APMA 1650

Hard Practice Midterm Exam 2 - Solutions

Problem 1. (Random Rectangle) Suppose that you want to generate a random rectangle with area 1 (pick your favorite unit). You do this by randomly selecting its height H uniformly at random between 1 and 2. The width W is automatically determined by the fact that the area is 1.

a. Compute $E[W]$ and $\text{Var}(W)$.

b. What is the probability density function $f_W(w)$ for W? (Be sure to specify its domain and check that it is valid)

Solution. Since the area of the rectangle is 1 we must have

$$HW = 1 \quad \Rightarrow \quad W = \frac{1}{H}.$$

Note that since W is an explicit function of H, the two random variables H and W are not jointly continuous (i.e. they don't have a joint PDF in the traditional sense).

a. By the problem statement we know that $H \sim \text{Uniform}(1, 2)$ and therefore has PDF

$$f_H(h) = \begin{cases} 1 & 1 \le h \le 2 \\ 0 & \text{otherwise} \end{cases}.$$

To find $E[W]$ and $\text{Var}(W)$ we use that $W = 1/H$ and apply LOTUS

$$E[W] = \int_1^2 \frac{1}{h}\,\mathrm{d}h = [\ln(h)]_1^2 = \ln(2).$$

Similarly

$$E[W^2] = \int_1^2 \frac{1}{h^2}\,\mathrm{d}h = \left[\frac{-1}{h}\right]_1^2 = \frac{1}{2},$$

therefore

$$\text{Var}(W) = E(W^2) - (EW)^2 = \frac{1}{2} - (\ln(2))^2$$

b. To find the distribution for $W = 1/H$, we apply the CDF method

$$F_W(w) = P(W \le w) = P(1/H \le w) = P(1/w \le H) = 1 - F_H(1/w)$$

Taking the derivative with respect to w gives

$$f_W(w) = \frac{\mathrm{d}}{\mathrm{d}w}F_W(w) = -\frac{\mathrm{d}}{\mathrm{d}w}(F_H(1/w)) = \frac{f_h(1/w)}{w^2} = \begin{cases} \frac{1}{w^2} & 1 \le 1/w \le 2 \\ 0 & \text{otherwise} \end{cases},$$

or more simply

$$f_W(w) = \begin{cases} \frac{1}{w^2} & 1/2 \leq w \leq 1 \\ 0 & \text{otherwise} \end{cases},$$

so that the range of W is clearly $R_W = [1/2, 1]$.

Problem 2. (Joint Conditioning) Let X and Y be continuous random variables with joint probability density

$$f_{XY}(x, y) = \begin{cases} cy & \text{if} \quad 0 \leq x \leq \infty \quad \text{and} \quad 0 \leq y \leq e^{-x} \\ 0 & \text{otherwise} \end{cases}.$$

a. Find the value c for which $f_{XY}(x, y)$ is a valid joint probability density.

b. Find $E[e^X]$.

c. Find the conditional probability density of X given $Y = y$.

Solution.

a. The normality condition implies

$$1 = c \int_0^\infty \int_0^{e^{-x}} y \, \mathrm{d}y \mathrm{d}x = \frac{c}{2} \int_0^\infty e^{-2x} \mathrm{d}x = \frac{c}{4}$$

Therefore $c = 4$.

b. By LOTUS, we have

$$E[e^X] = \int_0^\infty \int_0^{e^{-x}} 4ye^x \, \mathrm{d}y \mathrm{d}x = 2 \int_0^\infty e^x e^{-2x} \mathrm{d}x = 2 \int_0^\infty e^{-x} \mathrm{d}x = 2$$

c. To find the conditional density, we first compute the marginal $f_Y(y)$. Note that for $y \in (0, 1]$ and $x \geq 0$

$$0 \leq y \leq e^{-x} \quad \Leftrightarrow \quad 0 \leq x \leq -\ln(y)$$

so that the Y marginal is given by

$$f_Y(y) = \int_{-\infty}^\infty f_{XY}(x, y) \mathrm{d}x = \int_0^{-\ln(y)} 4y \mathrm{d}x = \begin{cases} -4y \ln(y) & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Therefore the conditional PDF is for $0 < y < 1$

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \begin{cases} -\frac{1}{\ln(y)} & 0 \leq x \leq -\ln(y) \\ 0 & \text{otherwise} \end{cases}$$

Note that this means that X conditioned on X is just a uniform random variable $(X|Y = y) \sim \text{Uniform}(0, -\ln(y))$.

Problem 3. (Jobs) Suppose a company has determined that the number of jobs per week, N, varies from week to week and has a Poisson distribution with rate λ. The number of hours to complete each job Y_i is Gamma(α, β). The total time to complete all jobs in a week is $T = \sum_{i=1}^{N} Y_i$. Note that T is the sum of a random number of random variables.

a. Find $E[T|N = n]$.

b. Find ET, the expected time it takes to complete all the jobs.

Solution.

a. Note that by the properties of conditional expectation and the mean of Gamma(α, β)

$$E[T|N = n] = E\left[\sum_{i=1}^{N} Y_i \middle| N = n\right] = \sum_{i=1}^{n} E[Y_i|N = n] = nE[Y_i] = \frac{n\alpha}{\beta}.$$

b. The expected time to complete all the jobs can then be computed by the law of iterated expectation

$$E[T] = E[E[T|N]] = E\left[\frac{N\alpha}{\beta}\right] = \frac{\alpha}{\beta}E[N] = \frac{\alpha\lambda}{\beta}.$$

Problem 4. (Tight)

a. Let X be a Bernoulli(p) random variable. Show that Markov's inequality is actually an equality for $a = 1$, i.e. $P(X \geq 1) = EX$.

b. Let X_1 and X_2 be two independent Bernoulli(p) RVs and let $X = X_1 - X_2$. Show that Chebyshev is an equality for $b = 1$, i.e. $P(|X - EX| \geq 1) = \text{Var}(X)$.

Solution.

a. Note that since X is Bernoulli,

$$P(X \geq 1) = P(X = 1) = p$$

while similarly

$$EX = P(X = 1) = p$$

Therefore we have $P(X \geq 1) = EX$.

b. For this, we note that X satisfies

$$X = \begin{cases} 1 & \text{with probability } P(X_1 = 1, X_2 = 0) = p(1-p) \\ 0 & \text{with probability } P(X_1 = 1, X_2 = 1) + P(X_1 = 0, X_2 = 0) = 2p^2 \\ -1 & \text{with probability } P(X_1 = 0, X_2 = 1) = p(1-p) \end{cases}$$

and $EX = EX_1 - EX_2 = p - p = 0$, so that

$$P(|X - EX| \geq 1) = P(|X| = 1) = 2p(1-p)$$

While using the properties of variance satisfies (since X_1 and X_2 are independent)

$$\text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) = 2p(1-p).$$

Therefore we have

$$P(|X - EX| \geq 1) = \text{Var}(X).$$

Problem 5. (Boxes) Suppose that m homeworks are distributed randomly into n homework dropboxes (each dropbox can have more than one homework assignment). Let X denote the number of empty dropboxes.

a. Find EX

b. Find $\text{Var}(X)$.

Solution. Before we begin, we let X_i be a Bernoulli random variable

$$X_i = \begin{cases} 1 & \text{if the ith dropbox is empty} \\ 0 & \text{otherwise} \end{cases}$$

So that

$$X = \sum_{i=1}^{n} X_i$$

a. Note that the probability that one of the students doesn't pick the ith dropbox is just $(1 - 1/n)$, therefore

$$EX_i = P(X_i = 1) = \left(1 - \frac{1}{n}\right)^m.$$

To find EX, we just use linearity of expectation

$$EX = \sum_{i=1}^{n} EX_i = n\left(1 - \frac{1}{n}\right)^m$$

b. Note that X_i are not independent. So to find the variance, we use the properties of covariance

$$\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \text{Cov}(X_i, X_j)$$

We know that since X_i is Bernoulli

$$\text{Var}(X_i) = \left(1 - \frac{1}{n}\right)^m \left(1 - \left(1 - \frac{1}{n}\right)^m\right) = \left(1 - \frac{1}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m}$$

while the covariance for $i \neq j$ is given by

$$\mathrm{Cov}(X_i, X_j) = E[X_i X_j] - (EX_i)(EX_j)$$

Note that $X_i X_j$ is another Bernoulli random variable

$$X_i X_j = \begin{cases} 1 & \text{if the ith and jth bins are empty} \\ 0 & \text{otherwise} \end{cases}$$

Note that if $X_i X_j = 1$, then both X_i and X_j must be 1. It follows from the law of total probability that

$$P(X_i X_j = 1) = P(X_i = 1 | X_j = 1) P(X_j = 1)$$

and that the conditional probability $P(X_i = 1 | X_j = 1)$ is just the probability that the ith bin is empty given that the jth bin is also empty is

$$P(X_i = 1 | X_j = 1) = \left(1 - \frac{1}{n-1}\right)^m.$$

Therefore the probability that m homeworks don't end up in two distinct bins is for $i \neq j$

$$E[X_i X_j] = P(X_i X_j = 1) = \left(1 - \frac{1}{n-1}\right)^m \left(1 - \frac{1}{n}\right)^m = \left(1 - \frac{2}{n}\right)^m,$$

where in the last line we used that

$$\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{n}\right) = \left(1 - \frac{2}{n}\right).$$

Therefore

$$\mathrm{Cov}(X_i, X_j) = \left(1 - \frac{2}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m}$$

It follows that the variance is

$$\begin{aligned}
\mathrm{Var}(X) &= n\mathrm{Var}(X_i) + n(n-1)\mathrm{Cov}(X_i, X_j) \\
&= n\left[\left(1 - \frac{1}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m}\right] + n(n-1)\left[\left(1 - \frac{2}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m}\right] \\
&= n\left(1 - \frac{1}{n}\right)^m \left(1 - n\left(1 - \frac{1}{n}\right)^m\right) + n(n-1)\left(1 - \frac{2}{n}\right)^m
\end{aligned}$$