Introduction to Statistics - Homework #2

Exercise 3.34

An airline charges the following baggage fees: $25 for the first bag and $35 for the second. Suppose 54% of passengers have no checked luggage, 34% have one piece of checked luggage, and 12% have two pieces. We assume a negligible portion of people check more than two bags.

(1) Build a probability model, compute the average revenue per passenger, and compute the corresponding standard deviation.

Solution:

The random variable X represents the baggage fee per passenger.

The expected value $E(X)$ is:

$$E(X) = (0 \times 0.54) + (25 \times 0.34) + (60 \times 0.12) = 0 + 8.5 + 7.2 = 15.7.$$

To compute variance, first find $E(X^2)$:

$$E(X^2) = (0^2 \times 0.54) + (25^2 \times 0.34) + (60^2 \times 0.12).$$

$$= 0 + (625 \times 0.34) + (3600 \times 0.12).$$

$$= 0 + 212.5 + 432 = 644.5.$$

Now, compute variance:

$$\text{Var}(X) = E(X^2) - (E(X))^2 = 644.5 - (15.7)^2 = 644.5 - 246.49 = 398.01.$$

Finally, compute the standard deviation:

$$\text{SD}(X) = \sqrt{398.01} \approx 19.95.$$

(2) About how much revenue should the airline expect for a flight of 120 passengers? With what standard deviation? Note any assumptions you make and if you think they are justified.

Solution:

The total revenue T for 120 passengers follows:

$$E(T) = 120 \times E(X) = 120 \times 15.7 = 1884.$$

Since each passenger's baggage fee is independent, the variance of total revenue is:

$$\text{Var}(T) = 120 \times \text{Var}(X) = 120 \times 398.01 = 47761.2.$$

The standard deviation is:

$$\text{SD}(T) = \sqrt{47761.2} \approx 218.54.$$

Exercise 3.47

Suppose we have n independent observations $X_1, X_2, \ldots, X_n$ from a distribution with mean μ and standard deviation σ. What is the expectation and variance of the sample mean of these n values:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}?$$

Solution:

By linearity of expectation:

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n\mu = \mu.$$

Since the X_i are independent, the variance of a sum equals the sum of the variances:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\sigma^2.$$

Using $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$, we get:

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}.$$

Extra problem 1

Suppose that X and Y have a joint probability density function (pdf)

$$f(x,y) = \begin{cases} \frac{3}{2}y^2, & 0 \le x \le 2,\ 0 \le y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

and the marginal pdfs

$$f_X(x) = \begin{cases} \frac{1}{2}, & 0 \le x \le 2 \\ 0, & \text{otherwise.} \end{cases}, \quad f_Y(y) = \begin{cases} 3y^2, & 0 \le y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

(1) Determine if X and Y are independent.

Solution:

X and Y are independent if and only if $f(x,y) = f_X(x)\, f_Y(y)$ for all x, y. We check:

$$f_X(x)\, f_Y(y) = \frac{1}{2} \cdot 3y^2 = \frac{3}{2}y^2 = f(x,y).$$

Since the factorization holds, X and Y are independent.

(2) Find $E(Y)$ and $\mathrm{Var}(Y)$.

Solution:

$$E(Y) = \int_0^1 y f_Y(y)\, dy = \int_0^1 y(3y^2)\, dy = 3\int_0^1 y^3\, dy = \frac{3}{4}.$$

$$E(Y^2) = \int_0^1 y^2 f_Y(y)\, dy = \int_0^1 y^2(3y^2)\, dy = 3\int_0^1 y^4\, dy = \frac{3}{5}.$$

$$\mathrm{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80}.$$

Extra problem 2

Suppose that X and Y have a joint probability mass function (pmf)

$$f(x,y) = \begin{cases} \frac{1}{12}(x+y), & x \in \{0,2\},\ y \in \{0,1,2\}, \\ 0, & \text{otherwise.} \end{cases}$$

(1) Find the marginal pmf of X, $f_X(x)$.

Solution: The joint probability distribution table is:

$x\backslash y$	**0**	**1**	**2**	$f_X(x)$
0	0	1/12	1/6	1/4
2	1/6	1/4	1/3	3/4
$f_Y(y)$	1/6	1/3	1/2	1

$$f_X(x) = \sum_{\forall y} f(x,y), \qquad f_X(0) = 0 + \tfrac{1}{12} + \tfrac{1}{6} = \tfrac{1}{4}, \qquad f_X(2) = \tfrac{1}{6} + \tfrac{1}{4} + \tfrac{1}{3} = \tfrac{3}{4}.$$

$$\text{Therefore, } f_X(x) = \begin{cases} 1/4, & x = 0, \\ 3/4, & x = 2, \\ 0, & \text{otherwise.} \end{cases}$$

(2) Find the marginal pmf of Y, $f_Y(y)$.

Solution: The marginal pmf of Y is obtained by summing each column:

$$f_Y(y) = \sum_{\forall x} f(x,y), \quad f_Y(0) = 0 + \tfrac{1}{6} = \tfrac{1}{6}, \cdots, \quad f_Y(2) = \tfrac{1}{6} + \tfrac{1}{3} = \tfrac{1}{2}.$$

$$\text{Therefore, } f_Y(y) = \begin{cases} 1/6, & y = 0, \\ 1/3, & y = 1, \\ 1/2, & y = 2, \\ 0, & \text{otherwise.} \end{cases}$$

(3) Find $E(X)$ and $E(Y)$.

Solution:

$$E(X) = \sum_{\forall x} x\, f_X(x) = 0 \cdot \tfrac{1}{4} + 2 \cdot \tfrac{3}{4} = \tfrac{3}{2},$$

$$E(Y) = \sum_{\forall y} y\, f_Y(y) = 0 \cdot \tfrac{1}{6} + 1 \cdot \tfrac{1}{3} + 2 \cdot \tfrac{1}{2} = \tfrac{4}{3}.$$

(4) Find $E(XY)$ and $\mathrm{Cov}(X,Y)$.

Solution: Since $xy = 0$ whenever $x = 0$ or $y = 0$, only the cells $(x,y) \in \{(2,1),(2,2)\}$ contribute to the sum. Reading $f(2,1) = 1/4$ and $f(2,2) = 1/3$ from the table above:

$$E(XY) = \sum_{\forall x,y} xy\, f(x,y) = (2)(1) \cdot \tfrac{1}{4} + (2)(2) \cdot \tfrac{1}{3} = \tfrac{1}{2} + \tfrac{4}{3} = \tfrac{11}{6}.$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)\,E(Y) = \tfrac{11}{6} - \tfrac{3}{2} \cdot \tfrac{4}{3} = \tfrac{11}{6} - 2 = -\tfrac{1}{6}.$$