Exam 1 2023

Mengxiang Jiang
Stat 610 Distribution Theory

September 30, 2025

Problem 1. Consider the following continuous cdf for a random variable W.

$$F_W(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \frac{x}{3} & \text{if } 0 < x < 1, \\ \frac{1}{3} & \text{if } 1 \leq x < 2, \\ \frac{2x}{3} - 1 & \text{if } 2 \leq x < 3, \\ 1 & \text{if } 3 \leq x. \end{cases}$$

Find the probability density function (pdf) and $\mathsf{E}(W)$.

Since $F_W(x)$ is continuous, we can take the derivative to get the pdf.

$$f_W(x) = \begin{cases} \frac{1}{3} & \text{if } 0 < x < 1, \\ 0 & \text{if } 1 \leq x < 2, \\ \frac{2}{3} & \text{if } 2 \leq x < 3, \\ 0 & \text{otherwise.} \end{cases}$$

Then we can calculate $\mathsf{E}(W)$ as follows.

$$\begin{aligned} \mathsf{E}(W) &= \int_{-\infty}^{\infty} x f_W(x) dx \\ &= \int_0^1 x \cdot \frac{1}{3} dx + \int_1^2 x \cdot 0 dx + \int_2^3 x \cdot \frac{2}{3} dx \\ &= \frac{x^2}{6}\bigg|_0^1 + 0 + \frac{x^2}{3}\bigg|_2^3 \\ &= \frac{1}{6} + 0 + (3 - \frac{4}{3}) = \frac{11}{6}. \end{aligned}$$

Problem 2. A population of 1000 school districts consists of 150 districts with 1 high school, 350 with 2 high schools, 300 with 3 high schools, and 200 with 4 high schools. A researcher selects two districts at random, without replacement. Given that the two districts have the same number of high schools, what is the chance they have x high schools, for each

$x = 1, 2, 3, 4$? Computable expressions suffice.

Let A be the event that the two selected districts have the same number of high schools. Let B_x be the event that the two selected districts have x high schools. We want to find $\mathsf{P}(B_x|A)$ for $x = 1, 2, 3, 4$. By the definition of conditional probability, we have

$$\mathsf{P}(B_x|A) = \frac{\mathsf{P}(B_x \cap A)}{\mathsf{P}(A)} = \frac{\mathsf{P}(B_x)}{\mathsf{P}(A)},$$

since $B_x \subseteq A$. We can calculate $\mathsf{P}(A)$ as follows.

$$\begin{aligned}
\mathsf{P}(A) &= \mathsf{P}(B_1) + \mathsf{P}(B_2) + \mathsf{P}(B_3) + \mathsf{P}(B_4) \\
&= \frac{\binom{150}{2}}{\binom{1000}{2}} + \frac{\binom{350}{2}}{\binom{1000}{2}} + \frac{\binom{300}{2}}{\binom{1000}{2}} + \frac{\binom{200}{2}}{\binom{1000}{2}} \\
&= \frac{\binom{150}{2} + \binom{350}{2} + \binom{300}{2} + \binom{200}{2}}{\binom{1000}{2}}.
\end{aligned}$$

Then we can calculate $\mathsf{P}(B_x|A)$ as follows.

$$\mathsf{P}(B_x|A) = \frac{\mathsf{P}(B_x)}{\mathsf{P}(A)} = \frac{\frac{\binom{n_x}{2}}{\binom{1000}{2}}}{\frac{\binom{150}{2}+\binom{350}{2}+\binom{300}{2}+\binom{200}{2}}{\binom{1000}{2}}} = \frac{\binom{n_x}{2}}{\binom{150}{2} + \binom{350}{2} + \binom{300}{2} + \binom{200}{2}},$$

where n_x is the number of districts with x high schools.

Problem 3. Suppose T has gamma(2,1) distribution.

(a) Determine the cumulative distribution function (cdf), for all real values, and provide an equation that the median must solve. (Do not try to solve it.)

(b) Find the pdf for $Y = T^{1/3}$.

(a) Since the pdf of T is given by

$$f_T(t) = \begin{cases} te^{-t} & \text{if } t > 0, \\ 0 & \text{otherwise,} \end{cases}$$

we can calculate the cdf of T as follows.

$$\begin{aligned}
F_T(t) &= \int_{-\infty}^{t} f_T(x)dx \\
&= \int_{0}^{t} xe^{-x}dx \quad (\text{since } f_T(x) = 0 \text{ for } x \leq 0) \\
&= -xe^{-x} - e^{-x}\Big|_0^t \quad (\text{by integration by parts}) \\
&= -te^{-t} - e^{-t} + 1 \\
&= 1 - e^{-t} - te^{-t}.
\end{aligned}$$

Thus the cdf of T is given by

$$F_T(t) = \begin{cases} 0 & \text{if } t \le 0, \\ 1 - e^{-t} - te^{-t} & \text{if } t > 0. \end{cases}$$

The median must satisfy the equation

$$F_T(m) = \frac{1}{2}.$$

(b) The pdf of T is given by

$$f_T(t) = \begin{cases} te^{-t} & \text{if } t > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since $Y = T^{1/3}$, we have $T = Y^3$ and $\frac{dT}{dY} = 3Y^2$. Then we can calculate the pdf of Y as follows.

$$f_Y(y) = f_T(y^3)\left|\frac{dT}{dY}\right| = f_T(y^3)\cdot 3y^2 = \begin{cases} 3y^5 e^{-y^3} & \text{if } y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Problem 4. Derive the moment generating function (mgf) for the geometric(p) pmf. (Recall that $\sum_{k=1}^{\infty} a^k = \frac{a}{1-a}$ for $|a| < 1$.)

The pmf of a geometric(p) random variable X is given by

$$f_X(x) = \begin{cases} p(1-p)^{x-1} & \text{if } x = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Then we can calculate the mgf of X as follows.

$$\begin{aligned} M_X(t) &= \mathsf{E}(e^{tX}) \\ &= \sum_{x=1}^{\infty} e^{tx} f_X(x) \\ &= \sum_{x=1}^{\infty} e^{tx} p(1-p)^{x-1} \\ &= \frac{p}{1-p}\sum_{x=1}^{\infty} \left[e^t(1-p)\right]^x \\ &= \frac{p}{1-p}\cdot\frac{e^t(1-p)}{1-e^t(1-p)} \quad (\text{since } |e^t(1-p)| < 1) \\ &= \frac{pe^t}{1-(1-p)e^t}. \end{aligned}$$