Statistical Analysis

Master in Statistics and Information Management

Professor *Ana Cristina Costa* (ccosta@novaims.unl.pt)

HOMEWORK

LU4: POINT ESTIMATION

1. Consider a population with distribution N(μ, σ) and a random sample $X_1, \dots, X_n$ from this population.

a) Show that $\bar{X}$ is a consistent estimator of μ.

b) Consider now two alternative estimators of μ, both unbiased, and their variances are the following:

$$\hat{\mu}_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \hat{\mu}_2 = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i$$

$$V[\hat{\mu}_1] = \frac{1}{n}\sigma^2 \qquad V[\hat{\mu}_2] = \frac{1}{n-1}\sigma^2$$

Show that $\hat{\mu}_1$ is more efficient than $\hat{\mu}_2$.

2. Consider a random sample $X_1, X_2, \dots, X_n$ from some population with mean μ and variance σ^2, and the following estimator of the mean of this population:

$$\hat{\mu} = 0{,}4X_1 + aX_3 + 0{,}3X_5 + 0{,}2X_n$$

a) Find the value of *a* that makes $\hat{\mu}$ an unbiased estimator.

b) Study $\hat{\mu}$ as to its consistency in mean square error.

c) For *a*=0,1, compare $\hat{\mu}$ with the following alternative estimator as to the relative efficiency:

$$\tilde{\mu} = \frac{1}{8}(2X_1 + 5X_2 - 3X_3 + 4X_n)$$

3. Consider a population with Poisson distribution. In order to estimate λ, the following two estimators were considered based on a random sample $X_1, \dots, X_n$ from this population:

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{and} \quad \tilde{\lambda} = \frac{X_1 + X_n}{2}$$

a) Study the bias and the consistency of the estimator $\tilde{\lambda}$.

b) Compare the efficiency of the two proposed estimators.

4. Consider the following estimator of the variance of a normal population with mean μ and variance σ^2:

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}$$

Taking into consideration that if $Z_1, Z_2, \dots, Z_n$ are iid N(0, 1) then $\sum_{i=1}^{n} {Z_i}^2 \sim \chi^2_{(n)}$:

a) Show that $\tilde{\sigma}^2$ is an unbiased estimator.

b) Compare the efficiency of $\tilde{\sigma}^2$ with the efficiency of $\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

c) Show that $\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is consistent.

5. Let X be the number of observed successes in a random sample of size n.

a) Show that the sample proportion of successes $\hat{P} = X/n$ is an unbiased estimator of the probability of success p.

b) Show that the sample proportion of failures $\hat{Q} = 1 - X/n$ is an unbiased estimator of the probability of failure q.

6. Consider two independent random samples of sizes n_1 and n_2 from the same population, and the following estimators of the population mean μ:

$$\hat{\mu}_1 = \frac{\bar{X}_1 + \bar{X}_2}{2} \quad \text{and} \quad \hat{\mu}_2 = \frac{1 + n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2},$$

where $\bar{X}_1$ and $\bar{X}_2$ are the sample means of the first and second samples, respectively.

a) Determine the bias of each of the estimators.

b) Determine the variance of each of the estimators.

Assuming that $n_2 = kn_1$ and $k \geq 1$ is an integer:

c) Compare the variance of the two estimators.

d) Determine the mean squared error of each of the estimators.

e) Compare the efficiency of the estimators when $n_1 = 10,\ n_2 = 40\ (k = 4)$ and $\sigma^2 = 1$.

f) Show that both estimators are consistent in mean square error.

SOLUTIONS

Question 2)

a) a = 0.1

b) $\hat{\mu}$ is not consistent in mean square error

c) $\hat{\mu}$ is more efficient than $\tilde{\mu}$

Question 3)

a) $\tilde{\lambda}$ is unbiased but not consistent

b) If $n < 2$, $\tilde{\lambda}$ is more efficient than $\hat{\lambda}$;

if $n = 2$, $\hat{\lambda}$ and $\tilde{\lambda}$ are equally efficient;

if $n > 2$, $\hat{\lambda}$ is more efficient than $\tilde{\lambda}$

Question 4)

b) $\tilde{\sigma}^2$ is more efficient than S^2

Question 6)

a) $bias\ (\hat{\mu}_1) = 0$; $bias\ (\hat{\mu}_2) = \frac{1}{n_1+n_2}$

b) $V(\hat{\mu}_1) = \frac{\sigma^2}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$; $V(\hat{\mu}_2) = \frac{\sigma^2}{n_1+n_2}$

c) $\hat{\mu}_2$ is more efficient

d) $EQM\ (\hat{\mu}_1) = \frac{\sigma^2}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$; $EQM\ (\hat{\mu}_2) = \frac{(n_1+n_2)\sigma^2+1}{(n_1+n_2)^2}$

e) Although biased, $\hat{\mu}_2$ is more efficient than $\hat{\mu}_1$ when k=4