Homework 3 Solutions

1. (10 points) *Messenger*

   a. (X, Y) is a random point on the rectangle 13x2 whose center is at the origin, in other words X and Y are independent uniform over $(-6.5, 6.5)$ and $(-1, 1)$, respectively, and the distance is $Z = |X| + |Y|$. Also note that the four quadrants are symmetric.

$$E(Z) = \int_{-1}^{1} \int_{-6.5}^{6.5} (|x| + |y|) \frac{1}{13} \frac{1}{2} dx dy \tag{1}$$
$$= 4 \int_{0}^{1} \int_{0}^{6.5} (x + y) \frac{1}{26} dx dy \tag{2}$$
$$= \frac{4}{26} \int_{0}^{1} (6.5^2/2 + 6.5y) dy = 4/26(6.5^2/2 + 6.5/2) = 3.75 \tag{3}$$

   b. Recall that the Markov inequality for a positive random variable is $\mathrm{P}(X > a) \leq E(X)/a$. Also clearly distance is positive:

$$\mathrm{P}(Z > 5) \leq 3.75/5 = 75\% \tag{4}$$

2. (10 points) *Pasta and rice*

   a. The constraints are $100 \leq \max\{X, R\} \leq 300$. The joint pdf is constant over that region. Figure 1 contains the diagram.

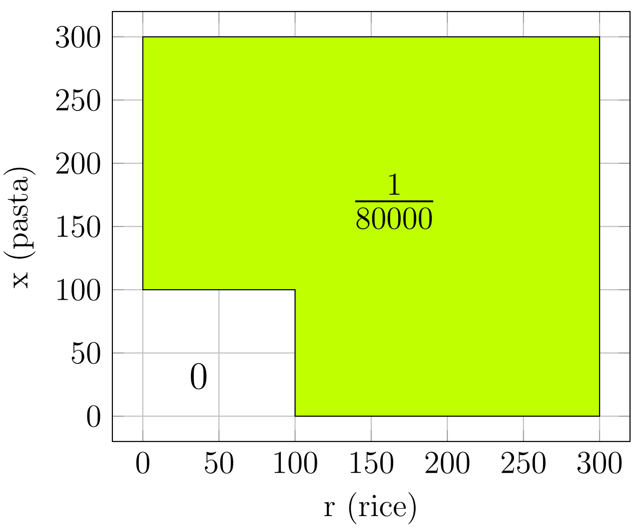

Figure 1: Joint pdf of R and X.

   b. We compute

$$\mathrm{E}(RX) = \int_{x=100}^{300} \int_{r=0}^{300} \frac{rx}{80000} \,\mathrm{d}x \,\mathrm{d}r + \int_{x=0}^{100} \int_{r=100}^{300} \frac{rx}{80000} \,\mathrm{d}x \,\mathrm{d}r \tag{5}$$
$$= \frac{1}{80000} \left(\frac{x^2}{2} \Big]_{100}^{300} \frac{r^2}{2} \Big]_{0}^{300} + \frac{x^2}{2} \Big]_{0}^{100} \frac{r^2}{2} \Big]_{100}^{300} \right) \tag{6}$$
$$= 25000. \tag{7}$$

$$E(X) = \int_{x=100}^{300} \int_{r=0}^{300} \frac{x}{80000} \, \mathrm{d}x \, \mathrm{d}r + \int_{x=0}^{100} \int_{r=100}^{300} \frac{x}{80000} \, \mathrm{d}x \, \mathrm{d}r \tag{8}$$

$$= \frac{1}{80000} \left(300 \frac{x^2}{2}]_{100}^{300} + 200 \frac{x^2}{2}]_{0}^{100} \ \frac{r^2}{2}]_{100}^{300} \right) \tag{9}$$

$$= 162.5. \tag{10}$$

By symmetry $E(R) = 162.5$. The covariance $\mathrm{Cov}(R, X) = E(RX) - E(R)E(X) = 1406.25$ so R and X are negatively correlated.

c. The variables are correlated, which implies they cannot be independent.

3. (10 points) *Restaurant*

a. Let N be the number of customers and C_i the money customer i spends. The money made on a given night is

$$M = \sum_{i=1}^{N} C_i. \tag{11}$$

We are interested in the mean of this random variable. Conditioned on $N = n$ the mean is

$$E(M|N = n) = \sum_{i=1}^{n} E(C_i|N = n) = n \, E(C_1|N = n), \tag{12}$$

as long as all the C_i have the same conditional mean given the number of customers. This assumption is implicit in what the management is saying.

Now, if N and C_i (i.e. the number of customers and what each customer spends) are **independent** then $E(C_i|N = n) = E(C_i) = 30$, so $E(M|N = n) = 30n$. This would imply that by iterated expectation

$$E(M) = E(E(M|N)) = E(N\mu) = E(30N) = 30 \, E(N) = 1200 \text{ dollars}. \tag{13}$$

In fact, if you want to be very rigorous, we don't need N and C_i to be independent, as long as $E(C_i|N = n) = E(C_i) = 30$, which is a less strong requirement.

b. Under the imagined scenario

$$E(N) = 100 \, P(N = 100) + 10 \, P(N = 10) \tag{14}$$

$$= 100 \, P(\text{good night}) + 10 \, (1 - P(\text{good night})) \tag{15}$$

$$= 90 \, P(\text{good night}) + 10 = 40, \tag{16}$$

so $P(\text{good night}) = 1/3$.

c. To compute the expected earned money per night under the assumptions we first compute the conditional expectation given N

$$E(M|N) = \begin{cases} 1000 & \text{if } N = 100 \\ 400 & \text{if } N = 10. \end{cases} \tag{17}$$

Now by iterated expectation

$$\mathrm{E}\left(\mathrm{E}\left(M|N\right)\right) = 1000\,\mathrm{P}\left(N = 100\right) + 400\,\mathrm{P}\left(N = 10\right) = 600 \text{ dollars.} \tag{18}$$

You are telling this story to show that the actual expected gain can be very different (in this case half!) from the management's back-of-the-envelope calculation if the money spent by each customer depends on the number of clients.

4. (10 points) *Copper*

a. Let us define the random variables C (amount of copper), X (price of copper) and V (value of the stored copper). From the problem statement $V \leq 2.5$ million dollars and $\mathrm{E}(V) = 2$ million dollars. Since we only know the mean of the random variable, we use Markov's inequality. Since we want to bound the probability that V is *smaller* than a certain quantity, we will apply the inequality to $-V$ or better still to $Y = 2.5$ million$-V$ which is nonnegative:

$$\mathrm{P}\left(V \leq 1 \text{ million}\right) = \mathrm{P}\left(Y > 1.5 \text{ million}\right) \tag{19}$$

$$\leq \frac{\mathrm{E}\left(Y\right)}{1.5 \text{ million}} = \frac{\mathrm{E}\left(2.5 \text{ million} - V\right)}{1.5 \text{ million}} = \frac{0.5 \text{ million}}{1.5 \text{ million}} = \frac{1}{3}. \tag{20}$$

The probability of the event of interest is at most 1/3.

b. Probably not, as the company will buy more copper when it is cheaper and sell when it is more expensive, so it seems plausible for C and X to have negative correlation.

c. In order to use the additional information we apply Chebyshev's inequality. For this we need the variance of V. First, note that by independence,

$$\mathrm{E}\left(V\right) = \mathrm{E}\left(XC\right) = \mathrm{E}\left(X\right)\mathrm{E}\left(C\right), \tag{21}$$

so $\mathrm{E}\left(C\right) = \mathrm{E}\left(V\right) / \mathrm{E}\left(X\right) = 4/9$ million. We can now compute

$$\mathrm{E}\left(V^2\right) = \mathrm{E}\left(X^2C^2\right) \tag{22}$$

$$= \mathrm{E}\left(X^2\right)\mathrm{E}\left(C^2\right) \text{ by independence} \tag{23}$$

$$= \left(\mathrm{Var}\left(X\right) + \mathrm{E}^2\left(X\right)\right)\left(\mathrm{Var}\left(C^2\right) + \mathrm{E}^2\left(C\right)\right) \tag{24}$$

$$= \left(0.2^2 + 4.5^2\right)\left(10000^2 + \left(\frac{4000000}{9}\right)^2\right) \tag{25}$$

$$\approx 4\,10^{12} + 9.93\,10^9. \tag{26}$$

$$\mathrm{Var}\left(V\right) = \mathrm{E}\left(V^2\right) - \mathrm{E}^2\left(V\right) \approx 9.93\,10^9. \tag{27}$$

Now,

$$\mathrm{P}\left(V \leq 1 \text{ million}\right) \leq \mathrm{P}\left(|V - \mathrm{E}\left(V\right)| > 1 \text{ million}\right) \tag{28}$$

$$\leq \frac{\mathrm{Var}\left(V\right)}{10^{12}} \approx 0.993\%. \tag{29}$$

A **much** sharper bound.

5. (10 points) *Law of conditional variance*

a. Given $X = x$, X is constant, and its value is x, therefore once x is fixed $\mathrm{Var}(Y|X = x)$ is a **number**. It represents the variance of the random variable Y given the information $X = x$. In other words, given two random variables X and Y, look at their joint density, condition X to the value x. This slice gives rise to a new distribution of the random variable $Y|(X = x)$, $\mathrm{Var}(Y|X = x)$ is precisely the variance of this random variable.

$$\mathrm{Var}(Y|X = x) = E((Y - E(Y|X = x))^2|X = x) \tag{30}$$

b. Now we don't keep x fixed but rather we treat it as a variable, for any x we assign a value $\mathrm{Var}(Y|X = x)$, this is a function from the range of X to real numbers, which we call by h, so $h(x) = \mathrm{Var}(Y|X = x)$. Since this function is defined on a probability space we can regard it as a **random variable** by $h(X) = \mathrm{Var}(Y|X)$.

c. Using iterated expectations we get:

$$\begin{aligned}
E[\mathrm{Var}(Y|X)] &= E[E(Y^2|X)] - E[E(Y|X)^2] && (31)\\
&= E(Y^2) - E[E(Y|X)^2] && (32)\\
\mathrm{Var}[E(Y|X)] &= E[E(Y|X)^2] - E[E(Y|X)]^2 && (33)\\
&= E[E(Y|X)^2] - E(Y)^2 && (34)\\
\implies E[\mathrm{Var}(Y|X)] + \mathrm{Var}[E(Y|X)] &= E(Y^2) - E(Y)^2 = \mathrm{Var}(Y) && (35)
\end{aligned}$$

Average of the variance of Y given X, plus the variance of the average of Y given X.

d. Let T be the time at which a runner gets injured. And let A be the random variable that describes the age group: $A = 1$ be the group of runners below 30, and $A = 2$ be the group above 30. So that $T|\{A = 1\} \sim exp(1)$ and $T|\{A = 2\} \sim exp(2)$. We are also given that $P(A = 2) = 0.2$. Also note that if $X \sim exp(\lambda)$ then $E(X^2) = Var(X) + E(X)^2 = 1/\lambda^2 + 1/\lambda^2 = 2/\lambda^2$.

$$\begin{aligned}
E(T) &= E(E(T|A)) = E(T|A = 1)P(A = 1) + E(T|A = 2)P(A = 2) && (36)\\
&= 1 \cdot 0.8 + 2 \cdot 0.2 = 1.2 && (37)\\
E(T^2) &= E(E(T^2|A)) = E(T^2|A = 1)P(A = 1) + E(T^2|A = 2)P(A = 2) && (38)\\
&= 2/1^2 \cdot 0.8 + 2/2^2 \cdot 0.2 = 1.7 && (39)\\
\implies \quad \mathrm{Var}(T) &= 0.26 && (40)
\end{aligned}$$