Homework Assignment 01

Numerical Statistics Fall, 2021

Jake Underland

2021-10-15

1.

Calculate the arithmetic mean $\bar{x_5}$ of a data $4, 6, 8, 9, 13$ and then express it in the form of an irreducible fraction or an integer.

Solution.

$$\bar{x_5} = \frac{4+6+8+9+13}{5} = \frac{40}{5} = 8$$

2.

Suppose we draw a sample of size 16 from a finite population of size 65, where sampling is without replacement. Calculate the finite population correction factor FPC and then express it in the form of a decimal to the third decimal places.

Solution.

$$FPC_{N,n} = \sqrt{\frac{N-n}{N-1}},\ N = 65,\ n = 16$$

$$FPC_{65,16} = \sqrt{\frac{65-16}{65-1}} = \sqrt{\frac{49}{64}} = \frac{7}{8} = 0.875$$

3.

Suppose $X_1, X_2, X_3, X_4, X_5 \overset{i.i.d.}{\sim} Ber(0.4)$. Calculate the probability of the outcome $\{X_1 = 0, X_2 = 0, X_3 = 1, X_4 = 0, X_5 = 1\}$ and then express it in the form of a decimal to the fifth decimal places.

Solution.

$$\begin{aligned} P(X_1 = 0, X_2 = 0, X_3 = 1, X_4 = 0, X_5 = 1) &= P(X_1 = 0)P(X_2 = 0)P(X_3 = 1)P(X_4 = 0)P(X_5 = 1) \\ &= p^2(1-p)^3 \\ &= (0.4)^2(0.6)^3 \\ &= 0.03456 \end{aligned}$$

4.

Consider a data $(x_1, x_2, x_3, x_4) = (-5, a, 1, b)$ $(-5 < a < 1 < b)$ whose arithmetic mean and variance are 2 and $\frac{77}{2}$, respectively. Find the real numbers a and b.

Solution.
From the formula for the arithmetic mean, we have

$$\begin{aligned}\frac{-5+a+1+b}{4} &= 2 \\ -5+a+1+b &= 8 \\ b &= 12-a\end{aligned}$$

Plug the value of b into the equation for variance:

$$\begin{aligned}\frac{(-5-2)^2+(a-2)^2+(1-2)^2+(12-a-2)^2}{4} &= \frac{77}{2} \\ 49+a^2-4a+4+1+100-20a+a^2 &= 154 \\ 2a^2-24a &= 0 \\ a(12-a) &= 0 \\ \implies \begin{cases} a=0 \\ b=12 \end{cases}, \begin{cases} a=12 \\ b=0 \end{cases}\end{aligned}$$

Since $-5 < a < 1 < b$, we can rule out the case where $a = 12, b = 0$, leaving us with

$$a = 0,\ b = 12 \ldots \square$$

5.

Suppose we have a data $x_1, x_2, \ldots, x_n$ of size n. Prove that the inequality

$$\sum_{i=1}^{n}(x_i - c)^2 \geq \sum_{i=1}^{n}(x_i - \bar{x})^2$$

holds for any real number c, where $\bar{x}$ is the arithmetic mean.

Solution.
In order to prove $\sum_{i=1}^{n}(x_i - c)^2 \geq \sum_{i=1}^{n}(x_i - \bar{x})^2$, we first show $\sum_{i=1}^{n}(x_i - c)^2 - \sum_{i=1}^{n}(x_i - \bar{x})^2 \geq 0$.

$$\begin{aligned}\sum_{i=1}^{n}(x_i - c)^2 - \sum_{i=1}^{n}(x_i - \bar{x})^2 &= \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i^2 - 2c\underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} + 2\bar{x}\underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} + nc^2 - n\bar{x}^2 \\ &= -2cn\bar{x} + 2n\bar{x}^2 + nc^2 - n\bar{x}^2 \\ &= n(\bar{x}^2 - 2c\bar{x} + c^2) \\ &= n(\bar{x} - c)^2 \geq 0\end{aligned}$$

Where the last inequality holds because $n \geq 0$ by definition and $(\bar{x} - c)^2$ is squared and thus at least 0. Then, we have proven that

$$\sum_{i=1}^{n}(x_i - c)^2 - \sum_{i=1}^{n}(x_i - \bar{x})^2 \geq 0$$

From which follows:

$$\sum_{i=1}^{n}(x_i - c)^2 \geq \sum_{i=1}^{n}(x_i - \bar{x})^2 \ldots \square$$