APMA 1650 - Spring 2021

Midterm Exam 2 - Solutions

Problem 2. **(Ever Given)** (20pts) In a given year, the proportion of time T that a particular canal is blocked has a probability density function

$$f_T(t) = \begin{cases} 2(1-t), & 0 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

The global economic cost of this blockage (in billions of dollars) is given by

$$C = 10T + 4T^2.$$

a. (4 pts) Find the mean of C.

b. (8 pts) Find the variance of C.

c. (8 pts) Use Chebyshev to find a value b such $P(C \leq b)$ is at least 0.75. (State your answer to two decimal places)

Solution:

a. The mean of C is given by

$$EC = E[10T + 4T^2] = 10ET + 4E[T^2].$$

Computing ET and $E[T^2]$ via LOTUS, we obtain

$$ET = \int_0^1 t2(1-t)\mathrm{d}t = 2\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{3}$$

and

$$E[T^2] = \int_0^1 t^2 2(1-t)\mathrm{d}t = 2\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{1}{6}$$

Therefore

$$EC = 10\left(\frac{1}{3}\right) + 4\left(\frac{1}{6}\right) = 4$$

b. To calculate the variance, we see that

$$\begin{aligned}
\mathrm{Var}(C) &= EC^2 - (EC)^2 \\
&= E(10T + 4T^2)^2 - 16 = E[100T^2 + 80T^3 + 16T^4] - 16 \\
&= 100E[T^2] + 80E[T^3] + 16E[T^4] - 16
\end{aligned}$$

and compute $E[T^3]$ and $E[T^4]$ via LOTUS

$$E[T^3] = \int_0^1 t^3 2(1-t)\mathrm{d}t = 2\left(\frac{1}{4} - \frac{1}{5}\right) = \frac{1}{10}$$

and

$$E[T^4] = \int_0^1 t^4 2(1-t)\mathrm{d}t = 2\left(\frac{1}{5} - \frac{1}{6}\right) = \frac{1}{15}$$

Therefore

$$\mathrm{Var}(C) = \frac{100}{6} + \frac{80}{10} + \frac{16}{15} - 16 = \frac{146}{15}$$

c. To apply Chebyshev we note that for $b \geq 4$

$$P(C \leq b) = P(C - 4 \leq b - 4) \geq P(|C - 4| \leq b - 4)$$

and therefore

$$P(C \leq b) \geq P(|C - 4| \leq b - 4) \geq 1 - \frac{\left(\frac{146}{15}\right)}{(b-4)^2}$$

we see that be get a bound of at least 0.75

$$1 - \frac{\left(\frac{146}{15}\right)}{(b-4)^2} \geq 0.75 \quad \Leftrightarrow \quad b \geq 4 + 2\sqrt{\frac{146}{15}} \approx 10.24$$

(Note we rounded up so that the lower bound is still valid)

Problem 3. (Quality Balls) (20 pts) A quality control process for a basketball manufacturing plant samples 10 finished basketballs a day and counts X the number of defective basketballs in that sample. The probability p that a given basketball is defective varies from day to day and is assumed to have a uniform distribution on $[0, 1/4]$.

a. (6 pts) For a given p, what is the conditional distribution of X? (You can give a name and appropriate parameters)

b. (6 pts) Find EX.

c. (8 pts) Find $\mathrm{Var}(X)$.

Solution:

a. For a given p, X is Binomial$(10, p)$

b. Using the law of iterated expectation and the fact that Binomial$(10, p)$ has mean $10p$

$$EX = E[E[X \mid p]] = E[10p] = \frac{10}{8} = \frac{5}{4}$$

since the mean of a Uniform$(0, 1/4)$ random variable is $1/8$.

c. To calculate the variance, we use the law of total variance and the fact that Binomial$(10, p)$ has mean $10p$ and variance $10p(1-p)$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Var}(E[X \,|\, p\,]) + E[\mathrm{Var}(X \,|\, p\,)] \\ &= \mathrm{Var}(10p) + E[10p(1-p)] \\ &= 100\mathrm{Var}(p) + 10E[p(1-p)]\end{aligned}$$

Note that for a uniform random variable $p \sim \mathrm{Uniform}(0, 1/4)$

$$\mathrm{Var}(p) = \frac{\left(\frac{1}{4}\right)^2}{12} = \frac{1}{192}$$

and

$$E[p(1-p)] = \int_0^{1/4} 4x(1-x)\mathrm{d}x = 4\left(\frac{\left(\frac{1}{4}\right)^2}{2} - \frac{\left(\frac{1}{4}\right)^3}{3}\right) = \frac{5}{48}.$$

Therefore

$$\mathrm{Var}(X) = 100\left(\frac{1}{192}\right) + 10\left(\frac{5}{48}\right) = \frac{25}{16}.$$

Problem 4. (Joint Ratio) (20 pts) Suppose that X and Y are jointly continuous random variables with joint distribution

$$f_{XY}(x, y) = \begin{cases} \frac{c}{x} & 0 < x \leq 1,\ 0 \leq y \leq 2x \leq 2 \\ 0 & \text{otherwise} \end{cases}.$$

a. (6 pts) Find the value c that makes $f_{XY}(x, y)$ a well defined joint density.

b. (6 pts) What are the marginals $f_X(x), f_Y(y)$?

c. (8 pts) Let $R = Y/X$. Find the CDF and PDF for R. Be sure to specify it's range. (Hint: write $P(R \leq r)$ as an area integral over a certain region)

Solution:

a. For normality, we require that

$$1 = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\mathrm{d}x\mathrm{d}y = \int_0^1 \left(\int_0^{2x} \frac{c}{x}\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 2c\mathrm{d}x = 2c$$

Therefore we need $c = \frac{1}{2}$.

b. The marginals are given for $x \in [0,1]$ by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_0^{2x} \frac{1}{2x}\mathrm{d}y = \frac{2x}{2x} = 1.$$

so

$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}.$$

Similarly for $0 < y \le 2$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x = \int_{y/2}^{1} \frac{1}{2x}\mathrm{d}x = -\frac{1}{2}\ln\left(\frac{y}{2}\right)$$

so

$$f_Y(y) = \begin{cases} -\frac{1}{2}\ln\left(\frac{y}{2}\right) & 0 < y \le 2 \\ 0 & \text{otherwise} \end{cases}.$$

c. Since we always have that $0 \le Y \le 2X$, it follows that $0 \le R = \frac{Y}{X} \le 2$. For $0 \le r \le 2$ the CDF is given by

$$F_R(r) = P(R \le r) = P(Y \le rX) = \int_0^1 \int_0^{rx} \frac{1}{2x}\mathrm{d}y\mathrm{d}x = \int_0^1 \frac{r}{2}\mathrm{d}x = \frac{r}{2}$$

Therefore

$$f_R(r) = \frac{\mathrm{d}}{\mathrm{d}r}F_R(r) = \begin{cases} \frac{1}{2} & 0 \le r \le 2 \\ 0 & \text{otherwise} \end{cases},$$

and so the range $R_R = [0,2]$.

Problem 5. (Standard Normality) (20 pts) Let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ be two independent normal random variables and let $Z = X + aY + b$, for two numbers $a, b \in \mathbb{R}$.

a. (6 pts) Find the mean and variance of Z.

b. (8 pts) Find the correlations, $\rho(Z,X)$ and $\rho(Z,Y)$.

c. (6 pts) Suppose that $\sigma_X^2 \le 1$. Using the fact that the sum of two normal random variables is again normal, find the values of a and b so that $Z \sim N(0,1)$.

Solution:

a. By linearity of expectation, the mean is

$$EZ = EX + aEY + b = \mu_X + a\mu_Y + b$$

Since X and Y are independent

$$\mathrm{Var}(Z) = \mathrm{Var}(X + aY + b) = \mathrm{Var}(X + aY) = \mathrm{Var}(X) + a^2\mathrm{Var}(Y) = \sigma_X^2 + a^2\sigma_Y^2$$

b. To find the correlations we first find the covariances

$$\text{Cov}(Z,X) = \text{Cov}(X + aY + b, X) = \text{Var}(X) + a\,\underbrace{\text{Cov}(Y,X)}_{=0} = \sigma_X^2$$

$$\text{Cov}(Z,Y) = \text{Cov}(X + aY + b, Y) = \underbrace{\text{Cov}(X,Y)}_{=0} + a\text{Var}(Y) = a\sigma_Y^2$$

Therefore

$$\rho(Z,X) = \frac{\text{Cov}(Z,X)}{\sqrt{\text{Var}(Z)\text{Var}(X)}} = \frac{\sigma_X}{\sqrt{\sigma_X^2 + a^2\sigma_Y^2}}$$

and

$$\rho(Z,Y) = \frac{\text{Cov}(Z,Y)}{\sqrt{\text{Var}(Z)\text{Var}(Y)}} = \frac{a\sigma_Y}{\sqrt{\sigma_X^2 + a^2\sigma_Y^2}}$$

c. For $Z \sim N(0,1)$, we need to solve

$$\mu_Z = \mu_X + a\mu_Y + b = 0$$

$$\sigma_Z^2 = \sigma_X^2 + a^2\sigma_Y^2 = 1.$$

The second equation gives

$$a = \frac{\sqrt{1-\sigma_X^2}}{\sigma_Y}$$

and upon substituting this into the first equation

$$b = -\mu_X - \frac{\mu_Y\sqrt{1-\sigma_X^2}}{\sigma_Y}$$

Problem 6. (Mix it up) (20 pts) Suppose a random variable Y has a CDF given by

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y^2 & 0 \le y < 1/2 \\ y & 1/2 \le y < 1 \\ 1 & y \ge 1. \end{cases}$$

a. (6 pts) Write $F_Y(y) = C(y) + D(y)$, where $C(y)$ is the continuous part and $D(y)$ is the discrete part (a staircase function).

b. (8 pts) Find the generalized PDF $f_Y(y)$.

c. (6 pts) What is the expected value of Y?

Solution:

a. There is a jump of size $(1/2) - (1/2)^2 = 1/4$ at $x = 1/4$ removing the jump gives

$$C(y) = \begin{cases} 0 & y < 0 \\ y^2 & 0 \le y < 1/2 \\ y - 1/4 & 1/2 \le y < 1 \\ 3/4 & y \ge 1 \end{cases},$$

which is now continuous. The staircase function just has one jump

$$D(y) = \frac{1}{4}u(y - 1/2).$$

b. Taking the derivative of $F_Y(y)$ gives the generalized PDF

$$f_Y(y) = \frac{1}{4}\delta(y - 1/2) + \begin{cases} 2y & 0 \le y \le 1/2 \\ 1 & 1/2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

c. The expectation is given by

$$\begin{aligned} EY &= \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y = \frac{1}{4}\left(\frac{1}{2}\right) + \int_0^{1/2} 2y^2\mathrm{d}y + \int_{1/2}^{1} y\,\mathrm{d}y \\ &= \frac{1}{8} + \frac{2}{3}\left(\frac{1}{2}\right)^3 + \frac{1}{2}\left[1 - \left(\frac{1}{2}\right)^2\right] = \frac{7}{12}. \end{aligned}$$