McGill University

MATH 323, Winter 2017

Assignment 5

1 April, 2017

- You do not have to submit this assignment.
- Try **every** problem yourself before looking at the solutions.

---

1. A drawer contains n pairs of socks of n different colors. (That is, there is a left sock and a right sock for each of those n colors, and hence, all $2n$ socks are distinguishable.) Let X denote the number of matching pairs in a sample of four socks selected without replacement.

   Find the pmf of X.

2. A total of $n := 3n_1 + 2n_2 + n_3$ distinct items labeled $1, \ldots, n$ are to be put into $n_1 + n_2 + n_3$ labeled boxes in the following way: the n items will be split into n_1 groups of three items each, n_2 groups of two items each, and n_3 groups of one item each, and each group of items will be placed into a single box.

   Assuming all possibilities to be equally likely, find the probability that item 1 and item 2 are not put in a box together (either by themselves or with a third item).

3. A group of 12 researchers consists of 4 faculty members from each of the three universities McGill, UdeM, and Concordia. From the group, 4 researchers are to be selected at random to form a committee.

   Find the probability that the committee will have representatives from all three universities.

4. Suppose $\lambda > 0$ and $p \in (0,1)$. The number of fish caught in a day by a fisherman follows a $\text{Poi}(\lambda)$ distribution. At the end of the day, each fish caught is examined independently and sent off to a market to be sold with probability p (and is discarded with probability $1-p$). Let X be the number of fish sent to the market at the end of the day.

   (a) Compute the mgf of X, and use that to identify its distribution.

   (b) Find the pmf of X directly.

5. Suppose $Y \sim \text{Gamma}(2,1)$ and for each $y > 0$, $X | Y = y \sim \text{Unif}[0, y]$.

   (a) Compute the mgf of X, and use that to identify its distribution.

(b) Find the pdf of X directly.

6. A heat source is placed in a room with a thermometer. The proportion of rise, X, in the thermometer scale follows a Beta$(T,1)$ distribution conditional on the fact that the heat source radiates heat for T units of time. T itself is a random variable having an Exp(β) distribution for some deterministic $\beta > 0$.

   Find the marginal distribution of X.

7. Suppose X_1 and X_2 are independent real-valued RVs with respective densities f_1 and f_2. Define

$$\overline{M} := \max\{X_1, X_2\}, \quad \text{and} \quad \underline{M} := \min\{X_1, X_2\}.$$

   (a) Find the cdf and density of $\overline{M}$.
   (b) Find the cdf and density of $\underline{M}$.

8. Suppose $U_1, U_2, U_3, \ldots$ are i.i.d. Unif$[0,1]$ random variables. Let

$$\overline{M}_n := \max\{U_1, \ldots, U_n\}, \quad \text{and} \quad \underline{M}_n := \min\{U_1, \ldots, U_n\}.$$

   (a) Find the cdf and density of $U_1 + U_2$.
   (b) Find the cdf and density of $U_1 - U_2$.
   (c) Find the cdf and density of $|U_1 - U_2|$.
   (d) Find the cdf and density of $U_1 U_2$.
   (e) Find the cdf and density of $\overline{M}_n$.
   (f) Find the cdf and density of $\underline{M}_n$.
   (g) Show that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\big(\overline{M}_n \le 1 - \varepsilon\big) = 0, \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}\big(\underline{M}_n \ge \varepsilon\big) = 0.$$

9. A magnet and a metal cube are placed in a hallway of length one unit. The locations of the two objects are independent and uniform over the length of the hallway. The attraction of the magnet can overcome the friction of the floor and draw the cube towards itself only if the distance between the two objects is less than $1/2$ unit.

   Find the probability that the cube will start moving towards the magnet.

10. Mary arrives at a restaurant at 12:00 and waits for Shelley and Frank to arrive. Shelley and Frank's arrival times are independent and are uniformly distributed over the time interval 12:00 and 13:00. Mary decides that she will leave after T hours if neither Shelley nor Frank shows up by then.

    (a) If T is a fixed number in $[0,1]$, find the probability that Mary will leave before meeting the others.

(b) If $T \sim \mathrm{Unif}[0,1]$, find the probability that Mary will leave before meeting the others.

11. The joint pmf of X and Y is given in the following table:

$Y \backslash X$	-2	-1	0	1
-3	$1/20$	0	$2/20$	$3/20$
0	0	$4/20$	$2/20$	0
1	$1/20$	$1/20$	$5/20$	$1/20$

(a) Find the marginal distribution of X.

(b) Find the marginal distribution of Y.

(c) Find the conditional distribution of X given Y.

(d) Find the conditional distribution of Y given X.

(e) Find $\mathbb{P}\big(X \le 0 \mid Y = 0\big)$.

(f) Find $\mathbb{P}\big(X \le 0 \mid Y \le 0\big)$.

(g) Find the distribution of $X+Y$, and compute $\mathbb{P}(X+Y > -1)$.

(h) Find $V(X)$.

(i) Find $V(Y)$.

(j) Find $V(X|Y)$.

(k) Find $V(Y|X)$.

(l) Find $\rho_{X,Y}$.

(m) Are X and Y independent?

(n) Find $\mathbb{E}\big[(X+Y)^2 \big| X\big]$.

(o) Find $\mathrm{Cov}(|X|, Y^2)$.

12. The joint density of X and Y is given by

$$f(x,y) = \begin{cases} c(x+y), & \text{if } 0 \le y \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find c.

(b) Find the marginal distribution of X.

(c) Find the marginal distribution of Y.

(d) Find the conditional distribution of X given Y.

(e) Find the conditional distribution of Y given X.

(f) Find $V(X)$.

(g) Find $V(Y)$.

(h) Find $V(X|Y)$.

(i) Find $V(Y|X)$.

(j) Find $\mathbb{P}\big(X \le 1/2 \mid Y = 1/4\big)$.

(k) Find $\mathbb{P}\big(X \le 1/2 \mid Y \le 1/4\big)$.

(l) Find $\rho_{X,Y}$.

(m) Are X and Y independent?

(n) Find the cdf and density of $X+Y$.

(o) Find $\mathbb{E}\big[(X+Y)^2 \big| X\big]$.

(p) Find $\mathrm{Cov}(|X|, Y^2)$.

13. Let $A \subset \mathbb{R}^2$. Define

$$\mathrm{Area}(A) := \iint_{(x,y)\in A} dxdy.$$

Recall from the lectures that for a region $A \subset \mathbb{R}^2$ with $\mathrm{Area}(A) < \infty$, (X,Y) is uniformly distributed over A (in short, $(X,Y) \sim \mathrm{Unif}(A)$) if the function $f(\cdot,\cdot)$ given by

$$f(x,y) := \begin{cases} 1/\mathrm{Area}(A), & \text{if } (x,y) \in A, \\ 0, & \text{otherwise} \end{cases}$$

is a joint density for (X,Y).

If $(X,Y) \sim \mathrm{Unif}(A)$ and $B \subset \mathbb{R}^2$, find $\mathbb{P}\big((X,Y) \in B\big)$.

14. If $(X,Y) \sim \mathrm{Unif}(A)$, find $\mathbb{P}\big((X,Y) \in B\big)$ when

(a) $A = [0,1] \times [0,1]$ and $B = \{(x,y) \,:\, x^2 + y^2 \le 1\}$;

(b) $A = \{(x,y) \,:\, x^2 + y^2 \le 1\}$ and $B = [0,1] \times [0,1]$;

(c) $A = \{(x,y) \,:\, y \ge 0,\ x^2 + y^2 \le 4\}$ and $B = \{(x,y) \,:\, x^2 + y^2 \ge 1\}$;

15. Suppose $g : \mathbb{R} \to [0,\infty)$ is a continuous function such that $\int_{\mathbb{R}} g(x)dx = 1$. Let

$$A := \{(x,y) : \ x \in \mathbb{R},\ 0 \le y \le g(x)\},$$

i.e., A is the part of the plane above the x-axis that falls under the graph of g. Suppose $(X,Y) \sim \mathrm{Unif}(A)$.

Find the density of X.

16. A mine detector is located at the center of a very large field. Assume that it has a detection radius of 1 unit, i.e., a mine within distance 1 of the detector gets spotted. A landmine is hidden in the field, and representing the field by $\mathbb{R}^2$ with the detector as the origin, the location of the landmine is a random variable with density

$$f(x,y) := \begin{cases} |x|/2, & \text{if } (x,y) \in [-1,1] \times [-1,1], \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the probability that the mine gets detected.

(b) If the mine gets detected, it is disarmed by military personnel causing no damage. Otherwise, it explodes, and the cost of repairing the damage caused by an explosion at (x, y) is x^2. Find the expected cost of repair.

17. Find the distribution of $g(X)$ if

 (a) $X \sim N(0,1)$ and $g(u) = e^u$ for $u \in \mathbb{R}$;

 (b) $X \sim \text{Unif}[-2,2]$ and $g(u) = u^2$ for $u \in \mathbb{R}$;

 (c) $X \sim \text{Unif}[-1,1]$ and $g(u) = |u|$ for $u \in \mathbb{R}$;

 (d) $X \sim \text{Unif}[0,1]$ and $g(u) = -\log(1-u)$ for $u < 1$.

18. Suppose X has density f.

 (a) Find the density of $aX + b$.

 (b) Find the density of X^{2n}, $n \in \{1,2,\ldots\}$.

19. Recall the following results from class: Suppose $X_1,\ldots,X_k$ are independent random variables.

 (a) If $X_i \sim N(\mu_i,\sigma_i^2)$, $1 \le i \le k$, then $\sum_{i=1}^k X_i \sim N\big(\sum_{i=1}^k \mu_i, \sum_{i=1}^k \sigma_i^2\big)$.

 (b) If $X_i \sim \text{Poi}(\lambda_i)$, $1 \le i \le k$, then $\sum_{i=1}^k X_i \sim \text{Poi}\big(\sum_{i=1}^k \lambda_i\big)$.

 (c) If $X_i \sim \text{Bin}(n_i, p)$, $1 \le i \le k$, then $\sum_{i=1}^k X_i \sim \text{Bin}\big(\sum_{i=1}^k n_i, p\big)$.

 (d) If $X_i \sim \text{Gamma}(\alpha_i, \beta)$, $1 \le i \le k$, then $\sum_{i=1}^k X_i \sim \text{Gamma}\big(\sum_{i=1}^k \alpha_i, \beta\big)$.

 Write down proofs of these results.

20. Recall the following result from class: If $X \sim N(\mu,\sigma^2)$, then for any $b \in \mathbb{R}$ and $a \neq 0$, $aX + b \sim N(a\mu + b, b^2\sigma^2)$.

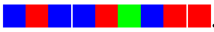
 (a) Prove this using mgfs.

 (b) Prove this by computing the pdf of $aX + b$ directly.

21. Show that $1 - \Phi(x) = \Phi(-x)$ for any $x \in \mathbb{R}$.

22. Measuring amount of money in thousands, the *gross annual salary* (i.e., salary before tax deductions) distributions in two towns A and B are respectively $N(25,16)$ and $N(30,9)$. The tax rates in town A and town B are respectively 30% and 20%.

 Find the probability that the sum of the *net salaries* (i.e., salary after tax deductions) of two individuals chosen independently from town A will be more than that of an individual chosen from town B. (Express your answer in terms of the standard Gaussian cdf $\Phi(\cdot)$.)

23. A total of $3n$ distinguishable balls are arranged in a straight line at random. Among the $3n$ balls, n are red, n are blue, and the rest are green. We say that $(i, i+1, i+2)$ is a color-block if the colors of the balls at the i-th, $i+1$-st and $i+2$-nd positions are different. Let Y_n denote the total number of color-blocks. For example, $Y_n = 3$ for the following arrangement: .

Clearly, $Y_n = \sum_{i=1}^{2n-2} X_i$, where $X_i = \mathbb{1}_{\{(i,i+1,i+2) \text{ is a color-block}\}}$.

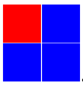
(a) Find $\mathbb{E}[X_i]$ for $1 \le i \le 2n-2$.

(b) Find $\mathbb{E}[X_i X_j]$ for $1 \le i < j \le 2n-2$.

(c) Find $\mathbb{E}[Y_n]$.

(d) Find $\mathrm{V}(Y_n)$.

24. Each of the 4 squares of a 2×2 grid is colored either red or blue independently with probability $1/2$ each. We say that two adjacent squares alternate colors if they have different colors. Let Y denote the total number of color alternating adjacent squares. For example, $Y = 2$ for the following configuration:

.

Express Y as a sum of 4 indicator random variables and use that to find

(a) $\mathbb{E}[Y]$, and

(b) $\mathrm{Var}(Y)$.

25. A factory manufactures light bulbs. The number of defective light bulbs produced in any given day follows a Poi(2) distribution. Assume that the operating conditions are independent on different days.

(a) Write down the exact distribution of the number of defective bulbs produced in a year (365 days).

(b) Using normal approximation, find the probability that at least 800 defective bulbs will be produced in a year. (Express your answer in terms of the standard Gaussian cdf $\Phi(\cdot)$.)

26. The lifetime of every light bulb in a lot of $10,000$ bulbs follows an exponential distribution with an average lifetime of four years. A bulb is considered defective if it burns out in its first year of use. Assume that the lifetimes of different bulbs are independent.

(a) Write down the exact distribution of the number of defective bulbs in the lot.

(b) Using normal approximation, find the probability that at least 2000 bulbs in the lot are defective. (Express your answer in terms of the standard Gaussian cdf $\Phi(\cdot)$.)

27. Consider a collection of 10^6 subatomic particles, each of which can be in one of six possible states–$S_1, \dots, S_6$. The probability that a given particle is in state S_i is $i/21$, $1 \le i \le 6$. Assume that the particles behave independently. For any $A \subset \{1, \dots, 6\}$, let X_A be the number of particles that are in state S_i for some $i \in A$.

   (a) Write down the joint pmf of $(X_{\{1\}}, \dots, X_{\{6\}})$.

   (b) Write down the pmf of $X_{\{1,3\}}$.

   (c) Write down the joint pmf of $(X_{\{4,6\}}, X_{\{1,2\}})$.

   (d) Using normal approximation, find the probability that at least 19×10^4 many particles are either in state S_1 or in state S_3. (Express your answer in terms of the standard Gaussian cdf $\Phi(\cdot)$.)

28. A random walker moves on the integer lattice $\mathbb{Z}$ in the following way: He starts his walk from the origin, and at each step, he moves one step to the right (resp. left) with probability $1/2$. The moves at different steps are independent. Let S_n denote the location of the walker after the n-th step, $n \ge 0$. (Thus, $S_0 = 0$, and S_1 is either $+1$ or -1 etc.)

   Use the CLT to show that for any $\delta, \varepsilon > 0$ small and $K > 0$ large,

$$\lim_{n\to\infty} \mathbb{P}\Big(\big|S_n\big| \ge \varepsilon n^{1/2+\delta}\Big) = 0, \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}\Big(\big|S_n\big| \le K n^{1/2-\delta}\Big) = 0.$$

   (We say that the distance of the random walker from the origin is typically of the order of $\sqrt{n}$ after n steps.)

29. Suppose $X \sim \text{Unif}[0,1]$. Let $X = .X_1X_2\dots$ be the binary expansion of X, i.e., $X_i = 0$ or 1 and

$$X = \sum_{i=1}^{\infty} \frac{X_i}{2^i}.$$

   Find the distribution of X_1.

30. Suppose $X_1 \perp\!\!\!\perp X_2$ and X_i has density f_i, $i = 1, 2$.

   Show that $X_1 + X_2$ is a continuous random variable. (Recall that by our definition, a random variable is continuous iff its cdf is a continuous function. Hence the above claim is equivalent to the claim that the cdf of $X_1 + X_2$ is continuous, i.e., $\mathbb{P}\big(X_1 + X_2 = y\big) = 0$ for all $y \in \mathbb{R}$.)

   Note: In fact, $X_1 + X_2$ has a density given by

$$f(y) := \int_{-\infty}^{\infty} f_1(u) f_2(y-u)\,du.$$

31. A particle's location in the plane is a random vector (X, Y), where X, Y are i.i.d. $N(0,1)$. Let R denote the distance of the particle from the origin.

   Find the density of R.

32. X and Y have a joint density $f(\cdot,\cdot)$ (resp. joint pmf $p(\cdot,\cdot)$) that satisfies

$$f(x,y) = f(y,x) \ \text{(resp. } p(x,y) = p(y,x)) \quad \text{for all} \quad x,y \in \mathbb{R}.$$

Show that $(X,Y) \stackrel{\mathrm{d}}{=} (Y,X)$, i.e., the random vectors (X,Y) and (Y,X) have the same joint distribution.

(Recall that two random vectors $\boldsymbol{X} = (X_1,\ldots,X_k)$ and $\boldsymbol{X}' = (X_1',\ldots,X_k')$ have the same joint distribution iff they have the same joint cdf, or equivalently if

$$\mathbb{P}\big(X_1 \le x_1,\ldots,X_k \le x_k\big) = \mathbb{P}\big(X_1' \le x_1,\ldots,X_k' \le x_k\big)$$

for all $x_1,\ldots,x_k \in \mathbb{R}$.)

33. Recall from the lectures that the sgn function is defined as

$$\operatorname{sgn}(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

(a) Suppose $X_1 \sim N(3,1)$, $X_2 \sim N(-2,4)$, and $X_3 \sim N(4,5)$. Further X_1, X_2, X_3 are independent.

Find the distribution of $\operatorname{sgn}\big(3X_1 + 2X_2 - X_3\big)$. (Express your answer in terms of the standard Gaussian cdf $\Phi(\cdot)$.)

(b) Suppose X and Y are i.i.d. discrete real-valued random variables with pmf $p(\cdot)$.

Find the distribution of $\operatorname{sgn}\big(X - Y\big)$.

(c) Suppose X and Y are i.i.d. real-valued continuous random variables having densities.

Find the distribution of $\operatorname{sgn}\big(X - Y\big)$.

34. Measuring time in microseconds, the lifetime of a class of subatomic particles has an exponential distribution with mean 1 time-unit. Each particle can be classified into one of four categories depending on their lifetime as follows:

(i) type I: if lifetime is less than 0.5 unit;
(ii) type II: if lifetime is between 0.5 and 1 unit;
(iii) type III: if lifetime is between 1 and 1.5 unit;
(iv) type IV: if lifetime is more than 1.5 unit.

Eight such particles are observed in a laboratory. Assume that the lifetimes of different particles are independent.

(a) Find the probability that 3 of them are of type I, and 2 of them are of type III.

(b) Find the probability that at least 2 of them are either of type I or of type II.

35. Suppose $X_1, X_2, \ldots$ are i.i.d. random variables with $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1^2] = 1$, and $a_1, \ldots, a_k$ are real numbers. Consider the time series $Y_t,\ t = 1, 2, 3, \ldots$, given by

$$Y_t := \sum_{j=1}^{k} a_j X_{t+j-1}.$$

Find the autocovariance function $\gamma(s,t) := \mathrm{Cov}(X_s, X_t)$ for $1 \le s \le t$.

"You cannot get an unfamiliar problem in the MATH 323 exam

if you are familiar with everything covered in class."

– William Wordsworth, 1801