CS698C: Sketching and Sampling for Big Data Analysis

Review Probability Exercises 1

1 Basic definitions

1. If the sample space is $\mathcal{C} = C_1 \cup C_2$ and if $P(C_1) = 0.8$ and $P(C_2) = 0.5$, find $P(C_1 \cap C_2)$.

2. If the sample space is $\mathcal{C} = \{c : -\infty < c < \infty\}$ and if $C \subset \mathcal{C}$ is a set for which the integral $\int_C \alpha e^{-|x|} dx$ exists, find the value of α for which the integrand becomes a probability set function.

3. Show that $P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2)$.

4. Show the inclusion-exclusion formula for $P(C_1 \cup C_2 \cup C_3)$.

5. A person has purchased 10 of 1000 tickets sold in a certain raffle. To determine the five prize winners, five tickets are to be drawn at random and without replacement. Calculate the probability that this person wins at least one prize. *Hint:* First calculate the probability that this person does not win a prize.

6. There are five red chips and three blue chips in a bowl. The red chips are numbered 1,2,3,4,5 respectively and the blue chips are numbered 1,2,3, respectively. If two chips are drawn at random and without replacement, find the probability that these chips either have the same number or the same color.

7. In a lot of 50 light bulbs, there are 2 bad bulbs. An inspector examines five bulbs, which are selected at random and without replacement.

   (a) Find the probability of at least one defective bulb among the five.

   (b) How many bulbs should be examined so that the probability of finding at least one bad bulb exceeds $\frac{1}{2}$?

8. Three plants C_1, C_2 and C_3 respectively produce 10%,40% and 50% of a company's output. C_1 produces 1% of its products as defective, C_2 has 3% of its products defective and C_3 has 4% of its products defective. One item that is selected at random from the company's output is observed to be defective. What is the posterior probability that this product came from plant C_1? What is the probability that a random selected output from the company is found defective.

9. Bowl 1 contains 6 red chips and 4 blue chips. Five of these 10 chips are selected at random and without replacement and put in bowl 2, which was originally empty. One chip is drawn at random from bowl 2. Given that this chip is blue, find the conditional probability that two red chips and three blue chips are transferred from bowl 1 to bowl 2.

10. Let C_1 and C_2 be independent events. Show that the following pairs of events are also independent: (a) C_1 and C_2^c, (b) C_1^c and C_2 and (c) C_1^c and C_2^c. *Hint:* In (a) write $P(C_1 \cap C_2^c) = P(C_2^c \mid C_1)P(C_1) = P(C_1)(1 - P(C_2 \mid C_1)) = P(C_1)(1 - P(C_2)) = P(C_1)P(C_2^c)$.

11. Say that $C_1, C_2, \ldots, C_k$ are independent events that have respective probabilities $p_1, p_2, \ldots, p_k$. Argue that the probability of at least one of $C_1, C_2, \ldots, C_k$ occurs is equal to $1-(1-p_1)(1-p_2)\ldots(1-p_k)$. *Hint:* Show and use the fact that $(C_1 \cup C_2 \cup \cdots \cup C_k)^c = C_1^c \cap C_2^c \cap \ldots C_k^c$.

12. A fair coin is tossed independently until the first head appears. What is the probability that the coin is tossed k times?

13. A (six-faced normal) die is cast independently until the first 6 appears. If the casting stops an odd number of times then Antu wins, otherwise Bantu wins.

    (a) Assuming the die is fair, what is the probability that Antu wins?

    (b) Let p denote the probability of a 6. Show that the game favours Antu, for all p, $0 < p < 1$.

14. Each bag in a large box contains 25 tulip bulbs. It is known that 60% of the bags contain bulbs for 5 red and 20 yellow tulips, while the remaining 40% of the bags contain bulbs for 15 red and 10 yellow tulips. A bag is selected at random and bulb taken at random from this bag is planted.

    (a) What is the probability that it will be a yellow tulip?

    (b) Given that it is yellow, what is the conditional probability it comes from a bag that contained 5 red and 20 yellow bulbs.

15. **Polya's Urn problem: Let an urn (bin) contain $B_0 \geq 1$ black balls and $W_0 \geq 1$ white balls. Draw a ball from the urn uniformly at random and then return the ball back to the urn, together with one additional ball of the same color. Let (B_n, W_n) denote the number of black and white balls in the urn after the nth draw-and-replacement.

    (a) Find the distribution of (B_n, W_n) after n draw-replacement steps.

16. *Binomial distribution.* Consider a biased coin with $P(H) = p$ and $P(T) = 1-p$. Consider an experiment where it is tossed n times. Let X be the random variable that counts the number of heads in the outcome of the experiment, $0 \leq X \leq n$. Derive the pmf for X, $p_X(x)$, $0 \leq x \leq n$.

17. *Cauchy Distribution.* The pdf $f_X(x) = \frac{1}{\pi(1+x^2)}$, $-\infty < x < \infty$ defines the Cauchy distribution. Show that this is indeed a probability density function by showing that $\int_{-\infty}^{\infty} f_X(x)dx = 1$. Another way to define this distribution is as follows. Define a random variable X to have a uniform distribution in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, so that $f_X(x) = \frac{2}{\pi}$, for x in this interval, and zero otherwise. Define $Y = \tan(X)$. Y has a Cauchy distribution.

18. A *median* of a distribution of a random variable X of the discrete or continuous type is a value of X such that $P(X < x) < \frac{1}{2}$ and $P(X \leq x) \geq \frac{1}{2}$. Find the median of the following distributions.

    (a) $p(x) = \binom{4}{x}(\frac{1}{4})^x(\frac{3}{4})^{4-x}$, $x = 0, 1, 2, 3, 4$, zero elsewhere.

    (b) $f_X(x) = \frac{2}{\pi(1+x^2)}$, $0 \leq x < \infty$.

2 Expectation, Variance and Moments

1. If the variance of a random variable X exists, show that $\mathrm{E}\left[X^2\right] \geq (\mathrm{E}\,[X])^2$.

2. Let the random variable X have mean μ, standard deviation σ and mgf $M(t)$, for $-h < t < h$. Show that

$$\mathrm{E}\left[\frac{X-\mu}{\sigma}\right] = 0, \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^2\right] = 1, \text{ and}$$
$$\mathrm{E}\left[\exp\left\{t\left(\frac{X-\mu}{\sigma}\right)\right\}\right] = \exp\left\{-\frac{\mu t}{\sigma}\right\} M\left(\frac{t}{\sigma}\right), -h\sigma < t < h\sigma \ .$$

3. Let X be a random variable such that $\mathrm{E}\left[(X-b)^2\right]$ exists for all real b. Show that $\mathrm{E}\left[(X-b)^2\right]$ is a minimum when $b = \mathrm{E}\,[X]$.

4. Let X have the mgf $M(t) = e^{t^2/2}$, $-\infty < t < \infty$. Using Taylor's series expansion of $M(t)$ around 0, show that

$$\mathrm{E}\left[X^{2k}\right] = (2k-1)(2k-3)\cdots(3)(1) = \frac{(2k)!}{2^k k!}, k = 1, 2, 3, \ldots$$
$$\mathrm{E}\left[X^{2k-1}\right] = 0, k = 1, 2, 3, \ldots$$

5. Let X be a random variable with mgf $M(t)$, $-h < t < h$ Prove that

$$P(X \geq a) \leq e^{-at} M(t), \quad 0 < t < h$$

and that

$$P(X \leq a) \leq e^{-at} M(t), \quad -h < t < 0 \ .$$

Hint: Use exponential moments method.

6. The mgf of X exists for all real values of t and is given by

$$M(t) = \frac{e^t - e^{-t}}{t}, t \neq 0, M(0) = 1 \ .$$

Use the results of the preceding exercise to show that $P(X \geq 1) = 0$ and $P(X \leq -1) = 0$. Note that here h is infinite.

7. Let X be a positive random variable $P(X \leq 0) = 0$. Argue the following. Note that (i) and (iii) are not necessarily true for general random variables.
   (a) $\mathrm{E}\,[1/X] \geq 1/\mathrm{E}\,[X]$.
   (b) $\mathrm{E}\,[-\log X] \geq -\log(\mathrm{E}\,[X])$.
   (c) For $q > p \geq 1$,
   $$\left(\mathrm{E}\,[X^q]\right)^{1/q} \geq \left(\mathrm{E}\,[X^p]\right)^{1/p}$$

8. **Let X be a general random variable (discrete or continuous) with $\mathrm{E}\,[X] = \mu$. [1]INSERT

[1]This will also be done in the course.

(a) Show that $\phi(t)$ defined as $\phi(t) = e^{-t}$, for $-\infty < t < \infty$ is a convex function. (Same is true for $\phi(t) = e^t$ and for $\phi(t) = e^{-\lambda t}$, for any $\lambda \in \mathbb{R}$.)

(b) Hence show that $\mathrm{E}\left[e^{-\lambda X}\right] \geq e^{-\lambda\mu}$, for any $-\infty < \lambda < \infty$.

(c) Show the following.

$$\mathbf{E}_X\left[e^{\lambda(X-\mu)}\right] = e^{-\lambda\mu}\mathbf{E}_X\left[e^{\lambda X}\right] \leq \mathbf{E}_{X'}\left[e^{-\lambda X'}\right]\mathbf{E}_X\left[e^{\lambda X}\right] = \mathbf{E}_{X,X'}\left[e^{\lambda(X-X')}\right]$$

where X' is an independent copy of the random variable X. (i.e., X and X' are independent but identically distributed). $X - X'$ has a symmetric distribution. (Show)

(d) *Let r be a Rademacher random variable with support $\{-1, 1\}$. That is, $\Pr[r = 1] = \Pr[r = -1] = \frac{1}{2}$.

$$\mathbf{E}_r\left[e^{\lambda r}\right] = \frac{e^\lambda + e^{-\lambda}}{2} \leq e^{\lambda^2/2}, \qquad \text{for } -\infty < \lambda < \infty.$$

(e) For any random variable X and an independent copy of X, show that $r(X - X')$ has the same distribution as $X - X'$. Using parts 3 and 5, show that

$$\begin{aligned}\mathrm{E}\left[e^{\lambda(X-\mu)}\right] \leq \mathbf{E}_{X,X'}\left[e^{\lambda(X-X')}\right] = \mathbf{E}_{X,X',r}\left[e^{\lambda r(X-X')}\right] &= \mathbf{E}_{X,X'}\left[\mathbf{E}_r\left[e^{\lambda r(X-X')}\right]\right] \\ &\leq \mathbf{E}_{X,X'}\left[e^{\lambda^2(X-X')^2/2}\right]\end{aligned}$$

(f) Suppose X has support only in the interval $[a, b]$. Hence, $(X - X')^2 \leq (b - a)^2$. This gives,

$$\mathrm{E}\left[e^{\lambda(X-\mu)}\right] \leq e^{\lambda^2(b-a)^2/2} \quad .$$

3 Random vector variables

1. Show that the function $F(x, y)$ that is equal to 1 provided $x + y \geq 1$ and that is equal to 0 provided $x + y \leq -1$, cannot be a distribution function of two random variables. *Hint:* Find four numbers $a < b$, $c < d$ so that

$$\Pr[a < X \leq b, c < Y \leq d] = F(b, d) - F(a, d) - F(b, c) + F(a, b)$$

is less than 0.

2. Let $g(x)$ be a pdf with $\int_0^\infty g(x)dx = 1$. Show that

$$f(x_1, x_2) = \frac{2g(\sqrt{x_1^2 + x_2^2})}{\pi\sqrt{x_1^2 + x_2^2}}, 0 < x_1 < \infty, 0 < x_2 < \infty,$$

and zero elsewhere is a pdf of two continuous-type X_1 and X_2. *Hint:* Use polar coordinates.

3. * Let X_1, X_2 be two random variables with the joint pmf $p(x_1, x_2)$, $(x_1, x_2) \in \mathcal{S}$, where, $\mathcal{S}$ is the support of X_1, X_2. Let $Y = g(X_1, X_2)$ be a function such that

$$\sum\sum_{(x_1,x_2)\in\mathcal{S}} |g(x_1, x_2)|\, p(x_1, x_2) < \infty.$$

Show that

$$\mathrm{E}[Y] = \sum\sum_{(x_1,x_2)\in\mathcal{S}} g(x_1, x_2)p(x_1, x_2) \quad .$$

3.1 Conditional Probability, Expectation and Variance

1. *A practice problem.* Let $f(x_1, x_2) = 21x^2x_2^3$, $0 < x_1 < x_2 < 1$, zero elsewhere, be the joint pdf of X_1 and X_2.

   (a) Find the conditional mean and variance of X_1, given $X_2 = x_2$, $0 < x_2 < 1$.

   (b) Find the distribution of $Y = \mathrm{E}[X_1 \mid X_2]$.

   (c) Determine $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$ and compare them to $\mathrm{E}[X_1]$ and $\mathrm{Var}[X_1]$.

2. *Practice problem.* Suppose X_1 and X_2 are random variables of the discrete type which have the joint pmf $p(x_1, x_2) = (x_1 + 2x_2)/18$, $(x_1, x_2) = (1,1), (1,2), (2,1), (2,2)$, and zeros elsewhere. Determine the conditional mean and variance of X_2, given $X_1 = x_1$, for $x_1 = 1$ or 2. Calculate $\mathrm{E}[3X_1 - 2X_2]$.

3. Let X_1 and X_2 be two random variables such that the conditional distributions and means exist. Show that:

   (a) $\mathrm{E}[X_1 + X_2 \mid X_2] = \mathrm{E}[X_1 \mid X_2] + X_2$.

   (b) $\mathrm{E}[u(X_2) \mid X_2] = u(X_2)$ and $\mathrm{Var}[u(X_2) \mid X_2] = 0$ (there is no randomness for $u(X_2) \mid X_2$.

4. Let the joint pdf of X and Y be given by

$$f(x,y) = \begin{cases} \frac{2}{(1+x+y)^3} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Calculate the marginal pdf of X and the conditional pdf of Y, given $X = x$.

   (b) For a fixed $X = x$, compute $\mathrm{E}[1 + X + Y \mid X = x]$ and use the result to compute $\mathrm{E}[Y \mid X]$.

5. Choose a random point from the interval $(0,1)$ uniformly and let the random variable X_1 be equal to the number corresponding to this point. Then choose a point uniformly at random from the interval $(0, x_1)$ where, x_1 is the experimental value obtained for X_1. Let the random variable X_2 be equal to the number which corresponds to this point.

   (a) Make assumptions about the marginal pdf $f_{X_1}(x_1)$ and the conditional pdf $f_{X_2|X_1}(x_2 \mid x_1)$.

   (b) Calculate $\Pr[X_1 + X_2 \geq 1]$.

   (c) Find the conditional expectation $\mathrm{E}[X_1 \mid X_2 = x_2]$.

6. *A memoryless distribution.* Let $f(x)$ and $F(x)$ denote the pdf and cdf of the distribution of a random variable X. The conditional pdf of X, given $X > x_0$, where, x_0 is a fixed number, is defined as

$$f(x \mid X > x_0) = \frac{f(x)}{1 - F(x_0)}, \quad x > x_0 \ .$$

   This has application in a problem of time until death, conditional on survival until time x_0.

   (a) Show that $f(x \mid X > x_0)$ is a pdf.

(b) Let $f(x) = e^{-x}$, $0 < x < \infty$ and zero elsewhere. Calculate $\Pr[X > 2 \mid X > 1]$. Show that $P(X > k \mid X > l) = P(X > k - l)$.

(*This is often called the memory less property of this distribution, given that an equipment is functional till time l, the probability that it would survive for another $k-l$ years equals the probability that the equipment would have survived $k-l$ years from the beginning. Memoryless $\equiv$ it doesn't matter how long until beginning has elapsed. Exponential distribution is often used to model the waiting time between two successive calls at a telecom switch.*).

3.2 Covariance and Correlation coefficient

1. Let X and Y be random variables with joint pdf $f(x, y)$. Show that if $Y = a + bX$ for $b \geq 0$ then, $\rho = 1$ and if $Y = a - bX$ for $b \geq 0$ then $\rho = -1$, irrespective of the value of a. (**Is it possible to have $\rho = 1$ when Y does not satisfy this property.)

2. Let $\sigma_1^2 = \sigma_2^2 = \sigma^2$ be the common variance of X_1 and X_2 and let ρ be the correlation coefficient of X_1 and X_2. Show that for $k > 0$,

$$\Pr\left[|(X - \mu_1) + (X_2 - \mu_2)| \geq k\sigma\right] \leq \frac{2(1+\rho)}{k^2} \quad .$$

3. **(*Log moment function.*) Let $\psi(t_1, t_2) = \log M(t_1, t_2)$, where, $M(t_1, t_2)$ is the mgf of X and Y. Show that

$$\left.\frac{\partial \psi(t_1, t_2)}{\partial t_i}\right|_{(t_1,t_2)=(0,0)}, \qquad \left.\frac{\partial^2 \psi(t_1, t_2)}{\partial t_i^2}\right|_{(t_1,t_2)=(0,0)}, \quad i = 1, 2,$$

and

$$\left.\frac{\partial^2 \psi(t_1, t_2)}{\partial t_1 \partial t_2}\right|_{(t_1,t_2)=(0,0)}$$

yield the expectations, the variances and the covariance of the two random variables.

3.3 Independent random variables

1. Let X and Y be random variables with pdf defined below. Argue in each case whether they are independent or not.

(a)

$$f(x, y) = \begin{cases} 6xy^2 & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(b)

$$f(x, y) = \begin{cases} 12xy^2 & 0 < x < 1, 0 < y < x \\ 0 & \text{otherwise.} \end{cases}$$

(c)

$$f(x, y) = \begin{cases} \frac{1}{\pi} & \left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2 \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

(d)

$$f(x,y) = \begin{cases} \frac{4}{\pi} & (x-1)^2 + (y+2)^2 < \frac{1}{4} \\ 0 & \text{otherwise.} \end{cases}$$

Find $f_X(x)$ and $f_Y(y)$.

2. Find the probability of the union of the events $a < X_1 < b, -\infty < X_2 < \infty$ and $-\infty < X_1 < \infty, c < X_2 < d$ if X_1 and X_2 are two independent variables with $\Pr[a < X_1 < b] = \frac{2}{3}$ and $\Pr[c < X_2 < d] = \frac{5}{8}$. Where is independence needed?

3. (Review.) Let X_1 and X_2 be random variables with joint pdf $f(x_1, x_2) = e^{-x_1 - x_2}$, $0 < x_1 < \infty, 0 < x_2 < \infty$ and zeros elsewhere. Show that X_1 and X_2 are independent and that $M(t_1, t_2) = (1-t_1)^{-1}(1-t_2)^{-1}$, $t_1 < 1, t_2 < 1$. Also show that

$$\mathrm{E}\left[e^{t(X_1+X_2)}\right] = \frac{1}{(1-t)^2}, \qquad t < 1 \ .$$

Find the mean and variance of $X_1 + X_2$.

4. Let X and Y be random variables with the support space consisting of four points $(0,0), (1,0), (1,1), (1,-1)$. Assign positive probabilities to these four points so that the correlation coefficient is equal to zero. Are X and Y independent?

5. Let X_1, X_2 be two discrete random variables with joint pmf $f(x_1, x_2) = (1/2)^{x_1+x_2}$, for $x_1, x_2 = 1, 2, \ldots, < infty$.

   (a) Show that they are independent.

   (b) Calculate the joint mgf $M(t_1, t_2)$ and note that it equals $M(t_1, 0)M(0, t_1)$.

   (c) Find the joint distribution of $X_1 + X_2$ and X_2.

   (d) Hence find the mean, variance and pmf of $X_1 + X_2$.

4 Multivariate Distributions

1. Let X, Y, Z have joint pdf $f(x,y,z) = \frac{2}{3}(x+y+z)$, $0 < x < 1, 0 < y < 1, 0 < z < 1$, and zero elsewhere.

   (a) Find the marginal pdfs of X, Y, Z, and their cdfs.

   (b) Are they independent?

   (c) Find the conditional distribution of X and Y given $Z = z$, and evaluate $\mathrm{E}[X + Y \mid Z = z]$.

   (d) Find the conditional distribution of X, given $Y = y$ and $Z = z$ and evaluate $\mathrm{E}[X \mid y, z]$.

2. Let X_1, X_2, X_3 and X_4 be four independent variables, each with pdf $f(x) = 3(1 - x^2)$, $0 < x < 1$ and zeros elsewhere. If $Y = \min(X_1, X_2, X_3, X_4)$, find the pdf and cdf of Y. Analogously, if $Z = \max(X_1, X_2, X_3, X_4)$, find the pdf and cdf of Y. *Hint:* Consider the events $Y > y$ and $Z < z$, respectively.

3. Let X_1, X_2, X_3 be iid with common pdf $f(x) = e^{-x}$, $x > 0$ and zero elsewhere. Evaluate $\Pr[X_1 < X_2 \mid X_1 < 2X_2]$ and $\Pr[X_1 < X_2 < X_3 \mid X_3 < 1]$.

4. Let X be a random vector with n variables $X_1, X_2, \ldots, X_n$. Show that $\mathrm{Cov}(X)$ is positive semi-definite, but not positive definite, iff the random variables are linearly related; that is, there exists $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and a scalar c such that $\Pr\left[a^T X + c = 0\right] = 1$. *Hint: $Cov(X)$ is positive semi-definite but not positive definite iff $Var\left[a^T X\right] = 0$, for some some $a \in \mathbb{R}^n$.*

5. Let X has the pdf $f(x) = \frac{1}{2}$, for $-1 < x < 1$ and zero elsewhere, find the pdf of $Y = X^2$.

6. If X has the pdf $f(x) = \frac{1}{4}$, for $-1 < x < 3$ and zero elsewhere, find the pdf of $Y = X^2$.

7. Let X_1, X_2 and X_3 be iid with common pdf $f(x) = e^{-x}, x > 0$ and zero for $x \leq 0$. Find the joint pdf of $Y_1 = X_1, Y_2 = X_1 + X_2, Y_3 = X_1 + X_2 + X_3$. Are these random variables mutually independent?

8. Let X_1, X_2, X_3 be iid with common pdf $f(x) = e^{-x}, x > 0$ and zero elsewhere. Find the joint pdf of $Y_1 = X_1/X_2, Y_2 = X_3/(X_1 + X_2), Y_3 = X_1 + X_2$. Are they mutually independent?

9. Let X_1, X_2, X_3, X_4 have the joint pdf $f(x_1, x_2, X_3, x_4) = 24$, $0 < x_1 < x_2 < x_3 < x_4 < 1$. Find the joint pdf of $Y_1 = X_1/X_2, Y_2 = X_2/X_3, Y_3 = X_3/X_4, Y_4 = X_4$ and show that they are mutually independent.

10. Let X_1, X_2, X_3 be iid with common mgf $M(t) = ((3/4 + (1/4)e^t)^2$, for all $t \in \mathbb{R}$.
    (a) Determine the probabilities, $\Pr[X_i = k]$, $k = 0, 1, 2$.
    (b) Find the mgf of $Y = X_1 + X_2 + X_3$ and then determine the probabilities, $P(Y_i = k)$, $k = 0, 1, \ldots, 6$.

11. Let $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$ and $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}$ be random variables. Recall that for any random k-dimensional vector Z, $\mathrm{Cov}(Z)$ is defined as $\mathrm{E}\left[(Z - \mathrm{E}[Z])(Z - \mathrm{E}[Z])^T\right]$. This definition is extended to coordinate wise covariance matrix of X with Y as follows:

$$\mathrm{Cov}(X, Y) = \mathrm{E}\left[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])^T\right]$$

Let Z be the $n + m$-dimensional vector $Z = \begin{bmatrix} X \\ Y \end{bmatrix}$. Show that,

(a)
$$\mathrm{Cov}(Z) = \begin{bmatrix} \mathrm{Cov}(X) & \mathrm{Cov}(X, Y) \\ \mathrm{Cov}(Y, X) & \mathrm{Cov}(Y) \end{bmatrix} .$$

(b) Given $p \times n$ matrix A of constant entries and $q \times m$ matrix B of constant entries, show that
$$\mathrm{Cov}(AX, BY) = A\mathrm{Cov}(X, Y)\, B^T.$$

(c) Define the $(p + q) \times (n + m)$ block matrix as follows.
$$D = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

(d) Write $\mathrm{Cov}(DZ)$ in terms of $\mathrm{Cov}(X), \mathrm{Cov}(Y)$ and $\mathrm{Cov}(X, Y)$.