W271-2 – Spring 2016 – **Lab 1**

Juanjo Carin, Kevin Davis, Ashley Levato, Minghu Song

January 14, 2016

Contents

Part I: Marginal, Joint, and Conditional Probabilities **2**

Question 1 2

Question 2 3

Part II: Random Variables, Expectation, Conditional Exp. **5**

Question 3 5

Question 4 10

Question 5 11

Question 6 12

Question 7 13

Part III: Statistical Estimation and Statistical Inference **13**

Question 8 13

Question 9 14

Question 10 14

Question 11 14

Question 12 14

Question 13 14

Part I: Marginal, Joint, and Conditional Probabilities

Figure 1: Venn Diagram

Question 1

In a team of data scientists, 36 are expert in machine learning, 28 are expert in statistics, and 18 are awesome. 22 are expert in both machine learning and statistics, 12 are expert in machine learning and are awesome, 9 are expert in statistics and are awesome, and 48 are expert in either machine learning or statistics or are awesome. Suppose you are in a cocktail party with this group of data scientists and you have an equal probability of meeting any one of them.

1. **What is the probability of meeting a data scientist who is an expert in both machine learning and statistics and is awesome?**

Let N be the size of the team of data scientists (i.e., the sample size), and M, S, A the event that a data scientist is either a machine learning expert, a statistics expert, or awesome.

$$
\begin{aligned}
\Pr(\mathbf{M} \cap \mathbf{S} \cap \mathbf{A}) &= \Pr(M) + \Pr(S \cap A) - \Pr(M \cup (S \cap A)) \\
&= \Pr(M) + \Pr(S \cap A) - \Pr((M \cup S) \cap (M \cup A))
\end{aligned}
$$

$$= \Pr(M) + \Pr(S \cap A) - ((\Pr(M \cup S) + \Pr(M \cup A) - \Pr((M \cup S) \cup (M \cup A))))$$
$$= \Pr(M) + \Pr(S \cap A) - (\Pr(M) - \Pr(S) + \Pr(M \cap S)) - (\Pr(M) - \Pr(A) + \Pr(M \cap A)) + \Pr(M \cup S \cup A)$$
$$= \Pr(M \cap S) + \Pr(M \cap A) + \Pr(M \cap A) - \Pr(M) - \Pr(S) - \Pr(A) + \Pr(M \cup S \cup A)$$
$$= \frac{22}{N} + \frac{12}{N} + \frac{9}{N} - \frac{36}{N} - \frac{28}{N} - \frac{18}{N} + \frac{48}{N} = \frac{(22+12+9)-(36-28-18)+48}{N} = \frac{43-82+48}{N}$$
$$= \mathbf{\frac{9}{N}}$$

2. **Suppose you meet a data scientist who is an expert in machine learning. Given this information, what is the probability that s/he is not awesome?**

$$\Pr(\mathbf{A^c}|\mathbf{M}) = 1 - \Pr(A|M) = 1 - \frac{\Pr(A \cap M)}{\Pr(M)} = 1 - \frac{\frac{12}{N}}{\frac{36}{N}} = 1 - \frac{12}{36} = 1 - \frac{1}{3} = \mathbf{\frac{2}{3} = 0.6667}$$

3. **Suppose the you meet a data scientist who is awesome. Given this information, what is the probability that s/he is an expert in either machine learning or statistics?**

$$\Pr(\mathbf{M} \cup \mathbf{S}|\mathbf{A}) = \frac{\Pr((M \cup S) \cap A)}{\Pr(A)} = \frac{\Pr((M \cap A) \cup (S \cap A))}{\Pr(A)}$$
$$= \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr((M \cap A) \cap (S \cap A)}{\Pr(S)} = \frac{\Pr(M \cap A) + \Pr(S \cap A) - \Pr(M \cap S \cap A)}{\Pr(S)}$$
$$= \frac{\frac{12}{N} + \frac{9}{N} - \frac{9}{N}}{\frac{18}{N}} = \frac{12+9-9}{18} = \frac{12}{18}$$
$$= \mathbf{\frac{2}{3} = 0.6667}$$

Question 2

Suppose for events A and B, $\Pr(A) = p \leq \frac{1}{2}, \Pr(B) = q$, where $\frac{1}{4} < q < \frac{1}{2}$. These are the only information we have about the events.

1. **What are the maximum and minimum possible values for $\Pr(A \cup B)$?**

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

The maximum value of $\Pr(A \cup B)$ occurs when $A \cap B$ is the smallest set possible. In this case, since $\Pr(A) \leq \frac{1}{2}$ and $\Pr(B) < \frac{1}{2}$, it would be $A \cap B = \varnothing$ (if A and B were disjoint sets, which might be the case), so:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = \Pr(A) + \Pr(B) - \Pr(\varnothing) = p + q - 0 = p + q$$

which could have a maximum value close to 1 (i.e., $A \cup B \approx \Omega$), in case $\Pr(A) = \frac{1}{2}$ and $\Pr(B) \approx \frac{1}{2}$.

The minimum value of $\Pr(A \cup B)$ occurs when $A \cup B$ is the largest set possible, A or B. In this case, since $\Pr(A)$ does not have a lower bound, that would happen when $A \subseteq B$ and (consequently) $A \cap B = A$, which would lead to:

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) = \Pr(B) = q$$

whose minimum value is greater than $\frac{1}{4}$.

In summary,

$$\mathbf{\frac{1}{4}} < \Pr(\mathbf{A} \cup \mathbf{B}) < \mathbf{1}$$

2. **What are the maximum and minimum possible values for $\Pr(A|B)$?**

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

If $B \subseteq A$ (which would imply that the lower bound for p is also $\frac{1}{4}$), then:

$$\Pr(A|B) = \frac{\Pr(B)}{\Pr(B)} = 1$$

As seen in the previous part, since $\Pr(A) \leq \frac{1}{2}$ and $Pr(B) < \frac{1}{2}$, it might occur that $A \cap B = \varnothing$, and hence:

$$\Pr(A|B) = \frac{\Pr(\varnothing)}{\Pr(B)} = \frac{0}{q} = 0$$

irrespective of the value of q.

In summary,

$$\mathbf{0} \leq \Pr(\mathbf{A}|\mathbf{B}) \leq \mathbf{1}$$

Part II: Random Variables, Expectation, Conditional Exp.

Question 3

Suppose the life span of a particular server is a continuous random variable, t, with a uniform probability distribution between 0 and k year, where $k \leq 10$ is a positive integer.

The server comes with a contract that guarantees a full or partial refund, depending on how long it lasts. Specifically, if the server fails in the first year, it gives a full refund denoted by θ. If it lasts more than 1 year but fails before $\frac{k}{2}$ years, the manufacturer will pay $x = \$A(k-t)^{1/2}$, where A is some positive constant equal to 2 if $t \leq \frac{k}{2}$. If it lasts between $\frac{k}{2}$ and $\frac{3k}{4}$ years, it pays $\frac{\theta}{10}$.

1. **Given that the server lasts for $\frac{k}{4}$ years without failing, what is the probability that it will last another year?**

$$\Pr\left(\mathbf{t \leq 1 + \frac{k}{4} \middle| t \geq \frac{k}{4}}\right) = \frac{\Pr\left(\left(t \leq 1 + \frac{k}{4}\right) \bigcap \left(t \geq \frac{k}{4}\right)\right)}{\Pr\left(t \geq \frac{k}{4}\right)} = \frac{\Pr\left(\frac{k}{4} \leq t \leq 1 + \frac{k}{4}\right)}{Pr\left(t \geq \frac{k}{4}\right)}$$

$$= \frac{\Pr\left(t \leq 1 + \frac{k}{4}\right) - \Pr\left(t \leq \frac{k}{4}\right)}{1 - \Pr\left(t \leq \frac{k}{4}\right)} = \frac{\frac{1}{k}\left(1 + \frac{k}{4} - \frac{k}{4}\right)}{1 - \frac{1}{k}\frac{k}{4}} = \frac{\frac{1}{k}}{\frac{3}{4}}$$

$$= \mathbf{\frac{4}{3k}}$$

2. **Compute the expected payout from the contract, $E(x)$.**

I tried 3 different approaches, all leading to the same result.

The first one is based on https://www.probabilitycourse.com/chapter4/4_3_1_mixed.php. Before computing the expected value of X, we need to compute its *cdf*.

t (the life span of the server) is a continuous variable with the following *pdf*:

$$f_t(t) = \begin{cases} \frac{1}{k} & 0 \leq t \leq \ k \\ 0 & \text{otherwise} \end{cases}$$

But X (the payout or refund) is not continuous:

$$X = g(t) = \begin{cases} \theta & 0 \leq t \leq \ 1 \\ A\sqrt{k-t} & 1 \leq t \leq \ \frac{k}{2} \\ \frac{\theta}{10} & \frac{k}{2} \leq t \leq \ \frac{3k}{4} \\ 0 & \text{otherwise} \end{cases}$$

(The value of X for $t = 1$ is $A\sqrt{k-1}$, which is not necessarily equal to θ; the value of X for $t = k/2$ is $A\sqrt{k/2}$, which again is not necessarily equal to $\theta/10$.)

First we compute the probability that X takes its possible discrete values:

$$\Pr(X = 0) = \Pr\left(t \geq \frac{3k}{4}\right) = 1 - F_t\left(\frac{3k}{4}\right) = 1 - \frac{1}{k}\frac{3k}{4} = 1 - \frac{3}{4} = \frac{1}{4}$$

$$\Pr\left(X = \frac{\theta}{10}\right) = \Pr\left(\frac{k}{2} \leq t \leq \frac{3k}{4}\right) = F_t\left(\frac{3k}{4}\right) - F_t\left(\frac{k}{2}\right) = \frac{1}{k}\frac{3k}{4} - \frac{1}{k}\frac{k}{2} = \frac{3}{4} - \frac{1}{2} = \frac{3-2}{4} = \frac{1}{4}$$

$$\Pr(X = \theta) = \Pr(0 \leq t \leq 1) = F_t(1) - F_t(0) = \frac{1}{k} - \frac{0}{k} = \frac{1}{k}$$

For $A\sqrt{\frac{k}{2}} \leq X \leq A\sqrt{k-1}$, we can compute the *cdf* of X as follows:

$$\begin{aligned}
F_X(x) &= \Pr(X \leq x) = \Pr\left(A\sqrt{k-t} \leq x\right) \\
&= \Pr\left(k - t \leq \left(\tfrac{x}{A}\right)^2\right) = \Pr\left(t \geq k - \left(\tfrac{x}{A}\right)^2\right) = 1 - \Pr\left(t \leq k - \left(\tfrac{x}{A}\right)^2\right) \\
&= 1 - \int_{t=0}^{k-\left(\frac{x}{A}\right)^2} \frac{1}{k} dt = 1 - \frac{1}{k}\left(k - \left(\frac{x}{A}\right)^2\right) = \frac{x^2}{kA^2}
\end{aligned}$$

So the overall expression for $F_X(x)$ is:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x < \frac{\theta}{10} \\ \frac{1}{2} & \frac{\theta}{10} \leq x \leq \ A\sqrt{\frac{k}{2}} \\ \frac{x^2}{kA^2} & A\sqrt{\frac{k}{2}} \leq x \leq \ A\sqrt{k-1} \\ 1 - \frac{1}{k} & A\sqrt{k-1} \leq x < \theta \\ 1 & x \geq \theta \end{cases}$$

Let's check that:

$$\int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} \frac{dF_X(x)}{dx} dx + \sum_{x_k} \Pr(X = x_k) = 1$$

$$\int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} \frac{2x}{kA^2} dx + \Pr(X = 0) + \Pr\left(X = \frac{\theta}{10}\right) + \Pr(X = \theta) = \left[\frac{x^2}{kA^2}\right]_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} + \frac{1}{4} + \frac{1}{4} + \frac{1}{k}$$

$$= \frac{(k-1) - \frac{k}{2}}{k} + \frac{1}{2} + \frac{1}{k} = \frac{\frac{k}{2} - 1}{k} + \frac{1}{2} + \frac{1}{k} = \frac{1}{2} - \frac{1}{k} + \frac{1}{2} + \frac{1}{k} = 1$$

Now we can compute the expected value of X as:

$$\begin{aligned}
\mathbf{E(X)} &= \int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} x\frac{dF_X(x)}{dx} dx + \sum_{x_k} x_k \Pr(X = x_k) \\
&= \int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} \frac{2x^2}{kA^2} dx + 0 \cdot \Pr(X = 0) + \frac{\theta}{10} \cdot \Pr\left(X = \frac{\theta}{10}\right) + \theta \cdot \Pr(X = \theta)
\end{aligned}$$

$$= \left[\frac{2x^2}{3kA^2} \right]_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} + 0 \cdot \frac{1}{4} + \frac{\theta}{10} \cdot \frac{1}{4} + \theta \cdot \frac{1}{k} = \mathbf{\frac{2A}{3k} \left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2} \right)^{\frac{3}{2}} \right) + \frac{\theta}{40} + \frac{\theta}{k}}$$

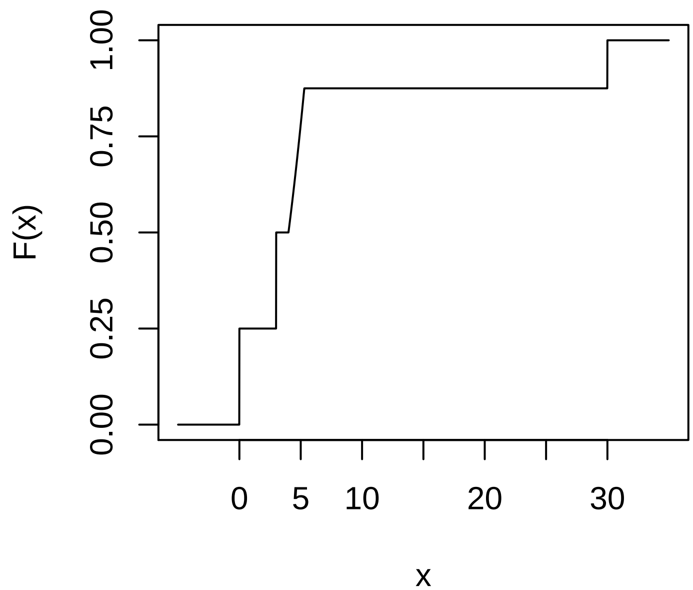

Figure 2: Approximate aspect of $F_X(x)$ (for $A = 2, k = 8, \theta = 30$)

Another way would be just applying the *law of the unconscious statistician* (https://en.wikipedia.org/wiki/Law_of_the_unconscious_statistician):

$$E\left[g(t)\right] = \int g(t) f_t(t) dt$$

$$\mathbf{E(X)} = \int_{t=0}^{k} g(t) f_t(t) dt$$

$$= \int_{t=0}^{1} \frac{g(x)}{k} dt + \int_{t=1}^{k/2} \frac{g(x)}{k} dt + \int_{t=k/2}^{3k/4} \frac{g(x)}{k} dt + \int_{t=3k/4}^{k} \frac{g(x)}{k} dt$$

$$= \frac{\theta}{k} [t]_{t=0}^{1} + \frac{A}{k} \left[\left(-\frac{2}{3} \right) (k-t)^{\frac{3}{2}} \right]_{t=1}^{k/2} + \frac{\theta/10}{k} [t]_{t=k/2}^{3k/4} + \frac{0}{k} [t]_{t=3k/4}^{k}$$

$$= \frac{\theta}{k} + \frac{2A}{3k} \left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2} \right)^{\frac{3}{2}} \right) + \frac{\theta}{10k} \frac{k}{4} + 0 = \mathbf{\frac{2A}{3k} \left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2} \right)^{\frac{3}{2}} \right) + \frac{\theta}{40} + \frac{\theta}{k}}$$

The last approach is based on http://homepage.stat.uiowa.edu/~nshyamal/22S175/DI.pdf. A *mixed distribution* can be decomposed into a continuous distribution and a discrete one:

$$\mathbf{F(x) = F_C(x) + F_D(x)}$$

each with weights p_C and p_D, respectively.

p_D is the sum of the heights of the jumps:

$F(0) - F(0^-) = \frac{1}{4}$

$F\left(\frac{\theta}{10}\right) - F\left(\frac{\theta^-}{10}\right) = \frac{1}{4}$

$F(\theta) - F(\theta^-) = \frac{1}{k}$

$$\Rightarrow p_D = \frac{1}{2} + \frac{1}{k} = \frac{k+2}{2k} \Rightarrow \mathbf{p_C = 1 - p_D} = \frac{k-2}{2k}$$

$$\Pr(X_D = x) = \begin{cases} \dfrac{\frac{1}{4}}{\frac{k+2}{2k}} = \dfrac{1}{4}\dfrac{2k}{k+2} = \dfrac{k}{2(k+2)} & x = 0 \\ \dfrac{\frac{1}{4}}{\frac{k+2}{2k}} = \dfrac{1}{4}\dfrac{2k}{k+2} = \dfrac{k}{2(k+2)} & x = \dfrac{\theta}{10} \\ \dfrac{\frac{1}{k}}{\frac{k+2}{2k}} = \dfrac{1}{k}\dfrac{2k}{k+2} = \dfrac{2}{k+2} & x = \theta \end{cases}$$

$$\mathbf{F_C(x)} = \frac{\mathbf{F(x) + p_D \cdot F_D(x)}}{\mathbf{p_C}}$$

$$\Rightarrow F_C(X_C = x) = \begin{cases} 0 & x < A\sqrt{\dfrac{k}{2}} \\ \dfrac{2x^2 - kA^2}{A^2(k-2)} & A\sqrt{\dfrac{k}{2}} \le x < A\sqrt{k-1} \\ 1 & x \ge A\sqrt{k-1} \end{cases}$$

$$\Rightarrow f_C(X_C = x) = \begin{cases} 0 & x < A\sqrt{\dfrac{k}{2}} \\ \dfrac{4x}{A^2(k-2)} & A\sqrt{\dfrac{k}{2}} \le x < A\sqrt{k-1} \\ 0 & x \ge A\sqrt{k-1} \end{cases}$$

$$E(X_D) = \sum_{x_k} x_k \Pr(X_D = x_k) = 0 \cdot \frac{k}{2(k+2)} + \frac{\theta}{10} \cdot \frac{k}{2(k+2)} + \theta \cdot \frac{2}{k+2}$$

$$= \frac{k\theta}{20(k+2)} + \frac{2\theta}{k+2} = \frac{(k+40)\theta}{20(k+2)}$$

$$E(X_C) = \int_{x=-\infty}^{\infty} x f_C(x) dx = \int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} \frac{4x^2}{A^2(k-2)} dx = \left[\frac{4x^3}{3A^2(k-2)}\right]_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} = \frac{4A}{3(k-2)}\left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2}\right)^{\frac{3}{2}}\right)$$

$$\mathbf{E(X) = p_D \cdot E(X_D) + p_C \cdot E(X_C)}$$

$$\mathbf{E(X)} = \frac{k+2}{2k}\cdot\frac{(k+40)\theta}{20(k+2)} + \frac{k-2}{2k}\cdot\frac{4A}{3(k-2)}\left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2}\right)^{\frac{3}{2}}\right)$$

$$= \mathbf{\frac{2A}{3k}\left((k-1)^{\frac{3}{2}} - \left(\frac{k}{2}\right)^{\frac{3}{2}}\right) + \frac{\theta}{40} + \frac{\theta}{k}}$$

3. **Compute the variance of the payout from the contract.**

The 3rd link in (2) mentions the expression for the variance of a mixed random variable:

$$\mathbf{Var(X) = (p_D \cdot Var(X_D) + p_C \cdot Var(X_C)) + p_D \cdot p_C \cdot (E(X_D) - E(X_C))^2}$$

We already computed $p_D, p_C, E(X_D)$, and $E(X_C)$. We'll compute now the variances of both components of X, omitting the final expression of $Var(X)$ (which is really complex and with a lot of terms):

$$Var(X_D) = \sum_{x_k}(x_k - E(X_D))^2 \Pr(x_k)$$

$$= \left(0 - \frac{(k+40)\theta}{20(k+2)}\right)^2 \frac{k}{2(k+2)} + \left(\frac{\theta}{10} - \frac{(k+40)\theta}{20(k+2)}\right)^2 \frac{k}{2(k+2)} + \left(\theta - \frac{(k+40)\theta}{20(k+2)}\right)^2 \frac{2}{k+2}$$

$$= \frac{(k+40)^2\theta^2}{400(k+2)^2}\cdot\frac{k}{2(k+2)} + \left(\frac{2\theta(k+2)-\theta(k+40)}{20(k+2)}\right)^2\cdot\frac{k}{2(k+2)} + \left(\frac{20\theta(k+2)-\theta(k+40)}{20(k+2)}\right)^2\cdot\frac{2}{k+2}$$

$$= \frac{k\theta^2(k+40)^2}{800(k+2)} + \frac{4k\theta^2(k+2)^2 + k\theta^2(k+40)^2 - 4k\theta^2(k+2)(k+40)}{800(k+2)^2}$$

$$+ \frac{1600\theta^2(k+2)r + 4\theta^2(k+40)^2 - 80\theta^2(k+2)(k+40)}{800(k+2)^3}$$

$$= \frac{2\theta^2(k+2)(k+40)^2 + 4\theta^2(k+400)(k+2)^2 - 4\theta^2(k+20)(k+2)(k+40)}{800(k+2)^3}$$

$$= \frac{\theta^2(k+40)((k+40)-2(k+20))}{400(k+2)^2} + \frac{\theta^2(k+400)}{200(k+2)}$$

$$= \frac{\theta^2(k+400)}{200(k+2)} - \frac{k\theta^2(k+40)}{400(k+2)^2} = \frac{\theta^2((k+400)(k+2) - k(k+40))}{400(k+2)^2}$$

$$= \frac{\theta^2(k^2+2k+400k+800-k^2-40k)}{400(k+2)^2} = \frac{\theta^2(362k+800)}{400(k+2)^2} = \frac{\theta^2(181k+400)}{200(k+2)^2}$$

$$Var(X_C) = \int_{x=-\infty}^{\infty} x^2 f_C(x)dx - (E(X_C))^2$$

$$= \int_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} \frac{4x^3}{A^2(k-2)}dx - (E(X_C))^2 = \left[\frac{x^4}{A^2(k-2)}\right]_{x=A\sqrt{k/2}}^{A\sqrt{k-1}} - (E(X_C))^2$$

$$= \frac{A^4(k-1)^2 - A^4\dfrac{k^2}{4}}{A^2(k-2)} - (E(X_C))^2 = \frac{A^2(k^2+1-2k-\dfrac{k^2}{4}}{k-2} - (E(X_C))^2$$

$$= \frac{A^2(3k^2+4-8k)}{4(k-2)} - \frac{16A^2}{9(k-2)^2}\left((k-1)^3 + \frac{k}{2} - 2\left(\frac{k}{2}\right)^{\frac{3}{2}}(k-1)^{\frac{3}{2}}\right)$$

Question 4

Continuous random variables X and Y have a joint distribution with probability density function $f(x,y) = 2e^{-x}e^{-2y}$ for $0 < x < \infty, 0 < y < \infty$ and 0 otherwise.

1. **Compute $\Pr(X > a, Y < b)$, where a, b are positive constants and $a < b$.**

$$\Pr(X > a, Y < b) = \int_{x=a}^{\infty} \int_{y=0}^{b} f(x,y) dx dy$$

$$= \int_{x=a}^{\infty} \int_{y=0}^{b} 2e^{-x}e^{-2y} dx dy = 2 \left(\int_{x=a}^{\infty} e^{-x} dx \right) \left(\int_{y=0}^{b} e^{-2y} dy \right)$$

$$= 2 \left[-e^{-x} \right]_{x=a}^{\infty} \left[-\frac{e^{-2y}}{2} \right]_{y=0}^{b} = \left[e^{-x} \right]_{x=a}^{\infty} \left[e^{-2y} \right]_{y=0}^{b} = \left(0 - e^{-a} \right) \left(e^{-2b} - 1 \right)$$

$$= \mathbf{e^{-a} \left(1 - e^{-2b} \right) = e^{-a} - e^{-a-2b}}$$

2. **Compute $\Pr(X < Y)$.**

$$\Pr(X < Y) = \int_{x=0}^{y} \int_{y=0}^{\infty} f(x,y) dx dy$$

$$= \int_{x=0}^{y} \int_{y=0}^{\infty} 2e^{-x}e^{-2y} dx dy = 2 \int_{y=0}^{\infty} \left(\int_{x=0}^{y} e^{-x} dx \right) e^{-2y} dy$$

$$= 2 \int_{y=o}^{\infty} \left[-e^{-x} \right]_{x=0}^{y} e^{-2y} dy = 2 \int_{y=0}^{\infty} \left(1 - e^{-y} \right) e^{-2y} dy = 2 \int_{y=0}^{\infty} \left(e^{-2y} - e^{-3y} \right) dy$$

$$= 2 \left[-\frac{e^{-2y}}{2} + \frac{e^{-3y}}{3} \right]_{y=0}^{\infty} = 2 \left[0 - \left(-\frac{1}{2} + \frac{1}{3} \right) \right] = 2 \left(\frac{1}{2} - \frac{1}{3} \right) = 1 - \frac{2}{3}$$

$$= \mathbf{\frac{1}{3}}$$

3. **Compute $\Pr(X < a)$.**

$$\Pr(X < a) = \int_{x=0}^{a} \int_{y=0}^{\infty} f(x,y) dx dy$$

$$= \int_{x=0}^{a} \int_{y=0}^{\infty} 2e^{-x}e^{-2y} dx dy = 2 \left(\int_{x=0}^{a} e^{-x} dx \right) \left(\int_{y=0}^{\infty} e^{-2y} dy \right)$$

$$= 2 \left[-e^{-x} \right]_{x=0}^{a} \left[-\frac{e^{-2y}}{2} \right]_{y=0}^{\infty} = \left[e^{-x} \right]_{x=0}^{a} \left[e^{-2y} \right]_{y=0}^{\infty} = \left(e^{-a} - 1 \right) (0 - 1)$$

$$= \mathbf{1 - e^{-a}}$$

Question 5

Let X be a random variable and x be a real number. A linear function of the squared deviation from x is another random variable, $Y = a + b(X - x)^2$, where a and b are some positive constant.

1. **Find the value of x that minimizes $E(Y)$. Show that your result is really the minimum.**

The *Law of the unconscious statistician* states that:

$$E\left[g(X)\right] = \int g(x) f_X(x) dx$$

So, if we call $\mu = E(X)$ and $\sigma^2 = Var(X)$ (and knowing that $Var(X) = E\left(X^2\right) - \left(E(X)\right)^2 = E\left(X^2\right) - \mu^2$):

$$\begin{aligned}
E(Y) &= \int \left(a + b(X - x)^2\right) f(X) dX \\
&= \int \left(a + b(X^2 + x^2 - 2xX\right) f(X) dX \\
&= (a + bx^2) \int f(X) dX - 2bx \int X f(X) dX + b \int X^2 f(X) dX \\
&= (a + bx^2) 1 - 2bx\mu + b(\sigma^2 - \mu^2) \\
&= bx^2 - 2b\mu x + \left(a + b(\sigma^2 + \mu^2)\right)
\end{aligned}$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx - 2b\mu = 2b(x - \mu) = 2b\left(x - E(X)\right)$$

And consequently:

$$\frac{dE(Y)}{dx} = 0 \Rightarrow \mathbf{x = E(X)}$$

2. **Find the value of $E(Y)$ for the choice of x you found in (1)?**

We just have to substitute in the last expression of $E(Y)

$$\begin{aligned}
\mathbf{E(Y)} &= \int \left(a + b(X - \mu)^2\right) f(X) dX \\
&= b\mu^2 - 2b\mu^2 + \left(a + b(\sigma^2 + \mu^2)\right) = -b\mu^2 + a + b\sigma^2 + b\mu^2 \\
&= \mathbf{a + b\sigma^2}
\end{aligned}$$

3. **Suppose $Y = ax + b(X - x)^2$. Find the values of x that minimizes $E(Y)$. Show that your result is really the minimum.**

$$\begin{aligned}
E(Y) &= \int \left(ax + b(X - x)^2\right) f(X) dX \\
&= \int \left(ax + b(X^2 + x^2 - 2xX\right) f(X) dX \\
&= (ax + bx^2) \int f(X) dX - 2bx \int X f(X) dX + b \int X^2 f(X) dX \\
&= (ax + bx^2) 1 - 2bx\mu + b(\sigma^2 - \mu^2)
\end{aligned}$$

$$= bx^2 + (a - 2b\mu)x + b(\sigma^2 + \mu^2)$$

Hence:

$$\frac{dE(Y)}{dx} = 2bx + (a - 2b\mu) = 2b(x - \mu) + a = 2b\,(x - E(X)) + a$$

And consequently:

$$\frac{dE(Y)}{dx} = 0 \Rightarrow \mathbf{x = E(X) - \frac{a}{2b}}$$

Question 6

Suppose X and Y are independent continuous random variables, where both of which are uniformly distributed between 0 and 1. Let random variable $Z = X + Y$.

1. **Choose a value of z between 0 and 2, and draw a graph depicting the region of the $X - Y$ plane for which Z is less than z.**

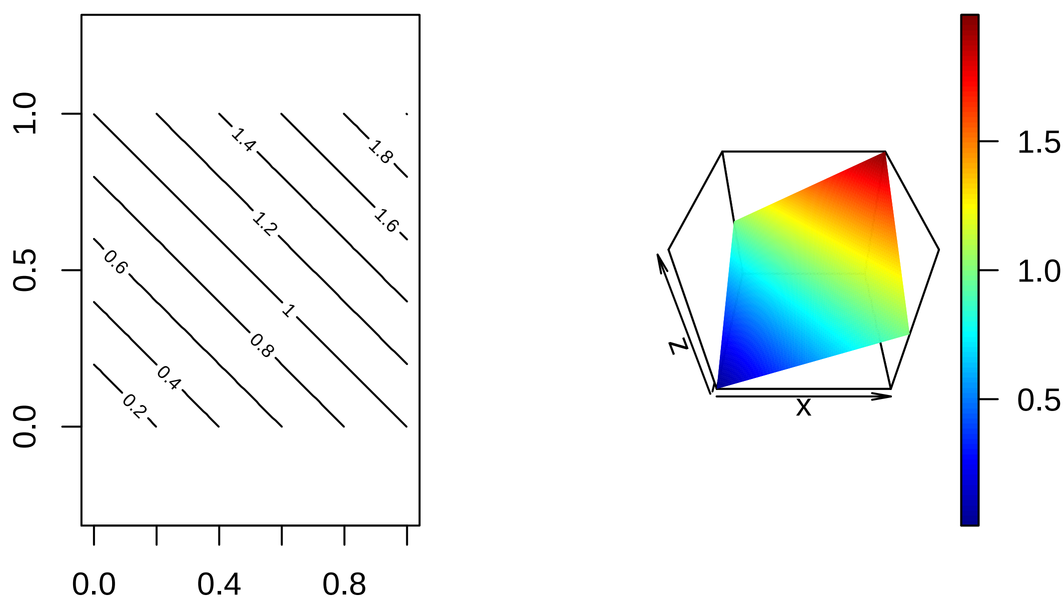

Figure 3: Some plot

2. **Derive the probability density function, $f(z)$.**

Question 7

In a casino, you pay the following game. A pair fair, ordinary 6-faced dices is rolled. If the sum of the dice is 2, 3, or 12, the house wins. If it is 7 or 11, you win. If it is any other number x, the house rolls the dice again until the sum is either 7 or x. If it is 7, the house wins. If it is x, you win. A game ends if one of the two players wins. Let Y be the number of rolls needed until the game ends.

1. **Is the expected number of rolls given that you win more than, equal to, or less than the expected number of rolls given that house wins (in a game)? The steps to arrive at your answer numerically need to be clearly shown.**

2. Suppose it takes \$20 to pay, and the payoff is \$100, \$80, \$60, \$40, \$0 if you win in the 1st, 2nd, 3rd, 4th, 5th round, respectively. That is, if you win in the 1st round, you are paid \$100 (so your net profit is \$80), if you win in the 2nd round, you are paid \$80, etc. Derive the expected payoff function of a game.

Part III: Statistical Estimation and Statistical Inference

In classical statistics, parameters are unknown constants whereas estimators are functions of samples and are random variables. The questions in this section are designed to clarify the relationship between parameters and estimators, and explore the properties that different estimators may have.

Question 8

Let $Y1, \ldots, Yn$ be n random variables, such that any two of them are uncorrelated, and all share the same mean μ and variance σ^2. Let Y be the average Y_i, which is also a random variable.

Define the class of linear estimators of μ by

$$W = \sum_{i=1}^{n} a_i Y_i$$

where the a_i are constants.

1. **What restriction on the a_i is needed for W to be an unbiased estimator of μ?**
2. **Find $Var(W)$.**
3. **Given a set of numbes $a_1, \ldots, a_n$, the following inequality holds:**

$$\frac{1}{n}\left(\sum_{i=1}^{n} a_i\right)^2 \leq \sum_{i=1}^{n} a_i^2$$

Use this inequality, along with the previous parts of this question, tho show that $Var(W) \geq Var(\bar{Y})$ whenever W is unbiased. We say that $\bar{Y}$ is the best linear unbiased estimator (BLUE).

Question 9

Let $\bar{Y}$ denote the average of n independent draws from a population distribution with mean μ and variance σ^2. Consider two alternative estimators of μ: $W_1 = \frac{n-1}{n}\bar{Y}$ and $W_2 = k\bar{Y}$, where $0 < k < 1$.

1. **Compute the biases of both W_1 and W_2. Which estimator is consistent?**
2. **Compute $Var(W_1)$ and $Var(W_2)$. Which estimator has lower variance?**

Question 10

Given a random sample $Y_1, Y_2, \ldots, Y_n$ from some distribution $F(\cdot)$ with mean μ and variance σ^2, where both μ and σ^2 are unknown parameters.

Let $\bar{Y}$ be the average of the sample. Consider the following estimator for σ^2:

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(Y - \bar{Y}\right)^2$$

1. **Show that $E(\bar{Y}) = E(Y_i)\ \forall i \in 1, 2, \ldots, n$**
2. **Show that $Var(\bar{Y}) = \frac{1}{n}Var(Y_i)\ \forall i \in 1, 2, \ldots, n$**
3. **Compute the expectation of $\widehat{\sigma^2}$ in terms of n and σ^2. In your derivation, make sure make use of the *i.i.d.* property and identify where you use it.**
4. **Is this an unbiased estimator for σ^2?**
5. **If not, what function of $\widehat{\sigma^2}$ produce an unbiased estimator?**

Question 11

Wooldridge's textbook: Appendix C, Question 4*i*, *ii*, *iii*.

Question 12

Wooldridge's textbook: Appendix C, Question 6.

Question 13

Wooldridge's textbook: Appendix C, Question 8. In addition, answer the following questions:

1. **Define Type I error.**
2. **What is the probability of Type I error of this test?**
3. **Define Type II error.**
4. **What is the probability of Type II error when using this decision rule, assuming the "true" population proportion is $\theta^* = 0.45$.**
5. **Define the power of the test (in general terms).**
6. **Calculate the power of this test, again assuming the "true" population proportion is $\theta^* = 0.45$.**