Probability and Statistics: Midterm Exam 2

Spring 2018

Name: ______________________________

This exam is scheduled for 75 minutes. Notes and other outside materials are not permitted. Non graphing calculators are allowed; if you do not have any, numerical formulas are enough.
Show all work to receive full credit, except where specified. The exam is worth 45 points.

Problem Number	Problem Points	Points Earned
MC	10	
TF	5	
FR1	10	
FR2	10	
FR3	10	
Total	45	

Multiple Choice

(2 points each) Circle the correct answer for each question. *You need not justify your answer, but one point of partial credit may be awarded.*

1 Consider a continuous random variable X with a standard normal distribution: $X \sim N(0,1)$. What is $P(X > 0.8)$?

Note: Tables for the cumulative distribution function Φ of $N(0,1)$ are provided at the end of the exam

(A) 0.53188

(B) 0.78814

(C) 0.46812

(D) 0.21186

(E) 0.42074

Tables

$$P(X > 0.8) = 1 - P(X \le 0.8) = 1 - \Phi(0.8) \overset{\downarrow}{=} 1 - 0.78814 = 0.21186$$

2 Consider a random variable X such that $\text{Var}(X) = 3$. We define $Y = 2X - 5$. What is $\text{Var}(Y)$?

(A) 7

(B) 12

(C) 6

(D) 1

(E) We need $E[X]$ to be able to compute the answer.

$$\text{Var}(Y) = 4\,\text{Var}(X) = 12$$

3 Consider two continuous random variables X and Y with joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} x+y & \text{for } 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $E[XY^2]$?

(A) $\frac{24}{59}$

(B) 7

(C) $\frac{1}{3}$

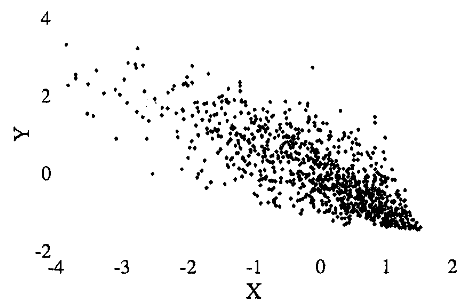
(D) $\frac{12}{57}$

(E) $\frac{17}{72}$

$$E[XY^2] = \int_0^1 \int_0^1 xy^2(x+y)\,dy\,dx = \int_0^1 \int_0^1 (x^2y^2 + xy^3)\,dy\,dx$$

$$= \int_0^1 \left[\frac{x^2}{3}y^3 + \frac{x}{4}y^4\right]_0^1 dx = \int_0^1 \left(\frac{x^2}{3} + \frac{x}{4}\right)dx = \left[\frac{x^3}{9} + \frac{x^2}{8}\right]_0^1 = \frac{1}{9} + \frac{1}{8} = \frac{17}{72}$$

4 The plot below was obtained by carrying out a simulation in which we generated 1000 realization of a pair of random variables (X,Y). What can you say about X and Y by looking at this plot?



(A) X and Y are not independent, but not correlated

(B) X and Y are positively correlated.

(C) X and Y are independent.

(D) X and Y are negatively correlated.

(E) The dataset is not sufficient to conclude about correlation and independence of X and Y.

5 Suppose that the number of calls per hour to the NYU IT help desk follows a Poisson process with rate $\lambda = 4$/hour. Suppose furthermore that 6 calls arrived in the first hour. What is the probability that there will be at most 1 call in the second hour?

(A) $5e^{-4}$

(B) $\frac{4^6}{6!}e^{-4}(1-5e^{-4})$

(C) $\frac{4^6}{6!}5e^{-8}$

(D) $1-5e^{-4}$

(E) $\frac{4^6}{6!}e^{-4}+5e^{-4}$

independence for intervals which do not overlap

$$P(N_{[1,2]} \le 1 \mid N_{[0,1]} = 6) = P(N_{[1,2]} \le 1)$$
$$= P(N_{[1,2]} = 0) + P(N_{[1,2]} = 1)$$
$$= e^{-4} + 4e^{-4} = 5e^{-4}$$

True or False

(1 point each) Indicate whether each statement is true or false. No partial credit will be given.

1 The covariance of two random variables is always a positive number.

(T) True ● False

Two random variables which are negatively correlated have a negative covariance.

2 If two random variables are independent, then they are uncorrelated.

● True (F) False

For independent random variables X and Y, $E[XY] = E[X]E[Y]$
$\Rightarrow Cov(X,Y) = E[XY] - E[X]E[Y] = 0$.

3 Consider two continuous random variables X and Y with joint cumulative distribution function

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{if } x < 0 \text{ or } y < 0 \\ \frac{1}{2}x^2y + \frac{1}{2}xy^3 & \text{if } 0 \le x, y \le 1 \\ \frac{1}{2}x^2 + \frac{1}{2}x & \text{if } 0 \le x \le 1 \text{ and } y \ge 1 \\ \frac{1}{2}y + \frac{1}{2}y^3 & \text{if } 0 \le y \le 1 \text{ and } x \ge 1 \\ 1 & \text{if } x > 1 \text{ and } y > 1 \end{cases}$$

The marginal cumulative distribution function F_X is given by

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

● True (F) False

$$F_X(x) = \lim_{y \to +\infty} F_{X,Y}(x,y) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2}x^2 + \frac{1}{2}x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } x > 1 \end{cases}$$

4 Two continuous random variables X and Y have the joint cumulative distribution function $F_{X,Y}$ given by

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-x})(1-e^{-y}) & \text{if } x \ge 0 \text{ and } y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X,Y} = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}(F_{X,Y})\right] = \frac{\partial}{\partial x}\left[(1-e^{-x})\frac{d}{dy}(1-e^{-y})\right]$$
$$= e^{-x}e^{-y}$$

X and Y thus have the following joint probability density function $f_{X,Y}$

$$f_{X,Y}(x,y) = \begin{cases} e^{-x} + e^{-y} & \text{if } x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(T) True ● False

5 The central limit theorem only applies to independent, identically distributed random variables which have finite mean.

● True (F) False

As we said in the lecture notes, the central limit theorem does not apply to random variables for which the mean does not exist or is infinite.

Free Response

Be sure to show all your work neatly and indicate your final answer where appropriate.

1 (10 points) Two discrete random variables X and Y have the following joint probability mass function, where a is a real number.

		X			
		0	1	2	
Y	-1	a	$2a$	a	$4a$
	0	0	a	a	$2a$
	1	$3a$	0	a	$4a$
		$4a$	$3a$	$3a$	

(A) (2 Points) What is the value of a?

$4a + 3a + 3a = 1 \Rightarrow a = \frac{1}{10}$

(B) (2 Points) What is the marginal probability mass function p_X of X? What is the marginal probability mass function p_Y of Y?

Reading the extra row and column in the table above, we have

$$p_X(x) = \begin{cases} \frac{2}{5} & \text{if } x = 0 \\ \frac{3}{10} & \text{if } x = 1 \\ \frac{3}{10} & \text{if } x = 2 \end{cases} \qquad p_Y(y) = \begin{cases} \frac{2}{5} & \text{if } y = -1 \\ \frac{1}{5} & \text{if } y = 0 \\ \frac{2}{5} & \text{if } y = 1 \end{cases}$$

(C) (1 Point) Compute $P(X = 1|Y = 0)$.

$P(X=1 | Y=0) = \frac{P(X=1, Y=0)}{P(Y=0)} = \frac{a}{2a} = \frac{1}{2}$

(D) (1 Point) Are X and Y independent random variables?

$P(X=1 | Y=0) = \frac{1}{2} \neq \frac{3}{10} = P(X=1)$ so X and Y are not independent.

(E) (2 points) What is the expected value of X? What is the expected value of Y?

$E[X] = 0 \cdot \frac{2}{5} + 1 \cdot \frac{3}{10} + 2 \cdot \frac{3}{10} = \frac{9}{10}$ $\qquad E[Y] = -1 \cdot \frac{2}{5} + 0 \cdot \frac{1}{5}$

(F) (2 Points) What is the covariance of X and Y?

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = E[XY]$$

XY takes values $-2, -1, 0, 1, 2$, with probability mass function

$$p_{XY}(x) = \begin{cases} \frac{1}{10} & \text{if } x = -2 \\ \frac{1}{5} & \text{if } x = -1 \\ \frac{3}{5} & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ \frac{1}{10} & \text{if } x = 2 \end{cases}$$

$$\Rightarrow E[XY] = \cancel{-2 \cdot \frac{1}{10}} + (-1) \cdot \frac{1}{5} + 0 \cdot \frac{3}{5} + 1 \cdot 0 + \cancel{2 \cdot \frac{1}{10}} = -\frac{1}{5}$$

Thus, $\mathrm{Cov}(X,Y) = -\frac{1}{5}$

2 (10 Points) You have invited 64 guests to a party. You need to make sandwiches for the guests. You believe that a guest might need 0, 1 or 2 sandwiches with probabilities $\frac{1}{4}, \frac{1}{2}$, and $\frac{1}{4}$ respectively. You assume that the number of sandwiches each guest needs is independent from other guests.

Use the central limit theorem to predict how many sandwiches you should make so that you are 95% sure that there is no shortage.

Note: Tables for the cumulative distribution function Φ of $N(0,1)$ are provided at the end of the exam

Let Y be the random variable for the total number of sandwiches the guests will need: $Y = \sum_{i=1}^{64} X_i$

with $E[X_i] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$, $Var(X_i) = E[X_i^2] - 1$
$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} - 1 = \frac{1}{2}$

We want to find the number s of sandwiches such that

$P(Y \leq s) \geq 0.95$

To use the central limit theorem, we shift and rescale Y:

$Z = \dfrac{Y - 64 \cdot 1}{\sqrt{64}\ \sigma_{X_i}} = \dfrac{\sqrt{2}}{8}(Y-64)$ has mean 0 and standard deviation 1

$$P(Y \leq s) \geq 0.95 \iff P\left(\frac{\sqrt{2}}{8}(Y-64) \leq \frac{\sqrt{2}}{8}(s-64)\right) \geq 0.95$$

$$\iff P\left(Z \leq \frac{\sqrt{2}}{8}(s-64)\right) \geq 0.95$$

$$\overset{\text{Tables for } \Phi}{\iff} \frac{\sqrt{2}}{8}(s-64) \geq 1.645$$

$$\iff s \geq 73.301$$

Thus, we need to make 74 sandwiches to be 95% sure that there is no shortage.

3 (10 points) Let X and Y be two independent random variables which are each uniformly distributed on the interval $(0,1)$. We define the random variables $W = \min(X,Y)$ and $Z = \max(X,Y)$. The purpose of this problem is to compute $\mathrm{Cov}(W,Z)$. We will do this through the following steps.

(A) (1 Point) What is the cumulative distribution function F_X of X? What is the cumulative distribution function F_Y of Y?

(B) (3 Points) What is the cumulative distribution function F_W of W?

Hint: Observe that because of the definition of W, for any real number w, $W > w$ if and only if $X > w$ and $Y > w$.

(C) (2 Points) What is the cumulative distribution function F_Z of Z?

Hint: Observe that because of the definition of Z, for any real number z, $Z \le z$ if and only if $X \le z$ and $Y \le z$.

(D) (4 Points) Compute $\mathrm{Cov}(W,Z)$.

(A) $f_X(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \Rightarrow F_X(x) = \begin{cases} 0 & \text{if } x \le 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$

Y is not different, so $F_Y(y) = \begin{cases} 0 & \text{if } y \le 0 \\ y & \text{if } 0 < y < 1 \\ 1 & \text{if } y \geq 1 \end{cases}$

(B) $\forall w \in \mathbb{R}$, $F_W(w) \overset{\text{Definition}}{=} P(W \le w) = 1 - P(W > w) = 1 - P(X > w, Y > w)$

$\overset{X \,\&\, Y \text{ are independent}}{=} 1 - P(X > w)P(Y > w)$

$= 1 - (1 - F_X(w))(1 - F_Y(w))$

Hence, $F_W(w) = \begin{cases} 0 & \text{if } w \le 0 \\ 1-(1-w)^2 = w(2-w) & \text{if } 0 < w < 1 \\ 1 & \text{if } w \geq 1 \end{cases}$

(C) $\forall z \in \mathbb{R}$, $F_Z(z) \overset{\text{Definition}}{=} P(Z \le z) = P(X \le z, Y \le z) \overset{X \,\&\, Y \text{ are independent}}{=} P(X \le z) P(Y \le z)$

$= F_X(z) F_Y(z)$

Hence $F_Z(z) = \begin{cases} 0 & \text{if } z \le 0 \\ z^2 & \text{if } 0 < z < 1 \\ 1 & \text{if } z \geq 1 \end{cases}$

(D) $Cov(W,Z) = E[WZ] - E[W]E[Z]$

Observe that since $W = \min(X,Y)$ and $Z = \max(X,Y)$,

$WZ = XY$

Hence, $E[WZ] = E[XY] = E[X]E[Y] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Now, from F_W and F_Z, we can compute

$$f_W(w) = \begin{cases} 0 & \text{if } w \leq 0 \\ 2-2w & \text{if } 0 < w < 1 \\ 0 & \text{if } w \geq 1 \end{cases}, \quad f_Z(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ 2z & \text{if } 0 < z < 1 \\ 0 & \text{if } z \geq 1 \end{cases}$$

Hence, $E[W] = \int_{-\infty}^{+\infty} w f_W(w)\,dw = \int_0^1 (2w - 2w^2)\,dw = \left[w^2 - \frac{2}{3}w^3\right]_0^1 = \frac{1}{3}$

$$E[Z] = \int_{-\infty}^{+\infty} z f_Z(z)\,dz = \int_0^1 2z^2\,dz = \left[\frac{2}{3}z^3\right]_0^1 = \frac{2}{3}$$

$$Cov(W,Z) = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$$