ECO 3150

ANSWERS TO ASSIGNMENT 2

K. Day Winter 2019

---

1. (a) Let M represent "Married," S "Single," $LT1$ "Less than one year," and $GE1$ "greater than or equal to 1 year." Then we need to find $P(M)$. Since $LT1$ and $GE1$ are mutually exclusive and collectively exhaustive events,

$$\begin{aligned} P(M) &= P(M \cap GE1) + P(M \cap LT1) \\ &= 0.64 + 0.13 \\ &= 0.77 \, . \end{aligned}$$

Thus the probability that a randomly-chosen salesperson was married is 0.77.

(b) For this question we need to find $P(LT1)$. Since M and S are mutually exclusive and collectively exhaustive events,

$$\begin{aligned} P(LT1) &= P(LT1 \cap M) + P(LT1 \cap S) \\ &= 0.13 + 0.06 \\ &= 0.19 \, . \end{aligned}$$

Thus the probability that a randomly-chosen salesperson left the job within a year is 0.19.

(c) We need to find $P(LT1 \,|\, S)$.

$$P(LT1 \,|\, S) = \frac{P(LT1 \cap S)}{P(S)} = \frac{P(LT1 \cap S)}{1 - P(M)} = \frac{0.06}{1 - 0.77} = \frac{0.06}{0.23} = 0.261 \, .$$

Therefore the probability that a randomly-chosen single salesperson left the job within a year is 0.261. (Note that the probability of a complement rule was needed to compute $P(S)$ in the denominator of this conditional probability.)

(d) We need to find $P(M \,|\, GE1)$.

$$P(M \,|\, GE1) = \frac{P(M \cap GE1)}{P(GE1)} = \frac{P(M \cap GE1)}{1 - P(GE1)} = \frac{0.64}{1 - 0.19} = \frac{0.64}{0.81} = 0.790 \, .$$

Therefore the probability that a randomly-chosen salesperson who stayed in the job for at least a year was married is 0.790.

(e) Marital status and job tenure will be independent if $P(MS_i \cap JT_j) = P(MS_i)P(JT_j)$ for all possible combinations of marital status (MS_i) and job tenure (JT_j). First consider $P(M \cap LT1)$:

$$P(M) \cdot P(LT1) = (0.77)(0.19) = 0.1463 \neq 0.13 = P(M \cap LT1) \, .$$

Thus marital status and job tenure are not statistically independent.

Note that other correct ways to answer this question would be to compare the answers to (a) and (d), or the answers to (b) and (c). They would be identical if marital status and job tenure were independent.

2. The actual profit earned by the store owner will depend on the number of copies of the paper requested by customers. If four copies are ordered, the total cost incurred by the store owner will be

$$TC = 4(0.70) = 2.80,$$

where cost is measured in dollars. The total revenue will be $0.90X$ if $X \leq 4$, and $4(0.90) = 3.60$ if more than four copies are requested (since all four copies on hand will be sold). Profit (π) will, in general, be given by

$$\pi = TR - TC - L\ ,$$

where L is the loss of goodwill. Since the loss of goodwill will be incurred only if the store owner receives 5 requests, the formula for profit given that the store owner has purchased 4 copies of the newspaper will be

$$\begin{aligned}\pi &= TR - TC - L \\ &= \begin{cases} 3.60 - 2.80 - 0.05 & \text{if } X = 5 \\ 0.90X - 2.80 & \text{if } X \leq 4 \end{cases} \\ &= \begin{cases} 0.75 & \text{if } X = 5 \\ 0.90X - 2.80 & \text{if } X \leq 4 \end{cases} .\end{aligned}$$

The following table shows the profit computed using the above formula, for each value of X, given that four copies have been ordered:

x	$P(x)$	π
0	0.12	-2.80
1	0.16	-1.90
2	0.18	-1.00
3	0.32	-0.01
4	0.14	0.80
5	0.08	0.75

Since profits are a function of the random variable X, expected profits can be computed using the general formula for the expectation of a function $g(X)$ of the random variable X. (Alternatively, the probability distribution of profits will be the same as that of X.) The formula is

$$E[\pi(x)] = \sum_x \pi(x)P(x)\ .$$

Applying this formula, we obtain

$$\begin{aligned}E[\pi(4)] &= (-2.8)(0.12) + (-1.9)(0.16) + (-1.0)(0.18) + (-0.1)(0.32) \\ &\quad +(0.08)(0.14) + (0.75)(0.08) = -0.68\ .\end{aligned}$$

Thus the expected profit if four newspapers will be ordered is -\$0.68; in other words, the owner will incur a loss of 68 cents.

3. (a) The formula for deriving the marginal distribution of X from the joint distribution of X and Y is

$$P_X(x) = \sum_y P(x, y) \ .$$

This function must be evaluated for each possible value of X:

$$\begin{aligned} P_X(0) &= P(0,0) + P(0,1) + P(0,2) + P(0,3) = 0.07 + 0.07 + 0.06 + 0.02 = 0.22 \\ P_X(1) &= P(1,0) + P(1,1) + P(1,2) + P(1,3) = 0.09 + 0.06 + 0.07 + 0.04 = 0.26 \\ P_X(2) &= P(2,0) + P(2,1) + P(2,2) + P(2,3) = 0.06 + 0.07 + 0.14 + 0.16 = 0.43 \\ P_X(3) &= P(3,0) + P(3,1) + P(3,2) + P(3,3) = 0.01 + 0.01 + 0.03 + 0.04 = 0.09 \ . \end{aligned}$$

The marginal probability function of X can be summarized in tabular form as

x	0	1	2	3
$P_X(x)$	0.22	0.26	0.43	0.09

It is easy to verify that these probabilities sum to 1 as is required.

The mean or expected value of X can be computed as follows:

$$\begin{aligned} E(X) &= \sum_x x P_X(x) \\ &= 0(0.22) + 1(0.26) + 2(0.43) + 3(0.09) \\ &= 1.39 \ . \end{aligned}$$

Therefore the mean number of tests taken by students per day is 1.39.

(b) The formula for deriving the marginal distribution of Y from the joint distribution of X and Y is

$$P_Y(y) = \sum_x P(x, y) \ .$$

This function must be evaluated for each possible value of Y:

$$\begin{aligned} P_Y(0) &= P(0,0) + P(1,0) + P(2,0) + P(3,0) = 0.07 + 0.09 + 0.06 + 0.01 = 0.23 \\ P_Y(1) &= P(0,1) + P(1,1) + P(2,1) + P(3,1) = 0.07 + 0.06 + 0.07 + 0.01 = 0.21 \\ P_Y(2) &= P(0,2) + P(1,2) + P(2,2) + P(3,2) = 0.06 + 0.07 + 0.14 + 0.03 = 0.30 \\ P_Y(3) &= P(0,3) + P(1,3) + P(2,3) + P(3,3) = 0.02 + 0.04 + 0.16 + 0.04 = 0.26 \ . \end{aligned}$$

The marginal probability function of Y can be summarized in the following table:

y	0	1	2	3
$P_Y(y)$	0.23	0.21	0.30	0.26

The mean or expected value of Y can be computed as follows:

$$\begin{aligned} E(Y) &= \sum_y y P_Y(y) \\ &= 0(0.23) + 1(0.21) + 2(0.30) + 3(0.26) \\ &= 1.59 \ . \end{aligned}$$

Therefore the mean number of snacks eaten by students per day is 1.59.

(c) The general formula for the conditional probability function of Y given $X = x = 3$ is

$$P_{Y\,|\,X}(y\,|\,x) = \frac{P(x,y)}{P_X(x)} \qquad \text{or} \qquad P_{Y\,|\,X}(y\,|\,3) = \frac{P(3,y)}{P_X(3)}\ .$$

This formula can be applied to compute each individual value of $P_{Y/X}(y/3)$ as follows:

$$\begin{aligned}
P_{Y\,|\,X}(0\,|\,3) &= \frac{P(3,0)}{P_X(3)} = \frac{0.01}{0.09} = 0.111 \\
P_{Y\,|\,X}(1\,|\,3) &= \frac{P(3,1)}{P_X(3)} = \frac{0.01}{0.09} = 0.111 \\
P_{Y\,|\,X}(2\,|\,3) &= \frac{P(3,2)}{P_X(3)} = \frac{0.03}{0.09} = 0.333 \\
P_{Y\,|\,X}(3\,|\,3) &= \frac{P(3,3)}{P_X(3)} = \frac{0.04}{0.09} = 0.444\ .
\end{aligned}$$

Thus the conditional probability function of Y, given $X = 3$, is

y	0	1	2	3
$P_{Y\,\|\,X}(y\,\|\,3)$	0.111	0.111	0.333	0.444

This function tells us the probability of obtaining each value of Y, given that $X = 3$.

(d) The covariance between X and Y is computed as follows:

$$\begin{aligned}
\text{cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\
&= \sum_x \sum_y xyP(x,y) - \mu_X\mu_Y \\
&= (0)(0)(0.07) + (0)(1)(0.07) + (0)(2)(0.06) + (0)(3)(0.02) \\
&+ (1)(0)(0.09) + (1)(1)(0.06) + (1)(2)(0.07) + (1)(3)(0.04) \\
&+ (2)(0)(0.06) + (2)(1)(0.07) + (2)(2)(0.14) + (2)(3)(0.16) \\
&+ (3)(0)(0.01) + (3)(1)(0.01) + (3)(2)(0.03) + (3)(3)(0.04) - (1.39)(1.59) \\
&= 0.3399\ .
\end{aligned}$$

Therefore the covariance between X and Y is 0.3399.

(e) The number of snacks and the number of tests will be independent of each other if $P(x,y) = P_X(x)P_Y(y)$ for all possible combinations of x and y. Consider first the combination $x = 0$, $y = 0$:

$$P_X(0)P_Y(0) = (0.22)(0.23) = 0.0506 \neq 0.07 = P(0,0)\ .$$

Thus the number of snacks and the number of tests are not independent, because the condition does not hold for the combination $x = 0$, $y = 0$.

4. By definition,

$$\text{var}(cX) = E\left\{[cX - E(cX)]^2\right\} .$$

First consider $E(cX)$. Since c is a constant, $E(cX) = cE(X)$. Thus

$$\begin{aligned}\text{var}(cX) &= E\left\{[cX - cE(X)]^2\right\} \\ &= E\left\{c^2[X - E(X)]^2\right\} \\ &= c^2E\left\{[X - E(X)]^2\right\} \\ &= c^2\text{var}(X)\end{aligned}$$

since $\text{var}(X) = E\{[X - E(X)]^2\}$ by definition.

This proves one part of the equality given in the question. To prove the second part, we can proceed as follows:

$$\begin{aligned}\text{var}(cX) &= E\left\{[cX - cE(X)]^2\right\} \\ &= E[c^2X^2 - 2c^2XE(X) + c^2(E(X))^2] \\ &= E(c^2X^2) - E[2c^2XE(X)] + E[c^2(E(X))^2] \\ &= c^2E(X^2) - 2c^2E(X)E(X) + c^2[E(X)]^2 \\ &\quad \text{since } c \text{ and } E(X) \text{ are constants} \\ &= c^2E(X^2) - 2c^2[E(X)]^2 + c^2[E(X)]^2 \\ &= c^2E(X^2) - c^2[E(X)]^2 .\end{aligned}$$

(Note that one could have proved this more quickly using the result that $\text{var}(X) = E(X^2) - \mu_X^2$.)

Therefore

$$\text{var}(cX) = c^2E(X^2) - c^2[E(X)]^2 = c^2\text{var}(X) .$$