Assignment 2

Benjamin Jakubowski

September 24, 2015

1. Spider on a Wall

a. Position of Spider on Wall

Let $F_{X,Y}(x,y)$ be the pdf. Then, since the spider spends twice as much time under the painting than it does on the rest of the wall,

$$\int_{y=6}^{8}\int_{x=4}^{6} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 2/3$$

The spider is equally likely to be anywhere under the painting, so $f_{X,Y}(x,y) = c$ and

$$\begin{aligned}\int_{y=6}^{8}\int_{x=4}^{6} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y &= 2/3\\ \int_{y=6}^{8}\int_{x=4}^{6} c\mathrm{d}x\mathrm{d}y &= 2/3\\ \int_{y=6}^{8} 2c\mathrm{d}y &= 2/3\\ 4c &= 2/3\end{aligned}$$

Thus $c = 1/6$. Now let S be the region of the wall not under the painting. Note $Area_S = 96$, and

$$\int_{\{(x,y)\in S\}} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 1/3$$

Since the spider is equally likely to be anywhere in S,

$$\begin{aligned}\int_{\{(x,y)\in S\}} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y &= 1/3\\ \int_{\{(x,y)\in S\}} k\mathrm{d}x\mathrm{d}y &= 1/3\\ k\int_{\{(x,y)\in S\}} 1\mathrm{d}x\mathrm{d}y &= 1/3\\ k * Area_S &= 1/3\\ k * 96 &= 1/3\end{aligned}$$

So $k = 1/288$. Thus, (with S defined as above),

$$f_{X,Y}(x,y) = \begin{cases} 1/6 & \text{if } 4 \leq x \leq 6, 6 \leq y \leq 8 \\ 1/288 & \text{if } (x,y) \in S \\ 0 & \text{otherwise} \end{cases}$$

B. Height of spider on wall

The pdf of the height is just the marginal pdf $f_Y(y)$,

$$\begin{aligned} f_Y(y) &= \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x \\ &= \int_0^{10} f_{X,Y}(x,y)\mathrm{d}x \end{aligned}$$

since $f_{X,Y}(x,y) = 0$ for $0 < x, 10 < x$.
Then, for $0 \leq y < 6$,

$$f_Y(y) = \int_0^{10} f_{X,Y}(x,y)\mathrm{d}x = \int_0^{10} 1/288\mathrm{d}x = 10/288$$

For $6 \leq y \leq 8$,

$$\begin{aligned} f_Y(y) &= \int_0^{10} f_{X,Y}(x,y)\mathrm{d}x \\ &= \int_0^{4} f_{X,Y}(x,y)\mathrm{d}x + \int_4^{6} f_{X,Y}(x,y)\mathrm{d}x + \int_6^{10} f_{X,Y}(x,y)\mathrm{d}x \\ &= \int_0^{4} 1/288\mathrm{d}x + \int_4^{6} 1/6\mathrm{d}x + \int_6^{10} 1/288\mathrm{d}x \\ &= 4/288 + 2/6 + 4/288 = 104/288 \end{aligned}$$

For $8 < y \leq 10$,

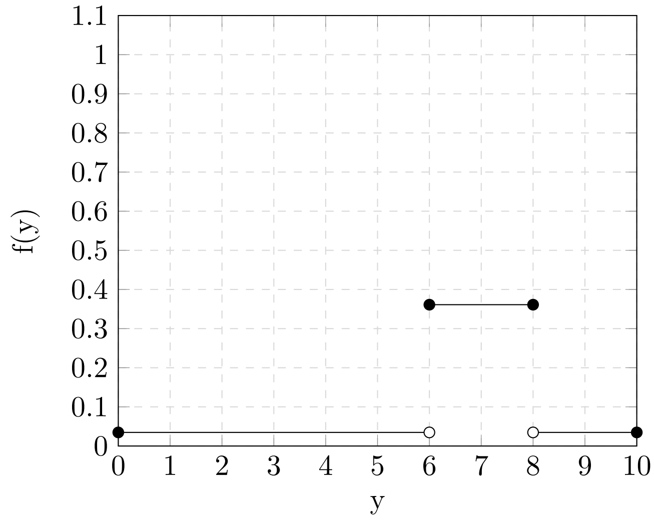
$$f_Y(y) = \int_0^{10} f_{X,Y}(x,y)\mathrm{d}x = \int_0^{10} 1/288\mathrm{d}x = 10/288$$

Thus,

$$f_Y(y) = \begin{cases} 10/288 & \text{if } 0 \leq y < 6 \\ 104/288 & \text{if } 6 \leq y \leq 8 \\ 10/288 & \text{if } 8 < y \leq 10 \end{cases}$$

This pdf is plotted below[1]:

[1] This plot was generated using code adapted from "Graphing in LaTeXusing PGF and TikZ" by Lauderdale and Gluck



c. CDF of height, given spider is visible

Recall the set S is the visible area of the wall, and that $P((x,y) \in S) = 1/3$. The conditional cdf of the height Y, given we see the spider (i.e. $Y \in S$), is given by:

$$f_{Y|Y \in S}(u) = \frac{\int_{y=0}^{u} \int_{\{x|(x,y) \in S\}} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y}{P((x,y) \in S)} = \frac{\int_{y=0}^{u} \int_{\{x|(x,y) \in S\}} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y}{1/3}$$

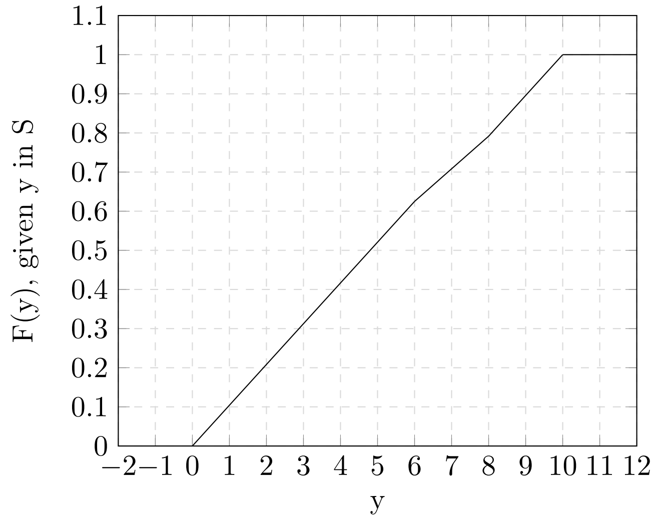
Thus

$$f_{Y|Y \in S}(u) = \begin{cases} \frac{\int_0^u f_Y(y)\mathrm{d}y}{1/3} & \text{if } 0 \le u < 6 \\ \frac{\int_0^6 f_Y(y)\mathrm{d}y + \int_6^u \int_0^4 f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y + \int_6^u \int_6^{10} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y}{1/3} & \text{if } 6 \le u \le 8 \\ \frac{\int_0^6 f_Y(y)\mathrm{d}y + \int_6^8 \int_0^4 f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y + \int_6^8 \int_6^{10} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y + \int_8^u f_Y(y)\mathrm{d}y}{1/3} & \text{if } 8 < u \le 10 \end{cases}$$

Evaluating these integrals yields

$$f_{Y|Y \in S}(u) = \begin{cases} \frac{10/288u}{96/288} = 10/96u & \text{if } 0 \le u < 6 \\ \frac{60/288 + 4/288(u-6) + 4/288(u-6)}{1/3} = \frac{8/288u + 12/288}{96/288} = 8/96u + 12/96 & \text{if } 6 \le u \le 8 \\ \frac{76/288 + 10/288(u-8)}{96/288} = 10/96u - 4/96 & \text{if } 8 < u \le 10 \end{cases}$$

This conditional cdf is plotted below:



2. Pizza Delivery

a. Wait time- Pat or Robbie get called

Let $X_P, X_R \sim Exp(\lambda)$ model the wait time until Pat and Robbie get calls, respectively. Then the wait time until one of them gets a call is $Z = \min\{X_P, X_R\}$. Then,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = 1 - P(z < Z) \\ &= 1 - P(z < X_P, z < X_R) \end{aligned}$$

Now, let's assume X_R and X_P are independent. This is a reasonable assumption if the market is large. At the lower limit for market size, imagine there is only a single customer who wants one pizza. Then, if they call Pat, they don't call Robbie (and vice versa), so obviously X_R and X_P are not independent. On the other hand, if the market is large and customers act independently of each other, X_R and X_P can reasonably be assumed to be independent. Then

$$\begin{aligned} F_Z &= 1 - P(z < X_P, z < X_R) \\ &= 1 - P(z < X_P)P(z < X_R) \\ &= 1 - (1 - F_{X_P}(z))(1 - F_{X_R}(z)) \\ &= 1 - (1 - F_{X_P}(z) - F_{X_R}(z) + F_{X_P}(z)F_{X_R}(z)) \\ &= F_{X_P}(z) + F_{X_R}(z) - F_{X_P}(z)F_{X_R}(z) \\ &= (1 - e^{-\lambda z}) + (1 - e^{-\lambda z}) - (1 - e^{-\lambda z})(1 - e^{-\lambda z}) \\ &= 1 - e^{-2\lambda z} \end{aligned}$$

Thus, $Z \sim Exp(2\lambda)$. This makes intuitive sense- if you expect calls to come into each store at a rate of λ, then you expect the combined rate of calls into both stores to be 2λ.

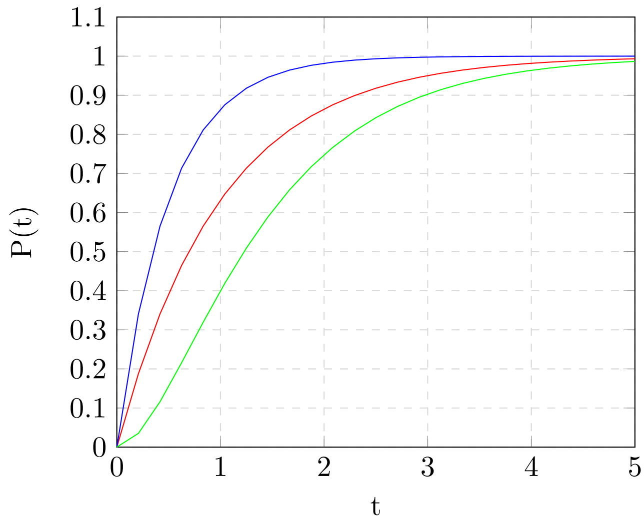
B. Wait time- Pat and Robbie get called

The distribution of the time until both Pat and Robbie have received a call is modeled by $W = \max\{X_R, X_P\}$. Thus,

$$\begin{aligned}
F_W(W) &= P(W \leq w) \\
&= P(X_P \leq w, X_R \leq w) \\
&= P(X_P \leq w)P(X_R \leq w) \text{ (again assuming independence)} \\
&= F_{X_P}(w)F_{X_R}(w) \\
&= 1 - 2e^{-\lambda w} + e^{-2\lambda w}
\end{aligned}$$

C. Comparing distributions

Color-coded plots of the CDFs (with $\lambda = 1$) are given below, with red for $F_X(t)$, blue for $F_Z(t)$, and green for $F_W(t)$.



These results make sense, when you observe the functions are ordered over all $0 < t$. Specifically, $F_Z(t) > F_X(t) > F_W(t)$. That's expected since, during any time interval $(0, t)$, the event {Pat gets a call $\cup$ Robbie gets a call} $\supseteq$ {Pat gets a call} $\supseteq$ {Pat gets a call $\cap$ Robbie gets a call} (note this is obviously also true if we swap the names "Pat" and "Robbie").

3. Cars

A. Proof of Equation (1)

To start, recall $P(\Omega) = 1$. Now let $X \sim B(n, 0.5)$. Then,

$$
\begin{aligned}
\sum_{k=0}^{n}\binom{n}{k}(1/2)^k(1/2)^{n-k} &= 1\\
\sum_{k=0}^{n}\binom{n}{k}(1/2)^n &= 1\\
(1/2)^n\sum_{k=0}^{n}\binom{n}{k} &= 1\\
\sum_{k=0}^{n}\binom{n}{k} &= 2^n
\end{aligned}
$$

B. PMF of the combined total number of mechanical problems

Let X_1, X_2 be IID poisson RVs, with parameter λ, modeling the number of mechanical failures in your car and your sister's car, respectively.
Then $Y = X_1 + X_2$ is the total number of mechanical failures. Then

$$
\begin{aligned}
P(Y=y) &= P(X_1+X_2=y)\\
&= \sum_{x_2=0}^{y} P(X_1 = y - x_2 | X_2 = x_2)P(X_2 = x_2)\\
&= \sum_{x_2=0}^{y} P(X_1 = y - x_2)P(X_2 = x_2) \text{ (by independence)}\\
&= \sum_{x_2=0}^{y} \frac{\lambda^{y-x_2}}{(y-x_2)!}e^{-\lambda}\frac{\lambda^{x_2}}{(x_2)!}e^{-\lambda}\\
&= \sum_{x_2=0}^{y} \frac{y!}{x_2!(y-x_2)!}\lambda^y e^{-2\lambda}\frac{1}{y!}\\
&= \lambda^y e^{-2\lambda}\frac{1}{y!}\sum_{x_2=0}^{y}\frac{y!}{x_2!(y-x_2)!} = \lambda^y e^{-2\lambda}\frac{1}{y!}\sum_{x_2=0}^{y}\binom{y}{x_2}\\
&= \lambda^y e^{-2\lambda}\frac{1}{y!}2^y\\
&= \frac{(2\lambda)^y}{y!}e^{-2\lambda}
\end{aligned}
$$

Thus, $Y \sim Pois(2\lambda)$. Again, this makes intuitive sense- the Poisson distribution models the number of events occurring in a time interval, given a known rate λ. If we're adding two IID Poisson RVs, we'd expect this rate to double, but the underlying generating process to be otherwise unchanged.

4. Race

a. Joint PDF of the positions of Mary and Hannah

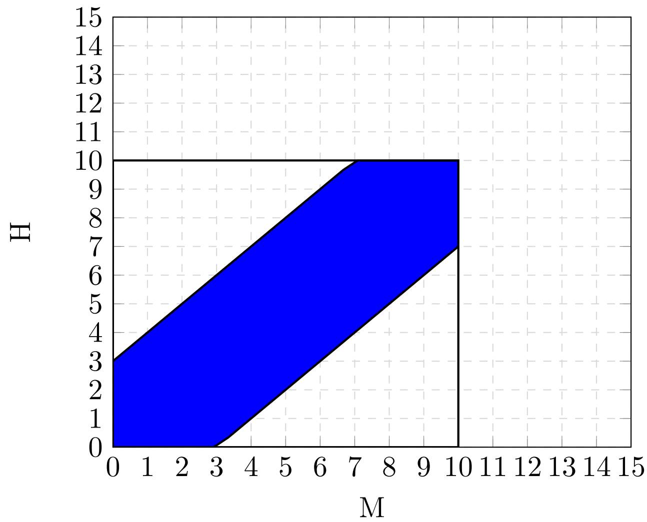
First, let $M \sim U(0, 10)$ and $H \sim U(0, 10)$ represent the locations of Mary and Hannah respectively. Since M and H are IID, $f_{M,H}(m, h) = f_M(m) f_H(h)$. Thus

$$f_{M,H}(m,h) = \begin{cases} \frac{1}{10-0}\frac{1}{10-0} = \frac{1}{100} & \text{if } 0 \leq H \leq 10, 0 \leq M \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

This joint PDF is shown below (in section B). The boxed area represents $F_{M,H}(m, h)$ while the blue area represents the area in which $D \leq d = 3$ (defined and discussed below).

b. Shading area corresponding to $P(D \leq d)$

First, let D represent the distance between Mary and Hannah. To illustrate the area corresponding to $D \leq d$, let's let $d = 3$. The corresponding area is shaded blue in the plot below.



c. Finding the PDF $F_D(d)$

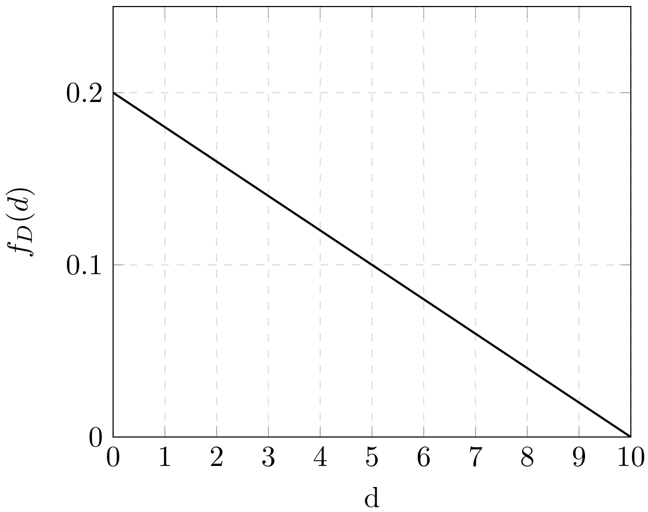
First, notice the area in $[0, 10] \times [0, 10]$ not covered by $D \leq d$ is $(10 - d)^2$. Thus,

$$F_D(d) = P(D \leq d) = 1 - P(D > d) = 1 - 1/100 * (10 - d)^2$$

Thus

$$f_D(d) = F'_D(d) = 1/5 - 1/50d$$

This CDF is plotted below:



Again, as a side note, observe this PDF makes intuitive sense- given M and H are IID uniform, it makes sense $f_D(d)$ has a maximum at $d = 0$.

5. Cheating at coin flips

a. Probability of heads/tails

First, let $X \sim$ Bernouilli(P) model the outcome of the coin flip. Let

$$X = \begin{cases} 1 & \text{if heads} \\ 0 & \text{if tails} \end{cases}$$

Then

$$P_X(x) = \begin{cases} (1-P) & \text{for x =0} \\ P & \text{for x= 1} \end{cases}$$

Since we suspect but do not know that the coin is biased, P is itself a $U(0.5, 1)$ random variable. This is reasonable- if the coin is biased, we can assume it is biased in Marvin's favor (so $0.5 \le P$), and (obviously) $P \le 1$. The probability the coin flip is heads under this model is:

$$\begin{aligned} P(X=1) &= \int_{0.5}^{1} P(X=1|P=p) f_P(p) \mathrm{d}p \\ &= \int_{0.5}^{1} p \cdot f_P(p) \mathrm{d}p \\ &= \int_{0.5}^{1} 2p \mathrm{d}p = p^2 \Big|_{p=0.5}^{1} = 1 - 0.25 = 0.75 \end{aligned}$$

Thus, under this model, the probability of heads is 0.75 and the probability of tails is 0.25. As a side note, this again makes intuitive sense- X is essentially an indicator variable, so $P(X=1)$ is just the expected value $\mathrm{E}[P] = 0.75$.

B. Updating our model

We will use Bayes rule to update our model. If the coin flip is heads ($X = 1$), then

$$f_{P|X}(p|1) = \frac{P(X=1|P=p) \cdot f_P(p)}{\int_{0.5}^{1} P(X=1|P=p) \cdot f_P(p) \mathrm{d}p}$$

Recall $f_P(p) = 2$ for $0.5 \leq p \leq 1$, $\int_{0.5}^{1} P(X=1|P=p) \cdot f_P(p) \mathrm{d}p = 0.75$, and $P(X=1|P=p) = p$. Thus, substitution yields

$$f_{P|X}(p|1) = \frac{p \cdot 2}{0.75} = 8/3p$$

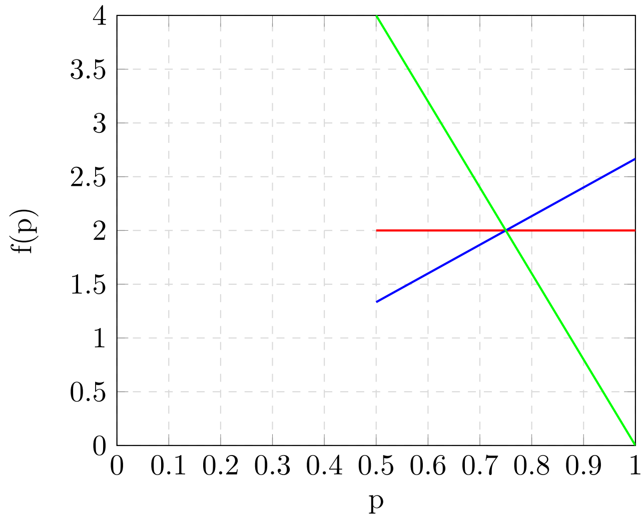
Next, if the coin flip is tails, ($X = 0$), then

$$\begin{aligned} f_{P|X}(p|0) &= \frac{P(X=0|P=p) \cdot f_P(p)}{\int_{0.5}^{1} P(X=0|P=p) \cdot f_P(p) \mathrm{d}p} \\ &= \frac{(1-p) \cdot 2}{0.25} \\ &= 8(1-p) \end{aligned}$$

Thus, the conditional PDF $f_{P|X}(p|x)$ is

$$f_{P|X}(p|x) = \begin{cases} 8/3p & \text{for x = 1 (flip 1 is heads)} \\ 8(1-p) & \text{for x = 0 (flip 1 is tails)} \end{cases}$$

Below are color-coded plots of these posterior PDFs, and the prior PDF $f_P(p)$, with red for $f_P(p)$, blue for $f_{P|X}(p|1)$, and green for $f_{P|X}(p|0)$.



These posterior PDFs make sense. If our first flip returns a head, then we become more confident the coin is in fact biased. On the other hand, if our first flip returns a tail, then we are less confident the coin is biased. To see this more clearly, we can consider the expected value of the two posterior distributions:

$$\mathrm{E}[P|X=0] = \int_{0.5}^{1} p \cdot 8(1-p)\mathrm{d}p = 2/3 = 0.\bar{6}$$
$$\mathrm{E}[P|X=1] = \int_{0.5}^{1} p \cdot 8/3p\mathrm{d}p = 7/9 = 0.\bar{7}$$

Recall the our prior expectation $\mathrm{E}[P] = 0.75$. Thus, if the first flip is a tails, we revise our expected value for the bias downwards considerably. On the other hand, if the first flip is a heads, we revise our expected value for the bias up only slightly. This is also visible in the plots of the PDFs- the change in the slope of the PDF if tails is much greater than the change in the slope if heads.