

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) Let random variable X represent the baggage fee per passenger. Find $E(X)$ and $SD(X)$.

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 $E(X^2)$:

$$\begin{aligned} E(X^2) &= (0^2 \times 0.54) + (25^2 \times 0.34) + (60^2 \times 0.12). \\ &= 0 + (625 \times 0.34) + (3600 \times 0.12). \\ &= 0 + 212.5 + 432 = 644.5. \end{aligned}$$

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:

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

(2) About how much revenue should the airline expect for a flight of 120 passengers? Given that each passenger's baggage fee is independent, what is the standard deviation of the revenue for a flight of 120 passengers?

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:

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

Exercise 3.47

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

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

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:

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

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

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

Extra problem 1

Suppose that X and Y have a joint probability density function

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

(1) Find the marginal pdf of X and Y , $f_X(x)$ and $f_Y(y)$.

Solution:

$$f_X(x) = \int_0^1 f(x, y) dy = \int_0^1 \frac{3}{2}y^2 dy = \frac{1}{2}, \quad 0 \leq x \leq 2.$$

$$f_Y(y) = \int_0^2 f(x, y) dx = \int_0^2 \frac{3}{2}y^2 dx = 3y^2, \quad 0 \leq y \leq 1.$$

(2) 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.

(3) Find $E(Y)$ and $\text{Var}(Y)$.

Solution:

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

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

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

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

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

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

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

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

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

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

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

Solution:

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

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

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

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:

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

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

Extra problem 3

Suppose that X and Y have a joint probability density function

$$f(x, y) = \begin{cases} e^{-x-y}, & x > 0, y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(1) Find the marginal densities $f_X(x)$ and $f_Y(y)$.

Solution:

$$f_X(x) = \int_0^{\infty} e^{-x-y} dy = e^{-x}, \quad x > 0.$$

$$f_Y(y) = \int_0^{\infty} e^{-x-y} dx = e^{-y}, \quad y > 0.$$

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

Solution:

We compute $E(X)$ by integration by parts:

$$E(X) = \int_0^{\infty} x e^{-x} dx.$$

Let $u = x$, $v = e^{-x}$.

$$\int_0^{\infty} x e^{-x} dx = \left[-x e^{-x} \right]_0^{\infty} + \int_0^{\infty} e^{-x} dx.$$

The boundary term vanishes because $\lim_{x \rightarrow \infty} x e^{-x} = 0$. Thus,

$$E(X) = 0 + \left[-e^{-x} \right]_0^{\infty} = 1.$$

Similarly, $E(Y) = 1$.

(3) Find $E(XY)$ and $\text{Cov}(X, Y)$.

Solution:

Since $f(x, y) = e^{-x}e^{-y} = f_X(x)f_Y(y)$, X and Y are independent. Thus,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

$$E(XY) = E(X)E(Y) = 1 \cdot 1 = 1.$$

(4) Find $P(X < Y)$.

Solution:

$$P(X < Y) = \int_0^{\infty} \int_{y=x}^{\infty} e^{-(x+y)} dy dx = \int_0^{\infty} e^{-x} \left[-e^{-y} \right]_{y=x}^{\infty} dx = \int_0^{\infty} e^{-x} e^{-x} dx$$