2 < 1$$

$$0 < 4p^2 - 4p + 1$$

The right side is a perfect square:

$$0 < (2p-1)^2$$

This inequality holds true for all values of p except when $2p - 1 = 0$, which occurs when $p = 1/2$. Since the problem explicitly states that $p \neq 1/2$, the limit L satisfies the condition:

$$L = 4p(1-p) < 1$$

Therefore, the series converges:

$$\sum_{k=1}^{\infty} P(A_{2k}) < \infty$$

By the First Borel-Cantelli Lemma, since the sum of the probabilities of the events A_{2k} converges, the probability that infinitely many of these events occur is zero:

$$P(\{A_{2k} \text{ i.o.}\}) = 0$$

Consequently, the probability that only a finite number of these events occur is 1:

$$P(\{A_{2k} \text{ f.o.}\}) = 1$$

This proves that with probability 1, the number of heads and tails will only be equal in the first $2k$ throws for a finite number of indices k.

Problem 3

Suppose $P\left(\left[0, \frac{8}{4+n}\right]\right) = \frac{2+e^{-n}}{6}$ for all $n = 1, 2, 3, \ldots$. What must $P(\{0\})$ be?

Solution

We define the sequence of events $A_n = \left[0, \frac{8}{4+n}\right]$. This is a **decreasing** sequence of sets since $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$.

The limit of the sets is the intersection:

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \left[0, \frac{8}{4+n}\right] = \{0\}$$

By the **Continuity of Probability for Decreasing Events**:

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

Substituting the given probability function:

$$P(\{0\}) = \lim_{n\to\infty} P\left(\left[0, \frac{8}{4+n}\right]\right) = \lim_{n\to\infty} \left(\frac{2+e^{-n}}{6}\right)$$

Since $\lim_{n\to\infty} e^{-n} = 0$:

$$P(\{0\}) = \frac{2+0}{6} = \frac{1}{3}$$

Problem 4

Suppose $P([0,1]) = 1$, but $P\left(\left[\frac{1}{n}, 1\right]\right) = 0$ for all $n = 1, 2, 3, \ldots$. What must $P(\{0\})$ be?

Solution

Let $B_n = \left[\frac{1}{n}, 1\right]$. This is an **increasing** sequence of sets: $B_1 \subseteq B_2 \subseteq B_3 \subseteq \ldots$.

1. **Calculate $P((0,1])$:** The limit of the sets is the union:

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \left[\frac{1}{n}, 1\right] = (0,1]$$

   By the **Continuity of Probability for Increasing Events**:

$$P((0,1]) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n)$$

   Since $P(B_n) = 0$ for all n:

$$P((0,1]) = \lim_{n\to\infty} 0 = 0$$

2. **Use Additivity:** The sample space $[0,1]$ can be written as the union of two disjoint sets:

$$[0,1] = \{0\} \cup (0,1]$$

   By the additivity axiom:

$$P([0,1]) = P(\{0\}) + P((0,1])$$

3. **Solve for $P(\{0\})$:**

$$1 = P(\{0\}) + 0$$

$$P(\{0\}) = 1$$

Probability Theory

Tutorial 4

1. Let the sample space be $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let $\mathcal{A}$ be the σ-algebra generated by the sets $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$, denoted as $\mathcal{A} = \sigma(A, B)$.

   (a) List all sets in $\mathcal{A}$.

   (b) Is the function $X(\omega)$ defined as:

   $$X(\omega) = \begin{cases} 2 & \text{if } \omega \in \{1, 2, 3, 4\} \\ 7 & \text{if } \omega \in \{5, 6\} \end{cases}$$

   a random variable over the measurable space $(\Omega, \mathcal{A})$?

   (c) Give an example of a function on Ω that is not a random variable over $(\Omega, \mathcal{A})$.

2. Let X be a random variable with $P(X > 0) > 0$. Prove that there is a $\delta > 0$ such that $P(X \geq \delta) > 0$.

3. A random variable X has a distribution function $F_X(x)$ given by:

   $$F_X(x) = \begin{cases} 0 & x < 2 \\ 1/2 & 2 \leq x < 4 \\ 3/4 & 4 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

   Calculate $P(X = 2)$, $P(X = 4)$, and $P(X = 6)$.

4. Let $\Omega = [0, 3]$ with probability uniform and $X(\omega)$ be the random variable given by:

   $$X(\omega) = \begin{cases} \omega + 1 & \text{si } 0 \leq \omega \leq 1 \\ 2 & \text{si } 1 < \omega \leq 1.5 \\ 3 & \text{si } 1.5 < \omega < 2 \\ -\omega + 4 & \text{si } 2 \leq \omega \leq 3 \end{cases}$$

   Find $F_X(t)$, the distribution function of X.

Problem 1

Let the sample space be $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let $\mathcal{A}$ be the σ-algebra generated by the sets $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$, denoted as $\mathcal{A} = \sigma(A, B)$.

Solution

1. List all sets in $\mathcal{A}$.
2. Is the function $X(\omega)$ defined as:

$$X(\omega) = \begin{cases} 2 & \text{if } \omega \in \{1, 2, 3, 4\} \\ 7 & \text{if } \omega \in \{5, 6\} \end{cases}$$

   a random variable over the measurable space $(\Omega, \mathcal{A})$?
3. Give an example of a function on Ω that is not a random variable over $(\Omega, \mathcal{A})$.
4. Show that there exists a probability measure P on $(\Omega, \mathcal{A})$ such that $P(A)$ is zero or one for all $A \in \mathcal{A}$, yet P is not a point mass.

Let the sample space be $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let $\mathcal{A} = \sigma(\{1, 2, 3, 4\}, \{3, 4, 5, 6\})$ be the σ-algebra generated by the sets $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5, 6\}$.

a) List all sets in $\mathcal{A}$

The σ-algebra $\mathcal{A}$ consists of all possible unions of its **atoms** (the smallest non-empty sets in $\mathcal{A}$), found by taking intersections of the generating sets and their complements:

- $C_1 = A \cap B^c = \{1, 2, 3, 4\} \cap \{1, 2\} = \{1, 2\}$
- $C_2 = A \cap B = \{1, 2, 3, 4\} \cap \{3, 4, 5, 6\} = \{3, 4\}$
- $C_3 = A^c \cap B = \{5, 6\} \cap \{3, 4, 5, 6\} = \{5, 6\}$

Since there are $k = 3$ non-empty atoms, $\mathcal{A}$ contains $2^3 = 8$ sets.

$$\mathcal{A} = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{1, 2, 3, 4\}, \{3, 4, 5, 6\}, \{1, 2, 5, 6\}, \{1, 2, 3, 4, 5, 6\}\}$$

(Note: $\{1, 2, 3, 4\} = A$, $\{3, 4, 5, 6\} = B$, $\{1, 2, 5, 6\} = (A \cap B)^c$).

b) Is the function $X(\omega)$ a random variable over $(\Omega, \mathcal{A})$?

The function is:

$$X(\omega) = \begin{cases} 2 & \text{if } \omega \in \{1, 2, 3, 4\} \\ 7 & \text{if } \omega \in \{5, 6\} \end{cases}$$

X is a random variable if $X^{-1}(E) \in \mathcal{A}$ for every Borel set E. We check the preimages of the values in the range $\{2, 7\}$:

1. $X^{-1}(\{2\}) = \{1, 2, 3, 4\} = A$. Since $A \in \mathcal{A}$.
2. $X^{-1}(\{7\}) = \{5, 6\} = C_3$. Since $C_3 \in \mathcal{A}$.

Since the preimages of all values in the range belong to $\mathcal{A}$, **YES, X is a random variable** over $(\Omega, \mathcal{A})$.

c) Give an example of a function on Ω that is not a random variable over $(\Omega, \mathcal{A})$

A function Y is not a random variable if it distinguishes between elements within the same atom. The set $\{1\}$ is not in $\mathcal{A}$ because $\mathcal{A}$ cannot separate $\omega = 1$ from $\omega = 2$.

Define the function $Y(\omega)$ as:

$$Y(\omega) = \begin{cases} 10 & \text{if } \omega = 1 \\ 0 & \text{if } \omega \in \{2, 3, 4, 5, 6\} \end{cases}$$

The preimage of 10 is:

$$Y^{-1}(\{10\}) = \{1\}$$

Since $\{1\} \notin \mathcal{A}$, the condition for Y to be a random variable fails.

Therefore, **Y is not a random variable** over $(\Omega, \mathcal{A})$.

Problem 2

Let X be a random variable with $P(X > 0) > 0$. Prove that there is a $\delta > 0$ such that $P(X \geq \delta) > 0$.

Solution

Method 1: Put $A = \{X > 0\}$ and $A_n = \{X \geq 1/n\}$ for all $n \in \mathbb{N}$. Then, $A_n \uparrow A$ (or A_n converges to A from below), and using Continuity of Probabilities, we have $\lim_{n\to\infty} P(A_n) = P(A)$. Given that $P(A) > 0$, it follows that there must exist some $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geq N$, we have $P(A_n) > 0$. In particular, $P(A_N) > 0$. We can therefore choose $\delta = 1/N > 0$. This completes the proof.

Method 2: Put $A = \{X > 0\}$ and $A_n = \{X \geq 1/n\}$ for all $n \in \mathbb{N}$. Then, $A = \bigcup_{n=1}^{\infty} A_n$. If for every $n \in \mathbb{N}$, we had $P(A_n) = 0$, then using a fundamental property of probability (countable subadditivity), we would have $P(A) = P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} 0 = 0$. This result, $P(A) = 0$, is a contradiction to our initial assumption that $P(X > 0) > 0$. Therefore, there must exist at least one $N \in \mathbb{N}$ such that $P(A_N) > 0$. We can then choose $\delta = 1/N > 0$. This concludes the proof.

Problem 3

A random variable X has a distribution function $F_X(x)$ given by:

$$F_X(x) = \begin{cases} 0 & x < 2 \\ 1/2 & 2 \leq x < 4 \\ 3/4 & 4 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

Calculate $P(X = 2)$, $P(X = 4)$, and $P(X = 6)$.

Solution

The point probability at a point t is calculated by the jump in the distribution function: $P(X = t) = F_X(t) - \lim_{x \to t^-} F_X(x)$.

1. **For $t = 2$:**
$$P(X = 2) = F_X(2) - \lim_{x \to 2^-} F_X(x) = \frac{1}{2} - 0 = \mathbf{\frac{1}{2}}$$
2. **For $t = 4$:**
$$P(X = 4) = F_X(4) - \lim_{x \to 4^-} F_X(x) = \frac{3}{4} - \frac{1}{2} = \frac{3}{4} - \frac{2}{4} = \mathbf{\frac{1}{4}}$$
3. **For $t = 6$:**
$$P(X = 6) = F_X(6) - \lim_{x \to 6^-} F_X(x) = 1 - \frac{3}{4} = \mathbf{\frac{1}{4}}$$

Problem 4

Let $\Omega = [0, 3]$ with probability uniform and $X(\omega)$ be the random variable given by:

$$X(\omega) = \begin{cases} \omega + 1 & \text{si } 0 \le \omega \le 1 \\ 2 & \text{si } 1 < \omega \le 1.5 \\ 3 & \text{si } 1.5 < \omega < 2 \\ -\omega + 4 & \text{si } 2 \le \omega \le 3 \end{cases}$$

Find $F_X(t)$, the distribution function of X.

Solution

The sample space is $\Omega = [0, 3]$ with a uniform probability measure.

- Total Length: $|\Omega| = 3$.
- Probability of any event $A \subseteq \Omega$: $P(A) = \frac{|A|}{3}$.

The random variable $X(\omega)$ is defined piecewise over Ω:

$$X(\omega) = \begin{cases} \omega + 1 & \text{if } 0 \le \omega \le 1 \quad (\text{Interval } I_1) \\ 2 & \text{if } 1 < \omega \le 1.5 \quad (\text{Interval } I_2) \\ 3 & \text{if } 1.5 < \omega < 2 \quad (\text{Interval } I_3) \\ -\omega + 4 & \text{if } 2 \le \omega \le 3 \quad (\text{Interval } I_4) \end{cases}$$

We analyze the range of X and the probability mass contributed by each interval I_i:

1. **Interval I_1 (Continuous):** $\omega \in [0, 1]$. Length $= 1$. $P(I_1) = 1/3$. $X(\omega) = \omega + 1$ covers the continuous range $[\mathbf{1}, \mathbf{2}]$.
2. **Interval I_2 (Point Mass at 2):** $\omega \in (1, 1.5]$. Length $= 0.5$. This contributes a **discrete mass** at $X = 2$: $P(X = 2 \text{ from } I_2) = P(I_2) = \frac{0.5}{3} = \mathbf{1/6}$.
3. **Interval I_3 (Point Mass at 3):** $\omega \in (1.5, 2)$. Length $= 0.5$. This contributes a **discrete mass** at $X = 3$: $P(X = 3) = P(I_3) = \frac{0.5}{3} = \mathbf{1/6}$.
4. **Interval I_4 (Continuous):** $\omega \in [2, 3]$. Length $= 1$. $P(I_4) = 1/3$. $X(\omega) = -\omega + 4$ covers the continuous range $[\mathbf{1}, \mathbf{2}]$ (since $X(2) = 2$ and $X(3) = 1$).

The overall range of X is $[\mathbf{1}, \mathbf{2}] \cup \{\mathbf{3}\}$. The minimum value is 1, and the maximum is 3.
We find $F_X(t) = P(X \le t)$ by summing the continuous and discrete probabilities up to t.

Case 1: $t < 1$

Since 1 is the minimum value X can take, $P(X \le t) = 0$.

$$F_X(t) = 0$$

Case 2: $1 \le t < 2$

Probability accumulates from the two continuous segments, I_1 and I_4.

- **Contribution from I_1 ($X = \omega + 1$):** $X \le t \implies \omega + 1 \le t \implies \omega \le t - 1$. Since $t \in [1, 2)$, $t - 1 \in [0, 1)$, so $\omega \in [0, t - 1]$.

$$P(I_1 \text{ part}) = P(0 \le \omega \le t - 1) = \frac{t-1}{3}$$

- **Contribution from I_4 ($X = -\omega + 4$):** $X \le t \implies -\omega + 4 \le t \implies \omega \ge 4 - t$. Since $t \in [1, 2)$, $4 - t \in [2, 3)$, so $\omega \in [4 - t, 3]$.

$$P(I_4 \text{ part}) = P(4 - t \le \omega \le 3) = \frac{3 - (4 - t)}{3} = \frac{t-1}{3}$$

The total accumulated probability is:

$$F_X(t) = \frac{t-1}{3} + \frac{t-1}{3} = \mathbf{\frac{2(t-1)}{3}}$$

Case 3: $2 \le t < 3$

The CDF must include the total continuous mass plus the discrete mass at $X = 2$.

- Continuous mass up to 2: 2/3.
- Discrete mass at $X = 2$ (from I_2): 1/6.

For $t \in [2, 3)$, the CDF is constant at the new accumulated value:

$$F_X(t) = \frac{2}{3} + \frac{1}{6} = \frac{4}{6} + \frac{1}{6} = \mathbf{5/6}$$

Case 4: $t \ge 3$

The CDF must include the remaining discrete mass at $X = 3$.

- Mass accumulated before $t = 3$: 5/6.
- Discrete mass at $X = 3$ (from I_3): 1/6.

The total accumulated probability is:

$$F_X(t) = \frac{5}{6} + \frac{1}{6} = 1$$

$$F_X(t) = \begin{cases} 0 & \text{if } t < 1 \\ \frac{2(t-1)}{3} & \text{if } 1 \le t < 2 \\ 5/6 & \text{if } 2 \le t < 3 \\ 1 & \text{if } t \ge 3 \end{cases}$$

Probability Theory

Tutorial 5

1. The Mathematics Final Exam is multiple choice (MC). It consists of 20 exercises, each with 4 options (only one option is correct). To **pass** the exam, two conditions must be met simultaneously:

   (a) The number of correct answers must be **greater than** the number of incorrect answers.

   (b) There must be a **minimum of 8** correct answers.

   The following piece of advice is common among students: *"Only answer if you know the answer."* Is this true? Of course, if one knew the answers to all questions, the most sensible thing is to answer the whole exam. So, we assume the student does not know the answers and answers all questions randomly and independently.

   Initially, let's compare two scenarios:

   A Answer **15** questions.

   B Answer **20** questions.

   Which of the two scenarios yields a higher probability of passing? Is there a better scenario? What happens if the only restriction is having 8 correct answers?

2. In a cage, there is one mouse and two chicks. The probability that each animal leaves the cage in the next hour is 0.3, and this event is independent for all animals. After one hour, the following random variables are defined:

   - X: The number of legs remaining in the cage. (mouse has 4 legs and the chicks 2 legs)
   - Y: The number of animals remaining in the cage.

   Find the joint probability mass function of (X, Y) and determine the marginal distributions . Are X and Y independent?

3. A retailer verifies that the demand for cars (X) is a random variable with a **Poisson distribution** with parameter $\lambda = 2$ cars per week. The retailer restocks every Monday morning to have 4 cars in stock ($S = 4$).

   The questions posed are:

   (a) What is the probability of selling out the entire stock during the week?

   (b) What is the probability that the retailer is unable to fulfill at least one order?

   (c) What is the minimum stock needed to ensure $P(\text{fulfilling all orders}) \geq 0.99$?

   (d) What is the distribution of the number of cars sold per week?

   (e) If the profit is 1000 € per car sold and the fixed cost is 800 €, what is the distributon of the Weekly Profit (P_G)?

Problem 1

The Mathematics Final Exam is multiple choice (MC). It consists of 20 exercises, each with 4 options (only one option is correct). To **pass** the exam, two conditions must be met simultaneously:

1. The number of correct answers must be **greater than** the number of incorrect answers.
2. There must be a **minimum of 8** correct answers.

The following piece of advice is common among students: *"Only answer if you know the answer."* Is this true? Of course, if one knew the answers to all questions, the most sensible thing is to answer the whole exam. So, we assume the student does not know the answers and answers all questions randomly and independently.

Initially, let's compare two scenarios:

A Answer **15** questions.

B Answer **20** questions.

Which of the two scenarios yields a higher probability of passing? Is there a better scenario? What happens if the only restriction is having 8 correct answers?

Solution

Since each exercise has 4 options and only 1 is correct, the probability of success (answering correctly at random) is:

$$p = P(\text{success}) = \frac{1}{4} = 0.25$$

Strategy A: Answering 15 Questions

Let S_{15} be the number of correctly answered questions out of 15 attempts.

$$S_{15} \sim B(n = 15, p = 1/4)$$

The two passing conditions are:

1. $S_{15} \geq 8$ (*Minimum of 8 correct*)
2. $S_{15} > (\text{Total answered}) - S_{15}$
   $S_{15} > 15 - S_{15} \implies 2S_{15} > 15 \implies S_{15} > 7.5$

The combined passing condition is $\mathbf{S_{15} \geq 8}$ (since S_{15} must be an integer).

$$\begin{aligned} P(\text{pass in Scenario A}) &= P(\{S_{15} \geq 8\} \cap \{S_{15} > 7.5\}) \\ &= P(S_{15} \geq 8) \\ &= \sum_{k=8}^{15} p_{S_{15}}(k) \\ &= 1 - P(S_{15} \leq 7) \\ &= 1 - F_{S_{15}}(7) \approx \mathbf{0.01729} \quad \text{(low!)} \end{aligned}$$

Strategy B: Answering 20 Questions

Let S_{20} be the number of correctly answered questions out of 20 attempts.

$$S_{20} \sim B(n = 20, p = 1/4)$$

The two passing conditions are:

1. $S_{20} \geq 8$ (*Minimum of 8 correct*)
2. $S_{20} > 20 - S_{20} \implies 2S_{20} > 20 \implies S_{20} > 10$

The combined passing condition is $\mathbf{S_{20} > 10}$ (or $\mathbf{S_{20} \geq 11}$), as it is the most restrictive requirement.

$$\begin{aligned} P(\text{pass in Scenario B}) &= P(\{S_{20} \geq 8\} \cap \{S_{20} > 10\}) \\ &= P(S_{20} > 10) \\ &= \sum_{k=11}^{20} p_{S_{20}}(k) \\ &= 1 - P(S_{20} \leq 10) \\ &= 1 - F_{S_{20}}(10) \approx \mathbf{0.003942} \quad \text{(even lower!)} \end{aligned}$$

Partial Conclusion: The probability of passing with Strategy A ($n = 15$) is ≈ 4.388 times greater than with Strategy B ($n = 20$). The higher probability of passing is achieved by answering **15** questions.

Intermediate Strategies: Is There a Better One?

Let S_n be the number of correctly answered questions out of n attempts, where $8 \leq n \leq 20$.

$$S_n \sim B(n, p = 1/4)$$

The general passing condition is: $S_n \geq 8$ and $S_n > n - S_n \implies S_n > n/2$.

$$\mathbf{P}(\text{pass}) = \mathbf{P}(\mathbf{S_n} > \max\{\mathbf{7}, \mathbf{n/2}\}) = \mathbf{1} - \mathbf{F_{S_n}}(\max\{\mathbf{7}, \lfloor \mathbf{n/2} \rfloor\})$$

Probability of Passing for Different n

The table shows the probability of passing for different values of n:

n	$\max\{7, \lfloor n/2 \rfloor\}$	Prob. to Pass
8	7	0.00002
9	7	0.00010
10	7	0.00041
11	7	0.00124
12	7	0.00282
13	7	0.00563
14	7	0.01026
15	**7**	**0.01730**
16	8	0.00754
17	8	0.01237
18	9	0.00543
19	9	0.00885
20	10	0.00394

The highest probability of passing is obtained by answering **15** questions.

Scenario without Incorrect Answer Restrictions

If the only restriction were having a minimum of 8 correct answers ($S_n \geq 8$), and the condition $S_n > n - S_n$ were eliminated:

$$P(\text{pass w/o restr.}) = P(S_n \geq 8) = 1 - F_{S_n}(7)$$

In this case, the probability is strictly increasing with n, making $\mathbf{n} = \mathbf{20}$ the best strategy. However, under the **actual system** with the double restriction, the optimal strategy is $n = 15$.

Problem 2

In a cage, there is one mouse and two chicks. The probability that each animal leaves the cage in the next hour is 0.3, and this event is independent for all animals. After one hour, the following random variables are defined:

- X: The number of legs remaining in the cage.
- Y: The number of animals remaining in the cage.

Find the joint probability mass function of (X, Y), determine the marginal distributions, check for independence, and calculate a conditional probability.

Solution

The cage contains 1 mouse (4 legs) and 2 chicks (2 legs each). The probability that any individual animal **stays** in the cage is $p = 1 - 0.3 = 0.7$. All outcomes are independent.

The total number of animals is 3 (1 mouse, 2 chicks). The maximum number of legs is $4(1)+2(2) = 8$.

The possible values for the random variables X and Y are:

- Y: Total number of animals remaining in the cage. $Y \in \{0, 1, 2, 3\}$.
- X: Total number of legs remaining in the cage. $X \in \{0, 2, 4, 6, 8\}$.

The entries $p_{X,Y}(x, y)$ are calculated based on the independent Bernoulli trials (Mouse stay ~ 0.7; Chick stay ~ 0.7).

Table 1: Joint PMF $p_{X,Y}(x, y)$ and Marginal PMFs

Legs $X \backslash$ Animals Y	**0**	**1**	**2**	**3**	**Marginal $p_X(x)$**
0	0.027	0	0	0	**0.027**
2	0	0.126	0	0	**0.126**
4	0	0.063	0.147	0	**0.210**
6	0	0	0.294	0	**0.294**
8	0	0	0	0.343	**0.343**
Marginal $p_Y(y)$	**0.027**	**0.189**	**0.441**	**0.343**	**1.000**

Marginal Probability Mass Functions

The marginal PMFs are obtained by summing the rows (for $p_X(x)$) and columns (for $p_Y(y)$) in the joint PMF table (Table 1).

Independence Check

X and Y are independent if and only if $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ for all possible pairs (x, y).

Let's test the pair $(X = 8, Y = 3)$:

- **Joint Probability:** $p_{X,Y}(8, 3) = 0.343$
- **Product of Marginals:** $p_X(8) \cdot p_Y(3) = (0.343) \cdot (0.343) \approx 0.1176$

Since $p_{X,Y}(8, 3) \neq p_X(8)p_Y(3)$, the variables **X** and **Y** are **NOT independent**.

Problem 3

A retailer verifies that the demand for cars (X) is a random variable with a **Poisson distribution** with parameter $\lambda = 2$ cars per week. The retailer restocks every Monday morning to have 4 cars in stock ($S = 4$).

The questions posed are:

1. What is the probability of selling out the entire stock during the week?
2. What is the probability that the retailer is unable to fulfill at least one order?
3. What is the minimum stock needed to ensure $P(\text{fulfilling all orders}) \geq 0.99$?
4. What is the distribution of the number of cars sold per week?
5. If the profit is 1000 € per car sold and the fixed cost is 800 €, what is the distributon of the Weekly Profit (P_G)?

Solution

Let X be the weekly demand. $X \sim \text{Poisson}(2)$. The PMF is $P(X = k) = \frac{e^{-2}2^k}{k!}$. Stock $S = 4$.

Probability of Selling Out $P(X \geq 4)$

$$P(X \geq 4) = 1 - P(X \leq 3)$$

$$P(X \leq 3) = e^{-2} \sum_{k=0}^{3} \frac{2^k}{k!} = e^{-2}\left(1 + 2 + 2 + \frac{4}{3}\right) = \frac{19}{3}e^{-2} \approx 0.8571$$

$$P(X \geq 4) = 1 - 0.8571 \approx \mathbf{0.1429}$$

Probability of Unable to Fulfill ($P(X > 4)$)

$$P(X > 4) = 1 - P(X \leq 4)$$

$$P(X = 4) = \frac{e^{-2}2^4}{4!} = \frac{2}{3}e^{-2} \approx 0.0902$$

$$P(X > 4) \approx 1 - (0.8571 + 0.0902) = 1 - 0.9473 \approx \mathbf{0.0527}$$

Minimum Stock (S) for $P(X \leq S) \geq 0.99$

Checking the CDF for X:

- $P(X \leq 5) \approx 0.9834$
- $P(X \leq 6) \approx 0.9955$

The minimum stock is **6** cars.

Distribution of cars Sold (V)

V is the number of cars sold, defined as $V = \min(X, 4)$. V is a **censored Poisson variable**. Its PMF is:

$$P(V = v) = \begin{cases} P(X = v) & \text{for } v \in \{0, 1, 2, 3\} \\ P(X \geq 4) & \text{for } v = 4 \\ 0 & \text{otherwise} \end{cases}$$

Probability Mass Function of the Profit ($P(G = g)$)

Since $G = 1000V - 800$ implies $V = \frac{g+800}{1000}$, the possible values for G are $g \in \{-800, 200, 1200, 2200, 3200\}$.

The PMF of G is:

$$P(G = g) = \begin{cases} P(X = k) = \frac{e^{-2}2^k}{k!} & \text{if } g = 1000k - 800, \text{ for } k \in \{0, 1, 2, 3\} \\ P(X \geq 4) = 1 - \frac{19}{3}e^{-2} & \text{if } g = 3200 \quad (k = 4) \\ 0 & \text{otherwise} \end{cases}$$

Probability Theory

Tutorial 6

1. In a certain human population, the Cephalic Index I (skull width expressed as a percentage of its length) is normally distributed among individuals. There is 58% with $I \leq 75$, 38% with $75 \leq I \leq 80$, and 4% with $I > 80$. Find the density function of the index, $f_I(i)$, and the probability $P(78 \leq I \leq 82)$.

2. Alice and Peter agreed to meet at 8 PM to go to the movies. Since they are not punctual, their arrival times X and Y can be assumed to be independent random variables with a uniform distribution between 8 PM and 9 PM.

   (a) What is the joint density of X and Y?

   (b) What is the probability that both arrive between 8:15 PM and 8:45 PM?

   (c) If both are willing to wait no more than 10 minutes for the other after they arrive, what is the probability that they miss each other?

Problem 1

In a certain human population, the Cephalic Index I (skull width expressed as a percentage of its length) is normally distributed among individuals. There is 58% with $I \leq 75$, 38% with $75 \leq I \leq 80$, and 4% with $I > 80$. Find the density function of the index, $f_I(i)$, and the probability $P(78 \leq I \leq 82)$.

Solution

The Cephalic Index I follows a Normal distribution: $I \sim N(\mu, \sigma^2)$. We need to find the mean (μ) and the standard deviation (σ). Let $Z = \frac{I-\mu}{\sigma}$ be the standard Normal variable $Z \sim N(0,1)$, with CDF $\Phi(z)$.

1. Finding μ and σ: We use the given cumulative probabilities:

1. $P(I \leq 75) = 0.58 \implies \Phi\left(\frac{75-\mu}{\sigma}\right) = 0.58$
2. $P(I \leq 80) = P(I \leq 75) + P(75 \leq I \leq 80) = 0.58 + 0.38 = 0.96$

   Using the standard Normal table (or calculator):

   - $\Phi(z_1) = 0.58 \implies z_1 \approx 0.20$
   - $\Phi(z_2) = 0.96 \implies z_2 \approx 1.75$

   This yields a system of linear equations:

$$
\begin{aligned}
(1) \quad & 75 - \mu = 0.20\sigma \\
(2) \quad & 80 - \mu = 1.75\sigma
\end{aligned}
$$

Subtracting equation (1) from equation (2):

$$(80-\mu) - (75-\mu) = 1.75\sigma - 0.20\sigma \implies 5 = 1.55\sigma$$

$$\sigma = \frac{5}{1.55} \approx 3.226$$

Substituting σ back into equation (1):

$$75 - \mu = 0.20(3.226) \implies 75 - \mu \approx 0.645$$

$$\mu = 75 - 0.645 \approx 74.355$$

2. The Density Function $f_I(i)$: The PDF of I is:

$$f_I(i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{i-\mu}{\sigma}\right)^2}$$

$$f_I(i) = \frac{1}{\sqrt{2\pi}(3.226)} e^{-\frac{1}{2}\left(\frac{i-74.355}{3.226}\right)^2}$$

3. Calculating $P(78 \leq I \leq 82)$: We standardize the values:

$$Z_{78} = \frac{78 - 74.355}{3.226} \approx \frac{3.645}{3.226} \approx 1.13$$

$$Z_{82} = \frac{82 - 74.355}{3.226} \approx \frac{7.645}{3.226} \approx 2.37$$

$$P(78 \leq I \leq 82) = P(1.13 \leq Z \leq 2.37) = \Phi(2.37) - \Phi(1.13)$$

Using the Z-table values:

$$P(78 \leq I \leq 82) \approx 0.9911 - 0.8708 = 0.1203$$

Problem 2

Alice and Peter agreed to meet at 8 PM to go to the movies. Since they are not punctual, their arrival times X and Y can be assumed to be independent random variables with a uniform distribution between 8 PM and 9 PM.

1. What is the joint density of X and Y?
2. What is the probability that both arrive between 8:15 PM and 8:45 PM?
3. If both are willing to wait no more than 10 minutes for the other after they arrive, what is the probability that they miss each other?

Solution

Let X and Y be the arrival times in minutes after 8:00 PM. Thus, $X, Y \sim U(0, 60)$. The individual PDF is $f_X(x) = f_Y(y) = \frac{1}{60}$ for $x, y \in [0, 60]$, and 0 otherwise.

a) What is the joint density of X and Y? Since X and Y are independent, the joint density is the product of the marginal densities:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = \frac{1}{60} \cdot \frac{1}{60} = \frac{1}{3600}$$

for $0 \le x \le 60$ and $0 \le y \le 60$. The density is 0 otherwise. The sample space is a square of area $60 \times 60 = 3600$.

b) What is the probability that both arrive between 8:15 PM and 8:45 PM? In minutes after 8:00 PM, this corresponds to $15 \le X \le 45$ and $15 \le Y \le 45$. The region of integration R is a square defined by $R = [15, 45] \times [15, 45]$. The area of this region is $A(R) = (45 - 15) \times (45 - 15) = 30 \times 30 = 900$. The probability is:

$$P(15 \le X \le 45, 15 \le Y \le 45) = \iint_R f_{X,Y}(x, y)\, dxdy = \frac{1}{3600} \cdot A(R) = \frac{900}{3600} = \frac{1}{4} = 0.25$$

c) If both are willing to wait no more than 10 minutes for the other after arriving, what is the probability that they miss each other? They miss each other if the difference between their arrival times is greater than 10 minutes. We are looking for $P(\text{Miss}) = P(|X - Y| > 10)$.

It is easier to calculate the probability that they **meet** ($P(\text{Meet})$), which is the complement:

$$P(\text{Meet}) = P(|X - Y| \le 10) \implies P(\text{Meet}) = P(-10 \le X - Y \le 10)$$

The region where they meet is $M = \{(x, y) : Y - 10 \le X \le Y + 10\}$ within the 60×60 square.

The probability is given by $\frac{\text{Area}(M)}{\text{Total Area}}$. The total area is 3600. The area where they **miss** ($|X - Y| > 10$) consists of two triangles, T_1 and T_2, in the corners of the square.

- T_1: $X - Y > 10$ (or $Y < X - 10$). Vertices are $(10, 0), (60, 0), (60, 50)$. Base $= 50$, Height $= 50$. $\text{Area}(T_1) = \frac{1}{2}(50^2) = 1250$.
- T_2: $Y - X > 10$ (or $Y > X + 10$). Vertices are $(0, 10), (0, 60), (50, 60)$. Base $= 50$, Height $= 50$. $\text{Area}(T_2) = \frac{1}{2}(50^2) = 1250$.

The total area where they miss is $A(\text{Miss}) = 1250 + 1250 = 2500$.

The probability that they miss each other is:

$$P(\text{Miss}) = \frac{A(\text{Miss})}{\text{Total Area}} = \frac{2500}{3600} = \frac{25}{36} \approx 0.6944$$

Probability Theory

Tutorial 7

1. Maria has three baby children: A (Ana), B (Ben), and C (Carl). She puts them to sleep at night. Because she sings beautifully, all three fall asleep as soon as she finishes the song. The number of consecutive hours each baby sleeps is a discrete random variable that takes the values 6 and 8 with probability 1/2 each. We assume that these three random variables are independent of each other. Find the distribution of the number of consecutive hours the mother sleeps, if she falls asleep at the same moment as her three children (she is very tired) and wakes up when the first of her three children wakes up.

2. A fair six-sided die is rolled. Let $S = \{1, 2, 3, 4, 5, 6\}$ be the sample space. Define two random variables X and Y based on the outcome of the roll:

$$X = \begin{cases} 1 & \text{if the outcome is less than 4} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad Y = \begin{cases} 1 & \text{if the outcome is 1} \\ 0 & \text{otherwise} \end{cases}$$

   (a) Calculate the joint probability mass function of (X, Y).

   (b) Calculate the marginal probability mass functions of X and Y.

   (c) Let $A = \{Y = 1\}$, $B = \{X = 1\}$ and $C = \{X = 0\}$. Calculate $P(A|B)$ and $P(A|C)$.

3. The alcohol percentage in a one-liter bottle of a certain beverage is a continuous random variable X with density:

$$f_X(x) = c(1 - x^2)\mathbf{1}_{(0,1)}(x)$$

   where $\mathbf{1}_{(0,1)}(x)$ is the indicator function for the interval $(0, 1)$.

   (a) What value must c take for f_X to truly be a probability density function?

   (b) Find the cumulative distribution function (CDF) of X.

   (c) If the alcohol content of the bottles is to be divided into two equally probable categories ("low" and "high"), what is the threshold value that defines the median

   (d) If a case of these beverages contains ten one-liter bottles, assuming the bottles are independent, calculate the probability that in a case there are at least 2 bottles with an alcohol percentage lower than 50%.

Problem 1

Maria has three baby children: A (Ana), B (Ben), and C (Carl). She puts them to sleep at night. Because she sings beautifully, all three fall asleep as soon as she finishes the song. The number of consecutive hours each baby sleeps is a discrete random variable that takes the values 6 and 8 with probability 1/2 each. We assume that these three random variables are independent of each other. Find the distribution of the number of consecutive hours the mother sleeps, if she falls asleep at the same moment as her three children (she is very tired) and wakes up when the first of her three children wakes up.

Solution

Let A, B, and C be the number of consecutive hours Ana, Ben, and Carl sleep, respectively. The PMF for each child is identical and independent:

$$P(A = a) = P(B = b) = P(C = c) = \begin{cases} 1/2 & \text{if } a, b, c \in \{6, 8\} \\ 0 & \text{otherwise} \end{cases}$$

Let M be the number of consecutive hours the mother sleeps. Since the mother wakes up when the first child wakes up, M is the minimum of the three sleeping times:

$$M = \min(A, B, C)$$

Since A, B, C can only take values 6 or 8, M can only take the values 6 or 8.

Calculating $P(M = 8)$ The mother sleeps for 8 hours if and only if all three children sleep for 8 hours.

$$M = 8 \iff A = 8 \text{ and } B = 8 \text{ and } C = 8$$

Due to the independence of A, B, and C:

$$\begin{aligned} P(M = 8) &= P(A = 8, B = 8, C = 8) \\ &= P(A = 8)P(B = 8)P(C = 8) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) \\ &= \frac{1}{8} \end{aligned}$$

Calculating $P(M = 6)$ Since M is a discrete variable that only takes values 6 or 8, $P(M = 6)$ is the complement of $P(M = 8)$:

$$P(M = 6) = 1 - P(M = 8) = 1 - \frac{1}{8} = \frac{7}{8}$$

The distribution of the number of consecutive hours the mother sleeps is:

$$P(M = m) = \begin{cases} 7/8 & \text{if } m = 6 \\ 1/8 & \text{if } m = 8 \\ 0 & \text{otherwise} \end{cases}$$

Probability Ana sleeps 8 hours: $P(A = 8) = \frac{1}{2}$ Probability the Mother sleeps 8 hours: $P(M = 8) = \frac{1}{8}$. Since $1/2 > 1/8$, the probability that Ana sleeps 8 consecutive hours is significantly higher than the probability that the mother sleeps 8 consecutive hours. The mother's sleep time is severely restricted by the minimum sleep time among her three children.

In this example, we see that what changes the distribution of the minimum is, of course, that the minimum is systematically smaller than the rest of the random variables. That is why, if we compare the number of hours Maria sleeps, we see that the probability of her sleeping 8 consecutive hours is much lower than the probability of Ana doing so. The same happens in the rest of the examples we saw, but this one is very simple, so we can clearly see what happens. In fact, we can make the following construction.

Let us make the following construction of random variables based on a uniform variable $U \sim U(0,1)$. Let:

$$X_1 = 6 \cdot I_{(0,1/2)}(U) + 8 \cdot I_{[1/2,1)}(U)$$
$$X_2 = 6 \cdot I_{(0,1/4)\cup[1/2,3/4)}(U) + 8 \cdot I_{[1/4,1/2)\cup[3/4,1)}(U)$$
$$X_3 = 6 \cdot I_{(0,1/8)\cup[2/8,3/8)\cup[4/8,5/8)\cup[6/8,7/8)}(U) + 8 \cdot I_{[1/8,2/8)\cup[3/8,4/8)\cup[5/8,6/8)\cup[7/8,1)}(U)$$

1. Verify that X_1, X_2, X_3 are random variables that have the same distribution as the sleeping hours from the previous Example.

2. Prove that they are independent random variables. (Hint: Try to perform the minimum amount of calculations possible to demonstrate independence).

3. Graph them on the same coordinate axis, and graph their minimum, $Y = \min\{X_1, X_2, X_3\}$, the number of consecutive sleeping hours of the mother.

We can see that that for each X_i, $P(X_i = 6) = 1/2$ and $P(X_i = 8) = 1/2$. Since $U \sim U(0,1)$, the probability of U falling into any set $S \subset (0,1)$ is equal to the length (Lebesgue measure) of S, denoted $|S|$. For example

- $P(X_1 = 6) = P(U \in (0,1/2)) = |(0,1/2)| = 1/2$.
- $P(X_1 = 8) = P(U \in [1/2,1)) = |[1/2,1)| = 1/2$.

Since X_i are discrete variables taking values in $\{6,8\}$, the variables are independent if and only if for all combinations $(x_1,x_2,x_3) \in \{6,8\}^3$

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = P(X_1 = x_1)P(X_2 = x_2)P(X_3 = x_3)$$

Given the symmetry, it is sufficient to prove independence for the event $X_i = 8$. Let $S_i = \{u \in (0,1) : X_i(u) = 8\}$. We left this as an exercise.

For these variables, the minimum function $Y(u) = \min\{X_1(u), X_2(u), X_3(u)\}$ will be equal to 8 only when $X_1(u) = 8$ AND $X_2(u) = 8$ AND $X_3(u) = 8$. This occurs on the interval $S_1 \cap S_2 \cap S_3 = [7/8, 1)$.

$$Y(u) = \begin{cases} 8 & \text{if } u \in [7/8,1) \\ 6 & \text{if } u \in (0,7/8) \end{cases}$$

This confirms that $P(Y = 8) = 1/8$, consistent with the calculation in the previous problem.

Problem 2

A fair six-sided die is rolled. Let $S = \{1,2,3,4,5,6\}$ be the sample space. Define two random variables X and Y based on the outcome of the roll:

$$X = \begin{cases} 1 & \text{if the outcome is less than 4} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad Y = \begin{cases} 1 & \text{if the outcome is 1} \\ 0 & \text{otherwise} \end{cases}$$

1. Calculate the joint probability mass function of (X, Y).

2. Calculate the marginal probability mass functions of X and Y.

3. Let $A = \{Y = 1\}$, $B = \{X = 1\}$ and $C = \{X = 0\}$. Calculate $P(A|B)$ and $P(A|C)$.

Solution

Variable Definitions and Events

The possible outcomes and their probabilities are $P(\text{outcome} = k) = 1/6$ for $k \in \{1, \ldots, 6\}$.

The events corresponding to X and Y are:

- $X = 1$: outcome is $\{1, 2, 3\}$. (3 outcomes)
- $X = 0$: outcome is $\{4, 5, 6\}$. (3 outcomes)
- $Y = 1$: outcome is $\{1\}$. (1 outcome)
- $Y = 0$: outcome is $\{2, 3, 4, 5, 6\}$. (5 outcomes)

The range of (X, Y) is the set of possible pairs: $\mathcal{S}_{X,Y} = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

We calculate the probability for each pair:

1. $\mathbf{p_{X,Y}(1, 1) = P(X = 1, Y = 1)}$: Outcome is < 4 (i.e., $\{1, 2, 3\}$) AND outcome is 1 (i.e., $\{1\}$).

$$P(X = 1, Y = 1) = P(\text{outcome} = 1) = \mathbf{1/6}$$

2. $\mathbf{p_{X,Y}(1, 0) = P(X = 1, Y = 0)}$: Outcome is < 4 (i.e., $\{1, 2, 3\}$) AND outcome is $\neq 1$ (i.e., $\{2, 3, 4, 5, 6\}$).

$$P(X = 1, Y = 0) = P(\text{outcome} \in \{2, 3\}) = 1/6 + 1/6 = \mathbf{2/6}$$

3. $\mathbf{p_{X,Y}(0, 1) = P(X = 0, Y = 1)}$: Outcome is ≥ 4 (i.e., $\{4, 5, 6\}$) AND outcome is 1 (i.e., $\{1\}$).

$$P(X = 0, Y = 1) = P(\emptyset) = \mathbf{0}$$

4. $\mathbf{p_{X,Y}(0, 0) = P(X = 0, Y = 0)}$: Outcome is ≥ 4 (i.e., $\{4, 5, 6\}$) AND outcome is $\neq 1$ (i.e., $\{2, 3, 4, 5, 6\}$).

$$P(X = 0, Y = 0) = P(\text{outcome} \in \{4, 5, 6\}) = 1/6 + 1/6 + 1/6 = \mathbf{3/6}$$

The range is $\mathcal{S}_{X,Y} = \{(0, 0), (1, 0), (1, 1)\}$. Note that $P(0, 1) = 0$, so $(0, 1)$ is not a possible value pair, though it is sometimes included in the domain. A strict range would exclude it.

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

- $\mathbf{p_X(1)}$: $P(X = 1) = p_{X,Y}(1, 0) + p_{X,Y}(1, 1) = 2/6 + 1/6 = \mathbf{3/6} = \mathbf{1/2}$
- $\mathbf{p_X(0)}$: $P(X = 0) = p_{X,Y}(0, 0) + p_{X,Y}(0, 1) = 3/6 + 0 = \mathbf{3/6} = \mathbf{1/2}$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

- $\mathbf{p_Y(1)}$: $P(Y = 1) = p_{X,Y}(0, 1) + p_{X,Y}(1, 1) = 0 + 1/6 = \mathbf{1/6}$
- $\mathbf{p_Y(0)}$: $P(Y = 0) = p_{X,Y}(0, 0) + p_{X,Y}(1, 0) = 3/6 + 2/6 = \mathbf{5/6}$

We use the definition $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

1. **Calculate** $P(A|B) = P(Y = 1|X = 1)$:
   - $A \cap B = \{Y = 1 \cap X = 1\}$. From section 1(a), $P(A \cap B) = P(X = 1, Y = 1) = 1/6$.

- $P(B) = P(X = 1)$. From section 2, $P(B) = 3/6$.

$$P(Y = 1|X = 1) = \frac{P(X = 1, Y = 1)}{P(X = 1)} = \frac{1/6}{3/6} = \mathbf{1/3}$$

2. **Calculate** $P(A|C) = P(Y = 1|X = 0)$:

- $A \cap C = \{Y = 1 \cap X = 0\}$. From section 1(c), $P(A \cap C) = P(X = 0, Y = 1) = 0$.
- $P(C) = P(X = 0)$. From section 2, $P(C) = 3/6$.

$$P(Y = 1|X = 0) = \frac{P(X = 0, Y = 1)}{P(X = 0)} = \frac{0}{3/6} = \mathbf{0}$$

Problem 3

The alcohol percentage in a one-liter bottle of a certain beverage is a continuous random variable X with density:

$$f_X(x) = c(1 - x^2)\mathbf{1}_{(0,1)}(x)$$

where $\mathbf{1}_{(0,1)}(x)$ is the indicator function for the interval $(0, 1)$.

1. What value must c take for f_X to truly be a probability density function?
2. Find the cumulative distribution function (CDF) of X.
3. If the alcohol content of the bottles is to be divided into two equally probable categories ("low" and "high"), what is the threshold value that defines the median
4. If a case of these beverages contains ten one-liter bottles, assuming the bottles are independent, calculate the probability that in a case there are at least 2 bottles with an alcohol percentage lower than 50%.

Solution

Finding the Constant c

For $f_X(x)$ to be a valid PDF, its integral over its support must equal 1:

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 c(1 - x^2)dx = 1$$

$$c\left[x - \frac{x^3}{3}\right]_0^1 = 1$$

$$c\left(1 - \frac{1}{3}\right) = 1 \quad \Rightarrow \quad c\left(\frac{2}{3}\right) = 1$$

Result: $c = \mathbf{3/2}$. The PDF is $f_X(x) = \frac{3}{2}(1 - x^2)$ for $0 < x < 1$.

Cumulative Distribution Function (CDF)

For $0 \le x \le 1$:

$$F_X(x) = \int_0^x f_X(t)dt = \int_0^x \frac{3}{2}(1 - t^2)dt = \frac{3}{2}\left[t - \frac{t^3}{3}\right]_0^x$$

$$\mathbf{F_X(x) = \frac{3x}{2} - \frac{x^3}{2} \quad \text{for } 0 \le x \le 1}$$

And $F_X(x) = 0$ for $x < 0$, $F_X(x) = 1$ for $x > 1$.

Threshold (Median)

The threshold m (median) must satisfy ¶$(X \le m) = F_X(m) = 0.5$.

$$\frac{3m}{2} - \frac{m^3}{2} = \frac{1}{2}$$

The problem simplifies to solving the cubic equation:

$$m^3 - 3m + 1 = 0 \quad \text{for } m \in (0, 1)$$

Binomial Distribution

The criterion is $P_p = \P(X < 0.50) = F_X(0.5)$.

$$P_p = F_X(1/2) = \frac{3(1/2)}{2} - \frac{(1/2)^3}{2} = \frac{3}{4} - \frac{1/8}{2} = \frac{3}{4} - \frac{1}{16} = \frac{11}{16}$$

Let Z be the number of bottles with $X < 0.50$. Since the bottles are independent, $Z \sim \text{Binomial}(n = 10, p = 11/16)$. The probability of failure is $q = 1 - p = 5/16$.

1) Probability of at least 2 bottles with $X < 0.50$

$$\P(Z \geq 2) = 1 - \P(Z < 2) = 1 - [\P(Z = 0) + \P(Z = 1)]$$

$$\P(Z = 0) = \binom{10}{0} p^0 q^{10} = \left(\frac{5}{16}\right)^{10}$$

$$\P(Z = 1) = \binom{10}{1} p^1 q^9 = 10 \left(\frac{11}{16}\right) \left(\frac{5}{16}\right)^9$$

$$\P(Z \geq 2) = 1 - \left[\left(\frac{5}{16}\right)^{10} + 10 \left(\frac{11}{16}\right) \left(\frac{5}{16}\right)^9\right]$$

Probability Theory

Tutorial 8

1. Let X be a random variable following the Binomial distribution, $X \sim \text{Binomial}(n, p)$, where n is the number of trials and p is the success probability. The Probability Mass Function (PMF) is given by:
$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \ldots, n$$
The expected value $E[X] = np$.

2. Let Y be a random variable following the Geometric distribution, $Y \sim \text{Geometric}(p)$, representing the number of trials k needed to observe the first success. The PMF is given by:
$$P(Y = k) = (1-p)^{k-1} p, \quad \text{for } k = 1, 2, 3, \ldots$$
The expected value $E[Y] = 1/p$.

3. Let X be a continuous random variable following the Gamma distribution, $X \sim \text{Gamma}(\alpha, \lambda)$, with shape parameter $\alpha > 0$ and rate parameter $\lambda > 0$. The Probability Density Function (PDF) is given by:
$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad \text{for } x > 0$$
The expected value $E[X] = \alpha/\lambda$.

4. A segment of unit length is cut at two points chosen independently and uniformly at random. We want to find the expected length of the segment **between the two cuts**.

Problem 1

Let X be a random variable following the Binomial distribution, $X \sim \text{Binomial}(n, p)$, where n is the number of trials and p is the success probability. The Probability Mass Function (PMF) is given by:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \dots, n$$

The expected value $E[X] = np$.

Solution

The summation begins from $k = 1$ since the term for $k = 0$ is zero $(0 \cdot P(X = 0) = 0)$.

$$E[X] = \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Using the identity $k\frac{n!}{k!} = n\frac{(n-1)!}{(k-1)!} = n\binom{n-1}{k-1}$:

$$E[X] = \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

Factor out n and one factor of p, noting that $p^k = p \cdot p^{k-1}$ and $n - k = (n-1) - (k-1)$:

$$E[X] = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

Let $j = k - 1$. The summation limits become $j = 0$ to $j = n - 1$:

$$E[X] = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j}$$

The summation represents the sum of all probabilities for a $\text{Binomial}(n-1, p)$ distribution, which must equal 1 by the law of total probability.

$$\mathbf{E[X] = np}$$

Problem 2

Let Y be a random variable following the Geometric distribution, $Y \sim \text{Geometric}(p)$, representing the number of trials k needed to observe the first success. The PMF is given by:

$$P(Y = k) = (1-p)^{k-1}p, \quad \text{for } k = 1, 2, 3, \ldots$$

The expected value $E[Y] = 1/p$.

Solution

Let $q = 1 - p$. The expectation is $E[Y] = p \sum_{k=1}^{\infty} kq^{k-1}$. We use the identity derived from differentiating the geometric series $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$:

$$\frac{d}{dq}\left(\sum_{k=0}^{\infty} q^k\right) = \sum_{k=1}^{\infty} kq^{k-1} = \frac{d}{dq}\left((1-q)^{-1}\right) = (-1)(1-q)^{-2}(-1) = \frac{1}{(1-q)^2}$$

Substituting this result back into $E[Y]$:

$$E[Y] = p \cdot \frac{1}{(1-q)^2}$$

Since $q = 1 - p$, we have $1 - q = p$:

$$E[Y] = p \cdot \frac{1}{p^2}$$

Thus:

$$\mathbf{E[Y] = \frac{1}{p}}$$

Problem 3

Let X be a continuous random variable following the Gamma distribution, $X \sim \text{Gamma}(\alpha, \lambda)$, with shape parameter $\alpha > 0$ and rate parameter $\lambda > 0$. The Probability Density Function (PDF) is given by:

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad \text{for } x > 0$$

The expected value $E[X] = \alpha/\lambda$.

Solution

Substitute the PDF into the expectation integral:

$$E[X] = \int_0^\infty x \left(\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \right) dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x} dx$$

We perform a change of variable on the integral, letting $t = \lambda x$. Then $x = t/\lambda$ and $dx = dt/\lambda$.

$$\int_0^\infty x^\alpha e^{-\lambda x} dx = \int_0^\infty \left(\frac{t}{\lambda} \right)^\alpha e^{-t} \frac{dt}{\lambda} = \frac{1}{\lambda^{\alpha+1}} \int_0^\infty t^\alpha e^{-t} dt$$

The integral $\int_0^\infty t^\alpha e^{-t} dt$ is the definition of $\Gamma(\alpha + 1)$.

$$E[X] = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{1}{\lambda^{\alpha+1}} \Gamma(\alpha + 1)$$

Simplify the exponential coefficients: $\frac{\lambda^\alpha}{\lambda^{\alpha+1}} = \frac{1}{\lambda}$. Apply the property $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$:

$$E[X] = \frac{1}{\lambda} \cdot \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)}$$

Canceling $\Gamma(\alpha)$ terms, the final result is:

$$\mathbf{E}[\mathbf{X}] = \frac{\alpha}{\lambda}$$

Problem 4

A segment of unit length is cut at two points chosen independently and uniformly at random. We want to find the expected length of the segment **between the two cuts**.

Solution

Let U_1 and U_2 be the positions of the two cut points, where $U_1 \sim U(0, 1)$ and $U_2 \sim U(0, 1)$ are independent random variables. The length of the segment between the two cuts is $|U_1 - U_2|$.

The expected value $E[|U_1 - U_2|]$ is calculated using the joint probability density function (PDF) $f_{U_1U_2}(u_1, u_2)$:

$$f_{U_1U_2}(u_1, u_2) = 1 \cdot I_{[0,1]}(u_1)I_{[0,1]}(u_2)$$

The expectation integral is:

$$E[|U_1 - U_2|] = \int_0^1 \int_0^1 |u_1 - u_2| f_{U_1U_2}(u_1, u_2) du_1 du_2 = \int_0^1 \int_0^1 |u_1 - u_2| du_1 du_2$$

We split the integral over the unit square $[0, 1] \times [0, 1]$ into two regions defined by the diagonal line $u_1 = u_2$:

1. Region R_1: $u_2 \leq u_1$, where $|u_1 - u_2| = u_1 - u_2$.
2. Region R_2: $u_1 > u_2$, where $|u_1 - u_2| = u_2 - u_1$.

Due to symmetry, the integral over R_1 is equal to the integral over R_2. Thus, we can calculate the integral over R_1 and multiply the result by 2:

$$E[|U_1 - U_2|] = 2 \iint_{R_1} (u_1 - u_2) du_2 du_1$$

We define the limits for Region R_1: $0 \leq u_2 \leq u_1$ and $0 \leq u_1 \leq 1$.

$$E[|U_1 - U_2|] = 2 \int_0^1 \left[\int_0^{u_1} (u_1 - u_2) du_2 \right] du_1$$

First, solve the inner integral with respect to u_2:

$$\int_0^{u_1} (u_1 - u_2) du_2 = \left[u_1 u_2 - \frac{u_2^2}{2} \right]_0^{u_1} = \left(u_1^2 - \frac{u_1^2}{2} \right) - (0) = \frac{u_1^2}{2}$$

Now, substitute this result into the outer integral and solve with respect to u_1:

$$E[|U_1 - U_2|] = 2 \int_0^1 \frac{u_1^2}{2} du_1 = \int_0^1 u_1^2 du_1$$

$$E[|U_1 - U_2|] = \left[\frac{u_1^3}{3} \right]_0^1 = \frac{1^3}{3} - \frac{0^3}{3}$$

The expected length of the segment between the two cuts is:

$$\mathbf{E[|U_1 - U_2|] = \frac{1}{3}}$$

Probability Theory

Tutorial 9

1. Let X be a continuous random variable with an Exponential distribution of parameter 1: $X \sim \mathrm{Exp}\,(1)$. Let $Y = \min\,(X, 3)$.

   (a) Calculate the expected value of Y, $E[Y]$.

   (b) Calculate $E[Xh(Y)]$, where $h(Y) = Y^2$.

2. A particle is initially located at the origin of the plane. At each time step, it takes a jump of length one in random, mutually independent directions. Let θ_i be i.i.d. random variables with $\theta_i \sim \mathrm{U}\,(0, 2\pi)$.

   Let $(X_i, Y_i) = (\cos(\theta_i), \sin(\theta_i))$ be the jump taken at the i-th time step.

   (a) Determine the position of the particle in the plane after n steps in terms of the defined variables.

   (b) Calculate the expected value of the squared distance from the origin to the particle's position after n steps.

   (c) Calculate the covariance between the coordinates that indicate the particle's position after n steps.

Problem 1

Let X be a continuous random variable with an Exponential distribution of parameter 1: $X \sim \text{Exp}(1)$. Let $Y = \min(X, 3)$.

1. Calculate the expected value of Y, $E[Y]$.
2. Calculate $E[Xh(Y)]$, where $h(Y) = Y^2$.

Solution 2

The PDF and CDF of $X \sim \text{Exp}(1)$ are:

$$f_X(x) = e^{-x} \quad \text{for } x > 0$$

$$F_X(x) = 1 - e^{-x} \quad \text{for } x > 0$$

The variable Y is defined as $Y = \min(X, 3)$. This means $Y = X$ if $X < 3$, and $Y = 3$ if $X \geq 3$.

Part a: Calculate $E[Y]$

The expected value of Y is calculated by splitting the integral over the two cases defining Y:

$$E[Y] = \int_0^\infty y f_Y(y) dy = \int_0^3 x f_X(x) dx + \int_3^\infty 3 f_X(x) dx$$

$$E[Y] = \underbrace{\int_0^3 x e^{-x} dx}_{\text{Case } Y=X} + \underbrace{3 \int_3^\infty e^{-x} dx}_{\text{Case } Y=3}$$

Term 2 (Right-hand integral):

$$3 \int_3^\infty e^{-x} dx = 3 \left[-e^{-x}\right]_3^\infty = 3(0 - (-e^{-3})) = 3e^{-3}$$

Term 1 (Left-hand integral): We use integration by parts, $\int u dv = uv - \int v du$, with $u = x$ ($du = dx$) and $dv = e^{-x} dx$ ($v = -e^{-x}$):

$$\int_0^3 x e^{-x} dx = \left[-x e^{-x}\right]_0^3 - \int_0^3 (-e^{-x}) dx$$

$$= (-3e^{-3} - 0) + \int_0^3 e^{-x} dx$$

$$= -3e^{-3} + \left[-e^{-x}\right]_0^3$$

$$= -3e^{-3} + (-e^{-3} - (-e^0))$$

$$= -3e^{-3} - e^{-3} + 1 = 1 - 4e^{-3}$$

Final Result for $E[Y]$:

$$E[Y] = (1 - 4e^{-3}) + 3e^{-3} = 1 - e^{-3}$$

$E[Y] = 1 - e^{-3}$.

Part b: Calculate $E[Xh(Y)]$ where $h(Y) = Y^2$

We need to calculate $E[XY^2]$. Again, we split the integral based on the definition of Y:

$$E[XY^2] = \int_0^3 X \cdot Y^2 f_X(x)dx + \int_3^\infty X \cdot Y^2 f_X(x)dx$$

- **Case 1 $(0 < X < 3)$:** $Y = X$, so $XY^2 = X^3$.
- **Case 2 $(X \geq 3)$:** $Y = 3$, so $XY^2 = X(3^2) = 9X$.

$$E[XY^2] = \underbrace{\int_0^3 x^3 e^{-x} dx}_{\text{Case } Y=X} + \underbrace{\int_3^\infty 9xe^{-x} dx}_{\text{Case } Y=3}$$

Term 2 (Right-hand integral):

$$9\int_3^\infty xe^{-x}dx$$

We use integration by parts, $\int xe^{-x}dx = -xe^{-x} - e^{-x}$.

$$9\left[-xe^{-x} - e^{-x}\right]_3^\infty = 9\left[(-0-0) - (-3e^{-3} - e^{-3})\right] = 9(4e^{-3}) = 36e^{-3}$$

Term 1 (Left-hand integral):

$$\int_0^3 x^3 e^{-x}dx$$

This requires repeated integration by parts, or using the definition of the incomplete Gamma function. We integrate: $\int x^3 e^{-x}dx = -e^{-x}(x^3 + 3x^2 + 6x + 6)$.

$$\left[-e^{-x}(x^3 + 3x^2 + 6x + 6)\right]_0^3$$
$$= \left(-e^{-3}(3^3 + 3(3^2) + 6(3) + 6)\right) - \left(-e^{-0}(0 + 0 + 0 + 6)\right)$$
$$= -e^{-3}(27 + 27 + 18 + 6) - (-6)$$
$$= 6 - 78e^{-3}$$

Final Result for $E[XY^2]$:

$$E[XY^2] = (6 - 78e^{-3}) + 36e^{-3} = 6 - 42e^{-3}$$

$E[Xh(Y)] = E[XY^2] = 6 - 42e^{-3}$.

Problem 2

A particle is initially located at the origin of the plane. At each time step, it takes a jump of length one in random, mutually independent directions. Let θ_i be i.i.d. random variables with $\theta_i \sim \mathrm{U}(0, 2\pi)$.
Let $(X_i, Y_i) = (\cos(\theta_i), \sin(\theta_i))$ be the jump taken at the i-th time step.

1. Determine the position of the particle in the plane after n steps in terms of the defined variables.
2. Calculate the expected value of the squared distance from the origin to the particle's position after n steps.
3. Calculate the covariance between the coordinates that indicate the particle's position after n steps.

Solution

Position after n steps

The particle starts at the origin $(0, 0)$. The position vector after n steps, denoted by $\mathbf{S}_n = (S_{n,x}, S_{n,y})$, is the sum of the n individual jump vectors (X_i, Y_i):

$$\mathbf{S}_n = \sum_{i=1}^{n} (X_i, Y_i)$$

The coordinates are:

$$S_{n,x} = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \cos(\theta_i)$$

$$S_{n,y} = \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \sin(\theta_i)$$

Expected squared distance from the origin $E[D_n^2]$

The squared distance from the origin after n steps is $D_n^2 = S_{n,x}^2 + S_{n,y}^2$. We need to calculate $E[D_n^2]$.
By the linearity of expectation:

$$E[D_n^2] = E[S_{n,x}^2] + E[S_{n,y}^2]$$

We first calculate the expected value for a single step i:

$$E[X_i] = E[\cos(\theta_i)] = \int_0^{2\pi} \cos(\theta) \frac{1}{2\pi} d\theta = 0$$

$$E[Y_i] = E[\sin(\theta_i)] = \int_0^{2\pi} \sin(\theta) \frac{1}{2\pi} d\theta = 0$$

Since $E[X_i] = 0$ and $E[Y_i] = 0$, the variance of a single step is $E[X_i^2]$ and $E[Y_i^2]$.

$$E[X_i^2] = E[\cos^2(\theta_i)] = \int_0^{2\pi} \cos^2(\theta) \frac{1}{2\pi} d\theta = \frac{1}{2\pi} \left[\frac{\theta}{2} + \frac{\sin(2\theta)}{4} \right]_0^{2\pi} = \frac{1}{2\pi}(\pi) = \frac{1}{2}$$

Similarly, $E[Y_i^2] = 1/2$.
The variance of the sum of independent variables is the sum of their variances:

$$E[S_{n,x}^2] = \mathrm{Var}(S_{n,x}) + (E[S_{n,x}])^2$$

Since $E[X_i] = 0$, $E[S_{n,x}] = \sum E[X_i] = 0$.

$$E[S_{n,x}^2] = \text{Var}(S_{n,x}) = \sum_{i=1}^{n} \text{Var}(X_i) = \sum_{i=1}^{n}(E[X_i^2] - (E[X_i])^2) = \sum_{i=1}^{n} \left(\frac{1}{2} - 0\right) = \frac{n}{2}$$

Similarly, $E[S_{n,y}^2] = n/2$.

Finally, the expected squared distance is:

$$E[D_n^2] = E[S_{n,x}^2] + E[S_{n,y}^2] = \frac{n}{2} + \frac{n}{2} = n$$

The expected value of the squared distance from the origin after n steps is n.

Covariance between coordinates $Cov(S_{n,x}, S_{n,y})$

We use the property that the covariance of sums of independent variables simplifies greatly:

$$\text{Cov}(S_{n,x}, S_{n,y}) = \text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \text{Cov}(X_i, Y_j)$$

Since the steps are independent, $\text{Cov}(X_i, Y_j) = 0$ for $i \neq j$. Thus, only the terms where $i = j$ remain:

$$\text{Cov}(S_{n,x}, S_{n,y}) = \sum_{i=1}^{n} \text{Cov}(X_i, Y_i)$$

We calculate the covariance for a single step i:

$$\text{Cov}(X_i, Y_i) = E[X_i Y_i] - E[X_i]E[Y_i]$$

Since $E[X_i] = 0$ and $E[Y_i] = 0$, the second term is zero. We only need $E[X_i Y_i]$:

$$E[X_i Y_i] = E[\cos(\theta_i)\sin(\theta_i)] = \frac{1}{2}E[\sin(2\theta_i)]$$

$$E[\sin(2\theta_i)] = \int_0^{2\pi} \sin(2\theta)\frac{1}{2\pi}d\theta = \frac{1}{2\pi}\left[-\frac{\cos(2\theta)}{2}\right]_0^{2\pi} = \frac{1}{4\pi}(-\cos(4\pi) + \cos(0)) = \frac{1}{4\pi}(-1+1) = 0$$

Since $E[X_i Y_i] = 0$, $\text{Cov}(X_i, Y_i) = 0$.

Therefore, the covariance between the coordinates after n steps is:

$$\text{Cov}(S_{n,x}, S_{n,y}) = \sum_{i=1}^{n} 0 = 0$$

The covariance between the final coordinates $S_{n,x}$ and $S_{n,y}$ is 0.

Probability Theory

Tutorial 10

1. Let X be a random variable with mean μ and standard deviation σ. Let

$$P(|X - \mu| < k\sigma) = P(\mu - k\sigma < X < \mu + k\sigma)$$

   (a) **Calculate the Bound (Chebyshev):** Determine the lower probability bound for the following values of k: $k = 1.5$, $k = 2$, and $k = 3$.

   (b) **Calculate the Exact Probability (Normal):** Assume the random variable X follows a Normal distribution. Calculate the exact probability $P(\mu - k\sigma < X < \mu + k\sigma)$ for the same values of k.

   (c) **Analysis:** Compare the results in a table and explain the difference between the Chebyshev bound and the exact probability.

2. Find the Moment Generating Function of: Exponential Distribution, $X \sim \text{Exponential}(\lambda)$. Poisson Distribution, $X \sim \text{Poisson}(\lambda)$

3. Let $X_1, X_2, \ldots, X_k$ be a sequence of k mutually independent random variables. Use the Moment Generating Function (MGF) technique to prove the following closure properties:

   (a) **Binomial Distribution:** If $X_i \sim Bin(n_i, p)$, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Binomial distribution.

   (b) **Gamma Distribution:** If $X_i \sim Gamma(\alpha_i, \beta)$, where β is the common scale parameter, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Gamma distribution.

   (c) **Normal Distribution :** If $X_i \sim N(\mu_i, \sigma_i^2)$, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Normal distribution.

1 Problem 1

Let X be a random variable with mean μ and standard deviation σ. Let

$$P(|X-\mu| < k\sigma) = P(\mu - k\sigma < X < \mu + k\sigma)$$

1. **Calculate the Bound (Chebyshev):** Determine the lower probability bound for the following values of k: $k = 1.5$, $k = 2$, and $k = 3$.

2. **Calculate the Exact Probability (Normal):** Assume the random variable X follows a Normal distribution. Calculate the exact probability $P(\mu - k\sigma < X < \mu + k\sigma)$ for the same values of k.

3. **Analysis:** Compare the results in a table and explain the difference between the Chebyshev bound and the exact probability.

2 Solutions

Step 1: The Chebyshev Bound

The lower bound on the probability is calculated directly using $1 - 1/k^2$. This bound holds true for **any** probability distribution with finite mean and variance.

1. **For $k = 1.5$:**
$$1 - \frac{1}{(1.5)^2} = 1 - \frac{1}{2.25} \approx 1 - 0.4444 = 0.5556 \text{ (or } 55.56\%)$$

2. **For $k = 2$:**
$$1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 1 - 0.25 = 0.75 \text{ (or } 75\%)$$

3. **For $k = 3$:**
$$1 - \frac{1}{(3)^2} = 1 - \frac{1}{9} \approx 1 - 0.1111 = 0.8889 \text{ (or } 88.89\%)$$

Step 2: The Exact Probability for the Normal Distribution

If X is normal distributed, the probability within the interval is calculated by standardizing X to the standard normal variable $Z = (X - \mu)/\sigma$.

$$P(\mu - k\sigma < X < \mu + k\sigma) = P(-k < Z < k)$$

Using the cumulative distribution function $\Phi(z)$, the probability is:

$$P(-k < Z < k) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

We use tabulated (or computed) values for $\Phi(k)$:

1. **For $k = 1.5$:**
$$P(-1.5 < Z < 1.5) = 2 \times \Phi(1.5) - 1 \approx 2 \times 0.9332 - 1 = 0.8664 \text{ (or } 86.64\%)$$

2. **For $k = 2$:**
$$P(-2 < Z < 2) = 2 \times \Phi(2.0) - 1 \approx 2 \times 0.9772 - 1 = 0.9544 \text{ (or } 95.44\%)$$

3. **For $k = 3$:**
$$P(-3 < Z < 3) = 2 \times \Phi(3.0) - 1 \approx 2 \times 0.99865 - 1 = 0.9973 \text{ (or } 99.73\%)$$

Multiple (k)	**Chebyshev's Bound ($1 - 1/k^2$)**	**Exact Normal Probability ($2\Phi(k) - 1$)**
1.5	55.56%	86.64%
2.0	75.00%	95.44%
3.0	88.89%	99.73%

In every case, the exact probability for the Normal distribution is significantly **higher** than the minimum bound guaranteed by Chebyshev's Inequality.

Chebyshev's Inequality: This is a **universal** and conservative bound. It uses no information about the distribution's shap, only the mean and variance. Therefore, it must be broad enough to be valid for highly irregular or skewed distributions.

Exact Normal Probability: This uses the full information of the Normal distribution (which is symmetric, unimodal, and well-behaved). Since the Normal distribution has most of its probability mass concentrated tightly around the mean, its probability of falling within the interval $\pm k\sigma$ is much higher than the minimum limit guaranteed by Chebyshev.

As k increases, both values approach 1, but the difference between them remains substantial, illustrating the cost of the **generality** of Chebyshev's bound.

Problem 2

Find the Moment Generating Function of: Exponential Distribution, $X \sim \text{Exponential}(\lambda)$. Poisson Distribution, $X \sim \text{Poisson}(\lambda)$

Solution

The probability density function (PDF) is $f(x) = \lambda e^{-\lambda x}$ for $x > 0$.

$$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx} f(x) dx = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$

$$M_X(t) = \lambda \int_0^\infty e^{-(\lambda - t)x} dx$$

For the integral to converge, we require $\lambda - t > 0$, or $t < \lambda$. Evaluating the integral:

$$M_X(t) = \lambda \left[\frac{e^{-(\lambda - t)x}}{-(\lambda - t)} \right]_0^\infty = \frac{\lambda}{\lambda - t} \left[-e^{-(\lambda - t)x} \right]_0^\infty$$

$$M_X(t) = \frac{\lambda}{\lambda - t} \left[0 - (-e^0) \right] = \frac{\lambda}{\lambda - \mathbf{t}}, \quad \text{for } t < \lambda$$

The PMF is $p(k) = \frac{e^{-\lambda} \lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$.

$$M_X(t) = E[e^{tX}] = \sum_{k=0}^\infty e^{tk} P(X = k) = \sum_{k=0}^\infty e^{tk} \frac{e^{-\lambda} \lambda^k}{k!}$$

Factoring out the constant $e^{-\lambda}$ and grouping terms:

$$M_X(t) = e^{-\lambda} \sum_{k=0}^\infty \frac{(\lambda e^t)^k}{k!}$$

Recognizing the Maclaurin series for e^a, $\sum_{k=0}^\infty \frac{a^k}{k!} = e^a$, with $a = \lambda e^t$:

$$M_X(t) = e^{-\lambda} e^{\lambda e^t} = \exp\{\lambda(\mathbf{e^t - 1})\}$$

3 Problem 3

Let $X_1, X_2, \ldots, X_k$ be a sequence of k mutually independent random variables. Use the Moment Generating Function (MGF) technique to prove the following closure properties:

1. **Binomial Distribution:** If $X_i \sim Bin(n_i, p)$, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Binomial distribution.
2. **Gamma Distribution:** If $X_i \sim Gamma(\alpha_i, \beta)$, where β is the common scale parameter, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Gamma distribution.
3. **Normal Distribution :** If $X_i \sim N(\mu_i, \sigma_i^2)$, the sum $Y = \sum_{i=1}^{k} X_i$ follows a Normal distribution.

4 Solutions

The fundamental property of MGFs for independent random variables is that the MGF of a sum of variables is the product of their individual MGFs. Let $Y = \sum_{i=1}^{k} X_i$.

$$M_Y(t) = E[e^{tY}] = E\left[e^{t\sum X_i}\right] = E\left[\prod e^{tX_i}\right] = \prod_{i=1}^{k} E[e^{tX_i}] = \prod_{i=1}^{k} M_{X_i}(t)$$

Sum of Binomials

Let $X_i \sim Bin(n_i, p)$ for $i = 1, \ldots, k$. Note that the success probability p must be the same for all X_i.

MGF of X_i:

$$M_{X_i}(t) = (pe^t + 1 - p)^{n_i}$$

MGF of the Sum Y:

$$\begin{aligned}
M_Y(t) &= \prod_{i=1}^{k} M_{X_i}(t) \\
&= \prod_{i=1}^{k} (pe^t + 1 - p)^{n_i} \\
&= (pe^t + 1 - p)^{n_1} \cdot (pe^t + 1 - p)^{n_2} \cdots (pe^t + 1 - p)^{n_k} \\
&= (pe^t + 1 - p)^{\sum_{i=1}^{k} n_i}
\end{aligned}$$

The resulting MGF, $M_Y(t)$, has the exact form of the MGF of a Binomial distribution with parameters $N = \sum_{i=1}^{k} n_i$ and p. Since the MGF uniquely determines the distribution, we conclude that:

$$Y \sim Bin\left(\sum_{i=1}^{k} n_i, p\right)$$

Sum of Gammas

Let $X_i \sim Gamma(\alpha_i, \beta)$ for $i = 1, \ldots, k$, where α_i is the shape parameter and β is the common scale parameter.

MGF of X_i:

$$M_{X_i}(t) = (1 - \beta t)^{-\alpha_i} \quad \text{for } t < 1/\beta$$

MGF of the Sum Y:

$$\begin{aligned}
M_Y(t) &= \prod_{i=1}^{k} M_{X_i}(t) \\
&= \prod_{i=1}^{k} (1-\beta t)^{-\alpha_i} \\
&= (1-\beta t)^{-\alpha_1} \cdot (1-\beta t)^{-\alpha_2} \cdots (1-\beta t)^{-\alpha_k} \\
&= (1-\beta t)^{-\sum_{i=1}^{k} \alpha_i}
\end{aligned}$$

The resulting MGF, $M_Y(t)$, has the exact form of the MGF of a Gamma distribution with parameters $\alpha = \sum_{i=1}^{k} \alpha_i$ and β. Since the MGF uniquely determines the distribution, we conclude that:

$$Y \sim Gamma\left(\sum_{i=1}^{k} \alpha_i, \beta\right)$$

Sum of Normals

Let $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \ldots, k$.

MGF of X_i:

$$M_{X_i}(t) = e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2}$$

MGF of the Sum Y:

$$\begin{aligned}
M_Y(t) &= \prod_{i=1}^{k} M_{X_i}(t) \\
&= \prod_{i=1}^{k} e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2} \\
&= e^{(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2)} \cdot e^{(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2)} \cdots e^{(\mu_k t + \frac{1}{2}\sigma_k^2 t^2)} \\
&= e^{\sum_{i=1}^{k} (\mu_i t + \frac{1}{2}\sigma_i^2 t^2)} \\
&= e^{\left(\sum_{i=1}^{k} \mu_i\right) t + \frac{1}{2}\left(\sum_{i=1}^{k} \sigma_i^2\right) t^2}
\end{aligned}$$

The resulting MGF, $M_Y(t)$, has the exact form of the MGF of a Normal distribution with parameters $\mu = \sum_{i=1}^{k} \mu_i$ (the sum of the means) and $\sigma^2 = \sum_{i=1}^{k} \sigma_i^2$ (the sum of the variances). Since the MGF uniquely determines the distribution, we conclude that:

$$Y \sim N\left(\sum_{i=1}^{k} \mu_i, \sum_{i=1}^{k} \sigma_i^2\right)$$

Probability Theory

Tutorial 11

1. Let $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \text{geometric probability})$ and $X(\omega) \equiv 0$. Let the sequence of random variables $(X_n)_{n=1}^{\infty}$ be defined as:

$$X_n(t) = n\mathbf{1}_{[0,1/n]}(t).$$

Show that $X_n \xrightarrow{\text{p.}} 0$, X_n does not converge in mean square and $X_n \xrightarrow{\text{a.s}} 0$

2. Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent random variables with $X_1 = X_2 \equiv 0$ and for $n > 2$

$$P(X_n = n) = P(X_n = -n) = \frac{1}{2n\log(n)}, \quad P(X_n = 0) = 1 - \frac{1}{n\log(n)}.$$

Show that the sample mean $M_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converges in probability but NOT almost surely.

Problem 1

Let $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \text{geometric probability})$ and $X(\omega) \equiv 0$. Let the sequence of random variables $(X_n)_{n=1}^{\infty}$ be defined as:

$$X_n(t) = n\mathbf{1}_{[0,1/n]}(t).$$

Show that $X_n \xrightarrow{\text{p.}} 0$, X_n does not converge in mean square and $X_n \xrightarrow{\text{a.s}} 0$

Solution

Let $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}([0,1]), \text{geometric probability})$ and $X(\omega) \equiv 0$. Let the sequence of random variables $(X_n)_{n=1}^{\infty}$ be defined as:

$$X_n(t) = n\mathbf{1}_{[0,1/n]}(t).$$

Convergence in Probability: $X_n \xrightarrow{\text{p.}} 0$. For any $\epsilon > 0$, we need to show $P(|X_n - 0| \geq \epsilon) \to 0$.

$$P(|X_n| \geq \epsilon) = P(\{t \in [0,1] \mid n\mathbf{1}_{[0,1/n]}(t) \geq \epsilon\}) = P([0,1/n]) = \frac{1}{n}.$$

Convergence in Mean Square: X_n does **not** converge in mean square. We compute the mean square error, $E[(X_n - 0)^2] = E[X_n^2]$.

$$E[X_n^2] = \int_0^1 (X_n(t))^2 \, dt = \int_0^{1/n} (n)^2 \, dt + \int_{1/n}^1 (0)^2 \, dt = \int_0^{1/n} n^2 \, dt = n^2 \cdot \frac{1}{n} = n.$$

Since $\lim_{n\to\infty} E[X_n^2] = \lim_{n\to\infty} n = \infty$, the sequence does not converge in mean square.

Almost Sure Convergence: $X_n \xrightarrow{\text{a.s.}} 0$. We analyze the pointwise convergence of the sequence.

For any fixed $t \in (0,1]$, there exists an integer N such that $1/N < t$. Then for all $n > N$, we have $1/n < 1/N < t$, which means $t \notin [0, 1/n]$. Therefore, $X_n(t) = 0$ for all $n > N$, and $\lim_{n\to\infty} X_n(t) = 0$.

The only point where this fails is at $t = 0$, where $X_n(0) = n$ for all n, and the sequence diverges.

The set of non-convergence is $\{\omega \in \Omega \mid \lim X_n(\omega) \neq 0\} = \{0\}$. The probability of this set is $P(\{0\}) = 0$. Since the set of convergence has probability $P(\Omega \setminus \{0\}) = 1$, the sequence converges almost surely.

Problem 2

Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent random variables with $X_1 = X_2 \equiv 0$ and for $n > 2$

$$P(X_n = n) = P(X_n = -n) = \frac{1}{2n\log(n)}, \quad P(X_n = 0) = 1 - \frac{1}{n\log(n)}.$$

Show that the sample mean $M_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converges in probability but NOT almost surely.

Solution

Convergence in Probability (WLLN)

To show that $M_n \xrightarrow{\text{p.}} 0$, we will prove that $\mathrm{E}[M_n^2] \to 0$ as $n \to \infty$ and then apply Chebyshev's inequality.

The expectation of X_n is:

$$E[X_n] = n \cdot \frac{1}{2n\log(n)} + (-n) \cdot \frac{1}{2n\log(n)} + 0 \cdot \left(1 - \frac{1}{n\log(n)}\right) = 0 \quad \text{for } n > 2.$$

For $n = 1, 2$, $X_n = 0$, so $E[X_n] = 0$ as well. Since $E[X_n] = 0$ for all n, we have $E[M_n] = 0$.

The variance of the sample mean is:

$$\mathrm{Var}(M_n) = \frac{1}{n^2}\sum_{k=1}^{n} Var(X) = \frac{1}{n^2}\sum_{k=1}^{n} E(X^2).$$

For $k > 2$, the second moment is:

$$E[X_k^2] = k^2 \cdot \frac{1}{2k\log(k)} + (-k)^2 \cdot \frac{1}{2k\log(k)} = \frac{k}{\log(k)}.$$

For $k = 1, 2$, $E[X_k^2] = 0$. Thus,

$$\mathrm{Var}(M_n) = \frac{1}{n^2}\sum_{k=3}^{n} \frac{k}{\log(k)}.$$

Since $\frac{k}{\log(k)}$ is an increasing function for $k \geq 3$, we have $\sum_{k=3}^{n} \frac{k}{\log(k)} \leq \sum_{k=3}^{n} \frac{n}{\log(n)} = (n-2)\frac{n}{\log(n)}$.

Therefore,

$$\mathrm{Var}(M_n) \leq \frac{1}{n^2}\frac{n(n-2)}{\log(n)} = \frac{n-2}{n\log(n)}.$$

As $n \to \infty$, $\frac{n-2}{n\log(n)} \to 0$. So, $\mathrm{Var}(M_n) \to 0$.

By Chebyshev's inequality, for any $\epsilon > 0$:

$$P(|M_n| > \epsilon) \leq \frac{\mathrm{Var}(M_n)}{\epsilon^2} \to 0.$$

Hence, $M_n \xrightarrow{\text{p.}} 0$.

Almost Sure Convergence (SLLN)

To show that $M_n \overset{\text{a.s.}}{\not\to} 0$, we use the second Borel-Cantelli Lemma. We will show that the event $|X_n| = n$ occurs infinitely often with probability 1.

Let's define the event $A_n = \{|X_n| = n\}$. The probability of this event is:

$$P(A_n) = P(X_n = n) + P(X_n = -n) = \frac{1}{2n\log(n)} + \frac{1}{2n\log(n)} = \frac{1}{n\log(n)}.$$

Now, we examine the series of these probabilities:

$$\sum_{n=3}^{\infty} P(A_n) = \sum_{n=3}^{\infty} \frac{1}{n\log(n)}.$$

This series diverges by the integral test:

$$\int_3^\infty \frac{1}{x\log(x)}dx = [\log(\log(x))]_3^\infty = \infty.$$

Since $\sum P(A_n) = \infty$ and the events A_n are independent, the second Borel-Cantelli Lemma states that $P(A_n \text{ occurs i.o.}) = 1$. This means that with probability 1, the event $|X_n| = n$ occurs for infinitely many values of n.

For these infinitely many occurrences, we have $\left|\frac{X_n}{n}\right| = \frac{n}{n} = 1$. This means that $\left|\frac{X_n}{n}\right|$ does not converge to 0.

Since almost sure convergence of the mean implies that $\frac{X_n}{n} \to 0$ almost surely, and we have shown that $\left|\frac{X_n}{n}\right|$ does not converge to 0 almost surely, it follows that the sequence of sample means does not converge almost surely.

Probability Theory

Tutorial 12

1. Let $(U_k)_{k=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d.) random variables, where each U_k follows a continuous uniform distribution on the interval $[0, 1]$. Let $M_n = \max\{U_1, U_2, \ldots, U_n\}$. Show that the scaled random variable $Z_n = n(1 - M_n)$ converges in distribution to a random variable Z that follows an exponential distribution with parameter $\lambda = 1$.

2. An instructor has 50 exams that will be graded in sequence. The times required to grade the exams are independent, with a common distribution that has a mean $\mu = 20$ minutes and a standard deviation $\sigma = 4$ minutes. Approximate the probability that the instructor will grade at least 25 of the exams in the first 450 minutes of work.

3. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with: $E[X_i] = \mu$, and $Var(X_i) = \sigma^2 < \infty$ and $E(X_i^4) < \infty$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.

   (a) $S_n^2 \xrightarrow{P} \sigma^2$ as $n \to \infty$.

   (b) $\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \xrightarrow{d} N(0, 1)$ as $n \to \infty$.

Problem 1

Let $(U_k)_{k=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d.) random variables, where each U_k follows a continuous uniform distribution on the interval $[0,1]$. Let $M_n = \max\{U_1, U_2, \ldots, U_n\}$. Show that the scaled random variable $Z_n = n(1-M_n)$ converges in distribution to a random variable Z that follows an exponential distribution with parameter $\lambda = 1$.

Solution

To prove convergence in distribution, we will find the limit of the cumulative distribution function (CDF) of Z_n as $n \to \infty$ and show that it matches the CDF of a standard exponential distribution. The CDF of an exponential distribution with parameter $\lambda = 1$ is $F_Z(t) = 1 - e^{-t}$ for $t \geq 0$ and $F_Z(t) = 0$ for $t < 0$.

The CDF of a single random variable U_k is:

$$F_{U_k}(t) = P(U_k \leq t) = \begin{cases} 0 & \text{if } t < 0 \\ t & \text{if } 0 \leq t \leq 1 \\ 1 & \text{if } t > 1 \end{cases}$$

Now, let's find the CDF of $Z_n = n(1 - M_n)$:

$$F_{Z_n}(t) = P(Z_n \leq t) = P(n(1 - M_n) \leq t).$$

We rearrange the inequality to work with M_n:

$$1 - M_n \leq \frac{t}{n} \implies M_n \geq 1 - \frac{t}{n}.$$

Instead of calculating $F_{Z_n}(t)$ directly, we can use the complementary CDF, $1 - F_{Z_n}(t) = P(Z_n > t)$, which is easier to manipulate in this case:

$$P(Z_n > t) = P(n(1 - M_n) > t) = P\left(M_n < 1 - \frac{t}{n}\right).$$

The event $\{M_n < 1 - \frac{t}{n}\}$ occurs if and only if every single random variable U_k is less than $1 - \frac{t}{n}$. Since the random variables are independent, we can write:

$$P\left(M_n < 1 - \frac{t}{n}\right) = \prod_{k=1}^{n} P\left(U_k < 1 - \frac{t}{n}\right) = \left(P\left(U_1 < 1 - \frac{t}{n}\right)\right)^n.$$

For a fixed $t > 0$ and for sufficiently large n, we have $0 < 1 - \frac{t}{n} < 1$. Thus, we can use the CDF of a single uniform random variable:

$$P\left(U_1 < 1 - \frac{t}{n}\right) = 1 - \frac{t}{n}.$$

Substituting this back into our expression for the probability:

$$P(Z_n > t) = \left(1 - \frac{t}{n}\right)^n.$$

We now take the limit as $n \to \infty$. This is a well-known limit for the exponential function:

$$\lim_{n\to\infty} P(Z_n > t) = \lim_{n\to\infty} \left(1 - \frac{t}{n}\right)^n = e^{-t}.$$

Thus, the limiting complementary CDF is $P(Z > t) = e^{-t}$ for $t \geq 0$. By extension, the limiting CDF for $t \geq 0$ is $F_Z(t) = P(Z \leq t) = 1 - P(Z > t) = 1 - e^{-t}$. For $t < 0$, Z_n is a non-negative random variable for large n, so $F_{Z_n}(t) = 0$, and thus $\lim_{n\to\infty} F_{Z_n}(t) = 0$.

The limiting CDF is:

$$F_Z(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 - e^{-t} & \text{if } t \geq 0 \end{cases}$$

This is the CDF of a standard exponential distribution with parameter $\lambda = 1$. Therefore, we conclude that $Z_n = n(1 - M_n)$ converges in distribution to an exponential random variable.

Problem 2

An instructor has 50 exams that will be graded in sequence. The times required to grade the exams are independent, with a common distribution that has a mean $\mu = 20$ minutes and a standard deviation $\sigma = 4$ minutes. Approximate the probability that the instructor will grade at least 25 of the exams in the first 450 minutes of work.

Solution:

Let X_i be the time (in minutes) required to grade the i-th exam. We are interested in the time it takes to grade the first $n = 25$ exams, which is the sum:

$$X = \sum_{i=1}^{25} X_i$$

The event that the instructor grades at least 25 exams in the first 450 minutes is equivalent to the time required to grade those first 25 exams being less than or equal to 450 minutes. We want to approximate $P\{X \leq 450\}$.

Since the X_i are independent and identically distributed (i.i.d.), the expectation and variance of the sum are:

$$E[X] = E\left[\sum_{i=1}^{25} X_i\right] = \sum_{i=1}^{25} E[X_i] = 25 \cdot \mu = 25(20) = 500 \text{ minutes}$$

The variance is:

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^{25} X_i\right) = \sum_{i=1}^{25} \text{Var}(X_i) = 25 \cdot \sigma^2 = 25(4^2) = 25(16) = 400$$

The standard deviation of X is $\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{400} = 20$ minutes.

Since $n = 25$ is a sufficiently large number of i.i.d. variables, we can approximate the distribution of the sum X as a normal distribution $N(E[X], \text{Var}(X)) = N(500, 400)$.

We standardize X to find the approximate probability:

$$P\{X \leq 450\} \approx P\left\{\frac{X - E[X]}{\sqrt{\text{Var}(X)}} \leq \frac{450 - 500}{\sqrt{400}}\right\}$$

$$P\{X \leq 450\} \approx P\left\{Z \leq \frac{-50}{20}\right\} = P\{Z \leq -2.5\}$$

where Z is a standard normal random variable.

Using the symmetry of the standard normal distribution and the standard normal CDF $\Phi(a)$:

$$P\{Z \leq -2.5\} = 1 - P\{Z \leq 2.5\} = 1 - \Phi(2.5)$$

Consulting a standard normal table or calculator for $\Phi(2.5)$:

$$\Phi(2.5) \approx 0.9938$$

Thus, the approximate probability is:

$$P\{X \leq 450\} \approx 1 - 0.9938 = 0.0062$$

The probability that the instructor will grade at least 25 exams in the first 450 minutes is approximately **0.0062**.

Problem 3

Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables with: $E[X_i] = \mu$, $Var(X_i) = \sigma^2 < \infty$ and $E(X_i^4) < \infty$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$.

1. $S_n^2 \xrightarrow{P} \sigma^2$ as $n \to \infty$.
2. $\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \xrightarrow{d} N(0,1)$ as $n \to \infty$.

Solution:

1. Proof of $S_n^2 \xrightarrow{P} \sigma^2$ The sample variance can be decomposed as:

$$S_n^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}_n^2\right)$$

By the **Weak Law of Large Numbers (WLLN)**:

- $\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{P} E[X^2]$
- $\bar{X}_n \xrightarrow{P} \mu \implies \bar{X}_n^2 \xrightarrow{P} \mu^2$ (by Continuous Mapping Theorem)

Since $\frac{n}{n-1} \to 1$ as $n \to \infty$:

$$S_n^2 \xrightarrow{P} 1 \cdot (E[X^2] - \mu^2) = \sigma^2$$

2. Proof of CLT with Estimated Variance By the standard **Central Limit Theorem**:

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

We rewrite the target expression:

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right) \cdot \left(\frac{\sigma}{S_n}\right)$$

From Part 1, we have $S_n^2 \xrightarrow{P} \sigma^2$. Applying the Continuous Mapping Theorem for $g(t) = \sqrt{\sigma^2/t}$:

$$\frac{\sigma}{S_n} \xrightarrow{P} 1$$

By **Slutsky's Theorem**, if $Y_n \xrightarrow{d} Y$ and $W_n \xrightarrow{P} c$:

$$Y_n W_n \xrightarrow{d} cY$$

Defining $Y_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$ and $W_n = \frac{\sigma}{S_n}$, we obtain:

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \xrightarrow{d} 1 \cdot Z \sim N(0,1)$$

Probability Theory

Tutorial 13

Question 1

In a box there are 8 coins, enumerated by $1, 2, \ldots, 8$. Let's assume the probability that the i-th coin falls showing heads (H) is $P(H|C_i) = \frac{i}{8}$. A coin is chosen at random and tossed, and it shows H. What is the conditional probability that this coin is marked with the number 6?

(A) 1/4 (B) 1/8 (C) 1/6 (D) 3/8 (E) 2/9

Question 2

Let $X \sim N(1, 4)$ be a Gaussian probability distribution with $\mu = 1$ and $\sigma = 2$ (since $\sigma^2 = 4$). Compute the value of the moment generating function $M_X(t) = E(e^{tX})$ at the point $t = 1$:

(A) e (B) e^2 (C) e^3 (D) e^4 (E) e^6

Question 3

We participate in an Olympic long-jump competition, in which each participant has three attempts to jump. Assume that:

- A third (1/3) of all jumps are invalid (foul).
- Valid jumps are uniformly distributed within the interval $[2, 3]$.
- Attempts are independent events.

The result (R) for each candidate is the **maximum** of their three attempts. What is the expected value ($E[R]$) for each participant's result?

(A) 19 (B) 20 (C) 21 (D) 23 (E) 2

Question 4

The following two questions are independent:

A) Three factories produce cars for the same company.

- Factory 1: Produces 50% of cars, defect rate 1%.
- Factory 2: Produces 20% of cars, defect rate 2%.
- Factory 3: Produces 30% of cars, defect rate 3%.

A car is purchased and found to be defective. What is the probability that it was produced in Factory 2?

B) Let X be a Poisson random variable with parameter λ. Derive its **moment-generating function** $M_X(t) = E[e^{tX}]$.

Question 5

The following two questions are independent:

A) **Prove:** Suppose the events A, B, and C are independent. Show that the events A and $(B \cup C)$ are also independent.

B) Is it possible to find a random variable X satisfying the identity:

$$P(X \geq E[X] + 4\sqrt{Var(X)}) = 0.3$$

If "Yes", provide a concrete example. If "No", explain using probability inequalities.

Question 6

Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables such that $X_n \sim U(0,1)$ for each n. Define $Y_n := -\log(X_n)$.

A) Compute the expectation value $E[Y_1]$. *Hint: Use the fact that the primitive of* $\log(x)$ *is* $x\log(x) - x$.

B) What is the joint density function of $(Y_1, Y_2, \dots, Y_n)$?

C) Compute the following limit:

$$\lim_{n\to\infty} P(X_1 \cdot X_2 \cdot \dots \cdot X_n \leq e^{-n})$$

Question 7

In an airplane there are 100 seats. Each of the 100 passengers receives a digital card showing a unique number between 1 and 100 (a bijection between passengers and seats). Due to an internet error, passengers obtain totally random seat numbers and sit down in a random, unsystematic fashion (a random permutation).

What is the expected number of passengers who end up sitting in the seat originally assigned to them by their digital card?

(A) 1 (B) 2 (C) 3 (D) 4 (E) 5

Question 8

In a box there are 3 balls, enumerated 1, 2, and 3. We randomly extract a first ball (X) and then a second ball (Y) without replacement. The correlation coefficient $\rho_{X,Y}$ between X and Y is:

(A) 0 (B) $-1/2$ (C) $-1/3$ (D) $-1/4$ (E) $-1/5$

Question 9

Consider a Gaussian probability vector (X, Y, Z) with the joint density function:

$$f_{X,Y,Z}(x,y,z) = C\exp\left(-\frac{1}{2}\left[(x-1)^2 + 5(y-2)^2 + (z+1)^2 + 4(y-2)(z+1)\right]\right)$$

for a suitable positive constant C.

Compute the expected value $E[(X+Y+Z)^2]$.

(A) 4 (B) 7 (C) 8 (D) 9 (E) 10

Question 10

Let X and Y be two random variables with a joint density function given by:

$$f_{X,Y}(x,y) = \begin{cases} 8xy & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

A) Are X and Y independent from each other? Justify your answer.

B) Compute the covariance $Cov(X,Y)$.

C) Let $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ be infinitely many i.i.d. copies of the pair (X,Y), and let A_m denote the event $\{X_m < Y_m < 1/m\}$. Compute the probability that A_m occurs for infinitely many indices m.

Solutions

Question 1

Using Bayes' Theorem:

$$P(C_6|H) = \frac{P(H|C_6)P(C_6)}{P(H)}$$

1. Since the coin is chosen at random: $P(C_i) = \frac{1}{8}$ for all i. 2. The total probability of H is:

$$P(H) = \sum_{i=1}^{8} P(H|C_i)P(C_i) = \sum_{i=1}^{8} \frac{i}{8} \cdot \frac{1}{8} = \frac{1}{64} \sum_{i=1}^{8} i$$

Using the sum formula $\frac{n(n+1)}{2}$:

$$P(H) = \frac{1}{64} \cdot \frac{8(9)}{2} = \frac{36}{64} = \frac{9}{16}$$

3. The conditional probability is:

$$P(C_6|H) = \frac{(6/8) \cdot (1/8)}{9/16} = \frac{6/64}{9/16} = \frac{6}{64} \cdot \frac{16}{9} = \frac{6}{36} = \frac{1}{6}$$

Correct Answer: (C)

Question 2

The Moment Generating Function (MGF) for a normal distribution $N(\mu, \sigma^2)$ is:

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

Given $\mu = 1$ and $\sigma^2 = 4$, we evaluate at $t = 1$:

$$M_X(1) = e^{(1)(1) + \frac{1}{2}(4)(1)^2}$$

$$M_X(1) = e^{1+2} = e^3$$

Correct Answer: (C)

Question 3

Let X_i be the result of the i-th attempt. This is a mixture distribution:

- $P(X_i = 0) = 1/3$ (foul jump).
- $X_i \sim U(2, 3)$ with probability $2/3$ (valid jump).

The Cumulative Distribution Function (CDF) of a single attempt X for $r \in [2, 3]$ is:

$$F_X(r) = P(X \le r) = P(\text{foul}) + P(\text{valid})P(\text{jump} \le r|\text{valid}) = \frac{1}{3} + \frac{2}{3}(r-2).$$

The result for each candidate is $R = \max(X_1, X_2, X_3)$. Its CDF is $F_R(r) = [F_X(r)]^3$. To find the expected value using the density, we differentiate $F_R(r)$ for $r \in (2, 3)$:

$$f_R(r) = \frac{d}{dr}\left[\frac{1}{3} + \frac{2}{3}(r-2)\right]^3 = 3\left[\frac{1}{3} + \frac{2}{3}(r-2)\right]^2 \cdot \frac{2}{3} = 2\left[\frac{1}{3} + \frac{2}{3}(r-2)\right]^2.$$

The expected value $E[R]$ is calculated as follows (the discrete part at $r = 0$ contributes 0):

$$E[R] = \int_2^3 r \cdot f_R(r)\, dr = \int_2^3 r \cdot 2\left[\frac{1}{3} + \frac{2}{3}(r-2)\right]^2 dr.$$

Using the substitution $u = \frac{1}{3} + \frac{2}{3}(r-2)$, we have $du = \frac{2}{3}dr$ and $r = \frac{3}{2}(u+1)$. The limits change from $[2, 3]$ to $[1/3, 1]$:

$$\begin{aligned}
E[R] &= \int_{1/3}^{1} \frac{3}{2}(u+1) \cdot 2u^2 \cdot \frac{3}{2}\, du \\
&= \frac{9}{2} \int_{1/3}^{1} (u^3 + u^2)\, du \\
&= \frac{9}{2} \left[\frac{u^4}{4} + \frac{u^3}{3} \right]_{1/3}^{1} \\
&= \frac{9}{2} \left[\left(\frac{1}{4} + \frac{1}{3} \right) - \left(\frac{1}{324} + \frac{1}{81} \right) \right] = \frac{23}{9} \approx 2.55.
\end{aligned}$$

Correct Answer: (E)

Question 4

A) Let D be the event "defective" and F_i be Factory i.

$$P(D) = (0.5 \cdot 0.01) + (0.2 \cdot 0.02) + (0.3 \cdot 0.03) = 0.005 + 0.004 + 0.009 = 0.018$$

$$P(F_2|D) = \frac{P(D|F_2)P(F_2)}{P(D)} = \frac{0.004}{0.018} = \frac{2}{9} \approx 0.222$$

B) $M_X(t) = \sum_{k=0}^{\infty} e^{tk} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$.

Question 5

A)

$$P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)$$

By independence: $P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) = P(A)[P(B) + P(C) - P(B \cap C)] = P(A)P(B \cup C)$.

B) No.

To determine if such a random variable X exists, we define $k = 4$ as the number of standard deviations ($\sigma = \sqrt{Var(X)}$) above the mean $\mu = E[X]$. We are looking for:

$$P(X \geq \mu + 4\sigma) = 0.3$$

By **Chebyshev's Inequality** (two-tailed version), we know that for any $k > 0$:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

For $k = 4$, this gives $P(|X - \mu| \geq 4\sigma) \leq \frac{1}{16} = 0.0625$. Since the event $\{X \geq \mu + 4\sigma\}$ is a subset of the event $\{|X - \mu| \geq 4\sigma\}$, its probability must be even smaller:

$$P(X \geq \mu + 4\sigma) \leq P(|X - \mu| \geq 4\sigma) \leq 0.0625$$

The required probability 0.3 is significantly greater than the maximum possible bound 0.0625.

Question 6

A) $E[Y_1] = \int_0^1 -\log(x)dx = -[x\log x - x]_0^1 = 1$.

B) $Y_n = -\log X_n \implies X_n = e^{-Y_n}$. The Jacobian is $|dx/dy| = e^{-y}$. $f_Y(y) = f_X(e^{-y})e^{-y} = 1 \cdot e^{-y}$ for $y > 0$. They are i.i.d. $Exp(1)$. Joint density is $\prod e^{-y_i} = e^{-\sum y_i}$.

C) $P(\prod X_i \le e^{-n}) = P(\sum \log X_i \le -n) = P(-\sum \log X_i \ge n) = P(\sum Y_i \ge n)$. Since $E[Y] = 1$ and $V(Y) = 1$

$$P(\frac{1}{n}\sum Y_i - 1 \ge 0) = P(\frac{\frac{1}{n}\sum Y_i - 1}{\frac{1}{\sqrt{n}}} \ge 0) \to 1 - \phi(0) = 1/2$$

(using CLT symmetry) .

Question 7

Let I_i be an indicator variable that passenger i sits in their correct seat. $P(I_i = 1) = 1/100$. Total expected seats $E[X] = E[\sum_{i=1}^{100} I_i] = \sum E[I_i] = 100 \cdot \frac{1}{100} = 1$.

Correct Answer: (A)

Question 8

$E[X] = E[Y] = 2$. $Var(X) = Var(Y) = 2/3$. $E[XY] = \frac{1}{3\cdot 2}(1\cdot 2+1\cdot 3+2\cdot 1+2\cdot 3+3\cdot 1+3\cdot 2) = \frac{22}{6} = \frac{11}{3}$. $Cov(X,Y) = \frac{11}{3} - 4 = -1/3$. $\rho = \frac{-1/3}{2/3} = -1/2$.

Correct Answer: (B)

Question 9

The exponent is $-\frac{1}{2}(x-1)^2 - \frac{1}{2}[5(y-2)^2 + (z+1)^2 + 4(y-2)(z+1)]$. Mean vector $\mu = (1, 2, -1)$. $E[X+Y+Z] = 1+2-1 = 2$. The quadratic form for (Y,Z) is $(y-2, z+1)\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} y-2 \\ z+1 \end{pmatrix}$. The covariance matrix $\Sigma_{Y,Z}$ is the inverse: $\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$. $Var(X) = 1, Var(Y) = 1, Var(Z) = 5, Cov(Y,Z) = -2, Cov(X,Y) = 0, Cov(X,Z) = 0$. $Var(X+Y+Z) = 1+1+5+2(0+0-2) = 7-4 = 3$. $E[W^2] = Var(W) + (E[W])^2 = 3 + 2^2 = 7$.

Question 10

A) No, they are not independent because the support $0 < x < y < 1$ is not a product space (the range of x depends on y).

B) $E[X] = \int_0^1\int_0^y 8x^2y\,dxdy = 8/15$; $E[Y] = 4/5$; $E[XY] = \int_0^1\int_0^y 8x^2y^2\,dxdy = 4/9$. $Cov(X,Y) = E[XY] - E[X]E[Y] = \frac{4}{9} - \frac{32}{75}$.

C) $P(A_m) = \int_0^{1/8}\int_0^y 8xy\,dxdy = (1/m)^4$. Since $\sum_m P(A_m) < \infty$, by Borel-Cantelli, the probability of infinitely many occurrences is 0.

Probability Theory

Homework 1 - Probability Spaces

1. Suppose $\Omega = \{1, 2, 3\}$ and $\mathcal{F}$ is a collection of all subsets of Ω. Find (with proof) necessary and sufficient conditions on the real numbers x, y, and z such that there exists a countably additive probability measure P on $\mathcal{F}$, with $x = P\{1, 2\}$, $y = P\{2, 3\}$, and $z = P\{1, 3\}$.

2. Let $\Omega = \{1, 2, 3, 4\}$, and let $\mathcal{J} = \{\{1\}, \{2\}\}$. Describe explicitly the σ-algebra $\sigma(\mathcal{J})$ generated by $\mathcal{J}$.

3. Let P and Q be two probability measures defined on the same sample space Ω and σ-algebra $\mathcal{F}$. Suppose that $P(A) = Q(A)$ for all $A \in \mathcal{F}$ with $P(A) \leq \frac{1}{2}$. Prove that $P = Q$. i.e. that $P(A) = Q(A)$ for all $A \in \mathcal{F}$.

4. A point (x, y) is to be selected from the square S containing all points (x, y) such that $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Suppose that the probability that the selected point will belong to each specified subset of S is equal to the area of that subset. Find the probability of each of the following subsets:

   (a) the subset of points such that $\frac{1}{2} < x + y < \frac{3}{2}$;
   (b) the subset of points such that $y \leq 1 - x^2$;
   (c) the subset of points such that $x = y$.

5. Consider a sample space Ω consisting of the 24 permutations of the numbers 1, 2, 3, and 4, all equally likely. We define $A_i = \{\omega \in \Omega \mid \text{in } \omega \text{ the number } i \text{ appears in the } i\text{-th position}\}$. Calculate:

   (a) $P(A_1 \cup A_2 \cup A_3 \cup A_4)$
   (b) $P((A_1 \cup A_2) \cap (A_3 \cup A_4))$
   (c) $P(A_1 \cup A_3)$
   (d) $P(A_1 \cup (A_3 \cap A_4))$

Probability Theory

Homework 2- Conditional and Independence

1. A certain person considers that he can drink and drive: usually he believes he has a negligible chance of being involved in an accident, whereas he believes that if he drinks two pints of beer, his chance of being involved in an accident on the way home is only one in five hundred. Assuming that he drives home from the same pub every night, having drunk two pints of beer, what is the chance that he is involved in at least one accident in one year? Are there any assumptions that you make in answering the question?

2. We have three identical boxes, and each box contains two drawers. One box contains a gold coin in each drawer, another contains a silver coin in each drawer, and the third a gold coin in one drawer and a silver coin in the other.
   (a) A box is randomly selected. What is the probability that the selected box contains coins of different metals?
   (b) A box is randomly selected, and upon randomly opening a drawer, we find a gold coin. What is the probability that the other drawer contains a silver coin?

3. Assuming the events A_1, A_2 and A_3 are independent and that $r = 0, 1, 2, 3$, obtain expressions in terms of $P(A_1), P(A_2)$ and $P(A_3)$ for the probabilities of the following events:
   (a) Exactly r events occur for $r = 0, 1, 2, 3$.
   (b) At least r events occur for $r = 0, 1, 2, 3$.
   (c) At most r events occur for $r = 0, 1, 2, 3$.

4. We throw a fair coin an infinite number of times, independently. Consider a fixed streak (H, H, H, H, H).
   (a) Let A_n be the event "the streak (H, H, H, H, H) appears between throw $n+1$ and $n+5$." Calculate $P(A_n)$ for each n.
   (b) Are A_n and A_{n+1} independent? Are A_n and A_{n+5} independent?
   (c) Show that the probability of the streak (H, H, H, H, H) appearing infinitely many times is 1.
   (d) Can you generalize the previous point to any fixed streak of length k?

Probability Theory

Homework 3 - Discrets Random Variables

1. Let X and Y be discrete random variables, each with support equal to $\{1, 2, 3, 4\}$. Given that $P(X = x, Y = y) = \dfrac{x+y}{80}$, calculate:

   (a) The marginal probability mass function. Are X and Y independent?

   (b) $P(X = Y)$

   (c) $P(XY = 6)$

   (d) $P(1 \leq X \leq 2, 2 < Y \leq 4)$

2. There are three radioactive sources: F_1, F_2, and F_3. The number of particles emitted by each source per hour is a random variable following a Poisson distribution $X_i \sim \text{Pois}(\lambda_i)$, with the following parameters:

   - Source F_1: $\lambda_1 = 2$
   - Source F_2: $\lambda_2 = 3$
   - Source F_3: $\lambda_3 = 4$

   A researcher chooses one source at random and observes that this source emits $X = 4$ particles in one hour. Find the probability that the researcher chose Source F_2.

3. Prove that if a random variable X has a geometric distribution, then they have the memoryless property: $P\{W \geq t + s \mid W \geq t\} = P\{W \geq s\}$ for all $t, s > 0$.

4. $X_1, \ldots, X_n$ are independent and have Poisson distributions $X_i \sim \text{Pois}(\lambda_i)$ for all i. Show that the sum $(X_1 + ... + X_n)$ follows a Poisson distribution $\text{Pois}(\lambda_1 + ... + \lambda_n)$. (Hint: First show that the sum of two independent Poisson variables is Poisson)

Probability Theory

Homework 4 - Continuous Random Variables

1. The height of a randomly selected man from a population is normal with $\mu = 178$cm and $\sigma = 8$cm. What proportion of men from this population are over 185cm tall? There are 2.54cm to an inch. What is their height distribution in inches? The heights of the women in this population are normal with $\mu = 165$ cm and $\sigma = 7$cm. What proportion of the women are taller than half of the men?

2. The distribution function of a continuous random variable X is given by

$$F_X(x) = \begin{cases} a & \text{if } x < 1 \\ 2b\sqrt{x} + c & \text{if } x \in [1, 4) \\ d & \text{if } x \geq 4. \end{cases}$$

   (a) Find a, b, c, d.

   (b) Find the density function of X.

   (c) Calculate $P(X^2 \leq 2)$.

3. Prove that if $U \sim U[0, 1]$ and $\lambda > 0$, then $X = -\frac{1}{\lambda} \log U$ satisfies $X \sim \text{Exp}(\lambda)$.

4. Assume (X, Y) has a uniform distribution on: $D = \{(x, y) \in R^2 | 0 \leq x \leq 1, 0 \leq y \leq x\}$. Find the distribution function of $Z = Y - X$. Find the density of the marginal variables X and Y.

Probability Theory

Homework 5 - Expected Value

1. Let X be a random variable with finite mean, and let $a \in \mathbb{R}$ be any real number. Prove that $E(\max(X, a)) \geq \max(E(X), a)$.

2. Let $f(x) = ax^2 + bx + c$ be a second degree polynomial function (where $a, b, c \in \mathbb{R}$ are constants).

   (a) Find necessary and sufficient conditions on a, b, and c such that the equation $E(f(\alpha X)) = \alpha^2 E(f(X))$ holds for all $\alpha \in \mathbb{R}$ and all random variables X.

   (b) Find necessary and sufficient conditions on a, b, and c such that the equation $E(f(X - \beta)) = E(f(X))$ holds for all $\beta \in \mathbb{R}$ and all random variables X.

3. Let X and Y be independent random variables with $V(X) = 3$ and $V(Y) = 2$.

   (a) Calculate $cov(Y - 2X, 3X - Y)$.

   (b) Let $Z = X + 4Y$. Find $\rho(X, Z)$.

   (c) If in addition $E(X) = 1$ and $E(Y) = 2$. Find $V(XY)$.

4. Let X be a continuous random variable with support $\text{Sop}(X) = (-\infty, +\infty)$. Calculate $E(F(X))$.

Probability Theory

Homework 6 - Convergences and Limit Theorem

1. Show that if the random variables W_n for each $n \in \mathbb{N}$ and $\theta \in (0,1)$ satisfy $E(W_n) = \theta + \frac{1-\theta}{n+1}$ anad $Var\,(W_n) = \frac{n(1-\theta)^2}{(n+1)^2(n+2)}$ then prove that $W_n \xrightarrow{p} \theta$.

2. Let $X_i, i \geq 1$ be i.i.d. random variables with density function $f_X(x) = \frac{2x}{5^2} I_{(0,5)}(x)$. Let $X^{(n)} = \max\{X_1, \ldots, X_n\}$.

   (a) If $n = 4$, calculate $P(X^{(n)} < 3)$.

   (b) Prove that $X^{(n)} \xrightarrow{p} 5$.

   (c) Let $W_n = n(5 - X^{(n)})$. Prove that $W_n \xrightarrow{\mathcal{D}} W$ where $W \sim \epsilon(2/5)$.

3. Let $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ be two sequences of r.v.s. Assume $|X_n - Y_n| \xrightarrow{P} 0$ and $X_n \xrightarrow{D} X$. Show that $Y_n \xrightarrow{D} X$.

4. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of r.v.s. Assume $X_n \xrightarrow{D} c$, where c is a constant. Show that $X_n \xrightarrow{P} c$.