(Fall 2022) 515502 Probability

Early Bird: 2022/12/8, 11am; Normal: 2022/12/9, 11am

Homework 3: Joint Distributions, Bivariate Normal, and MGF

Submission Guidelines: Please combine all your write-ups (photos/scanned copies are acceptable; please make sure that the electronic files are of good quality and reader-friendly) into one single .pdf file and submit the file via E3.

Problem 1 (Joint Distributions of Two Random Variables)

(6+6+6+6=24 points)

Let X, Y be two random variables with the joint CDF

$$F_{XY}(t,u) = \begin{cases} 1 - \exp(-t) - \exp(-u) + \exp(-(t+u+\theta t u)), & \text{if } t > 0, u > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\theta \in [0,1]$. Please try to derive the following properties of X and Y.

- (a) Find the marginal CDF of X, Y. For which values of θ (if any) are X, Y independent?
- (b) Find the joint PDF of X, Y.
- (c) Find the marginal PDF of both X and Y.
- (d) Could you explain why we require θ to be in [0,1]?

Problem 2 (Moment Generating Functions)

(12+10=22 points)

- (a) Let X be a continuous uniform random variable between -1 and 3. Find the MGF of X (denoted by $M_X(t)$) and use the derived $M_X(t)$ to find E[X] and Var[X]. (Hint: When evaluating the first-order and second-order derivatives of $M_X(t)$, you may need to leverage the L'Hôpital's rule)
- **(b)** Let Y be a discrete random variable with PMF

$$p_Y(k) = \begin{cases} \frac{6}{\pi^2 k^2}, & \text{if } k \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

Show that the MGF of Y (denoted by $M_Y(t)$) does NOT exist, i.e., there exists no interval of the form $(-\delta, \delta)$ (with $\delta > 0$) such that $M_Y(t)$ exists. (Hint: Show that $M_Y(t)$ is not finite on $t \in (0, \infty)$)

Problem 3 (Use MGFs to Find Distributions)

(8+8=16 points)

In the following subproblems, please use the MGFs to determine the distribution of random variables.

- (a) Let X_1 and X_2 be two random variables with $M_{X_1}(t) = \exp\left[7(e^t 1)\right]$ and $M_{X_2}(t) = \frac{(e^{2t} e^t)}{t}$. Find the distributions of X_1 and X_2 . (Note: Please clearly express the distributions in terms of either CDF/PDF/PMF)
- (b) Suppose X and Y are i.i.d. Poisson random variables with rate $\lambda = 1$ and observation window T = 1. Use MGFs to determine whether X + 2Y is also a Poisson random variable?

Problem 4 (Bivariate Normal)

(12+12=24 points)

(a) Let Z and W be two independent standard normal random variables. Let X_1 and X_2 be defined as

$$X_1 = \sigma_1 Z + \mu_1,$$

 $X_2 = \sigma_2 (\rho Z + \sqrt{1 - \rho^2} W) + \mu_2,$

where $\sigma_1, \sigma_2 > 0$, μ_1, μ_2 are finite real numbers, and $\rho \in (-1, 1)$. Show that the joint PDF of X_1, X_2 is bivariate normal, i.e., for all $x_1, x_2 \in \mathbb{R}$

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\bigg[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)} \bigg].$$

(Hint: Leverage the theorem of linear transformation of two random variables in Lecture 19)

(b) Let X, Y be bivariate normal, and the marginal distributions of X, Y are both $\mathcal{N}(0, 1)$. The correlation coefficient between X and Y is ρ . Define $Z_1 = X + Y$ and $Z_2 = X - Y$. Please show the following two things: (i) Show that Z_1 and Z_2 are also bivariate normal. (ii) Find the joint PDF of Z_1 and Z_2 without using calculus. (Hint: Leverage the covariance properties of bivariate normal)

Problem 5 (Conditional Expectation)

(8+8+8=24 points)

Suppose we consider a random experiment with a green die and an orange die as follows:

- Both dice are fair, 6-sided dice. A green die is rolled until it lands 1 for the first time. An orange die is rolled until it lands 6 for the first time.
- Let T_1 be the sum of the values of the rolls of the green die (including the 1 at the end) and T_6 be the sum of the values of the rolls of the orange die (including the 6 at the end).

Suppose Rafael and Novak are debating whether $E[T_1] = E[T_6]$ or $E[T_1] < E[T_6]$. Here are their arguments:

• Rafael's argument: We shall have $E[T_1] = E[T_6]$. By Law of Iterated Expectation (LIE), the expected sum of the rolls of a die is the expected number of rolls times the expected value of one roll, and each of these factors is the same for the two dice. In more detail, let N_1 be the number of rolls of the green die and N_6 be the number of rolls of the orange die. By LIE and linearity, we have

$$E[T_1] = E[E[T_1|N_1]] = E[3.5 \cdot N_1] = 3.5 \cdot E[N_1].$$

By a same argument, for the orange die, we also have $E[T_1] = 3.5 \cdot E[N_6]$, which equals $3.5 \cdot E[N_1]$.

- Novak's argument: Actually, we shall have $E[T_1] = E[T_6]$. While the expected number of rolls is indeed the same for the two dice, but the key difference is that we know the last roll is a 1 for the green die and a 6 for the orange die. The expected totals are the same for the two dice excluding the last roll of each, and then including the last roll makes $E[T_1] < E[T_6]$.
- (a) Do you agree to Rafael's argument? If no, what is the issue with Rafael's argument? Please clearly explain your thoughts.
- (b) Do you agree to Novak's argument? If no, what is the issue with Novak's argument? Please clearly explain your thoughts.
- (b) Give your own derivations of $E[T_1]$ and $E[T_6]$.