Exam 1 2023

Mengxiang Jiang
Stat 610 Distribution Theory

October 1, 2025

Problem 1. The mean and variance of $X \sim \text{binomial}(n, p)$ are np and $np(1-p)$, respectively. What is $\mathsf{E}(X(X+1))$? Try to simplify.

$$\begin{aligned}
\mathsf{E}(X(X+1)) &= \mathsf{E}(X^2 + X) \\
&= \mathsf{E}(X^2) + \mathsf{E}(X) \\
&= \mathsf{Var}(X) + [\mathsf{E}(X)]^2 + \mathsf{E}(X) \\
&= np(1-p) + (np)^2 + np \\
&= n^2p^2 + np.
\end{aligned}$$

Problem 2. Let $Y \sim \text{uniform}(1, 3)$.

(a) Find the pdf for $V = \sqrt{Y}$.

(b) Find $\mathsf{E}(V)$ and $v_{0.50}$ (the 0.50-quantile for V). Are they the same value?

(a)

$$f_Y(y) = \begin{cases} \frac{1}{3-1} = \frac{1}{2} & \text{if } 1 < y < 3, \\ 0 & \text{otherwise.} \end{cases}$$

Since $V = \sqrt{Y}$, we have $Y = V^2$ and $\frac{dY}{dV} = 2V$. By the change of variables formula, we have

$$f_V(v) = f_Y(v^2)\left|\frac{dY}{dV}\right| = f_Y(v^2) \cdot 2v.$$

Note that $1 < y < 3$ is equivalent to $1 < v^2 < 3$, or $1 < v < \sqrt{3}$. Thus,

$$f_V(v) = \begin{cases} \frac{1}{2} \cdot 2v = v & \text{if } 1 < v < \sqrt{3}, \\ 0 & \text{otherwise.} \end{cases}$$

(b)

$$
\begin{aligned}
\mathsf{E}(V) &= \int_{-\infty}^{\infty} v f_V(v) dv \\
&= \int_{1}^{\sqrt{3}} v \cdot v dv \\
&= \int_{1}^{\sqrt{3}} v^2 dv \\
&= \left.\frac{v^3}{3}\right|_1^{\sqrt{3}} \\
&= \frac{(\sqrt{3})^3}{3} - \frac{1^3}{3} = \frac{3\sqrt{3}-1}{3}.
\end{aligned}
$$

To find $v_{0.50}$, we first find the cdf of V.

$$
F_V(v) = \begin{cases} 0 & \text{if } v \leq 1, \\ \int_1^v t dt = \left.\frac{t^2}{2}\right|_1^v = \frac{v^2}{2} - \frac{1}{2} = \frac{v^2-1}{2} & \text{if } 1 < v < \sqrt{3}, \\ 1 & \text{if } v \geq \sqrt{3}. \end{cases}
$$

The 0.50-quantile must satisfy the equation

$$F_V(v_{0.50}) = 0.50.$$

Thus,

$$\frac{v_{0.50}^2 - 1}{2} = 0.50 \implies v_{0.50}^2 - 1 = 1 \implies v_{0.50}^2 = 2 \implies v_{0.50} = \sqrt{2}.$$

Note that $\mathsf{E}(V) = \frac{3\sqrt{3}-1}{3} \approx 1.40$ and $v_{0.50} = \sqrt{2} \approx 1.41$. They are not the same value.

Problem 3. 80% of connections to a certain server are of Type A and the other 20% are of Type B. Let T be the lifetime of a connection.
For Type A, $T \sim$ exponential(1): $\mathsf{P}(T \leq t|\text{Type A}) = 1 - e^{-t}$.
For Type B, $T \sim$ gamma(2, .5): $\mathsf{P}(T \leq t|\text{Type B}) = 1 - (1 + 2t)e^{-2t}$.

(a) Provide the cdf and pdf for a randomly chosen connection. That is, give the unconditional $\mathsf{P}(T \leq t)$ and then determine the pdf.

(b) The moment generating function for T is $M_T(s) = \frac{0.8}{1-s} + \frac{0.2}{(1-0.5s)^2}$ for $s < 1$. Use this to find $\mathsf{E}(T)$.

(a) By the law of total probability, we have

$$\mathsf{P}(T \leq t) = \mathsf{P}(T \leq t|\text{Type A})\mathsf{P}(\text{Type A}) + \mathsf{P}(T \leq t|\text{Type B})\mathsf{P}(\text{Type B}).$$

$$
\begin{aligned}
\mathsf{P}(T \le t) &= (1 - e^{-t}) \cdot 0.8 + [1 - (1 + 2t)e^{-2t}] \cdot 0.2 \\
&= 0.8 - 0.8e^{-t} + 0.2 - 0.2(1 + 2t)e^{-2t} \\
&= 1 - 0.8e^{-t} - 0.2(1 + 2t)e^{-2t},
\end{aligned}
$$

Since this is a cdf, the pdf can be found by differentiating the cdf:

$$
\begin{aligned}
f_T(t) &= \frac{d}{dt}\mathsf{P}(T \le t) \\
&= \frac{d}{dt}\left[1 - 0.8e^{-t} - 0.2(1 + 2t)e^{-2t}\right] \\
&= 0 + 0.8e^{-t} - 0.2\frac{d}{dt}\left[(1 + 2t)e^{-2t}\right] \\
&= 0.8e^{-t} - 0.2\left[(1 + 2t)(-2e^{-2t}) + 2e^{-2t}\right] \quad \text{(by product rule)} \\
&= 0.8e^{-t} - 0.2\left[-2(1 + 2t)e^{-2t} + 2e^{-2t}\right] \\
&= 0.8e^{-t} - 0.2\left[(-2 - 4t + 2)e^{-2t}\right] \\
&= 0.8e^{-t} - 0.2(-4te^{-2t}) = 0.8e^{-t} + 0.8te^{-2t}.
\end{aligned}
$$

(b) Using the given mgf, we can find $\mathsf{E}(T)$ as follows.

$$
\begin{aligned}
\mathsf{E}(T) &= M_T'(0) \\
&= \frac{d}{ds}\left[\frac{0.8}{1 - s} + \frac{0.2}{(1 - 0.5s)^2}\right]_{s=0} \\
&= 0.8\frac{d}{ds}(1 - s)^{-1} + 0.2\frac{d}{ds}(1 - 0.5s)^{-2}\bigg|_{s=0} \\
&= 0.8(-(1 - s)^{-2})(-1) + 0.2(-2)(1 - 0.5s)^{-3}(-0.5)\bigg|_{s=0} \\
&= 0.8(1 - s)^{-2} + 0.2(1 - 0.5s)^{-3}\bigg|_{s=0} \\
&= 0.8(1 - 0)^{-2} + 0.2(1 - 0)^{-3} = 0.8 + 0.2 = 1.
\end{aligned}
$$

If we use the pdf found in part (a), we can also find $\mathsf{E}(T)$ as follows.

$$\begin{aligned}
\mathsf{E}(T) &= \int_{-\infty}^{\infty} t f_T(t) dt \\
&= \int_{0}^{\infty} t(0.8e^{-t} + 0.8te^{-2t}) dt \\
&= 0.8 \int_{0}^{\infty} te^{-t} dt + 0.8 \int_{0}^{\infty} t^2 e^{-2t} dt \\
&= 0.8 \cdot 1 + 0.8 \cdot \frac{1}{4} \int_{0}^{\infty} (2t)^2 e^{-2t} dt \quad \left(\text{using } \int_{0}^{\infty} y^m e^{-y} dy = m!\right) \\
&= 0.8 + 0.8 \cdot \frac{1}{4} \int_{0}^{\infty} y^2 e^{-y} dy \quad (\text{using } y = 2t) \\
&= 0.8 + 0.8 \cdot \frac{1}{4} \cdot 2! \quad \left(\text{using } \int_{0}^{\infty} y^m e^{-y} dy = m!\right) \\
&= 0.8 + 0.8 \cdot \frac{2}{8} = 0.8 + 0.2 = 1.
\end{aligned}$$

Problem 4. For a giveaway promotional event, a store randomly selects 40 pairs of shoes from its stock of 1000 pairs. The stock consists of 400 Hoka pairs, 350 New Balance pairs and 250 Saucony pairs. Give expressions for (a) the chance that x Hoka pairs, y New Balance pairs and z Saucony pairs are selected and for (b) the chance at least w Hoka or New Balance pairs are selected. Be complete.

(a) By the multivariate hypergeometric distribution, we have

$$\mathsf{P}(X = x, Y = y, Z = z) = \frac{\binom{400}{x}\binom{350}{y}\binom{250}{z}}{\binom{1000}{40}},$$

where X, Y and Z are the number of Hoka, New Balance and Saucony pairs selected, respectively, and x, y and z are nonnegative integers such that $x+y+z = 40$, $x \leq 400$, $y \leq 350$ and $z \leq 250$.

(b) Let A be the event that at least w Hoka or New Balance pairs are selected. Then $A = B \cup C$, where B is the event that at least w Hoka pairs are selected and C is the event that at least w New Balance pairs are selected. By the inclusion-exclusion principle, we have

$$\mathsf{P}(A) = \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(B \cap C).$$

We can calculate $\mathsf{P}(B)$, $\mathsf{P}(C)$ and $\mathsf{P}(B \cap C)$ as follows.

$$\mathsf{P}(B) = \sum_{x=w}^{40} \sum_{y=0}^{40-x} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}},$$

$$\mathsf{P}(C) = \sum_{y=w}^{40} \sum_{x=0}^{40-y} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}},$$

$$\mathsf{P}(B \cap C) = \sum_{x=w}^{40} \sum_{y=w}^{40-x} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}}.$$

Thus,

$$\begin{aligned}
\mathsf{P}(A) &= \mathsf{P}(B) + \mathsf{P}(C) - \mathsf{P}(B \cap C) \\
&= \sum_{x=w}^{40} \sum_{y=0}^{40-x} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}} \\
&\quad + \sum_{y=w}^{40} \sum_{x=0}^{40-y} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}} \\
&\quad - \sum_{x=w}^{40} \sum_{y=w}^{40-x} \frac{\binom{400}{x}\binom{350}{y}\binom{250}{40-x-y}}{\binom{1000}{40}}.
\end{aligned}$$