HW7

2024-11-03

Outline Q1: DONE - ish Q2: DONE Q3: DONE Q4: DONE Q5: DONE - ish Q6: DONE Q7: DONE

Q1

Consider continuous random vector (X, Y) where $X \sim N(0,1)$ (standard normal) and $Y|X = x \sim N(x,1)$ (the conditional distribution of Y given X= x is normal with mean x and variance 1).

(a)

Find the joint pdf of (X, Y).

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

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

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

(b)

Find the marginal distribution of Y.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{\left(-\frac{x^2}{2} - \frac{(y-x)^2}{2}\right)} dx$$

Simplifying the expression in the exponent, we have:

$$-\frac{x^2}{2} - \frac{(y-x)^2}{2} = -\frac{x^2}{2} - \frac{y^2 - 2xy + x^2}{2} = -\frac{y^2}{2} - \frac{1}{2}\left[x^2 - 2xy\right] = -\frac{y^2}{2} - \frac{1}{2}(x-y)^2$$

$$f_Y(y) = \frac{1}{2\pi} e^{\left(-\frac{y^2}{2}\right)} \int_{-\infty}^{\infty} e^{\left(-\frac{(x-y)^2}{2}\right)} dx$$

Looking at just the term(s) in the inside of the integral, we have:

$$\int_{-\infty}^{\infty} e^{\left(-\frac{(x-y)^2}{2}\right)} dx = \sqrt{2\pi}$$

Such that we can evaluate the full expression now as:

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

This pdf is the standard normal! Such that we say the marginal of Y given above describes the random variable Y as:

$$Y \sim N(0,1)$$

(c)

Consider random vector (V, W) where $W \sim N(0,2)$ (normal with mean 0 and variance 2). How should the conditional distribution $V \mid W = w$ be chosen so that (V, W) has the same distribution as (X, Y)?

First, consider the joint distribution of (X, Y); we know this is a multivariate normal distribution. As such, given $W \sim N(0,2)$, $V \sim N(0,1)$, we know that the random variable $V \mid W = w$ must also be normally distributed. Our goal then is to obtain the parameters of this normal distribution to uniquely characterize its distribution. To that end:

Consider the random vector (X, Y), made up of the Mean vector and Covariance Matrix:

Recall also that we know the Covariance Matrix as $Y|X = x \sim N(x,1)$, meaning Cov(X, Y) = 1 = Cov(Y, X). Taking this into account then:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\right)$$

As $\mu_V = 0$:

$$E(V \mid W = w) = \mu_V + \frac{Cov(V,W)}{Var(W)}(w - \mu_W) = \mu_V + \frac{Cov(V,W)}{Var(W)}(w - 0) = \frac{1}{2}w$$

$$E(V \mid W = w) = \mu_V + \frac{Cov(V,W)}{Var(W)}w = 0 + \frac{Cov(V,W)}{Var(W)}w = \frac{1}{2}w$$

As Var(V) = 1

$$Var(V|W = w) = Var(V) - \frac{Cov(V,W)^2}{Var(W)} = 1 - \frac{1^2}{2} = \frac{1}{2}$$

Furthermore, as we have found the mean and variance of V given W = w, we may describe its distribution as:

$$V \mid W = w \sim N\left(\frac{1}{2}w, \frac{1}{2}\right)$$

Q2: 4.7 Casella & Berger

A woman leaves for work between 8AM and 8:30AM and takes between 40 and 50 minutes to get there. Let the random variable X denote her time of departure, and the random variable Y the travel time. Assuming that these variables are independent and uniformly distributed, find the probability the the woman arrives at work before 9AM.

$$X \sim U(0, 30), Y \sim U(40, 50)$$

Given X and Y are both Uniformly distributed and independent, then:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \frac{1}{30}(\frac{1}{50-40}) = \frac{1}{300}$$

Define the random variable Z = X + Y, then we want to know the probability Z < 60 (1 hour, 60 minutes starting at the earliest time 8:00AM):

$$P(Z < 60) = P(X+Y < 60) = \int_{40}^{50}\int_{0}^{60-y} f_{X,Y}(x, y)dxdy = \int_{40}^{50}\int_{0}^{60-y} \frac{1}{300}dxdy = \int_{40}^{50} \frac{60-y}{300}dy = -\frac{(y-120(y))}{600}\Big|_{y=40}^{50} = \frac{1}{2}$$

Q3: 4.10 Casella & Berger

The random pair (X, Y) has the distribution:

	X		
	1	2	3
Y 2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
3	$\frac{1}{6}$	0	$\frac{1}{6}$
4	0	$\frac{1}{3}$	0

Figure 1: Img

(a)

Show that X and Y are dependent.

To show that two variables are not independent (are dependent), we need to show that the product of the marginal probabilities is not equal to joint probability. To that end:

$$P(X = 1) = \frac{1}{12} + \frac{1}{6} = \frac{3}{12} = \frac{1}{4}$$

$$P(Y = 4) = \frac{1}{3}$$

$$P(X = 1, Y = 4) = 0$$

$$P(X = 1, Y = 4) = 0 \neq \frac{1}{12} = \frac{1}{4}(\frac{1}{3}) = P(X = 1)P(Y = 4)$$

(b)

Give a probability table for random variables U and V that have the same marginals as X and Y but are independent.

We need probabilities of the rows to all be equal to one another. We need P(U = 1) = P(U = 3), and P(V = 2) = P(V = 3) = P(V = 4). To satisfy these conditions we have:

	U		
V	1	2	3
2	1/12	1/6	1/12
3	1/12	1/6	1/12
4	1/12	1/6	1/12

Q4

Suppose X and Y are independent random variables, where both have the same (marginal) geometric(p) distribution for $0 < p < 1$. Find the conditional distribution of X given $X + Y = k$ (integer $k \geq 2$).

$$P(X = x | S = k) = \frac{P(X = x, S = k)}{P(S = k)}$$

For the top part of the equation:

$$P(X = x, Y = k - x) = P(X = x)P(Y = k - x)$$

And:

$$P(X = x) = (1 - p)^{x-1}p$$

And:

$$P(Y = k - x) = (1 - p)^{k-x-1}p$$

Given X and Y are independent:

$$P(X = x, S = k) = (1 - p)^{x-1}p(1 - p)^{k-x-1}p = P(X = x, S = k) = (1 - p)^{k-2}p^2$$

Consider then the bottom of the initial formula:

$$P(S = k) = \sum_{x=1}^{k-1} P(X = x, Y = k - x) = (k - 1)(1 - p)^{k-2}p^2$$

Returning the overall, conditional formula, we have:

$$P(X = x | S = k) = \frac{P(X = x, S = k)}{P(S = k)} = \frac{(1 - p)^{k-2}p^2}{(k - 1)(1 - p)^{k-2}p^2} = \frac{1}{k - 1}$$

For x = 1, 2, ..., k - 1

We then conclude

$$X | (X + Y = k) \sim U(1, k - 1)$$

Q5: 4.27 Casella & Berger

Let $X \sim n(\mu, \sigma^2)$, and let $Y \sim n(\gamma, \sigma^2)$. Suppose X and Y are independent. Define: U = X + Y and V = X - Y. Show that U and V are independent normal random variables. Find the distribution of each of them.

Hey, so I am not entirely sure why this problem uses n to represent the normal instead of N as we usually do, but for this problem I just stayed (or attempted to stay) consistent with their format.

$$U \sim n(\mu + \gamma, 2\sigma^2)$$

$$V \sim n(\mu - \gamma, 2\sigma^2)$$

For X and Y independent, we may write the joint pdf as the product of the marginals:

$$f_{X,Y}(x,y) = f_X(x)f_Y(f) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(y-\gamma)^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2}e^{\frac{-\left((x-\mu)^2+(y-\gamma)^2\right)}{2\sigma^2}}$$

$$U = X + Y \text{ and } V = X - Y \rightarrow X = \frac{1}{2}(U+V), Y - \frac{1}{2}(U-V)$$

We then have the Jacobian:

$$J = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$|J| = \left| \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix} \right| = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = |-\frac{1}{4} - \frac{1}{4}| = |-\frac{1}{2}| = \frac{1}{2}$$

And we have the new support of U, V as:

$$(u,v) \in (-\infty, \infty)\text{x}(-\infty, \infty)$$

Using the above relations, we have:

$$f_{UV}(u,v) = \frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[\left(\frac{u+v}{2}-\mu\right)^2+\left(\frac{u-v}{2}-\gamma\right)^2\right]}\frac{1}{2}$$

$$f_{UV}(u,v) = \frac{1}{4\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[2\left(\frac{u}{2}\right)^2-u(\mu+\gamma)+\left(\frac{\mu+\gamma}{2}\right)^2+2\left(\frac{v}{2}\right)^2-v(\mu-\gamma)+\left(\frac{\mu-\gamma}{2}\right)^2\right]}$$

$$f_{UV}(u,v) = g(u)\frac{1}{4\pi\sigma^2}e^{-\frac{1}{2(2\sigma^2)}(u-(\mu+\gamma))^2}h(v)e^{-\frac{1}{2(2\sigma^2)}(v-(\mu-\gamma))^2}$$

Such that we have shown:

For:

$$f_U(u) = g(u)\frac{1}{4\pi\sigma^2}e^{-\frac{1}{2(2\sigma^2)}(u-(\mu+\gamma))^2}$$

And:

$$f_V(v) = h(v)e^{-\frac{1}{2(2\sigma^2)}(v-(\mu-\gamma))^2}$$

6

$$f_{UV}(u,v) = f_U(u)f_V(v)$$

and we conclude that U and V are independent.

Q6: 4.42 Casella & Berger

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.

Said differently, we want to find: $\rho_{XY,Y}$

$$\rho_{XY,Y} = \frac{Cov(XY,Y)}{\sigma_{XY}\sigma_Y} = \frac{E(XY^2) - \mu_X\mu_Y}{\sigma_{XY}\sigma_Y} = \frac{EXEY^2 - \mu_X\mu_Y\mu_Y}{\sigma_{XY}\sigma_Y}$$

$$\rho_{XY,Y} = \frac{\mu_X(\sigma_Y^2 + \mu_Y^2) - \mu_X\mu_Y\mu_Y}{\sigma_{XY}\sigma_Y} = \frac{\mu_X(\sigma_Y^2)}{\sigma_{XY}\sigma_Y}$$

With note of X and Y being independent, we then have the denominator:

$$\sigma_{XY}^2 = E(XY)^2 - [E(XY)]^2 = EX^2EY^2 - (EX)^2(EY)^2 = (Var(X) + E[X]^2)(Var(Y) + E[Y]^2) - (EX)^2(EY)^2$$

Evaluating with known quantities, we have:

$$\sigma_{XY}^2 = (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 = \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2 + \mu_X^2\mu_Y^2 - \mu_X^2\mu_Y^2 = \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2$$

Substituting into the original equation gives us:

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

Q7

Suppose that random variables X_1, X_2, X_3 have joint pdf $f(x_1, x_2, x_3) = 6$ for $0 < x_1 < x_2 < x_3 < 1$.

(a)

Are X_1, X_2, X_3 independent? Give an intuitive explanation.

No they are not independent, as knowing the value of one of the random variables limits the range of possible values that the other random variables can take.

(b)

Find the marginal pdf of X_2 and identify the distribution of X_2 as a member of an important family of distributions.

Let us first calculate the marginal of X_2:

$$f_{X_2}(x_2) = \int_0^{x_2} \int_{x_2}^1 6 \, dx_3 \, dx_1 = \int_{x_2}^1 6 \, dx_3 = 6(1 - x_2)$$

$$f_{X_2}(x_2) = \int_0^{x_2} 6(1 - x_2) \, dx_1 = 6(1 - x_2) \int_0^{x_2} 1 \, dx_1 = 6(1 - x_2)x_2$$

As $0 < x_2 < 1$, the range of x_2 gives us a clue as to the distribution of X_2. Specifically:

$$X_2 \sim Beta(2, 2)$$

(c)

Find the conditional pdf $f(x_1, x_3 | x_2)$ of X_1, X_3 given $X_2 = x_2 \in (0, 1)$.

$$f(x_1, x_3 \mid x_2) = \frac{f(x_1, x_2, x_3)}{f_{X_2}(x_2)} = \frac{6}{6x_2(1 - x_2)} = \frac{1}{x_2(1 - x_2)}$$

(d)

Show that X_1, X_3 are independent given $X_2 = x_2 \in (0, 1)$ (i.e., think of what must be true of the conditional pdf of $f(x_1, x_3 | x_2)$ in this case).

Let us start by considering the joint pdf:

$$f(x_1, x_3 \mid x_2) = \frac{1}{x_2(1 - x_2)}$$

Let us turn then to the marginals, with the goal of showing the product of the marginals is equal to the joint as specified. To that end:

i.

$$f_{X_1|X_2}(x_1 \mid x_2) = \int_{x_2}^{1} f(x_1, x_3 \mid x_2)\, dx_3 = \int_{x_2}^{1} \frac{1}{x_2(1 - x_2)}\, dx_3$$

$$f_{X_1|X_2}(x_1 \mid x_2) = \frac{1}{x_2(1 - x_2)} \int_{x_2}^{1} 1\, dx_3 = \frac{1}{x_2(1 - x_2)}(1 - x_2) = \frac{1}{x_2}$$

ii.

$$f_{X_3|X_2}(x_3 \mid x_2) = \int_{0}^{x_2} f(x_1, x_3 \mid x_2)\, dx_1 = \int_{0}^{x_2} \frac{1}{x_2(1 - x_2)}\, dx_1$$

$$f_{X_3|X_2}(x_3 \mid x_2) = \frac{1}{x_2(1 - x_2)} \int_{0}^{x_2} 1\, dx_1 = \frac{1}{x_2(1 - x_2)} x_2 = \frac{1}{1 - x_2}$$

Taking the results of i. and ii., we then have:

$$f_{X_1|X_2}(x_1 \mid x_2) = \frac{1}{x_2}$$

$$f_{X_3|X_2}(x_3 \mid x_2) = \frac{1}{1 - x_2}$$

$$f(x_1, x_3 \mid x_2) = \frac{1}{x_2(1 - x_2)} = f_{X_1|X_2}(x_1 \mid x_2) f_{X_3|X_2}(x_3 \mid x_2)$$

Since the conditional joint pdf can be written as a product of the marginal conditional pdfs, X_1 and X_3 are independent given $X_2 = x_2$.

(e)

Find the covariance of X_1 and X_3 given $X_2 = x_2$.

From part (d), we know that X_1 and X_3 are independent given $X_2 = x_2$, so we know that conditional on/given $X_2 = x_2$ that the covariance of X_1 and X_3 would be zero, or:

$$Cov(X_1, X_3 | X_2 = x_2) = 0$$