STAT 260 Spring 2023: Assignment 7

Due: Friday March 24th BEFORE 11:59pm PT to Crowdmark

Please read the instructions below and in the Written Assignment 6 assignment on Crowdmark.

For full marks, your work must be neatly written, and contain enough detail that it is clear how you arrived at your solutions. Messy, poorly formatted work will receive deductions, or may not be graded at all.

Talking to your classmates about assigned work is a healthy practice that is encouraged. However, in the end, each person is expected to write their own solutions, in their own words, and in a way that reflects their own understanding. Assignment questions are not to be posted to homework "help" websites.

Late policy: Late assignments will be accepted until the final cutoff of 11:59pm on Sunday March 26th. Solutions submitted within 1 hour of the Friday deadline will have a 5% late penalty automatically applied within Crowdmark. Solutions submitted after 1 hour of the Friday deadline but before the final Sunday cutoff will have a 20% late penalty applied. Solutions submitted after the final Sunday cutoff will be graded for feedback, but marks will not be awarded.

1. As part of a promotion, a restaurant offers fixed options for a \$20, a \$30, a \$40, and a \$50 meal. Let the random variable X represent the cost of the meal a customer orders, and let the random varialbe Y represent how much the customer leaves as a tip. The joint distribution for X and Y is given as:

		X				
	$p(x,y)$	20	30	40	50	
Y	3	0.14	0.09	0.02	0.03	.28
	5	0.11	0.08	0.06	0.05	.30
	7	0.02	0.04	0.07	0.10	.23
	10	0.01	0.01	0.05	0.12	.19
		.28	.22	.20	.30	

(a) **[1 mark]** If a person orders a meal that is \$30 or greater, what is the probability that they leave at least a \$7 tip?

$$P(Y \geq 7 \mid X \geq 30) = \frac{P(Y \geq 7 \cap X \geq 30)}{P(X \geq 30)}$$

$$= \frac{.04 + .07 + .10 + .01 + .05 + .12}{.22 + .20 + .30}$$

$$= \frac{0.39}{0.72} = 0.5417$$

(b) **[1 mark]** Calculate the covariance of X and Y. (That is, calculate $Cov(X, Y)$.) Using your covariance answer, what can you say about the independence of variables X and Y?

x	20	30	40	50
$f(x)$	.28	.22	.20	.30

y	3	5	7	10
$f(y)$	.28	.30	.23	.19

$$E(X) = 20(.28) + 30(.22) + 40(.20) + 50(.30) = 35.2$$

$$E(Y) = 3(.28) + 5(.30) + 7(.23) + 10(.19) = 5.85$$

		X			
	$X \cdot Y$	20	30	40	50
Y	3	60	90	120	150
	5	100	150	200	250
	7	140	210	280	350
	10	200	300	400	500

$$E(XY) = 60(.14) + 90(.09) + 120(.02) + 150(.03)$$
$$+ 100(.11) + 150(.08) + 200(.06) + 250(.05)$$
$$+ 140(.02) + 210(.04) + 280(.07) + 350(.10)$$
$$+ 200(.01) + 300(.01) + 400(.05) + 500(.12)$$
$$= 221.7$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 221.7 - (35.2)(5.85) = 15.78$$

Since $cov(X,Y) \neq 0$, X and Y are not independent.

(c) **[1 mark]** Calculate the correlation coefficient of X and Y. (That is, calculate $\rho = Corr(X, Y)$).

$$E(X^2) = 20^2(.28) + 30^2(.22) + 40^2(.20) + 50^2(.30) = 1380$$

$$V(X) = E(X^2) - (E(X))^2 = 1380 - (35.2)^2 = 140.96$$

$$E(Y^2) = 3^2(.28) + 5^2(.30) + 7^2(.23) + 10^2(.19) = 40.29$$

$$V(Y) = E(Y^2) - (E(Y))^2 = 40.29 - 5.85^2 = 6.0675$$

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{15.78}{\sqrt{140.96}\sqrt{\quad 0675}}$$

$$= 0.5396$$

2. **[1 mark]** A bakery puts in a weekly order to three different flour suppliers, but with recent problems in the supply chain they don't always receive the full amount that was ordered. For $i = 1, 2, 3$, let X_i be the weight of flour (in kg) that is received from supplier i. Suppose the weights received from each supplier are normally distributed, the weights from each supplier are independent, and it is known that

$$\mu_{X_1} = 2000, \qquad \sigma_{X_1} = 100, \qquad \mu_{X_2} = 4000, \qquad \sigma_{X_2} = 300, \qquad \mu_{X_3} = 3000, \qquad \sigma_{X_3} = 250$$

Suppose that flour from supplier 1 costs \$2.50 per kg, flour from supplier 2 costs \$3 per kg, and flour from supplier 3 costs \$4 per kg.

What is the probability that the total weekly cost of the received flour is greater than \$31,000?

total cost $= 2.5X_1 + 3X_2 + 4X_3$

X_1, X_2, X_3 normal $\Rightarrow 2.5X_1 + 3X_2 + 4X_3$ is also normal

Want $P(2.5X_1 + 3X_2 + 4X_3 > 31000)$.

$E(2.5X_1 + 3X_2 + 4X_3) = 2.5E(X_1) + 3E(X_2) + 4E(X_3)$

$= 2.5(2000) + 3(4000) + 4(3000) = 29000$

$V(2.5X_1 + 3X_2 + 4X_3) = (2.5)^2 V(X_1) + 3^2 V(X_2) + 4^2 V(X_3)$

$= (2.5)^2 (100)^2 + 3^2 (300)^2 + 4^2 (250)^2 = 1872500$

$Z = \dfrac{\text{r.v.} - \text{expected value}}{\text{st. dev of r.v.}}$

$P(2.5X_1 + 3X_2 + 4X_3 > 31000) = P\left(Z > \dfrac{31000 - 29000}{\sqrt{1872500}}\right)$

$= P(Z > 1.46) = 1 - P(Z \leq 1.46)$

$= 1 - 0.9279$

$= 0.0721$

3. **[1 mark]** In a particular town, bedrock in the ground has a mean depth of 23 meters with a standard deviation of 8.41 meters. A gravity survey is used to determine the depth of the bedrock in an area that is to be excavated. When the gravity survey is performed at a random sample of 39 locations, what is the probability that the mean bedrock depth is greater than 25.5 meters?

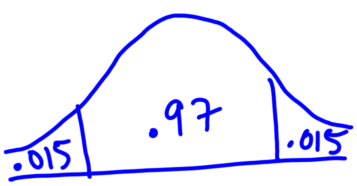
$n = 39 \geq 30$ so $\bar{x}$ is normal

$$P(\bar{X} > 25.5)$$

$$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{25.5 - 23}{8.41/\sqrt{39}}\right) = P(Z > 1.86)$$

$$= 1 - P(Z \leq 1.86) = 1 - 0.9686 = 0.0314$$

4. **[1 mark]** Soil pH refers to the degree of acidity or akalinity of the soil. A farmer has had some trouble growing crops this season and suspects that there is a problem with the pH level of their soil. A collection of 47 soil samples shows a mean pH level of 8.2 and a standard deviation of 1.09. Calculate a 97% confidence interval for μ, the mean pH level of the soil on the farm.



$$z_{\alpha/2} = z_{.015} = 2.17$$

$$\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n} = 8.2 \pm 2.17 \cdot 1.09/\sqrt{47}$$

$$= [7.8550, \quad 8.5450]$$