Contents

1 hw1

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

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

Hint. The first one: $x \in (\cup_i A_i)^c \iff x \notin \cup_i A_i$ (it means x does not belong to any of them) $\iff x \notin A_i$ for any $i \iff x \in A_i^c$ for any $i \iff x \in \cap_i A_i^c$.

2. *a)* 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.

Hint. Verify the definition of a σ-field. Since $\Omega \in \mathcal{F}_1$ and $\Omega \in \mathcal{F}_2$, then $\Omega \in \mathcal{F}_1 \cap \mathcal{F}_2$ etc...

b) 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 $\mathcal{A}$

3. Suppose that eight distinct envelopes are placed at random in three distinct mailboxes.

a) In how many different ways this can be done?

b) What is the probability that every mailbox receives at least one envelope? *Hint. Consider finding probability of the complementary event: "at least one mailbox is empty" = "exactly one empty" or "exactly two empty".*

4. How many different letter arrangements of length 4 (four letter "words") can be made using the letters MOTTO?

5. (a) In how many ways can 3 boys and 3 girls sit in a row?

(b) In how many ways can 3 boys and 3 girls sit in a row if the boys and the girls are each to sit together?

(c) In how many ways if only the boys must sit together?

(d) In how many ways if no two people of the same sex are allowed to sit together?

Hint. It is like arranging books on a shelf (problem considered in class: check Content folder on blackboard).

2 hw2

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.

2. If 4 married couples are arranged in a row, find the probability that no husband sits next to his wife. *Hint.* Let $A_k =$ "kth couple sits together", $k = 1, 2, 3, 4$. Apply inclusion/exclusion principle and think about the book arrangement on the shelf:

$$\begin{aligned} & \mathbf{P}\left(\text{at least one couple sits together}\right) \\ = \; & \mathbf{P}\left(A_1 \cup A_2 \cup A_3 \cup A_4\right) = \sum_{i=1}^{4} \mathbf{P}\left(A_i\right) - \sum_{i<j} \mathbf{P}\left(A_i A_j\right) + \sum_{i<j<k} \mathbf{P}\left(A_i A_j A_k\right) - \mathbf{P}\left(A_1 A_2 A_3 A_4\right) \\ = \; & \sum_{i=1}^{4} \mathbf{P}\left(A_i\right) - \mathbf{P}\left(A_1 A_2\right) - \mathbf{P}\left(A_1 A_3\right) - \mathbf{P}\left(A_1 A_4\right) - \mathbf{P}\left(A_2 A_3\right) - \mathbf{P}\left(A_2 A_4\right) - \mathbf{P}\left(A_3 A_4\right) \\ & + \mathbf{P}\left(A_1 A_2 A_3\right) + \mathbf{P}\left(A_1 A_2 A_4\right) + \mathbf{P}\left(A_1 A_3 A_4\right) + \mathbf{P}\left(A_2 A_3 A_4\right) - \mathbf{P}\left(A_1 A_2 A_3 A_4\right). \end{aligned}$$

For instance with $i < j$, $A_i A_j =$ "the ith couple and the jth couple go together, the remaining 4 people are arranged in any order".

3. *Answering a question take into consideration all the information revealed before it.*

A man has five coins, two of which are double-headed, one is double-tailed, and two are normal.

He shuts his eyes, picks a coin at random, and tosses it. What is the probability that the lower face of the coin is a head?

He opens his eyes and sees that the coin is showing heads; what is the probability that the lower face is a head?

He shuts his eyes again, and tosses the coin again. What is the probability that the lower face is a head?

He opens his eyes and sees that the coin is showing heads; what is the probability that the lower face is a head?

He discards this coin, picks another at random, and tosses it. What is the probability that it shows heads?

4. 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;

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

5. 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 $C \subseteq \Omega$, $\#C \leq 13$, and 13 is a prime number.

3 hw3

1. In each packet of Corn Flakes may be found a plastic bust of one of the last five Vice-Chancellors of Cambridge University, the probability that any given packet contains any specific Vice-Chancellor being 1/5, independently of all other packets. Show that the probability that each of the last three Vice-Chancellors is obtained in a bulk purchase of six packets is

$$1 - 3\left(\frac{4}{5}\right)^6 + 3\left(\frac{3}{5}\right)^6 - \left(\frac{2}{5}\right)^6.$$

Hint. Inclusion/exclusion?

2. Jane has 3 children, each of which is equally likely to be a boy or a girl independently of the others. Consider the events:

$$\begin{aligned} A &= \text{"all the children are of the same sex"}, \\ B &= \text{"there is at most one boy"}, \\ C &= \text{"the family includes a boy and a girl"}. \end{aligned}$$

(a) Show that A is independent of B, and that B is independent of C.

(b) Is A independent of C?

(c) Do these results hold if Jane has four children?

3. There are two roads from A to B and two roads from B to C. Each of the four roads is blocked by snow with probability p, independently of the others. Find the probability that there is an open road from A to B given that there is no open route from A to C.

If, in addition, there is a direct road from A to C, this road being blocked with probability p independently of the others, find the required conditional probability.

4. Consider a gambler G ruin problem where he starts with k dollars, $0 < k < N$. A fair coin is tossed repeatedly. G wins \$1 if H, and loses \$1 if T. The game stops in two cases: either G is ruined or G reaches the desired amount \$$N$. Show that the game stops with probability 1. Hint. Besides the ruin probability p_k consider the probability $\bar{p}_k$ to reach N. Besides the equation (6) on p.17 for p_k, write a similar equation for $\bar{p}_k$ and solve it (what are $\bar{p}_0, \bar{p}_N$?). Find $p_k + \bar{p}_k$.

5. Consider the following strategy playing the roulette. Bet \$1 on red. If red appears (which happens with probability 18/38), then take \$1 and stop playing for the day. If red does not appear, then bet additional \$1 on red each of the following two rounds, and then stop playing for the day no matter the outcome. Let X be the net gain (a negative gain means a loss).

(a) What are possible values of X? Find $\mathbf{P}(X = k)$ for all possible values k; sketch the distribution function (cdf) of X.

(b) Compute $\mathbf{P}(X > 0)$, the probability of net win. Is it a good strategy?

4 hw4

1. 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 $\varepsilon > 0$,

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

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

2. A communication channel transmits a signal as sequence of digits 0 and 1. The probability of incorrect reception of each digit is p. To reduce the probability of error at reception, 0 is transmitted as 00000 (five zeroes) and 1, as 11111. Assume that the digits are received independently and the majority decoding is used. Compute the probability of receiving the signal incorrectly if the original signal is (a) 0; (b) 101. Evaluate the probabilities when $p = 0.2$.

3. a) Let U be a r.v. with distribution function

$$F_U(x) = \mathbf{P}(U \leq u) = \begin{cases} 0 & \text{if } u < 0, \\ u & \text{if } 0 \leq u \leq 1, \\ 1 & \text{if } u > 1. \end{cases}$$

We say U is uniformly distributed in the interval $[0, 1]$; note that

$$\mathbf{P}(a < U \leq b) = \frac{b-a}{1} = b - a, 0 \leq a < b \leq 1.$$

Let F be a distribution function which is continuous and strictly increasing (note that range of F is $(0, 1)$) with an inverse $F^{-1} : (0, 1) \to \mathbf{R}$, i.e., F^{-1} is strictly increasing continuous, $F\left(F^{-1}(x)\right) = x, x \in (0, 1)$, and $F^{-1}(F(x)) = x, x \in \mathbf{R}$.

Show that $X = F^{-1}(U)$ is a r.v. having distribution function F, i.e. $F = F_X$.

b) Let X be a r.v. with a continuous distribution function F. Find expression for the distribution functions of the following random variables:

(i) X^2; (ii) $\sqrt{X}$; (iii) $F(X)$, assuming F is strictly increasing; (iv) $G^{-1}(F(X))$, assuming F and G are strictly increasing.

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

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

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

Hint:

$$\{Y \leq a\} = \{X < a\} \cup \{X = a\} = \{X \leq a\};$$

What is $\{Y \leq y\}$ in terms of X if $a < y < b$?

5. a) A robot shoots at a random angle $\Theta \in (-\pi/2, \pi/2)$ so that

$$\mathbf{P}(a < \Theta \leq b) = \frac{b-a}{\pi}, -\pi/2 < a < b < \pi/2.$$

Find the df and pdf of Θ.

b) The robot is one mile away from the front line and shoots at it at a random angle Θ (the same as in part a)). Find the df and pdf of the location X on the front line that is hit.

5 hw5

1. 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

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

2. a) Let X and Y have joint distribution function F. Show that for any $a<b, c<d$,

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

Find (in terms of F) the probability

$$\mathbf{P}(X=b, Y=d).$$

Comment. This equality implies that if F is a joint distribution function, then for any $a<b, c<d$,

$$F(b,d)-F(a,d)-F(b,c)+F(a,c)\geq 0. \tag{5.1}$$

If F is twice continuously differentiable around the rectangle $(a,b]\times(c,d]$,using Taylor formula, we have

$$\begin{aligned} 0 &\leq F(b,d)-F(a,d)-F(b,c)+F(a,c) \\ &= \int_0^1\int_0^1 \frac{\partial^2 F(a+s(b-a), c+r(d-c))}{\partial x \partial y} dr\,ds\,(d-c)(b-a). \end{aligned}$$

Hence

$$\frac{\partial^2 F(a,c)}{\partial x \partial y}\geq 0, \tag{5.2}$$

and $\frac{\partial^2 F(x,y)}{\partial x \partial y}\geq 0$ in the rectangle above guaranties that (5.1) holds.

b) Is the function $F(x,y)=1-e^{-xy}, 0\leq x,y<\infty$, the joint distribution function of some pair of r.v.?

Comment. In order for F to be joint distribution function, besides the properties (a), (b), (c) of Lemma 5, p.39, the inequality (5.1) must hold as well, which translates into (5.2), and $\frac{\partial^2 F(x,y)}{\partial x \partial y}\geq 0$ if F is twice continuously differentiable around (x,y).

3. You roll a fair die repeatedly. If it shows 1, you must stop, but you may choose to stop at any prior time. Your score is the number shown by the die on the final roll. Consider the following strategy $S(k)$: stop the first time that the die shows k or greater, $k = 2, 3, 4, 5, 6$.

a) What is the probability that "k or greater" shows up before 1 in the nth roll, $k \geq 2$? What is the probability that "k or greater" ($k = 2, 3, 4, 5, 6$) shows up before 1?

b) Let X_k be the score when the strategy S_k was used. Find $\mathbf{E}(X_k), k = 2, 3, 4, 5, 6$. Which strategy yields the highest expected score?

4. 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 their joint probability mass function and:

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

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

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

5. You pay $\$2^8 = \$256$ to enter and play the following game: A fair coin is tossed until H shows up, and if X is the number of tosses that was needed, you are paid $\$2^X$ for instance, if $X = 10$, then you are paid $\$2^{10}$).

Your win/loss $W = 2^X - 256$. What is $\mathbf{E}(W)$? What is the probability that $W \geq 0$?

6 hw6

1. Let $\mathbf{E}(X) = \mu_1 . \mathbf{E}(Y) = \mu_2$,$\mathrm{Var}(X) = \sigma_1^2$,$\mathrm{Var}(Y) = \sigma_2^2$. Assume

$$Y = aX + Z,$$

and $\mathrm{Cov}(X, Z) = 0$.

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

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.

3. A building has 10 floors above the street level. 20 of people enter at street level and board an elevator. They choose floors independently, equally likely at random. Let T be the number of stops the elevator must make.

Find $\mathbf{E}(T)$and $\mathrm{Var}(T)$. Hint. Use the indicators X_i of the event that at least one person selects floor i.

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 an expression for the variance of the total. Recall

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

5. An urn contains N balls, m of which are red. A random sample of n balls is withdrawn without replacement from the urn. Assume $n < m < N$. Show that the number X of red balls in this sample has the mass function

$$\mathbf{P}(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, k = 0, \ldots, n.$$

This is called the hypergeometric distribution with parameters N, m, n.Show further that if N, m and $N - m$ approach ∞ in such a way that $m/N \to p$ and $(N-m)/N \to 1-p$, then

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

Comment. It shows that for small n and large N, the distribution of X barely depends on whether or not the balls are replaced in the urn immediately after their withdrawal.

7 hw7

1. 51 passengers bought tickets on a 51-seat carriage. One seat was reserved for each passenger. The first 50 passengers took the seats at random so that all 51! possible seating arrangements (with one empty seat) are equally likely. The last passenger insisted on taking the assigned seat. If that seat is occupied, then the passenger in that seat has to move to the corresponding assigned seat, and so on. Compute the expected value of the number M of passengers who have to change their seats. Hint. One way could be to denote E_n the expected value of the number of passengers who have to change their seats in n seat plane (instead of 51), and write a recursion for E_n in terms of E_{n-1}, and going "down" get a general formula for E_n without knowing the distribution of $M = M_n$. Another way is by simply finding the distribution of $M = M_n$ explicitly.

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

$$Y = g\,(X) + W,$$

and W and X are independent. Show that

$$\mathbf{P}\,(Y = y|X = x) = \mathbf{P}\,(g\,(x) + W = y).$$

b) Let

$$X = Y + U,$$

and Y and U are independent. Assume the pmf $f_Y\,(y)$ and $f_U\,(u)$ are known. Find the joint pmf $f\,(x, y)$ of X and Y, and

$$\mathbf{P}\,(Y = y|X = x)$$

in terms of f_U and f_Y.

5. Let $S_n = X_1 + \ldots + X_n, n \geq 1$, be a random walk, where $\mathbf{E}\,(X_k) = \mu$, var$(X_k) = \sigma^2$, and $X_1, X_2, \ldots$ are independent.

(a) Find $\mathbf{E}\,(S_n)$ and var(S_n); (b) Let $n > m \geq 1$. Find cov(S_n, S_m) and the correlation coefficient $\rho\,(S_n, S_m)$. Hint: $S_n = S_m + \sum_{i=1}^{n-m} X_{m+i}$.

Find $\lim_{n\to\infty}$cov(S_n, S_m) and $\lim_{n\to\infty}\rho\,(S_n, S_m)$.

8 hw8

1. 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 at N, and let $p_k = \mathbf{P}(W|S_0 = k)$. Show that, if the particle starts at k with $0 < k < N$, the conditional probability that the first step is rightwards, given W, equals $\frac{pp_{k+1}}{p_k}$. Deduce that the expected duration of the walk, conditional on W when $S_0 = k$, satisfies the equation

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

Take $J_0 = 0$ as a boundary condition.

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

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

2. 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. Let $S_0 = k, 0 < k < N$.

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

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

where D_k is the mean number of steps until absorption and p_k is the probability of absorption at 0 (the gambler ruin probability). 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?

3. Consider a symmetric simple random walk S_n with $S_0 = 0, p = q = 1/2$. Let $\tau_0 = \min\{n \geq 1 : S_n = 0\}$ be the time of the first return of the walk to its starting point.

a) Show that

$$\mathbf{P}(\tau_0 = 2n) = \frac{1}{2n-1}\binom{2n}{n}2^{-2n} = \frac{1}{2n-1}\mathbf{P}(S_{2n} = 0).$$

Hint. Use total probability law with $\Omega = \{S_1 = 1\} \cup \{S_1 = -1\}$, hitting time theorem (14), p. 79 with space homogeneity property of r. walk.

b) Show that

$$\lim_{n\to\infty}\frac{\mathbf{P}(S_{2n} = 0)}{1/\sqrt{\pi n}} = \lim_{n\to\infty}\frac{\mathbf{P}(S_1 \ldots S_{2n} \neq 0)}{1/\sqrt{\pi n}} = 1.$$

Hint. Stirling's formula:

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

4. 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.

5. 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, $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$.

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

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

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

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

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 $\mathbf{E}(S_{\sigma_k})$ and $\mathrm{var}(S_{\sigma_k})$. Hint: $\sigma_k = \min\{\tau_k, \tau_{-k}\}$ and $\mathbf{P}(\tau_k < \infty) = \mathbf{P}(\tau_{-k} < \infty) = 1$. What values S_{σ_k} and $S_{\sigma_k}^2$ take?

9 hw9

1. Let S_n be a simple r.w.

a) Assume $S_0 = 0$. Let $\tau_0 = \min\{k > 0 : S_k = 0\}$, the return time back to zero. Show that

$$1 = \sum_{j=0}^{n} \mathbf{P}(\tau_0 > j)\, \mathbf{P}(S_{n-j} = 0).$$

Hint. Think about $\mathbf{P}(\sigma_n = j)$, where $\sigma_n = \max\{k \le n : S_k = 0\}$.

b) Let $S_0 = a > 0$, $p = q = 1/2$. Let $\tau_0 = \min\{k \ge 0 : S_k = 0\}$, the hitting time of zero. What does the reflection principle say about

$$\mathbf{P}(S_n = j, \tau_0 \le n | S_0 = a)$$

with $j, n \ge 0$?

2. a) Let Θ be uniform on $(0, \pi)$, and $a > 0$. Find the pdf of $Y = a \cos \Theta$. Hint. Find pdf 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.

c) Let X, Y be independent random variables with common distribution function F and pdf 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 $U = \min\{X, Y\}$.

3. Let $X_1, \ldots, X_n$ be positive independent identically distributed continuous random variables for which $\mathbf{E}(X_1^{-1})$ exists. Show that, if $m < n$, then $\mathbf{E}\left(\frac{S_m}{S_n}\right) = \frac{m}{n}$, where $S_k = X_1 + \ldots + X_k$.

Hint. The random variables $\frac{X_i}{S_n} = \frac{X_i}{X_1 + \ldots + X_n}, i = 1 \ldots, n$, are identically distributed. First think what is $\mathbf{E}\left(\frac{X_i}{S_n}\right)$?

4. Let $X \ge 0$ be continuous with pdf f. Show that

$$\mathbf{E}(X^r) = \int_0^\infty r x^{r-1} \mathbf{P}(X > x)\, dx = \int_0^\infty r x^{r-1} [1 - F(x)]\, dx$$

for any $r \ge 1$ for which the expectation is finite. With $r = 1$, we have $\mathbf{E}(X) = \int_0^\infty \mathbf{P}(X > x)\, dx$. Hint. Use pdf for the probability and change the order of integration.

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

(a) $X_1 < X_2 < X_3 < X_4$.

(b) $X_1 > X_2 < X_3 < X_4$. Hint. One way is to enumerate all possibilities and rewrite this event as a disjoint union of (a) type. The other way (without thinking): with probability 1,

$$\begin{aligned} I_{\{X_1 > X_2 < X_3 < X_4\}} &= I_{\{X_1 > X_2\}} I_{\{X_2 < X_3\}} I_{\{X_3 < X_4\}} \\ &= I_{\{X_1 > X_2\}} \left(1 - I_{\{X_2 > X_3\}}\right) \left(1 - I_{\{X_3 > X_4\}}\right). \end{aligned}$$

10 hw10

1. *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 (arrangement) 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 probability 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 directly (without using the joint df find in a)). Hint. First, show that the df of $X_{(k)}$ is

$$F_{X_{(k)}}(x) = \mathbf{P}\left(X_{(k)} \le x\right) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j \left[1 - F(x)\right]^{n-j} :$$

note $\sum_{j=1}^{n} I_{\{X_j \le x\}}$ is a binomial r.v.

c) Find the joint density function of the order statistics of n independent uniform variables in $(0, T)$. Hint: use a)

2. a) Let X, Y be independent standard normal. What is joint pdf of (X, Y)? Show that $R = X^2 + Y^2$ and $V = \frac{X}{\sqrt{X^2+Y^2}}$ are independent. Hint. All the values of V are between -1 and 1. Write $\mathbf{P}(R \le r, V \le v)$ as a double integral and use polar coordinates for it.

b) Let X, Y, Z be independent standard normal. Show that

$$\frac{X + YZ}{\sqrt{1 + Z^2}}$$

is standard normal. Hint: condition on Z.

3. Let (X_1, X_2) be standard bivariate normal with $\rho = 3/5$. Let (Y_1, Y_2) be the midterm and final exam scores of a randomly selected student. Assume

$$Y_1 = 80 + 3X_1, Y_2 = 75 + 2X_2.$$

Given a student got 90 in the midterm exam,

(a) What is the conditional expectation and conditional variance of her final exam score? Hint. Probably easier to reduce the question to (X_1, X_2) but also (Y_1, Y_2) is a normal bivariate.

(b) What is the conditional probability that she got more than 75 in the final exam? Express the probability in terms of $\Phi(x)$, the df of a standard normal r.v.

4. Let X, Y be normal bivariate r.v. with $\mu_1 = \mu_2 = 0$, variances σ_1^2, σ_2^2 and correlation coefficient ρ.

a) Write what are $\mathbf{E}(X|Y), var(X|Y)$?

b) Show that

$$\begin{aligned} \mathbf{E}(X|X+Y=z) &= \frac{\sigma_1^2 + \rho\sigma_1\sigma_2}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2} z, \\ \operatorname{var}(X|X+Y=z) &= \frac{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2}. \end{aligned}$$

Hint. $(X, X+Y)$is normal bivariate: apply a).

5. Let $Y = X + \varepsilon Z$, where $X \sim N\left(0, \sigma_1^2\right), Z \sim N(0,1)$ are independent (X is a random "target", εZ is a "noise", Y is what we observe and register). Find the best mean square estimate of X based on Y (recall it is $\hat{X} = \mathbf{E}(X|Y)$). Find the mean square error $\mathbf{E}\left[\left(X - \hat{X}\right)^2\right]$.