

Statistics

Homework 3 – Estimation

Abraham Murciano

April 17, 2020

1. (a) Given a sample x_1, \dots, x_n , we are to prove that $\forall k \in \{1, \dots, n\}$,

$$\hat{\mu}_k = \frac{\sum_{i=1}^k x_i}{k}$$

is an unbiased estimator of the average μ . To demonstrate this, we must show that $E[\hat{\mu}_k] = \mu$.

$$E[\hat{\mu}_k] = \frac{\sum_{i=1}^k E[x_i]}{k} = \frac{\sum_{i=1}^k \mu}{k} = \frac{k\mu}{k} = \mu$$

- (b) The larger the value of k , the better the estimator $\hat{\mu}_k$ is. This is because when an estimator's variance from its parameter is small, it is a better estimator. The variance of $\hat{\mu}_k$ is

$$\text{Var}(\hat{\mu}_k) = \text{Var}\left(\frac{\sum_{i=1}^k x_i}{k}\right) = \frac{1}{k^2} \text{Var}\left(\sum_{i=1}^k x_i\right)$$

It is clear from this that the larger k is, the smaller $\frac{1}{k^2}$ is, so the smaller $\text{Var}(\hat{\mu}_k)$ is.

2. With respect to a random variable X having mean μ and variance σ^2 , we have 2 independent samples of sizes n_1 and n_2 respectively, whose means are \bar{x}_1 and \bar{x}_2 . We are given the following estimator.

$$\hat{\mu} = a\bar{x}_1 + (1-a)\bar{x}_2$$

- (a) We are to show that $\hat{\mu}$ is an unbiased estimator for μ for all values of a .

To achieve this we must show that $E[\hat{\mu}] = \mu$.

$$\begin{aligned} E[\hat{\mu}] &= E[a\bar{x}_1 + (1-a)\bar{x}_2] \\ &= E[a\bar{x}_1] + E[(1-a)\bar{x}_2] \\ &= aE[\bar{x}_1] + (1-a)E[\bar{x}_2] \\ &= a\mu + (1-a)\mu \\ &= a\mu + \mu - a\mu \\ &= \mu \end{aligned}$$

- (b) The value of a for which we have the best possible estimator (within this class of estimators) occurs when the variance of the estimator is lowest.

$$\begin{aligned} \text{Var}(\hat{\mu}) &= \text{Var}(a\bar{x}_1 + (1-a)\bar{x}_2) \\ &= a^2 \text{Var}(\bar{x}_1) + (1-a)^2 \text{Var}(\bar{x}_2) \\ &= a^2\sigma^2 + (1-2a+a^2)\sigma^2 \\ &= 2\sigma^2 a^2 - 2\sigma^2 a + \sigma^2 \end{aligned}$$

Now we must differentiate this and set the derivative equal to zero in order to find the value of a which results in the smallest variance.

$$\begin{aligned}\frac{d \operatorname{Var}(\hat{\mu})}{da} &= 4\sigma^2 a - 2\sigma^2 = 0 \\ 4\sigma^2 a &= 2\sigma^2 \\ a &= \frac{2\sigma^2}{4\sigma^2} = \frac{1}{2}\end{aligned}$$

To show that it is indeed a minimum and not a maximum, we must now take the second derivative and check that it is positive when $a = \frac{1}{2}$.

$$\frac{d^2 \operatorname{Var}(\hat{\mu})}{da^2} = 4\sigma^2 \geq 0$$

4. We are given a random variable X with a mean μ and a variance σ^2 , as well as the following two estimators for μ .

$$\begin{aligned}\hat{\mu}_1 &= \frac{\sum_{i=1}^7 x_i}{7} \\ \hat{\mu}_2 &= \frac{2x_1 - x_6 + x_4}{2}\end{aligned}$$

- (a) $\hat{\mu}_1$ is unbiased since $E[\hat{\mu}_1] = \mu$.

$$E[\hat{\mu}_1] = E\left[\frac{\sum_{i=1}^7 x_i}{7}\right] = \frac{\sum_{i=1}^7 E[x_i]}{7} = \frac{7E[X]}{7} = E[X] = \mu$$

$\hat{\mu}_2$ is also unbiased since $E[\hat{\mu}_2] = \mu$.

$$E[\hat{\mu}_2] = E\left[\frac{2x_1 - x_6 + x_4}{2}\right] = \frac{2E[x_1] - E[x_6] + E[x_4]}{2} = \frac{2E[X]}{2} = E[X] = \mu$$

- (b) Of these two estimators, the preferred one would be the one with the smaller variance.

$$\begin{aligned}\operatorname{Var}(\hat{\mu}_1) &= \operatorname{Var}\left(\frac{\sum_{i=1}^7 x_i}{7}\right) = \frac{1}{49} \sum_{i=1}^7 \operatorname{Var}(x_i) = \frac{1}{49} 7\sigma^2 = \frac{\sigma^2}{7} \\ \operatorname{Var}(\hat{\mu}_2) &= \operatorname{Var}\left(\frac{2x_1 - x_6 + x_4}{2}\right) = \frac{\operatorname{Var}(2x_1) - \operatorname{Var}(x_6) + \operatorname{Var}(x_4)}{4} = \frac{4\sigma^2 - \sigma^2 + \sigma^2}{4} = \sigma^2\end{aligned}$$

Since $\operatorname{Var}(\hat{\mu}_1) \leq \operatorname{Var}(\hat{\mu}_2)$, $\hat{\mu}_1$ is a better estimator for μ .

9. Given that a sample of size n was taken from a continuous random variable, X , which has the gamma distribution that follows, we are to find the maximum likelihood estimator for λ .

$$f(x) = \begin{cases} \frac{e^{-\lambda x} \lambda^m x^{m-1}}{\Gamma(m)} & \text{if } x > 0 \wedge \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

To do so, the likelihood function must first be found. The likelihood function, $L(x_1, \dots, x_n)$, is the product of the probability function at each observed value of x .

$$L(x_1, \dots, x_n) = \prod f(x_i) = \prod \frac{e^{-\lambda x_i} \lambda^m x_i^{m-1}}{\Gamma(m)} = \frac{\lambda^{mn} e^{-\lambda \sum x_i} (\prod x_i)^{m-1}}{\Gamma(m)^n}$$

Subsequently we must find the point for λ which causes the likelihood function to have the highest value. This can be achieved by differentiating the likelihood function and setting it equal to zero, or doing the same to its logarithm. This can be done since the function $f(x) = \ln(x)$ is monotonically increasing.

$$\begin{aligned} \ln(L) &= \ln \left(\frac{\lambda^{mn} e^{-\lambda \sum x_i} (\prod x_i)^{m-1}}{\Gamma(m)^n} \right) \\ &= \ln \left(\lambda^{mn} e^{-\lambda \sum x_i} (\prod x_i)^{m-1} \right) - \ln(\Gamma(m)^n) \\ &= \ln(\lambda^{mn}) + \ln(e^{-\lambda \sum x_i}) + \ln \left((\prod x_i)^{m-1} \right) - n \ln(\Gamma(m)) \\ &= mn \ln(\lambda) - \lambda \sum x_i + (m-1) \ln \left(\prod x_i \right) - n \ln(\Gamma(m)) \end{aligned}$$

Now we can differentiate and solve for λ when the derivative is set to zero.

$$\frac{d \ln(L)}{d\lambda} = \frac{mn}{\lambda} - \sum x_i = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{mn}{\sum x_i}$$

To ensure that this is a maximum likelihood estimator and not a minimum likelihood estimator, we must differentiate it once more and verify that the result is negative. (What follows is in fact negative since both m and n are positive.)

$$\frac{d^2 \ln(L)}{d\lambda^2} = -\frac{mn}{\lambda^2} \leq 0$$

10. X has the following discrete distribution. A sample of size n is taken from it. We are to find a maximum likelihood estimator $\hat{\theta}$ and show that it is not biased.

$$f(x; \theta) = \begin{cases} \frac{1}{2} - \theta & \text{if } x = 0 \\ \frac{1}{2} + \theta & \text{if } x = 1 \end{cases}$$

The likelihood function is as follows. Let k be the number of occurrences in which $x = 0$.

$$L(x_1, \dots, x_n) = \prod f(x_i; \theta) = \left(\frac{1}{2} - \theta \right)^k \left(\frac{1}{2} + \theta \right)^{n-k}$$

We then find the logarithm of this function, differentiate it with respect to θ , and solve for θ when $\ln(L) = 0$.

$$\begin{aligned} \ln(L) &= \ln \left(\left(\frac{1}{2} - \theta \right)^k \left(\frac{1}{2} + \theta \right)^{n-k} \right) \\ &= \ln \left(\left(\frac{1}{2} - \theta \right)^k \right) + \ln \left(\left(\frac{1}{2} + \theta \right)^{n-k} \right) \\ &= k \ln \left(\frac{1}{2} - \theta \right) + (n-k) \ln \left(\frac{1}{2} + \theta \right) \end{aligned}$$

$$\begin{aligned}
\frac{d \ln(L)}{d\theta} &= 2 \left(\frac{k}{2\theta - 1} + \frac{n - k}{2\theta + 1} \right) = 0 \\
\frac{k}{2\theta - 1} + \frac{n - k}{2\theta + 1} &= 0 \\
\frac{k}{2\theta - 1} &= \frac{k - n}{2\theta + 1} \\
k(2\theta + 1) &= (k - n)(2\theta - 1) \\
2k\theta + k &= 2k\theta - k - 2n\theta + n \\
2n\theta &= n - 2k \\
\theta &= \frac{n - 2k}{2n} = \frac{1}{2} - \frac{k}{n} \\
\Rightarrow \hat{\theta} &= \frac{1}{2} - \frac{k}{n}
\end{aligned}$$

Now that $\hat{\theta}$ has been found, we must show that it is unbiased. To do so, we show that $E[\hat{\theta}] = \theta$.

$$E[\hat{\theta}] = \frac{1}{2} - \frac{E[k]}{n} = \frac{1}{2} - \frac{nP(X=0)}{n} = \frac{1}{2} - \frac{n(\frac{1}{2} - \theta)}{n} = \frac{1}{2} - \left(\frac{1}{2} - \theta\right) = \theta$$