Assignment 2

Akshad Shyam

17 March 2022

Q.108 UGC 2017

Let $X_1, X_2, ..., X_n$ be a random sample from $f_\theta(x)$, a probability density function or a probability mass function. Defined

$$s_n^2 = \frac{\sum_{n=1}^{n}(X_i - \overline{X})^2}{n-1}, \quad \overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

Then s_n^2 is unbiased for θ if

1) $f_\theta(x) = \frac{e^{-\theta}\theta^x}{x!}; x > 0$

2) $f_\theta(x) = \frac{e^{\frac{-x^2}{2\theta}}}{\sqrt{2\pi\theta}} - \infty < x < \infty, \theta > 0$

3) $f_\theta(x) = \frac{e^{\frac{-x}{\theta}}}{\theta} x > 0, \theta > 0$

4) $f_\theta(x) = \theta e^{-\theta x} x > 0, \theta > 0$

Answer: Option 1 and Option 2 are correct.

1 Solution

Regardless of the distribution, the expected value of the sample variance is equal to the population variance. Hence the sample variance is said to be an unbiased estimator of population variance.

$$E[s^2] = \sigma^2 \tag{1}$$

Proof:

Let $X_1, X_2..., X_n$ be n independent observations from a population with mean μ and variance σ^2

$$E[\sum X_i] = \sum E[X_{i]} \tag{2}$$

1

$$E[cX] = cE[X] \tag{3}$$

We also know that,

$$\sigma^2 = E[X^2] - E[X]^2; \ \ E[X^2] = \sigma^2 + \mu^2 \tag{4}$$

If we take n samples from this population, then the sample mean would be μ and sample variance would be $\frac{\sigma^2}{n}$. Then, according to 4,

$$E[\overline{X}^2] = \frac{\sigma^2}{n} + \mu^2 \tag{5}$$

Now,

$$E[s_n^2] = E[\frac{\sum_{n=1}^{n}(X_i - \overline{X})^2}{n-1}]$$

$$E[\sum_{n=1}^{n}(X_i - \overline{X})^2] = E[\sum(X_i^2 - 2X_i\overline{X} + \overline{X}^2)]$$

$$= E[\sum(X_i^2) - 2\overline{X}\sum(X_i) + \sum(\overline{X}^2)]$$

$$= E[\sum(X_i^2) - 2\overline{X}n\overline{X} + n\overline{X}^2]$$

$$= E[\sum(X_i^2) - n\overline{X}^2]$$

$$= [\sum(E[X_i^2]) - E[n\overline{X}^2]]$$

Now, substituting 4 and 5,

$$= [\sum(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)]$$

$$= [n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2] = (n-1)\sigma^2$$

$$E[s_n^2] = \frac{(n-1)\sigma^2}{(n-1)} = \sigma^2 \tag{6}$$

Hence the sample variance is said to be an unbiased estimator of the population variance.

2 Checking Options

To check if s_n^2 is an unbiased estimator of θ, θ should be the population variance of the distribution

2.1 Option 1

$f_\theta(x) = \frac{e^{-\theta}\theta^x}{x!}$

This is a probability distribution following poisson distribution

Mean = E[X] = θ

Variance = $E[X^2] - [E[X]]^2$

$$E[X^2] = \sum x^2 P(X = x)$$
$$= \sum x^2 \frac{e^{-\theta}\theta^x}{x!}$$
$$= \sum x \frac{e^{-\theta}\theta^x}{(x-1)!}$$
$$= \theta e^{-\theta}[\sum x \frac{\theta^{x-1}}{(x-1)!}]$$
$$= \theta e^{-\theta}[\sum (x-1) \frac{\theta^{x-1}}{(x-1)!} + \sum (1) \frac{\theta^{x-1}}{(x-1)!}]$$
$$= \theta e^{-\theta}[\theta[\sum \frac{\theta^{x-2}}{(x-2)!}] + \sum (1) \frac{\theta^{x-1}}{(x-1)!}]$$

According to taylor series expansion,

$$= \theta e^{-\theta}[\theta[e^\theta] + e^\theta] = \theta^2 + \theta$$
$$Variance = \theta^2 + \theta - (\theta)^2 = \theta$$

Hence Option 1 is correct

2.2 Option 2

$f_\theta(x) = \frac{e^{\frac{-x^2}{2\theta}}}{\sqrt{2\pi\theta}}$

Option 2 is a probability distribution function following the normal distribu-

tion.

$$\text{Mean} = 0; \text{Variance} = E[X^2] - [E[X]]^2 = E[X^2]$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) \mathrm{d}x$$

$$= \int_{-\infty}^{\infty} x^2 \frac{e^{\frac{-x^2}{2\theta}}}{\sqrt{2\pi\theta}} \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi\theta}} \int_{-\infty}^{\infty} x^2 e^{\frac{-x^2}{2\sqrt{\theta^2}}} \mathrm{d}x$$

Taking $z = \frac{x^2}{2\theta}; x\mathrm{d}x = \theta\mathrm{d}z$

$$= \frac{1}{\sqrt{2\pi\theta}} \int_{-\infty}^{\infty} \sqrt{2\theta z} e^{-z} \theta \mathrm{d}z$$

$$= \frac{2\theta}{\sqrt{\pi}} \int_{0}^{\infty} \sqrt{z} e^{-z} \mathrm{d}z$$

According to gamma function, $\Gamma(x) = \int z^{x-1} e^{-z} \mathrm{d}z$
Since x = 3/2 here $\Gamma(n/2) = \frac{(n-2)!!\sqrt{\pi}}{2^{\frac{n-1}{2}}} = \sqrt{\pi}/2$
Substituting this value,

$$\frac{2\theta}{\sqrt{\pi}} \times \frac{\sqrt{\pi}}{2}$$

$$= \theta$$

Hence Option 2 is correct

2.3 Option 3

$f_\theta(x) = \frac{e^{\frac{-x}{\theta}}}{\theta}$
Option 3 is a probability distribution function following the exponential distribution.

Mean $= E[X]$

$$E[X] = \int_0^\infty x \frac{e^{\frac{-x}{\theta}}}{\theta} \mathrm{d}x$$

Taking $y = \frac{x}{\theta}; \mathrm{d}x = \theta \mathrm{d}y$

$$= \theta \int_0^\infty y e^{-y} \mathrm{d}y$$
$$= \theta[-ye^{-y} - e^{-y}]_0^\infty = \theta$$

Variance $= E[X^2] - [E[X]]^2$

$$E[X^2] = \int_0^\infty x^2 \frac{e^{\frac{-x}{\theta}}}{\theta} \mathrm{d}x$$

Taking $y = \frac{x}{\theta}; \mathrm{d}x = \theta \mathrm{d}y$

$$= \theta^2 \int_0^\infty y^2 e^{-y} \mathrm{d}y$$
$$= \theta^2[-y^2 e^{-y} - 2ye^{-y} - 2e^{-y}]_0^\infty = 2\theta^2$$
$$Variance = 2\theta^2 - \theta^2 = \theta^2$$

Option 3 is incorrect.

2.4 Option 4

$f_\theta(x) = \theta e^{-\theta x}$
Option 4 is a probability distribution function following exponential distribution.
Mean $= E[X]$

$$E[X] = \int_0^\infty x\theta e^{-\theta x} \mathrm{d}x$$

Taking $y = x\theta; \mathrm{d}x = \frac{\mathrm{d}y}{\theta}$

$$= \frac{1}{\theta} \int_0^\infty y e^{-y} \mathrm{d}y$$

$$= \frac{1}{\theta} [-y e^{-y} - e^{-y}]_0^\infty = \frac{1}{\theta}$$

Variance $= E[X^2] - [E[X]]^2$

$$E[X^2] = \int_0^\infty x^2 \theta e^{-\theta x} \mathrm{d}x$$

Taking $y = x\theta; \mathrm{d}x = \frac{\mathrm{d}y}{\theta}$

$$= \frac{1}{\theta^2} \int_0^\infty y^2 e^{-y} \mathrm{d}y$$

$$= \frac{1}{\theta^2} [-y^2 e^{-y} - 2y e^{-y} - 2e^{-y}]_0^\infty = \frac{2}{\theta^2}$$

$$Variance = \frac{2}{\theta^2} - \frac{1}{\theta^2} = \frac{1}{\theta^2}$$

Option 4 is incorrect