Assignment 3

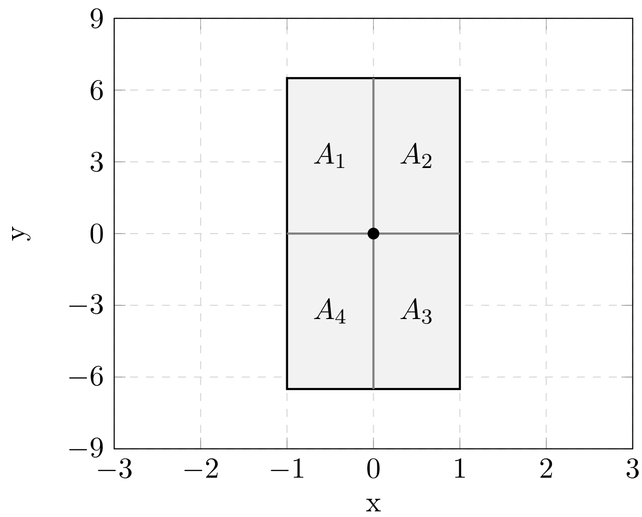
Benjamin Jakubowski

October 2, 2015

1. Messenger

a. Expected Manhattan Distance

First, assuming the delivery points are uniformly spread in all of Manhattan, let $X \sim U(-1, 1)$ and $Y \sim U(-6.5, 6.5)$ represent the x and y coordinate of the delivery point. This model is depicted in the plot below:



Next, note the definition of D (and uniformity in the distribution of delivery points) implies:

$$E[D|(x, y) \in A_i] = E[D] \text{ for } i, j \in \{1, 2, 3, 4\}$$

Thus, to proceed we can next limit our attention to simply A_2. Then

$$\begin{aligned} E[D] &= E[D|(x, y) \in A_i] \\ &= E[X + Y|(x, y) \in A_i] \\ &= E[X|(x, y) \in A_i] + E[Y|(x, y) \in A_i] \end{aligned}$$

Since the joint PDF is uniform, the marginals are also uniform. Since the expected value of a $U(a, b)$ random variable is $1/2 \cdot (b - a)$,

$$
\begin{aligned}
E[D] &= E[X|(x, y) \in A_i] + E[Y|(x, y) \in A_i] \\
&= 1/2(1 - 0) + 1/2(6.5 - 0) = 4.25
\end{aligned}
$$

Thus, $E[D] = 4.25$.

B. Bound on P(D) > 5

Recall Markov's Inequality states, for a non-negative RV X:

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Thus,

$$P(D \geq 5) \leq \frac{E[D]}{a} = \frac{4.25}{5} = .85$$

Thus, the probability a messenger has to travel more than 5 miles is less than .85.

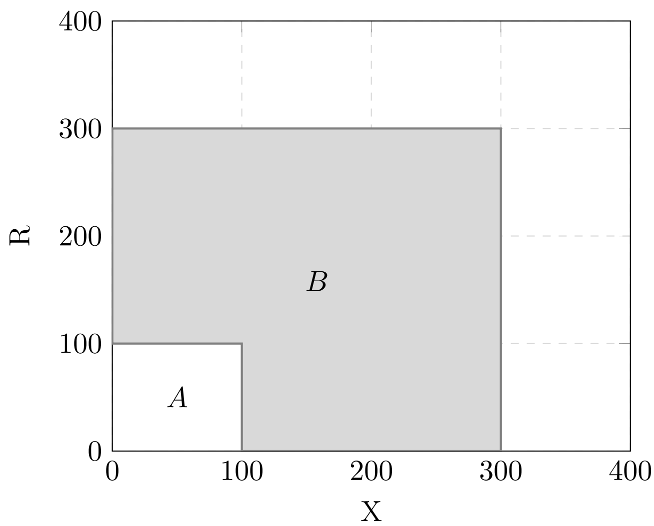
2. Pasta and Rice

A. Joint PDF of X and R

To start, information provided by the cook implies:

$$
\begin{aligned}
&P(X \geq 100 \cup R \geq 100) = 1 \\
\text{and, equivalently } &P(X < 100 \cap R < 100) = 0
\end{aligned}
$$

Moreover, $P(X \leq 300) = 1$ and $P(R \leq 300) = 1$. Based on this information (and the assumption the joint pdf is constant), we can construct the joint pdf of X and R $f_{X,R}(x, r)$ as shown below:



We can see

$$
\begin{aligned}
P((x,r) \in A) &= 0 \\
P((x,r) \in B) &= 1 \\
f_{X,R}(x,r) &= c \neq 0 \text{ for } (x,r) \in B \\
f_{X,R}(x,r) &= 0 \text{ for } (x,r) \in A.
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\int_B f_{X,R}(x,r)\mathrm{d}x\mathrm{d}r &= 1 \\
\int_B c\mathrm{d}x\mathrm{d}r &= 1 \\
c \cdot \int_B \mathrm{d}x\mathrm{d}r &= 1 \\
c = \frac{1}{Area_B} = \frac{1}{300^2 - 100^2} &= \frac{1}{80000}
\end{aligned}
$$

B. Correlation of X and R

To determine if X and R are uncorrelated, we need to see if $Cov(X,R) = 0$. Recall $Cov(X,R) = E(XR) - E(X)E(R)$. Thus, let's start by finding $E(XR)$:

$$
\begin{aligned}
E(XR) &= \int_0^{300}\int_0^{300} x \cdot y \cdot f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{100}^{300}\int_{100}^{300} x \cdot y \frac{1}{80000}\mathrm{d}x\mathrm{d}y + \int_0^{100}\int_{100}^{300} x \cdot y \frac{1}{80000}\mathrm{d}x\mathrm{d}y + \int_{100}^{300}\int_0^{100} x \cdot y \frac{1}{80000}\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{80000}\Big[\int_{100}^{300} [1/2x^2y]\big|_{x=100}^{300}\mathrm{d}y + \int_0^{100} [1/2x^2y]\big|_{x=100}^{300}\mathrm{d}y + \int_{100}^{300} [1/2x^2y]\big|_{x=0}^{100}\mathrm{d}y\Big] \\
&= \frac{1}{80000}\Big[\int_{100}^{300} (\frac{1}{2} \cdot 80000y)\mathrm{d}y + \int_0^{100} (\frac{1}{2} \cdot 80000y)\mathrm{d}y + \int_{100}^{300} (\frac{1}{2} \cdot 10000y)\mathrm{d}y\Big] \\
&= \frac{1}{80000}\Big[(\frac{1}{4} \cdot 80000y^2)\big|_{y=100}^{300} + (\frac{1}{4} \cdot 80000y^2)\big|_{y=0}^{100} + (\frac{1}{4} \cdot 10000y^2)\big|_{y=100}^{300}\Big] \\
&= \frac{1}{80000}\Big[\frac{1}{4} \cdot 80000^2 + \frac{1}{4} \cdot 80000 \cdot 10000 + \frac{1}{4} \cdot 10000 \cdot 80000)\Big] \\
&= \frac{1}{4}(80000 + 10000 + 10000) = 25000
\end{aligned}
$$

Next, note $E(X) = E(R)$ (by symmetry). Thus, $E(X)E(R) = (E(x))^2$. To find $E(X)$, we first determine the marginal PDF $f_X(x)$:

$$
f_X(x) = \begin{cases} \int_{100}^{300} \frac{1}{80000}\mathrm{d}y = \frac{1}{400} & \text{for } 0 \leq x < 100 \\ \int_0^{300} \frac{1}{80000}\mathrm{d}y = \frac{3}{800} & \text{for } 100 \leq x \leq 300 \end{cases}
$$

Thus,

$$E(X) = \int_0^{300} x f_X(x)\mathrm{d}x = \int_0^{100} x \cdot \frac{1}{400}\mathrm{d}x + \int_{100}^{300} x \cdot \frac{3}{800}\mathrm{d}x$$
$$= \frac{1}{400} \cdot \frac{1}{2} \cdot 100^2 + \frac{3}{800} \cdot \frac{1}{2} \cdot (300^2 - 100^2) = 162.5$$

Hence, $E(X)E(R) = E(X)^2 = 162.5^2 = 26406.25$.
Thus, $Cov(X,R) = 25000 - 26406.25 \neq 0$, so X and R are correlated.

c. Independence of X and R

X and R are clearly not independent. This follows directly from the correlation of X and R (since correlation implies dependence).

3. Restaurant

a. Responding to the management

Recall management stated *"on average we have 40 customers every night and on average each customer spends 30 dollars, so on average we make 1200 dollars per night."*
To model the problem probabilistically, let C be the number of customers on a single night and D be the number of dollars spent by a single customer. Then $E(C) = 40$ and $E(D) = 30$, so $E(C)E(D) = 1200$.
Now recall

$$Cov(C,D) = E(CD) - E(C)E(D)$$

so

$$Cov(C,D) + E(C)E(D) = E(CD)$$

Therefore, $E(CD) = 1200$ if and only if $Cov(C,D) = 0$.
Thus, the statement *"on average we make 1200 dollars per night"* is true if C and D are uncorrelated. As a side note, since independence implies uncorrelation, the statement is also obviously true under the condition of independence.

b. Probability of good night

Now, let

$$C = \begin{cases} 100 & \text{on a good night} \\ 0 & \text{on a bad night} \end{cases}$$

$$P_C(c) = \begin{cases} p & \text{for } c = 100 \\ (1-p) & \text{for } c = 10 \end{cases}$$

Then

$$E(C) = 40 = 100p + 10(1-p) \qquad = 100p + 10 - 10p = 90p + 10$$

Thus, $30 = 90p$ and $p = 1/3$.

C. Average amount each customer spends

If $p = 1/3$, then

$$\begin{aligned} E(D) &= E(D|C=10)P_C(10) + E(D|C=100)P_C(100) \\ &= 40 \cdot 2/3 + 10 \cdot 1/3 = 30 \end{aligned}$$

So $E(D) = 30$.

D. Average nightly revenue

The average nightly revenue is $E(CD)$.

$$\begin{aligned} E(CD) &= E(CD|C=10)P_C(c) + E(CD|C=100)P_C(100) \\ &= 10 \cdot 40 \cdot 2/3 + 10 \cdot 100 \cdot 1/3 \\ &= 600 \end{aligned}$$

Thus, under these assumptions, the expected (or average) nightly revenue is \$600. This is why you're telling this story to the management- you're showing them that the expected nightly revenue can be significantly different than the product of the expected number of customers (per night) and the expected revenue (per customer). More concretely, they might make a lot less money per night than they expect, and they may need to change their business plan accordingly.

4. Copper

A. Upper bound on desired probability

First, let's define some random variables:

$$\begin{aligned} C &= \text{Dollar value of the copper in stock} \\ D &= \text{Price of copper (dollars per pound)} \\ W &= \text{Amount of copper in stock (pounds)} \end{aligned}$$

In addition, by definition $W = C/D$.
Next, note the company has provided us with three important pieces of information:

$$\begin{aligned} E(C) &= 2000000 \\ P(D \leq 5) &= 1 \\ P(500000 < W) &= 0 \end{aligned}$$

Now, since $C = W \cdot D$, and $W \leq 500000, D \leq 5$ with probability 1, $C \leq 2500000$ with probability 1.
Now , given $E(C) = 2000000$ and $C \leq 2500000$ with probability 1, we state our bound:

$$P(C \leq 1000000) \leq \frac{1}{3}$$

To see this, note that $P(C \leq 1000000)$ is maximized by the distribution:

$$C_{max} = \begin{cases} 1000000 & \text{with probability } 1/3 \\ 2500000 & \text{with probability } 2/3 \end{cases}$$

First, note this distribution satisfies the constraint

$$E(C) = 1000000 \cdot 1/3 + 2500000 \cdot 2/3 = 2000000$$

Since we're placing an upper bound on $P(C \leq 1000000)$, we now don't need to consider any candidate distribution C_{cand} with $P(C_{cand} \leq 1000000) < 1/3$.

Next, note (interestingly, but unnecessarily for the purpose of proving our bound) that any C_{cand} with $P(C_{cand} \leq 1000000) = 1/3$ must be equal in distribution to C_{max}. If it were not, then $P_{C_{cand}}(c) > 0$ for some $c < 1000000$, so $E(C_{cand}) < 2000000$.

Finally (for similar reasons), for any C_{cand} with $P(C_{cand} \leq 1000000) > 1/3$, it must be true that $E(C_{cand}) < 2000000$.

Therefore, $P(C \leq 1000000) \leq 1/3$

B. Independence of price and amount of copper

It is *not* sensible to model the price of copper and the stock of copper as independent random variables. This judgement is based on the assumption the company tries to maximize profit by buying low and selling high. When the price D increases, the company would sell, decreasing the amount of copper in stock. When the price decreases, they would buy, increasing the stock of copper. Therefore, we'd expect $Cov(W, D) < 0$. Regardless of sign, $Cov(W, D) \neq 0$ so independence is not a sensible assumption.

C. Improved bound on desired probability

Again, we're interested in $P(C < 1000000)$. However, this time we're given additional information, namely $E(D) = 4$ and $\sigma_D = 0.2$ (so $Var(D) = 0.04$).
To start, note we're (unfortunately) assuming W and D are independent. Thus, $E(C) = E(WD) = E(W)E(D)$. Thus

$$E(D) = E(C)/E(D) = 2000000/4.5$$

Thus, we now know the following:

RV	E()	Var()
C	2000000	
W	2000000/4.5	10000
D	4.5	.04

In order to improve on our bound, our strategy will be to:

1. Determine $Var(C)$
2. Use Chebyshev's Inequality to bound $P(C \leq 1000000)$

First, to determine $Var(C)$

$$\begin{aligned} Var(C) &= E(C^2) - (E(C))^2 \\ &= E(W^2D^2) - (E(WD))^2 \\ &= E(W^2)E(D^2) - (E(W)E(D))^2 \\ &= E(W^2)E(D^2) - E(W)^2E(D)^2 \end{aligned}$$

But

$$\begin{aligned} E(W^2) &= Var(W) + E(W)^2 \\ E(D^2) &= Var(D) + E(D)^2 \end{aligned}$$

So substitution yields

$$\begin{aligned} Var(C) &= E(W^2)E(D^2) - E(W)^2E(D)^2 \\ &= (Var(W) + E(W)^2)(Var(D) + E(D)^2) - E(W)^2E(D)^2 \\ &= Var(W)Var(D) + Var(D)E(W)^2 + Var(W)E(D)^2 + E(W)^2E(D)^2 - E(W)^2E(D)^2 \end{aligned}$$

So

$$\begin{aligned} Var(C) &= Var(W)Var(D) + Var(D)E(W)^2 + Var(W)E(D)^2 \\ &= .04 \cdot 10000 + .04 \cdot (2000000/4.5)^2 + 10000 \cdot (4.5)^2 \\ &\approx 7901437468 \end{aligned}$$

Now, recall Chebyshev's inequality gives a bound on the distance between a random variable X and it's mean:

$$P(|X - E(X)| > a) \leq \frac{Var(X)}{a^2}$$

Thus, we have

$$P(|C - 2000000| > 100000) \leq \frac{7901437468}{(1000000)^2} = 0.0079$$

Therefore $P(C \leq 1000000) \leq 0.0079$.

5. Law of Conditional Variance

a. Interpreting $Var(Y|X = x)$

$Var(Y|X = x)$ is a function of x. It maps the value $X = x$ onto the variance of Y about its mean over the plane defined by $X = x$. To see this, recall the conditional distribution $f_{Y|X}(Y|X = x)$ is the distribution of Y, holding $X = x$. You can think of it as observing the joint pdf $f_{X,Y}$ over just the plane $X = x$. Then the conditional variance $Var(Y|X = x)$ measures the variance of Y in just this slice of the pdf.

b. Interpreting $Var(Y|X)$

$Var(Y|X) = h(X)$ is a random variable (since a function of a random variable is itself a random variable). It takes the random variable X from and maps it to $Var(Y|X)$.

c. Proof of Law of Conditional Variance

The law of conditional variance states:

$$Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$$

To see this equivalence, first recall

$$\begin{aligned} Var(Y) &= E(Y^2) - (E(Y))^2 \\ &= E(E(Y^2|X)) - (E(E(Y|X)))^2 \text{ \textit{(by the law of iterated expectations)}} \end{aligned}$$

But $E(Y^2|X) = Var(Y|X) + (E(Y|X))^2$, so

$$\begin{aligned} Var(Y) &= E(Var(Y|X) + E(Y|X)^2) - (E(E(Y|X)))^2 \\ &= E(Var(Y|X)) + E(E(Y|X)^2) - (E(E(Y|X)))^2 \text{ \textit{(by the linearity of expectations)}} \end{aligned}$$

But $E(E(Y|X)^2) - (E(E(Y|X)))^2 = Var(E(Y|X))$, so

$$Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$$

This, in essence, decomposes the variance of Y into two parts. The first, $E(Var(Y|X))$ describes how much on average Y deviates from its conditional mean $E(Y|X = x)$ (i.e. how spread Y is from its conditional mean, averaged over all possible values $X = x$). The second, $Var(E(Y|X))$, describes how much the conditional means $E(Y|X = x)$ deviate from the marginal mean $E(Y)$. Thus, in a sense, the first term describes on average how much variability is left in Y if X is given, while the second term describes how much variability there is in your expected value of Y if X is given.

D. Mean and SD of Time of Injury

First, let T be the exponential random variable modeling time at injury. Then let L be the parameter of the exponential random variable T. Then

$$L = \begin{cases} 1 & \text{if } Age_{runner} < 30 \\ 2 & \text{if } Age_{runner} \geq 30 \end{cases}$$

$$P_L(l) = \begin{cases} .8 & \text{for } l = 1 \\ .2 & \text{for } l = 2 \end{cases}$$

Next, note $T \sim Exp(L)$. Thus, $E(T|L=l) = \frac{1}{l}$ and $Var(T|L=l) = \frac{1}{l^2}$.
Then, by the law of iterated expectations

$$\begin{aligned} E(T) &= E(E(T|L)) \\ &= \sum_{l \in R_L} E(T|L=l) \cdot P_L(l) \\ &= E(T|L=l) \cdot P_L(1) + E(T|L=2) \cdot P_L(2) \\ &= \frac{1}{1} \cdot .8 + \frac{1}{2} \cdot .2 = .9 \end{aligned}$$

Next, we will apply the law of conditional variance,

$$Var(T) = E(Var(T|L)) + Var(E(T|L))$$

We will need to first evaluate the two terms on the right side of this expression. First,

$$\begin{aligned} E(Var(T|L)) &= \sum_{l \in R_L} Var(T|L=l) \cdot P_L(l) \\ &= Var(T|L=l) \cdot P_L(1) + Var(T|L=2) \cdot P_L(2) \\ &= \frac{1}{1^2} \cdot .8 + \frac{1}{2^2} \cdot .2 = .85 \end{aligned}$$

Second,

$$\begin{aligned} Var(E(T|L)) &= \sum_{l \in R_L} (E(T|L=l) - E(E(T|L)))^2 \cdot P_L(l) \\ &= \sum_{l \in R_L} (E(T|L=l) - E(T))^2 \cdot P_L(l) \\ &= \sum_{l \in R_L} (E(T|L=l) - .9)^2 \cdot P_L(l) \\ &= (\frac{1}{1} - .9)^2 \cdot .8 + (\frac{1}{2} - .9)^2 \cdot .2 = .328 \end{aligned}$$

Thus $E(T) = .9$ and $Var(T) = .85 + .328 = 1.178$.