QBS 120 - Problem Set 2

Rob Frost

Grading: problems 2 (3.20) and 6 (4.49), 5 pts for each. See detailed grading criteria below.

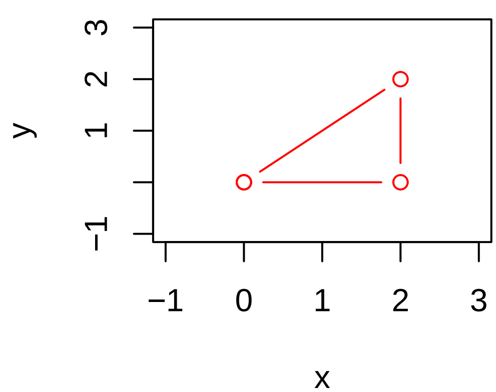
1. (Based on Rice, Chapter 3, Problem 18): Let X and Y have the joint density function $f(x,y) = kxy, 0 \leq y \leq x \leq 2$ **or 0 elsewhere.**

A **Describe the region over which the density if positive and use it in determining limits of integration to answer the following questions.**

In this case, the density is positive within a triangle defined by the following vertices:

```
>  (data = data.frame(x=c(0,2,2), y=c(0,0,2)))

  x y
1 0 0
2 2 0
3 2 2
```



B **Find k.**

We know that the joint density must integrate to 1 over the region defined in a) (note the limits of integration). Solving this multiple integration problem will give us k.

$$\int_0^2 \int_0^x kxy\,dy\,dx = 1$$
$$\int_0^2 \left[kxy^2/2\right]_0^x dx = 1$$
$$\int_0^2 kx^3/2\,dx = 1$$
$$\left[kx^4/8\right]_0^2 = 1$$
$$k = 0.5$$

C **Find the marginal densities of X and Y.**

To find the marginal density of X, we integrate the joint over Y:

$$\begin{aligned} f_X(x) &= \int_0^x \frac{xy}{2} dy \\ &= \left[xy^2/4\right]_0^x \\ &= x^3/4 \end{aligned}$$

Likewise, we can compute the marginal density of Y by integrating the joint density over X. The limits of integration are a bit tricky in this case: for a given y, x can range from y to 2 (not from 0 to 2).

$$\begin{aligned} f_Y(y) &= \int_y^2 \frac{xy}{2} dx \\ &= \left[x^2y/4\right]_y^2 \\ &= y - \frac{y^3}{4} \end{aligned}$$

D **Find the conditional densities of Y given X and X given Y.**

By definition, the conditional density of Y|X, defined for $0 \leq y \leq x$, is:

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} \\ &= \frac{xy/2}{x^3/4} \\ &= \frac{2y}{x^2} \end{aligned}$$

Similarly, the conditional density of X|Y, defined for $y \leq x \leq 2$, is:

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{xy/2}{y - y^3/4} \\ &= \frac{2x}{4 - y^2} \end{aligned}$$

2. (Based on Rice, Chapter 3, Problem 20) If X_1 is uniform on [0,1], and, conditional on X_1, X_2 is uniform on $[X_1, 2]$, find the joint and marginal distributions of X_1 and X_2.

As a standard uniform RV, we know that the density of X_1 is:

$$f_{X_1}(x_1) = 1, 0 \leq x_1 \leq 1$$

Grading: 1 pt to get $f_{X_1}(x_1)$

We are also told that the conditional density of X_2 is $U(X_1, 2)$:

$$f_{X_2|X_1}(x_2|x_1) = 1/(2 - x_1), 0 \leq x_1 \leq x_2 \leq 2$$

The joint density can be found directly from these by definition:

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2|X_1}(x_2|x_1) \\ &= 1 * 1/(2 - x_1) \\ &= 1/(2 - x_1), 0 \leq x_1 \leq 1, x_1 \leq x_2 \leq 2 \end{aligned}$$

Grading: 2 pts to get $f_{X_1,X_2}(x_1, x_2)$. Partial credit if the setup looks valid but they don't get the correct answer.

We were given the marginal density of X_1, $f_{X_1}(x_1) = 1$). For the marginal density of X_2, can find by integrating the joint density over X_1. For the limits of integration, note that for a given x_2, x_1, which must be smaller than x_2, is bound between 0 and x_2.

$$\begin{aligned} f_{X_2}(x_2) &= \int_0^{x_2} f_{X_1,X_2}(x_1, x_2) dx_1 \\ &= \int_0^{x_2} 1/(2 - x_1) dx_1 \\ &= \big[- ln(2 - x_1) \big]_0^{x_2} \\ &= -ln(2 - x_2) + ln(2) \end{aligned}$$

Grading: 2 pts to get $f_{X_2}(x_2)$. Partial credit if the setup looks valid but they don't get the correct answer.

3. (Optional, Rice, Chapter 3, Problem 53) Consider forming a random rectangle in two ways. Let U_1, U_2, U_3 be independent standard normal variables. One rectangle has sides U_1 and U_2 and the other is square with side U_3. Find the probability that the area of the square is greater than the area of the other rectangle.

This problem involves both functions of multiple RVs as well as joint distributions. The first step is to define RVs that represent the areas of the rectangle and square. Let $A_R = U_1 * U_2$ be the area of the rectangle and let $A_S = U_3^2$ be the area of the square. We need to find the densities of A_R and A_S and, using those marginal densities, the joint density. The desired probability will be found by integrating the joint density over the region where $A_S > A_R$.

To find the density of A_R, we'll adapt the approach from the book for the distribution of a quotient. From page 98, the density of $Z = Y/X$ is given by:

$$f_Z(z) = \int_{-\infty}^{\infty} |x| f_{X,Y}(x, xz) dx$$

The trick here is to realize that $Y = XZ$ and make this substitution to get an integral in just one variable. For the product $Z = XY$, we have Y = Z/X and the density is given by:

$$f_Z(z) = \int_{-\infty}^{\infty} 1/|x| f_{X,Y}(x, z/x) dx$$

So, for A_R this becomes ((to get the limits of integration, we know that the max is 1 (U_1 is standard normal) and the min, achieved when $U_2 = 1$, will be A_R):

$$\begin{aligned}
f_{A_R}(a_r) &= \int_{a_r}^{1} 1/|u_1| f_{U_1,U_2}(u_1, a_r/u_1) du_1 \\
&= \int_{a_r}^{1} 1/|u_1| f_{U_1}(u_1) f_{U_2}(a_r/u_1) du_1 && \text{joint density factors since independent} \\
&= \int_{a_r}^{1} 1/u_1 * 1 * 1 du_1 && \text{sub in U(0,1) density, integral over pos. so remove abs} \\
&= \left[log(u_1)\right]_{a_r}^{1} \\
&= -log(a_r)
\end{aligned}$$

To find the density of A_S, we'll start by defining the CDF:

$$\begin{aligned}
F_{A_S}(a_s) &= P(A_S < a_s) && \text{defn of CDF} \\
&= P(U_3^2 < a_s) && \text{sub in value of } A_S \\
&= P(U_3 < \sqrt{a_s}) && \text{manipulate} \\
&= F_{U_3}(\sqrt{a_s}) && \text{this equals the CDF of } U_3 \\
&= \sqrt{a_s}
\end{aligned}$$

Get the density via differentiation:

$$\begin{aligned}
f_{A_S}(a_s) &= \frac{\delta}{\delta a_s} F_{A_S}(a_s) \\
&= \frac{\delta}{\delta a_s} \sqrt{a_s} \\
&= \frac{1}{2\sqrt{a_s}}
\end{aligned}$$

Since A_R and A_S are independent, their joint density factors into the marginal densities. To find the desired probability, we want to integrate this joint density over the region where $A_S > A_R$. In (A_R, A_S) coordinates, this is the triangle with vertices (0,0), (0,1) and (1,1). We will therefore use the limits of (0,1) for A_S and the limits of $(0, a_s)$ (i.e., for a given A_S, A_R must be smaller). The probability is therefore given by:

$$
\begin{aligned}
P(A_S > A_R) &= \int_0^1 \int_0^{a_s} f_{A_R}(a_r) f_{A_S}(a_s) da_r da_s \\
&= \int_0^1 \int_0^{a_s} \frac{-log(a_r)}{2\sqrt{a_s}} da_r da_s \\
&= \int_0^1 \frac{1}{2\sqrt{a_s}} \big[- a_r log(a_r) + a_r \big]_0^{a_s} da_s \\
&= \int_0^1 \frac{a_s(-log(a_s) + 1)}{2\sqrt{a_s}} da_s \\
&= \int_0^1 (\frac{\sqrt{a_s}}{2} - \frac{\sqrt{a_s} log(a_s)}{2}) da_s \\
&= \lim_{x \to 0} \int_x^1 (\frac{\sqrt{a_s}}{2} - \frac{\sqrt{a_s} log(a_s)}{2}) da_s \\
&= \lim_{x \to 0} \big[a_s^{3/2}/3 + -(a_s^{3/2} log(a_s))/3 + 2a_s^{3/2}/9 \big]_x^1 \\
&= 1/3 - 0 + 2/9 - 0 \\
&= 5/9
\end{aligned}
$$

Integration by parts fun!

4. (Based on Rice, Chapter 3, Problem 71) Let $X_1, ..., X_n$ be independent RVs all with the same density f. Find an expression for the probability that the interval $[X_{(1)}, \infty)$ encompasses at least 100v% of the probability mass of density f.

For the interval $[X_{(1)}, \infty)$ to encompass at least 100v% of the probability mass of density f, we need $X_{(1)}$ to be less than some value a such that $P(X > a) = v$ or $1 - P(X < a) = v$ or $1 - F_X(a) = v$. The value of a itself is given the inverse CDF: $a = F_X^{-1}(1 - v)$.

The probability that $X_{(1)} \leq a$ can be found using the CDF of $X_{(x)} : F_{X_{(1)}}(x_{(1)})$. Since $X_{(1)}$ is the smallest order statistic the CDF is:

$$
F_{X_{(1)}}(x_{(1)}) = 1 - (1 - F_X(x_{(1)}))^n
$$

The desired probability is therefore given by:

$$
\begin{aligned}
P(X_{(1)} \leq a) &= F_{X_{(1)}}(a) \\
&= 1 - (1 - F_X(a))^n \\
&= 1 - (1 - F_X(F_X^{-1}(1 - v)))^n \\
&= 1 - v^n
\end{aligned}
$$

5. (Based on Rice, Chapter 4, Problem 31) Let X be uniformly distributed on the interval [1,4]. Find E[1/X]. Is E[1/X] = 1/E[X]?

This problem involves the expectation of a function of a continuous RV. We know that for the general function g(X) of RV X, E[g(X)] takes the form:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

In this case, $g(X) = 1/X, 1 \leq X \leq 4$ so the expectation becomes:

$$\begin{aligned} E[1/X] &= \int_1^4 1/X * f(x)dx \\ &= \int_1^4 1/(3X)dx && \text{f(x) for U(1,4) is just 1/(4-1) or 1/3} \\ &= \left[ln(x)/3\right]_1^4 \\ &= ln(4)/3 \end{aligned}$$

For 1/E[X], we'll first compute the expectation of X from first principles (as the center of mass of the probability density, we know it is in the middle of the uniform range: 2.5):

$$\begin{aligned} E[X] &= \int_1^4 xf(x)dx \\ &= \int_1^4 x/3dx \\ &= \left[x^2/6\right]_1^4 \\ &= 16/6 - 1/6 \\ &= 15/6 = 2.5 \\ 1/E[X] &= 0.4 \end{aligned}$$

So, does $E[1/X] = 1/E[X]$? No! In general, $E[g(X)] \neq g(E[X])$.

6. (based on Rice, Chapter 4, Problem 49) Two independent measurements, X and Y, are taken of a quantity μ. E[X] = E[Y] = μ, but σ_x and σ_y are unequal. The two measurements are combined by means of a weighted average to give:

$$Z = \alpha X + (1 - \alpha)Y$$

where α is a scalar and $0 \leq \alpha \leq 1$.

A **Show that** $E[Z] = \mu$

We want to find the expectation of a linear function of RVs. We can use the results in Theorem A from section 4.1.2. Specifically for jointly distributed RVs $X_1, ..., X_n$ and $Y = a\sum_{i=1}^{n} b_i X_i$, E[Y] is a linear function of the $E[X_i]$ (expectation of sum is sum of expectations):

$$E[Y] = a\sum_{i=1}^{n} b_i E[X_i]$$

Plugging in $Z = \alpha X + (1 - \alpha)Y$, we get:

$$
\begin{aligned}
E[Z] &= \alpha E[X] + (1-\alpha)E[Y] \\
&= \alpha\mu + (1-\alpha)\mu \\
&= \mu
\end{aligned}
$$

Grading: 1 pt to compute the correct expectation using Theorem A. Half credit if they use the correct theorem but make a mistake somewhere.

B **If X and Y are not independent, what is $E[Z]$?**

The result is the same ($E[Z] = \mu$) since Theorem A holds regardless of whether the RVs are independent or dependent.

Grading: 1/2 pt to note that the answer is the same regardless of independence.

C **What is Var(Z)? Does this result hold if X and Y are not independent?**

To find Var(Z), we'll make use of two results regarding the variance of linear functions of RVs:

- Theorem A in section 4.2 that $Var(a + bX) = b^2 Var(X)$
- Corollary B in section 4.3 that the variance of a sum of independent RVs is equal to the sum of the variances.

$$
\begin{aligned}
Var(Z) &= Var(\alpha X + (1-\alpha)Y) && \text{given} \\
&= Var(\alpha X) + Var((1-\alpha)Y) && \text{per Corollary B} \\
&= \alpha^2 Var(X) + (1-\alpha)^2 Var(Y) && \text{per Theorem A} \\
&= \alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 && \text{plug in vars of X and Y} \\
&= \alpha^2\sigma_X^2 + (1-2\alpha+\alpha^2)\sigma_Y^2
\end{aligned}
$$

Unlike expectation, this result for Var(Z) does not hold if X and Y are dependent.

Grading: 1.5 points total. 1 pt to compute the correct variance (half credit if approach looks valid but answer is wrong). 1/2 point to correctly note that the result does not hold if X and Y are dependent.

D **Find α in terms of σ_X and σ_Y to minimize Var(Z).**

To minimize Var(Z) we take the derivative of Var(Z) wrt to α and find where it is 0 (extrema) and a minimum. Now compute the derivative:

$$
\begin{aligned}
\frac{d}{d\alpha} Var(Z) &= \frac{d}{d\alpha}\left[\alpha^2\sigma_X^2 + (1-2\alpha+\alpha^2)\sigma_Y^2\right] \\
&= \frac{d}{d\alpha}\left[\alpha^2(\sigma_X^2+\sigma_Y^2) - 2\alpha\sigma_Y^2 + \sigma_Y^2\right] \\
&= 2\alpha(\sigma_X^2+\sigma_Y^2) - 2\sigma_Y^2
\end{aligned}
$$

To find minimum, set equal to 0 and solve for α:

$$2\alpha(\sigma_X^2 + \sigma_Y^2) - 2\sigma_Y^2 = 0$$
$$\alpha = \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$$

Since the second derivative of Var(Z) is positive, we know this is a minimum.

Grading: 1 pt to compute the correct value. Half credit if the approach is generally correct but answer is wrong.

E **Under what circumstances is it better to use the average $(X+Y)/2$ than either X or Y alone?**

Note that $(X+Y)/2$ is Z with $\alpha = 1/2$. Note also that the expected value is the same in all cases, per the result from part a). So, the only difference will be the variance and a smaller variance is typically desirable (less uncertainty in the outcome). Using the result for Var(Z) from part b) we know the variance when $\alpha = 1/2$ is:

$$\begin{aligned} Var(Z) &= \alpha^2\sigma_X^2 + (1 - 2\alpha + \alpha^2)\sigma_Y^2 \\ &= \sigma_X^2/4 + (1 - 1 + 1/4)\sigma_Y^2 \\ &= \sigma_X^2/4 + \sigma_Y^2/4 \end{aligned}$$

When is this less than $Var(X) = \sigma_X^2$? In that case, we'd rather use Z than X alone to minimize the variance:

$$\begin{aligned} \sigma_X^2/4 + \sigma_Y^2/4 &< \sigma_X^2 \\ \frac{3\sigma_X^2}{4} &> \sigma_Y^2/4 \\ \sigma_X^2 &> 3\sigma_Y^2 \\ \sigma_X^2/\sigma_y^2 &> 1/3 \end{aligned}$$

When would we prefer to use Z vs Y alone? That will be when $Var(Z) < Var(Y)$. By symmetry, we know that will be when $\sigma_Y^2/\sigma_X^2 > 1/3$ or equivalently $\sigma_X^2/\sigma_Y^2 < 3$.

Combining these inequalities gives us the range of σ_X^2/σ_Y^2 values when we would prefer to use the simple average

$$1/3 < \sigma_X^2/\sigma_y^2 < 3$$

Grading: 1 pt for the correct answer. Half credit if the approach is generally correct but answer is wrong.

7. (Based on Rice, Chapter 4, Problem 57) If X and Y are independent random variables, find E[XY] and Var(XY) in terms of the means and variances of X and Y.

Find E[XY]: Per Corollary A from Section 4.1.1: if X and Y are independent, then E[XY] = E[X]E[Y]:

$$E[XY] = E[X]E[Y] = \mu_X \mu_Y$$

Find Var(XY): By definition, Var(XY) can be expressed in terms of the expectations as:

$$Var(XY) = E[(XY)^2] - E[XY]^2$$

We already found that $E[XY] = E[X]E[Y]$. To find $E[(XY)^2] = E[X^2Y^2]$, remember that functions of independent RV are also independent. So, $E[X^2Y^2] = E[X^2]E[Y^2]$ per Corollary A. Plugging these in we get:

$$\begin{aligned} Var(XY) &= E[X^2]E[Y^2] - (E[X]E[Y])^2 && \text{by Corollary A} \\ &= E[X^2]E[Y^2] - E[X]^2E[Y]^2 \end{aligned}$$

$E[X]^2$ and $E[Y]^2$ are just the squares of the marginal expectations, so these are OK to include. However, what to with $E[X^2]E[Y^2]$? The trick here is to recognize that we can re-express $E[X^2]$ as $Var(X) + E[X]^2$, which includes just marginal variance and expectation terms. Using the μ and σ^2 notation, this therefore becomes:

$$\begin{aligned} Var(XY) &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 \\ Var(XY) &= \sigma_X^2\sigma_Y^2 + \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 \end{aligned}$$