Frequently Asked Questions

1. When do you say that an estimator is unbiased to the population parameter?
   Answer:
   The expected value (mean) of the estimator's sampling distribution is equal to the underlying population parameter; that is, there is no upward or downward bias. An estimator t of a parameter θ is said to be unbiased if E (t) = θ. Hence any estimate is said to be unbiased if the average value of the estimator of the population parameter is equal to the actual value of the parameter.
   Unbiased estimators determine the tendency, on the average, for the statistics to assume values close to the parameter of interest.

2. State the conditions that have to be satisfied by an estimator to be called a consistent estimator
   Answer:

   Suppose T_n is an estimator of an unknown parameter Ө. Tn is said to be consistent for Ө if $T_n \xrightarrow{P} \theta$ as $n \to \infty$

   That is $P[|T_n - \theta| < \varepsilon] \to 1$ as $n \to \infty$

   OR $P[|T_n - \theta| > \varepsilon] \to 0$ as $n \to \infty$

   A necessary and sufficient condition for consistency is E (Tn) = Ө and

   V (Tn) $\to 0$ as $n \to \infty$

   Consistency is a limiting property. Moreover several consistent estimators may exist for the same parameter

3. Suppose X and Y are independent random variables with the same unknown mean μ. Both X and Y have variance as 36. Let T=aX+bY be an estimator of μ
   i) Show that T is an unbiased estimator of μ if a+b=1
   ii) If a =1/ 3 and b = 2/3 what is the variance of T?
   iii) If a=1/2 and b=1/2 what is the variance of T?
   iv) What choices of a and b minimizes the variances of T subject to the requirement that T is the unbiased estimator of μ?
   Answer:

   X and Y are independent random variables with the same unknown mean μ.

   Then E(X) =E(Y) =μ

Both X and Y have variance as 36 then V(X) =V(Y) =36
Given T= aX+bY
i.T is an unbiased estimator of μ if and only if
E (T) =μ which implies E (aX+bY) =μ
=a E(X) +bE(Y) =aμ+bμ=μ= (a+b) μ=μ
L.H.S will be equal to R.H.S if a+b=1
Hence T is an unbiased estimator of μ if a+b=1

ii) V (T) = V (aX+bY) = $a^2V(X) + b^2V(Y) = a^2 .36+b^2.36=36(a^2+ b^2)$
If a = 1/3 and b = 2/3 then V (T) =36(1/9+4/9) =20

III. If a=1/2 and b=1/2 then V(T)=36(1/4+1/4)=18

IV. T is unbiased for μ if E (T) =μ. From (i) T is an unbiased estimator of μ if a+b=1
When a = 1/3 and b = 2/3, a+b=1 But V (T) =36(1/9+4/9) =20

When a=1/2 and b=1/2 then a+b=1 but V (T) =36(1/4+1/4) =18

Hence V (T) is minimum for a=1/2 and b=1/2

Therefore when a=b=1/2 variance of T is minimum subject to the requirement that T is the unbiased estimator of μ

4. The observations $x_1,x_2,..,x_n$ represents the random sample from a Uniform distribution over the interval (0,θ) where θ is the unknown parameter. Suppose statistic t= $\bar{x}$ is the mean value of the sample observations is unbiased for the population mean . Find the value of k so that kt is an unbiased estimator of θ.
Answer:
Given observations x1,x2,..,xn represents the random sample from a Uniform distribution over the interval (0,θ) where θ is the unknown parameter
Xi~U(0, θ)
Then we know that the population mean = θ/2 then E($\bar{x}$)=θ/2
kt is an unbiased estimator of θ if and only if
E(kt)= θ which implies k E(t) = θ implies kE($\bar{x}$)=θ
Then k (θ/2)= θ which implies k=2

5. If two samples of sizes 10 and 14 are drawn from the same population and have the variances 12.25 and 9 respectively. Calculate the unbiased estimate of the population variance.
Answer:

Given n1=10 ,n2=14, s_1^2=12.25 and s_2^2=9

For the samples of sizes n1 and n2 drawn from the population

$$\frac{(n_1-1)s_1^2}{\sigma^2} \sim \chi^2(n_1-1) \quad \text{and} \quad \frac{(n_2-1)s_2^2}{\sigma^2} \sim \chi^2(n_2-1)$$

$$\frac{(n_1-1)s_1^2}{\sigma^2} + \frac{(n_2-1)s_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2)$$

Hence
$$E[\frac{(n_1-1)s_1^2}{\sigma^2} + \frac{(n_2-1)s_2^2}{\sigma^2}] = (n_1+n_2-2)$$
and

$$E[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{\sigma^2}] = (n_1+n_2-2) \Rightarrow E[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)}] = \sigma^2$$

An unbiased estimator of the population variance is given by

$$\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)} = \frac{9*12.25+13*9}{10+14-2} = 10.33$$

6. Write a note on efficiency of an estimate.
   Answer:
   While there are many unbiased estimates for the same parameter of the population, the most efficient has a sampling distribution with the smallest variance. Our aim is to get such an estimate which has the least standard error. An estimate with least standard error is said to be an efficient estimate of the population parameter.

7. The sample 2,4,6,1,9,4,8,2,4,6,5,10 drawn from a population with density function f(x) = 1/θ, 0<x<θ. Obtain an unbiased estimate of population parameter θ
   Answer:
   Given f(x) = 1/θ, 0<x<θ then X~U(0,θ) for which
   E(X)= θ/2 and V(X)= θ^2/12
   Suppose t is an unbiased estimator of the population mean then E(t)= population mean because for the given density function population mean is θ/2
   If x is unbiased for θ then $\bar{x}$ is also unbiased for θ

   E(X)= θ/2 implies E(2x) =θ or E($2\bar{x}$)= θ

Hence $2\bar{x}$ is an unbiased estimator of the given population's parameter θ which is given by 2*5.08=10.16

8. If X1,X2,X3 are the three independent observations drawn from a Normal population with mean μ and variance σ^2. Find the relative efficiency of T1 with respect to T2. Which is more efficient where
   T1=3 X1+2 X2-4 X3
   T2=3X1-5X2+3X3
   Answer:
   If T1 and T2 are the two unbiased estimators then one of them must be more efficient and it has the less variance than the other. Here
   E(T1) = E(3 X1+2 X2-4 X3)=3E(X1)+2E(X2)-4E(X3)=3μ+ 2μ- 4μ= μ because E(Xi)= μ for all i=1,2,3

   Therefore since E(T1)= μ, T1 is unbiased for μ
   E(T2)= E(3X1-5X2+3X3)= 3E(X1)-5E(X2)+3E(X3)=3 μ-5 μ+3 μ= μ

   Therefore since E(T2)= μ, T2 is unbiased for μ
   Hence T1 and T2 are unbiased estimators of μ
   V(T1)= V(3 X1+2 X2-4 X3)=9 V(X1)+4V(X2)+16 V(X3) =29 σ^2
   V(T2)= = V(3X1-5X2+3X3)=9V(X1)+25V(X2)+9 V(X3)= 43 σ^2

   V(T1)<V(T2)
   Therefore T1 is more efficient
   Relative efficiency of T1 with respect to t2 is
   V(T2)/V(T1)= 43 σ^2 / 29 σ^2 =43/29>1

9. Briefly explain asymptotical unbiased estimators and relative efficiency of an estimator
   Answer:

   Suppose Tn is an estimator of an unknown parameter Θ then Tn is said to be unbiased for Θ if E(Tn)=Θ

   Tn is said to be asymptotically unbiased for Θ if E (Tn) =Θ as $n \to \infty$

   If a consistent estimator exists whose sampling variance is less than that of any other consistent estimator it is said to be most efficient and it provides a standard for the measurement of efficiency of a statistic.

   Further a relative efficiency of T1 with respect to T2 is defined as

   E (T1, T2) = V (T1)/V (T2)

a) If E(T1,T2)=1 then the both T1 and T2 are equally efficient
b) If E(T1,T2)>1 then T1 is more efficient thanT2
c) If E(T1,T2)<1 then T2 is more efficient thanT1

10. Derive an unbiased estimator for 1/θ from the density function.

Answer:

Given $f(x, \theta) = \theta e^{-\theta x}$,x>0, θ>0

Then $E(x) = \int_0^\infty x f(x)dx = \int_0^\infty x\theta e^{-\theta x} dx = 1/\theta$

$E(\bar{x}) = \frac{1}{n} E(\sum xi) = E(xi) = \frac{1}{\theta}$

Hence sample mean $\bar{x}$ is an unbiased estimator for 1/θ

$E(\bar{x}) = \frac{1}{\theta} \rightarrow \frac{1}{\theta}$ as $n \rightarrow \infty$

$V(\bar{x}) = V(\frac{1}{n}\sum xi) = \frac{1}{n^2}\sum V(xi) = \frac{V(xi)}{n}$

But $V(xi) = E(xi^2) - (E(xi))^2$

$E(xi^2) = \int_0^\infty x^2 f(x)dx = \int_0^\infty x^2 \theta e^{-\theta x} dx = 2/\theta^2$

Therefore $V(xi) = (2/\theta^2) - (1/\theta)^2 = 1/\theta^2$

$$\lim_{n\to\infty} V(\bar{x}) = \lim_{n\to\infty} \frac{V(xi)}{n} = \lim_{n\to\infty} \frac{1}{\theta^2 n} \rightarrow 0$$

Therefore $\bar{x}$ is consistent for 1/θ

The mean $\bar{x}$ for the data 0.16, 0.84.0.63,1.26,0.28,1.48,0.05 and 1.24 is ∑x/n=5.94/8=0.7425

11. While sampling from a Cauchy distribution with p.d.f $f(x) = \frac{1}{\pi}\frac{1}{1+(x-\mu)^2}, -\infty < x < \infty$

Show that a sample mean is not a consistent estimator of the population mean. But the sample median is a consistent estimator of the population mean (median). Hence obtain an consistent estimate of the population mean for the following sample of size 8

-1.7, 58.7,-3.33,-77.3,-0.65, 1.48, -9.6 and 1.77

Answer:

Let us make use of 2 results

1) $\bar{x}$ is distributed as x
2) For any population , for large sample size sample median is distributed as Normal about its population median with variance $\frac{1}{4(n-1)}\frac{1}{[f(populationmedian)]^2}$

Let x1,x2,..,xn be a random sample of size n drawn from a Cauchy population with mean μ . For a Cauchy population , population mean = population median = μ. We have to show that $\bar{x}$ does not converge in probability to μ as n tends to ∞.

$$P\left[\left|\bar{x}-\mu\right|<\varepsilon\right]= \quad P\left[-\varepsilon<\bar{x}-\mu<\varepsilon\right] \quad \Rightarrow P\left[\mu-\varepsilon<\bar{x}<\mu+\varepsilon\right]$$

$$\int_{\mu-\varepsilon}^{\mu+\varepsilon}\frac{1}{\pi}\frac{1}{1+(x-\mu)^2}dx=\int_{-\varepsilon}^{+\varepsilon}\frac{1}{\pi}\frac{1}{1+(z)^2}dz=\frac{2}{\pi}\tan^{-1}\varepsilon\neq 1 \text{ unless } \varepsilon=\infty$$

Hence $P\left[\left|\bar{x}-\mu\right|<\varepsilon\right]$ does not tend to 1 in probability as n tends to infinity

Therefore a sample mean is not an consistent estimator of population mean μ. Let m denote the sample median then

M ~N(μ, $\frac{1}{4(n-1)}\frac{1}{[f(\mu)]^2}$)

E(m) = μ $\rightarrow \mu$ as $n\rightarrow\infty$

V(m) = $\frac{1}{4(n-1)}\frac{1}{[\pi]^2}=\frac{\pi^2}{4(n-1)}\rightarrow 0$ as $n\rightarrow\infty$

Therefore a sample median m tends to population median or mean μ in probability as n tends to infinity.

Hence for the given set of observations median can be obtained as follows. Arranging the observations in an ascending order

-77.3, -9.6, -3.33,-1.7, -0.65, 1.48, 1.77, 58.7

Median = Value of ((n+1)/2) th observation = Value of (4.5) th observations=-1.175

Hence an consistent estimate of the population mean for the given sample of size 8 is -1.175

12. X_1, X_2, X_3 is a random sample of size 3 from a population with mean μ and variance σ^2. T_1, T_2 and T_3 are the estimators used to estimate the mean value μ where

    $T_1 = X_1+X_2-X_3$

    $T_2=2X_1+3X_2-4X_3$ and

    $T_3= (\lambda X_1+X_2+X_3)/3$.

    a) Find whether T1 and T2 are unbiased estimators
    b) Find the value of λ such that T3 is an unbiased estimator of μ
    c) With this value of λ is T3 a consistent estimator?
    d) Which is an efficient estimator?

Answer:

Since X1,X2,X3 is a random sample of size 3 from a population with mean μ and variance σ^2 E(xi)=μ and V(Xi)= σ^2 and COV(Xi,Xj)=0 for all i not equal to j

Given T1= X1+X2-X3 and T2=2X1+3X2-4X3

E (T1) = E(X1+X2-X3) =E(X1) +E(X2)-E(X3) =μ+ μ- μ= μ because E (Xi) = μ for all i=1, 2, 3

Therefore since E (T1) = μ, T1 is unbiased for μ

E (T2) = E (2X1+3X2-4X3) =) = 2E(X1) +3E(X2)-4E(X3) =2 μ+3 μ-4 μ= μ

Therefore since E (T2) = μ, T2 is unbiased for μ

Hence T1 and T2 are unbiased estimators of μ

b) If T3 is unbiased for μ then E (T3) = μ

E (T3) = μ implies E ((λX1+X2+X3)/3.)= μ

[Λ E(X1) +E(X2) +E(X3)]/3= μ

[Λ μ+ μ+μ]/3= μ implies λ μ+ 2μ = 3μ implies λ μ=μ implies λ =1

c) With λ =1, T3= (X1+X2+X3)/3= $\overline{X}$

Since sample mean is a consistent estimator of the population mean μ by weak law of large numbers, T3 is a consistent estimator of μ

d) An efficient estimator is the one which has the least variance

V (T1) = V(X1+X2-X3) = V(X1) +V(X2) + V(X3) = 3 σ^2

V (T2) = V(2X1+3X2-4X3)=4V(X1)+9V(X2)+16 V(X3)= 29 σ^2

V(T3) = V((X1+X2+X3)/3)= V(X1+X2+X3)/9

=3 σ^2/9= σ^2/3

Hence we can observe that V(T3) < V(T1) < V(T2)

Therefore T3 is an efficient estimator of μ

13. Show that when sampling from a Normal population a sample variance $s^2 = \frac{1}{n}\sum(xi-\bar{x})^2$ is biased for the population variance but is asymptotically unbiased. Also find the unbiased estimator of the variance and hence Obtain an unbiased estimate of the population variance σ^2 from the following sample of size 8

74.1,77.2, 74.4, 74, 73.8, 79.3, 75.8, 82.8 from a Normal population.

Answer:

Given $s^2 = \frac{1}{n}\sum(xi-\bar{x})^2$

We know that $\frac{ns^2}{\sigma^2} \sim \chi^2(n-1) \Rightarrow E(\frac{ns^2}{\sigma^2}) = n-1 \Rightarrow E(s^2) = \frac{n-1}{n}\sigma^2 \neq \sigma^2$

Therefore s^2 is biased for σ^2

But $\lim_{n\to\infty} E(s^2) = \lim_{n\to\infty} (\frac{n-1}{n})\sigma^2 \to \sigma^2$

Therefore s^2 is asymptotically unbiased for σ^2

$E(s^2) = \frac{n-1}{n}\sigma^2 \Rightarrow E(\frac{n}{n-1}s^2) = \sigma^2$

Therefore $\frac{n}{n-1}s^2$ is an unbiased estimator of σ^2

A sample variance

$$s^2 = \frac{1}{n}\sum(xi-\bar{x})^2 = \frac{1}{n}[\sum xi^2 - n\bar{x}^2] = \frac{1}{8}[46798.42 - 8*(76.425)^2] = 9.02$$

Therefore $\frac{n}{n-1}s^2 = \frac{8*9.02}{7} = 10.31$ is an unbiased estimate of the population variance

14. If x1, x2... xn are random observations on a Bernoulli variable X taking values 1 with probability p and value 0 with probability (1-p). Then show that $\bar{x}(1-\bar{x})$ is the consistent estimator of p(1-p)

Answer:

Since x1, x2...xn are i.i.d Bernoulli Variables with parameter p, ∑xi is Binomial with parameters n and p

That is ∑xi~B(n,p)

Then E(∑xi)=np and V(∑xi)=npq where q=1-p

E(∑xi)=np implies E($\bar{x}) = p \to p$ as $n \to \infty$

$$V(\bar{x}) = V(\frac{\sum xi}{n}) = \frac{1}{n^2} V(\sum xi) = \frac{npq}{n^2} \to 0 \text{ as } n \to \infty$$

Hence $\bar{x}$ is a consistent estimator of p . But $\bar{x}(1-\bar{x})$ being polynomial in $\bar{x}$ is a continuous function of $\bar{x}$. Since $\bar{x}$ is a consistent estimator of p by invariance property of consistent estimators $\bar{x}(1-\bar{x})$ is a consistent estimator of p(1-p)

15. For the sample of 10 observations from a Normal population with mean μ and variance σ^2 Obtain an unbiased estimates of population mean and variance

5, 7, 15, 20, 22, 17, 19, 25, 21 and 14

Answer:

$$\bar{x} = \frac{\sum xi}{n} = 16.5$$

$$s^2 = \frac{1}{n} \sum (xi - \bar{x})^2 = \frac{\sum xi^2}{n} - (\bar{x})^2 = \frac{395}{10} - (16.5)^2 = 37.25$$

An unbiased estimator of the population mean μ is the sample mean hence an unbiased estimate of population mean is 16.5

But an unbiased estimator of the population variance is given by $\frac{n}{n-1}s^2$. Hence an unbiased estimate of the population variance is 10*37.25/9=41.39