PES University, Bangalore

(Established under Karnataka Act No. 16 of 2013)

UE19CS203 – STATISTICS FOR DATA SCIENCE

Unit - 3 - Probability Distributions

QB SOLVED

Principles of Point Estimation - Mean Squared Error

Exercises for Section 4.9

1. Let $X_1 \ldots\ldots X_n$ be a random sample from $N(\mu, \sigma^2)$ population. For any constant k > 0, define $\widehat{\sigma_k} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{k}$. Consider $\hat{\sigma}_k{}^2$ as an estimator of σ^2.

   a) Compute bias of $\hat{\sigma}_k{}^2$ in terms of k. [Hint: The sample variance s^2 is unbiased and $\hat{\sigma}_k{}^2 = (n-1)s^2/k$].
   b) Compute the variance of $\hat{\sigma}_k{}^2$ in terms of k. [Hint: $\sigma_{s^2}^2 = 2\sigma^4/(n-1)$, and $\hat{\sigma}_k{}^2 = (n-1)s^2/k$].
   c) Compute mean squared error of $\hat{\sigma}_k{}^2$ in terms of k.
   d) For what value of k is the mean squared error of $\hat{\sigma}_k{}^2$ minimized?

   [Text Book Exercise – Section 4.9 – Q. No. 4 – Pg. No. 284,285]
   Solution:

   a) Compute bias of $\hat{\sigma}_k{}^2$ in terms of k. [Hint: The sample variance s^2 is unbiased and $\hat{\sigma}_k{}^2 = (n-1)s^2/k$].

   Given that $\widehat{\sigma_k} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{k}$

   The sample variance is given by, $s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

   Multiply and divide $\widehat{\sigma_k}$ by n-1,

   $$= \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \cdot \frac{n-1}{k} = \frac{(n-1)s^2}{k}$$

   It is given that sample variance is unbiased, that is $\mu_{s^2} = \sigma^2$, so bias is,

$$\mu_{\widehat{\sigma_k^2}} - \sigma^2 = \mu_{\frac{(n-1)s^2}{k}} - \sigma^2$$

$$= \frac{n-1}{k}\mu_{s^2} - \sigma^2$$

$$= \frac{n-1}{k}\sigma^2 - \sigma^2$$

$$Bias = \frac{n-k-1}{k}\sigma^2$$

b) Compute the variance of $\hat{\sigma}_k{}^2$ in terms of k. [Hint: $\sigma_{s^2}^2 = 2\sigma^4/(n-1)$, and $\hat{\sigma}_k{}^2 = (n-1)s^2/k$].

It is given that, sample variance s^2 is equal to $2\sigma^4/n-1$. The variance is,

$$\sigma^2{}_{\widehat{\sigma_k{}^2}} = \sigma^2{}_{\frac{(n-1)s^2}{k}}$$

$$= \frac{(n-1)^2}{k^2}\ \sigma^2{}_{s^2}$$

$$= \frac{2\sigma^4(n-1)^2}{(n-1)k^2}$$

$$Variance = \frac{2(n-1)}{k^2}\sigma^4$$

c) Compute mean squared error of $\hat{\sigma}_k{}^2$ in terms of k.

The mean squared error of $\hat{\sigma}_k{}^2$ is,

$$MSE_{\hat{\sigma}_k{}^2} = (\mu_{\widehat{\sigma_k^2}} - \sigma^2)^2 + \sigma^2{}_{\widehat{\sigma_k{}^2}}$$

$$= \frac{(n-k-1)^2}{k^2}\sigma^4 + \frac{2(n-1)}{k^2}\sigma^4$$

$$= \frac{n^2+k^2-2nk-2k-1}{k^2}\sigma^4$$

$$MSE_{\hat{\sigma}_k{}^2} = \frac{n^2+k^2-2nk+2k-1}{k^2}\sigma^4$$

d) For what value of k is the mean squared error of $\hat{\sigma}_k{}^2$ minimized?

To find the value that minimize the $MSE_{\hat{\sigma}_k{}^2}$. We will take derivative of k and set it equal to 0.

$$= \frac{d}{dk}\left(\frac{n^2 + k^2 - 2nk + 2k - 1}{k^2}\sigma^4\right)$$

$$= \frac{2(nk - k - n^2 + 1)}{k^3}\sigma^4$$

$$\frac{2(nk - k - n^2 + 1)}{k^3}\sigma^4 = 0$$

$$nk - k - n^2 + 1 = 0$$

$$k(n-1) = n^2 - 1$$

$$k = \frac{n^2 - 1}{n - 1}$$

$$k = n + 1.$$