STA2001 Tutorial 8

1. 4.1-6. The torque required to remove bolts in a steel plate is rated as very high, high, average, and low, and these occur about 25%, 35%, 20%, and 20% of the time, respectively. Suppose $n = 31$ bolts are rated; what is the probability of rating 9 very high, 10 high, 7 average, and 5 low? Assume independence of the 31 trials.

   Solution:

   We define the following random variables first:

   X : the number of bolt rated very high

   Y : the number of bolt rated high

   Z : the number of bolt rated average

   $K = n - X - Y - Z$: the number of bolt rated low

   Furthermore, we know that

$$P_X = P(X = 1) = 0.25, \ P_Y = P(Y = 1) = 0.35$$

$$P_Z = P(Z = 1) = 0.20, \ P_K = 1 - P_X - P_Y - P_Z = 0.20$$

   By recognizing this experiment takes a similar form with the binomial distribution (independent trials but with 4 different possible outcomes), the pmf is given by

$$\begin{aligned} f(x, y, z, k) &= P(X = x, Y = y, Z = z, K = n - x - y - z) \\ &= \frac{n!}{x!y!z!(n-x-y-z)!} P_X^x P_Y^y P_Z^z (1 - P_X - P_Y - P_Z)^{n-x-y-z} \\ &= \frac{31!}{9!10!7!5!} 0.25^9 0.35^{10} 0.2^7 0.2^5 \\ &= 0.0045. \end{aligned}$$

   In this question, these random variables follow the so-called multinomial distribution. Mention the homework problem: the marginal of multinomial distribution is binominal

2. 4.2-7 Let the joint pmf of X and Y be

$$f(x,y) = 1/4$$

where $(x,y) \in S = \{(0,0),(1,1),(1,-1),(2,0)\}$.

(a) Are X and Y independent?

(b) Calculate $\text{cov}(X,Y)$ and ρ .

This exercise also illustrates the fact that dependent random variables can have a correlation coefficient of zero.

Solution:

(a) Given the joint pmf $f(x,y)$, we could compute the marginal pmf:

$$f_X(x) = \begin{cases} \frac{1}{4}, & x = 0,2 \\ \frac{1}{2}, & x = 1 \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{4}, & y = 1,-1 \\ \frac{1}{2}, & y = 0 \end{cases}$$

Since

$$f_X(1)f_Y(1) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \neq \frac{1}{4} = f(1,1)$$

Therefore, X and Y are not independent.

(b) Since we have the marginal pmf, we could compute

$$\mu_X = E(X) = \sum_x x f_X(x) = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1$$

$$\mu_Y = E(Y) = \sum_y y f_Y(y) = 1 \times \frac{1}{4} - 1 \times \frac{1}{4} + \frac{1}{2} \times 0 = 0$$

From the joint pmf, we have

$$\begin{aligned} \mu_{XY} &= E(XY) \\ &= \sum_{(x,y)\in S} xyf(x,y) \\ &= 0 \times 0 \times \frac{1}{4} + 1 \times (-1) \times \frac{1}{4} + 1 \times 1 \times \frac{1}{4} + 2 \times 0 \times \frac{1}{4} \\ &= 0 \end{aligned}$$

Therefore, the covariance and coefficient of correlation are given by

$$\text{Cov}(X,Y) = E(XY) - \mu_X\mu_Y = 0 - 1 \times 0 = 0$$

$$\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = 0$$

3. 4.2-8. A certain raw material is classified as to moisture content X (in percent) and impurity Y (in percent). Let X and Y have the joint pmf given by

y\x	1	2	3	4
1	0.05	0.05	0.15	0.1
2	0.1	0.2	0.3	0.05

(a) Find the marginal pmfs, the means, and the variances of X and Y, respectively.

(b) Find the covariance and the correlation coefficient of X and Y.

(c) If additional heating is needed with high moisture content and additional filtering with high impurity such that the additional cost is given by the function $C = 2X + 10Y^2$ in dollars, find $E(C)$.

Solution:

(a) Given the joint pmf, we could compute the marginal pmf as follows:

$$f_X(x) = \begin{cases} 0.15, & x = 1, 4 \\ 0.25, & x = 2 \\ 0.45, & x = 3 \end{cases} \qquad f_Y(y) = \begin{cases} 0.35, & y = 1 \\ 0.65, & y = 2 \end{cases}$$

Thus,

$$\begin{aligned} \mu_X &= \sum_{x=1}^{4} x f_X(x) = 1 \times 0.15 + 2 \times 0.25 + 3 \times 0.45 + 4 \times 0.15 = 2.6 \\ \mu_Y &= \sum_{y=1}^{2} y f_Y(y) = 1 \times 0.35 + 2 \times 0.65 = 1.65 \\ \sigma_X^2 &= \sum_{x=1}^{4} x^2 f_X(x) - 2.6^2 \\ &= 1^2 \times 0.15 + 2^2 \times 0.25 + 3^2 \times 0.45 + 4^2 \times 0.15 - 2.6^2 = 0.84 \\ \sigma_Y^2 &= \sum_{y=1}^{2} y^2 f_Y(y) - 1.65^2 = 1^2 \times 0.35 + 2^2 \times 0.65 - 1.65^2 = 0.2275 \end{aligned}$$

(b) From the given table (joint pmf), we could easily compute

$$\begin{aligned} E(XY) &= 1 \times 1 \times 0.05 + 1 \times 2 \times 0.1 + 2 \times 1 \times 0.05 + 2 \times 2 \times 0.2 \\ &\quad + 3 \times 1 \times 0.15 + 3 \times 2 \times 0.3 + 4 \times 1 \times 0.1 + 4 \times 2 \times 0.05 \\ &= 4.2 \end{aligned}$$

Hence, the covariance and the coefficient of correlation are given by

$$\mathrm{Cov}(X, Y) = E(XY) - \mu_X \mu_Y = 4.2 - 2.6 \times 1.65 = -0.09$$

$$\rho = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{-0.09}{\sqrt{0.84}\sqrt{0.2275}} = -0.20588$$

(c) Utilize the linearity property of expectation, we have

$$
\begin{aligned}
E(2X + 10Y^2) &= 2E(X) + 10E(Y^2) \\
&= 2 \times 2.6 + 10 \times E(Y^2) \\
&= 5.2 + 10 \times (1^2 \times 0.35 + 2^2 \times 0.65) \\
&= 5.2 + 10 \times 2.95 \\
&= 34.7
\end{aligned}
$$