

Point Estimation

Basic Concepts of Point Estimation

A point estimate of a parameter θ , denoted by $\hat{\theta}$, is a single number that can be considered as a possible value for θ . Since it is computed from the sample $X = (X_1, \dots, X_n)$, it is a function of X , that is, $\hat{\theta} = \hat{\theta}(X)$.

Some simple examples are:

- (i) If X_1, \dots, X_n is from $B(1, p)$ (Bernoulli data), then $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$, the sample proportion of success.
- (ii) If X_1, \dots, X_n is a random sample from a continuous population $F(x)$ with mean μ and variance σ^2 , then the commonly used estimators of μ and σ^2 are

$$\hat{\mu} = \bar{X}; \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = S^2.$$

Some other estimators of μ are the sample median, the trimmed mean, etc.

Lecture 10: Point Estimation

MSU-STT-351-Sum-19A

(P. Vellaisamy: MSU-STT-351-Sum-19A)

Probability & Statistics for Engineers

1/31

(P. Vellaisamy: MSU-STT-351-Sum-19A)

Probability & Statistics for Engineers

2/31

Point Estimation

Next, we discuss some properties of the estimators.

(i) The Unbiased Estimators

Definition: An estimator $\hat{\theta} = \hat{\theta}(X)$ for the parameter θ is said to be unbiased if $E(\hat{\theta}(X)) = \theta$ for all θ .

Result: Let X_1, \dots, X_n be a random sample on $X \sim F(x)$ with mean μ and variance σ^2 . Then the sample mean \bar{X} and the sample variance S^2 are unbiased estimators of μ and σ^2 , respectively.

Proof: (i) Note that

$$E(\bar{X}_n) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n}(n\mu) = \mu.$$

(ii) Note

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Point Estimation

Then

$$\begin{aligned} E((n-1)S^2) &= E\left(\sum_{i=1}^n (X_i - \bar{X})^2\right) \\ &= E\left(\sum_{i=1}^n X_i^2 - n(\bar{X})^2\right) \\ &= nE(X_1^2) - nE(\bar{X}^2) \\ &= n(\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right) \\ &= (n-1)\sigma^2, \end{aligned}$$

using $E(X_1^2) = \text{Var}(X_1) + (E(X_1))^2$ and $E(\bar{X}^2) = \text{Var}(\bar{X}) + (E(\bar{X}))^2$.

Thus,

$$E(S^2) = \sigma^2.$$

(P. Vellaisamy: MSU-STT-351-Sum-19A)

Probability & Statistics for Engineers

3/31

(P. Vellaisamy: MSU-STT-351-Sum-19A)

Probability & Statistics for Engineers

4/31

Point Estimation

Example 1 (Ex. 4): Let X and Y denote the strength of concrete beams and cylinders. The following data are obtained

$X : 5.9, 7.2, 7.3, 6.3, 8.1, 6.8, 7.0, 7.6, 6.8, 6.5, 7.0, 6.3, 7.9, 9.0,$
 $8.2, 8.7, 7.8, 9.7, 7.4, 7.7, 9.7, 7.8, 7.7, 11.6, 11.3, 11.8, 10.7,$

$Y : 6.1, 5.8, 7.8, 7.1, 7.2, 9.2, 6.6, 8.3, 7.0, 8.3, 7.8, 8.1,$
 $7.4, 8.5, 8.9, 9.8, 9.7, 14.1, 12.6, 11.2,$

Suppose $E(X) = \mu_1$, $V(X) = \sigma_1^2$; $E(Y) = \mu_2$, $V(Y) = \sigma_2^2$.

(a) Show that $\bar{X} - \bar{Y}$ is an unbiased estimator of $\mu_1 - \mu_2$. Calculate it for the given data.

(b) Find the variance and standard deviation (standard error) of the estimator in Part(a), and then compute the estimated standard error.

(c) Calculate an estimate of the ratio σ_1/σ_2 of the two standard deviations.

(d) Suppose a single beam X and a single cylinder Y are randomly selected. Calculate an estimate of the variance of the difference $X - Y$.

Point Estimation

Solution: (a) $E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2$. Hence, the unbiased estimate based on the given data is

$$\bar{X} - \bar{Y} = 8.141 - 8.575 = 0.434$$

(b) $V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \sigma_1^2 + \sigma_2^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. Thus,

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{V(\bar{X} - \bar{Y})} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

An estimate would be

$$S_{\bar{X}-\bar{Y}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{(1.666)^2}{27} + \frac{(2.104)^2}{20}} = 0.5687.$$

Note S_1 is not an unbiased estimator of σ_1 . Similarly, S_1/S_2 is not an unbiased estimator of σ_1/σ_2 .

Point Estimation

(c) An estimate of σ_1/σ_2 is (this is a biased estimate)

$$\frac{S_1}{S_2} = \frac{1.660}{2.104} = 0.7890.$$

(d) Note that

$$V(X - Y) = V(X) + V(Y) = \sigma_1^2 + \sigma_2^2.$$

$$\text{Hence, } \sigma_1^2 + \sigma_2^2 = (1.66)^2 + (2.104)^2 = 7.1824$$

Example 2 (Ex 8): In a random sample of 80 components of a certain type, 12 are found to be defective.

- (a) Give a point estimate of the proportion of all non-defective units.
- (b) A system is to be constructed by randomly selecting two of these components and connecting them in series. Estimate the proportion of all such systems that work properly.

Point Estimation

Solution: (a) With p denoting the true proportion of non-defective components,

$$\hat{p} = \frac{80 - 12}{80} = 0.85$$

(b) $P(\text{system works}) = p^2$, since the system works if and only if both components work. So, an estimate of this probability is

$$\hat{p} = \left(\frac{68}{80}\right)^2 = .723$$

Variances of estimators

The unbiased estimators are not in general unique. Given two unbiased estimators, it is natural to choose the one with less variance. In some cases, depending on the form of $F(x|\theta)$, we can find the unbiased estimator with minimum variance, called the MVUE. For instance, in the $N(\mu, 1)$ case, the MVUE of μ is \bar{X} .

Point Estimation

Example 3 (Ex 10): Using a rod of length μ , you lay out a square plot whose length of each side is μ . Thus, the area of the plot will be μ^2 (unknown). Based on n independent measurements X_1, \dots, X_n of the length, estimate μ^2 . Assume that each X_i has mean μ and variance σ^2 .

- (a) Show that \bar{X}^2 is not an unbiased estimator for μ^2 .
(b) For what value of k is the estimator $\bar{X}^2 - kS^2$ unbiased for μ^2 ?

Solution: (a) Note $E(\bar{X}^2) = \text{Var}(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2$. So, the bias of the estimator \bar{X}^2 is $E(\bar{X}^2 - \mu^2) = \frac{\sigma^2}{n}$. Also, \bar{X}^2 tends to overestimate μ^2 .

(b) Also,

$$E(\bar{X}^2 - kS^2) = E(\bar{X}^2) - kE(S^2) = \mu^2 + \frac{\sigma^2}{n} - k\sigma^2.$$

Hence, with $k = 1/n$, $E(\bar{X}^2 - kS^2) = \mu^2$.

Point Estimation

The Standard Error of an Estimator

It is useful to report the **standard error** of the estimator, in addition to its value. Unfortunately, it depends on the unknown parameters, and hence its estimate is usually used.

For a binomial model the estimator $\hat{p} = S_n/n$ of p , has the standard deviation $\sqrt{\frac{p(1-p)}{n}}$ which depends on p (unknown).

To estimate μ based on a random sample from a normal distribution, we use the estimator \bar{X} , whose standard deviation $\frac{\sigma}{\sqrt{n}}$ which depends on another unknown parameter σ .

Using estimates of p and σ , we obtain

$$\text{s.e.}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \text{s.e.}(\bar{X}) = \frac{s}{\sqrt{n}}.$$

Point Estimation

Example 4 (Ex 12): Suppose fertilizer-1 has a mean yield per acre of μ_1 with variance σ^2 , whereas the expected yield for fertilizer-2 is μ_2 with the same variance σ^2 . Let S_i^2 denote the sample variances of yields based on sample sizes n_1 and n_2 , respectively, of the two fertilizers. Show that the pooled (combined) estimator

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

is an unbiased estimator of σ^2 .

Solution:

$$\begin{aligned} E\left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right] &= \frac{(n_1 - 1)}{n_1 + n_2 - 2} E(S_1^2) + \frac{(n_2 - 1)}{n_1 + n_2 - 2} E(S_2^2) \\ &= \frac{(n_1 - 1)}{n_1 + n_2 - 2} \sigma^2 + \frac{(n_2 - 1)}{n_1 + n_2 - 2} \sigma^2 \\ &= \sigma^2. \end{aligned}$$

Point Estimation

Method of Estimation

It is desirable to have **some general** methods of estimation which yield estimators with some good properties. One of the classical methods is the method of moments (MoM), though it is not frequently used these days. The maximum likelihood (ML) method is one of the popular methods and the resulting maximum likelihood estimators (MLEs) have several finite and large sample properties.

The method of moments

Early in the development of statistics, the moments of a distribution (mean, variance, skewness, kurtosis) were discussed in depth, and estimators were formulated by equating the sample moments (i.e., \bar{x}, s^2, \dots) to the corresponding population moments, which are functions of the parameters. The number of equations should be equal to the number of parameters.

Point Estimation

Example 1: Consider the exponential distribution, $E(\lambda)$, with density

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then $E[X] = 1/\lambda$, and so solving $\bar{X} = \frac{1}{\lambda}$, we obtain MoM as $\hat{\lambda} = (1/\bar{X})$.

Drawbacks of MoM estimators

(i) A drawback of the MoM estimators is that it is difficult to solve the associated equations. Consider the parameters α and β in a Weibull distribution (see pp. 181-183). In this case, we need to solve

$$\mu = \beta \Gamma\left(1 + \frac{1}{\alpha}\right), \quad \sigma^2 = \beta^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^2 \right],$$

which is not an easy one.

(ii) Since MoM estimators use only a few population moments and their sample counterparts, the resulting estimators may sometimes be unreasonable, as in the following example.

(P. Vellaisamy: MSU STT 351 Sum 19A)

Probability & Statistics for Engineers

13/31

Point Estimation

Example 5: Suppose X_1, \dots, X_n is a random sample from uniform $U(0, \theta)$ distribution. Then solving $E(X) = \frac{\theta}{2} = \bar{X}$, we get MoM estimator as $\hat{\theta} = 2\bar{X}$. It is possible $\hat{\theta} > \max(X_i)$, while each $X_i < \theta$.

Example 6 (Ex 22): Let X denote the proportion of allotted time that a randomly selected student spends working on a certain aptitude test. Suppose the pdf of X is

$$f(x; \theta) = \begin{cases} (\theta + 1)x^\theta, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

where $-1 < \theta$.

A random sample of ten students yields data

$$x_1 = 0.92, x_2 = 0.79, x_3 = 0.90, x_4 = 0.65, x_5 = 0.86,$$

$$x_6 = 0.47, x_7 = 0.73, x_8 = 0.97, x_9 = 0.94, x_{10} = 0.77.$$

(P. Vellaisamy: MSU STT 351 Sum 19A)

Probability & Statistics for Engineers

14/31

Point Estimation

(a) Obtain MoM estimator and find it from the above data.

(b) Obtain MLE of θ , and compute it for the given data.

Solution: (a)

$$E(X) = \int_0^1 x(\theta + 1)x^\theta dx = \frac{\theta + 1}{\theta + 2} = 1 - \frac{1}{\theta + 2}.$$

So, the moment estimator $\hat{\theta}$ is the solution to $\bar{X} = 1 - \frac{1}{\hat{\theta} + 2}$, yielding

$$\hat{\theta} = \frac{1}{1 - \bar{X}} - 2.$$

For the given data, $\bar{X} = 0.80$, $\hat{\theta} = 5 - 2 = 3$.

Point Estimation

Maximum Likelihood Estimators

The ML method, introduced by R.A. Fisher, is based on the likelihood function of unknown parameter.

Definition: Let $X = (X_1, \dots, X_n)$ be a random sample from $f(x|\theta)$. Then the joint density

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f(x_i|\theta) = L(\theta|x) \quad (\text{viewed as a function of } \theta)$$

is called the "likelihood function" of θ , for an observed $X = x = (x_1, \dots, x_n)$.

An estimate $\hat{\theta}(x)$ that maximizes the $L(\theta|x)$ is called a maximum likelihood estimate of θ . Also, the estimator $\hat{\theta}(X) = \hat{\theta}(X_1, \dots, X_n)$ is called the maximum likelihood estimator (MLE) of θ . Here, θ may be a vector. This method yields estimators that have many desirable properties; both finite as well as large sample properties. The basic idea to find an estimator $\hat{\theta}(x)$ which is the most likely given the data $X = (X_1, \dots, X_n)$.

(P. Vellaisamy: MSU STT 351 Sum 19A)

Probability & Statistics for Engineers

15/31

(P. Vellaisamy: MSU STT 351 Sum 19A)

Probability & Statistics for Engineers

16/31

Point Estimation

Example 7: Consider the density, discussed in Example 6,

$$f(x; \theta) = \begin{cases} (\theta + 1)x^\theta, & 0 \leq x \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Obtain the MLE of θ and compute it for the data given there.

Solution: Note the likelihood function is

$$f(x_1, \dots, x_n; \theta) = (\theta + 1)^n (x_1 x_2 \dots x_n)^\theta.$$

So, the log-likelihood is

$$n \ln(\theta + 1) + \theta \sum \ln(x_i).$$

Taking $\frac{d}{d\theta}$ and equating to 0 yields $\frac{n}{\theta + 1} = -\sum \ln(x_i)$. Solve for θ to get

$$\hat{\theta} = -\frac{n}{\sum \ln(x_i)} - 1.$$

Taking $\ln(x_i)$ for each given x_i yields ultimately $\hat{\theta} = 3.12$.

(P. Vellaisamy: MSU-STT351-Sum-19A)

Probability & Statistics for Engineers

Point Estimation

Example 8: Let $X \sim B(1, p)$, Bernoulli distribution, with pmf

$$P(X = x|p) = p(x|p) = p^x (1-p)^{1-x}, \quad x = 0, 1,$$

where $p = P(X = 1)$. Find the MLE of p , based on X_1, \dots, X_n .

Solution: Aim is to estimate the population proportion p based on a random sample $X = (X_1, \dots, X_n)$ of size n . Note X_1, \dots, X_n are independent and identically distributed random variables.

For $x_i \in \{0, 1\}$, we have the joint pmf of X_1, \dots, X_n is (using independence)

$$\begin{aligned} P(X_1 = x_1, \dots, X_n = x_n) &= P(X_1 = x_1) \dots P(X_n = x_n) \\ &= p^{x_1} (1-p)^{1-x_1} \dots p^{x_n} (1-p)^{1-x_n} \\ &= p^{\sum x_i} (1-p)^{n-\sum x_i}, \end{aligned}$$

since X_i 's have identical pmf.

Point Estimation

Write the above density as a function of p , the likelihood function is

$$L(p|x) = p^{\sum x_i} (1-p)^{n-\sum x_i} = p^{s_n} (1-p)^{n-s_n},$$

where $s_n = \sum x_i$.

Choose an estimator that maximizes $L(p|x)$. Take

$$\ell = \ln L = s_n \ln p + (n - s_n) \ln(1-p).$$

Now

$$\begin{aligned} \frac{\partial \ln L}{\partial p} &= 0 \Rightarrow \frac{s_n}{p} - \frac{n - s_n}{1-p} = 0 \\ &\Rightarrow \hat{p} = \frac{s_n}{n} = p, \end{aligned}$$

the sample mean (proportion).

Also, it can be shown that

$$\frac{\partial^2 \ell}{\partial p^2}|_{\hat{p}} < 0.$$

Hence, $\hat{p} = S_n/n = \sum_{i=1}^n X_i$, the sample proportion, is the MLE of p .

(P. Vellaisamy: MSU-STT351-Sum-19A)

Probability & Statistics for Engineers

Point Estimation

Example 9: Let X_1, \dots, X_n be a random sample from $N(\mu, \sigma^2)$, where both mean μ and σ^2 are unknown. Find the MLE's of μ and σ^2 .

Solution: Let $\underline{\theta} = (\mu, \sigma^2)$. Then

$$f(x_i|\underline{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2} = (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2}.$$

Hence, the joint density is

$$\begin{aligned} f(x_1, \dots, x_n|\underline{\theta}) &= f(x_1|\underline{\theta}) f(x_2|\underline{\theta}) \dots f(x_n|\underline{\theta}) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_i^n \left(\frac{x_i - \mu}{\sigma} \right)^2} \\ &= L(\mu, \sigma^2|x). \end{aligned}$$

(P. Vellaisamy: MSU-STT351-Sum-19A)

Probability & Statistics for Engineers

Point Estimation

Then

$$\begin{aligned}\ell = \ln L(\mu, \sigma^2 | x) &= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.\end{aligned}$$

Then, for all $\sigma^2 > 0$,

$$\frac{\partial \ln L}{\partial \mu} = 0 \Rightarrow \hat{\mu} = \bar{x}.$$

Substituting $\hat{\mu} = \bar{x}$ in $\ell(\mu, \sigma^2)$, we get

$$\ell(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Point Estimation

Then

$$\frac{\partial \ell(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Hence,

$$\frac{\partial \ell(\mu, \sigma^2)}{\partial \sigma^2} = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Also, the Hessian matrix of second order partial derivatives of $\ell(\bar{x}, \sigma^2)$, calculated at $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2$, can be shown to be nonnegative definite.

Therefore, $\hat{\mu}$ and $\hat{\sigma}^2$ are the MLEs of μ and σ^2 .

Point Estimation

Example 10: Let X_1, \dots, X_n be a random sample from exponential density

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Find the MLE of λ .

Solution: The joint density of X_1, \dots, X_n (likelihood function) is

$$f(x|\lambda) = \prod_{i=1}^n f(x_i|\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_i x_i}.$$

Hence,

$$\begin{aligned}L(\lambda|x) &= \lambda^n e^{-n\lambda \bar{x}} \\ \Rightarrow \ell &= \log(L) = n \ln(\lambda) - n\lambda \bar{x}; \\ \frac{\partial \ell}{\partial \lambda} &= 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{x}}.\end{aligned}$$

Thus, the MLE of λ is $\hat{\lambda} = \frac{1}{\bar{x}}$.

Point Estimation

Example 11 (Ex 29): Suppose n time headways X_1, \dots, X_n in a traffic flow follow a shifted-exponential with pdf

$$f(x|\lambda, \theta) = \begin{cases} \lambda e^{-\lambda(x-\theta)}, & x \geq \theta; \\ 0, & \text{otherwise.} \end{cases}$$

(a) Obtain the MLE's of θ and λ .

(b) If $n = 10$ time headway observations are

$$3.11, .64, 2.55, 2.20, 5.44, 3.42, 10.39, 8.93, 17.82, 1.30$$

calculate the estimates of θ and λ .

Point Estimation

Solution: (a) The joint pdf of X_1, \dots, X_n is

$$\begin{aligned} f(x_1, \dots, x_n | \lambda, \theta) &= \prod_{i=1}^n f(x_i | \lambda, \theta) \\ &= \begin{cases} \lambda^n e^{-\lambda \sum_{i=1}^n (x_i - \theta)}, & x_1 \geq \theta, \dots, x_n \geq \theta \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Notice that $x_1 \geq \theta, \dots, x_n \geq \theta$ iff $\min(x_i) \geq \theta$, and also

$$-\lambda \sum_{i=1}^n (x_i - \theta) = -\lambda \sum_{i=1}^n x_i + n\lambda\theta.$$

Point Estimation

Let now $\min_{1 \leq i \leq n}(x_i) = x_{(1)}$. Then the likelihood function is

$$L(\lambda, \theta | x) = \begin{cases} \lambda^n e^{n\lambda\theta - \lambda \sum_{i=1}^n x_i}, & x_{(1)} \geq \theta \\ 0, & \text{otherwise.} \end{cases}$$

Consider first the maximization with respect to θ . Note first the likelihood is zero for $\theta > x_{(1)}$. Also, it is increasing in θ for $\theta \leq x_{(1)}$ and hence attains the maximum when $\theta = x_{(1)}$. Hence, the MLE of θ is $\hat{\theta} = x_{(1)}$.

Now, substituting $\hat{\theta}$ in likelihood function

$$L(\lambda, \hat{\theta} | x) = \lambda^n e^{n\lambda x_{(1)} - \lambda \sum_{i=1}^n x_i} = \lambda^n e^{-n \sum_{i=1}^n (x_i - x_{(1)})}.$$

Point Estimation

This implies (taking log on both sides)

$$\begin{aligned} \ell(\lambda, \hat{\theta} | x) &= \ln(L(\lambda, \hat{\theta} | x)) = n \ln(\lambda) - n \sum_{i=1}^n (x_i - x_{(1)}) \\ &\Rightarrow \frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n (x_i - x_{(1)}) = 0. \end{aligned}$$

Solving for λ , the MLE of λ as

$$\hat{\lambda} = \frac{n}{\sum(x_i - x_{(1)})}.$$

(b) From the data, $\hat{\theta} = \min(x_i) = .64$ and $\sum_{i=1}^n x_i = 55.80$. hence,

$$\hat{\lambda} = \frac{10}{55.80 - 6.4} = 0.202.$$

Point Estimation

Properties of MLE's

(i) For large n , the MLE $\hat{\theta}(X)$ is asymptotically normal, unbiased, and has variance smaller than any other estimator.

(ii) Invariance property: If $\hat{\theta}$ is an MLE of θ , then $g(\hat{\theta})$ is an MLE of $g(\theta)$ for any function g .

Example 12: Let X_1, \dots, X_n be a random sample from exponential distribution $E(\lambda)$ with parameter λ . Find the MLE of the mean of the distribution.

Solution: As seen in Example 10, the MLE of λ is $\hat{\lambda} = \frac{1}{\bar{x}}$. Then the MLE of $g(\lambda) = \frac{1}{\lambda} = E(X_i)$ is

$$\hat{g}(\lambda) = \frac{1}{\hat{\lambda}} = \bar{x},$$

using the invariance property of the MLE.

Point Estimation

Example 11 (Ex 26): The following data represents shear strength (X) of the test spot weld 392, 376, 401, 367, 389, 362, 409, 415, 358, 375.

(a) Assuming that X is normally distributed, estimate the true average shear strength and standard deviation of shear strength using the method of maximum likelihood.

(b) Obtain the MLE of $P(X \leq 400)$.

Solution: (a) The MLE's of μ and σ^2 are

$$\hat{\mu} = \bar{X}; \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} S^2.$$

Hence, the MLE of σ is $\hat{\sigma} = \sqrt{\frac{n-1}{n} S^2}$.

Point Estimation

From the data given: $\hat{\mu} = \bar{X} = 384.4$; $S^2 = 395.16$. So,

$$\frac{1}{n} \sum (X_i - \bar{X})^2 = \hat{\sigma}^2 = \frac{9}{10} (395.16) = 355.64 \text{ and } \hat{\sigma} = \sqrt{355.64} = 18.86.$$

(b) Let $\theta = P(X \leq 400)$. Then

$$\begin{aligned} \theta &= P\left(\frac{X - \mu}{\sigma} \leq \frac{400 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{400 - \mu}{\sigma}\right) \quad (\text{note } Z \sim N(0, 1)) \\ &= \Phi\left(\frac{400 - \mu}{\sigma}\right). \end{aligned}$$

The MLE of θ , by invariance property, is

$$\hat{\theta} = \Phi\left(\frac{400 - \hat{\mu}}{\hat{\sigma}}\right) = \Phi\left(\frac{400 - 384.4}{18.86}\right) = 0.7881.$$