Problem Set #3

Danny Edgel
Econ 709: Economic Statistics and Econometrics I
Fall 2020

September 27, 2020

Collaborated with Sarah Bass, Emily Case, Michael Nattinger, and Alex Von Hafften

Question 1

A random point (X, Y) is distributed uniformly on the square with vertices $(1, 1)$, $(1, -1)$, (-1.1), and $(-1, -1)$. That is, the joint PDF is $f(x, y) = 1/4$ on the square and $f(x, y) = 0$ outside the square. Determine the probability of:

(a) $\mathbf{X^2 + Y^2 < 1}$

$P(X^2 + Y^2 < 1)$ is the area of the circle inscribed within the square. Therefore, $P(X^2 + Y^2 < 1) = \frac{\pi}{4}$.

(b) $|\mathbf{X} + \mathbf{Y}| < \mathbf{2}$

$P(|X + Y| < 2) = P(-2 < X + Y < 2)$. Note that $|X + Y| = 2$ only if $X = Y = -1$ or $X = Y = 1$. Since X and Y are continuous, $P(X = 1) = P(Y = 1) = P(X = -1) = P(Y = -1) = 0$. Therefore, $P(|X + Y| < 2) = 0$.

Question 2

Let the joint PDF of X and Y be given by $f(x, y) = g(x)h(y)$ $\forall x, y \in \mathbb{R}$. Let a denote $\int_{-\infty}^{\infty} g(x)dx$ and b denote $\int_{-\infty}^{\infty} h(x)dx$.

(a) **What conditions should a and b satisfy in order for $f(x, y)$ to be a bivariate PDF?**

If f is a bivariate PDF, then $\int_{-\infty}^{\infty} f(x, y)dxdy = 1$. Then,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)dxdy = \left(\int_{-\infty}^{\infty} g(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)dy\right) = ab$$

Thus, $ab = 1$ if f is a bivariable PDF.

(b) **Find the marginal PDF of X and Y.**

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\infty}^{\infty} g(x)h(y)dy = bg(x)$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} g(x)h(y)dx = ah(y)$$

(c) **Show that X and Y are independent.**

X and Y are independent if and only if $f(x,y) = f_X(x)f_Y(y)$. From (a) and (b), we can derive:

$$f_X(x)f_Y(y) = ag(x)bh(y) = g(x)h(y) = f(x,y)$$

Thus, X and Y are independent.

Question 3

Let the joint PDF of X and Y be given by

$$\mathbf{f(x,y)} = \begin{cases} cxy & \text{if } x, y \in [0,1],\ x+y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) **Find the value of c such that $f(x,y)$ is a joint PDF.**

If f is a PDF, then $\int_{-\infty}^{\infty} f(x,y)dxdy = 1$. Thus,

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy &= 1 \\
\int_0^1 \int_0^{1-x} cxydydx &= 1 \\
\int_0^1 cx\frac{1}{2}[y^2]_0^{1-x}dx &= 1 \\
\int_0^1 cx\frac{1}{2}(1-x)^2dx &= 1 \\
c\left[\frac{1}{2}x^2 - \frac{2}{3}x^3 + \frac{1}{4}x^4\right]_0^1 &= 2 \\
\frac{1}{12}c &= 2 \\
c &= 24
\end{aligned}$$

(b) **Find the marginal distributions of X and Y.**

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 \int_0^{1-x} cxydydx = \frac{1}{2}cx(1-x)^2$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^1 \int_0^{1-y} cxydxdy = \frac{1}{2}cy(1-y)^2$$

(c) **Are X and Y independent? Compare your answer to Problem 2 and discuss.**

X and Y are independent if and only if $f(x,y) = f_X(x)f_Y(y)$. From (a) and (b), we can derive:

$$f_X(x)f_Y(y) = \frac{1}{4}c^2x(1-x)^2y(1-y)^2 \neq cxy$$

Thus, X and Y are **not** independent.

Question 4

Show that any random variable is uncorrelated with a constant.

Let X be a random variable and a be a constant. Then,

$$Cov(X,a) = E[Xa] - E[X]E[a] = aE[X] - aE[X] = 0$$

Thus, $Corr(X,a) = Cov(X,a)/\sqrt{Var(X)Var(a)} = 0$, so X and a are uncorrelated.

Question 5

Let X and Y be independent random variables with means μ_X, μ_Y, and variances σ_X^2, σ_Y^2. Find an expression for the correlation of XY and Y in terms of these means and variances.

Given the definition of correlation and covariance, we have:

$$Corr(XY, Y) = \frac{E(XY^2) - E(XY)E(Y)}{\sqrt{Var(XY)Var(Y)}}$$

Separately, since X and Y are independent, we can solve:

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

Thus, we can write the correlation as:

$$Corr(XY, Y) = \frac{\mu_X\sigma_Y^2}{\sqrt{(\sigma_X^2\sigma_Y^2 - \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2)\sigma_Y^2}} = \frac{\mu_X\sigma_Y}{\sqrt{\sigma_X^2\sigma_Y^2 - \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2}}$$

Question 6

Prove the following: For any random vector $(X_1, X_2, ..., X_n)$,

$$\mathbf{Var}\left(\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{X_i}\right) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Var(X_i)} + \mathbf{2}\sum_{\mathbf{i \le i < j \le n}} \mathbf{Cov(X_i, X_j)}$$

We can prove this by induction:

Proof.

1. *Base step:* Let $n = 2$. Then,

$$Var(X_1+X_2) = Var(X_1)+Var(X_2)+2Cov(X_1, X_2) = \sum_{i=1}^{2} Var(X_i)+\sum_{1\le i\le j\le 2} Cov(X_i, X_j)$$

2. *Induction step:* Let $n = n$ and assume $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i \leq i < j \leq n} Cov(X_i, X_j)$. Then,

$$Var\left(\sum_{i=1}^{n+1} X_i\right) = Var\left(\sum_{i=1}^{n} X_i\right) + Var(X_{n+1}) + 2Cov\left(\sum_{i=1}^{n} X_i, X_{n+1}\right)$$
$$= \sum_{i=1}^{n} Var(X_i) + Var(X_{n+1}) + 2\sum_{i \leq i < j \leq n} Cov(X_i, X_j) + 2Cov\left(\sum_{i=1}^{n} X_i, X_{n+1}\right)$$
$$Cov\left(\sum_{i=1}^{n} X_i, X_{n+1}\right) = Cov(X_1, X_{n+1}) + ... + Cov(X_n, X_{n+1}) = \sum_{i=1}^{n} Cov(X_i, X_{n+1})$$
$$\therefore Var\left(\sum_{i=1}^{n+1} X_i\right) = \sum_{i=1}^{n+1} Var(X_i) + 2\sum_{i \leq i < j \leq n+1} Cov(X_i, X_j) \blacksquare$$

Question 7

Suppose that X and Y are joint normal, i.e., they have the joint PDF:

$$\mathbf{f(x,y)} = \frac{\mathbf{1}}{\mathbf{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}}\exp(-(\mathbf{2(1-\rho^2)})^{-\mathbf{1}}(\mathbf{x^2/\sigma_X^2 - 2\rho xy/\sigma_X\sigma_Y + y^2/\sigma_Y^2}))$$

(a) **Derive the marginal distribution of X and Y, and observe that both are normal distributions.**

The marginal distribution of X is defined as $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ and the marginal distribution of y is defined as $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$. We can solve:

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp(-(2(1-\rho^2))^{-1}(x^2/\sigma_X^2 - 2\rho xy/\sigma_X\sigma_Y + y^2/\sigma_Y^2))dy$$
$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{\exp(-(2(1-\rho^2))^{-1}(x^2/\sigma_X^2 + (y/\sigma_Y - \rho x/\sigma_X)^2 - \rho^2x^2/\sigma_X^2))}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}}dy$$
$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\exp\left(\frac{x^2/\sigma_X^2 - \rho^2x^2/\sigma_X^2}{-(2(1-\rho^2))}\right)\int_{-\infty}^{\infty} \frac{1}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}}dy$$
$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_X^2}}\int_{-\infty}^{\infty} \frac{1}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}}e^{\frac{(y-\sigma_Y\rho x/\sigma_X)^2}{-(2\sigma_Y^2(1-\rho^2))}}dy$$

Where the definite integral is a normal distribution with mean $\frac{\sigma_Y\rho x}{\sigma_X}$ and standard deviation $\sigma_Y\sqrt{1-\rho^2}$. Thus, the integral is equal to one and:

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_X^2}}$$

Using the same simplifying process, we can solve:

$$f_Y(y) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp(-(2(1-\rho^2))^{-1}(x^2/\sigma_X^2 - 2\rho xy/\sigma_X\sigma_Y + y^2/\sigma_Y^2))dx$$

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_Y^2}} \int_{-\infty}^{\infty} \frac{1}{\sigma_X\sqrt{2\pi}\sqrt{1-\rho^2}} e^{\frac{(x-\sigma_X\rho y/\sigma_Y)^2}{-(2\sigma_X^2(1-\rho^2))}} dx$$

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_Y^2}}$$

Clearly, each of these marginal distributions are normal distributions with a zero mean.

(b) **Derive the conditional distribution of Y given $X = x$, Observe that it is also a normal distribution.**

Since both X and Y are continuous random variables, $f_{Y|X}(y|x) = \frac{f_{Y,X}(y,x)}{f_X(x)}$. Then, we can derive:

$$f_{Y|X}(y|x) = \frac{\sigma_X\sqrt{2\pi}\exp\left(-\frac{x^2/\sigma_X^2 - 2\rho xy/\sigma_X\sigma_y + y^2/\sigma_Y^2}{2(1-\rho^2)}\right)}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}\exp\left(-\frac{x^2}{2\sigma_X^2}\right)}$$

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{\sigma_X^2(x^2/\sigma_X^2 - 2\rho xy/\sigma_X\sigma_Y + y^2/\sigma_Y^2) + (1-\rho^2)x^2}{2\sigma_X^2(1-\rho^2)}\right)$$

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{\left(\frac{\sigma_X}{\sigma_Y}y - \rho x\right)^2}{2\sigma_X(1-\rho^2)}\right)$$

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{\frac{\left(y - \frac{\sigma_Y}{\sigma_X}\rho x\right)^2}{2\sigma_Y^2(1-\rho^2)}}$$

Thus, the conditional distribution of Y on $X = x$ is normal with mean $\frac{\sigma_Y}{\sigma_X}\rho x$ and standard deviation $\sigma_Y\sqrt{1-\rho^2}$.

(c) **Derive the joint distribution of (X, Z) where $Z = (Y/\sigma_Y) - (\rho X/\sigma_))$, and then show that X and Z are independent.**

Solving $Z\ Y/\sigma_Y - \rho X/\sigma_X$ for Y yields $Y = \sigma_Y Z + \sigma_Y\rho X/\sigma_X$. Now, let $g : \begin{pmatrix} X \\ Y \end{pmatrix} \to \begin{pmatrix} X \\ Y/\sigma_Y - \rho X/\sigma_X \end{pmatrix}$ represent the mapping from $\begin{pmatrix} X \\ Y \end{pmatrix}$ to $\begin{pmatrix} X \\ Z \end{pmatrix}$. g has the inverse mapping $g^{-1} : \begin{pmatrix} X \\ Z \end{pmatrix} \to \begin{pmatrix} X \\ \sigma_Y Z + \sigma_Y\rho X/\sigma_X \end{pmatrix}$, which has Jacobian matrix J:

$$J = \begin{pmatrix} 1 & 0 \\ \frac{\sigma_Y}{\sigma_X}\rho & \sigma_Y \end{pmatrix}$$

Then, we can solve for $f(x, z)$:

$$
\begin{aligned}
f(z, x) &= f_{X,Y}(g^{-1}(x, z))|J| = f_{X,Y}(x, \sigma_Y Z + \sigma_Y \rho X/\sigma_X)\sigma_Y \\
&= \frac{\sigma_Y}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{\frac{x^2}{\sigma_X^2} - \frac{2\rho x(\sigma_Y z + \sigma_Y \rho x/\sigma_X}{\sigma_X\sigma_Y} + \frac{(\sigma_Y z + \sigma_Y \rho x/\sigma_X)^2}{\sigma_Y^2}}{2(1-\rho^2)}\right) \\
&= \left(\frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}}\right) \exp\left(\frac{\frac{x^2}{\sigma_X^2} - \frac{\sigma_Y}{\sigma_Y}\left(\frac{2\rho xz + 2\rho^2 x^2/\sigma_X}{\sigma_X}\right) + \frac{\sigma_Y^2}{\sigma_Y^2}\left(z + \frac{\rho x}{\sigma_X}\right)^2}{2(1-\rho^2)}\right) \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma_X}\right)\left(\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\right) \exp\left(\frac{\frac{x^2}{\sigma_X^2} - \frac{2\rho xz}{\sigma_X} - \frac{2\rho^2 x^2}{\sigma_X^2} + z^2 + \frac{2\rho xz}{\sigma_X} + \frac{\rho^2 x^2}{\sigma_X^2}}{2(1-\rho^2)}\right) \\
f_{Z,X}(z, x) &= \left(\frac{1}{\sqrt{2\pi}\sigma_X}\right) e^{-\frac{x^2}{2\sigma_X^2}} \left(\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\right) e^{-\frac{z^2}{2(1-\rho^2)}}
\end{aligned}
$$

We can see that the joint distribution of X and Z is the marginal distribution of X, multiplied by a normal distribution of Z with mean zero and standard deviation $\sqrt{1-\rho^2}$. Since the joint distribution is the product of each variable's marginal distribution, X and Y are independent.

Question 8

Consider a function $g : \mathbb{R} \to \mathbb{R}$. Recall that the inverse image of a set A, denoted $g^{-1}(A)$, is $g^{-1}(A) = \{x \in \mathbb{R} : g(x) \in A\}$. Let there be two functions, $g_1 : \mathbb{R} \to \mathbb{R}$ and $g_2 : \mathbb{R} \to \mathbb{R}$. Let X and Y be two random variables that are independent. Suppose that g_1 and g_2 are both Borel-measurable, which means that $g_1^{-1}(A)$ and $g_2^{-1}(A)$ are both in the Borel σ-field whenever A is in the Borel σ-field. Show that the two random variables $Z := g_1(X)$ and $W := g_2(Y)$ are independent. (Hint: use the 1st or the 2nd definition of independence.)

To show that Z and W are independent, I will show that the joint CDF of Z and W is equal to the product of their respective CDFs. Remebering that X and Y are independent and that any point in the Borel σ-field of Z or W must also be in the Borel σ-field of X or Y:

$$
\begin{aligned}
P(Z \leq z, W \leq w) &= P(g_1(X) \leq z, g_2(Y) \leq w) \\
&= P(X \leq g_1^{-1}(z), Y \leq g_2^{-1}(w)) \\
&= P(X \leq g_1^{-1}(z))P(Y \leq g_2^{-1}(w)) \\
&= P(g_1(X) \leq z)P(g_2(Y) \leq w) \\
P(Z \leq z, W \leq w) &= P(Z \leq z)P(W \leq w)
\end{aligned}
$$

Thus, Z and W are independent.