STA2001 Tutorial 11

1. 5.3-7 The distributions of incomes in two cities follow the two Pareto-type pdfs

$$f(x) = \frac{2}{x^3}, \quad 1 < x < \infty, \quad \text{and} \quad g(y) = \frac{3}{y^4}, \quad 1 < y < \infty,$$

respectively (Suppose that X and Y are independent). Here one unit represents \$20,000. One person with income is selected at random from each city. Let X and Y be their respective incomes. Compute $P(X < Y)$.

Solution:

Since X and Y are independent, the joint pdf is given by

$$f(x, y) = f_X(x) g_Y(y) = 6x^{-3}y^{-4}$$

where $1 < x < \infty,\ 1 < y < \infty.$

Thus $P(X < Y)$ is computed as

$$\begin{aligned} P(X < Y) &= \int_1^\infty \int_1^y 6x^{-3}y^{-4} dx\, dy \\ &= \int_1^\infty -3x^{-2}y^{-4}\Big|_1^y dy \\ &= \frac{3}{5}y^{-5} - y^{-3}\Big|_1^\infty \\ &= \frac{2}{5} \end{aligned}$$

2. 5.3-8 Suppose two independent claims are made on two insured homes, where each claim has pdf

$$f(x) = \frac{4}{x^5}, \quad 1 < x < \infty,$$

in which the unit is \$1000. Find the expected value of the larger claim.

Hint: If X_1 and X_2 are the two identical and independent claims and $Y = \max(X_1, X_2)$, then

$$G(y) = P(Y \le y) = P(X_1 < y)P(X_2 < y) = [P(X \le y)]^2.$$

Find $g(y) = G'(y)$ and $E(Y)$.

Solution:

From the hint, we have

$$G(y) = [P(X \le y)]^2 = \left[\int_1^y \frac{4}{x^5} dx\right]^2 = \left[1 - \frac{1}{y^4}\right]^2, \quad 1 < y < \infty$$

Take the derivative of $G(y)$ with respect to y, we have the probability density $g(y)$,

$$g(y) = G'(y) = 2\left(1 - \frac{1}{y^4}\right)\left(\frac{4}{y^5}\right) = \frac{8}{y^5} - \frac{8}{y^9}$$

Thus the expectation of Y is given by

$$\begin{aligned} E(Y) &= \int_1^\infty y8(y^{-5} - y^{-9})dy \\ &= \int_1^\infty 8(y^{-4} - y^{-8})dy \\ &= \left(-\frac{8}{3}y^{-3} + \frac{8}{7}y^{-7}\right)\Big|_1^\infty \\ &= \frac{32}{21} \end{aligned}$$

3. 5.3-20. Let X and Y be independent random variables with nonzero variances. Find the correlation coefficient of $W = XY$ and $V = X$ in terms of the means and variances of X and Y.

Solution:

Note that $W = XY$ and $V = X$, and X and Y are independent. Let μ_X, μ_Y and σ_X^2, σ_Y^2 be the means and variances. Therefore,

$$\begin{aligned}
\operatorname{Cov}(W, V) &= E(WV) - E(W)E(V) \\
&= E(XYX) - E(XY)E(X) \\
&= E(X^2)\mu_Y - \mu_X\mu_Y\mu_X \\
&= (\sigma_X^2 + \mu_X^2)\mu_Y - \mu_X^2\mu_Y \\
&= \sigma_X^2\mu_Y
\end{aligned}$$

$$\begin{aligned}
\sigma_W\sigma_V &= \sqrt{\sigma_W^2\sigma_V^2} \\
&= \sqrt{(E(W^2) - (E(W))^2)\,\sigma_X^2} \\
&= \sqrt{(E(X^2Y^2) - (E(XY))^2)\,\sigma_X^2} \\
&= \sqrt{(E(X^2)E(Y^2) - \mu_X^2\mu_Y^2)\,\sigma_X^2} \\
&= \sqrt{((\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2)\,\sigma_X^2} \\
&= \sqrt{\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2} \cdot \sigma_X
\end{aligned}$$

The coefficient of correlation is given by

$$\rho = \frac{\operatorname{Cov}(W, V)}{\sigma_W\sigma_V} = \frac{\sigma_X\mu_Y}{\sqrt{\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2}}$$

4. The number of people who enter an elevator on the ground floor, denoted as X, is a Poisson random variable with mean λ. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where the others get off, compute the expected number of stops that the elevator will make in order to discharge all of its passengers.

   - Poisson pmf: $p_X(n) = \frac{e^{-\lambda}\lambda^n}{n!}, n \geq 0$
   - $\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$

   Solution. Let Y denote the number of stops the elevator make. Define the indicator random variable

$$I_k = \begin{cases} 1 & \text{if there exist a person getting off at } k\text{-th floor} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

   we have $Y = \sum_{k=1}^{N} I_k$.

   We have

$$\begin{aligned} \Pr\left(I_k = 0 \mid X = n\right) &= \Pr(\text{ no person choose floor } k \mid X = n) \\ &= \left(1 - \frac{1}{N}\right)^n \end{aligned} \tag{2}$$

   Hence by Law of Total Probability

$$\begin{aligned} \Pr\left(I_k = 0\right) &= \sum_{n=0}^{\infty} \Pr\left(I_k = 0 \mid X = n\right) p_X(n) \\ &= \sum_{n=0}^{\infty} \left(1 - \frac{1}{N}\right)^n \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda} \sum_{n=0} \frac{[\lambda(1 - 1/N)]^n}{n!} \\ &= e^{-\lambda} e^{\lambda(1-1/N)} = e^{-\lambda/N} \end{aligned} \tag{3}$$

   Thus $\mathbb{E}\left[I_k\right] = \Pr\left(I_k = 1\right) = 1 - e^{-\lambda/N}$. Therefore by linearity of expectation,

$$\mathbb{E}[Y] = \sum_{k=1}^{N} \mathbb{E}\left[I_k\right] = N\left(1 - e^{-\lambda/N}\right) \tag{4}$$

   The indicator function is a widely-used technique for its linearity of expectation, see more interesting examples in e.g.

   https://web.stanford.edu/class/archive/cs/cs161/cs161.1236/Resources/prob.pdf