Math 376 LM (22081) Fall 2021

Test 1 Solutions

1- (a) Let X_1 and X_2 be independent -nonconstant- rv's with $E(X_1) = E(X_2) = 0$ and for $i = 1, 2$
$E(X_i^2) = A_i, \quad E(X_i^3) = B_i, \quad E(X_i^4) = C_i.$

Let $Z = X_1 + X_2$. Show that $E(Z^2) = A_1 + A_2$ and $E(Z^3) = B_1 + B_2$. Compute $E(Z^4)$ and show, in particular, that $E(Z^4) \neq C_1 + C_2$.

(b) If X_1 and X_2 are independent continuous rv's show that covariance $Cov(X_1, X_2)$ equals zero when it is defined and describe when it is not defined.

(c) Assume that X is a continuous rv with density function f satisfying $f(x) \begin{cases} > 0 \text{ if } |x| < 1, \\ = 0 \text{ otherwise.} \end{cases}$ and $f(-x) = f(x)$ for all x, i.e. f is an even function. Show that $Cov(X, X^2) = 0$ but X and X^2 are not independent (Hint; Consider the events $\{X < a\}$ and $\{X^2 < a^2\}$ with $0 < a < 1$).

(a)By the binomial theorem, or direct multiplication:

$$E(Z^2) = E(X_1^2 + 2X_1X_2 + X_2^2) = A_1 + 2E(X_1X_2) + A_2,$$

$$E(Z^3) = E(X_1^2 + 3X_1^2X_2 + 3X_1X_2^2 + X_2^3) = B_1 + 3E(X_1^2X_2) + 3E(X_1X_2^2) + B_2,$$

$$E(Z^4) = E(X_1^4 + 4X_1^3X_2 + 6X_1^2X_2^2 + 4X_1X_2^3 + X_2^4)$$
$$= C_1 + 4E(X_1^3X_2) + 6E(X_1^2X_2^2) + 4E(X_1X_2^3) + C_2.$$

By independence,

$$E(X_1X_2) = E(X_1)E(X_2) = 0, \quad E(X_1^2X_2) = E(X_1^2)E(X_2) = 0,$$
$$E(X_1X_2^2) = E(X_1)E(X_2^2) = 0,$$
$$E(X_1^3X_2) = E(X_1^3)E(X_2) = 0, \quad E(X_1X_2^3) = E(X_1)E(X_2^3) = 0,$$
$$E(X_1^2X_2^2) = E(X_1^2)E(X_2^2) = A_1A_2.$$

Because the rv's are nonconstant, $A_1, A_2 > 0$. Therefore,

$$E(Z^2) = A_1 + A_2, \quad E(Z^3) = B_1 + B_2,$$
$$E(Z^4) = C_1 + 6A_1A_2 + C_2 > C_1 + C_2.$$

(b) The covariance requires that the integral $\int\int |x_1x_2|f_{X_1X_2}(x_1,x_2)dx_1dx_2$ be finite. This will imply that that X_1 and X_2 have finite means. In the independent case, when the means are finite, $E(X_1X_2) = E(X_1)E(X_2)$ and so the covariance $E(X_1X_2) - E(X_1)E(X_2)$ is defined and equals zero.

(c) Because $xf(x)$ and $x^3f(x)$ are odd functions of x, $E(X) = E(X^3) = 0$. Hence,

$$Cov(X, X^2) = E(X^3) - E(X)E(X^2) = 0$$

.

If $X^2 < a^2$ then $X < a$. That is, the event $\{X^2 < a^2\}$ is contained in the event $\{X < a\}$. Hence,

$$P(X^2 < a^2 \text{ and } X < a) = P(X^2 < a^2) > P(X^2 < a^2) \cdot P(X < a).$$

Alternatively, $P(X < a|X^2 < a^2) = 1 \neq P(X < a)$.

2- An urn contains N balls of which 3 are red. The balls are drawn one at a time. Let X_1 be the first time a red ball is drawn.

(a) Assume the drawing is without replacement. What is the range of X_1 and what is its pmf? (Hint: it helps to think of the balls as numbered from 1 to N.)

(b) Assume instead that the drawing is with replacement. What is the range of X_1 and what is its pmf?

(a) The range of X_1 is $\{1, 2, ..., N - 2\}$.

The first $k - 1$ balls are all blue with probability $\binom{N-3}{k-1} \div \binom{N}{k-1}$. Assuming this, the probability of the k^{th} ball being red is $3 \div (N - (k - 1))$. So

$$P(X = k) = (N-3)!(N-(k-1))!3/(N-3-(k-1))!N!(N-(k-1)) = \frac{3(N-k)(N-k-1)}{N(N-1)(N-2)}.$$

Alternatively, $P(X = k) =$

$$\frac{(N-3)(N-3-1)\dots(N-3-(k-2))3}{N(N-1)\dots(N-k+1)} = \frac{3(N-k)(N-k-1)}{N(N-1)(N-2)}.$$

(b) With replacement $X_1 - 1 \sim Geom(p)$ with $p = 3/N$. So the range of X_1 consists of all positive integers and $P(X_1 = k) = (\frac{N-3}{N})^{k-1}\frac{3}{N}$.

3- Assume that $X \sim Exp(\lambda)$, that is, it is an exponential distribution with parameter λ. Define $Y = [X]$ so that Y is the largest integer less than or equal to X.

(a) Describe the range of the rv Y.

(b) Compute the cdf and pmf for the random variable Y and show $Y \sim Geom(p)$ for suitable p.

(c) Compute the expected value of Y.

(d) Show that the limit of $E(X) \div E(Y)$ as $\lambda \to 0$ is equal to 1.

(a) Since $X > 0$, the range of the discrete rv Y consists of the non-negative integers $n = 0, 1, \dots$.

(b) For $n = 0, 1, \dots,$

$$F_Y(n) = P(Y \leq n) = P(X < n+1) = \int_0^{n+1} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda(n+1)},$$

$$f_Y(n) = F_Y(n) - F_Y(n-1) = P(n \leq X < n+1) =$$

$$e^{-\lambda n} - e^{-\lambda(n+1)} = (1 - e^{-\lambda})e^{-\lambda n}.$$

This is $Geom(1 - e^{-\lambda})$, that is, geometric with $p = 1 - e^{-\lambda}, q = e^{-\lambda}$.

(c)

$$E(Y) = \sum_{n=0}^{\infty} n f_Y(n) = 0 \cdot (1 - e^{-\lambda 1}) +$$

$$1 \cdot (e^{-\lambda 1} - e^{-\lambda(2)}) + 2 \cdot (e^{-\lambda 2} - e^{-\lambda 3}) + 3 \cdot (e^{-\lambda 3} - e^{-\lambda 4}) + \dots$$

$$= \sum_{n=0}^{\infty} n(e^{-\lambda n} - e^{-\lambda(n+1)}) = \sum_{n=1}^{\infty} e^{-\lambda n} = \frac{e^{-\lambda}}{1 - e^{-\lambda}}.$$

Alternatively, since $Y \sim Geom(p)$, $E(Y) = q/p = 1/(e^{\lambda} - 1)$

(d) $E(X) = \frac{1}{\lambda}$. So

$$E(X) \div E(Y) = \frac{e^{\lambda} - 1}{\lambda}.$$

Use L'Ȟopital's Rule or observe that this is the difference quotient for $\frac{d}{dx} e^x|_{x=0} = 1$.

4- Let X, Y be rv's with joint density $f_{X,Y}(x, y) = C(x + y)$ for $0 < y < x < 1$ and $= 0$ otherwise.

(a) Determine the constant C so that this defines a joint pdf.

(b) Compute the marginal distributions for X and Y and determine whether X and Y are independent (justifying your answer).

(c) For $Z = Y/X$, compute the cdf and the pdf of Z.

(a), (b) $f_X(x) = C\int_0^x f_{X,Y}(x,y)dy = C\int_0^x x + ydy = C\frac{3x^2}{2})$ for $0 < x < 1$.

$$C\int_0^1\int_0^x x + ydydx = C\int_0^1 \frac{3x^2}{2}dx = \frac{C}{2}$$

and so $C = 2$.
$f_Y(y) = 2\int_y^1 f_{X,Y}(x,y)dx = 2\int_y^1 x+ydx = 1+2y-3y^2$ for $0 < y < 1$.
They are not independent because the joint density is not the product of the marginals.

(c) Z satisfies $0 < Z < 1$. For $0 < z < 1$,

$$F_Z(z) = P(Z \leq z) = 2\int_{x=0}^{x=1}\int_{y=0}^{y=zx}(x+y)dydx =$$
$$\int_{x=0}^{x=1} x^2(2z+z^2)dx = \frac{2z+z^2}{3}$$

Hence, the pdf $f_Z(z) = \frac{2}{3}(1+z)$. for $0 < z < 1$ and equal to 0 otherwise.

5- $X_1, \ldots, X_n$ is a sample with distribution with exponential distribution $Exp(\theta)$ and it is known that $\theta \leq 2$. Hence, the range of the parameter we wish to estimate is $0 < \theta \leq 2$. Compute the MLE for θ.

$f_X(x) = \theta e^{-\theta x}$ and so $L(\theta) = \theta^n exp(-\theta\sum_{i=1}^n x_i)$ and

$$\ell(\theta) = n\ln(\theta) - \theta\sum_{i=1}^n x_i$$
$$\frac{\partial\ell}{\partial\theta} = n\theta^{-1} - \sum_{i=1}^n x_i$$

So $\frac{\partial\ell}{\partial\theta} = 0$ when $\theta = n \div \sum_{i=1}^n x_i$ with $\frac{\partial^2\ell}{\partial\theta^2} = -n\theta^{-2} < 0$.
However, θ is restricted to interval $(0, 2]$. If $n \div \sum_{i=1}^n x_i \leq 2$ then this is the point at which maximum ℓ is maximum. If $n \div \sum_{i=1}^n /x_i > 2$ then $\frac{\partial\ell}{\partial\theta} > 0$ in the range $(0, 2]$ and the maximum value occurs occurs at the right end-point $\theta = 2$. Thus, the MLE is

$$\hat{\theta} = minimum(2, 1/\bar{X}).$$

6- Let X be an rv with density

$$f(x,\theta)\begin{cases} = & \frac{2}{\theta^2}x \text{ for } \ 0 < x < \theta, \\ = & 0 \ \text{ otherwise} \end{cases} \quad .$$

$X_1, \ldots, X_n$ is an iid sample with distribution that of X.

(a) Compute the MLE $\hat{\theta}$ for θ.

(b) Compute the density function and expected value for $Z = \hat{\theta}$. In particular, determine whether $\hat{\theta}$ is unbiased.

(a)

$$L(\theta) = (2^n\theta^{-2n})\prod_i x_i I_{[0,\theta]}(x_i) = (2^n\theta^{-2n})I_{[0,\theta]}(max(x_1, ..., x_n))\prod_i x_i.$$

Since $2^n\theta^{-2n}$ is decreasing in θ, the MLE $\hat{\theta} = Z = \max(X_1, \ldots, X_n)$.

(b) The range of Z is $(0, \theta)$ and so the range of Z is $(0, 1)$. $F_X(x) = x^2/\theta^2$ for $0 < x < \theta$. So $F_Z(z) = z^{2n}/\theta^{2n}$ for $0 < z < \theta$. The density $f_Z(z) = 2nz^{2n-1}/\theta^{2n}$ and the expected value is $\frac{2n}{2n+1}$. Hence the MLE is biased.