ECON 6190

Problem Set 1

Gabe Sekeres

September 10, 2024

1. $\overline{X}$ is described as follows:

   (a) The sampling distribution of $\overline{X}$ is:

   $$P\{\overline{X} = x_i\} = \begin{cases} \frac{1}{8} & x_i = 0 \\ \frac{3}{8} & x_i = \frac{1}{3} \\ \frac{3}{8} & x_i = \frac{2}{3} \\ \frac{1}{8} & x_i = 1 \end{cases}$$

   Note that this is the same distribution as the probability mass function of $Y \sim \text{Binomial}(3, \frac{1}{2})$ when x_i is multiplied by 3.

   (b) The mean of $\overline{X}$ is $\frac{1}{2}$

   (c) The variance of $\overline{X}$ is:

   $$Var(\overline{X}) = \sum_i P\{\overline{X} = x_i\} \cdot (x_i - \mu)^2 = \left(\frac{1}{8} \cdot \frac{1}{4}\right) + \left(\frac{3}{8} \cdot \frac{1}{36}\right) + \left(\frac{3}{8} \cdot \frac{1}{36}\right) + \left(\frac{1}{8} \cdot \frac{1}{4}\right) = \frac{1}{12}$$

2. There are $\binom{5}{2} = 10$ possible choices of two families. The set of possible income means is

   $$\{1.5, 2, 2.5, 2.5, 3, 3, 3.5, 3.5, 4, 4.5\}$$

   which are obtained by choosing, respectively

   $$\{(1,2), (1,3), (1,4), (2,3), (1,5), (2,4), (2,5), (3,4), (3,5), (4,5)\}$$

   Thus, the sampling distribution of the sample mean is

   $$P\{\overline{X} = x_i\} = \begin{cases} \frac{1}{10} & x_i = 1.5 \\ \frac{1}{10} & x_i = 2 \\ \frac{1}{5} & x_i = 2.5 \\ \frac{1}{5} & x_i = 3 \\ \frac{1}{5} & x_i = 3.5 \\ \frac{1}{10} & x_i = 4 \\ \frac{1}{10} & x_i = 4.5 \end{cases}$$

3. We have that $f(x) = f(-x) \; \forall \, x \in \mathbb{R}$.

(a) ***Proof.***

$$\begin{aligned} F(-x) &= \int_{-\infty}^{x} f(-t)dt \\ &= \int_{-\infty}^{x} f(t)dt \\ &= F(x) \\ &= P\{X \le x\} \\ &= 1 - P\{X > x\} \\ &= 1 - F(x) \end{aligned}$$

Where the last line follows from the fact that X is a continuous random variable, so $P\{X > x\} = P\{X \ge x\}$. $\square$

(b) ***Proof.*** From above, we have that

$$F(-x) = 1 - F(x)$$

Taking $x = 0$, we get that

$$F(0) = 1 - F(0)$$

which implies that

$$P\{X \le 0\} = 1 - P\{X \le 0\} = P\{X \ge 0\}$$

And since the total density of X above and below 0 is equivalent, it must be the case that $\mathbb{E}[X] = 0$. $\square$

4. Show that $\mathbb{E}[s] \le \sigma$, where $s = \sqrt{s^2}$ and s^2 is the sample variance.

Proof. We have from the notes that $\mathbb{E}[s^2] = \sigma^2$. From Jensen's Inequality, we have that $(\mathbb{E}[s])^2 \le \mathbb{E}[s^2]$, so we have that $(\mathbb{E}[s])^2 \le \sigma^2$. Since all values are non-negative, we can take the square root and get $\mathbb{E}[s] \le \sigma$. $\square$

5. Show that $\min_a \mathbb{E}\,|X - a| = \mathbb{E}\,|X - m|$, where m is the median of X.

Proof. Note that:

$$\min_a \mathbb{E}\,|X - a| = \min_a \left(\int_{-\infty}^{a} |x - a| dF(x) \right)$$

Taking derivatives with respect to a, to find the point at which the function is minimized, we get that

$$\begin{aligned} \frac{\partial}{\partial a} \left(\int_{-\infty}^{a} |x - a| dF(x) \right) &= \int_{-\infty}^{a} -1 dF(x) + \int_{a}^{\infty} 1 dF(x) \\ &= -F(a) + (1 - F(a)) \\ &= 1 - 2F(a) = 0 \end{aligned}$$

This implies that the function is minimized when $F(a) = \frac{1}{2}$. Precisely, it is minimized at a, where $P\{X \le a\} = \frac{1}{2}$, which is exactly the median. Thus, $\min_a \mathbb{E}\,|X - a| = \mathbb{E}\,|X - m|$. $\square$

6. Let X be a random variable with conditional density

$$f(x \mid \theta) = \begin{cases} \frac{1}{\theta} & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

Usually we treat parameter θ as a constant. Now suppose $\theta > 0$ is treated as a random variable with density

$$g(\theta) = \begin{cases} \theta e^{-\theta} & \theta > 0 \\ 0 & \theta \leq 0 \end{cases}$$

where we use notation θ as both the random variable and the specific values it can take.

(a) We have that the joint density is $f(x, \theta) = f(x \mid \theta) g(\theta) = e^{-\theta}$. To find the marginal distribution, we simply integrate $f(x, \theta)$ with respect to θ. We get

$$\int_{-\infty}^{\infty} f(x, \theta) d\theta = \int_{-\infty}^{x} f(x, \theta) d\theta + \int_{x}^{\infty} f(x, \theta) d\theta = 0 + \int_{x}^{\infty} f(x, \theta) d\theta + 0$$

Thus,

$$f(x) = \int_{x}^{\infty} e^{-\theta} d\theta = \left(-e^{-\theta}\right|_{x}^{\infty} = 0 - (-e^{-x}) = e^{-x}$$

(when $x < \theta$, and $f(x) = 0$ otherwise).

(b) To find the conditional density, recall that from the definition of joint densities, $f(x, \theta) = g(\theta \mid x) f(x)$. Thus, we have that

$$g(\theta \mid x) = \frac{f(x, \theta)}{f(x)} = \frac{e^{-\theta}}{e^{-x}} = e^{x - \theta}$$

(when $x < \theta$, and $g(\theta \mid x) = 0$ otherwise).

(c) We have that $\mathbb{E}[(\theta - a)^2 \mid X = x] = \mathbb{E}[\theta^2 - 2a\theta + a^2 \mid X = x] = \mathbb{E}[\theta^2 \mid X = x] - 2a\,\mathbb{E}[\theta \mid X = x] + a^2$. This resolves to

$$\int_{-\infty}^{\infty} g(\theta^2 \mid x) f(x) d\theta - 2a \int_{-\infty}^{\infty} g(\theta \mid x) f(x) d\theta + a^2$$

which is

$$\int_{-\infty}^{\infty} e^{x - \theta^2} e^{-x} d\theta - 2a \int_{-\infty}^{\infty} e^{x - \theta} e^{-x} d\theta + a^2 = \int_{-\infty}^{\infty} e^{-\theta^2} d\theta - 2a \int_{-\infty}^{\infty} e^{-\theta} d\theta + a^2$$