Problem Set 3

Daniel Shapiro

9/13/2022

Question 1 Background:

Suppose that X and Y are identical and independently distributed (i.i.d.) random variables distributed $\mathcal{N}(\mu_X, \sigma_X^2)$ and $\mathcal{N}(\mu_Y, \sigma_Y^2)$, respectively. Note that σ_X represents the standard deviation of X, and σ_X^2 the variance. Find the following, expressed in terms of μ and σ:

1a) $E(7X - 6Y + 12)$

Additivity and homogeneity are two properties of expectations. Thus, we can write this expression as $7E[X] - 6E[Y] + 12$, or $7\mu_X - 6\mu_Y + 12$.

1b) Var$(X + 5Y)$

We know that $Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$. Inserting a 1 and 5 as a and b, we get $\sigma_X^2 + 25\sigma_Y^2 + 10Cov(X, Y)$. $Cov(X, Y) = E[XY] - E[X]E[Y] = E[(X - \mu_X)(Y - \mu_Y)]$. So then we get: $Var(X + 5Y) = \sigma_X^2 + 25\sigma_Y^2 + 10E[(X - \mu_X)(Y - \mu_Y)]$.

Because we know that the two variables are so-called "i.i.d." variables, we know that the co-variance is 0. Thus, we can remove the last term entirely and end up with $Var(X + 5Y) = \sigma_X^2 + 25\sigma_Y^2$.

1c) $E(5X^2 - 12XY + 16Y^2)$

First, we can split these up according to properties shown in 1a. So we get $5E[X^2] - 12E[XY] + 16E[Y^2]$.

Now, we can already deal with the first and last bits of this expression. We know that $\sigma_X^2 = E[X^2] - \mu_X^2$ and $\sigma_Y^2 = E[Y^2] - \mu_Y^2$. Thus, $E[X^2] = \sigma_X^2 + \mu_X^2$ and $E[Y^2] = \sigma_Y^2 + \mu_Y^2$. So we can set this as $5\sigma_X^2 + 5\mu_X^2 + 12E[XY] + 16\sigma_Y^2 + 16\mu_Y^2$.

Finally, the question tells us that X and Y are independent random variables. So we know therefore that $E[XY] = E[X]E[Y]$. We can write this as $\mu_X \mu_Y$.

Our final expression: $5\sigma_X^2 + 5\mu_X^2 + 12\mu_X\mu_Y + 16\sigma_Y^2 + 16\mu_Y^2$.

Question 2 Background:

Papers like this one have found that married couples tend to sort along ideological lines ("assortative mating"). Say we know the true distribution of ideology among heterosexual married couples along a scale of 1 (very conservative) to 4 (very liberal). The joint distribution for X (the woman's ideology) and Y (the man's ideology) is:

	X			
Y	1	2	3	4
1	.12	.05	.01	.01
2	.08	.17	.05	.02
3	.02	.01	.23	.05
4	.02	.02	.04	.10

2a) What is the expected value of X and the expected value of Y?

To find the expected value of X, we can run the formula: $E[X] = \Sigma(x)P(x)$. So we multiply the value x by each corresponding probability and add them up. So $E[X] = ((1*.12)+(1*.08)+...+(4*.10))$. The final answer is **2.45**.

To find the expected value of Y, we can run the same formula, but for Y instead of X. So we multiply the value y by each corresponding probability and add them up. So $E[Y] = ((1*.12)+(1*.05)+...+(4*.10))$. The final answer is **2.48**.

2b) What are the variances of X and of Y?

The variance of X or $(Var(X))$ can be defined as $\Sigma(x-\mu_x)^2 p(x)$. The variance of Y or $(Var(Y))$ can be defined as $\Sigma(y-\mu_y)^2 p(y)$. So we already know that $\mu_x = 2.45$ and $\mu_y = 2.48$; now, we need to find the probability of X at each point and the probability of Y at each point. To do this, we can just add up the proper rows and columns.

For X, we get: (.24 (when x = 1), .25 (when x = 2), .33 (when x = 3), and .18 (when x = 4)).

For Y, we get: (.19 (when y = 1), .32 (when y = 2), .31 (when y = 3), and .18 (when y = 4)).

Now, equation time:

$Var(X) = (1-2.45)^2 * .24 + (2-2.245)^2 * .25 + (3-2.45)^2 * .33 + (4-2.45^2 * .18)$. This ends up equaling **1.0875**.

$Var(Y) = (1-2.48)^2 * .19 + (2-2.248)^2 * .32 + (3-2.48)^2 * .31 + (4-2.48^2 * .18)$. This ends up equaling **0.9896**.

2c) What is the covariance of X and Y?

We know that the covariance of X and Y can be defined by the function:

$Cov(X,Y) = E[XY] - E[X]E[Y]$

First, we know $E[X]E[Y]$ – it's 2.48*2.45. This is 6.076.

To find E[XY], we need to multiply each value of Y and X by their probability, and then add everything together. Like so:

$((1*1*.12)+(1*2*.08)+(1*3*.03)+...+(4*4*.10))$.

After a somewhat time-consuming process, we end up with 6.70.

Then, we subtract: 6.70-6.076 = **0.624**.

2d) What is the correlation between X and Y?

The function for correlation is:

$Cor[X,Y] = \frac{Cov[X,Y]}{\sqrt{V[X]V[Y]}}$.

Luckily, we know all of these values already. Plugging in our numbers, we get:

$Cor[X,Y] = \frac{0.624}{\sqrt{1.0875*0.9896}} = \mathbf{0.6015}$.

2e) Find $E[Y|X=x]$ for $x = 1,2,3,4$. Find the probability mass function of the random variable $E[Y|X]$.

For purposes of this question, we will set E[X|Y] = Z. So for this question, we need to take each value of X=x and multiply it by the given probability value divided by the marginal probability of Y=y. For $E[Y|X=1]$, we thus get the following:

$((1*\frac{.12}{.24}) + (2*\frac{.08}{.24}) + (3*\frac{.02}{.24}) + (4*\frac{.02}{.24})) = 1.75$

For $E[Y|X=2]$, we get:

$((1*\frac{.05}{.25}) + (2*\frac{.17}{.25}) + (3*\frac{.01}{.25}) + (4*\frac{.02}{.25})) = 2$

For $E[Y|X=3]$, we get:

$((1*\frac{.01}{.33}) + (2*\frac{.05}{.33}) + (3*\frac{.23}{.33}) + (4*\frac{.04}{.33})) = 2.91$

For $E[Y|X=4]$, we get:

$((1*\frac{.01}{.18}) + (2*\frac{.02}{.18}) + (3*\frac{.05}{.18}) + (4*\frac{.10}{.18})) = 3.33$

Writing out the probability mass function is relatively easy if we've already figured out Z at x = 1,2,3,4. Below:

$$p(z) = \begin{cases} .24 & \text{if Z = 1.75} \\ .25 & \text{if Z = 2} \\ .33 & \text{if Z = 2.91} \\ .18 & \text{if Z = 3.33} \end{cases}$$

2f) Find $Var(Y|X)$.