Contents

1 Week 1: 8/22- 8/26

1. The following identities are true:

$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C), \\ A \cap (B \cap C) &= (A \cap B) \cap C, \\ A \backslash (B \cap C) &= (A \backslash B) \cup (A \backslash C). \end{aligned}$$

Hint. Let us show

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \tag{1.1}$$

Indeed, $x \in A \cap (B \cup C) \Leftrightarrow x \in A$ and ($x \in B$ or $x \in C$)$\Leftrightarrow$ ($x \in A$ and $x \in B$) or ($x \in A$ and $x \in C$) $\Leftrightarrow x \in (A \cap B) \cup (A \cap C)$.

Answer. The equality $A \cap (B \cap C) = (A \cap B) \cap C$ is obvious. In order to prove

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

we show that the complements of the sets on both sides are equal. By DeMorgan's law and (1.1),

$$\begin{aligned} (A \cup (B \cap C))^c &= A^c \cap \left(B^c \cup C^c\right) \\ &= \left(A^c \cap B^c\right) \cup \left(A^c \cap C^c\right), \end{aligned}$$

and

$$\begin{aligned} ((A \cup B) \cap (A \cup C))^c &= (A \cup B)^c \cup (A \cup C)^c \\ &= \left(A^c \cap B^c\right) \cup \left(A^c \cap C^c\right). \end{aligned}$$

Now, by de Morgan's law and (1.1) again,

$$\begin{aligned} A \backslash (B \cap C) &= A \cap (B \cap C)^c = A \cap \left(B^c \cup C^c\right) \\ &= (A \cap B^c) \cup (A \cap C^c) = (A \backslash B) \cup (A \backslash C). \end{aligned}$$

2. Let $\{A_i, i \in I\}$ be a collection of sets. Prove De Morgan's Laws:

$$(\cup_i A_i)^c = \cap_i A_i^c, \ (\cap_i A_i)^c = \cup_i A_i^c.$$

3. a) Let $\mathcal{F}$ be a σ-field, $A, B \in \mathcal{F}$. Show that $A \backslash B$ and $A \triangle B \in \mathcal{F}$.

b) Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be σ-fields of subsets of Ω. Show that $\mathcal{F}_1 \cap \mathcal{F}_2$, the collection of subsets of Ω that belong to both $\mathcal{F}_1$ and $\mathcal{F}_2$, is σ-field.

4. Let $\mathcal{A}$ be a collection of subsets of Ω, and let $\mathcal{F}_i, i \in I$, be all σ-fields that contain $\mathcal{A}$. Show that $\mathcal{F} = \cap_i \mathcal{F}_i$ is a σ-field.

Comment. Note that $\mathcal{P}(\Omega)$, the σ-field of all subsets of Ω, is among $\mathcal{F}_i$. The collection $\mathcal{F} = \cap_i \mathcal{F}_i$ is called the smallest σ-field containing $\mathcal{A}$.

5. Describe the sample spaces for the following experiments:

(a) Two balls were drawn without replacement from an urn which originally contained two red and two black balls.

Answer. Mark balls in the urn: $B_1, B_2, R_1.R_2$. If drawn one by one, and going with ordered pairs, then Ω consists of $4 \cdot 3 = 12$ ordered pairs:

$$\{R_1R_2, R_2R_1, R_1B_1, R_1B_2, R_2B_1, R_2B_2, B_1B_2, B_2B_1, B_1R_1, B_1R_2, B_2R_1, B_2R_2\}.$$

If drawn loosely as group of two, then Ω consists of $\binom{4}{2} = 6$ unordered pairs:

$$\{\{R_1, R_2\}, \{R_1, B_1\}, \{R_2, B_1\}, \{R_1, B_2\}, \{R_2, B_2\}, \{B_1, B_2\}\}.$$

(b) A coin is tossed three times.

Answer. $\Omega = \{HHH, HHT, HTT, HTH, TTT, TTH, THT, THH\}$ consists of $2^3 = 8$ different outcomes.

6. A fair die is thrown twice. What is the probability that:

(a) a six turns up exactly once?

Answer. The event A consists of all $(6, i)$ and $(j, 6)$ with $1 \leq i, j \leq 5$. Hence $\mathbf{P}(A) = 10/36 = 5/18$.

(b) both numbers are odd?

Answer. The event A consists of all (i, j) with $i, j \in \{1, 3, 5\}$; $\#A = 3^2 = 9$ and $\mathbf{P}(A) = 9/36 = 1/4$.

(c) the sum of the scores is 4?

Answer. $A = \{(1, 3), (2, 2), (3, 1)\}$, and $\mathbf{P}(A) = 3/36 = 1/12$.

(d) the sum of the scores is divisible by 3?

Answer. $A = \{(1, 2), (1, 5), (2, 1), (2, 4), (3, 3), (3, 6), (4, 2), (4, 5), (5, 1), (5, 4), (6, 3), (6.6)\}$, and

$$\mathbf{P}(A) = 12/36 = 1/3.$$

7. A fair coin is thrown repeatedly. What is the probability that on the nth throw:

(a) a head appears for the first time?

Answer. The sample space Ω consists of all "words" of length n made by H and T: $\#\Omega = 2^n$. The event $A =$" H appears 1st on the nth throw" is a single word with $n - 1$ T before H at the end:

$$\mathbf{P}(A) = 1/2^n = 2^{-n}.$$

(b) the number of heads and tails to date are equal?

Answer. If n is odd the probability is zero. If n is even, then there are $\binom{n}{n/2}$ different "words" with $n/2$ H:

$$\mathbf{P}(A) = \binom{n}{n/2} 2^{-n}.$$

(c) exactly two heads have appeared altogether to date?

Answer. There are $\binom{n}{2}$ different words with exactly two H :

$$\mathbf{P}(A) = \binom{n}{2} 2^{-n}.$$

(d) at least two heads have appeared to date?

Answer. Let B="at most one H" consists of one word with all T in it and n words with a single H in them: $\mathbf{P}(B) = (n+1)\,2^{-n}$ and $\mathbf{P}(A) = 1 - (n+1)\,2^{-n}$.

8. Show that the probability that exactly one of the events A and B occurs is

$$\mathbf{P}(A) + \mathbf{P}(B) - 2\mathbf{P}(A \cap B).$$

(note: it is the event $A \triangle B = (A \backslash B) \cup (B \backslash A)$)

Answer. Since $A \triangle B$ is the union of two disjoint sets,

$$A \triangle B = (A \backslash B) \cup (B \backslash A) = (A \backslash (A \cap B)) \cup (B \backslash (A \cap B)),$$

and $A \cap B \subseteq A$, $A \cap B \subseteq B$, we have

$$\begin{aligned} \mathbf{P}(A \triangle B) &= \mathbf{P}(A \backslash (A \cap B)) + \mathbf{P}(B \backslash (A \cap B)) \\ &= \mathbf{P}(A) - \mathbf{P}(A \cap B) + \mathbf{P}(B) - \mathbf{P}(A \cap B). \end{aligned}$$

9. a) A fair coin is tossed n times. What is the probability of H in the last toss?

Answer. The sample space Ω consists of all "words" of length n made by H and T: $\#\Omega = 2^n$. The event A="H in the last toss" consists of all words of length n formed by H and T with H at the end; $\#A = 2^{n-1}$ and

$$\mathbf{P}(A) = \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$

Similarly, probability of H in any toss is 1/2.

b) An urn contains 9 whites and one red ball. All ten balls are randomly drawn out without replacement one by one. What is the probability that:

(i) the red ball is taken out first?

Answer. The event A="red ball is taken out first" consists of all orderings of 10 balls with red ball first: $\#A = (n-1)!$. Hence

$$\mathbf{P}(A) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

(ii) the red ball is taken out last?

Answer. The event A="red ball is taken out last" consists of all orderings of 10 balls with red ball last: $\#A = (n-1)!$. Hence

$$\mathbf{P}(A) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

(ii) the red ball is taken in the kth draw ($1 \le k \le 10$)?

Answer. The event A="red ball is taken out last" consists of all orderings of 10 balls with red ball in the kth position: $\#A = (n-1)!$. Hence

$$\mathbf{P}(A) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

10. Six cups and saucers come in pairs: there are two cups and saucers which are red, two white, and two with stars on. If the cups are placed randomly onto the saucers (one each), find the probability that no cup is upon a saucer of the same pattern.

Hint. Put the saucers in a row. For instance, $RRWWSS$. The sample space Ω is the space of all cup $(R_1, R_2, W_1, W_2, S_1, S_2)$ arrangements along the row of saucers $RRWWSS$.

Answer. Put the saucers in a row. For instance, $RRWWSS$. The sample space Ω is the space of all cup $(R_1, R_2, W_1, W_2, S_1, S_2)$ arrangements along the row of saucers $RRWWSS$: $\#\Omega = 6!$ Let A="no cup is upon a saucer of the same pattern". Then $A = A_1 \cup A_2$, where A_1 are all arrangements where one pattern cups go to the same different pattern of saucers, and A_2 are all arrangements where one pattern cups go to the different patterns of saucers.

Note A_1 consists of $W_1W_2S_1S_2R_1R_2$ and $S_1S_2R_1R_2W_1W_2$ and their arrangements with switching R_1 and R_2, W_1 and W_2, S_1 and S_2 : $\#A_1 = 2 \cdot 2! \cdot 2!2! = 16$.

Counting A_2. For R_1 we have 4 choices (4 positions in $WWSS$ part), for R_2 we have remaining 2 choices: total number of R arrangements is $4 \cdot 2 = 8$. After both R are placed we have 4 ways to arrange both W. After R and W are placed there are only 2 ways to arrange remaining S. Hence $\#A_2 = 8 \cdot 4 \cdot 2 = 64$, and $\#A = \#A_1 + \#A_2 = 16 + 64 = 80$. Thus $\mathbf{P}(A) = 80/6! = 1/9$.

*2nd answer ("*systemic approach": inclusion/exclusion principle) Now we denote R_i = "ith red cup is put on red saucer", $i = 1, 2$. Similarly, we use $S_i, W_i, i = 1, 2$, notation. It is a version of the matching problem, and

$$\mathbf{P}(\text{no match}) = 1 - \mathbf{P}(\text{at least one match}) = 1 - \mathbf{P}(R_1 \cup R_2 \cup W_1 \cup W_2 \cup S_1 \cup S_2).$$

We apply inclusion/exclusion principle for

$$\begin{aligned} & \mathbf{P}(R_1 \cup R_2 \cup W_1 \cup W_2 \cup S_1 \cup S_2) \\ = \; & \mathcal{S}_1 - \mathcal{S}_2 + \mathcal{S}_3 - \mathcal{S}_4 + \mathcal{S}_5 - \mathcal{S}_6, \end{aligned}$$

where

$$\mathcal{S}_1 = \mathbf{P}(R_1) + \mathbf{P}(R_2) + \mathbf{P}(W_1) + \mathbf{P}(W_2) + \mathbf{P}(S_1) + \mathbf{P}(S_2) = 6 \cdot \frac{2 \cdot 5!}{6!},$$

and $\mathcal{S}_i$ is the sum of all probabilities of intersections of all distinct event groups of size i from the whole collection of six events $\{R_1, R_2, W_1, W_2, S_1, S_2\}$, $i = 2, 3, 4, 5, 6$. For instance, S_2 is the sum of probabilities of all distinct pairs out of $\{R_1, R_2, W_1, W_2, S_1, S_2\}$.

Computation of $\mathcal{S}_2$. There are 3 distinct pairs of the same pattern, and all three probabilities are the same: for instance $\mathbf{P}(R_1R_2) = \frac{2!4!}{6!}$.

There are 12 distinct pairs of different pattern cups and all 12 probabilities are the same, like $\mathbf{P}(R_1W_2) = \frac{2 \cdot 2 \cdot 4!}{6!}$. Hence

$$\mathcal{S}_2 = 3 \cdot \frac{2!4!}{6!} + 12 \cdot \frac{2 \cdot 2 \cdot 4!}{6!}.$$

Computation of $\mathcal{S}_3$. There are 12 triplets with 1 pair of the same pattern and single other cup, and all 12 probabilities are the same, like $\mathbf{P}(R_1R_2S_1) = \frac{2!2 \cdot 3!}{6!}$.

There are 8 distinct triplets with all 3 cups having different pattern, and all 8 probabilities are the same, like $\mathbf{P}(R_1W_2S_1) = \frac{2 \cdot 2 \cdot 2 \cdot 3!}{6!}$. Hence

$$\mathcal{S}_3 = 12 \cdot \frac{2!2 \cdot 3!}{6!} + 8 \cdot \frac{2 \cdot 2 \cdot 2 \cdot 3!}{6!}.$$

Computation of $\mathcal{S}_4$. There are 3 distinct groups of 4 consisting of two pairs, all 3 probabilities are the same, like $\mathbf{P}(R_1R_2S_1S_2) = \frac{2!2!2!}{6!}$.

There are 12 distinct groups of 4 consisting of one pair and two single cups of different pattern, all 12 probabilities are the same, like $\mathbf{P}\left(R_1 R_2 S_1 W_2\right) = \frac{2!2 \cdot 2 \cdot 2!}{6!}$. Hence

$$\mathcal{S}_4 = 3 \cdot \frac{2!2!2!}{6!} + 12 \cdot \frac{2!2 \cdot 2 \cdot 2!}{6!}.$$

Computation of $\mathcal{S}_5$. There are 6 distinct groups of 5 consisting of two pairs and a single cup of different pattern, all 6 probabilities are the same, like $\mathbf{P}\left(R_1 R_2 S_1 S_2 W_1\right) = \frac{2!2!2}{6!}$:

$$\mathcal{S}_5 = 6 \cdot \frac{2!2!2}{6!}.$$

Finally,

$$\mathcal{S}_6 = \mathbf{P}\left(R_1 R_2 W_1 W_2 S_1 S_2\right) = \frac{2!2!2!}{6!}$$

Thus

$$\begin{aligned}
& \mathbf{P}\,(\text{at least one match}) \\
= \;& 6 \cdot \frac{2 \cdot 5!}{6!} - \left(3 \cdot \frac{2!4!}{6!} + 12 \cdot \frac{2 \cdot 2 \cdot 4!}{6!}\right) \\
& + \left(12 \cdot \frac{2!2 \cdot 3!}{6!} + 8 \cdot \frac{2 \cdot 2 \cdot 2 \cdot 3!}{6!}\right) \\
& - \left(3 \cdot \frac{2!2 \cdot 2!}{6!} + 12 \cdot \frac{2!2 \cdot 2 \cdot 2!}{6!}\right) + 6 \cdot \frac{2!2!2}{6!} - \frac{2!2!2!}{6!} \\
= \;& \frac{8}{9}, \mathbf{P}\,(\text{no match}) = 1 - \frac{8}{9} = \frac{1}{9}.
\end{aligned}$$

11. You choose k of the first n positive integers, and a lottery chooses a random subset L of the same size (k numbers in it). What is the probability that:

(a) L includes no consecutive integers? *Hint about counting nonconsecutive integers*. Let $n = 7, k = 3$. (i) Put $7 - 3 = 4$ white balls in a row with spaces between them, in the beginning and at the end (there are 5 spaces). (ii) Choose 3 spaces and put 3 black balls there. Number all balls from the left to the right. For instance, if the 2nd, fourth and fifth space were chosen, then L consists of the numbers $2, 5, 7$. Realize that number of ways to have non consecutive integers in L equals to the number of ways to choose 3 spaces among 5 available.

Answer. As a sample space Ω take all possible groups of k numbers from $\{1, 2, \ldots, n\}$: $\#\Omega = \binom{n}{k}$. Let A="L includes no consecutive integers". In order to count outcomes in A, imagine we put $n-k$ white balls in a row with spaces between them, also one space before the first ball and one space after the last ball: there are $n - k + 1$ spaces. Choose k spaces and place black balls there: there are $\binom{n-k+1}{k}$ different ways to do it. Write the numbers $1, 2, \ldots, n$ from the left to the right on the balls in the row. The numbers written on the black balls are numbers in L (no consecutive integers). Thus $\#A = \binom{n-k+1}{k}$, and

$$\mathbf{P}\,(A) = \frac{\binom{n-k+1}{k}}{\binom{n}{k}}.$$

Comment. We were choosing a group of k numbers. The counting of the outcomes in A goes with a row of the balls because we are choosing from the set of numbers $\{1, 2, 3, \ldots, n\}$ that has an order (one number is bigger than the other and A cannot contain consecutive numbers).

(b) L includes exactly one pair of consecutive integers? Hint. Like (a) but think about one pair of consecutive integers as one entity.

Answer. The sample space Ω is the same as in (a): $\#\Omega = \binom{n}{k}$. Let $A =$"L includes exactly one pair of consecutive integers". In order to count the outcomes in A, imagine we put $n-k$ white balls in a row with spaces between them, also one space before the first ball and one space after the last ball: there are $n-k+1$ spaces. Take $k-2$ black balls and a narrow box with two black balls. First put the box into one of the spaces: there are $n-k+1$ ways to do it. Then choose $k-2$ spaces from the remaining $n-k$ spaces and put the black balls there: there are $\binom{n-k}{k-2}$ different ways. Write the numbers from 1 to n on the balls including those in the box. The k numbers on the black balls are the numbers in L (there is exactly one pair of consecutive numbers). Thus

$$\#A = (n-k+1)\binom{n-k}{k-2} = (k-1)\binom{n-k+1}{k-1}$$

and

$$\mathbf{P}(A) = \frac{(k-1)\binom{n-k+1}{k-1}}{\binom{n}{k}}.$$

Question. If you play lottery, (a) or (b) strategy would be better?

(c) the numbers in L are drawn in increasing order? Hint. Any ordering of k numbers is equally likely.

Answer. Since any ordering of k numbers is equally likely, the probability is $1/k!$.

(d) your choice of numbers is the same as L?

Answer. The probability is, obviously,

$$\frac{1}{\binom{n}{k}}.$$

(e) there are exactly l of your numbers matching members of L?

Answer. In your collection of k numbers there are $\binom{k}{l}$ different groups of size l that could match with l numbers in L; you particular group of l numbers must be in L but the remaining $k-l$ numbers in L do not match; you remove from $\{1, 2, \dots, n\}$ the whole your group of k numbers and choose $k-l$ numbers from the remaining $n-k$ numbers in $\binom{n-k}{k-l}$ different ways. The probability is

$$\frac{\binom{k}{l}\binom{n-k}{k-l}}{\binom{n}{k}}.$$

12. 10% of the surface of a sphere is colored blue, the rest is red. Show that it is possible to inscribe a cube in S with all its vertices red. Hint: select a random inscribed cube and think what is the probability that a kth vertex is blue. Then estimate from above the probability that at least one vertex is blue.

Answer. The inscribed cube has 8 vertices. Let B_k be the event that kth vertex is blue, $1 \le k \le 8$. Then

$$\mathbf{P}(\text{at least one vertex is blue}) = \mathbf{P}\left(\cup_{k=1}^{8} B_k\right) \le \sum_{k=1}^{8} \mathbf{P}(B_k) \le 8 \cdot 0.1 = 0.8.$$

Hence $\mathbf{P}(\text{all vertices red}) = 1 - \mathbf{P}(\text{at least one vertex is blue}) \ge 0.2$.

2 Week 2: 8/29- 9/2

1. A school offers three language classes: Spanish (S), French (F), and German (G). There are 100 students total, of which 28 take S, 26 take F, 16 take G, 12 take both S and F, 4 take both S and G, 6 take both F and G, and 2 take all three languages.

(1) Compute the probability that a randomly selected student (a) is not taking any of the three language classes (hint: inclusion/exclusion); (b) takes exactly one of the three language classes.

(2) Compute the probability that, of two randomly selected students, at least one takes a language class.

Answer. (a) We compute firs, using inclusion/exclusion,

$$\begin{aligned} & \mathbf{P}\left(\text{takes at least one of language classes}\right) \\ = & \mathbf{P}\left(S \cup F \cup G\right) = \mathbf{P}\left(S\right) + \mathbf{P}\left(F\right) + \mathbf{P}\left(G\right) \\ & -\mathbf{P}\left(SF\right) - \mathbf{P}\left(FG\right) - \mathbf{P}\left(SG\right) + \mathbf{P}\left(SFG\right) \\ = & 0.28 + 0.26 + 0.16 - 0.12 - 0.04 - 0.06 + 0.02 = 0.5 \end{aligned}$$

Hence $\mathbf{P}\left(\text{not taking any of the three language classes}\right) = 1 - 0.5 = 0.5$.

(b) Looking at the Venn's diagram,

$$\begin{aligned} \mathbf{P}\left(\text{only F}\right) &= \mathbf{P}\left(F\right) - \mathbf{P}\left(SF\right) - \mathbf{P}\left(SG\right) + \mathbf{P}\left(SFG\right) \\ &= 0.26 - 0.12 - 0.06 + 0.02 = 0.1, \\ \mathbf{P}\left(\text{only S}\right) &= 0.28 - 0.12 - 0.04 + 0.02 = 0.14, \\ \mathbf{P}\left(\text{only G}\right) &= 0.16 - 0.04 - 0.06 + 0.02 = 0.08. \end{aligned}$$

Thus

$$\mathbf{P}\left(\text{exactly one}\right) = 0.1 + 0.14 + 0.08 = 0.32.$$

(2) Compute the probability that, of two randomly selected students, at least one takes a language class.

Answer. There are 50 students taking at least one of three language classes, and remaining 50 doe not take any. By inclusion/exclusion,

$$\begin{aligned} & \mathbf{P}\left(\text{at least one of them takes a language class}\right) \\ = & \mathbf{P}\left(\text{1st takes}\right) + \mathbf{P}\left(\text{2nd takes}\right) - \mathbf{P}\left(\text{both take}\right) \\ = & 0.5 + 0.5 - \frac{\binom{50}{2}}{\binom{100}{2}} = \frac{149}{198} = 0.752\,53. \end{aligned}$$

Also, by counting directly:

$$\begin{aligned} & \mathbf{P}\left(\text{at least one of them takes a language class}\right) \\ = & \frac{\binom{50}{2} + \binom{50}{1}\binom{50}{1}}{\binom{100}{2}} = \frac{149}{198} == 0.752\,53. \end{aligned}$$

2. A rare disease affects one person in 10^3. A test for the disease shows positive with probability 0.99 when applied to an ill person, and with probability 0.01 when applied to a healthy person. What is the probability that you have the disease given that the test shows negative?

Answer. Notation: $A =$ "you have a disease", $B =$ "test shows negative". By Bayes,

$$\begin{aligned}\mathbf{P}(A|B) &= \frac{\mathbf{P}(B|A)\mathbf{P}(A)}{\mathbf{P}(B|A)\mathbf{P}(A)+\mathbf{P}(B|A^c)\mathbf{P}(A^c)} \\ &= \frac{0.01\cdot 0.001}{0.01\cdot 0.001+0.99\cdot 0.999} = 1.0111\times 10^{-5}.\end{aligned}$$

3. English and American spellings are *rigour* and *rigor*, respectively. At a certain hotel, 40% of guests are from England and the rest are from America. A guest at the hotel writes the word (as either *rigour* or *rigor*), and a randomly selected letter from that word turns out to be a vowel. Compute the probability that the guest is from England.

Answer. Denoting $B =$ "selected letter is a vowel", $A =$ "guest from England", by Bayess,

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(B|A)\mathbf{P}(A)}{\mathbf{P}(B)}.$$

By total probability law,

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(B|A)\mathbf{P}(A)+\mathbf{P}\left(B|A^c\right)\mathbf{P}\left(A^c\right) \\ &= \frac{3}{6}\cdot 0.4+\frac{2}{5}\cdot 0.6 = 0.44.\end{aligned}$$

So,

$$\mathbf{P}(A|B) = \frac{\frac{3}{6}\cdot 0.4}{0.44} = 0.45455.$$

4. In a certain community, 36% of all the families have a dog and 30% have a cat. Of those families with a dog, 22% also have a cat. Compute the probability that a randomly selected family

(a) has both a dog and a cat;

(b) has a dog given that it has a cat.

Hint. Interpret 22% as conditional probability.

(a) has both a dog and a cat;

Answer. Notation: $D =$" family has a dog", $C =$"family has a cat". We have $\mathbf{P}(D) = 0.36, \mathbf{P}(C) = 0.3, \mathbf{P}(C|D) = 0.22$.

By multiplication rule, $0.22\cdot 0.36 = 0.0792$

$$\mathbf{P}(DC) = \mathbf{P}(C|D)\mathbf{P}(D) = 0.22\cdot 0.36 = 0.0792.$$

(b) has a dog given that it has a cat.

Answer. By definition,

$$\mathbf{P}(D|C) = \frac{\mathbf{P}(CD)}{\mathbf{P}(C)} = \frac{0.0792}{0.3} = 0.264.$$

5. (Galton's paradox) You flip three fair coins. At least two are alike, and it is an even chance that the third is a head or a tail. Therefore $\mathbf{P}(\text{all alike})=\frac{1}{2}$. Do you agree? (all alike means all heads or all tails).

Answer. No. The sample space Ω of all words of length 3 made of H and T: $\#\Omega = 2^3 = 8$ and all outcomes are equally likely. Since "all alike"$=\{HHH, TTT\}$,

$$\mathbf{P}(\text{all alike}) = \frac{2}{2^3} = \frac{1}{4}.$$

6. The event A is said to be repelled by the event B if $\mathbf{P}(A|B) < \mathbf{P}(A)$, and to be attracted by B if $\mathbf{P}(A|B) > \mathbf{P}(A)$. Show that if B attracts A, then A attracts B, and B^c repels A. If A attracts B, and B attracts C, does A attract C?

Answer. If B attracts A, then

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} > \mathbf{P}(A). \tag{2.2}$$

Hence

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} > \mathbf{P}(B),$$

i.e., A attracts B. On the other hand, by (2.2),

$$\mathbf{P}\left(A|B^c\right) = \frac{\mathbf{P}(A \cap B^c)}{\mathbf{P}(B^c)} = \frac{\mathbf{P}(A) - \mathbf{P}(A \cap B)}{\mathbf{P}(B^c)} < \mathbf{P}(A),$$

i.e., B^c repells A.

If A attracts B, and B attracts C, does A attract C? Not necessarily. For example, a fair die is rolled, A=" score is 3 or 4 or 5", B=" score is ≥ 4", C=" score is even". Then

$$\begin{aligned}
\mathbf{P}(B|A) &= \frac{2}{3} > \mathbf{P}(B) = \frac{1}{2}, \\
\mathbf{P}(C|B) &= \frac{2}{3} > \mathbf{P}(C) = \frac{1}{2}, \\
\mathbf{P}(C|A) &= \frac{1}{3} < \frac{1}{2} = \mathbf{P}(C).
\end{aligned}$$

7. Calculate the probability that a hand of 13 cards dealt from a normal shuffled pack of 52 contains exactly two kings and one ace. What is the probability that it contains exactly one ace given that it contains exactly two kings?

Answer. $\mathbf{P}(\text{exactly one ace and two kings}) = \frac{\binom{4}{2}\binom{4}{1}\binom{44}{10}}{\binom{52}{13}} = \frac{301\,587}{3215\,975} = 9.377\,8 \times 10^{-2}$.

$\mathbf{P}(\text{exactly one ace}|\text{ exactly two kings}) = \frac{\binom{4}{2}\binom{4}{1}\binom{44}{10}}{\binom{4}{2}\binom{48}{11}} = \frac{2849}{6486} = 0.439\,25$.

8. A woman has n keys, of which two will open her door. (a) If she tries the keys at random, discarding those that do not work,what is the probability that she will open the door on her kth try? (b) What if she does not discard previously tried keys? What is the probability of no right key in k tries ($k \geq 1$)? If $n = 9$, how many tries are needed to be 90% sure that the door is opened?

9. Three prisoners are informed by their jailer that one of them has been chosen at random to be executed and the other two are to be freed. Prisoner A asks the jailer to tell him privately which of his fellow prisoners will be set free, claiming that there would be no harm in divulging this information because he already knows that at least one of the two will go free. The jailer refuses to answer the question, pointing out that if A knew which of his fellow prisoners were to be set free, then his own probability of being executed would rise from 1/3 to 1/2 because he would then be one of two prisoners. What do you think of the jailer's reasoning?

Answer. It is the "same" as the class problem about an award behind 3 closed doors. Consider A="death sentence for A", B="death sentence for B", C="death sentence for C",

F_B=" jailer says B goes free", F_C =" jailer says C goes free".

We know $P(A) = P(B) = P(C) = \frac{1}{3}$, $P(F_B|A) = P(F_C|A) = \frac{1}{2}$, $P(F_B|B) = 0$, $P(F_C|B) = 1$, $P(F_B|B) = 0$, $P(F_C|C) = 0$, $P(F_B|C) = 1$.

Jailer is wrong: by Bayes,

$$\begin{aligned} P(A|F_B) &= \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{2}\frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{1}{3} = \mathbf{P}(A), \\ P(A|F_C) &= \frac{1}{3} = \mathbf{P}(A). \end{aligned}$$

As we found in class with the award, A and F_B, also A and F_C, are independent.

10. Some form of prophylaxis is said to be 90 per cent effective at prevention during one year's treatment. If the degrees of effectiveness in different years are independent, show that the treatment is more likely than not to fail within 7 years.

Answer. Since effectiveness in different years are independent, the treatment is effective 7 years in a row with probability $0.9^7 = 0.478\,30 < 50\%$.

11. The color of a person's eyes is determined by a single pair of genes. If they are both blue-eyed genes, then the person will have blue eyes; if they are both brown-eyed genes, then the person will have brown eyes; and if one of them is a blue-eyed gene and the other a brown-eyed gene, then the person will have brown eyes. (Because of the latter fact, we say that the brown-eyed gene is dominant over the blue-eyed one.) A newborn child independently receives one eye gene from each of its parents, and the gene it receives from a parent is equally likely to be either of the two eye genes of that parent. Suppose that Smith and both of his parents have brown eyes, but Smith's sister has blue eyes.

(a) What is the probability that Smith possesses a blueeyed gene? Hint: "Smith and both of his parents have brown eyes, but Smith's sister has blue eyes" means both Smith's parents had mixed genes.

Answer. "Smith and both of his parents have brown eyes, but Smith's sister has blue eyes" means both Smith's parents had mixed genes. Considfer the following events:

A="Smith has blue gene", B="his eyes are brown", Bl_1="blue gene from father", Bl_2="blue gene from mother", Br_1="brown from father", Br_2="brown gene from mother".

Then

$$\begin{aligned} P(A|B) &= \frac{P(AB)}{P(B)} = \frac{P(Bl_1 Br_2 \cup Bl_2 Br_1)}{P(Bl_1 Br_2 \cup Bl_2 Br_1 \cup Br_1 Br_2)} \\ &= \frac{\frac{1}{4} + \frac{1}{4}}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \frac{2}{3}; \end{aligned}$$

note that $P(B) = \frac{3}{4}$, $P(AB) = \frac{1}{2}$.

(b) Suppose that Smith's wife has blue eyes. What is the probability that their first child will have blue eyes?

Answer. Let C_1 =”1st child has blue eyes”, B_1=”1st child brown eyes”. Then

$$\begin{aligned}
P(C_1|B) &= \frac{P(C_1 B)}{P(B)} = \frac{P(C_1 BA)}{P(B)} = \frac{P(C_1|BA)P(A|B)P(B)}{P(B)} \\
&= P(C_1|BA)P(A|B) = \frac{1}{2}\frac{2}{3} = \frac{1}{3}, \\
P(B_1|B) &= 1 - P(C_1|B) = \frac{2}{3}.
\end{aligned}$$

(c) If their first child has brown eyes, what is the probability that their next child will also have brown eyes?

Answer. Let B_2=”2nd child brown eyes”. Then

$$\begin{aligned}
P(B_2|B_1 B) &= \frac{P(B_2 B_1 B)}{P(B_1 B)} = \frac{P(B_2 B_1 BA) + P(B_2 B_1 BA^c)}{P(B_1 B)} \\
&= \frac{P(B_2 B_1|BA)P(BA) + P(B_2 B_1|BA^c)P(BA^c)}{P(B_1|B)P(B)} \\
&= \frac{(\frac{1}{2}\frac{1}{2})\frac{1}{2} + 1 \cdot \frac{1}{4}}{\frac{2}{3}\frac{3}{4}} = \frac{3}{4}.
\end{aligned}$$

12. Three fair dice have different colors: red, blue, and yellow. These three dice are rolled and the face value of each is recorded as R, B, Y, respectively.

(a) Compute the probability that $B < Y < R$, given all the numbers are different;

Answer. Let A = ”all different”, D = ”$B < Y < R$”. We find $\#A = 6 \cdot 5 \cdot 4 = 120$, and $\#D = \binom{6}{3} = 20$. Hence

$$\mathbf{P}(D|A) = \frac{\#D}{\#A} = \frac{20}{120} = \frac{1}{6} = \frac{1}{3!}.$$

(b) Compute the probability that $B < Y < R$.

Answer. The sample space $\Omega = \{(i, j, k) : 1 \le i, j, k \le 6\}$, the number of outcomes of 3 experiments with 3 outcomes each is $\#\Omega = 6^3$,and $\#D = \binom{6}{3} = 20$. Hence,

$$\mathbf{P}(D) = \frac{\#D}{\#\Omega} = \frac{20}{6^3} = \frac{5}{54}.$$

13. Let $\Omega = \{1, 2, \ldots, 13\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, the set of all subsets, and all outcomes are equally likely: $\mathbf{P}(A) = \frac{\#A}{13}$, $A \in \mathcal{F}$. Show that if A, B are independent, then at least one of them is either $\emptyset$ or Ω. Hint. For any $C \subseteq \Omega$, $\#C \le 13$, and 13 is a prime number.

Answer. Let A and B be independent. By definition of independence,

$$\frac{\#(A \cap B)}{13} = \frac{\#A}{13}\frac{\#B}{13}$$

or

$$13 \cdot \#(A \cap B) = (\#A) \cdot (\#B).$$

Since 13 is a prime number, this can happen iff $\#A = 0$ or $\#B = 0$ (which means at least one of them is empty) or $\#A = \#B = 13$ (which means $A = B = \Omega$).

Comment. In the considered probablity space $(\Omega, \mathcal{F}, \mathbf{P})$, there are non trivial independent events at all.

14. (Conditioning of the conditional probability) Recall conditional probability is a probability. For the events A, B, C with $\mathbf{P}(B \cap C) > 0$, denote $\tilde{\mathbf{P}}(A)$ the conditional probability $\mathbf{P}(A|B)$:

$$\tilde{\mathbf{P}}(A) = \mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Show that

a)

$$\tilde{\mathbf{P}}(A|C) = \mathbf{P}(A|BC) = \mathbf{P}(A|B \cap C).$$

Answer. By definition,

$$\tilde{\mathbf{P}}(A|C) = \frac{\tilde{\mathbf{P}}(AC)}{\tilde{\mathbf{P}}(C)} = \frac{\mathbf{P}(AC|B)}{\mathbf{P}(C|B)} = \frac{\frac{\mathbf{P}(ACB)}{\mathbf{P}(B)}}{\frac{\mathbf{P}(BC)}{\mathbf{P}(B)}} = \frac{\mathbf{P}(ACB)}{\mathbf{P}(CB)} = \mathbf{P}(A|BC).$$

b)

$$\mathbf{P}(C|AB) = \frac{\mathbf{P}(BC|A)}{\mathbf{P}(B|A)}.$$

If B and C are conditionally independent ($\mathbf{P}(BC|A) = \mathbf{P}(B|A)\mathbf{P}(C|A), \mathbf{P}(BC|A^c) = \mathbf{P}(B|A^c)\mathbf{P}(C|A^c)$) , then

$$\mathbf{P}(C|AB) = \mathbf{P}(C|A), \mathbf{P}\left(C|A^c B\right) = \mathbf{P}\left(C|A^c\right).$$

Answer. By definition and multiplication rule,

$$\mathbf{P}(C|AB) = \frac{\mathbf{P}(BCA)}{\mathbf{P}(BA)} = \frac{\mathbf{P}(BC|A)\mathbf{P}(A)}{\mathbf{P}(B|A)\mathbf{P}(A)} = \frac{\mathbf{P}(BC|A)}{\mathbf{P}(B|A)}.$$

If B and C are conditionally independent, then

$$\mathbf{P}(C|AB) = \frac{\mathbf{P}(BC|A)}{\mathbf{P}(B|A)} = \frac{\mathbf{P}(B|A)\mathbf{P}(C|A)}{\mathbf{P}(B|A)} = \mathbf{P}(C|A).$$

c)

$$\begin{aligned} \mathbf{P}(C|B) &= \mathbf{P}(CA|B) + \mathbf{P}\left(CA^c|B\right) = \\ &= \mathbf{P}(C|AB)\mathbf{P}(A|B) + \mathbf{P}\left(C|A^c B\right)\mathbf{P}\left(A^c|B\right) \end{aligned}$$

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(C|AB)\mathbf{P}(A|B)}{\mathbf{P}(C|AB)\mathbf{P}(A|B) + \mathbf{P}(C|A^c B)\mathbf{P}(A^c|B)}.$$

If B and C are conditionally independent ($\mathbf{P}(BC|A) = \mathbf{P}(B|A)\mathbf{P}(C|A), \mathbf{P}(BC|A^c) = \mathbf{P}(B|A^c)\mathbf{P}(C|A^c)$) , then (using b))

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(C|A)\mathbf{P}(A|B)}{\mathbf{P}(C|A)\mathbf{P}(A|B) + \mathbf{P}(C|A^c)\mathbf{P}(A^c|B)}.$$

Answer. By definition,

$$\begin{aligned}\mathbf{P}(A|BC) &= \frac{\mathbf{P}(ABC)}{\mathbf{P}(BC)} \\ &= \frac{\mathbf{P}(C|AB)\,\mathbf{P}(A|B)\,\mathbf{P}(B)}{\mathbf{P}(C|AB)\,\mathbf{P}(A|B)\,\mathbf{P}(B)+\mathbf{P}(C|A^cB)\,\mathbf{P}(A^c|B)\,\mathbf{P}(B)} \\ &= \frac{\mathbf{P}(C|AB)\,\mathbf{P}(A|B)}{\mathbf{P}(C|AB)\,\mathbf{P}(A|B)+\mathbf{P}(C|A^cB)\,\mathbf{P}(A^c|B)}\end{aligned}$$

If B and C are conditionally independent ($\mathbf{P}(BC|A) = \mathbf{P}(B|A)\,\mathbf{P}(C|A)\,, \mathbf{P}(BC|A^c) = \mathbf{P}(B|A^c)\,\mathbf{P}(C|A^c)$) , then (using b))

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(C|A)\,\mathbf{P}(A|B)}{\mathbf{P}(C|A)\,\mathbf{P}(A|B)+\mathbf{P}(C|A^c)\,\mathbf{P}(A^c|B)}.$$

d) Also,

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(BC|A)\,\mathbf{P}(A)}{\mathbf{P}(BC|A)\,\mathbf{P}(A)+\mathbf{P}(BC|A^c)\,\mathbf{P}(A^c)}.$$

If B and C are conditionally independent ($\mathbf{P}(BC|A) = \mathbf{P}(B|A)\,\mathbf{P}(C|A)\,, \mathbf{P}(BC|A^c) = \mathbf{P}(B|A^c)\,\mathbf{P}(C|A^c)$) , then we have

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(B|A)\,\mathbf{P}(C|A)\,\mathbf{P}(A)}{\mathbf{P}(B|A)\,\mathbf{P}(C|A)\,\mathbf{P}(A)+\mathbf{P}(B|A^c)\,\mathbf{P}(C|A^c)\,\mathbf{P}(A^c)}$$

Answer. Indeed,

$$\mathbf{P}(A|BC) = \frac{\mathbf{P}(ABC)}{\mathbf{P}(BC)} = \frac{\mathbf{P}(BC|A)\,\mathbf{P}(A)}{\mathbf{P}(BC|A)\,\mathbf{P}(A)+\mathbf{P}(BC|A^c)\,\mathbf{P}(A^c)}.$$

3 Week 3: 9/7- 9/9

1. (1st step analysis) Independent trials, consisting of rolling a pair of fair dice, are performed. What is the probability that an outcome of 5 appears before an outcome of 7 when the outcome of a roll is the sum of the dice? Hint. Let $A =$"5 appears before 7". Apply total probability law to the partition $\Omega = B_1 \cup B_2 \cup B_3$, where B_1= "the first trial results in a 5", B_2= "1st trial results in a 7", B_3= "first trial results in neither a 5 nor a 7". Then solve equation for $\mathbf{P}(A)$.

Answer. Let $A =$"5 appears before 7". We apply total probability law to the partition $\Omega = B_1 \cup B_2 \cup B_3$, where B_1= "the first trial results is a 5", B_2= "1st trial results is a 7", B_3= "first trial results is neither a 5 nor a 7":

$$\mathbf{P}(A) = \mathbf{P}(A|B_1)\,\mathbf{P}(B_1) + \mathbf{P}(A|B_2)\,\mathbf{P}(B_2) + \mathbf{P}(A|B_3)\,\mathbf{P}(B_3).$$

Since $B_1 = \{(1,4),(2,3),(3,2),(4,1)\}$, $B_2 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, we find by counting

$$\mathbf{P}(B_1) = \frac{4}{36} = \frac{1}{9}, \mathbf{P}(B_2) = \frac{6}{36} = \frac{1}{6}, \mathbf{P}(B_3) = 1 - \frac{1}{9} - \frac{1}{6} = \frac{13}{18}.$$

Obviously, $\mathbf{P}(A|B_1) = 1$, $\mathbf{P}(A|B_2) = 0$. Since the trials are independent, $\mathbf{P}(A|B_3) = \mathbf{P}(A)$. Hence, by total probability law above,

$$\mathbf{P}(A) = (1)\,\frac{1}{9} + \mathbf{P}(A)\,\frac{13}{18}.$$

Solving for $\mathbf{P}(A)$:

$$\mathbf{P}(A) = \frac{18}{5}\frac{1}{9} = \frac{2}{5}.$$

2. (1st step analysis) a) Two players A,B take turns to roll a die; they do this in the order ABAB...

(i) Find the probability that, A is the first to throw a 6;

(ii) Find the probability that the first 6 to appear is thrown by A, the second 6 to appear is thrown by B.

Hint for both: (i), (ii). Total probability with coditioning with respect to D_1=" 6 shows up in the first roll", D_2 =" 6 does not show up in the first two rolls".

Answer. (i) Let E = "A is the 1st", D_1=" 6 shows up in the first roll", D_2 =" 6 does not show up in the first two rolls". Total probability formula,

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(E|D_1)\,\mathbf{P}(D_1) + \mathbf{P}(E|D_2)\,\mathbf{P}(D_2) \\ &= 1 \cdot \mathbf{P}(D_1) + \mathbf{P}(E)\,\mathbf{P}(D_2) = \frac{1}{6} + \mathbf{P}(E)\left(\frac{5}{6}\right)^2.\end{aligned}$$

Solving, for $\mathbf{P}(E)$ we find

$$\mathbf{P}(E) = \frac{36}{6 \cdot 11} = \frac{6}{11}.$$

(ii) Let E = "1st by A, 2nd by B", D_1=" 6 shows up in the first roll", D_2 =" 6 does not show up in the first two rolls". Total probability formula, and by (i) above,

$$\begin{aligned}
\mathbf{P}(E) &= \mathbf{P}(E|D_1)\,\mathbf{P}(D_1) + \mathbf{P}(E|D_2)\,\mathbf{P}(D_2) \\
&= \mathbf{P}(\text{B starts and the ist 6 by him})\,\frac{1}{6} + \mathbf{P}(E)\,(5/6)^2 \\
&= \frac{6}{11}\cdot\frac{1}{6} + \mathbf{P}(E)\,(5/6)^2\,,
\end{aligned}$$

and solving, for $\mathbf{P}(E)$ we find

$$\mathbf{P}(E) = \frac{(6/11)\,(1/6)}{1-(5/6)^2} = \frac{36}{121}.$$

b) Three players A,B,C take turns to roll a die; they do this in the order ABCABC...

(i) Show that the probability that, of the three players, A is the first to throw a 6, B the second, and C the third, is $216/1001$.

Hint. Total probability with coditioning with respect to D_1 ="6 shows up in the first and in the 2nd roll", D_2 ="6 in the first but not in the second or 3rd", D_3=" no 6 the first three rolls"

(ii) Show that the probability that the first 6 to appear is thrown by A, the second 6 to appear is thrown by B, and the third 6 to appear is thrown by C, is $46656/753571$.

Answer. (i) Let E = "A is the 1st, B the second, and C the third", D_1 ="6 shows up in the first and in the 2nd roll", D_2 ="6 in the first but not in the second or 3rd", D_3=" no 6 the first three rolls". By total probability formula,

$$\begin{aligned}
\mathbf{P}(E) &= \mathbf{P}(E|D_1)\,\mathbf{P}(D_1) + \mathbf{P}(E|D_2)\,\mathbf{P}(D_2) + \mathbf{P}(E|D_3)\,\mathbf{P}(D_3) \\
&= 1\cdot\mathbf{P}(D_1) + \mathbf{P}(\text{B first, C the second among B and C only})\,\mathbf{P}(D_2) \\
&\quad + \mathbf{P}(E)\,\mathbf{P}(D_3)\,.
\end{aligned}$$

By part a), $\mathbf{P}(\text{B first, C the second among B and C only}) = 6/11$, and

$$\mathbf{P}(E) = \left(\frac{1}{6}\right)^2 + \frac{6}{11}\cdot\frac{1}{6}\left(\frac{5}{6}\right)^2 + \mathbf{P}(E)\left(\frac{5}{6}\right)^3 = \frac{216}{1001}.$$

(ii) Show that the probability that the first 6 to appear is thrown by A, the second 6 to appear is thrown by B, and the third 6 to appear is thrown by C, is $46656/753571$.

Answer. Let E = "1st by A, the second by B, and the third by C". Then

$$\mathbf{P}(E) = \mathbf{P}(\text{3rd by C}|\text{1st by A, second by B})\,\mathbf{P}(\text{second by B}|\text{1st by A})\,\mathbf{P}(\text{1st by A})\,.$$

Now, directly,

$$\mathbf{P}(\text{1st by A}) = \sum_{k=0}^{\infty}(5/6)^{(3k+1)-1}\,(1/6) = \frac{36}{91};$$

alternatively, using 1st step analysis,

$$\begin{aligned}
\mathbf{P}(\text{1st by A}) &= \mathbf{P}(\text{1st by A}|\text{6 in 1st roll})\,\mathbf{P}(\text{6 in first roll}) \\
&\quad + \mathbf{P}(\text{1st by A}|\text{ no 6 in the first three rolls})\,\mathbf{P}(\text{no 6 in the first three rolls}) \\
&= 1\cdot\frac{1}{6} + \mathbf{P}(\text{1st by A})\left(\frac{5}{6}\right)^3,
\end{aligned}$$

and solving $\mathbf{P}$ (1st by A) $= \frac{36}{91}$. Also,

$$\begin{aligned}\mathbf{P}(\text{second by B}|\text{1st by A}) &= \mathbf{P}(\text{B starts and 1st by him}) = \frac{36}{91},\\ \mathbf{P}(\text{3rd by C}|\text{1st by A, second by B}) &= \mathbf{P}(\text{C starts and 1st by him}) = \frac{36}{91}.\end{aligned}$$

Thus $\mathbf{P}(E) = \left(\frac{36}{91}\right)^3 = \frac{46\,656}{753\,571}. = 6.191\,3 \times 10^{-2}$

3. Let $A_k, k \geq 1$, be a sequence of events. Show that

$$\mathbf{P}\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbf{P}(A_n).$$

Answer. By finite subaditivity property, for any events $A_1, A_2, \ldots$ and any m,

$$\mathbf{P}\left(\cup_{n=1}^{m} A_n\right) \leq \sum_{n=1}^{m} \mathbf{P}(A_n).$$

The sequence $\cup_{n=1}^{m} A_n, m \geq 1$, is increasing and $\lim_m \cup_{n=1}^{m} A_n = \cup_{n=1}^{\infty} A_n$. By continuity of probability,

$$\mathbf{P}\left(\cup_{n=1}^{\infty} A_n\right) \leq \lim_m \sum_{n=1}^{m} \mathbf{P}(A_n) = \sum_{n=1}^{\infty} \mathbf{P}(A_n).$$

4. Let $A_k, k \geq 1$, be events such that $\mathbf{P}(A_k) = 1$ for all k. Show that $\mathbf{P}\left(\cap_{k=1}^{\infty} A_k\right) = 1$.

5. At least one of the events $A_k, 1 \leq k \leq n$, is certain to occur, but certainly no more than two occur. If $\mathbf{P}(A_k) = p, \mathbf{P}\left(A_k \cap A_j\right) = q, k \neq j$, show that $p \geq 1/n$ and $q \leq 2/n$.

Answer. Since at least one of the events $A_k, 1 \leq k \leq n$, is certain to occur,

$$\mathbf{P}\left(\cup_{k=1}^{n} A_k\right) = 1.$$

Since certainly no more than two occur, probability of any intersection of mor than two events is zero. Therefore, by inclusion-exclusion principle,

$$\begin{aligned}1 &= \mathbf{P}\left(\cup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbf{P}(A_k) - \sum_{i<j} \mathbf{P}\left(A_i \cap A_j\right)\\ &= np - \binom{n}{2} q = np - \frac{n(n-1)}{2} q,\end{aligned}$$

that is

$$p - \frac{n-1}{2} q = \frac{1}{n}.$$

Hence $p \geq p - \frac{n-1}{2} q = 1/n$ and

$$\begin{aligned}\frac{n-1}{2} q &= p - 1/n \leq 1 - \frac{1}{n} = \frac{n-1}{n},\\ \frac{q}{2} &\leq \frac{1}{n}, q \leq \frac{2}{n}.\end{aligned}$$

6. Let $A_1, A_2, \ldots$ be a sequence of events. Define

$$B_n = \cup_{m=n}^{\infty} A_m, C_n = \cap_{m=n}^{\infty} A_m,$$

that is $B_n =$ "at least one A_m after n happens", $C_n =$ "all $A_m, m \geq n$ happen. Note that $C_n \subseteq A_n \subseteq B_n$, the sequence B_n is decreasing and the sequence C_n is increasing with limits. We denote

$$\begin{aligned} B &= \lim_n B_n = \cap_{n=1}^{\infty} B_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m, \\ C &= \lim_n C_n = \cup_{n=1}^{\infty} C_n = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m. \end{aligned}$$

Show that

a) $B = \limsup_n A_n = \{\omega \in \Omega : \omega \in A_n$ for infinitely many values of $n\} =$ "infinitely many A_n happen" and

$$\mathbf{P}\left(\limsup_n A_n\right) = \mathbf{P}(B) = \lim_n \mathbf{P}\left(\cup_{m=n}^{\infty} A_m\right) \geq \lim_n \sup_{m \geq n} \mathbf{P}(A_m) =: \limsup_n \mathbf{P}(A_n).$$

Here $\limsup_n \mathbf{P}(A_n) = \lim_{n\to\infty} \sup_{m \geq n} \mathbf{P}(A_m)$ is the upper limit of the sequence of numbers $\mathbf{P}(A_n)$.

Recall for a sequence of numbers a_n,

$$\limsup_n a_n = \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n \geq 1} \sup_{k \geq n} a_k, \quad \liminf_n a_n = \lim_{n\to\infty} \inf_{k \geq n} a_k = \sup_{n\geq 1} \inf_{k\geq n} a_k.$$

Limit of a sequence of numbers a_n exists iff $\limsup_n a_n = \liminf_n a_n$

Answer. Indeed, $\omega \in B = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m \Leftrightarrow \omega \in \cup_{m=n}^{\infty} A_m$ for all $n \Leftrightarrow$ for each $n \geq 1$ there is $m \geq n$ so that $\omega \in A_m$.

The sequence of events $B_n = \cup_{m=n}^{\infty} A_m, n \geq 1$, is decreasing, and by definition

$$\limsup_n A_n = \lim_n B_n = \cap_{n=1}^{\infty} B_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m = B.$$

By continuity of probability and definition of the upper limit of numbers (see (**??**)),

$$\mathbf{P}\left(\limsup_n A_n\right) = \mathbf{P}(B) = \lim_{n\to\infty} \mathbf{P}\left(\cup_{m=n}^{\infty} A_m\right) \geq \lim_{n\to\infty} \sup_{k\geq n} \mathbf{P}(A_k) = \limsup_n \mathbf{P}(A_n),$$

because for every $k \geq n$ we have $A_k \subseteq \cup_{m=n}^{\infty} A_m$, $\mathbf{P}\left(\cup_{m=n}^{\infty} A_m\right) \geq \mathbf{P}(A_k)$ and $\mathbf{P}\left(\cup_{m=n}^{\infty} A_m\right) \geq \sup_{k\geq n} \mathbf{P}(A_k)$.

b) $C = \{\omega \in \Omega : \omega \in A_n$ for all but finitely many values of $n\} =$ "all A_n except a finite number of them happen" and

$$\mathbf{P}\left(\liminf_n A_n\right) = \mathbf{P}(C) = \lim_n \mathbf{P}\left(\cap_{m=n}^{\infty} A_m\right) \leq \lim_n \inf_{m\geq n} \mathbf{P}(A_m) =: \liminf_n \mathbf{P}(A_n).$$

Answer. Indeed, $\omega \in C = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m \Leftrightarrow$ there is $n \geq 1$ so that $\omega \in A_m$ for all $m \geq n$.

The sequence of events $C_n = \cap_{m=n}^{\infty} A_m, n \geq 1$, is increasing, and by definition

$$\liminf_n A_n = \lim_n C_n = \cup_{n=1}^{\infty} C_n = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m = C.$$

By continuity of probability and definition of the lower limit of numbers (see (**??**)),

$$\mathbf{P}\left(\liminf_n A_n\right) = \mathbf{P}(C) = \lim_{n\to\infty} \mathbf{P}\left(\cap_{m=n}^{\infty} A_m\right) \le \lim_{n\to\infty} \inf_{k\ge n} \mathbf{P}(A_k) = \liminf_n \mathbf{P}(A_n) .,$$

because for every $k \ge n$ we have $\cap_{m=n}^{\infty} A_m \subseteq A_k$, $\mathbf{P}\left(\cap_{m=n}^{\infty} A_m\right) \le \mathbf{P}(A_k)$ and $\mathbf{P}\left(\cap_{m=n}^{\infty} A_m\right) \le \inf_{k\ge n} \mathbf{P}(A_k)$.

c)

$$\mathbf{P}\left(\liminf_n A_n\right) \le \mathbf{P}\left(\limsup_n A_n\right),$$

and if $\liminf_n A_n = \limsup_n A_n = A$, then

$$\mathbf{P}(A) = \lim_n \mathbf{P}(A_n).$$

Answer. Indeed for any sequence of numbers $a_n, n \ge 1$, $\liminf_n a_n \le \limsup_n a_n$. Hence by a) and b),

$$\mathbf{P}\left(\liminf_n A_n\right) \le \liminf_n \mathbf{P}(A_n) \le \limsup_n \mathbf{P}(A_n) \le \mathbf{P}\left(\limsup_n A_n\right). \tag{3.3}$$

Thus, by (3.3), if $\liminf_n A_n = \limsup_n A_n = A$, then

$$\mathbf{P}(A) = \liminf_n \mathbf{P}(A_n) = \limsup_n \mathbf{P}(A_n) = \lim_n \mathbf{P}(A_n).$$

7. A coin with $\mathbf{P}(H) = p, \mathbf{P}(T) = q = 1 - p$, is tossed repeatedly (indefinitely). Let $H_k =$ "H in the kth toss", $T_k =$ "T in the kth toss". Assume all tosses are independent.

(a) Find $\mathbf{P}$ (at least one H after n) $= \mathbf{P}\left(\cup_{m=n}^{\infty} H_m\right) = 1 - \mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right)$.

Hint. Recall, by continuity of probability, $\mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right) = \lim_{l\to\infty} \mathbf{P}\left(\cap_{m=n}^{n+l} T_m\right)$.

(b) Find probability of infinitely many H's.

Hint. $\mathbf{P}$ (infinitely many $H's$) $= \mathbf{P}\left(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} H_m\right)$: use the previous part (a) and cotinuity of probability.

8. Airlines find that each passenger who reserves a seat fails to turn up with probability $1/10$ independently of other passengers. So TW airlines always sell 10 tickets for their 9 seat airplane while BA sell 20 tickets for their 18 seat airplane. Which is more often overbooked? Hint. X, the number of passengers that show up for TW plane, and Y, the number of passengers that show up for BA plane, are binomial r.v.

Answer. Let X be the number of people among 10 that do not show up for TW flight, and Y be the number of people among 20 that fails to turn up for BA flight. Since $X \sim Bin(10, 0.1)$ and $Y \sim Bin(20, 0.1)$,

$$\begin{aligned}
\mathbf{P}(\text{TW flight overbooked}) &= \mathbf{P}(X = 0) = 0.9^{10} \approx 0.35, \\
\mathbf{P}(\text{BA flight overbooked}) &= \mathbf{P}(X \le 1) = \mathbf{P}(X = 0) + \mathbf{P}(X = 1) \\
&= 0.9^{20} + \binom{20}{1} 0.1 \cdot 0.9^{19} \approx 0.39
\end{aligned}$$

9. Eight pawns are placed randomly on the chessboard, no more than one to a square. What is the probability that

(a) they are in a straight line (do not forget the diagonals)?

Answer. The chessboard has 64 squares; Ω is any possible configuration of 8 pawns on it: $\#\Omega = \binom{64}{8}$. Let A="8 pawns in one line". Now, 8 pawns in one line could be while placed in one of 8 rows or in one of eight columns or along two diagonals: $\#A = 8 + 8 + 2 = 18$, and

$$\mathbf{P}(A) = \frac{18}{\binom{64}{8}}.$$

(b) no two are in the same row or column?

Answer. Let B="no two are in the same row or column". The first pawn can be put in the first row in 8 ways; when it is done we have 7 choices for the 2nd pawn in the second row, 6 choices for the 3rd pawn etc.: $\#B = 8 \cdot 7 \cdot \ldots . 1 = 8!$, and

$$\mathbf{P}(B) = \frac{8!}{\binom{64}{8}}.$$

10. Three coins each show heads with probability 3/5 and tails otherwise. The first coin counts 10 points for a head and 2 for a tail, the second counts 4 points for both head and tail, and the third counts 3 points for the head and 20 for a tail. You and your opponent choose a coin. You both know which is which and you cannot choose the same coin. The chosen coins are tossed and the larger score wins \$$10^{10}$. Would you prefer to be the first to pick the coin or the second? Hint. Compare the probabilities that "1st coin beats the second", "1st coin beats the 3rd", "2nd coin beats the 3rd".

Answer. We find

$$\begin{aligned}
\mathbf{P}(\text{"1st coin beats the second"}) &= \mathbf{P}(H_1) = \frac{3}{5} > \frac{1}{2}, \\
\mathbf{P}(\text{"1st coin beats the 3rd"}) &= \mathbf{P}(H_1 \text{ and } H_3) = \frac{3}{5} \cdot \frac{3}{5} = \frac{9}{25} < \frac{1}{2}, \\
\mathbf{P}(\text{"2nd coin beats the 3rd"}) &= \mathbf{P}(H_3) = \frac{3}{5} > 1/2.
\end{aligned}$$

We prefer to be the second: If opponent picks "1", we pick "3"; If opponent picks "2", we pick "1"; If opponent picks "3", we pick "2".

4 Week 4: 9/12- 9/16

1. a) Show that if F and G are distribution functions and $0 < \lambda < 1$, then $\lambda F + 1 - \lambda G$ is a distribution function. Is the product FG a distribution function? Hint. A function $F : \mathbf{R} \to [0, 1]$ is a distribution function if the rpoperties a), b), c) of Lemma (6) on p. 28 hold.

Answer. A function $H : \mathbf{R} \to [0, 1]$ is a distribution function if the properties (a), (b),(c) of Lemma (6) on p. 28 hold:

(a) Since $\lim_{x\to\infty} F(x) = \lim_{x\to\infty} G(x) = 1, \lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} G(x) = 0$, by limit laws,

$$\begin{aligned} & \lim_{x\to\infty} [\lambda F(x) + (1-\lambda) G(x)] \\ &= \lambda \lim_{x\to\infty} F(x) + (1-\lambda) \lim_{x\to\infty} G(x) = \lambda + 1 - \lambda = 1, \\ \lim_{x\to\infty} [F(x) G(x)] &= \lim_{x\to\infty} F(x) \lim_{x\to\infty} G(x) = (1)(1) = 1, \end{aligned}$$

and

$$\begin{aligned} & \lim_{x\to-\infty} [\lambda F(x) + (1-\lambda) G(x)] \\ &= \lambda \lim_{x\to-\infty} F(x) + (1-\lambda) \lim_{x\to-\infty} G(x) = 0, \\ \lim_{x\to-\infty} [F(x) G(x)] &= \lim_{x\to-\infty} F(x) \lim_{x\to-\infty} G(x) = 0. \end{aligned}$$

(b) If $x < y$, then $F(x) \le F(y)$ and $G(x) \le G(y)$. Hence

$$\begin{aligned} \lambda F(x) + (1-\lambda) G(x) &\le \lambda F(y) + (1-\lambda) G(y), \\ F(x) G(x) &\le F(y) G(y). \end{aligned}$$

(c) Since $F(x+) = F(x), G(x+) = G(x), x \in \mathbf{R}$, by limit laws with $h \downarrow 0$, for $x \in \mathbf{R}^d$,

$$\begin{aligned} \lambda F(x+h) + (1-\lambda) G(x+h) &\to \lambda F(x+) + (1-\lambda) G(x+) \\ &= \lambda F(x) + (1-\lambda) G(x), \\ F(x+h) G(x+h) &\to F(x+) G(x+) = F(x) G(x). \end{aligned}$$

Both, $\lambda F + (1-\lambda) G$ and FG are distribution functions.

b) A random variable X has distribution function F. What is the distribution function of $Y = aX + b$, where a and b are real constants?

Answer. For $a > 0$,

$$\begin{aligned} F_Y(y) &= \mathbf{P}(Y \le y) = \mathbf{P}(aX + b \le y) = \mathbf{P}\left(X \le \frac{y-b}{a}\right) \\ &= F\left(\frac{y-b}{a}\right), y \in \mathbf{R}. \end{aligned}$$

If $a = 0$, then $Y = b$, and

$$F_Y(y) = \begin{matrix} 0 & \text{if } y < b, \\ 1 & \text{if } y \ge b. \end{matrix}$$

For $a < 0$,

$$\begin{aligned} F_Y(y) &= \mathbf{P}(Y \le y) = \mathbf{P}(aX + b \le y) = \mathbf{P}\left(X \ge \frac{y-b}{a}\right) \\ &= 1 - \mathbf{P}\left(X < \frac{y-b}{a}\right) = 1 - F\left(\frac{y-b}{a}-\right), y \in \mathbf{R}. \end{aligned}$$

2. Let X have a distribution function (cdf)

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{1}{2}x & \text{if } 0 \le x \le 2, \\ 1 & \text{if } x > 2, \end{cases}$$

and let $Y = X^2$. Find

(a) $\mathbf{P}\left(\frac{1}{2} \le X \le \frac{3}{2}\right)$; (b) $\mathbf{P}(1 \le X < 2)$; (c) $\mathbf{P}(Y \le X)$; (d)$\mathbf{P}(X \le 2Y)$.

(e) the distribution function of $Z = \sqrt{X}$.

Answer. Since F is continuous (sketch the graph), (a)

$$\mathbf{P}\left(\frac{1}{2} \le X \le \frac{3}{2}\right) = F\left(\frac{3}{2}\right) - F\left(\frac{1}{2}\right) = \frac{1}{2}\frac{3}{2} - \frac{1}{2}\frac{1}{2} = \frac{1}{2},$$

(b)

$$\begin{aligned} \mathbf{P}(1 \le X < 2) &= F(2) - F(1) \\ &= 1 - \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

(c) "$Y \le X$"$\Leftrightarrow$"$X^2 - X \le 0$"$\Leftrightarrow$"$X(X-1) \le 0$"$\Leftrightarrow$"$0 \le X \le 1$" (sketch the parabola $y = x(x-1)$). Hence

$$\mathbf{P}(Y \le X) = \mathbf{P}(0 \le X \le 1) = F(1) - F(0) = \frac{1}{2}.$$

(d) "$X \le 2Y$"$\Leftrightarrow$"$X - 2X^2 \le 0$"$\Leftrightarrow$"$X(1-2X) \le 0$"$\Leftrightarrow$"$X \le 0$ or $X \ge 1/2$" (sketch the parabola $y = x(1-2x)$). Hence

$$\begin{aligned} \mathbf{P}(Y \le X) &= \mathbf{P}(X \le 0 \text{ or } X \ge 1/2) = \mathbf{P}(X \le 0) + \mathbf{P}(X \ge 1/2) \\ &= F(0) + (1 - F(1/2)) = 1 - \frac{1}{4} = \frac{3}{4}. \end{aligned}$$

(e) Note that $\mathbf{P}(X \ge 0) = 1$. Hence $\mathbf{P}(Z \ge 0) = 1$, and for $z \ge 0$,

$$F_Z(z) = \mathbf{P}(Z \le z) = \mathbf{P}\left(\sqrt{X} \le z\right) = \mathbf{P}\left(X \le z^2\right) = F\left(z^2\right).$$

Thus

$$F_Z(z) = \begin{matrix} 0 & \text{if } z < 0, \\ \frac{1}{2}z^2 & \text{if } 0 \le z \le \sqrt{2}, \\ 1 & \text{if } z > \sqrt{2}. \end{matrix}$$

3. Let X have a distribution function (cdf)

$$F(x) = \begin{cases} 0 & \text{if } x < -1, \\ 1-p & \text{if } -1 \le x < 0, \\ 1-p+\frac{1}{2}xp & \text{if } 0 \le x \le 2, \\ 1 & \text{if } x > 2. \end{cases}$$

Sketch the function and find: (a) $\mathbf{P}(X = -1)$, (b) $\mathbf{P}(X = 0)$, (c) $\mathbf{P}(X \ge 1)$.

Answer. A sketch:

(a), (b), (c)

$$\begin{aligned} \mathbf{P}(X = -1) &= F(-1) - F(-1-) = (1-p) - 0 = 1-p, \\ \mathbf{P}(X = 0) &= F(0) - F(0-) = (1-p) - (1-p) = 0, \\ \mathbf{P}(X \ge 1) &= 1 - F(1-) = 1 - F(1) = 1 - (1 - 1/2p) = \frac{1}{2}p. \end{aligned}$$

4. Each toss of a coin results in a head with probability p. The coin is tossed until the first H appears. Let X be the total number of tosses. What is $\mathbf{P}(X > m)$? Find the distribution function (cdf) of X.

Answer. Assuming independence,

$$\mathbf{P}(X > m) = \mathbf{P}(\text{all T in } m \text{ tosses}) = (1-p)^m.$$

Hence, denoting $[x]$ the integer part of $x \ge 1$, we have for $x \ge 1$,

$$\mathbf{P}(X \le x) = \mathbf{P}(X \le [x]) = (1-p)^{[x]}.$$

Thus

$$F(x) = \mathbf{P}(X \le x) = \begin{cases} 0 & \text{if } x < 1, \\ (1-p)^{[x]} & \text{if } x \ge 1. \end{cases}$$

Also, we can describe it as $F(x) = 0$ if $x < 0$, and $F(x) = (1-p)^k$ if $k \le x < k+1, k = 1, 2, \ldots$; we denoted $[x]$ the integer part of x.

5. Buses arrive at ten minutes intervals starting at noon. A man arrives at the bus stop a random number X minutes after noon, where X has distribution function (cdf)

$$F(x) = \mathbf{P}(X \le x) = \begin{cases} 0 & \text{if } x < 0, \\ x/60 & \text{if } 0 \le x \le 60, \\ 1 & \text{if } x > 60. \end{cases}$$

What is the probability that he waits less than 5 minutes for a bus?

Answer. X is uniform in $(0, 60)$. For A= "he waits less than 5 minutes" we have

$$\begin{aligned} A &= \{5 \le X \le 10\} \cup \{15 \le X \le 20\} \cup \{25 \le X \le 30\} \\ &\quad \cup \{35 \le X \le 40\} \cup \{45 \le X \le 50\} \cup \{55 \le X \le 60\}. \end{aligned}$$

Since F is a continuos function, and all the sets in the union are disjoint, $|A| = 30$ ($|A|$ is the length of A), and

$$\mathbf{P}(\text{less than 5 minutes}) = \frac{|A|}{60} = \frac{30}{60} = \frac{1}{2}.$$

6. A coin is tossed repeatedly and heads turns up on each toss with probability p. Let H_n and T_n be the numbers of heads and tails in n tosses. Show that for each $\varepsilon > 0$,

$$\mathbf{P}\left(2p-1-\varepsilon \le \frac{1}{n}(H_n - T_n) \le 2p-1+\varepsilon\right) \to 1$$

as $n \to \infty$. Hint. $H_n + T_n = n$.

7. Let $X_k, k \ge 1$, be observations which are independent and identically distributed with unknown distribution function F. Describe and justify a method for estimating $F(x)$.

Hint. For $x \in \mathbf{R}$, consider a sequence $A_k = \{X_k \le x\}, k \ge 1$, of independent events. What is their probability $\mathbf{P}(A_k)$? Apply Bernoulli theorem.

Answer. For $x \in \mathbf{R}$, consider a sequence $A_k = \{X_k \le x\}, k \ge 1$, of independent events. Then $\mathbf{P}(A_k) = \mathbf{P}(X_k \le x) = F(x), k \ge 1$. Consider the random function

$$F_n(x) = \frac{\sum_{k=1}^n I_{\{X_k \le x\}}}{n} = \frac{\sum_{k=1}^n I_{A_k}}{n}, x \in \mathbf{R};$$

$F_n(x), x \in \mathbf{R}$, is called empirical distribution function. By Bernoulli theorem, $F_n(x) \to F(x)$ as $n \to \infty$ with probability 1.

8. a) Show that if X, Y are random variables on a given probability space $(\Omega, \mathcal{F}, \mathbf{P})$, then so are $X + Y$. Show that the set of all random variables on $(\Omega, \mathcal{F}, \mathbf{P})$ constitutes a vector space: if X, Y are random variables and $a, b \in \mathbf{R}$, then $aX + bY$ is a r.v. If Ω is finite, write down a basis for this space.

b) Let $X_1, X_2, \ldots$ be r.v. Show that $\sup_n X_n$ and $\inf_n X_n$ are r.v. Show that the limit $\limsup_n X_n$ and $\liminf_n X_n$ are r.v.

Answer. a) Let X, Y be r.v.: $\{X \le x\}, \{Y \le y\} \in \mathcal{F}$ for all $x, y \in \mathbf{R}$. Let $\mathbb{Q}$ be the countable set of all rational numbers. Then for any $z \in \mathbf{R}$,

$$\{X + Y \le z\} = \cap_{n=1}^\infty \cup_{r \in \mathbb{Q}} \left(\{X \le r\} \cap \left\{Y \le z - r + \frac{1}{n}\right\}\right) \in \mathcal{F}.$$

If $\Omega = \{\omega_1, \ldots, \omega_n\}$ is finite, and $X : \Omega \to \mathbf{R}$, then

$$X(\omega) = \sum_{k=1}^n X(\omega_n) I_{\{\omega_k\}}(\omega),$$

i.e. $\left\{I_{\{\omega_k\}}, k = 1, \ldots, n\right\}$ is the basis: any r.v. is a linear combination of $I_{\{\omega_k\}}, k = 1, \ldots, n$.

b) For any x,

$$\left\{\sup_n X_n \le x\right\} = \cap_{n=1}^\infty \{X_n \le x\} \in \mathcal{F},$$

and $\inf_n X_n = -\sup_n(-X_n)$ is a r.v. as well.

9. A deck of 52 cards is dealt to 4 people. J is one of them. Let X be the number of aces J gets.

(a) Find the probability mass function of X.

Answer. The range of X is $\{0, 1, 2, 3, 4\}$. There are $\binom{52}{13,13,13,13}$ different ways to divide a deck of 52 cards to 4 people, 13 cards each. For a given $k \in \{0, 1, 2, 3, 4\}$, let us count in how many ways J could get k aces:

(i) there are $\binom{4}{k}$ ways to choose specific k aces (recall $\binom{4}{0} = \binom{4}{4} = 1$).

(ii) remaining $13 - k$ cards should come from the deck without aces: there are $\binom{48}{13-k}$ ways of that.

(ii) after J gets 13 cards, there are $\binom{39}{13,13,13}$ different ways to divide the remaining 39 cards to 3 people, 13 cards each.

Hence by multiplication principle, there are $\binom{4}{k}\binom{48}{13-k}\binom{39}{13,13,13}$ J could get 4 aces, and

$$\begin{aligned}\mathbf{P}(X = k) &= \frac{\binom{4}{k}\binom{48}{13-k}\binom{39}{13,13,13}}{\binom{52}{13,13,13,13}} \\ &= \frac{\binom{4}{k}\binom{48}{13-k}39!13!}{52!} = \frac{\binom{4}{k}\binom{48}{13-k}}{\binom{52}{13}}.\end{aligned}$$

We find

k	0	1	2	3	4
$\mathbf{P}(X = k)$	$\frac{6327}{20\,825} = 0.303\,82$	$\frac{9139}{20\,825} = 0.438\,85$	$\frac{4446}{20\,825} = 0.213\,49$	$\frac{858}{20\,825} = 0.041\,2$	$\frac{11}{4165} = 0.00264\,1$

Note that

$$\begin{aligned}&\mathbf{E}(X) \\ =\ & (1)\frac{9139}{20\,825} + (2)\frac{4446}{20\,825} + (3)\frac{858}{20\,825} + (4)\frac{11}{4165} = 1.\end{aligned}$$

(b) What is the probability that J gets 3 or 4 aces?

Answer. $\mathbf{P}$ (J gets 3 or 4 aces) $= p(3) + p(4) = \frac{11}{4165} + \frac{858}{20\,825} = \frac{913}{20\,825} = 0.04\,384\,2$.

(c) What is the probability that J gets at least once three or four aces in 100 independent deals.

Answer. Denoting $a = \mathbf{P}$ (J gets 3 or 4 aces) $= \frac{913}{20\,825}$, we have

$$\begin{aligned}\mathbf{P}\text{ (J does not get 3 or 4 aces)} &= \mathbf{P}\text{ (J gets at most 2 aces)} \\ &= 1 - a = \frac{19\,912}{20\,825} = 0.956\,16,\end{aligned}$$

and

$$\begin{aligned}&\mathbf{P}\text{ (J gets at least once three or four aces in 100 games)} \\ =\ & 1 - \mathbf{P}\text{ (J never gets three or four aces in 100 games)} \\ =\ & 1 - (1-a)^{100} = 0.9887.\end{aligned}$$

10. (Truncation) Let X be a r.v. with distribution function $F(x)$. Define a truncated r.v. Y as

$$Y = \begin{cases} a & \text{if } x < a, \\ X & \text{if } a \le x \le b, \\ b & \text{if } x > b. \end{cases}$$

Write the distribution function F_Y of Y using F. How F_Y behave as $a \to -\infty, b \to \infty$?

11. a) Show that if f, g are prob. density functions and $\lambda \in [0,1]$, then $\lambda f + (1-\lambda) g$ is a density. Is the product fg a density function? Recall h is a (probability) density function if $h \geq 0$ and $\int_{-\infty}^{\infty} h(x)\, dx = 1$.

Answer. By definition, $\lambda f + (1-\lambda) g \geq 0$, and $\int_{-\infty}^{\infty} f dx = \int_{-\infty}^{\infty} g dx = 1$. Hence

$$\begin{aligned} \int_{-\infty}^{\infty} [\lambda f + (1-\lambda) g] dx &= \lambda \int_{-\infty}^{\infty} f dx + (1-\lambda) \int_{-\infty}^{\infty} g dx \\ &= \lambda + (1-\lambda) = 1. \end{aligned}$$

b) For what constant c, d the following is a density function:

$$f(x) = \begin{cases} c x^{-d} & \text{if } x > 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the corresponding distribution function.

Answer. The function $f \geq 0$ if $c > 0$, and

$$\int_{-\infty}^{\infty} f dx = c \int_{1}^{\infty} x^{-d} dx$$

converges if and only if $d > 1$ (Calculus II), and for $d > 1$,

$$c \int_{1}^{\infty} x^{-d} dx = c \frac{x^{-d+1}}{-d+1} |_{1}^{\infty} = c \frac{1}{d-1} = 1$$

if $c = d - 1$. Thus f is a pdf if $c = d - 1, d > 1$.

Distribution function, for $x > 1$,

$$F(x) = \int_{-\infty}^{x} f(y)\, dy = (d-1) \int_{1}^{x} y^{-d} dy = 1 - x^{-d+1}.$$

Thus

$$F(x) = \int_{-\infty}^{x} f(y)\, dy = \begin{cases} 0 & \text{if } x \leq 1, \\ 1 - x^{1-d} & \text{if } x > 1. \end{cases}$$

12. Death probability playing Russian roulette is 0.167. Let X be the number of games until the player is dead, equivalently, X is the player lifetime (assuming time is measured by the number of games).

(a) Find the pmf of X and the probability $\mathbf{P}(X > n)$ (it is the probability to be alive at time moment n).

Answer. The range of X is the set $\{1, 2, \ldots\}$, X is geometric with parameter $p = 0.167$: it is like waiting for the first H repeatedly tossing a coin with $\mathbf{P}(H) = p$. Hence

$$\mathbf{P}(X = k) = q^{k-1} p, k = 1, 2, \ldots$$

with $q = 1 - p$. The pmf is

$$\begin{aligned} f(x) &= q^{x-1} p, x = 1, 2, \ldots, \\ f(x) &= 0 \text{ otherwise.} \end{aligned}$$

Now,

$$\mathbf{P}\left(\text{player is alive at time } n\right) = \mathbf{P}\left(X > n\right) = q^{n}, n = 1, 2, \ldots.$$

(b) Show that $\mathbf{P}\left(X > n + m | X > n\right) = \mathbf{P}\left(X > m\right)$. Why is it called "memoryless" property of X?

Answer. By definition and part a),

$$\begin{aligned}
\mathbf{P}\left(X > n + m | X > n\right) &= \frac{\mathbf{P}\left(X > n + m, X > n\right)}{\mathbf{P}\left(X > n\right)} \\
&= \frac{\mathbf{P}\left(X > n + m\right)}{\mathbf{P}\left(X > n\right)} = \frac{q^{n+m}}{q^{n}} = q^{m} = \mathbf{P}\left(X > m\right).
\end{aligned}$$

X is "memoryless": to be alive after n games does not decrease the probability to survive the next m games; it is the same as if no games were played (the death threat does not accumulate).

13. Which of the following are distribution functions? For those that are find the corresponding density function f.

(a)

$$F\left(x\right) = \begin{cases} 1 - e^{-x^{2}} & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

(b)

$$F\left(x\right) = \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Hint. Check whether $f\left(x\right) = F'\left(x\right)$ could be a pdf, i.e. $f \geq 0, \int_{-\infty}^{\infty} f\left(x\right) dx = 1$ and $F\left(x\right) = \int_{-\infty}^{x} f\left(u\right) du$.

14. For what values of the constant C do the following define mass functions on the positive integers $1, 2, \ldots$?

(a) Geometric: $f\left(k\right) = C2^{-k}, k \in \{1, 2, \ldots\}$.

(b) Logarithmic: $f\left(k\right) = C2^{-k}/k, k \in \{1, 2, \ldots\}$.

(c) Inverse square: $f\left(k\right) = Ck^{-2}, k \in \{1, 2, \ldots\}$.

(d) $f\left(k\right) = C2^{k}/k!, k \in \{1, 2, \ldots\}$.

Hint. Since $f\left(k\right) \geq 0$ with any $C > 0$, check for what $C > 0$, $\sum_{k=1}^{\infty} f\left(k\right) = 1$.

Answer. Since $f\left(k\right) \geq 0$ with any $C > 0$, we check for what $C > 0$, $\sum_{k=1}^{\infty} f\left(k\right) = 1$.

(a)

$$C \sum_{k=1}^{\infty} 2^{-k} = C \frac{2^{-1}}{1 - 1/2} = C = 1.$$

(b) Since $\frac{2^{-k}}{k} \leq 2^{-k}, k \geq 1$, the series $\sum_{k=1}^{\infty} \frac{2^{-k}}{k}$ converges. Hence

$$C \sum_{k=1}^{\infty} \frac{2^{-k}}{k} = 1$$

implies that $C = \left(\sum_{k=1}^{\infty} \frac{2^{-k}}{k}\right)^{-1}$. To find a "nice" expression for C, consider for $x \in (0, 1)$, $h\left(x\right) = \sum_{k=1}^{\infty} \frac{x^{k}}{k} = \sum_{k=1}^{\infty} \int_{0}^{x} y^{k-1} dx = \int_{0}^{x} \left(\sum_{k=1}^{\infty} y^{k-1}\right) dx = \int_{0}^{x} \frac{1}{1-y} dx = -\ln\left(1 - y\right)|_{0}^{x} = -\ln\left(1 - x\right)$, and $h\left(1/2\right) = -\ln\left(1/2\right) = \ln 2$. Hence $C = \left(\ln 2\right)^{-1}$.

(c) The series $\sum_{k=1}^{\infty}\frac{1}{k^2}$ converges as p-series, and $\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{1}{6}\pi^2$:

$$C=\left(\sum_{k=1}^{\infty}\frac{1}{k^2}\right)^{-1}=\frac{6}{\pi^2}.$$

(d) For any $x\in\mathbf{R}$, $\sum_{k=1}^{\infty}\frac{x^k}{k!}=e^x-1$. Hence

$$C\sum_{k=1}^{\infty}\frac{2^k}{k!}=C(e^2-1)=1$$

gives $C=\left(e^2-1\right)^{-1}$.

15. We toss n coins, and each one shows heads with probability p, independently of each of the others. Each coin which shows heads is tossed again. What is the mass function of the number of heads resulting from the second round of tosses.

Answer. Let X_1 be the number of H in the first toss of n coins, and X_2 be the number of H in the 2nd toss of X_1 coins. Note that $X_1\sim Bin\,(n,p)$, and $X_2\sim Bin\,(k,p)$ given $X_1=k\geq 1$. Now, denoting $q=1-p$,

$$\begin{aligned}\mathbf{P}\left(X_2=0|X_1=0\right)\mathbf{P}\left(X_1=0\right) &= \mathbf{P}\left(X_1=0\right)=q^n,\\ \mathbf{P}\left(X_2=j|X_1=0\right)\mathbf{P}\left(X_1=0\right) &= 0,\, j\geq 1.\end{aligned}$$

Obviously, $\Omega=\cup_{k=0}^{n}\{X_1=k\}$, and $\{X_1=n\}$ are disjoint. By total probability formula,

$$\begin{aligned}\mathbf{P}\left(X_2=j\right) &= \sum_{k=j}^{n}\mathbf{P}\left(X_2=j|X_1=k\right)\mathbf{P}\left(X_1=k\right)\\ &= \sum_{k=j}^{n}\binom{k}{j}p^j q^{k-j}\binom{n}{k}p^k q^{n-k}=\sum_{k=j}^{n}\frac{k!}{(k-j)!j!}\frac{n!}{k!(n-k)!}p^{j+k}q^{n-j}\\ &= \sum_{k=j}^{n}\binom{n}{n-k,k-j,j}p^{j+k}q^{n-j}=\binom{n}{j}p^{2j}\left(1-p^2\right)^{n-j},\end{aligned}$$

i.e. X_2 is binomial$\left(n,p^2\right)$.

5 Week 5: 9/19- 9/23

1. a) If U and V are jointly continuous, show that $\mathbf{P}(U = V) = 0$. Hint. If f is their joint pdf, then $\mathbf{P}((U, V) \in D) = \int\int_D f(u, v)\, dv du$. Recall the geometric meaning of the double integral: it represents the volume under the graph $z = f(x, y), (x, y) \in D$.

Comment. Let X be uniformly distributed on (0, 1), and take $Y = X$. Then X is continuous, Y is continuous, and $\mathbf{P}(X = Y) = 1$. Hence (X, Y) cannot be jointly continuous: in general, continuity of the marginal distributions does not imply the joint continuity.

Answer. Let $D = \{(u, v) \in \mathbf{R}^2 : u = v\}$. If f is their joint pdf, then

$$\mathbf{P}(U = V) = \mathbf{P}((U, V) \in D) = \int\int_D f(u, v)\, dv du.$$

The solid under graph of $z = f(x, y), (x, y) \in D$, is a piece of a plane: its volume is zero. Thus $\int\int_D f(u, v)\, dv du = 0$.

Comment. Let X be uniformly distributed on (0, 1), and take $Y = X$. Then X is continuous, Y is continuous, and $\mathbf{P}(X = Y) = 1$. Hence (X, Y) cannot be jointly continuous: in general, continuity of the marginal distributions does not imply the joint continuity.

b) Let U and V be jointly continuous with joint pdf $f(x, y)$. Show the marginal distributions are continuous with pdf

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} f(u, v)\, dv, u \in \mathbf{R}, \\ f_V(v) &= \int_{-\infty}^{\infty} f(u, v)\, du, v \in \mathbf{R}. \end{aligned}$$

Answer. By definition of the joint pdf, for $u \in \mathbf{R}$, and rewriting the double integral as an iterated one,

$$\begin{aligned} F_U(u) &= \mathbf{P}(U \le u) = \mathbf{P}(U \le u, -\infty < V < \infty) \\ &= \int_{-\infty}^{u}\int_{-\infty}^{\infty} f(x, y)\, dy dx = \int_{-\infty}^{u}\left(\int_{-\infty}^{\infty} f(x, y)\, dy\right) dx = \int_{-\infty}^{u} g(x)\, dx, \end{aligned}$$

where

$$g(x) = \left(\int_{-\infty}^{\infty} f(x, y)\, dy\right), x \in \mathbf{R}.$$

Since $f \ge 0$, we have $g \ge 0$ as well, i.e. $g = f_U$. Similarly, we find that

$$f_V(v) = \int_{-\infty}^{\infty} f(x, v)\, dx, v \in \mathbf{R}.$$

Comment. Joint continuity implies continuity of the marginal distributions.

2. A fair coin is tossed twice. Let X be the number of heads, and let W be the indicator function of the event $\{X = 2\}$. Find $\mathbf{P}(X = x, W = w)$ for all appropriate values of x and w.

Answer. We have

$$\begin{aligned} \{X = 0, W = 1\} &= \emptyset, \{X = 0, W = 0\} = \{X = 0\}, \\ \{X = 1, W = 1\} &= \emptyset, \{X = 1, W = 0\} = \{X = 1\}, \\ \{X = 2, W = 1\} &= \{X = 2\}, \{X = 2, W = 0\} = \emptyset, \end{aligned}$$

and

$$\begin{aligned}
\mathbf{P}(X=0,W=1) &= 0, \mathbf{P}(X=0,W=0) = \mathbf{P}(X=0) = (1/2)^2 = 1/4; \\
\mathbf{P}(X=1,W=1) &= 0, \mathbf{P}(X=1,W=0) = \mathbf{P}(X=1) = \binom{2}{1}(1/2)^2 = 1/2; \\
\mathbf{P}(X=2,W=1) &= \mathbf{P}(X=2) = (1/2)^2 = 1/4, \mathbf{P}(X=2,W=0) = 0.
\end{aligned}$$

3. Let X be a Bernoulli random variable, so that $\mathbf{P}(X=0) = 1-p, \mathbf{P}(X=1) = p$. Let $Y = 1-X$ and $Z = XY$. Find $\mathbf{P}(X=x, Y=y)$ and $\mathbf{P}(X=x, Z=z)$ for all $x, y, z \in \{0,1\}$.

Answer. First

$$\begin{aligned}
\mathbf{P}(X=0,Y=0) &= \mathbf{P}(X=0,X=1) = 0, \\
\mathbf{P}(X=0,Y=1) &= \mathbf{P}(X=0) = 1-p, \\
\mathbf{P}(X=1,Y=0) &= \mathbf{P}(X=1) = p, \mathbf{P}(X=1,Y=1) = 0.
\end{aligned}$$

Since the values of X are 0 or 1,$X^2 = X$, and $Z = XY = X(1-X) = X - X^2 = 0$,

$$\begin{aligned}
\mathbf{P}(X=0,Z=0) &= \mathbf{P}(X=0) = 1-p, \\
\mathbf{P}(X=1,Z=0) &= \mathbf{P}(X=1) = p, \\
\mathbf{P}(X=1,Z=1) &= \mathbf{P}(X=0,Z=1) = 0
\end{aligned}$$

4. The random variables X and Y have joint distribution function

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{if } x < 0, \\ (1-e^{-x})\left(\frac{1}{2} + \frac{1}{\pi}\tan^{-1} y\right) & \text{if } x \geq 0. \end{cases}$$

Show that X and Y are (jointly) continuously distributed. Hint. If X and Y are (jointly) continuously distributed, then their joint pdf is $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$ if f is continuous at (x,y). If we figure out what could be $f(x,y)$, check that

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dv\,du.$$

5. Let X and Y have joint distribution function F. Show that

$$\begin{aligned}
& \mathbf{P}(a < X \leq b, c < Y \leq d) \\
= {} & F(b,d) - F(a,d) - F(b,c) + F(a,c).
\end{aligned}$$

6. Which of the following are distribution functions? For those that are find the corresponding density function f.

(a)

$$F(x) = \begin{cases} 1 - e^{-x^2} & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Answer. The function F is continuous including zero: $\lim_{x\to 0} F(x) = 0$. Hence (c) of Lemma (6) holds, p. 28; also, F is nondecreasing: (b) holds. Finally,

$$\lim_{x\to-\infty} F(x) = 0, \lim_{x\to\infty} F(x) = 1.$$

The pdf is

$$f(x) = F'(x) = \begin{cases} 2xe^{-x^2} & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is nonnegative and, using substitution,

$$\int_{-\infty}^{x} f(y)\, dy = F(x), x \in \mathbf{R}.$$

(b)

$$F(x) = \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Hint. Check whether $f(x) = F'(x)$ could be a pdf, i.e. $f \geq 0, \int_{-\infty}^{\infty} f(x)\, dx = 1$ and $F(x) = \int_{-\infty}^{x} f(u)\, du$.

Answer. The function F is continuous including zero: $\lim_{x \to 0+} F(x) = \lim_{x \to 0-} F(x) = 0$. Hence (c) of Lemma (6) holds, p. 28; also, F is nondecreasing: (b) holds. Finally,

$$\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1.$$

The pdf is

$$f(x) = F'(x) = \begin{cases} x^{-2}e^{-1/x} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Using substitution we check that

$$F(x) = \int_{-\infty}^{x} f(y)\, dy, x \in \mathbf{R}.$$

7. For what values of the constant C do the following define mass functions on the positive integers $1, 2, \ldots$?

(a) Geometric: $f(k) = C2^{-k}, k \in \{1, 2, \ldots\}$.
(b) Logarithmic: $f(k) = C2^{-k}/k, k \in \{1, 2, \ldots\}$.
(c) Inverse square: $f(k) = Ck^{-2}, k \in \{1, 2, \ldots\}$.
(d) $f(k) = C2^{k}/k!, k \in \{1, 2, \ldots\}$.
Hint. Since $f(k) \geq 0$ with any $C > 0$, check for what $C > 0$, $\sum_{k=1}^{\infty} f(k) = 1$.
Answer. Since $f(k) \geq 0$ with any $C > 0$, we check for what $C > 0$, $\sum_{k=1}^{\infty} f(k) = 1$.
(a)

$$C \sum_{k=1}^{\infty} 2^{-k} = C \frac{2^{-1}}{1 - 1/2} = C = 1.$$

(b) Since $\frac{2^{-k}}{k} \leq 2^{-k}, k \geq 1$, the series $\sum_{k=1}^{\infty} \frac{2^{-k}}{k}$ converges. Hence

$$C \sum_{k=1}^{\infty} \frac{2^{-k}}{k} = 1$$

implies that $C = \left(\sum_{k=1}^{\infty} \frac{2^{-k}}{k}\right)^{-1}$. To find a "nice" expression for C, consider for $x \in (0,1)$, $h(x) = \sum_{k=1}^{\infty} \frac{x^k}{k} = \sum_{k=1}^{\infty} \int_0^x y^{k-1} dx = \int_0^x \left(\sum_{k=1}^{\infty} y^{k-1}\right) dx = \int_0^x \frac{1}{1-y} dx = -\ln(1-y)|_0^x = -\ln(1-x)$, and $h(1/2) = -\ln(1/2) = \ln 2$. Hence $C = (\ln 2)^{-1}$.

(c) The series $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges as p-series, and $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{1}{6}\pi^2$:

$$C = \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^{-1} = \frac{6}{\pi^2}.$$

(d) For any $x \in \mathbf{R}$, $\sum_{k=1}^{\infty} \frac{x^k}{k!} = e^x - 1$. Hence

$$C \sum_{k=1}^{\infty} \frac{2^k}{k!} = C(e^2 - 1) = 1$$

gives $C = \left(e^2 - 1\right)^{-1}$.

8. We toss n coins, and each one shows heads with probability p, independently of each of the others. Each coin which shows heads is tossed again. What is the mass function of the number of heads resulting from the second round of tosses.

9. Let X and Y be independent random variables, each taking the values -1 and 1 with probability $1/2$, and let $Z = XY$. Show that X, Y, and Z are pairwise independent. Are they independent?

Answer. We find mass function of Z :

$$\begin{aligned} \mathbf{P}(Z=1) &= \mathbf{P}(\{X=Y=1\} \cup \{X=Y=-1\}) = (1/2)^2 + (1/2)^2 = 1/2. \\ \mathbf{P}(Z=-1) &= \mathbf{P}(\{X=1, Y=-1\} \cup \{X=-1, Y=1\}) = 1/2. \end{aligned}$$

Now,

$$\begin{aligned} \mathbf{P}(X=1, Z=1) &= \mathbf{P}(X=1, Y=1) = (1/2)^2 = \mathbf{P}(X=1)\,\mathbf{P}(Z=1), \\ \mathbf{P}(X=1, Z=-1) &= \mathbf{P}(X=1, Y=-1) = (1/2)^2 = \mathbf{P}(X=1)\,\mathbf{P}(Z=-1), \\ \mathbf{P}(X=-1, Z=1) &= \mathbf{P}(X=-1, Y=-1) = (1/2)^2 = \mathbf{P}(X=-1)\,\mathbf{P}(Z=1), \\ \mathbf{P}(X=-1, Z=-1) &= \mathbf{P}(X=-1, Y=1) = (1/2)^2 = \mathbf{P}(X=-1)\,\mathbf{P}(Z=-1). \end{aligned}$$

Hence X and Z are independent. By symmetry, Y and Z are independent as well. On the other hand,

$$0 = \mathbf{P}(X=1, Y=-1, Z=1) \neq \mathbf{P}(Z=1)\,\mathbf{P}(Y=-1)\,\mathbf{P}(X=1) = 1/8.$$

They are not independent: three r.v. Recall X, Y, Z are independent if they are pairwise independent and $\{X=x\}, \{Y=y\}, \{Z=z\}$ are independent for all x, y, z.

10. Let X and Y be independent random variables taking values in the positive integers and having the same mass function $f(x) = 2^{-x}$ for $x \in \{1, 2, \ldots\}$. Find joint probability mass function and:

(a) $\mathbf{P}(\min\{X, Y\} \le x)$. Hint. Find $\mathbf{P}(\min\{X, Y\} > x)$.

(b) $\mathbf{P}(Y > X)$; (c) $\mathbf{P}(X = Y)$, (d) $\mathbf{P}(X \ge kY)$, for a given positive integer k;

(e) $\mathbf{P}(X \text{ divides } Y)$. Hint. X divides Y means $Y = lX, l = 1, 2, \ldots$. Answer is a series.

11. Is it generally true that $\mathbf{E}(1/X) = 1/\mathbf{E}(X)$? Is it ever true that $\mathbf{E}(1/X) = 1/\mathbf{E}(X)$?

Answer. Is it generally true that $\mathbf{E}(1/X) = 1/\mathbf{E}(X)$? No. Let $X \sim$Bernoulli$(1/2)$. Then

$$\begin{aligned}
\mathbf{E}\left(\frac{1}{1+X}\right) &= \frac{1}{2}\frac{1}{2} + 1\frac{1}{2} = 3/4, \mathbf{E}(1+X) = \frac{1}{2} + 2\frac{1}{2} = \frac{3}{2}, \\
\mathbf{E}\left(\frac{1}{1+X}\right) &\neq \frac{1}{\mathbf{E}(1+X)}.
\end{aligned}$$

Is it ever true that $\mathbf{E}(1/X) = 1/\mathbf{E}(X)$? Sometimes. Let X have mass function: $f(-1) = \frac{1}{9}$, $f\left(\frac{1}{2}\right) = \frac{4}{9}$, $f(2) = \frac{4}{9}$. We find

$$\begin{aligned}
\mathbf{E}(1/X) &= (-1)\frac{1}{9} + (2)\frac{4}{9} + \frac{1}{2}\frac{4}{9} = 1, \\
\mathbf{E}(X) &= (-1)\frac{1}{9} + \frac{1}{2}\frac{4}{9} + 2\cdot\frac{4}{9} = 1.
\end{aligned}$$

12. (a) Let X and Y be independent discrete random variables, and let $g, h : \mathbf{R} \to \mathbf{R}$. Show that $g(X)$ and $h(Y)$ are independent.

(b) Show that X, Y are independent if and only if $f_{X,Y}(x, y)$ can be factorized as the product $g(x)h(y)$ of a function of x alone and a function of y alone.

(c) If X and Y are independent and $g, h : \mathbf{R} \to \mathbf{R}$, show that $\mathbf{E}[g(X)h(Y)] = \mathbf{E}(g(X))\mathbf{E}(h(Y))$ whenever these expectations exist.

Answer. (a) For any u, v,

$$\{g(X) = u, h(Y) = v\} = \cup_{g(x)=u, h(y)=v} \{X = x, Y = y\}.$$

Thus

$$\begin{aligned}
\mathbf{P}(g(X) = u, h(Y) = v) &= \sum_{g(x)=u,h(y)=v} \mathbf{P}(X = x, Y = y) = \sum_{g(x)=u,h(y)=v} \mathbf{P}(X = x)\mathbf{P}(Y = y) \\
&= \sum_{g(x)=u} \mathbf{P}(X = x) \sum_{h(y)=v} \mathbf{P}(Y = y) = \mathbf{P}(g(X) = u)\mathbf{P}(h(Y) = v).
\end{aligned}$$

(b) Show that X, Y are independent if and only if $f_{X,Y}(x, y)$ can be factorized as the product $g(x)h(y)$ of a function of x alone and a function of y alone.

Answer. Let X, Y be independent. Then

$$\begin{aligned}
f_{X,Y}(x, y) &= \mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x)\mathbf{P}(Y = y) \\
&= f_X(x) f_Y(y), x, y \in \mathbf{R}.
\end{aligned}$$

Let $f_{X,Y}(x, y) = f(x)g(y), x, y \in \mathbf{R}$. Then

$$\begin{aligned}
f_X(x) &= \mathbf{P}(X = x) = \mathbf{P}\left(\cup_y \{Y = y\} \cap \{X = x\}\right) \\
&= \sum_y \mathbf{P}(X = x, Y = y) = \sum_y f(x)g(y) = f(x)\sum_y g(y) = cf(x), x \in \mathbf{R}.
\end{aligned}$$

Note that $1 = \sum_x f_X(x) = \sum_x f(x) \sum_y g(y) = c\sum_x f(x)$, i.e.,

$$\sum_x f(x) = c^{-1}.$$

Similarly,

$$f_Y(y) = g(y)\sum_x f(x) = c^{-1}g(y), y \in \mathbf{R}.$$

Hence

$$\begin{aligned} f_{X,Y}(x,y) &= \mathbf{P}(X = x, Y = y) = (cf(x))\left(c^{-1}g(y)\right) \\ &= f_X(x) f_Y(y) = \mathbf{P}(X = x)\mathbf{P}(Y = y), x, y \in \mathbf{R}. \end{aligned}$$

13. Given 10000 married couples, compute the probability that, in at least one of them, (a) both husband and wife were born on April 30; (b) both husband and wife were born on the same day. Use binomial and its Poisson approximation. Is approximation accurate?

14. a) Number X of goals in a game of a hockey player is Poisson. He scores at least one goal in roughly half of his games. Find Poisson parameter λ of X.

Hint. Bernoulli theorem: interpret "He scores at least one goal in roughly half of his games" as $\mathbf{P}(X \geq 1) = 1/2$.

b) Find probability that he scores a hattrick (three goals) in a game.

Comment: we could look at this probability as the percentage of games where he scores a hattrick.

Answer. a)We find that

$$\frac{1}{2} = \mathbf{P}(X \geq 1) = 1 - \mathbf{P}(X = 0) = 1 - e^{-\lambda}.$$

Hence

$$\begin{aligned} e^{-\lambda} &= 1 - \frac{1}{2} = \frac{1}{2}, -\lambda = \ln\frac{1}{2} = -\ln 2, \\ \lambda &= \ln 2. \end{aligned}$$

b) Now,

$$\mathbf{P}(X = 3) = e^{-\ln 2}\frac{(\ln 2)^3}{3!} = \frac{1}{2}\frac{(\ln 2)^3}{3!} = 0.027752$$

15. A box contains b blue and r red balls (total number of balls in the box is $n = b + r$).

All balls are removed at random one by one and arranged in a row. Let X_i be the number of red balls between the $(i-1)$th and ith blue ball drawn, $i = 2, \ldots, b$; Let X_1 be the number of red balls until the first blue ball shows up, and X_{b+1} be the number of red balls after the last blue ball drawn. Consider the random vector $X = (X_1, \ldots X_{b+1})$. The range of X are all the vectors $(k_1, \ldots, k_{b+1})$ with nonnegative integer components $k_1, \ldots, k_{b+1}$ such that $k_1 + \ldots + k_{b+1} = r$. For such a vector $(k_1, \ldots, k_{b+1})$ with $k_i \geq 0$, and $k_1 + \ldots + k_{b+1} = r$, find

$$\mathbf{P}(X = (k_1, \ldots, k_{b+1})) = \mathbf{P}(X_1 = k_1, \ldots, X_{b+1} = k_{b+1}).$$

Note $X_i = 0$ means there are no red balls between the $(i-1)$th and ith blue balls if $i = 2, \ldots, b$; $X_1 = 0$ means the row starts with blue ball; $X_{b+1} = 0$ means the last ball is blue. Hint. The computation is simple. Do not overthink.

6 Week 6: 9/26- 9/30

1. Consolidated Products sells extreme-sports bikes and plaster casts for broken bones. Average number of extreme-sports bikes sold in a week is 10 with the standard deviation of 3 bikes. Average number of plaster casts sold in a week is 12 with the standard deviation of 4. Company finds that the extreme-sports bike and plaster cast sales have a positive correlation of 0.4.

a) Find the mean and standard deviation of the total number of items (extreme sports bikes plus plaster casts sold in a week).

b) Suppose the profit per extreme-sports bike (in hundreds of dollars) is 5 and the profit per cast is 2. Find the mean and standard deviation of the total profit.

Answer. Let X = # of bikes sold, Y = # of casts sold. The correlation $\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$ and Cov(X,Y)=$\rho\sigma_X\sigma_Y = 0.4 \cdot 3 \cdot 4 = 4.8$.

a)

$$\begin{aligned} Var(X+Y) &= Var(X) + Var(Y) + 2Cov(X,Y) \\ &= 3^2 + 4^2 + 2 \cdot 4.8 = 34.6. \end{aligned}$$

b) The profit

$$R = 5X + 2Y,\ \mathbf{E}(R) = 5\mathbf{E}(X) + 2\mathbf{E}(Y) = 5 \cdot 10 + 2 \cdot 12 = 74,$$

and

$$\begin{aligned} Var(R) &= Var(5X+2Y) = 5^2 Var(X) + 2^2 Var(Y) + 2 \cdot 5 \cdot 2Cov(X,Y) \\ &= 25 \cdot 3^2 + 4 \cdot 4^2 + 20 \cdot 4.8 = 385.0, \sigma_R = \sqrt{385} = 19.621. \end{aligned}$$

2. A biased coin is tossed n times with probability p of H. A run is a sequence of throws which result in the same outcome. For example, HHTHTTH contains 5 runs. Show that the expected number of runs is $1 + 2(n-1)p(1-p)$. Find the variance of the number of runs.

4. An urn contains n balls numbered 1, 2, …, n. We remove k balls at random without replacement, and add up their numbers. Find the mean and the variance of the total.

Answer. Let A_i = "the ball numbered i is among k. I X is the sum, then $X = \sum_{i=1}^{n} i I_{A_i}$,

$$\begin{aligned} X^2 &= \sum_{i=1}^{n} i^2 I_{A_i} + 2\sum_{i<j} ij I_{A_i A_j}, \\ \mathbf{E}(X) &= \sum_{i=1}^{n} i\mathbf{P}(A_i), \mathbf{E}\left(X^2\right) = \sum_{i=1}^{n} i^2 \mathbf{P}(A_i) + 2\sum_{i<j} ij\mathbf{P}\left(A_i A_j\right). \end{aligned}$$

We find, by counting,

$$\begin{aligned} \mathbf{P}(A_i) &= \frac{k(n-1)\dots(n-1-(k-1)+1)}{n(n-1)\dots(n-1-(k-1)+1)} = \frac{k}{n}, \\ \mathbf{P}\left(A_i A_j\right) &= \frac{k(k-1)(n-2)\dots(n-2-(k-2)+1)}{n(n-1)\dots(n-1-(k-1)+1)} = \frac{k}{n}\frac{k-1}{n-1} \end{aligned}$$

$$\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{P}(A_i) = \sum_{i=1}^{n} i\frac{k}{n} = \frac{n(n+1)}{2}\frac{k}{n} = \frac{k(n+1)}{2} \text{ etc.}$$

5. Of the $2n$ people in a given collection of n couples, exactly m die. Assuming that the m people have been picked at random, with what probability each couple survives? Find the mean number of surviving couples.

Answer. The sample space Ω consists of all different groups of m out of $2n$: $\#\Omega = \binom{2n}{m}$. Let A="a particular couple survives". Then $\#A = \binom{2n-2}{m}$, and

$$\begin{aligned} p &= \mathbf{P}(A) = \frac{\binom{2n-2}{m}}{\binom{2n}{m}} = \frac{(2n-m)(2n-m-1)}{2n(2n-1)} \\ &= \left(1 - \frac{m}{2n}\right)\left(1 - \frac{m}{2n-1}\right). \end{aligned}$$

Let A_i = "ith couple survives", $i = 1, \ldots, n$. Then for the number of surviving couples $X = \sum_{i=1}^{n} I_{A_i}$, we have

$$\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{P}(A_i) = n\left(1 - \frac{m}{2n}\right)\left(1 - \frac{m}{2n-1}\right).$$

6. Let X be Poisson r.v.: $\mathbf{P}(X = n) = p_n(\lambda) = e^{-\lambda}\lambda^n/n!, n \geq 0$. Show that

$$\mathbf{P}(X \leq n) = 1 - \int_0^{\lambda} p_n(t)\, dt.$$

Hint. Integrate by parts once $\int_0^{\lambda} e^{-t}\frac{t^n}{n!}dt$; alternatively, look at $\frac{d}{d\lambda}p_n(\lambda)$: what is $\frac{d}{d\lambda}\mathbf{P}(X \leq n)$?

Answer. Integrating by parts,

$$\begin{aligned} \int_0^{\lambda} p_n(t)\, dt &= \int_0^{\lambda} e^{-t}\frac{t^n}{n!}dt = -e^{-t}\frac{t^n}{n!}|_0^{\lambda} + \int_0^{\lambda} \frac{t^{n-1}}{(n-1)!}e^{-t}dt \\ &= \int_0^{\lambda} p_{n-1}(t)\, dt - e^{-\lambda}\frac{\lambda^n}{n!} = \ldots \\ &= \int_0^{\lambda} p_0(t)\, dt - \sum_{i=1}^{n} e^{-\lambda}\frac{\lambda^j}{j!} = 1 - e^{-\lambda} - \sum_{i=1}^{n} e^{-\lambda}\frac{\lambda^j}{j!}, \end{aligned}$$

and

$$1 - \int_0^{\lambda} p_n(t)\, dt = e^{-\lambda}\sum_{j=0}^{n}\frac{\lambda^j}{j!} = \mathbf{P}(X \leq n).$$

7. Let $X_1, \ldots, X_n$ be independent and suppose X_k is Bernoulli with parameter p_k. Let $Y = X_1 + \ldots + X_n$. Show that

$$\mathbf{E}(Y) = \sum_{k=1}^{n} p_k, \mathrm{Var}(Y) = \sum_{k=1}^{n} p_k(1 - p_k).$$

Show that if $\mathbf{E}(Y)$ is fixed ($\mathbf{E}(Y) = \sum_{k=1}^{n} p_k = c$), Var$(Y)$ is maximal when $p_1 = \ldots = p_n$, that is to say, the variation in the sum is greatest when individuals are more alike. Hint. Calculus, Lagrange multipliers.

Answer. Since X_k ~Bernoulli(p_k), $E(X_k) = p_k$,var$(X_k) = p_k(1-p_k)$. Then $\mathbf{E}(Y) = \sum_{k=1}^{n} \mathbf{E}(X_k) = \sum_{k=1}^{n} p_k$, and because of independence,

$$\text{Var}(Y) = \sum_{k=1}^{n} \text{var}(X_k) = \sum_{k=1}^{n} p_k(1-p_k).$$

Hence we have to maximize the function of n variables $p = (p_1, \ldots, p_n)$:

$$H(p) = \sum_{k=1}^{n} p_k(1-p_k)$$

subject to the restriction $g(p) = \sum_{k=1}^{n} p_k - c = 0$. For $p = (p_1, \ldots, p_k)$ at which max is achieved the following must hold for some Lagrange multiplier λ :

$$\frac{\partial H(p)}{\partial p_k} = \lambda \frac{\partial g(p)}{\partial p_k}, g(p) = 0.$$

or, differentiating,

$$\begin{aligned} 1 - 2p_k &= \lambda, k = 1, \ldots, n, \\ \sum_{k=1}^{n} p_k &= c. \end{aligned}$$

Thus $n - 2c = n\lambda$, $\lambda = 1 - 2c/n$, and $1 - 2p_k = 1 - 2c/n$ or

$$p_k = c/n, k = 1, \ldots, n.$$

8. A system is called a "k out of n" system if it contains n components and it works whenever k or more of these components are working. Suppose that each component is working with probability p, independently of the other components, and let X_i be the indicator function of the event that component i is working. Then $X = \sum_{i=1}^{n} X_i$ is the number of components that are working. What is the distribution of X? What is the probability that the system works.

Answer. We can look at X as the total number of successes in n independent trials with a success probability p. Hence X ~bin(n, p), and

$$\mathbf{P}(X \geq k) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}.$$

9. Show that if Var$(X) = 0$ then X is almost surely constant; that is, there exists a $d \in \mathbf{R}$ such that $\mathbf{P}(X = d) = 1$. Hint. First show that if $\mathbf{E}(X^2) = 0$ then $\mathbf{P}(X = 0) = 1$.

10. (Continuation of #15 of week 5, #1 of hw5) A box contains b blue and r red balls (total number of balls in the box is $n = b + r$).

a) All balls are removed at random one by one and arranged in a row. Let X_i be the number of red balls between the $(i-1)$th and ith blue ball drawn, $i = 2, \ldots, b$; Let X_1 be the number of red balls until the first blue ball shows up, and X_{b+1} be the number of red balls after the last blue ball drawn. Consider the random vector $X = (X_1, \ldots X_{b+1})$. The range of X are all the vectors $(k_1, \ldots, k_{b+1})$ with nonnegative integer components $k_1, \ldots, k_{b+1}$ such that $k_1 + \ldots + k_{b+1} = r$. For such a vector $(k_1, \ldots, k_{b+1})$ with $k_i \geq 0$, and $k_1 + \ldots + k_{b+1} = r$, find

$$\mathbf{P}\left(X = (k_1, \ldots, k_{b+1})\right) = \mathbf{P}\left(X_1 = k_1, \ldots, X_{b+1} = k_{b+1}\right).$$

b) Find $\mathbf{E}(X_1), \ldots, \mathbf{E}(X_{b+1})$. Hint. The pmf of X in a) is symmetric in $(k_1, \ldots, k_{b+1})$: all the components $X_1, \ldots, X_{b+1}$ are identically distributed.

c) Let Y_i be the number of balls needed to be removed until the ith blue ball shows up, $i = 1, \ldots, b$. Find $\mathbf{E}(Y_i)$.

d) Find the pmf of X_1 and Y_1. What are the pmf of $X_1, \ldots, X_{b+1}$?

11. A population of N animals has had a number n of its members captured, marked, and released. Let X be the number of animals it is necessary to recapture (without re-release) in order to obtain m marked animals. Show that

$$\mathbf{P}(X = k) = \frac{n}{N} \frac{\binom{n-1}{m-1}\binom{N-n}{k-m}}{\binom{N-1}{k-1}}$$

and find $\mathbf{E}(X)$. This distribution has been called *negative hypergeometric*. Hint. See the previous problem **10**.

12. If one picks a numerical entry at random from an almanac, or the annual accounts of a corporation, the first two significant digits, X, Y, are found to have approximately the joint mass function

$$f(x, y) = \log_{10}\left(1 + \frac{1}{10x + y}\right), 1 \leq x \leq 9, 0 \leq y \leq 9,$$

x, y are integers. Find the mass function of X and an approximation to its mean. [A heuristic explanation for this phenomenon may be found in the second of Feller's volumes (1971).] Hint. A sum of logarithms is log of the product, $1 + \frac{1}{10x+y} = \frac{10x+y+1}{10x+y}$.

Answer. For $1 \leq x \leq 9$ (x is an integer),

$$\begin{aligned} f_X(x) &= \sum_{y=0}^{9} f(x, y) = \sum_{y=0}^{9} \log_{10}\left(1 + \frac{1}{10x + y}\right) \\ &= \log_{10}\left[\prod_{y=0}^{9}\left(1 + \frac{1}{10x + y}\right)\right] = \log_{10}\left[\prod_{y=0}^{9} \frac{10x + y + 1}{10x + y}\right]. \end{aligned}$$

Since

$$\prod_{y=0}^{9} \frac{10x + y + 1}{10x + y} = \frac{10x + 1}{10x} \frac{10x + 2}{10x + 1} \ldots \frac{10x + 10}{10x + 9} = \frac{10x + 10}{10x} = 1 + \frac{1}{x},$$

we have

$$f_X(x) = \log_{10}\left(1 + \frac{1}{x}\right), 1 \leq x \leq 9,$$

and $\mathbf{E}(X) = \sum_{x=1}^{9} x \log_{10}\left(1+\frac{1}{x}\right) = 3.4402$.

13. A fair coin was tossed twice. Let X_i be the number of heads in the ith toss, $i = 1, 2$. Let $X = X_1 + X_2, Y = X_1 - X_2$.

a) Find the joint probability mass function of X and Y. Are X and Y independent?

Answer. Range of X is $\{0, 1, 2\}$, Range of Y is $\{-1, 0, 1\}$. The joint pmf is

$$\begin{aligned} p(0,-1) &= 0, p(0,0) = \mathbf{P}(X_1 = 0, X_2 = 0) = \frac{1}{4}, p(0,1) = 0, \\ p(1,-1) &= \mathbf{P}(X_1 = 0, X_2 = 1) = \frac{1}{4}, p(1,0) = 0, p(1,1) = \frac{1}{4}, \\ p(2,-1) &= 0, p(2,0) = \mathbf{P}(X_1 = 1, X_2 = 1) = \frac{1}{4}, p(2,1) = 0. \end{aligned}$$

In the form of a table:

$X\backslash Y$	-1	0	1	$p_X(x)$
0	0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$p_Y(y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	

X and Y are not independent: $p(0,-1) \neq p_X(0)\, p_Y(-1)$.

b) Find $\mathrm{Cov}(X, Y)$ and $\rho(X, Y)$.

Answer. By linearity of covariance,

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathrm{Cov}(X_1 + X_2, X_1 - X_2) = \mathrm{Cov}(X_1, X_1) - \mathrm{Cov}(X_2, X_2) \\ &= \mathrm{Var}(X_1) - \mathrm{Var}(X_2) = 1/4 - 1/4 = 0. \end{aligned}$$

14. Let (X, Y) be discrete with a symmetric pmf $f(x,y) = \mathbf{P}(X = x, Y = y)$:

$$f(x,y) = f(y,x) \text{ for all } x, y \in \mathbf{R}.$$

Show that X and Y are identically distributed, that is $f_X(z) = f_Y(z)$ for all $z \in \mathbf{R}$.

Answer. By symmetry and relabeling, for any $z \in \mathbf{R}$,

$$\begin{aligned} f_X(z) &= \sum_y f(z,y) = \sum_y f(y,z) = \sum_u f(u,z), \\ f_Y(z) &= \sum_x f(x,z) = \sum_u f(u,z). \end{aligned}$$

15. Let X_n have binomial(n, p) distribution.

(a) Find $\mathbf{E}\left(\frac{1}{1+X_n}\right)$. Simplify your answer so it does not involve a sum up to n, n + 1, etc.

Answer. It is expectation of a function of X_n :

$$\begin{aligned} \mathbf{E}\left(\frac{1}{1+X_n}\right) &= \sum_{k=0}^{n} \frac{1}{1+k}\binom{n}{k} p^k q^{n-k} = \frac{1}{n+1}\sum_{k=0}^{n} \frac{1+n}{1+k}\binom{n}{k} p^k q^{n-k} \\ &= \frac{1}{n+1}\frac{1}{p}\sum_{k=0}^{n} \frac{(n+1)!}{(k+1)!\,(n+1-(k+1))!} p^{k+1} q^{(n+1)-(k+1)} \\ &= \frac{1}{p(n+1)}\left[1-(1-p)^{n+1}\right]. \end{aligned}$$

(b) Suppose $p = p_n$ and $np_n \to \lambda$ as $n \to \infty$, with $\lambda \in (0, \infty)$. Find $\lim_{n\to\infty} \mathbf{E}\frac{1}{1+X_n}$. Is it the same as $\lim_n \frac{1}{\mathbf{E}(1+X_n)}$?

Answer. We have

$$\begin{aligned} & \mathbf{E}\left(\frac{1}{1+X_n}\right) \\ = \quad & \frac{1}{p\,(n+1)}\left[1-(1-p)^{n+1}\right] = \frac{1}{np_n\frac{(n+1)}{n}}\left[1-\left(1-\frac{np_n}{n}\right)^n(1-p_n)\right] \\ \to \quad & \frac{1}{\lambda}\left(1-e^{-\lambda}\right) \text{ as } n \to \infty. \end{aligned}$$

Since X_n is binomial(n, p_n), $\mathbf{E}\,(X_n) = np_n \to \lambda$ as $n \to \infty$, and

$$\lim_n \frac{1}{\mathbf{E}\,(1+X_n)} = \frac{1}{1+\lambda} \neq \frac{1}{\lambda}\left(1-e^{-\lambda}\right).$$

16. *Coupons*. Every package of some intrinsically dull commodity includes a small and exciting plastic object. There are c different types of object, and each package is equally likely to contain any given type. You buy one package each day.

Find the expected number of different types of plastic objects you can collect in n days. What is the variance of that number?

7 Week 7: 10/3-10/7

1. a) The expected number of accidents at an industrial facility is 5 per week. The average number of injured people in each accident is 2.5. Assuming all the independence you need, compute the expected number of injured workers per week.

Answer. Let N be the number of accidents per week, $X_1, X_2, ...$ be the number of injured in those accidents. Assume independence of N and all $X_1, X_2, \ldots$The number of injured per week is

$$X = \sum_{k=1}^{N} X_k,$$

assuming $\sum_{k=1}^{N} X_k = 0$ if $N = 0$. Then

$$\begin{aligned}
\mathbf{E}(X|N=n) &= \mathbf{E}\left(\sum_{k=1}^{N} X_k|N=n\right) = \mathbf{E}\left(\sum_{k=1}^{n} X_k|N=n\right) = \mathbf{E}\left(\sum_{k=1}^{n} X_k\right) \\
&= \sum_{k=1}^{n} \mathbf{E}(X_k) = n \cdot 2.5 = 2.5n, n = 0, 1, \ldots.
\end{aligned}$$

Hence $\mathbf{E}(X|N) = 2.5N$, and

$$\mathbf{E}(X) = \mathbf{E}\{\mathbf{E}(X|N)\} = 2.5\mathbf{E}(N) = 2.5 \cdot 5 = 12.5.$$

b) The life time of the light bulb is characterized by the mean value μ and standard deviation σ. There are two types of light bulbs in a box, with the corresponding parameters μ_1, σ_1 and μ_2, σ_2. The proportion of type-1 bulbs in the box is p. A bulb is selected at random. Denote by X the life time of the this bulb. Compute the expected value and the variance of X.

Answer. Denoting by T the type of the bulb, we have

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= \mathbf{E}\left[\mathbf{E}\left(X^2|T\right)\right], \mathbf{E}(X) = \mathbf{E}[\mathbf{E}(X|T)], \\
\mathbf{E}[\mathbf{E}(X|T)] &= \mathbf{E}(X|T=1)\,\mathbf{P}(T=1) + \mathbf{E}(X|T=2)\,\mathbf{P}(T=2), \\
& \ldots
\end{aligned}$$

Then

$$\begin{aligned}
\mu &= \mathbf{E}(X) = \mathbf{E}[\mathbf{E}(X|T)] = \mathbf{E}(X|T=1)\,\mathbf{P}(T=1) + \mathbf{E}(X|T=2)\,\mathbf{P}(T=2) \\
&= \mu_1 p + \mu_2 q, \text{ where } q = 1 - p.
\end{aligned}$$

Now,

$$\mathrm{Var}(X) = \mathbf{E}\left(X^2\right) - \mu^2.$$

So we find first

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= \mathbf{E}\left[\mathbf{E}\left(X^2|T\right)\right] = \mathbf{E}\left(X^2|T=1\right)\mathbf{P}(T=1) + \mathbf{E}\left(X^2|T=2\right)\mathbf{P}(T=2) \\
&= \left(\sigma_1^2 + \mu_1^2\right) p + \left(\sigma_2^2 + \mu_2^2\right) q.
\end{aligned}$$

Then

$$\begin{aligned}
\text{Var}(X) &= \mathbf{E}\left(X^2\right) - \mu^2 \\
&= \left(\sigma_1^2 + \mu_1^2\right) p + \left(\sigma_2^2 + \mu_2^2\right) q - (\mu_1 p + \mu_2 q)^2 \\
&= pq\left(\mu_1 - \mu_2\right)^2 + p\sigma_1^2 + q\sigma_2^2.
\end{aligned}$$

2. We define the conditional variance, var$(Y|X)$, as a random variable

$$\text{var}(Y|X) = \mathbf{E}\left[(Y - \mathbf{E}(Y|X))^2 |X\right].$$

Show that

$$\text{var}(Y) = \mathbf{E}(\text{var}(Y|X)) + \text{var}(\mathbf{E}(Y|X)).$$

Hint. Use both theorems of 3.7.

3. A factory has produced n robots, each of which is faulty with probability p. To each robot a test is applied which detects the fault (if present) with probability δ (it passes all good robots). Let X be the number of faulty robots, and Y the number detected as faulty. Assume the usual independence.

(a) What is the probability that a robot passed is in fact faulty?

(b) Let Z be the number of passed faulty robots. Given $Y = k$, what is the distribution of Z? What is $\mathbf{E}(Z|Y)$?

(c) Show that

$$\mathbf{E}(X|Y) = \frac{np(1-\delta) + (1-p)Y}{1 - p\delta}.$$

4. a) Let X be geometric (waiting time or lifetime): $\mathbf{P}(X = k) = (1-p)^{k-1} p, k \in \{1, 2, \ldots\}$. Show that $\mathbf{P}(X = n + k|X > n) = \mathbf{P}(X = k), k \in \{1, 2, \ldots\}$. Why do you think that this is called the 'lack of memory' property? Find

$$\mathbf{E}(X - n|X > n) = \sum_{k=1}^{\infty} k\mathbf{P}(X - n = k|X > n).$$

Recall $\mathbf{E}(X) =$?

Answer. First, $\mathbf{P}(X > n) = (1-p)^n, n \geq 1$. By definition, for $k \geq 1$,

$$\begin{aligned}
\mathbf{P}(X = n + k|X > n) &= \frac{\mathbf{P}(X = n + k, X > n)}{\mathbf{P}(|X > n)} = \frac{\mathbf{P}(X = n + k)}{\mathbf{P}(|X > n)} \\
&= \frac{(1-p)^{n+k-1} p}{(1-p)^n} = (1-p)^{k-1} p.
\end{aligned}$$

Thus, given $X > n$, $X - n$ has the same geometric distribution as X. If X were our lifetime, i.e. X is our death moment, then given we survived n time units (given $X > n$), the remaining lifetime $X - n$ has the same distribution as the initial X: it is forgotten that we already lived n time units.

Since, given $X > n$, $X - n$ has the same geometric distribution as X,

$$\mathbf{E}(X - n|X > n) = \mathbf{E}(X) = \frac{1}{p}.$$

b) Show that the sum of two independent binomial variables, $\text{bin}(m, p)$ and $\text{bin}(n, p)$ respectively, is $\text{bin}(m+n, p)$. Explain without computing: think about standard interpretation of a binomial r.v.

Answer. Let $X \sim\text{bin}(m, p)$ and $Y \sim\text{bin}(n, p)$ be independent. We can interpret X as the number of heads in m independent tosses of a coin with $P(H) = p$, and Y as the number of heads in n independent tosses of a coin with $P(H) = p$. Then $X + Y$ is the number of heads in $n + m$ independent tosses of a coin with $P(H) = p$, that is $X + Y \sim\text{bin}(m+n, p)$.

5. Let X be binomial(k, p) and Y be binomial(l, p). Assume X and Y are independent, and let $Z = X + Y$. Show that the conditional distribution of X given $Z = n$ is a hypergeometric distribution (see #5 of hw6).

Answer. According to #4b) above, $Z = X + Y \sim\text{bin}(N = k + l, p)$. Thus for $i \leq n \leq k \leq N = k + l$, by definition, denoting $q = 1 - p$,

$$\begin{aligned}
\mathbf{P}(X = i|X + Y = n) &= \frac{\mathbf{P}(X = i, X + Y = n)}{\mathbf{P}(X + Y = n)} = \frac{\mathbf{P}(X = i, Y = n - i)}{\mathbf{P}(X + Y = n)} \\
&= \frac{\binom{k}{i} p^i q^{k-i} \binom{l}{n-i} p^{n-i} q^{l-(n-i)}}{\binom{N}{n} p^n q^{N-n}} = \frac{\binom{k}{i}\binom{l}{n-i}}{\binom{N}{n}}.
\end{aligned}$$

Distribution is hypergeometric: population size N, two different parts ($N = k + l$) with the sample of size n.

6. Show the following:

(a) $\mathbf{E}(aY + bZ|X) = a\mathbf{E}(Y|X) + b\mathbf{E}(Z|X)$ for $a, b \in \mathbf{R}$.

(b) $\mathbf{E}(Y|X) \geq 0$ if $Y \geq 0$;

(c) $\mathbf{E}(c|X) = c$;

(d) if X and Y are independent then $\mathbf{E}(Y|X) = \mathbf{E}(Y)$;

(e) ('pull-through property') $\mathbf{E}(Yg(X)|X) = g(X)\mathbf{E}(Y|X)$ for any suitable function g;

(f) ('tower property') $\mathbf{E}[\mathbf{E}(Y|X, Z)|X] = \mathbf{E}(Y|X) = \mathbf{E}[\mathbf{E}(Y|X)|X, Z]$.

Answer. Let $f(x, y, z)$ be the joint mass function of X, Y, Z. Then

$$f_{X,Y}(x, y) = \sum_z f(x, y, z), \; f_{Z,X} = \sum_y f(x, y, z),$$

and

$$\begin{aligned}
& \mathbf{E}(aY + bZ|X = x) \\
= & \sum_{y,z} (ay + bz)\mathbf{P}(Y = y, Z = z|X = x) = \sum_{y,z} (ay + bz) \frac{\mathbf{P}(Y = y, Z = z, X = x)}{\mathbf{P}(X = x)} \\
= & a \sum_y y \frac{\sum_z \mathbf{P}(Y = y, Z = z, X = x)}{\mathbf{P}(X = x)} + b \sum_z bz \sum_y \frac{\mathbf{P}(Y = y, Z = z, X = x)}{\mathbf{P}(X = x)} \\
= & a \sum_y y \frac{\mathbf{P}(Y = y, X = x)}{\mathbf{P}(X = x)} + b \sum_z z \frac{\mathbf{P}(Z = z, X = x)}{\mathbf{P}(X = x)} \\
= & a \sum_y y\mathbf{P}(Y = y|X = x) + b \sum_z z\mathbf{P}(Z = z|X = x) = a\mathbf{E}(Y|X = x) + b\mathbf{E}(Z|X = x)
\end{aligned}$$

(b) $\mathbf{E}(Y|X) \geq 0$ if $Y \geq 0$;

Answer. Assume Y assumes only nonnegative values. Then

$$\mathbf{E}(Y|X=x) = \sum_y y\mathbf{P}(Y=y|X=x) = \sum_{y\geq 0} y\frac{\mathbf{P}(Y=y, X=x)}{\mathbf{P}(X=x)} \geq 0.$$

(c) $\mathbf{E}(c|X) = c$;
Answer. Obvious.
(d) if X and Y are independent then $\mathbf{E}(Y|X) = \mathbf{E}(Y)$;
Answer. Indeed, for any x such that $\mathbf{P}(X=x) > 0$,

$$\begin{aligned}\mathbf{E}(Y|X=x) &= \sum_y y\mathbf{P}(Y=y|X=x) = \sum_y y\frac{\mathbf{P}(Y=y, X=x)}{\mathbf{P}(X=x)} \\ &= \sum_y y\frac{\mathbf{P}(Y=y)\mathbf{P}(X=x)}{\mathbf{P}(X=x)} = \sum_y y\mathbf{P}(Y=y) = \mathbf{E}(Y).\end{aligned}$$

(e) ('pull-through property') $\mathbf{E}(Yg(X)|X) = g(X)\mathbf{E}(Y|X)$ for any suitable function g;
Answer. Indeed, for any $u \in \mathbf{R}$ such that $\mathbf{P}(X=u) > 0$,

$$\begin{aligned}\mathbf{E}(Yg(X)|X=u) &= \sum_{y,x} yg(x)\mathbf{P}(Y=y, X=z|X=u) \\ &= \sum_{y,x} yg(x)\frac{\mathbf{P}(Y=y, X=x, X=u)}{\mathbf{P}(X=u)} \\ &= \sum_y yg(u)\frac{\mathbf{P}(Y=y, X=u)}{\mathbf{P}(X=u)} = g(u)\sum_y y\mathbf{P}(Y=y|X=u) \\ &= g(u)E(Y|X=u)\end{aligned}$$

(f) ('tower property') $\mathbf{E}\{\mathbf{E}(Y|X,Z)|X\} = \mathbf{E}\{Y|X\} = \mathbf{E}\{\mathbf{E}(Y|X)|X,Z\}$.
Answer. Similarly as in (e) we can show that

$$\begin{aligned}\mathbf{E}(g(X,Y)|X=x) &= \mathbf{E}(g(x,Y)|X=x), x \in \mathbf{R}, \\ \mathbf{E}(h(X,Y,Z)|X=x, Z=z) &= \mathbf{E}(h(x,Y,z)|X=x, Z=z), x \in \mathbf{R}.\end{aligned} \tag{7.4}$$

First, for any x, z,

$$\mathbf{E}(Y|X=x, Z=z) = \sum_y y\mathbf{P}(Y=y|X=x, Z=z),$$

and, by (7.4),

$$\begin{aligned}&\mathbf{E}\{\mathbf{E}(Y|X,Z)|X=x\} \\ =& \sum_z \mathbf{E}(Y|X=x, Z=z)\mathbf{P}(Z=z|X=x) = \sum_z\sum_y y\mathbf{P}(Y=y|X=x, Z=z)\mathbf{P}(Z=z|X=x) \\ =& \sum_z\sum_y y\frac{\mathbf{P}(Y=y, X=x, Z=z)}{\mathbf{P}(X=x, Z=z)}\frac{\mathbf{P}(Z=z, X=x)}{\mathbf{P}(X=x)} = \sum_z\sum_y y\frac{\mathbf{P}(Y=y, X=x, Z=z)}{\mathbf{P}(X=x)} \\ =& \sum_y y\sum_z \frac{\mathbf{P}(Y=y, X=x, Z=z)}{\mathbf{P}(X=x)}\sum_y y\frac{\mathbf{P}(Y=y, X=x)}{\mathbf{P}(X=x)} = \sum_y y\mathbf{P}(Y=y|Z=x) = \mathbf{E}(Y|X=x),\end{aligned}$$

i.e. $\mathbf{E}\{\mathbf{E}(Y|X,Z)|X\} = \mathbf{E}(Y|X)$. On the other hand, by (7.4),

$$\mathbf{E}\{\mathbf{E}(Y|X)|X=x, Z=z\} = \mathbf{E}(Y|X=x) \text{ for all } x, z,$$

i.e. $\mathbf{E}\{\mathbf{E}(Y|X)|X,Z\} = \mathbf{E}(Y|X)$.

7. The lifetime of a machine (in days) is a random variable T with mass function f. Given that the machine is working after t days, what is the mean subsequent lifetime of the machine when:

$f(x) = (N+1)^{-1}$ for $x \in \{0, 1, 2, \ldots, N\}$.

Answer. For $N - t \geq k \geq 1$,

$$\mathbf{P}(T = t+k|T > t) = \mathbf{P}(T - t = k|T > t) = \frac{\mathbf{P}(T = t+k)}{\mathbf{P}(T > t)} = \frac{\frac{1}{N+1}}{\frac{N-t}{N+1}} = \frac{1}{N-t}.$$

Hence

$$\mathbf{E}(T - t|T > t) = \frac{1}{N-t}\sum_{k=1}^{N-t} k = \frac{1}{2}(N - t + 1).$$

8. Let X, Y have the joint mass function

$$f(x, y) = \frac{C}{(x+y-1)(x+y)(x+y+1)}, x, y \in \{1, 2, \ldots\}.$$

Find the mass function of $V = X - Y$.

Answer. The mass function of $X, -Y$ is

$$\begin{aligned} g(x,y) &= \mathbf{P}(X = x, -Y = y) = \mathbf{P}(X = x, Y = -y) = f(x, -y) \\ &= \frac{C}{(x-y-1)(x-y)(x-y+1)}, x, -y \in \{1, 2, \ldots\}. \end{aligned}$$

Hence for $v \in \mathbf{Z} = \{0, \pm 1, \ldots\}$,

$$\begin{aligned} \mathbf{P}(V = v) &= \mathbf{P}(X + (-Y) = v) \\ &= \sum_x g(x, v - x) = \sum_x f(x, x - v) \\ &= \sum_{x \geq 1, x \geq v+1} \frac{C}{(2x - v - 1)(2x - v)(2x - v + 1)}. \end{aligned}$$

For $v \leq 0$,

$$\begin{aligned} f_V(v) &= \sum_{x=1}^{\infty} \frac{C}{(2x-v-1)(2x-v)(2x-v+1)} = \sum_{2x-v=2-v}^{\infty} \frac{C}{(2x-v-1)(2x-v)(2x-v+1)} \\ &= \sum_{z=2-v}^{\infty} \frac{C}{(z-1)z(z+1)} = \frac{C}{2}\sum_{z=2-v}^{\infty}\left[\left(\frac{1}{z-1} - \frac{1}{z}\right) - \left(\frac{1}{z} - \frac{1}{z+1}\right)\right] \\ &= \frac{C}{2}\left(\frac{1}{1-v} - \frac{1}{2-v}\right) \end{aligned}$$

For $v > 0$,

$$\begin{aligned}
f_V(v) &= \sum_{x=1+v}^{\infty} \frac{C}{(2x-v-1)(2x-v)(2x-v+1)} = \sum_{2x-v=2+v}^{\infty} \frac{C}{(2x-v-1)(2x-v)(2x-v+1)} \\
&= \sum_{z=2+v}^{\infty} \frac{C}{(z-1)z(z+1)} = \frac{C}{2}\sum_{z=2+v}^{\infty}\left[\left(\frac{1}{z-1}-\frac{1}{z}\right)-\left(\frac{1}{z}-\frac{1}{z+1}\right)\right] \\
&= \frac{C}{2}\left(\frac{1}{1+v}-\frac{1}{2+v}\right)
\end{aligned}$$

Hence

$$f_V(v) = \frac{C}{2}\left(\frac{1}{1+|v|}-\frac{1}{2+|v|}\right), v \in \mathbf{R}.$$

9. Let X, Y be independent geometric random variables with respective parameters α, β. Show that

$$\mathbf{P}(X+Y=z) = \frac{\alpha\beta}{\alpha-\beta}\left\{(1-\beta)^{z-1}-(1-\alpha)^{z-1}\right\}.$$

Answer. We have

$$\begin{aligned}
f_X(x) &= (1-\alpha)^{x-1}\alpha, x \geq 1, \\
f_Y(y) &= (1-\beta)^{y-1}\beta, y \geq 1,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{P}(X+Y=z) &= \sum_{x\geq 1, z-x\geq 1} f_X(x) f_Y(z-x) = \sum_{x=1}^{z-1} f_X(x) f_Y(z-x) \\
&= \sum_{x=1}^{z-1}(1-\alpha)^{x-1}(1-\beta)^{z-x-1}\alpha\beta = \alpha\beta\frac{(1-\beta)^z}{(1-\alpha)(1-\beta)}\sum_{x=1}^{z-1}\left(\frac{1-\alpha}{1-\beta}\right)^x \\
&= \alpha\beta\frac{(1-\beta)^z}{(1-\alpha)(1-\beta)}\frac{\frac{1-\alpha}{1-\beta}-\left(\frac{1-\alpha}{1-\beta}\right)^z}{1-\frac{1-\alpha}{1-\beta}} = \alpha\beta\frac{(1-\beta)^{z-1}-(1-\alpha)^{z-1}}{\alpha-\beta}.
\end{aligned}$$

10. Let $\{X_k, 1 \leq k \leq n\}$ be independent geometric random variables with parameter p. Show that $Z = \sum_{k=1}^{n} X_k$ has a negative binomial distribution. [Hint: No calculations are necessary.]

Answer. Consider a sequence of independent trials with probability of a success p. We can interpret X_1 as the waiting time for the first success, X_i is the time between $i-1$ and ith success, $i = 2, \ldots, n$. Then $Z = \sum_{k=1}^{n} X_k$ is the waiting time for the nth success. We know it has the negative binomial distribution.

11. Consider two coins: coin 1 shows heads with probability p_1 and coin 2 shows heads with probability p_2. Each coin is tossed repeatedly. Let T_i be the time of first heads for coin i, and define the event $A = \{T_1 < T_2\}$.

(a) Find $\mathbf{P}(A)$.

(b) Find $\mathbf{P}(T_1 = k|A)$ for all $k \geq 1$.

Answer. (a) Since $A = \cup_{j=1}^{\infty} \{j < T_2, T_1 = j\}$,

$$\begin{aligned}
\mathbf{P}(A) &= \sum_{j=1}^{\infty} \mathbf{P}(T_1 = j)\,\mathbf{P}(T_2 > j) = \sum_{j=1}^{\infty} q_2^j \mathbf{P}(T_1 = j) = \sum_{j=1}^{\infty} q_2^j q_1^{j-1} p_1 = \frac{p_1}{q_1} \sum_{j=1}^{\infty} (q_1 q_2)^j \\
&= \frac{p_1}{q_1} \frac{q_1 q_2}{1 - q_1 q_2}
\end{aligned}$$

(b) Since $A \{T_1 = k\} = \{k < T_2, T_1 = k\}$,

$$\begin{aligned}
\mathbf{P}(T_1 &= k|A) = \frac{\mathbf{P}(A \cap \{T_1 = k\})}{\mathbf{P}(A)} = \frac{P(k < T_2, T_1 = k)}{\mathbf{P}(A)} = \frac{\mathbf{P}(k < T_2)\,\mathbf{P}(T_1 = k)}{\mathbf{P}(A)} \\
&= \frac{q_2^k q_1^{k-1} p_1}{\frac{p_1}{q_1} \frac{q_1 q_2}{1 - q_1 q_2}} = (q_1 q_2)^{k-1} (1 - q_1 q_2)\,, k = 1, 2, \ldots,
\end{aligned}$$

that is given A, T_1 is geometric with $p = 1 - q_1 q_2$.

12. *Voter paradox.* Let X, Y, Z be discrete random variables with the property that their values are distinct with probability 1. Let $a = \mathbf{P}(X > Y)\,, b = \mathbf{P}(Y > Z)\,, c = \mathbf{P}(Z > X)\,.$

(a) Show that $\min\{a, b, c\} \le 2/3$.

Answer. Since

$$\begin{aligned}
&\mathbf{P}(X > Y) + \mathbf{P}(Y > Z) + \mathbf{P}(Z > X) \\
= \;&\mathbf{E}\left(I_{\{X > Y\}} + I_{\{Y > Z\}} + I_{\{Z > X\}}\right) \le 2,
\end{aligned}$$

it follows that $\min\{a, b, c\} \le 2/3$.

(b) Show that, if X, Y, Z are independent and identically distributed, then $a = b = c = 1/2$.

Answer. Let $f(u) = f_X(u) = f_Y(u) = f_Z(u)\,, u \in \mathbf{R}$. Then, relabeling summation variables,

$$\mathbf{P}(X < Y) = \sum_{x < y} f(x) f(y) = \sum_{x > y} f(x) f(y) = \mathbf{P}(X > Y) = \frac{1}{2}.$$

Similarly, $\mathbf{P}(X > Z) = \mathbf{P}(X < Z) = \frac{1}{2}$, and $\mathbf{P}(Y > Z) = \mathbf{P}(Y < Z)\,.$

(c) Find $\min\{a, b, c\}$ and $\sup_p \min\{a, b, c\}$ when $\mathbf{P}(X = 0) = 1$, and Y, Z are independent with

$$\begin{aligned}
\mathbf{P}(Z = 1) &= \mathbf{P}(Y = -1) = p, \\
\mathbf{P}(Z = -2) &= \mathbf{P}(Y = 2) = 1 - p.
\end{aligned}$$

Answer. We find

$$\begin{aligned}
a &= \mathbf{P}(X > Y) = \mathbf{P}(Y = -1) = p, \\
b &= \mathbf{P}(Y > Z) = \mathbf{P}(Y = -1, Z = -2) + \mathbf{P}(Y = 2) \\
&= \mathbf{P}(Y = -1)\,\mathbf{P}(Z = -2) + \mathbf{P}(Y = 2) = p(1 - p) + 1 - p = 1 - p^2, \\
c &= \mathbf{P}(Z > X) = \mathbf{P}(Z = 1) = p.
\end{aligned}$$

Hence

$$\min\{a, b, c\} = \min\left\{p, 1 - p^2\right\} = \begin{cases} p = a = c & \text{if } 0 \le p \le \frac{-1+\sqrt{5}}{2} \\ b = 1 - p^2 & \frac{-1+\sqrt{5}}{2} \le p \le 1. \end{cases}$$

Thus

$$\sup_p \min\{a, b, c\} = \frac{-1+\sqrt{5}}{2} = 0.618\,03.$$

[Part (a) is related to the observation that, in an election, it is possible for more than half of the voters to prefer candidate A to candidate B, more than half B to C, and more than half C to A].

13. *Mutual information.* Let X, Y be discrete random variables with joint mass function $f(x, y)$.

(a) Show that $\mathbf{E}(\log f_X(X)) \geq \mathbf{E}(\log f_Y(X))$.

(b) Show that the mutual information

$$I(X, Y) := \mathbf{E}\left(\log\left\{\frac{f(X, Y)}{f_X(X)\, f_Y(Y)}\right\}\right)$$

satisfies $I(X, Y) \geq 0$ with equality if and only if X, Y are independent. Hint:

(i) $\log y \leq y - 1$, $y>0$, with equality if and only if $y = 1$. In particular, $\log y \geq 1 - \frac{1}{y}$, $y>0$.

(ii) If $Z \geq 0$ and $\mathbf{E}(Z) = 0$, then $Z = 0$ with probability 1.

(c) Show that if $Y = X$, then $I(X, X) = -\mathbf{E}(\log f_X(X))$.

Comment. I measures how much knowing one of these variables reduces uncertainty about the other. $I(X, X)$ is called the entropy of X. The entropy reflects how much information we learn on average from an observation of X.

Answer. Let $l(y) = \log y + 1 - y$, $y > 0$. Then $l(1) = 0$, $l'(y) = \frac{1}{y} - 1 \leq 0$ if $y > 1$, and $l'(y) > 0$ if $y \in (0, 1)$. Hence $l(y) \leq 0$, $y > 0$, and $l(y) = 0$ iff $y = 1$. Equivalently, replacing y by $1/y$,

$$\log y + \frac{1}{y} - 1 \geq 0,\ y > 0, \tag{7.5}$$

and $\log y + \frac{1}{y} - 1 = 0$ iff $y = 1$.

(a) We assume that $\mathbf{P}(Y = x) > 0$ if $\mathbf{P}(X = x) > 0$, and both expected values make sense. By inequality (7.5),

$$\begin{aligned}
&\mathbf{E}(\log f_X(X)) - \mathbf{E}(\log f_Y(X)) \\
= \;& \mathbf{E}[\log f_X(X) - \log f_Y(X)] = \mathbf{E}\left(\log \frac{f_X(X)}{f_Y(X)}\right) \\
\geq \;& 1 - \mathbf{E}\frac{f_Y(X)}{f_X(X)} = 1 - \sum_x \frac{f_Y(x)}{f_X(x)} \mathbf{P}(X = x) \geq 1 - \sum_x f_Y(x) = 0.
\end{aligned}$$

(b) Assume expected value exists. By inequality (7.5),

$$\begin{aligned}
&\mathbf{E}\left(\log\left\{\frac{f(X, Y)}{f_X(X)\, f_Y(Y)}\right\}\right) \\
\geq \;& 1 - \mathbf{E}\left(\frac{f_X(X)\, f_Y(Y)}{f(X, Y)}\right) = 1 - \sum_{x,y} \frac{f_X(x)\, f_Y(y)}{f(x, y)} \mathbf{P}(X = x, Y = y) \\
\geq \;& 1 - \sum_{x,y} f_X(x)\, f_Y(y) = 1 - \sum_x f_X(x) \sum_y f_Y(y) = 0.
\end{aligned}$$

Obviously, if X, Y are independent, then $f(x, y) = f_X(x)\, f_Y(y)$, and $I(X, Y) = \log 1 = 0$.

Assume now that $I(X,Y) = 0$. Consider a nonnegative r.v. $Z = \log\left\{\frac{f(X,Y)}{f_X(X)f_Y(Y)}\right\} + 1 - \frac{f_X(X)f_Y(Y)}{f(X,Y)} \geq 0$. Then, assuming $f(x,y) \neq 0$ if $f_X(x)\, f_Y(y) \neq 0$,

$$\mathbf{E}(Z) = I(X,Y) + 1 - \mathbf{E}\left(\frac{f_X(X)\, f_Y(Y)}{f(X,Y)}\right) = 0.$$

Hence $Z = \log\left\{\frac{f(X,Y)}{f_X(X)f_Y(Y)}\right\} + 1 - \frac{f_X(X)f_Y(Y)}{f(X,Y)} = 0$, equivalently, $\frac{f_X(X)f_Y(Y)}{f(X,Y)} = 1$ with probability 1.

(c) Show that if $Y = X$, then $I(X,X) = -\mathbf{E}(\log f_X(X))$.

Answer. Let $Y = X$. The joint distribution function in this case,

$$f(x,y) = \mathbf{P}(X = x, X = y) = \begin{cases} f_X(x) & \text{if } x = y, \\ 0 & \text{if } x \neq y \end{cases}.$$

Hence, since $X = Y$,

$$\begin{aligned} \mathbf{E}\left(\log\left\{\frac{f(X,Y)}{f_X(X)\, f_Y(Y)}\right\}\right) &= \mathbf{E}\left(\log\left\{\frac{f_X(X)}{f_X(X)\, f_X(X)}\right\}\right) \\ &= --\mathbf{E}(\log f_X(X)). \end{aligned}$$

8 Week 8-9: 10/10-10/12, 10/17 -10/19

1. Let τ_k be the time which elapses before a simple random walk is absorbed at either of the absorbing barriers at 0 and N, having started at k where $0 \leq k \leq N$. Show that $\mathbf{P}(\tau_k < \infty) = 1$.

Answer. See Remark 1 in class note of 10/11, and answer to #1 of hw3.

2. For simple random walk S_n with absorbing barriers at 0 and N, let W be the event that the particle is absorbed at 0 rather than N, and let $p_k = \mathbf{P}(W|S_0 = k)$, $0 < k < N$. Show that, if the particle starts at k, the conditional probability that the first step is rightwards, given W, equals $\frac{pp_{k+1}}{p_k}$. Deduce that the mean (expected) duration of the walk, given W, satisfies the equation

$$pp_{k+1}J_{k+1} - p_k J_k + (p_k - pp_k) J_{k-1} = -p_k, 0 < k < N.$$

Take $J_0 = 0$ as a boundary condition. Find J_k in the symmetric case $p = 1/2$.

Hint. Let Y_k be the duration of the walk (number of steps until absorption) when $S_0 = k$. Then $J_k = \mathbf{E}(Y_k|W)$, and

$$\begin{aligned} J_k &= \mathbf{E}(Y_k|\text{1st step rightwards, given } W)\,\mathbf{P}(\text{1st step rightwards}|W) \\ &\quad + \mathbf{E}(Y_k|\text{1st step leftwards, given } W)\,\mathbf{P}(\text{1st step lefttwards}|W). \end{aligned}$$

Answer. We find

$$\begin{aligned} \mathbf{P}(\text{1st step rightwards}|W) &= \frac{\mathbf{P}(\text{1st step rightwards}, W)}{\mathbf{P}(W)} \\ &= \frac{\mathbf{P}(W|\text{1st step rightwards})\,\mathbf{P}(\text{1st step rightwards})}{\mathbf{P}(W)} = \frac{p_{k+1}p}{p_k}. \end{aligned}$$

So,

$$\mathbf{P}(\text{1st step leftwards}|W) = 1 - \frac{p_{k+1}p}{p_k}.$$

Let Y_k be the duration of the walk (number of steps until absorption) when $S_0 = k$. Then $J_k = \mathbf{E}(Y_k|W)$, and

$$\begin{aligned} J_k &= \mathbf{E}(Y_k|\text{1st step rightwards, } W)\,\mathbf{P}(\text{1st step rightwards}|W) \\ &\quad + \mathbf{E}(Y_k|\text{1st step leftwards, } W)\,\mathbf{P}(\text{1st step lefttwards}|W). \end{aligned}$$

Now,

$$\begin{aligned} \mathbf{E}(Y_k|\text{1st step rightwards, } W) &= 1 + J_{k+1}, \\ \mathbf{E}(Y_k|\text{1st step leftwards, } W) &= 1 + J_{k-1}, \end{aligned}$$

and

$$J_k = (1 + J_{k+1})\frac{p_{k+1}p}{p_k} + (1 + J_{k-1})\left(1 - \frac{p_{k+1}p}{p_k}\right).$$

Multiplying by p_k,

$$pJ_k = p_{k+1}pJ_{k+1} + (p_k - p_{k+1}p) J_{k-1} + p_k.$$

3. Consider a simple random walk on the set $\{0, 1, \ldots, N\}$ in which each step is to the right with probability p or to the left with probability $q = 1 - p$. Absorbing barriers are placed at 0 and N.

Show that the number X of positive steps of the walk before absorption satisfies

$$\mathbf{E}(X) = \frac{1}{2}\{D_k - k + N(1 - p_k)\},$$

where D_k is the mean number of steps until absorption and p_k is the probability of absorption at 0. Hint. If Z_k is the number of steps until absorption, and Y is the number of negative steps until absorption, then $D_k = \mathbf{E}(X + Y)$. What can you say about $k + X - Y$? How many values it takes?

4. A coin is tossed repeatedly, heads turning up with probability p on each toss. Player A wins the game if m heads appear before n tails have appeared, and player B wins otherwise. Let $p_{m,n}$ be the probability that A wins the game. Set up a difference equation for the $p_{m,n}$.

What are the boundary conditions?

Answer. Let E_{mn} be the event that m heads appear before n tails. Then

$$\begin{aligned} p_{mn} &= \mathbf{P}(E_{mn}) = \mathbf{P}(E_{mn} |\text{ 1st H})\,\mathbf{P}(\text{1st H}) + \mathbf{P}(E_{mn} |\text{ 1st T})\,\mathbf{P}(\text{1st T}) \\ &= p_{m-1,n}\cdot p + p_{m,n-1}\cdot(1-p),\, m, n \geq 1. \end{aligned}$$

Boundary conditions: $p_{0n} = 1$, $p_{m0} = 0$.

5. For a symmetric simple random walk starting at 0, show that the mass function of the maximum satisfies $\mathbf{P}(M_n = r) = \mathbf{P}(S_n = r) + \mathbf{P}(S_n = r + 1)$ for $r \geq 0$. Hint: see (13), p. 78.

Answer. According to (13), p. 78, for $r \geq 0$,

$$\mathbf{P}(M_n \geq r) = \mathbf{P}(S_n = r) + 2\mathbf{P}(S_n \geq r + 1).$$

Since $\{M_n \geq r + 1\} \subseteq \{M_n \geq r\}$, and $\mathbf{P}(S_n \geq r + 1) = \mathbf{P}(S_n = r + 1) + \mathbf{P}(S_n \geq r + 2)$,

$$\begin{aligned} \mathbf{P}(M_n = r) &= \mathbf{P}(M_n \geq r) - \mathbf{P}(M_n \geq r + 1) \\ &= \mathbf{P}(S_n = r) + 2\mathbf{P}(S_n \geq r + 1) \\ &\quad -\mathbf{P}(S_n = r + 1) - 2\mathbf{P}(S_n \geq r + 2) \\ &= \mathbf{P}(S_n = r) + \mathbf{P}(S_n = r + 1). \end{aligned}$$

6. Let S_n be symmetric simple r.w., $p = q = 1/2$. Let $S_0 = 0$. According to (2), p. 76,

$$\mathbf{P}(S_{2k} = 0) = \binom{2k}{k} 2^{-2k}.$$

Show that

a)

$$\lim_{k\to\infty} \frac{\mathbf{P}(S_{2k} = 0)}{1/\sqrt{\pi k}} = 1.$$

Hint. Use Stirling's formula:

$$\lim_{k\to\infty} \frac{k!}{k^k e^{-k}\sqrt{2\pi k}} = 1.$$

b)

$$\mathbf{E}\left(\frac{\sum_{k=n+1}^{2n} I_{\{S_{2k}=0\}}}{2n}\right) \to 0$$

as $n \to \infty$.

Answer. By part a),

$$\begin{aligned}
\mathbf{E}\left(\frac{\sum_{k=n+1}^{2n} I_{\{S_{2k}=0\}}}{2n}\right) &= \frac{\sum_{k=n+1}^{2n} \mathbf{P}\left(S_{2k}=0\right)}{2n} \sim \frac{1}{2n}\sum_{k=n+1}^{2n}\frac{1}{\sqrt{\pi k}} \\
&= \frac{1}{2\sqrt{n}}\sum_{k=n+1}^{2n}\frac{1}{\sqrt{\pi\frac{k}{n}}}\frac{1}{n} \sim \frac{1}{2\sqrt{n}}\int_1^2 \frac{1}{\sqrt{\pi x}}dx = \frac{\sqrt{2}-1}{\sqrt{\pi n}} \to 0
\end{aligned}$$

as $n \to \infty$.

7. Let S_n be symmetric simple r.w. ($p = q = 1/2$), and $S_0 = 0$, i.e.,

$$S_n = X_1 + \ldots + X_n, n \geq 1,$$

where X_i are independent identically distributed, $\mathbf{P}\left(X_i = 1\right) = \mathbf{P}\left(X_i = -1\right) = 1/2$.

a) Show that $\bar{S}_n = -S_n, n \geq 0$, is symmetric r.w. as well, that is the sequences $\{S_n, n \geq 0\}$, and $\{-S_n, n \geq 0\}$ are identically distributed. Hint: X_i and $-X_i$ have identical mass functions, and $-X_i$ are independent.

b) For $b \neq 0$, set $\tau_b = \tau_b(S) = \min\{n > 0 : S_n = b\}$. Show that

$$\mathbf{P}\left(\tau_b < \tau_{-b}\right) = \mathbf{P}\left(\tau_{-b} < \tau_b\right) = 1/2.$$

Hint. For any $a \neq 0$, $\mathbf{P}\left(\tau_a < \infty\right) = 1$. Since the sequences $\{S_n, n \geq 0\}$, and $\{-S_n, n \geq 0\}$ are identically distributed,

$$\mathbf{P}\left(\tau_b(S) < \tau_{-b}(S)\right) = \mathbf{P}\left(\tau_b(-S) < \tau_{-b}(-S)\right),$$

where

$$\begin{aligned}
\tau_b(S) &= \min\{n > 0 : S_n = b\}, \tau_b(-S) = \min\{n > 0 : -S_n = b\}, \\
\tau_{-b}(S) &= \min\{n > 0 : S_n = -b\}, \tau_{-b}(-S) = \min\{n > 0 : -S_n = -b\}.
\end{aligned}$$

c) Let $\sigma_k = \min\{n > 0 : S_n \notin (-k, k)\}$. Find $\mathbf{E}\left(S_{\sigma_k}\right)$ and $\operatorname{var}\left(S_{\sigma_k}\right)$. Hint: $\sigma_k = \min\{\tau_k, \tau_{-k}\}$ and $\mathbf{P}\left(\tau_k < \infty\right) = \mathbf{P}\left(\tau_{-k} < \infty\right) = 1$. What values S_{τ_k} and $S_{\tau_k}^2$ take?

8. Consider a symmetric simple random walk S_n with $S_0 = 0$. Let $T = \min\{n \leq 1 : S_n = 0\}$ be the time of the first return of the walk to its starting point. Show that

$$\mathbf{P}\left(T = 2n\right) = \frac{1}{2n-1}\binom{2n}{n}2^{-2n}.$$

Hint. Use total probability law with $\Omega = \{S_1 = 1\} \cup \{S_1 = -1\}$:1st step analysis?

9. Consider a random walk $S_n = a + X_1 + \ldots + X_n, n \geq 1, S_0 = a$ with an integer a and i.i.d. X_i such that $\mathbf{P}\left(X_i = 1\right) = p, \mathbf{P}\left(X_i = -1\right) = 1 - p$, $p \in (0, 1)$. For the integers i, j, let $E_{i,j}$ be the event that j is reached when $S_0 = i$, and $p_{i,j} = \mathbf{P}\left(E_{i,j}\right)$. Let $a > 0$.

a) Show that

$$p_{a,0} = \left(p_{1,0}\right)^a.$$

Hint. use conditioning and space/time homogeneity of the random walk; $E_{i,j}$ means $S_0 = i$ and $S_n = j$ for some n.

Answer. Since $E_{a,0} \subseteq E_{a,a-1}$, by multiplication formula and space time homogeneity of the random walk,

$$\begin{aligned} p_{a,0} &= \mathbf{P}\left(E_{a,0}\right) = \mathbf{P}\left(E_{a,0}|E_{a,a-1}\right)\mathbf{P}\left(E_{a,a-1}\right) \\ &= p_{a-1,0}\mathbf{P}\left(E_{a,a-1}\right) = p_{a-1,0}p_{1,0} = \ldots = p_{1,0}^a. \end{aligned}$$

b) Find $p_{1,0}$. Hint. Having in mind that $S_0 = 1$, and conditioning on the first step, derive an equation for $p_{1,0}$, and solve it.

Answer. Conditioning on the first step, and using part a),

$$\begin{aligned} p_{1,0} &= \mathbf{P}\left(E_{1,0}|X_1 = 1\right)\mathbf{P}\left(X_1 = 1\right) + \mathbf{P}\left(E_{1,0}|X_1 = -1\right)\mathbf{P}\left(X_1 = -1\right) \\ &= p_{2,0}p + (1)(1-p) = p_{1,0}^2 p + (1-p). \end{aligned}$$

By solving quadratic equation $px^2 - x + 1 - p = 0$ we find that $x = 1$ or $x = \frac{q}{p} < 1$ if $p > 1/2$.

10. Let $S_n, n \geq 0$, be symmetric simple random walk. Let $T = \min\{n : S_n = 0\}$, and write $\mathbf{P}_a$ for probabilities when the walk starts at $S_0 = a$. By basic probabilities for S_n,$n \geq 0$, we mean probabilities of the form $\mathbf{P}_0(S_n = k), \mathbf{P}_0(S_n \geq k)$, or $\mathbf{P}_0(S_n \leq k)$, all of which corresponding to starting at $S_0 = 0$.

(a) For $a \geq 1, i \geq 1, n \geq 1$, express $\mathbf{P}_a(S_n = i, T \leq n)$ and $\mathbf{P}_a(S_n = i, T > n)$ in terms of finitely many basic probabilities. Hint. Reflection principle and space homogeneity.

(b) For $a \geq 1, i \geq 1, n \geq 1$, show that

$$\mathbf{P}_a\left(T > n\right) = \mathbf{P}_a\left(S_1 \ldots S_n \neq 0\right) = \sum_{j=1-a}^{a} \mathbf{P}_0\left(S_n = j\right).$$

(c) You may take as given that $\mathbf{P}_0(S_n = j) \sim 1/\sqrt{\frac{\pi}{2}n}$ for each fixed $j \in \mathbf{Z}$; here $\sim$ means the ratio converges to 1. Use this to find c, α such that $\mathbf{P}_a\left(T > n\right) \sim c/n^\alpha$ as $n \to \infty$, where $a > 0$. Does c or α depend on a? Hint. Consider even n?

9 Week 10: 10/24- 10/28

1. a) Let $X \sim \Gamma(\lambda, n)$, and $Y \sim \Gamma(\lambda, m)$ be independent. Show that $X + Y \sim \Gamma(\lambda, n+m)$.
Hint. The pdf of $V \sim \Gamma(\lambda, k)$ is

$$f_V(x) = \frac{(\lambda x)^{k-1}}{(k-1)!} \cdot \lambda e^{-\lambda x}, x > 0.$$

It is known that

$$\int_0^1 u^j (1-u)^{n-j} \, du = \frac{j!(n-j)!}{(n+1)!}.$$

b) Let $V \sim \Gamma(\lambda, n)$. Find $\mathbf{E}(V)$, $\mathrm{Var}(V)$. *Hint.* V is a sum of independent exponential r.v.

2. Let $V \sim \Gamma(\lambda, n)$. Show that

$$\mathbf{P}(V > t) = \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}, t > 0.$$

Hint. Integrate by parts, induction.
Answer. Let $V \sim \Gamma(\lambda, n)$. Then

$$\begin{aligned}
\mathbf{P}(V > t) &= \int_t^\infty \frac{(\lambda x)^{n-1}}{(n-1)!} \lambda e^{-\lambda x} dx = -\int_t^\infty \frac{(\lambda x)^{n-1}}{(n-1)!} d e^{-\lambda x} \\
&= -\frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} |_t^\infty + \int_t^\infty \frac{(n-1)(\lambda x)^{n-2}}{(n-1)!} \lambda e^{-\lambda x} dx \\
&= \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} + \int_t^\infty \frac{(\lambda x)^{n-2}}{(n-2)!} \lambda e^{-\lambda x} dx.
\end{aligned}$$

Continuing,

$$\begin{aligned}
\mathbf{P}(V > t) &= \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} + \ldots + \frac{\lambda t}{1!} e^{-\lambda t} + \int_t^\infty \lambda e^{-\lambda x} dx \\
&= \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t} + \ldots + \frac{\lambda t}{1!} e^{-\lambda t} + e^{-\lambda t}.
\end{aligned}$$

3. I am selling my house, and have decided to accept the first offer exceeding \$K. Assuming that offers are independent random variables with common distribution function F, what is the expected number of offers received before I sell the house.

Answer. Probability that an offer X_n exceeds K is $p = \mathbf{P}(X_n > K) = 1 - \mathbf{P}(X_n \le K) = 1 - F(K)$. The the number Y of independent offers until the first >K is made is geometric with parameter p, i.e.,

$$\mathbf{E}(Y) = \frac{1}{p} = \frac{1}{1 - F(K)}.$$

4. Let X, Y be independent random variables with common distribution function F and density function f. Show that $V = \max\{X, Y\}$ has distribution function $\mathbf{P}(V \le x) = F(x)^2$ and density function $f_V(x) = 2f(x)F(x)$. Find the density function of $U = \min\{X, Y\}$. Hint. Compute $\mathbf{P}(U > x)$

5. The annual rainfall figures in Bandrika are independent identically distributed continuous random variables $\{X_r, r \geq 1\}$. Find the probability that:

(a) $X_1 < X_2 < X_3 < X_4$. Hint. Use symmetry: all orderings of X_1, X_2, X_3, X_4 are equally likely.

(b) $X_1 > X_2 < X_3 < X_4$. Hint. Rewrite this event as a union of (a) type.

6. Let $\{X_r, r \geq 1\}$ be independent and identically distributed with distribution function F satisfying $F(y) < 1$ for all y, and let $Y(y) = \min\{k : X_k > y\}$. Show that

$$\lim_{y\to\infty} \mathbf{P}(Y(y) \leq \mathbf{E}[Y(y)]) = 1 - e^{-1}.$$

Hint. Find $\mathbf{P}(Y(y) > n)$ first. What is the distribution of $Y(y)$? What is $\mathbf{E}[Y(y)]$?

Answer. First, by independence,

$$\mathbf{P}(Y(y) > n) = \mathbf{P}(X_1 \leq y, \ldots, X_n \leq y) = F(y)^n.$$

We can identify distribution of $Y(y)$ without computing: $Y(y)$ is defined as the moment when the first "success" happens: the sequence X_k are outcomes of independent trials and we label "success" if $\{X_k > y\}$ and "failure" if $\{X_k \leq y\}$. Probability of success is $p = \mathbf{P}(X_k > y) = 1 - F(y)$, and $Y(y)$ is geometric(p). Alternatively, we could compute:

$$\begin{aligned}\mathbf{P}(Y(y) = n) &= \mathbf{P}(Y(y) > n-1) - \mathbf{P}(Y(y) > n) \\ &= F(y)^{n-1} - F(y)^n = F(y)^{n-1}(1 - F(y)),\end{aligned}$$

i.e. $Y(y)$ is geometric with parameter $p = 1 - F(y)$. Hence $\mathbf{E}(Y(y)) = \frac{1}{p} = \frac{1}{1-F(y)}$. So, denoting $[a]$ the integral part of a,

$$\begin{aligned}\mathbf{P}(Y(y) \leq \mathbf{E}[Y(y)]) &= \mathbf{P}\left(Y(y) \leq \left[\frac{1}{1-F(y)}\right]\right) \\ &= 1 - \mathbf{P}\left(Y(y) > \left[\frac{1}{1-F(y)}\right]\right) \\ &= 1 - F(y)^{\left[\frac{1}{1-F(y)}\right]}.\end{aligned}$$

Now,

$$F(y)^{\frac{1}{1-F(y)}} \leq F(y)^{\left[\frac{1}{1-F(y)}\right]} \leq F(y)^{\frac{1}{1-F(y)}-1},$$

and

$$1 \leq \frac{F(y)^{\left[\frac{1}{1-F(y)}\right]}}{F(y)^{\frac{1}{1-F(y)}}} \leq \frac{1}{F(y)} \to 1$$

as $y \to \infty$. Hence

$$\begin{aligned}\lim_{y\to\infty} F(y)^{\left[\frac{1}{1-F(y)}\right]} &= \lim_{y\to\infty} F(y)^{\frac{1}{1-F(y)}} \\ &= \lim_{y\to\infty} \{1 - (1 - F(y))\}^{\frac{1}{1-F(y)}} = e^{-1},\end{aligned}$$

because $\frac{1}{1-F(y)} \to \infty$ as $y \to \infty$, and Calculus 2:

$$\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^x = e^{-1}.$$

7. (a) Let Θ be uniform on $(0,\pi)$, and $a \in \mathbf{R}$. Find the density of $Y = a\cos\Theta$. Hint. Find density of $X = \cos\Theta$ first.

(b) Let U be a continuous r.v. with pdf

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, x \in \mathbf{R}$$

(it is standard Cauchy r.v.). Show that U and $1/U$ have the same distribution.

8. Let f, g be density functions of X and Y.

a) Show that $\alpha f + (1-\alpha) g$ is a density function for $\alpha \in [0,1]$.

Answer. Indeed $f, g \geq 0$ and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} g(x)\,dx = 1.$$

Hence $\alpha f + (1-\alpha) g \geq 0$, and

$$\begin{aligned}
&\int_{-\infty}^{\infty} [\alpha f(x) + (1-\alpha) g(x)]\,dx \\
= \quad &\alpha \int_{-\infty}^{\infty} f(x)\,dx + (1-\alpha)\int_{-\infty}^{\infty} g(x)\,dx = \alpha + (1-\alpha) = 1,
\end{aligned}$$

i.e., $\alpha f + (1-\alpha)$ is a pdf.

b) Let Z be bern(α) with $\alpha \in [0,1]$ independent of X and Y. What is the distribution function of $U = ZX + (1-Z)Y$? What is the density function of U?

Answer. Since $\Omega = \{Z=1\} \cup \{Z=0\}$, and Z is independent, for any $u \in \mathbf{R}$

$$\begin{aligned}
F_U(u) &= \mathbf{P}(U \leq u) \\
&= \mathbf{P}(ZX + (1-Z)Y \leq u, Z=1) + \mathbf{P}(ZX + (1-Z)Y \leq u, Z=0) \\
&= \mathbf{P}(X \leq u, Z=1) + \mathbf{P}(Y \leq u, Z=0) = \alpha\mathbf{P}(X \leq u) + (1-\alpha)\mathbf{P}(Y \leq u) \\
&= \alpha\int_{-\infty}^{u} f(x)\,dx + (1-\alpha)\int_{-\infty}^{u} g(x)\,dx = \int_{-\infty}^{u} [\alpha f(x) + (1-\alpha) g(x)]\,dx.
\end{aligned}$$

So, $\alpha f + (1-\alpha) g$ is the pdf of U.

9. *Survival.* Let X be a positive random variable with density function f and distribution function F with values in $[0,1)$. Define the hazard function

$$H(t) = -\log[1 - F(t)], t \geq 0,$$

survival function

$$S(t) = \mathbf{P}(X > t) = 1 - F(t), t \geq 0,$$

and the hazard rate

$$r(t) = \lim_{h\downarrow 0}\frac{\mathbf{P}(t < X \leq t+h | X > t)}{h}.$$

Comment. Possible interpretation: X is death moment; $S(t) = \mathbf{P}(X > t)$ is probability to be alive at time t; $\mathbf{P}(t < X \le t + h | X > t)$ is probability to die (quickly if h is small) in time interval $(t, t+h]$ given being alive at t.

Show that:

(a)

$$r(t) = H'(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}.$$

Answer. By definition, we have

$$S(t) = \mathbf{P}(X > t) = 1 - F(t) = S(t) = e^{-H(t)}, t \ge 0.$$

Hence

$$\begin{aligned} & \frac{\mathbf{P}(t < X \le t + h | X > t)}{h} \\ = \ & \frac{1}{h} \frac{\mathbf{P}(t < X \le t + h)}{\mathbf{P}(X > t)} = \frac{1}{h} \frac{S(t) - S(t+h)}{S(t)} = \frac{1}{h} \frac{\exp\{-H(t)\} - \exp\{-H(t+h)\}}{\exp\{-H(t)\}} \\ = \ & \exp\{H(t)\} \frac{\exp\{-H(t)\} - \exp\{-H(t+h)\}}{h}. \end{aligned}$$

We recognize

$$\begin{aligned} & \lim_{h \to 0} \frac{\exp\{-H(t)\} - \exp\{-H(t+h)\}}{h} \\ = \ & -\lim_{h \to 0} \frac{\exp\{-H(t+h)\} - \exp\{-H(t)\}}{h} = -\frac{d}{dt}\left(e^{-H(t)}\right) = H'(t) e^{-H(t)}. \end{aligned}$$

So,

$$\frac{\mathbf{P}(t < X \le t + h | X > t)}{h} \to H'(t) = r(t)$$

as $h \to 0$. Finally,

$$H'(t) = -\frac{d}{dt}(\log[1 - F(t)]) = \frac{F'(t)}{1 - F(t)} = \frac{f(t)}{S(t)}, t \ge 0.$$

(b) If $r(t)$ increases with t, then $H(t)/t$ increases with t. Hint. $H(t)/t$ increases if $\frac{d}{dt}\left(\frac{H(t)}{t}\right) \ge 0$.

Answer. By fundamental calculus theorem,

$$H(t) = \int_0^t H'(s)\,ds = \int_0^t r(s)\,ds;$$

Thus

$$\begin{aligned} \frac{d}{dt}\left(\frac{H(t)}{t}\right) &= \frac{d}{dt}\left(\frac{\int_0^t r(s)\,ds}{t}\right) = \frac{t r(t) - \int_0^t r(s)\,ds}{t^2} \\ &= \frac{\int_0^t [r(t) - r(s)]\,ds}{t^2} \ge 0, t > 0, \end{aligned}$$

if $r(t)$ increases with t, and $H(0) = 0$. Therefore $H(t)/t$ increases with t.

10. *Order statistics*. Let $X_1, \ldots, X_n$ be independent identically distributed variables with a common pdf f. Such a collection is called a random sample. For each $\omega \in \Omega$, arrange the sample values $X_1(\omega), \ldots, X_n(\omega)$ in non-decreasing order $X_{(1)}(\omega), \ldots, X_{(n)}(\omega)$, where (1), (2),..., (n) is a (random) permutation of 1, 2,..., n. The new variables $X_{(1)}, \ldots, X_{(n)}$ are called the order statistics.

a) Show, by a symmetry argument, that the joint distribution function of the order statistics satisfies

$$\begin{aligned}
& \mathbf{P}\left(X_{(1)} \le y_1, \ldots, X_{(n)} \le y_n\right) \\
= \;& n!\mathbf{P}\left(X_1 \le y_1, \ldots, X_n \le y_n, X_1 < \ldots < X_n\right) \\
= \;& \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_n} \chi(x_1, \ldots, x_n) f(x_1) \ldots f(x_n)\, dx_1 \ldots dx_n,
\end{aligned}$$

where

$$\chi(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } x_1 < \ldots < x_n \\ 0 & \text{otherwise.} \end{cases}$$

Note $X_{(1)} = \min_{1 \le k \le n} X_k$, $X_{(n)} = \max_{1 \le k \le n} X_k$.

Hint. We have

$$\left\{X_{(1)} \le y_1, \ldots, X_{(n)} \le y_n\right\} = \cup_{j_1,\ldots,j_n} \left\{X_{j_1} \le y_1, \ldots, X_{j_n} \le y_n, X_{j_1} < \ldots < X_{j_n}\right\},$$

where the union is taken over all possible different orderings (permutations) $j, \ldots, j_n$ of $\{1, \ldots, n\}$. All the sets in the union are disjoint, and there are $n!$ of them. With probaility 1,

$$\Omega = \cup_{j_1,\ldots,j_n} \left\{X_{j_1} < \ldots < X_{j_n}\right\}.$$

b) Find the marginal density function of the kth order statistic $X_{(k)}$ of a sample with size n:

(i) by integrating the result in a);

(ii) directly. Hint. First, find the df of $X_{(k)}$

$$F_{X_{(k)}}(x) = \mathbf{P}\left(X_{(k)} \le x\right) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j \left[1 - F(x)\right]^{n-j}$$

Note $\sum_{j=1}^{n} I_{\{X_j \le x\}}$ is a binomial r.v.

11. Find the joint density function of the order statistics of n independent uniform variables in $(0, T)$.

12. Let $X_1, \ldots, X_n$ be independent and uniformly distributed on $(0, 1)$, with order statistics $X_{(1)}, \ldots, X_{(n)}$.

Show that, for fixed k, the density function of $nX_{(k)}$ converges as $n \to \infty$, and find and identify the limit function.

Answer. By hw the pdf of $X_{(k)}$ is

$$f(x) = k\binom{n}{k} x^{k-1} (1 - x)^{n-k}, 0 < x < 1.$$

Hence the pdf of $nX_{(k)}$ is

$$\begin{aligned}
\frac{1}{n} f\left(\frac{x}{n}\right) &= \frac{k}{n}\binom{n}{k}\left(\frac{x}{n}\right)^{k-1}\left(1-\frac{x}{n}\right)^{n-k} \\
&= \frac{x^{k-1}}{(k-1)!}\left(1-\frac{x}{n}\right)^{n} \frac{(n-1)\ldots(n-(k-1)+1)}{n^{k-1}}\ldots \\
&\to \frac{x^{k-1}}{(k-1)!}e^{-x} \text{ as } n \to \infty
\end{aligned}$$

which is the pdf of gamma$(1, k)$-r.v.

13. a) Let X take nonnegative integer values. Show that

$$\mathbf{E}(X) = \sum_{n=0}^{\infty} \mathbf{P}(X > n).$$

Answer. Since X takes nonnegative integer values, we have

$$X = \sum_{n=0}^{\infty} I_{\{n<X\}},$$

and

$$\mathbf{E}(X) = \sum_{n=0}^{\infty} \mathbf{E}\left(I_{\{n<X\}}\right) = \sum_{n=0}^{\infty} \mathbf{P}(n < X).$$

b) Let M be the minimum value seen in 4 die rolls. Find $\mathbf{E}(M)$. You do not need to simplify to one number, just get an expression with numbers.

Answer. Let X_i be the score in the ith roll. Then for $n = 0, \ldots, 5$, we have

$$\mathbf{P}(X_i > n) = \frac{6-n}{6} = 1 - \frac{n}{6}, \mathbf{P}(M > n) = \mathbf{P}(\text{every } X_i > n) = \left(1 - \frac{n}{6}\right)^4$$

By part a),

$$\mathbf{E}(M) = \sum_{n=0}^{5}\left(1 - \frac{n}{6}\right)^4.$$

14. A population of N animals has had a number n of its members captured, marked, and released. Let X be the number of animals it is necessary to recapture (without re-release) in order to obtain m marked animals. Show that

$$\mathbf{P}(X = k) = \frac{n}{N}\frac{\binom{n-1}{m-1}\binom{N-n}{k-m}}{\binom{N-1}{k-1}}$$

and find $\mathbf{E}(X)$. This distribution has been called *negative hypergeometric*. Hint. Look at #10 of week 6 and the corresponding homework.

Answer. Think of Ω as set of all sequences of captured animals of length k: $\#\Omega = N(N-1)\ldots(N-k+1) = \binom{N}{k}k!$. Let $k \geq m$, $B =$"kth captured animal is mth marked".

Then kth animal marked can be n different ways, the sequence of $k-1$ animal before the kth must contain $m-1$ marked and the remaining $k-m$ unmarked. That can be done in $\binom{n-1}{m-1}\binom{N-n}{k-m}(k-1)!$ different ways. By multiplication principle,

$$
\begin{aligned}
\#B &= n\binom{n-1}{m-1}\binom{N-n}{k-m}(k-1)! \\
\mathbf{P}(X=k) &= \frac{n\binom{n-1}{m-1}\binom{N-n}{k-m}(k-1)!}{\binom{N}{k}k!} \\
&= \frac{n\binom{n-1}{m-1}\binom{N-n}{k-m}}{\binom{N}{k}k} = \frac{n}{N}\frac{\binom{n-1}{m-1}\binom{N-n}{k-m}}{\binom{N-1}{k-1}}.
\end{aligned}
$$

The r.v. X is Y_m from #1 of hw6: $\mathbf{E}(X) = \mathbf{E}(Y_m) = m \cdot \frac{N+1}{n+1}$.

10 Week 11: 10/31- 11/4

1. Let $Y = X + U$, where X, U are independent continuous r.v. Show that (X, Y) has joint pdf $f(x, y) = f_X(x) f_U(y - x)$. Find the conditional pdf of Y given $X = x$ (the form of the conditional pdf shows that given $X = x$, $Y = x + U$ in distribution.

3. Let X be a standard normal variable and for $a > 0$, define the random variable Y by

$$Y_a = \begin{cases} X & \text{if } |X| < a, \\ -X & \text{if } |X| \geq a. \end{cases}$$

(a) Verify that Y_a is a standard normal r.v.

Answer. Let $F(x) = \mathbf{P}(X \leq x), x \in \mathbf{R}$, be the distribution function of X. For $y \leq -a$,

$$\mathbf{P}(Y_a \leq y) = \mathbf{P}(-X \leq y) = \mathbf{P}(X \leq y) = F(y);$$

For $y \in [-a.a]$,

$$\begin{aligned} \mathbf{P}(Y_a \leq y) &= \mathbf{P}(Y_a \leq -a) + \mathbf{P}(-a < Y_a \leq y) = F(-a) + \mathbf{P}(-a < X \leq y) \\ &= F(-a) + F(y) - F(-a) = F(y), \end{aligned}$$

and for $y > a$,

$$\begin{aligned} \mathbf{P}(Y_a \leq y) &= \mathbf{P}(Y_a \leq a) + \mathbf{P}(a < Y_a \leq y) = F(a) + \mathbf{P}(a < -X \leq y) \\ &= F(a) + \mathbf{P}(a < X \leq y) = F(a) + F(y) - F(a) = F(y). \end{aligned}$$

So, we found that $F(y) = \mathbf{P}(Y_a \leq y), y \in \mathbf{R}$: Y_a is standard normal.

(b) Write $\rho(a) = \mathbf{E}(XY_a)$, the correlation coefficient, as a difference of two integrals using pdf $f(x)$ of X.

Answer. Since

$$Y_a = XI_{\{|X|<a\}} - XI_{\{|X|\geq a\}},$$

we have

$$\begin{aligned} \rho(a) &= \operatorname{cov}(X, Y_a) = \mathbf{E}(XY_a) = \mathbf{E}\left[X^2 I_{\{|X|<a\}} - X^2 I_{\{|X|\geq a\}}\right] \\ &= \frac{2}{\sqrt{2\pi}} \int_0^a x^2 e^{-x^2/2} dx - \frac{2}{\sqrt{2\pi}} \int_a^\infty x^2 e^{-x^2/2} dx, a \geq 0. \end{aligned}$$

(c) Is there a value of a for which $\rho(a) = 0$? Hint. Intermediate value theorem (Calculus I), look at the expression in (b).

Answer. The function $\rho(a), a \geq 0$, is continuous (as a function of a variable integration limit), $\rho(0) = -1$ and $\lim_{a \to \infty} \rho(a) = 1$. Therefore, by intermediate value theorem (Calculus 1) there is $a_0 > 0$ so that $\rho(a_0) = 0$.

Comment. Note that the standard normal variables X and Y_{a_0} are uncorrelated. Nevertheless

1. they are not independent: $\mathbf{P}(|X| \leq a_0, |Y_{a_0}| \leq a_0) = \mathbf{P}(|X| \leq a_0) \neq \mathbf{P}(|X| \leq a_0) \mathbf{P}(|Y_{a_0}| \leq a_0)$;
2. $X + Y_{a_0} = 2XI_{\{|X|\leq a_0\}}$ is not normal r.v.

(d) Is the pair (X, Y_a) a bivariate normal? Explain.

Answer. No: for instance $\mathbf{P}(X = Y_a) = \mathbf{P}(|X| \leq a) > 0$ if $a > 0$: (X, Y_a) cannot have a joint pdf. Look at the comment to (c) as well.

4. Find the density function of $Z = X + Y$ when X, Y have joint density function

$$f(x, y) = \frac{1}{2}(x + y) e^{-(x+y)}, x > 0, y > 0.$$

Answer. We have

$$f(x, y) = \frac{1}{2}(x + y) e^{-(x+y)} I_{(0,\infty)}(x) I_{(0,\infty)}(y), x, y \in \mathbf{R}.$$

For $z > 0$,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f(x, z - x) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2}(x + z - x) e^{-(x+z-x)} I_{(0,\infty)}(x) I_{(0,\infty)}(z - x) dx \\ &= \int_0^z \frac{1}{2} z e^{-z} dx = \frac{1}{2} z^2 e^{-z}, z > 0. \end{aligned}$$

5. Suppose X and Y are independent continuous random variables with uniform distribution on $(0, 1)$.

(a) For $a > 0$, find pdf of $V = aY$. How is V distributed? Find the density function of $X + 2Y$.

Answer. Range of $V = (0, a)$. For $v \in (0, a)$,

$$\begin{aligned} \mathbf{P}(V \leq v) &= \mathbf{P}(aY \leq v) = \mathbf{P}\left(Y \leq \frac{v}{a}\right) = \frac{v}{a}, \\ f_V(v) &= \frac{1}{a}, 0 < v < a, \end{aligned}$$

that is V is uniform in $(0, 1)$. Hence $U = 2Y$ is uniform in $(0, 2)$. Range of $X + 2Y = X + U$ is $(0, 3)$, and for $z \in (0, 3)$,

$$\begin{aligned} f_{X+U}(z) &= \int_{-\infty}^{\infty} f_X(x) f_U(z - x) dx = \int_{\max\{0,z-2\}}^{\min\{1,z\}} 1 \cdot \frac{1}{2} dx \\ &= \frac{1}{2}(\min\{1, z\} - \max\{0, z - 2\}), \end{aligned}$$

because $f_X(x) f_U(z - x) \neq 0$ iff $0 < x < 1, 0 < z - x < 2$, equivalently, $\max\{0, z - 2\} < x < \min\{1, z\}$. Hence

$$f_{X+U}(z) = \begin{cases} \frac{z}{2}, & 0 < z < 1 \\ \frac{1}{2} & 1 < z < 2 \\ \frac{1-(z-2)}{2} = \frac{3-z}{2} & 2 < z < 3 \end{cases}$$

(b) Find the joint density function for $X - Y, X + Y$.

Answer. Consider $T = (X - Y, X + Y) = H(X, Y)$, where $t = (t_1, t_2) = H(x, y)$ is linear map

$$t_1 = x - y, t_2 = x + y.$$

Solving for given (t_1, t_2), we find the inverse $(x, y) = G(t_1, t_2) = H^{-1}(t_1, t_2)$:

$$x = \frac{1}{2}t_1 + \frac{1}{2}t_2, y = -\frac{1}{2}t_1 + \frac{1}{2}t_2, \text{ and Jacobian } J(t) = \frac{1}{2}.$$

Range of $T = S = \left\{(t_1, t_2) : 0 < \frac{1}{2}t_1 + \frac{1}{2}t_2 < 1, 0 < -\frac{1}{2}t_1 + \frac{1}{2}t_2 < 1\right\}$, equivalently,

$$S = \{(t_1, t_2) : -t_1 < t_2 < 2 - t_1, t_1 < t_2 < 2 + t_1\}.$$

It is a rectangle in (t_1, t_2) -plane. Hence

$$f_V(t_1, t_2) = \frac{1}{2} I_S(t_1, t_2), (t_1, t_2) \in \mathbf{R}^2;$$

V is uniform in S.

6. Let X be standard normal. Find $\mathbf{E}(X|X > 0)$. Hint: the conditional pdf of X given the event $\{X > 0\}$ is

$$h(x) = \frac{d}{dx}[\mathbf{P}(X \leq x|X > 0)], x > 0.$$

Answer. We have

$$\begin{aligned} \mathbf{P}(X \leq x|X > 0) &= \frac{\mathbf{P}(X \leq x, X > 0)}{\mathbf{P}(X > 0)} = 2\mathbf{P}(0 < X \leq x) \\ &= 2\left[\Phi(x) - \frac{1}{2}\right] \text{ if } x > 0. \end{aligned}$$

The conditional pdf is

$$\varphi(x) = \frac{d}{dx}\mathbf{P}(X \leq x|X > 0) = 2\phi(x), x > 0,$$

where ϕ is the pdf of standard normal. Thus

$$\mathbf{E}(X|X > 0) = 2\frac{1}{\sqrt{2\pi}} \int_0^\infty x e^{-x^2/2} dx = -\sqrt{\frac{2}{\pi}} e^{-x^2/2}|_0^\infty = \sqrt{\frac{2}{\pi}}.$$

7. A coin-making machine produces quarters in such way that, for each coin, the probability U to turn up heads is uniform in $(0, 1)$. A coin pops out (randomly) and is flipped 10 times. Let X be number of heads in those 10 tosses.

a) Find $\mathbf{E}(X)$ and var(X). Hint. $X \sim$bin$(10, U)$.

Answer. Given $U = u$, X is binomial(n, u) :

$$\begin{aligned} \mathbf{E}(X|U = u) &= 10u, \mathrm{Var}(X|U = u) = 10u(1 - u), \\ \mathbf{E}(X|U) &= 10U, \mathrm{Var}(X|U) = 10U(1 - U). \end{aligned}$$

Then

$$\begin{aligned} \mathbf{E}(X) &= 10\mathbf{E}(U) = 10 \cdot \frac{1}{2} = 5, \\ \mathrm{Var}(X) &= \mathbf{E}[\mathrm{Var}(X|U)] + \mathrm{Var}(\mathbf{E}(X|U)) \\ &= 10\mathbf{E}[U(1 - U)] + 10^2 \mathrm{Var}(U) \\ &= 10\int_0^1 u(1 - u) du + 10^2 \frac{1}{12} = 10 \cdot \frac{1}{6} + \frac{100}{12} = \frac{120}{12} = 10. \end{aligned}$$

b) What are

$$\begin{aligned}\mathbf{P}(X=j|U=u), j &= 0,\ldots,10 \text{ and}\\ \mathbf{P}(X=j|U), j &= 0,\ldots,10?\end{aligned}$$

Answer. Given $U=u$, X is binomial(n,u) :

$$\begin{aligned}\mathbf{P}(X=j|U=u) &= \binom{10}{j}u^j(1-u)^{10-j}, j=0,\ldots,10 \text{ and}\\ \mathbf{P}(X=j|U) &= \binom{10}{j}U^j(1-U)^{10-j}, j=0,\ldots,10.\end{aligned}$$

c) Find $\mathbf{P}(X=j), j=0,\ldots,10$, the distribution of X. Hint: $\mathbf{P}(X=j)=\mathbf{E}[\mathbf{P}(X=j|U)]$, $j=0,\ldots,10$; It is known that

$$\int_0^1 u^j(1-u)^{n-j}\,du=\frac{j!(n-j)!}{(n+1)!}.$$

Answer. Using condition,

$$\begin{aligned}\mathbf{P}(X=j) &= \mathbf{E}[\mathbf{P}(X=j|U)]=\binom{10}{j}\int_0^1 u^j(1-u)^{10-j}\,du\\ &= \frac{10!}{j!(10-j)!}\frac{j!(10-j)!}{11!}=\frac{1}{11}, j=0,\ldots,10.\end{aligned}$$

That is X is uniform in $\{0,1,\ldots,10\}$.

d) For $k=0,1,\ldots,10, 0\le v\le 1$, find

$$\mathbf{P}(U\le v|X=k),$$

and conditional pdf of U given $X=k$. What is $\mathbf{E}(U|X=k)$ and $\mathbf{E}(U|X)$?

Hint. Note

$$\mathbf{P}(X=k,U\le v)=\mathbf{E}\left[\mathbf{P}(X=k|U)\,I_{\{U\le v\}}\right]=\int_0^v \mathbf{P}(X=k|U=u)\,du,$$

and conditional pdf of U given $X=k$ is

$$f(v|k)=\frac{d}{dv}\mathbf{P}(U\le v|X=k).$$

Answer. Since

$$\mathbf{P}(U\le v|X=k)=\frac{\mathbf{P}(X=k,U\le v)}{\mathbf{P}(X=k)}=\frac{\mathbf{E}\left[\mathbf{P}(X=k|U)\,I_{\{U\le v\}}\right]}{\frac{1}{11}}=11\int_0^v \mathbf{P}(X=k|U=u)\,du,$$

the conditional pdf of U given $X=k$ is

$$f(v|k)=\frac{d}{dv}\mathbf{P}(U\le v|X=k)=11\mathbf{P}(X=k|U=v), 0<v<1.$$

Hence

$$\begin{aligned}
\mathbf{E}(U|X=k) &= \int_0^1 vf(v|k)\,dv = 11\int_0^1 v\mathbf{P}(X=k|U=v)\,dv \\
&= 11\int_0^1 \binom{10}{k} v^{k+1}(1-v)^{10-k}\,dv = 11\cdot\frac{10!}{k!\,(10-k)!}\frac{(k+1)!\,(10-k)!}{12!} \\
&= \frac{k+1}{12}, \mathbf{E}(U|X) = \frac{X+1}{12}.
\end{aligned}$$

8. Let a random variable X be normal $N\left(\mu,\sigma^2\right)$, and let the conditional distribution of Y given X be normal $N\left(a+bX,\sigma_1^2\right)$. Recall $f_{Y|X}(y|x) = f(x,y)/f_X(x)$, where $f(x,y)$ is the joint pdf of X,Y.

a) find the joint pdf of X,Y. Is (X,Y) normal bivariate?

Answer. We have

$$\begin{aligned}
f_{Y|X}(y|x) &= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{1}{2}\frac{(y-a-bx)^2}{\sigma_1^2}\right\} = \frac{f(x,y)}{f_X(x)}, \\
f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\},
\end{aligned}$$

and by definition of the conditional pdf

$$\begin{aligned}
f(x,y) &= f_{Y|X}(y|x)\,f_X(x) \\
&= \frac{1}{2\pi\sigma_1\sigma}\exp\left\{-\frac{1}{2}\left[\left(\frac{y-a-bx}{\sigma_1}\right)^2 + \left(\frac{x-\mu}{\sigma}\right)^2\right]\right\}.
\end{aligned}$$

Note that,

$$\begin{aligned}
&\left(\frac{y-a-bx}{\sigma_1}\right)^2 + \left(\frac{x-\mu}{\sigma}\right)^2 \\
&= \left(\frac{y-a-b\mu-b(x-\mu)}{\sigma_1}\right)^2 + \left(\frac{x-\mu}{\sigma}\right)^2 \\
&= \left(\frac{y-a-b\mu}{\sigma_1}\right)^2 - 2b\sigma\frac{(y-a-b\mu)}{\sigma_1^2}\left(\frac{x-\mu}{\sigma}\right) + \left(\frac{x-\mu}{\sigma}\right)^2 + b^2\left(\frac{x-\mu}{\sigma_1}\right)^2 \\
&= \left(\frac{y-a-b\mu}{\sqrt{b^2\sigma^2+\sigma_1^2}}\right)^2\frac{b^2\sigma^2+\sigma_1^2}{\sigma_1^2} - 2b\sigma\left(\frac{y-a-b\mu}{\sqrt{b^2\sigma^2+\sigma_1^2}}\right)\frac{\sqrt{b^2\sigma^2+\sigma_1^2}}{\sigma_1^2}\left(\frac{x-\mu}{\sigma}\right) \\
&\quad + \left(\frac{x-\mu}{\sigma}\right)^2 + \frac{b^2\sigma^2}{\sigma_1^2}\left(\frac{x-\mu}{\sigma}\right)^2 \\
&= \frac{b^2\sigma^2+\sigma_1^2}{\sigma_1^2}\times \\
&\quad\times\left\{\left(\frac{y-a-b\mu}{\sqrt{b^2\sigma^2+\sigma_1^2}}\right)^2 - 2\frac{b\sigma}{\sqrt{b^2\sigma^2+\sigma_1^2}}\left(\frac{y-a-b\mu}{\sqrt{b^2\sigma^2+\sigma_1^2}}\right)\left(\frac{x-\mu}{\sigma}\right) + \left(\frac{x-\mu}{\sigma}\right)^2\right\}
\end{aligned}$$

and, denoting $\rho = \frac{b\sigma}{\sqrt{b^2\sigma^2+\sigma_1^2}}$,

$$\sigma_1\sigma = \sigma\sqrt{b^2\sigma^2+\sigma_1^2}\sqrt{\frac{\sigma_1^2}{b^2\sigma^2+\sigma_1^2}} = \sigma\sqrt{b^2\sigma^2+\sigma_1^2}\sqrt{1-\rho^2}.$$

So, (X, Y) is normal bivariate with parameters

$$\begin{aligned} \mathbf{E}(X) &= \mu, \mathbf{E}(Y) = a + b\mu, \\ \mathrm{var}(X) &= \sigma^2, \mathrm{var}(Y) = b^2\sigma^2 + \sigma_1^2, \\ \rho &= \frac{b\sigma}{\sqrt{b^2\sigma^2+\sigma_1^2}}. \end{aligned}$$

Comment. In hw11, we will go an alternative and easier way (using joint mgf function of (X, Y)) to determine that (X, Y) is normal bivariate.

b) Find the marginal distribution of Y and the correlation coefficient of X and Y.

Answer. We found in a) that (X, Y) is normal bivariate and all parameters. In particular,

$$\rho(X, Y) = \rho = \frac{b\sigma}{\sqrt{b^2\sigma^2+\sigma_1^2}},$$

and $Y \sim N\left(a + b\mu, b^2\sigma^2 + \sigma_1^2\right)$.

9. Let $Y = c_1X_1 + \ldots + c_nX_n$. Show that

$$\mathrm{var}(Y) = \sum_{i=1}^{n} c_i^2 \mathrm{var}(X_i) + 2\sum_{i<j} c_ic_j \mathrm{cov}\left(X_i, X_j\right) = cBc',$$

where $c = (c_1, \ldots, c_n)$ is the row vector of coefficients, c' is the transpose of c, i.e. a column vector of the coefficients, and $B = \left(b_{ij}\right)$ is the $n \times n$ symmetric matrix with b_{ij} =cov$\left(X_i, X_j\right)$.

10. Let $X = (X_1, X_2)$ be normal bivariate with parameters $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho$. Let B be 2×2-matrix with b_{ij} =cov$\left(X_i, X_j\right)$. Show that the joint pdf

$$f(x) = \frac{1}{2\pi\sqrt{\det B}} \exp\left\{-\frac{1}{2}(x-\mu) B^{-1} (x-\mu)'\right\}, x = (x_1, x_2) \in \mathbf{R}^2, \mu = (\mu_1, \mu_2),$$

and $(x-\mu)'$ is the transpose of $(x-\mu)$, i.e. a column vector.

Answer. The joint pdf of (X_1, X_2) was given as

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2\left(1-\rho^2\right)} Q(x)\right\}, x = (x_1, x_2),$$

with

$$Q(x) = \left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right).$$

The covariance matrix of it is

$$B = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

its determinant

$$\delta = \det(B) = \sigma_1^2\sigma_2^2(1-\rho)^2.$$

Since the inverse

$$B^{-1} = \begin{pmatrix} \frac{\sigma_2^2}{\delta} & \frac{-\rho\sigma_1\sigma_2}{\delta} \\ \frac{-\rho\sigma_1\sigma_2}{\delta} & \frac{\sigma_1^2}{\delta} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_1^2(1-\rho)^2} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_1^2(1-\rho)^2} \end{pmatrix},$$

it follows that

$$\frac{Q(x)}{1-\rho^2} = (x-\mu)B^{-1}(x-\mu)'.$$

That is $X = (X_1, X_2) \sim N(\mu, B)$, where $\mu = (\mu_1, \mu_2)$

11. Let X, Y be independent standard normal. Let $W = X - Y, Z = X + Y$.

a) Are W and Z independent?

Answer. Since (W, Z) is normal bivariate (transformation matrix is invertible), we find that

$$\mathrm{Cov}(W,Z) = \mathrm{Cov}(X-Y, X+Y) = \mathrm{Var}(X) - \mathrm{Var}(Y) = 0.$$

So, W and Z are independent.

b) Find $\mathbf{E}(X+2Y|Z)$.

Answer. We have

$$\begin{aligned} \mathbf{E}(X+2Y|Z) &= \mathbf{E}(X+Y|Z) + \mathbf{E}(Y|Z) = \mathbf{E}(Z|Z) + \mathbf{E}(Y|Z) \\ &= Z + \mathbf{E}(Y|Z). \end{aligned}$$

Now, $(Y,Z) = (Y, X+Y)$ is normal bivariate we find its parameters:

$$\begin{aligned} \mu_1 &= \mathbf{E}(Y) = 0 = \mu_2 = \mathbf{E}(Z), \\ \sigma_1^2 &= \mathrm{Var}(Y) = 1, \sigma_2^2 = \mathrm{Var}(Z) = 1 + 1 = 2, \\ \rho &= \rho(Y,Z) = \frac{\mathrm{Cov}(Y, X+Y)}{\sqrt{2}} = \frac{1}{\sqrt{2}}. \end{aligned}$$

Hence $Y = \rho\frac{\sigma_1}{\sigma_2}Z + U = \frac{1}{2}Z + U$, where $U \sim N\left(0, \sigma_1^2\left(1-\rho^2\right)\right)$ is independent of Z. So, $\mathbf{E}(Y|Z) = \frac{1}{2}Z$, and

$$\mathbf{E}(X+2Y|Z) = Z + \frac{1}{2}Z = \frac{3}{2}Z.$$

12. Let X, Y be continuous positive with joint pdf $f(x,y), x, y > 0$. Let $U = X$ and $V = X/Y$.

a) Find joint pdf of U, V.

Answer. Consider the mapping T of $\{(x,y) : x > 0, y > 0\}$ to $\{(u,v) : u > 0, v > 0\}$ given by $u = x, v = x/y$. Its inverse T^{-1}, by solving for x, y,

$$x = u, y = \frac{u}{v}.$$

Its Jacobian is

$$J\left(u,v\right)=\left|\begin{array}{cc}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{array}\right|=\left|\begin{array}{cc}1 & 1/v\\ 0 & -\frac{u}{v^{2}}\end{array}\right|=-\frac{u}{v^{2}}.$$

Hence the pdf of (U,V) is

$$f_{(U,V)}\left(u,v\right)=\left|J\left(u,v\right)\right|f\left(x\left(u,v\right),y\left(u,v\right)\right)=\frac{u}{v^{2}}f\left(u,\frac{u}{v}\right),u,v>0.$$

b) Find pdf of $V=X/Y$.

Answer. The pdf of V is

$$f_{V}\left(v\right)=\int_{-\infty}^{\infty}f_{(U,V)}\left(u,v\right)du=\int_{0}^{\infty}\frac{u}{v^{2}}f\left(u,\frac{u}{v}\right)du,v>0.$$

13. Let X,Y be independent continuous r.v. with pdf

$$f\left(x\right)=x/2\text{ for }x\in\left(0,2\right).$$

Find pdf or df (choose either one) of

a) $Z=X+Y$; b) $Z=\min\left\{X,Y\right\}$; c) $Z=X/Y$.

Answer. a) We have $f\left(x\right)=\frac{x}{2}I_{(0,2)}\left(x\right),x\in\mathbf{R}$. The set of possible values of Z is $(0,4)$. The pdf of Z is

$$\begin{aligned}f_{Z}\left(z\right) &= \int_{-\infty}^{\infty}f\left(x\right)f\left(z-x\right)dx=\int_{0}^{2}\frac{x}{2}\frac{\left(z-x\right)}{2}I_{(0,2)}\left(z-x\right)dx\\ &= \int_{\max\{0,z-2\}}^{\min\{2,z\}}\frac{x}{2}\frac{\left(z-x\right)}{2}dx=\frac{1}{4}\left(z\frac{x^{2}}{2}-\frac{x^{3}}{3}\right)|_{\max\{0,z-2\}}^{\min\{2,z\}}\\ &= \frac{1}{4}\left(z\frac{x^{2}}{2}-\frac{x^{3}}{3}\right)|_{x=\min\{2,z\}}-\frac{1}{4}\left(z\frac{x^{2}}{2}-\frac{x^{3}}{3}\right)|_{x=\min\{0,z-2\}}.\end{aligned}$$

Computing

$$\begin{aligned}f_{Z}\left(z\right) &= \frac{1}{4}\left(z\frac{z^{2}}{2}-\frac{z^{3}}{3}\right)=\frac{1}{24}z^{3}\text{ if }z\in\left(0,2\right),\\ f_{Z}\left(z\right) &= \frac{1}{4}\left(2z-\frac{8}{3}\right)-\frac{1}{4}\left(z\frac{\left(z-2\right)^{2}}{2}-\frac{\left(z-2\right)^{3}}{3}\right)\\ &= \frac{1}{4}\left(2z-\frac{8}{3}\right)-\frac{1}{4}\left(z-2\right)^{2}\left(\frac{z}{6}+\frac{2}{3}\right)\text{ if }z\in\left(2,4\right).\end{aligned}$$

b) The range of Z is $(0,2)$. For $z\in\left(0,2\right)$,

$$\begin{aligned}1-F_{Z}\left(z\right) &= \mathbf{P}\left(Z>z\right)=\mathbf{P}\left(\min\left\{X,Y\right\}>z\right)=\mathbf{P}\left(Y>z,X>z\right)\\ &= \mathbf{P}\left(X>z\right)\mathbf{P}\left(Y>z\right)=\left(\int_{z}^{2}\frac{x}{2}dx\right)^{2}=\left(\frac{x^{2}}{4}|_{z}^{2}\right)^{2}=\left(1-\frac{z^{2}}{4}\right)^{2},\end{aligned}$$

and

$$F_{Z}\left(z\right)=1-\left(1-\frac{z^{2}}{4}\right)^{2},z\in\left(0,2\right),F_{Z}\left(z\right)=1,z\geq 2,F_{Z}\left(z\right)=0,z\leq 0.$$

c) The joint pdf of (X, Y) is $f(x, y) = \frac{x}{2}\frac{y}{2}, x, y \in (0, 2)$. According to the previous exercise, the pdf

$$\begin{aligned}
f_Z(z) &= \int_0^\infty \frac{x}{z^2} f\left(x, \frac{x}{z}\right) dx \\
&= \int_0^2 \frac{x}{z^2}\frac{x}{2}\frac{x}{2z} I_{(0,2)}\left(\frac{x}{z}\right) dx = \int_0^{\min\{2,2z\}} \frac{x}{z^2}\frac{x}{2}\frac{x}{2z} dx \\
&= \frac{1}{4z^3}\int_0^{\min\{2,2z\}} x^3 dx = \frac{1}{4z^3}\frac{x^4}{4}|_0^{\min(2,2z)}.
\end{aligned}$$

Thus

$$\begin{aligned}
f_Z(z) &= \frac{1}{16} 16z = z \text{ if } z \in (0, 1), \\
f_Z(z) &= \frac{1}{4z^3}\frac{16}{4} = \frac{1}{z^3} \text{ if } z \geq 1.
\end{aligned}$$

14. a) Let $Z = (Z_1, \ldots, Z_n)$ be the standard normal in $\mathbf{R}^d$. Find the pdf of $R = \sqrt{Z_1^2 + \ldots + Z_n^2}$. $R^2 = Z_1^2 + \ldots + Z_n^2$.

b) Let X be uniform in the ball of radius a, that is its pdf $f(x) = \frac{1}{\omega_n a^n}, |x| \leq a$. Find f_R and f_{R^2}.

Answer. General answer was given in the class note of 11/5:

If $f(x)$ is the joint pdf of $X = (X_1, \ldots, X_n)$, and $f(x) = f(|x|)$ (f depends only on $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$), then

$$\begin{aligned}
f_R(r) &= n\omega_n f(r) r^{n-1}, r > 0, \\
f_{R^2}(z) &= \frac{\pi^{n/2}}{\Gamma(n/2)} f\left(\sqrt{z}\right) z^{n/2-1}, z > 0,
\end{aligned}$$

where $\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ is the volume of the unit ball.

a) We have

$$f(z) = f_Z(z) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{|z|^2}{2}}, z \in \mathbf{R}^n,$$

and $f(z)$ depends only on $|z|$.

$$\begin{aligned}
|S_{n-1}| &= n\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \\
\omega_n &= \frac{2\pi^{n/2}}{n\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(1 + n/2)}.
\end{aligned}$$

In this case,

$$\begin{aligned}
f_R(r) &= \frac{2\pi^{n/2}}{\Gamma(n/2)} \frac{1}{(2\pi)^{n/2}} e^{-r^2/2} r^{n-1} \\
&= \frac{1}{2^{n/2-1}\Gamma(n/2)} e^{-r^2/2} r^{n-1}, r > 0,
\end{aligned}$$

and

$$
\begin{aligned}
f_{R^2}(z) &= \frac{1}{(2\pi)^{n/2}} \frac{\pi^{n/2}}{\Gamma(n/2)} e^{-\frac{z}{2}} z^{n/2-1} = \frac{1}{2^{n/2}} \frac{1}{\Gamma(n/2)} e^{-\frac{z}{2}} z^{n/2-1} \\
&= \frac{1}{\Gamma(n/2)} \left(\frac{z}{2}\right)^{\frac{n}{2}-1} \left(\frac{1}{2} e^{-\frac{z}{2}}\right), z > 0,
\end{aligned}
$$

i.e. it has $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$-distribution which is also called χ^2 -distribution with $n-1$ degrees of freedom.

b) For X, uniform in the ball of radius a, $f(x) = f_X(x) = \frac{1}{\omega_n a^n}$, $|x| \le a$, depends on $|x|$ only as well. So,

$$
\begin{aligned}
f_R(r) &= n\omega_n \frac{1}{w_n a^n} r^{n-1} = \frac{n r^{n-1}}{a^n}, 0 < r \le a, \\
f_{R^2}(r) &= \frac{n r^{\frac{n}{2}-1}}{2a^n}, 0 < r \le a^2,
\end{aligned}
$$

15. Suppose (X, Y) has joint density of the form $f(x, y) = g\left(\sqrt{x^2 + y^2}\right)$ for $(x, y) \in \mathbf{R}^2$, for some function g. Show that $Z = Y/X$ has the Cauchy density $h(t) = \frac{1}{\pi} \frac{1}{1+t^2}$, $t \in \mathbf{R}$.

HINT: Polar coordinates.

Answer. For $z \in \mathbf{R}$, using polar coordinates in the double integral, $x = r\cos\theta$, $y = r\sin\theta$,

$$
\begin{aligned}
\mathbf{P}(Z \le z) &= \mathbf{P}\left(\frac{Y}{X} \le z\right) = \int_{\mathbf{R}^2} I_{\{\frac{y}{x} \le z\}} g\left(\sqrt{x^2+y^2}\right) dx dy \\
&= \int_0^{2\pi} \int_0^\infty I_{\{\tan\theta \le z\}} g(r) r dr d\theta = \left(\frac{1}{2\pi} \int_0^{2\pi} I_{\{\tan\theta \le z\}} d\theta\right) 2\pi \int_0^\infty g(r) r dr \\
&= \frac{1}{2\pi} \int_0^{2\pi} I_{\{\tan\theta \le z\}} d\theta = \frac{\pi + 2\tan^{-1} z}{2\pi} = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} z, z \in \mathbf{R}.
\end{aligned}
$$

because looking at the graph $y = \tan\theta$, $\theta \in [0, 2\pi]$,

$$
\int_0^{2\pi} I_{\{\tan\theta \le z\}} d\theta = \pi + 2\tan^{-1} z.
$$

Then

$$
f_Z(z) = \frac{1}{\pi} \frac{1}{1+z^2}, z \in \mathbf{R}.
$$

It is the pdf of Cauchy r.v.

16. Consider Bernoulli trials with success probability $p \in (0, 1)$. Let p_n be the probability of an odd number of successes in n trials.

(a) Express p_n in terms of p_{n-1}. Hint. 1st step analysis.

Answer. Let X_n be number of successes in n trials. Then

$$
\begin{aligned}
p_n &= \mathbf{P}(X_n \text{ is odd}) = \mathbf{P}(X_n \text{ is odd} | X_1 = 1) \mathbf{P}(X_1 = 1) + \mathbf{P}(X_n \text{ is odd} | X_1 = 0) \mathbf{P}(X_1 = 0) \\
&= (1 - p_{n-1}) p + p_{n-1} q = p + (q - p) p_{n-1}
\end{aligned}
$$

(b) Based on (a), for what value λ does $p_{n-1} = \lambda$ imply $p_n = \lambda$?

Answer. If $p_{n-1} = \lambda$, then

$$p_n = p + (q - p)\,\lambda = \lambda$$

if

$$p = \lambda\,(1 + p - q) = 2\lambda p, \lambda = 1/2.$$

(c) Show that $\lim_n p_n = \lambda$, the value you found in (b). HINT: Write p_n as $\lambda + \varepsilon_n$, for the λ you found in (b).

Answer. We have

$$\begin{aligned} p_n &= p + (q - p)\,p_{n-1}, \\ \lambda &= p + (q - p)\,\lambda. \end{aligned}$$

Subtracting the 2nd equation from the first, and going "down",

$$\begin{aligned} p_n - \lambda &= (q - p)\,(p_{n-1} - \lambda) = (q - p)^2\,(p_{n-2} - \lambda) = \ldots (q - p)^{n-1}\,(p_1 - \lambda) \\ &= (q - p)^{n-1}\,(p - \lambda) \end{aligned}$$

We get

$$p_n = \frac{1}{2} + (q - p)^{n-1}\left(p - \frac{1}{2}\right) \to \frac{1}{2} \text{ as } n \to \infty.$$

11 Week 12 -13: 11/7- 11/9, 11/14 - 11/18

1. Let X, Y be independent standard normal $N(0,1)$ random variables.

(a) Find a for which $U = X + 2Y$ and $V = aX + Y$ are independent.

Answer. Let us find for what a, U and V are uncorrelated:

$$\mathrm{Cov}(U,V) = \mathrm{Cov}(X+2Y, aX+Y) = a\mathrm{Var}(X) + 2\mathrm{Var}(Y) = a + 2 = 0, a = -2.$$

Since

$$(U,V) = (X+2Y, -2X+Y) = (X,Y)\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix},$$

(U,V) is normal bivariate with $\mathrm{Cov}(U,V) = 0 : U, V$ are independent.

(b) Find $\mathbf{E}(XY|X+2Y=a)$ for all $a \in \mathbf{R}$. HINT: Use (a).

Answer. Following the hint: $X = U - 2(V+2X)$, $X = \frac{1}{5}U - \frac{2}{5}V$,

$Y = V + 2X = V + \frac{2}{5}U - \frac{4}{5}V = \frac{2}{5}U + \frac{1}{5}V$

$XY = \left(\frac{1}{5}U - \frac{2}{5}V\right)\left(\frac{2}{5}U + \frac{1}{5}V\right) = \frac{2}{25}U^2 - \frac{3}{25}UV - \frac{2}{25}V^2$

and

$$\begin{aligned}\mathbf{E}(XY|U) &= \frac{2}{25}U^2 - \frac{2}{25}\mathbf{E}\left(V^2\right) \\ &= \frac{2}{25}U^2 - \frac{2}{25}\mathrm{Var}(V) = \frac{2}{25}U^2 - \frac{2}{25}\cdot 5 = \frac{2}{25}U^2 - \frac{2}{5}.\end{aligned}$$

`2nd answer(without knowing (a) at all).` $\mathrm{Var}(U) = 5$, $\mathrm{Cov}(X,U) = 1$, $\rho(X,U) = \frac{1}{\sqrt{5}}$, $\mathrm{Cov}(Y,U) = 2$, $\rho(Y,U) = \frac{2}{\sqrt{5}}$. Then

$$\begin{aligned}X &= \frac{1}{\sqrt{5}}\frac{1}{\sqrt{5}}U + Z_1 = \frac{1}{5}U + Z_1, \\ Y &= \frac{2}{\sqrt{5}}\frac{1}{\sqrt{5}}U + Z_2 = \frac{2}{5}U + Z_2\end{aligned}$$

Then

$$XY = \left(\frac{1}{5}U + Z_1\right)\left(\frac{2}{5}U + Z_2\right) = \frac{2}{5}UZ_1 + \frac{1}{5}UZ_2 + Z_1Z_2 + \frac{2}{25}U^2,$$

and

$$\mathbf{E}(XY|U) = \frac{2}{25}U^2 + \mathbf{E}(Z_1Z_2) = \frac{2}{25}U^2 + Cov(Z_1Z_2) = \frac{2}{25}U^2 - \frac{2}{5},$$

because

$$Cov(Z_1,Z_2) = Cov\left(X - \frac{1}{5}U, Y - \frac{2}{5}U\right) = -\frac{2}{5} - \frac{2}{5} + \frac{2}{25}\cdot 5 = -\frac{2}{5},$$

2. The joint generating function of X, Y is defined as the function

$$G_{X,Y}(s,t) = \mathbf{E}\left(s^X t^Y\right).$$

Show, without justifying, that $G_X(s) = G_{(X,Y)}(s,1)$ and $G_Y(t) = G_{(X,Y)}(1,t)$, and

$$\mathbf{E}(XY) = \frac{\partial^2 G_{(X,Y)}(s,t)}{\partial s \partial t}|_{s=t=1}.$$

3. Number of cars N arriving at a McDonalds drive-up window is Poisson(λ). The number of passengers in these cars are independent random variables X_i each equally likely to be one, two, three or four.

a) Find the probability generating function of N,say $G_N(s) = \mathbf{E}\left(s^N\right)$.

Answer. By definition

$$G_N(s) = \mathbf{E}\left(s^N\right) = \sum_{n=0}^{\infty} s^n \frac{\lambda^n}{N!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(s\lambda)^n}{n!} = e^{-\lambda} e^{s\lambda} = e^{\lambda(s-1)}, s \in \mathbf{R}.$$

b) Find the moment generating function of $X = X_i$,say $M_X(t) = \mathbf{E}\left(e^{tX}\right)$.

Answer. By definition,

$$M_X(t) = \mathbf{E}\left(e^{tX}\right) = \sum_{n=1}^{4} e^{tn} \frac{1}{4} = \frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4}, t \in \mathbf{R}.$$

c) Find the moment generating function of the total number of passengers passing by the drive up window in a given day. Hint. That number $S = \sum_{i=1}^{N} X_i$.

Answer. The number of passengers $S = \sum_{i=1}^{N} X_i$, and, using a) and b),

$$\begin{aligned} M_S(t) &= \mathbf{E}\left(e^{tS}\right) = \mathbf{E}\left[\mathbf{E}\left(e^{tS}|N\right)\right] \\ &= \sum_{n=0}^{\infty} \mathbf{E}\left(e^{tS}|N = n\right) \mathbf{P}(N = n) \\ &= \sum_{n=0}^{\infty} \mathbf{E}\left(e^{t\sum_{i=1}^{n} X_i}\right) \mathbf{P}(N = n) \\ &= \sum_{n=0}^{\infty} M_X(t)^n \mathbf{P}(N = n) = G_N(M_X(t)) \\ &= \exp\left\{\lambda\left(\frac{e^t + e^{2t} + e^{3t} + e^{4t}}{4} - 1\right)\right\}, t \in \mathbf{R}. \end{aligned}$$

4. Assume that Y and $X_1, \dots$ are independent, $\mathbf{P}(Y = n) = 2^{-n}, n \geq 1$; X_i are independent identically distributed, $\mathbf{P}(X_i > t) = e^{-\pi t}, t > 0, i \geq 1$.

Let $S_n = \sum_{i=1}^{n} X_i, n \geq 1$, and $Z = S_Y = \sum_{i=1}^{N} X_i$.

a) Find $\mathbf{E}(Z)$.

b) Find the probability generating function $\mathbf{E}\left(s^Y\right)$.

c) Find the moment generating function $\mathbf{E}(\exp\{\beta X\})$.

d) Find the moment generating function $\mathbf{E}(\exp\{\beta Z\})$.

e) Find $\mathbf{E}\left(Z^3\right)$.

5. a) Let $g(u) = \mathbf{E}\left(u^S\right)$ is the generating function of a nonnegative integer valued r.v. S such that $\mathbf{P}(S > 0) > 0$. Let T be distributed as S, given $S > 0$. Express $h(u) = \mathbf{E}\left(u^T\right)$, the generating function of T in terms of g.

Answer. T is distributed as S, given $S > 0$ means

$$\mathbf{P}(T \leq t) = \mathbf{P}(S \leq t | S > 0) = \frac{\mathbf{P}(0 < S \leq t)}{\mathbf{P}(S > 0)}, t > 0.$$

We have

$$\begin{aligned} h(u) &= \mathbf{E}\left(u^T\right) = \mathbf{E}\left(u^S | S>0\right) = \frac{\mathbf{E}\left(u^S, S>0\right)}{\mathbf{P}(S>0)} = \frac{\mathbf{E}\left[u^S - u^S I_{\{S=0\}}\right]}{\mathbf{P}(S>0)} = \frac{\mathbf{E}\left(u^S\right) - E\left(u^S I_{\{S=0\}}\right)}{1-\mathbf{P}(S=0)} \\ &= \frac{\mathbf{E}\left(u^S\right) - E\left(I_{\{S=0\}}\right)}{1-\mathbf{P}(S=0)} = \frac{\mathbf{E}\left(u^S\right) - \mathbf{P}(S=0)}{1-\mathbf{P}(S=0)} = \frac{g(u)-g(0)}{1-g(0)} \end{aligned}$$

In the following parts b) and c) of this problem, N is a nonnegative integer valued r.v. with generating function $f(u) = \mathbf{E}\left(u^N\right)$, and S is the number of heads in N tosses of a $p \in (0,1)$ coin, with all coin tosses having probability p of coming up heads, independently of each other and N.

b) Write the generating function $g(u) = \mathbf{E}\left(u^S\right)$ of S.

Answer. Given $N = n$, S is bin(n, p) and

$$\mathbf{E}\left(u^S | N = n\right) = (q+up)^n,$$

and, denoting $q = 1-p$,

$$\begin{aligned} g(u) &= \mathbf{E}\left(u^S\right) = \mathbf{E}\left[\mathbf{E}\left(u^S | N\right)\right] = \sum_{n=0}^{\infty} \mathbf{E}\left(u^S | N=n\right) \mathbf{P}(N=n) \\ &= \sum_{n=0}^{\infty} (q+up)^n \mathbf{P}(N=n) = f(q+up). \end{aligned}$$

c) Now combine parts a) and b): What is the probability generating function of the number T of heads in N tosses of a p-coin, conditional on at least one head, when N has generating function f?

Answer. We have by a), and b),

$$\begin{aligned} \mathbf{E}\left(u^T\right) &= \mathbf{E}\left(u^S | S>0\right) = \frac{g(u)-g(0)}{1-g(0)} \\ &= \frac{f(q+up)-f(q)}{1-f(q)} \end{aligned}$$

Now following questions can be answered independently:

d) Suppose someone claims that for $\alpha \in (0,1)$, the function $f(u) = 1-(1-u)^\alpha$ is generating function of a nonnegative integer valued r.v. N. What properties of f must you check? Is the hypothesis $\alpha > 0$ used? What happens if $\alpha = 0, 1$ and $\alpha > 1$?

Answer. Moment generating function is power series

$$\mathbf{E}\left(u^N\right) = \sum_{n=0}^{\infty} u^n p_n$$

converging absolutely for $|s| \leq 1$ with **nonnegative** coefficients at s^n whose sum is 1:

$$\sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \mathbf{P}(N=n) = 1.$$

By Calculus 2, binomial series,

$$\begin{aligned}
(1-u)^{\alpha} &= \sum_{n=0}^{\infty}\binom{\alpha}{n}x^{n}=1-\alpha u+\frac{\alpha(\alpha-1)}{2!}u^{2}-\frac{\alpha(\alpha-1)(\alpha-2)}{3!}u^{3}+\ldots \\
f(u) &= 1-(1-u)^{\alpha}=\alpha u-\frac{\alpha(\alpha-1)}{2!}u^{2}+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}u^{3}-\ldots,|u|\le 1.
\end{aligned}$$

All coefficients at u^n are nonnegative for $\alpha\in(0,1)$ and

$$1=f(1)=\alpha-\frac{\alpha(\alpha-1)}{2!}+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}-\frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{4!}\ldots.$$

Therefore, indeed, $f(u)$ is generating function.

For $\alpha=0$, $f(u)=1$ which would mean that $p_0=1, p_k=0, k\ge 1$, that is $N=0$ with probability 1.

For $\alpha=1$, $f(u)=1-(1-u)=u$, we have $p_1=1, p_k=0$ if $k\ne 1$. So, $N=1$ with probability 1 in this case.

For $\alpha>1$, not all coefficients at u^n are positive: $f(u)$ cannot be generating function in this case.

e) Combine parts a)-d) and suppose $\alpha\in(0,1)$, N has generating function $f(u)=1-(1-u)^{\alpha}$, and T is the number of heads in N tosses of a p-coin, given at least one head. Do N, T have the same distribution?

Answer. According to part c),

$$\begin{aligned}
\mathbf{E}\left(u^{T}\right) &= \frac{f(q+up)-f(q)}{1-f(q)}=\frac{(1-q)^{\alpha}-(1-q-up)^{\alpha}}{(1-q)^{\alpha}} \\
&= 1-\frac{(1-q-up)^{\alpha}}{(1-q)^{\alpha}}=1-\left(\frac{1-q-up}{1-q}\right)^{\alpha}=1-(1-u)^{\alpha}=f(u).
\end{aligned}$$

Since generating functions coincide, N and T have the same distribution.

6. Fix $p\in(0,1)$ and consider independent Poisson random variables $X_k, k\ge 1$ with

$$\mathbf{E}(X_k)=\frac{p^k}{k}.$$

Verify that the sum $\sum_{k=1}^{\infty}kX_k$ converges with probability one and determine the distribution of a r.v. $Y=\sum_{k=1}^{\infty}kX_k$. Hint. Compute generating functions of $X_k,\ kX_k,$ and Y.

Answer. First,

$$\mathbf{E}(Y)=\mathbf{E}\left(\sum_{k=1}^{\infty}kX_k\right)=\sum_{k=1}^{\infty}k\mathbf{E}(X_k)=\sum_{k=1}^{\infty}k\frac{p^k}{k}=\sum_{k=1}^{\infty}p^k=\frac{p}{1-p}<\infty.$$

Hence $\sum_{k=1}^{\infty}kX_k<\infty$ with probability 1.

We find generating functions:

$$\begin{aligned}
G_{X_k}(s) &= \mathbf{E}\left(s^{X_k}\right)=\exp\left\{\frac{p^k}{k}(s-1)\right\}, \\
G_{kX_k}(s) &= \mathbf{E}\left(s^{kX_k}\right)=\mathbf{E}\left(\left(s^k\right)^{X_k}\right)=\exp\left\{\frac{p^k}{k}\left(s^k-1\right)\right\},
\end{aligned}$$

and, by independence,

$$\begin{aligned} G_Y(s) &= \mathbf{E}\left(s^Y\right) = \prod_{k=1}^{\infty} \exp\left\{\frac{p^k}{k}\left(s^k-1\right)\right\} \\ &= \exp\left\{\sum_{k=1}^{\infty}\frac{p^k}{k}\left(s^k-1\right)\right\} \end{aligned}$$

Recall Calculus 2:

$$\sum_{n=1}^{\infty}\frac{x^n}{n} = -\ln(1-x),\, |x|<1.$$

Hence, denoting $q = 1-p$,

$$\begin{aligned} \sum_{k=1}^{\infty}\frac{p^k}{k}\left(s^k-1\right) &= \sum_{k=1}^{\infty}\frac{(sp)^k}{k} - \sum_{k=1}^{\infty}\frac{p^k}{k} \\ &= -\ln(1-sp) + \ln(1-p) = \ln\frac{q}{1-sp}, \end{aligned}$$

and

$$G_Y(s) = \exp\left\{\ln\frac{q}{1-sp}\right\} = \frac{q}{1-sp}.$$

Recall generating function of U ~geometric(q) :

$$G_U(s) = \mathbf{E}\left(s^U\right) = \sum_{n=1}^{\infty} s^n p^{n-1} q = \frac{sq}{1-sp}.$$

Thus

$$G_{U-1}(s) = \mathbf{E}\left(s^{U-1}\right) = \frac{\mathbf{E}\left(s^U\right)}{s} = \frac{q}{1-sp},$$

that is Y and $U-1$ are identically distributed.

7. Let X, Y be independent, identically distributed with exponential pdf $f(x) = e^{-x}, x>0$. Let $Z = X - Y$.

Calculate the mgf of X and the mgf of Z. Calculate the pdf of Z.

8. Consider a branching process with immigration: each generation is supplemented by an "immigrant" with probability p. This means that the size Z_n of the n-th generation satisfies

$$Z_n = I_n + \sum_{k=1}^{Z_{n-1}} X_k,$$

where $I_n = 1$ with probability p and $I_n = 0$ otherwise; the number of children X_k of the kth person in the nth generation are independent identically distributes with generating function $G(s)$. We assume that Z_{n-1}, I_n and X_k are independent. Let $G_n(s) = G_{Z_n}(s)$ and $\mu_n = \mathbf{E}(Z_n)$.

(a) Show that $G_n(s) = [ps + (1-p)]\, G_{n-1}(G(s))$. Hint: condition on Z_{n-1}.

(b) Show that $\mu_n = p + \mu_{n-1}\mu$;

(c) If $\mu_n \to \mu_\infty$ as $n \to \infty$, what is μ_∞ in terms of p and μ.

9. Let Z_n be the size of nth generation in an ordinary branching process with $Z_0 = 1, \mathbf{E}(Z_1) = \mu$ and var$(Z_1) > 0$. Show that $\mathbf{E}(Z_n Z_m) = \mu^{n-m}\mathbf{E}\left(Z_m^2\right)$ for $m \le n$. Hence find the correlation coefficient $\rho(Z_n, Z_m)$ in terms of μ.

Answer. We have

$$Z_n = X_1 + \ldots + X_{Z_m} = \sum_{i=1}^{Z_m} X_i,$$

where X_i is the number of descendents of ith individual of mth generation in the nth generation, meaning X_i is the number of descendents of ith individual of mth generation in $n-m$ generations: X_i and Z_{n-m} are identically distributed.

Hence

$$\mathbf{E}(Z_n Z_m) = \mathbf{E}\left[\sum_{i=1}^{Z_m} X_i Z_m\right] = \mathbf{E}\left[\mathbf{E}\left(\sum_{i=1}^{Z_m} X_i Z_m | Z_m\right)\right] = \mathbf{E}\left(Z_m^2 \mu^{n-m}\right) = \mu^{n-m}\mathbf{E}\left(Z_m^2\right)$$

because

$$\begin{aligned}
\mathbf{E}\left(\sum_{i=1}^{Z_m} X_i Z_m | Z_m = k\right) &= \mathbf{E}\left(k\sum_{i=1}^{k} X_i | Z_m = k\right) = k^2\mathbf{E}(Z_{n-m}) = k^2\mu^{n-m}, \\
\mathbf{E}\left(\sum_{i=1}^{Z_m} X_i Z_m | Z_m\right) &= Z_m^2\mu^{n-m},
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{cov}(Z_n, Z_m) &= \mathbf{E}(Z_n Z_m) - \mathbf{E}(Z_n)\mathbf{E}(Z_m) = \mu^{n-m}\mathbf{E}\left(Z_m^2\right) - \mu^{n+m}, \\
\rho(Z_n, Z_m) &= \frac{\operatorname{cov}(Z_n, Z_m)}{\sqrt{\operatorname{var}(Z_n)\operatorname{var}(Z_m)}}
\end{aligned}$$

By Lemma (2) of 5.4, for $\mu = 1, n \ge m$,

$$\begin{aligned}
\operatorname{cov}(Z_n, Z_m) &= m\sigma^2 + 1 - 1 = m\sigma^2, \\
\rho(Z_n, Z_m) &= \frac{m\sigma^2}{\sqrt{n\sigma^2 m\sigma^2}} = \sqrt{\frac{m}{n}}.
\end{aligned}$$

10. Let a random variable X be normal $N\left(\mu, \sigma^2\right)$, and let the conditional distribution of Y given X be normal $N\left(a + bX, \sigma_1^2\right)$.

a) Find the joint mgf of (X, Y) defined as

$$M(s,t) = \mathbf{E}\left(e^{sX+tY}\right);$$

(simplify the answer), and the mgf of Y.

Hint. Condition on X. Do not integrate: recall what is the mgf of a normal r.v.

b) Is (X, Y) normal bivariate? Hint. The joint mgf of normal bivariate $(X_1, X_2) \sim N(\mu, B)$ is the function

$$\phi(t) = \exp\left\{it\mu' + \frac{1}{2}tBt'\right\}, t = (t_1, t_2) \in \mathbf{R}^2.$$

12 Week 14 -15: 11/21, 11/28 -12/2

1. Let $\phi_{(X,Y)}(s,t), s,t \in \mathbf{R}$, be the joint characteristic function of X, Y:

$$\phi_{(X,Y)}(s,t) = \mathbf{E}\left(e^{isX+itY}\right).$$

Show that $\phi_X(s) = \phi_{(X,Y)}(s,0), s \in \mathbf{R}$, and $\phi_Y(t) = \phi_{(X,Y)}(0,t), t \in \mathbf{R}$, and, formally, without justifying,

$$\mathbf{E}(XY) = -\frac{\partial^2 G_{(X,Y)}(s,t)}{\partial s \partial t}|_{s=t=0}.$$

2. Let $X = (X_1, \ldots, X_d)$ be multivariate normal $N(\mu, B)$.

(i) Show that

$$\begin{aligned} \phi_X(t) &= \exp\left\{t\mu' - \frac{1}{2}tBt'\right\} \\ &= \exp\left\{\sum_{j=1}^{d} it_j\mu_j - \frac{1}{2}\sum_{k,j=1}^{d} b_{kj}t_kt_j\right\}, t = (t_1, \ldots, t_d) \in \mathbf{R}^d. \end{aligned}$$

Hint. We have $B = D'D$ for some $d \times d$ -matrix D, and $X = ZD + \mu$, where $Z = (Z_1, \ldots, Z_d)$ is r. vector of independent $Z_i \sim N(0,1)$: find ϕ_Z first and use definition of ϕ_X.

(ii) Let $Y = c_1X_1 + \ldots + c_2X_d = cX'$ with $c = (c_1, \ldots, c_d) \neq 0$. Find ϕ_Y. Derive from the form of ϕ_Y that Y is normal.

3. Let $X_1, X_2, \ldots$ be a sequence of independent uniform r.v. in $(0,1)$, and let $Y_n = \min_{1 \le k \le n} X_k$. Show that

$$nY_n \overset{D}{\to} X \sim \exp(1), \text{ and } Y_n \overset{D}{\to} 0.$$

4. Let $X_1, X_2, \ldots$ be independent identically distributed having the moment generating function $M_X(t), -\infty < t < \infty$. Let N be an integer-valued random variable (that is, N takes values$0, 1, 2, \ldots$) with moment generating function $M_N(t), -\infty < t < \infty$. Assume that N is independent of all X_k and define $S = \sum_{k=1}^{N} X_k$, assuming $S = 0$ if $N = 0$.

(a) Find $\mathbf{E}\left(e^{tS}|N = n\right), n = 0, 1, 2, \ldots$, and $\mathbf{E}\left(e^{tS}|N\right)$.

(b) Use (a) to confirm that the random variable S has the moment generating function

$$M_S(t) = M_N(\ln M_X(t)), -\infty < t < \infty.$$

(c) Use (b) to derive the formula $\mathbf{E}(S) = \mu_N\mu_X$, where $\mu_N = \mathbf{E}(N), \mu_X = \mathbf{E}(X_i)$. Find $\mathrm{Var}(S)$.

5. a) Let H_n be number of H in n independent tosses of a p-coin. Apply CLT to approximate $\mathbf{P}\left(a < \frac{H_n}{n} < b\right), 0 < a < b < 1$, for large n.

b) Let Y_n be Poisson(n). Apply CLT to approximate $\mathbf{P}(a < Y_n < b), 0 < a < b$, for large n. Hint, $Y_n = X_1 + \ldots + X_n$, where X_k are independent Poisson(1).

6. It is well known that infants born to mothers who smoke tend to be small and prone to a range of ailments. It is conjectured that also they look abnormal. Nurses were shown selections of photographs of babies, one half of whom had smokers as mothers; the nurses were asked to judge from a baby's appearance whether or not the mother smoked. In 1500 trials the correct answer was given 910 times. Is the conjecture plausible? If so, why?

Hint. Assume the probability that a nurse judgement is correct with probability 1/2, and let X_n be the number of correct answers in n independent judgements. Estimate approximately $\mathbf{P}(X_{1500} \geq 910)$?

7. A sequence of biased coins is flipped; the chance that the kth coin shows a head is U_k, where U_k is a random variable taking values in $(0, 1)$. Let X_n be the number of heads after n flips. Does X_n obey the central limit theorem when the U_k are independent and identically distributed?

8. Let X_n have the binomial distribution $\text{bin}(n, U)$, where U is uniform in $(0, 1)$ (look at hw12). Show that

$$\frac{X_n}{n+1} \xrightarrow{D} U.$$

Hint: enough to show the convergence of characteristic functions.

Comment. CLT does not hold for X_n.

9. Let X_n have distribution function

$$F_n(x) = x - \frac{\sin(2n\pi x)}{2n\pi}, x \in [0, 1].$$

(a) Show that F_n is indeed a distribution function, and that X_n has a density function.

(b) Show that, as $n \to \infty$, F_n converges to the uniform distribution function, but that the density function of f_n does not converge to the uniform density function.

10. You have a choice to roll a fair die either 100 times or 1000 times. For each of the following outcomes, state whether it is more likely with 100 rolls, or with 1000 rolls. Justify your answer, but you do not need to give a full formal proof.

(a) The number 1 shows on the die between 15% and 20% of the time.

(b) The number showing is at most 3, at least half the time.

(c) The number showing is 2 or 5, at least half the time.