Name: ______________________________ Tutorial group: T1

Matriculation number: | | | | | | | | | |

NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER I 2016/17

MH2500– Probability and Introduction to Statistics

18 October 2016 Test 3 40 minutes

INSTRUCTIONS

1. Do not turn over the pages until you are told to do so.
2. Write down your name, tutorial group, and matriculation number.
3. This test paper contains **FOUR (4)** questions and comprises **FIVE (5)** printed pages.
4. Answer **all** questions. The marks for each question are indicated at the beginning of each question.
5. You are allowed three sides of an A4 size paper as cheat sheet.

For graders only	Question	1	2	3	4	Bonus	Total
	Marks						

QUESTION 1. **(8 marks)**

If X and Y have the joint density function

$$f(x,y) = \begin{cases} x + \frac{5}{2}y, & 0 \le x \le y \le 1; \\ 0, & \text{elsewhere.} \end{cases}$$

Let $Z = Y - X$, find $P(Z \ge \frac{1}{2})$. Leave your answer as a fraction or to three significant figures.

[Answer:]

- Method 1

  We first find the region R which is the intersection of $0 \le x \le y \le 1$ and $y - x \ge \frac{1}{2}$. R is a triangle formed by $(0, \frac{1}{2})$, $(0, 1)$, and $(\frac{1}{2}, 1)$. The probability $P(Z \ge \frac{1}{2})$ is the integration of the density function $f(x,y)$ over region R,

$$P(Z \ge \frac{1}{2}) = \iint_R f(x,y)dxdy = \int_{\frac{1}{2}}^{1} \int_0^{y-\frac{1}{2}} (x + \frac{5}{2}y)dxdy$$

$$= \int_{\frac{1}{2}}^{1} \left(\frac{x^2}{2} + \frac{5}{2}xy \right) \Big|_{x=0}^{x=y-\frac{1}{2}} dy = \int_{\frac{1}{2}}^{1} \left(3y^2 - \frac{7}{4}y + \frac{1}{8} \right) dy$$

$$= \left(y^3 - \frac{7}{8}y^2 + \frac{1}{8}y \right) \Big|_{y=\frac{1}{2}}^{y=1} = \frac{1}{4} - \frac{-1}{32} = \frac{9}{32} (= 0.28125)$$

- Method 2

  Then density function of Z is

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, x+z)dx = \begin{cases} \int_0^{1-z} \left[x + \frac{5}{2}(x+z) \right] dx, & 0 \le z \le 1; \\ 0, & \text{elsewhere.} \end{cases}$$

  further

$$f_Z(z) = \begin{cases} -\frac{3}{4}z^2 - z + \frac{7}{4}, & 0 \le z \le 1; \\ 0, & \text{elsewhere.} \end{cases}$$

  So that

$$P(Z \ge \frac{1}{2}) = \int_{\frac{1}{2}}^{1} [-\frac{3}{4}z^2 - z + \frac{7}{4}]dz = \frac{9}{32} (= 0.28125).$$

QUESTION 2. **(10 marks)**

If X and Y have the joint density function

$$f(x,y) = \begin{cases} x + 4y, & 0 \le y \le x \le 1; \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal densities of X and Y.

(b) Are X and Y independent?

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

(d) Find the joint cumulative distribution function of X and Y.

[Answer:]

a.

$$f_X(x) = \int_0^x (x+4y)dy = (xy + y^2)\Big|_{y=0}^{y=x} = 3x^2, \qquad 0 \le x \le 1$$

$$f_Y(y) = \int_y^1 (x+4y)dx = \left(\frac{1}{2}x^2 + 4xy\right)\Big|_{x=y}^{x=1} = -\frac{9}{2}y^2 + 4y + \frac{1}{2}, \qquad 0 \le y \le 1.$$

b. Since $f(x,y) = x + 4y \ne (3x^2) \times (-\frac{9}{2}y^2 + 4y + \frac{1}{2}) = f_X(x)f_Y(y)$, X and Y are not independent.

c.

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{x+4y}{3x^2}, \qquad 0 \le y \le x \le 1; \quad x \ne 0.$$

d. The joint CDF of X and Y is defined as $F(u,v) = P(x \le u, y \le v)$

- When $0 \le v \le u \le 1$,

$$\begin{aligned} F(u,v) &= \int_0^v \int_y^u (x+4y)dxdy = \int_0^v \left(\frac{1}{2}x^2 + 4xy\right)\Big|_{x=y}^{x=u} dy \\ &= \int_0^v \left(-\frac{9}{2}y^2 + 4uy + \frac{u^2}{2}\right) dy = \left[-\frac{3}{2}y^3 + 2uy^2 + \frac{u^2}{2}y\right]\Big|_{y=0}^{y=v} \\ &= -\frac{3}{2}v^3 + 2uv^2 + \frac{u^2v}{2}, \qquad 0 \le v \le u \le 1. \end{aligned}$$

  Changing the variables back to x and y, we have the CDF as

$$F(x,y) = -\frac{3}{2}y^3 + 2xy^2 + \frac{x^2y}{2}, \qquad 0 \le y \le x \le 1.$$

- When $0 \le u \le 1$ and $v \ge u$,

$$F(u,v) = \int_0^u \int_y^u (x+4y)dxdy = \int_0^u \left(\frac{1}{2}x^2 + 4xy\right)\Big|_{x=y}^{x=1} dy$$

$$= \int_0^u \left(-\frac{9}{2}y^2 + 4uy + \frac{u^2}{2}\right) dy = \left[-\frac{3}{2}y^3 + 2uy^2 + \frac{u^2}{2}y\right]\Big|_{y=0}^{y=u}$$

$$= u^3, \qquad 0 \le u \le 1.$$

Changing the variables back to x and y, we have the CDF as

$$F(x,y) = x^3, \qquad 0 \le x \le 1; y \ge x.$$

- When $0 \le v \le 1$ and $u \ge 1$

$$F(u,v) = \int_0^v \int_y^1 (x+4y)dxdy = \int_0^v \left(\frac{1}{2}x^2 + 4xy\right)\Big|_{x=y}^{x=1} dy$$

$$= \int_0^v \left(-\frac{9}{2}y^2 + 4y + \frac{1}{2}\right) dy = \left[-\frac{3}{2}y^3 + 2y^2 + \frac{1}{2}y\right]\Big|_{y=0}^{y=v}$$

$$= -\frac{3}{2}v^3 + 2v^2 + \frac{v}{2}, \qquad 0 \le v \le 1.$$

Changing the variables back to x and y, we have the CDF as

$$F(x,y) = -\frac{3}{2}y^3 + 2y^2 + \frac{y}{2}, \qquad 0 \le y \le 1; x \ge 1.$$

- When $v \ge 1$ and $u \ge 1$, $F(u,v) = 1$, so

$$F(x,y) = 1.$$

- When $v \le 0$ or $u \le 0$, $F(u,v) = 1$, then

$$F(x,y) = 0.$$

QUESTION 3. **(8 marks)**

If X_1, X_2, and X_3 are independent random variables, each with the same exponential density function $f(x)$,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Find the joint density of $X_{(1)}$ and $X_{(3)}$. Note that $X_{(1)} \leq X_{(2)} \leq X_{(3)}$.

[Answer:] Let $g(v, u)$ be the joint density function of $X_{(1)}$ and $X_{(3)}$. Using the differential argument, we can have

$$g(v, u)dvdu = \binom{3}{1} f(v)dv \binom{2}{1} f(u)du \binom{1}{1} [F(u) - F(v)],$$

and therefore $g(v, u) = 6f(v)f(u)[F(u) - F(v)]$. Note that $F(x)$ is the cumulative distribution function of the exponential random variables X_i's. $F(x)$ takes the following form,

$$F(x) = \int_0^x f(\tau)d\tau = \int_0^x \lambda e^{-\lambda\tau} d\tau = [-e^{-\lambda\tau}]|_{\tau=0}^{\tau=x} = 1 - e^{-\lambda x}, \quad x \geq 0.$$

Explicitly, $g(v, u)$ can be written as

$$g(v, u) = 6f(v)f(u)[F(u) - F(v)] = 6\lambda^2 e^{-\lambda(v+u)}[e^{-\lambda v} - e^{-\lambda u}], \quad 0 \leq v \leq u.$$

Equivalently,

$$g(v, u) = 6\lambda^2[e^{-\lambda(2v+u)} - e^{-\lambda(v+2u)}], \quad 0 \leq v \leq u.$$

QUESTION 4. **(8 marks)**

X is a binomial random variable with parameters n and p.

(a) Find E(X).

(b) Let $Y = X(X-1)$, find E(Y).

(c) Let $Z = X^2$, find E(Z). (Hint: Z=X+Y.)

(d) Find Var(X).

[Answer]

(a) The binomial distribution can be viewed as a linear combination of n independent Bernoulli distributions. The expected value of a Bernoulli random variable with parameter p is p. So $E(X) = np$.

We can also work with the definition of expectation. For a binomial random variable X with parameters n and p, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Hence

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}, \end{aligned}$$

and let $j = k-1$

$$\begin{aligned} E(X) &= \sum_{j=0}^{n-1} \frac{n!}{j!(n-1-j)!} p^{j+1} (1-p)^{n-1-j} \\ &= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} \end{aligned}$$

$\frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j}$ is the probability of a binomal random variable with parameters $n-1$ and p and j successes. So, the sum from 0 to $n-1$ is 1. This gives $E(X) = np$.

(b) If we once again view a binomal distribution as a linear combination of n Bernoulli random variables, Var(X) should be easily obtained as Var(X) $= np(1-p)$, given the fact that the variance of a Bernoulli random variable is $p(1-p)$. Further, $\mathrm{Var}(X) = E(X^2) - [E(X)]^2$, we have $E(X^2) = \mathrm{Var}(X) + [E(X)]^2 = np(1-p) + n^2p^2$. Therefore $E[X(X-1)] = E(X^2) - E(X) = np(1-p) + n^2p^2 - np = n(n-1)p^2$.

Alternatively, we can use the definition of expectation to find $E[X(X-1)]$ as follows,

$$
\begin{aligned}
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^k (1-p)^{n-k},
\end{aligned}
$$

and let $j = k-2$

$$
\begin{aligned}
E[X(X-1)] &= \sum_{j=0}^{n-2} \frac{n!}{j!(n-2-j)!} p^{j+2} (1-p)^{n-2-j} \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \frac{(n-2)!}{j!(n-2-j)!} p^j (1-p)^{n-2-j}
\end{aligned}
$$

$\frac{(n-2)!}{j!(n-2-j)!} p^j (1-p)^{n-2-j}$ is the probability of a binomal random variable with parameters $n-2$ and p and j successes. So, the sum from 0 to $n-2$ is 1. This gives $E[X(X-1)] = n(n-1)p^2$.

(c) If the binomial distribution is considered as a linear combination of n independent Bernoulli distrubtions, we have already solved all the problems of Question 4. Otherwise, $E(X^2) = E[X(X-1)+X] = E[X(X-1)] + E(X) = n(n-1)p^2 + np$.

(d) $\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$.