MATHEMATICAL STATISTICS - 1 | CHAPTER 6.1

POINT ESTIMATION

A family of pdf's : $\{f(x;\theta): \theta \in \Omega\}$, Ω: Parameter Space

An arbitrary member of pdf's: $f(x;\theta)$, $\theta \in \Omega$

A specific member of pdf's ; $f(x;\theta_0)$

For example, we may have the family

$\{Normal\ (\theta;1): \theta \in \Omega\}$ $\Omega:(-\infty, \infty)$

An arbitrary member is $Normal(\theta;1) : \theta \in \Omega$

A (specific) member of pdf's $Normal(0;1)$

Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution that has pdf which is one member of $\{f(x;\theta): \theta \in \Omega\}$

That is, our sample arises from pdf $f(x;\theta : \theta \in \Omega$

Our problem is that of defining a statistics (point estimator)

$$\boxed{Y_1 = u_1(X_1, X_2, \ldots, X_n)}$$

so that the number $y_1 = u_1(x_1, x_2, \ldots x_n)$ will be a good point estimate of θ

Maximum Likelihood Estimation (MLE)

Let $X_1, X_2, \ldots, X_n$ is a random Saple from $f(x;\theta)$

likelihood function is the joint distribution of the r.s.

$L(\theta) = f(x_1;\theta) . f(x_2;\theta) . \ldots . f(x_n;\theta)$

which value of θ maximizes $L(\theta)$? That is $\hat{\theta}_{MLE}$!

To find MLE, we have the following steps:

(i) Find $L(\theta) = f(x_1;\theta) . f(x_2;\theta) . \dots . f(x_n;\theta)$

(ii) $L(\theta)$ and its logarithm are maximized for the same value of θ. Take $\ln L(\theta)$

(iii) Find $\frac{\partial \ln L(\theta)}{\partial\theta}$ and solve $\frac{\partial \ln L(\theta)}{\partial\theta} = 0$

The solution is maximum likelihood estimator.

Example: let $X_i \overset{i.i.d}{\sim}$ Bernoulli (θ), $i = 1, 2, \dots, n$, $0 \le \theta \le 1$
Find mle of θ.

Answer: Remember, Bernoulli is a single trial of Binomial distribution (namely Binomial with $n=1$). The pdf is;

$$f(x_i;\theta) = \theta^{x_i}(1-\theta)^{1-x_i}; \quad x_i = 0, 1 \; ; \; 0 \le \theta \le 1$$

(i) $L(\theta) = \theta^{x_1} . (1-\theta)^{1-x_1} . \theta^{x_2} . (1-\theta)^{1-x_2} . \dots . \theta^{x_n}(1-\theta)^{1-x_n}$

$$= \theta^{\Sigma x_i} . (1-\theta)^{n-\Sigma x_i}$$

(ii) $\ln L(\theta) = \Sigma x_i . \ln\theta + (n - \Sigma x_i) . \ln(1-\theta)$

(iii) $\frac{d \ln L(\theta)}{d\theta} = \frac{\Sigma x_i}{\theta} - \frac{n - \Sigma x_i}{1-\theta} = 0$

$$(1-\theta).\Sigma x_i = \theta.(n - \Sigma x_i)$$

$$\Sigma x_i - \cancel{\theta.\Sigma x_i} = n\theta - \cancel{\theta.\Sigma x_i}$$

$$\boxed{\hat{\theta}_{MLE} = \frac{\Sigma x_i}{n} = \overline{X}}$$

Example: let $X_i \overset{i.i.d}{\sim}$ Normal (θ_1, θ_2), $i = 1, 2, \dots, n$

$$\Omega : -\infty < \theta_1 < \infty \; : \; \theta_2 > 0$$

Find mle.'s of θ_1 and θ_2.

Answer: $f(x_i;\underbrace{\theta_1}_{\mu},\underbrace{\theta_2}_{\sigma^2}) = \frac{1}{\sqrt{2\pi\theta_2}} \cdot \exp\left[-\frac{1}{2}\cdot\frac{(x_i-\theta_1)^2}{\theta_2}\right]$

(i) $L(\theta_1,\theta_2) = (2\pi\theta_2)^{-n/2} \cdot \exp\left[-\frac{1}{2}\sum\frac{(x_i-\theta_1)^2}{\theta_2}\right]$

(ii) $\ln L(\theta_1,\theta_2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\theta_2 - \frac{1}{2}\sum\frac{(x_i-\theta_1)^2}{\theta_2}$

(iii) $\frac{\partial \ln(\theta_1,\theta_2)}{\partial\theta_1} = -\frac{1}{2\theta_2}\cdot(-2)\cdot\sum(x_i-\theta_1) = 0$

$$\sum(x_i-\theta_1) = 0$$

$$\sum x_i - n\cdot\theta_1 = 0$$

$$\boxed{\hat{\theta}_{1(MLE)} = \frac{\sum x_i}{n} = \bar{X}}$$

$$\frac{\partial \ln(\theta_1,\theta_2)}{\partial\theta_2} = -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2}\cdot\sum(x_i-\theta_1)^2 = 0$$

$$\frac{\sum(x_i-\theta_1)^2}{2\theta_2^2} = \frac{n}{2\theta_2}$$

$$\hat{\theta}_{2(MLE)} = \frac{\sum(x_i-\hat{\theta}_{1(MLE)})^2}{n} = S^2$$

Note that, in the final step, we have replaced θ_1 with its MLE because $y_1(x_1,\dots,x_n)$ statistics should NOT contain any unknown parameter.

Unbiased Estimator: An estimator $\hat{\theta}$ is unbiased if $\boxed{E(\hat{\theta}) = \theta}$.

Consistent Estimator: An estimator $\hat{\theta}$ is consistent if $\hat{\theta} \xrightarrow[P]{} \theta$.

Note that unbiasedness is stronger than consistency.

Example: Check if the mle's of Normal distribution are
(i) Unbiased (ii) Consistent.

Answer: Remember; $X_i \sim Normal(\theta_1, \theta_2)$ and mle's are:

$$\hat{\theta}_1 = \bar{X} = \frac{\sum x_i}{n}; \quad \hat{\theta}_2 = S^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

For Mean: $\mu = \theta_1$

$$E(\hat{\theta}_1) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n} \cdot \left(E(X_1) + E(X_2) + \dots + E(X_n)\right) = \frac{1}{n} \cdot n \cdot \theta_1 = \theta_1$$

$\hat{\theta}_1$ is an unbiased estimator of θ_1.

Remember; $Y_n \xrightarrow{p} \mu$ iff (i) $E(Y_n) \xrightarrow[n \to \infty]{} \mu$

(ii) $Var(Y_n) = \sigma_n^2 < \infty$ (Finite $\forall n$)

(iii) $Var(Y_n) \xrightarrow[n \to \infty]{} 0$

Then, we have;

(i) $E(\hat{\theta}_1) = \theta$

(ii) $Var(\hat{\theta}_1) = \frac{\sigma^2}{n}$ is finite $\forall n$

(iii) $\lim_{n \to \infty} Var(\hat{\theta}_1) = \lim_{n \to \infty} \left(\frac{\sigma^2}{n}\right) = 0$

So; $\hat{\theta}_1 \xrightarrow{p} \theta$ and $\hat{\theta}_1$ is a consistent estimator.

For Variance; $\sigma^2 = \theta_2$ (Remember $W = \frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}$ and $E(W) = n-1$, $Var(W) = 2.(n-1)$)

$$E(\hat{\theta}_2) = E(S^2) = \frac{\sigma^2}{n} E\left(\frac{n \cdot S^2}{\sigma^2}\right) = \frac{\sigma^2}{n} \cdot (n-1)$$

$E(\hat{\theta}_2) = \frac{n-1}{n} \cdot \sigma^2 \neq \theta_2$ so, $\hat{\theta}_2$ is NOT an unbiased estimator of θ.

However,

(i) $E(\hat{\theta}_2) = \frac{n-1}{n} \cdot \theta_2 \xrightarrow[n \to \infty]{} \theta_2$

(ii) $Var(\hat{\theta}_2) = Var\left[\frac{\sigma^2}{n} \cdot \frac{n \cdot S^2}{\sigma^2}\right] = \frac{\sigma^4}{n^2} \cdot Var(W) = \frac{\sigma^4}{n^2} \cdot (2n-2) < \infty \ \forall n$

(iii) $\lim_{n \to \infty} Var(\hat{\theta}_2) = \lim_{n \to \infty} \left(\frac{2n-2}{n^2} \cdot \sigma^4\right) = 2\sigma^4 \cdot \lim_{n \to \infty} \left(\frac{n-1}{n^2}\right) = 0$

So, $\hat{\theta}_2$ is a consistent estimator of θ_2

*Note that, unbiasedness is NOT a general property of MLE but we can obtain unbiased estimators using MLE.

To illustrate; $E(\hat{\theta}_2) = \frac{n-1}{n} \cdot \theta_2$

$$E\left(\frac{n}{n-1} \cdot \hat{\theta}_2\right) = \theta_2$$

So; $\frac{n}{n-1} \cdot \hat{\theta}_2 = \frac{n}{n-1} \cdot S^2 = \frac{n}{n-1} \cdot \frac{\Sigma(X_i - \bar{X})}{n} = \frac{\Sigma(X_i-\bar{X})}{n-1} = S^{*2}$

is an unbiased estimator of θ_2.

*In ALL cases of practical interest, MLEs are consistent.

6.1. Let $X_1, X_2, \ldots, X_n$ represent a random sample from each of the distributions having the following probability density functions:

(a) $f(x;\theta) = \theta^x e^{-\theta}/x!$, $x = 0, 1, 2, \ldots$, $0 \le \theta < \infty$, zero elsewhere, where $f(0;0) = 1$.

(b) $f(x;\theta) = \theta x^{\theta-1}$, $0 < x < 1$, $0 < \theta < \infty$, zero elsewhere.

(c) $f(x;\theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty$, $0 < \theta < \infty$, zero elsewhere.

(d) $f(x;\theta) = \frac{1}{2}e^{-|x-\theta|}$, $-\infty < x < \infty$, $-\infty < \theta < \infty$.

(e) $f(x;\theta) = e^{-(x-\theta)}$, $\theta \le x < \infty$, $-\infty < \theta < \infty$, zero elsewhere.

In each case find the m.l.e. $\hat{\theta}$ of θ.

6.2. Let $X_1, X_2, \ldots, X_n$ be i.i.d., each with the distribution having p.d.f. $f(x;\theta_1,\theta_2) = (1/\theta_2)e^{-(x-\theta_1)/\theta_2}$, $\theta_1 \le x < \infty$, $-\infty < \theta_1 < \infty$, $0 < \theta_2 < \infty$, zero elsewhere. Find the maximum likelihood estimators of θ_1 and θ_2.

6.1) a) $X \sim$ Poisson (θ)

$$f(x_i;\theta) = \frac{e^{-\theta} \cdot \theta^{x}}{x!}$$

$$L(\theta) = \frac{e^{-\theta} \cdot \theta^{x_1}}{x_1!} \cdot \frac{e^{-\theta} \cdot \theta^{x_2}}{x_2!} \cdot \ldots \cdot \frac{e^{-\theta} \cdot \theta^{x_n}}{x_n!} = \frac{e^{-n\theta} \cdot \theta^{\Sigma x_i}}{\prod x_i!}$$

$$\ln L(\theta) = -n \cdot \theta + \Sigma X_i \cdot \ln\theta - \Sigma \ln x_i!$$

$$\frac{d \ln L(\theta)}{d\theta} = -n + \frac{\Sigma X_i}{\theta} = 0 \Rightarrow \hat{\theta}_{MLE} = \frac{\Sigma X_i}{n} = \bar{X}$$

b) $X \sim Beta(\alpha = \theta ; \beta = 1)$

$$f(x_i ; \theta) = \theta \cdot x^{\theta - 1}$$

$$L(\theta) = \theta \cdot x_1^{\theta-1} \cdot \theta \cdot x_2^{\theta-1} \cdot \ldots \cdot \theta \cdot x_n^{\theta-1} = \theta^n \cdot (\Pi x_i)^{\theta-1}$$

$$\ln L(\theta) = n \cdot \ln \theta + (\theta - 1) \cdot \sum \ln x_i$$

$$\frac{d \ln L(\theta)}{d\theta} = \frac{n}{\theta} + \sum \ln X_i = 0 \Rightarrow \hat{\theta}_{MLE} = \frac{-\sum \ln X_i}{n} = \underline{\underline{-\overline{\ln X}}}$$

(Note that; $X \sim Beta(\alpha ; \beta) \Rightarrow f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \cdot x^{\alpha - 1} \cdot (1 - x)^{\beta - 1}$

remember; $\Gamma(n) = (n-1)!$ for n integer and

$\frac{\Gamma(\theta + 1)}{\Gamma(\theta)\Gamma(1)} = \frac{\theta!}{(\theta - 1)! \cdot 0!} = \theta$ is true even if θ is NOTEN)

c) $X \sim Exponential(\theta)$

$$f(x_i ; \theta) = \frac{1}{\theta} \cdot e^{-x_i/\theta}$$

$$L(\theta) = \frac{1}{\theta} \cdot e^{-x_1/\theta} \cdot \frac{1}{\theta} \cdot e^{-x_2/\theta} \cdot \ldots \cdot \frac{1}{\theta} \cdot e^{-x_n/\theta} = \frac{1}{\theta^n} \cdot e^{-\frac{1}{\theta} \cdot \sum x_i}$$

$$\ln L(\theta) = -n \cdot \ln \theta - \frac{1}{\theta} \cdot \sum x_i$$

$$\frac{d \ln L(\theta)}{d\theta} = -\frac{n}{\theta} + \frac{\sum X_i}{\theta^2} = 0$$

$$\frac{\sum X_i}{\theta^2} = \frac{n}{\theta} \Rightarrow \hat{\theta}_{MLE} = \frac{\sum X_i}{n} = \underline{\underline{\bar{X}}}$$

d) $X \sim$ Double Exponential $(\mu=\theta;\ \beta=1)$

$$f(x_i;\theta)=\frac{1}{2}\cdot e^{-|x-\theta|}$$

$$L(\theta)=\frac{1}{2}\cdot e^{-|x_1-\theta|}\cdot\frac{1}{2}\cdot e^{-|x_2-\theta|}\cdot\ \dots\ \cdot\frac{1}{2}\cdot e^{-|x_n-\theta|}$$

$$=\left(\frac{1}{2}\right)^n\cdot e^{-\sum|x_i-\theta|}$$

$$\ln L(\theta)=-n\ln 2-\sum|x_i-\theta|$$

$$\frac{d\ln L(\theta)}{d\theta}=\frac{d}{d\theta}\left[-\sum|x_i-\theta|\right] \text{ is independent of } \theta.$$

However, $\sum|X_i-\theta|$ is minimum (and so, $e^{-\sum|X_i-\theta|}$ is maximum) for $\theta=\text{Median}(X_i)=\tilde{X}$

So; $\hat{\theta}_{MLE}=\tilde{X}$

(Note that: $X\sim$ Double Exponential $(\mu;\beta)$

$$f(x)=\frac{1}{2\beta}\exp\left(-\frac{|x-\mu|}{\beta}\right)\quad \text{Mean}: \mu;\ \text{Variance}: 2\beta^2$$

e) $f(x_i;\theta)=e^{-(x_i-\theta)}\qquad \theta\le x_i<\infty$

$$L(\theta)=e^{-(x_1-\theta)}\cdot e^{-(x_2-\theta)}\cdot\ \dots\ \cdot e^{-(x_n-\theta)}=e^{-\sum(x_i-\theta)}$$

$$\ln L(\theta)=-\sum(X_i-\theta)=-\sum X_i+n\cdot\theta$$

$\frac{d\ln L(\theta)}{d\theta}$ is independent of θ.

However, $\theta\le x_i \Rightarrow \sum(x_i-\theta)$ is minimized (and so, $e^{-\sum(x_i-\theta)}$ is maximized) when θ is maximum.

Max. Value of θ is $X_{(1)} \Rightarrow \hat{\theta}_{MLE}=X_{(1)}$ → First ORDER Statistics.

6.2) $X \sim$ Two-Parameter Exponential $(\theta_1; \theta_2)$

$$f(x_i; \theta_1, \theta_2) = \frac{1}{\theta_2} \cdot e^{-(x-\theta_1)/\theta_2} \quad \theta_1 \leq x < \infty;\ \theta_1 \in \mathbb{R},\ \theta_2 > 0$$

$$L(\theta_1, \theta_2) = \theta_2^{-n} \cdot e^{-\frac{1}{\theta_2}\sum(x_i - \theta_1)}$$

$$\ln L(\theta_1, \theta_2) = -n \ln \theta_2 - \frac{1}{\theta_2} \cdot \sum(x_i - \theta_1)$$

As we have seen in 6.1(e), $\hat{\theta}_{MLE} = X_{(1)}$ → First order Statistics

$$\frac{d}{d\theta_2} \ln(\theta_1, \theta_2) = -\frac{n}{\theta_2} + \frac{1}{\theta_2^2} \sum(x_i - \theta_1) = 0$$

$$\hat{\theta}_{2(MLE)} = \frac{\sum(x_i - \hat{\theta}_1)}{n} = \frac{\sum(x_i - x_{(1)})}{n}$$

Example: let $X \sim$ Uniform (θ)

$$f(x; \theta) = \frac{1}{\theta} \quad 0 < x \leq \theta,\ \theta > 0$$

Find the MLE of θ. Is it unbiased? If NOT, find an unbiased estimator of θ using the estimator you've found. Also check if MLE of θ is consistent.

Answer: $f(x_i; \theta) = \frac{1}{\theta} \quad 0 < x \leq \theta,\ \theta > 0$

$$L(\theta) = \frac{1}{\theta} \cdot \frac{1}{\theta} \cdot \ldots \cdot \frac{1}{\theta} = \frac{1}{\theta^n} \quad 0 < x_i \leq \theta$$

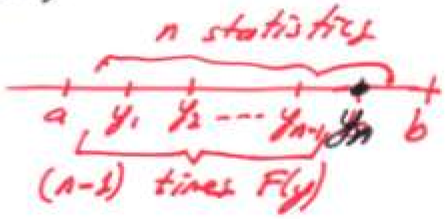
$\ln L(\theta) = -n \cdot \ln \theta$ and $\frac{d \ln(\theta)}{d\theta} = \frac{-n}{\theta} = 0$ has NO solution for θ.

However, $L(\theta)$ is maximized when θ is minimum.
Since $\theta \geq x_i$, minimum value of θ is $X_{(n)}$ → n^{th} order statistics.

So; $\hat{\theta}_{MLE} = X_{(n)}$

Remember, n^{th} order statistics have pdf

$$g_n(y) = \cancel{n!}\,n \cdot \frac{[F(y)]^{n-1}}{\cancel{(n-1)!}} \cdot f(y)$$



Then, $X_i \sim Uniform(\theta)$, $i = 1, 2, \ldots, n$

$$f(x) = \frac{1}{\theta} \qquad F(x) = \frac{x}{\theta} \quad \text{and} \quad g_n(x) = n \cdot \left(\frac{x}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} = \frac{n \cdot x^{n-1}}{\theta^n}$$

$$E(X_{(n)}) = \int_0^\theta x\, g_n(x)\,dx = \int_0^\theta \frac{n \cdot x^{n-1+1}}{\theta^n}dx = \frac{n}{\theta^n} \cdot \int_0^\theta x^n dx = \frac{n}{\theta^n} \cdot \frac{x^{n+1}}{n+1}\Bigg]_0^\theta = \frac{n}{n+1} \cdot \theta$$

$E(X_{(n)}) = \frac{n}{n+1} \cdot \theta \neq \theta$, $X_n = \hat{\theta}_{MLE}$ is NOT an unbiased estimator of θ.

$E\left[\frac{n+1}{n} \cdot X_{(n)}\right] = \theta$ so; $\hat{\theta} = \frac{n+1}{n} \cdot X_{(n)}$ is an Unbiased Estimator of θ.

To check for Consistency, we need $Var(\hat{\theta})$

$$E(X_{(n)}^2) = \int_0^\theta x^2 \cdot \frac{n \cdot x^{n-1}}{\theta^n}dx = \frac{n}{\theta^n} \cdot \frac{x^{n+2}}{n+2}\Bigg]_0^\theta = \frac{n}{n+2} \cdot \theta^2$$

$$Var(\hat{\theta}) = E(X_{(n)}^2) - E^2(X_{(n)}) = \frac{n}{n+2}\theta^2 - \frac{(n)^2}{(n+1)^2} \cdot \theta^2$$

$$= n\theta^2\left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right] = n\theta^2\left[\frac{\cancel{n^2}+\cancel{2n}+1-\cancel{n^2}-\cancel{2n}}{(n+2)(n+1)^2}\right] = \frac{n}{(n+2)(n+1)^2}\theta^2$$

we have; (i) $E(\hat{\theta}_{MLE}) = E(X_{(n)}) = \frac{n}{n+1} \cdot \theta \xrightarrow[n \to \infty]{} \theta$

(ii) $Var(\hat{\theta}_{MLE}) = \frac{n}{(n+2)(n+1)^2} \cdot \theta^2 < \infty$ for all n

(iii) $\lim_{n \to \infty} Var(\hat{\theta}_{MLE}) = 0$. So $\hat{\theta} \xrightarrow[P]{} \theta$ and $\hat{\theta}$ is a consistent estimator of θ.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* The following Example Shows MLE is NOT always unique.

Ex: $X \sim$ Uniform $(\theta ; \theta+1)$

$f(x;\theta) = 1$ Find MLE of θ.

Answer; $f(x_i;\theta) = 1 \quad \theta \leq x_i \leq \theta+1$

$L(\theta) = 1.1. \ldots .1 = 1 \quad \theta \leq x_i \leq \theta+1$

$\theta \leq x_i \qquad x_i \leq \theta+1$

① $\theta \leq X_{(1)} \qquad X_{(n)} \leq \theta+1$

② $X_{(n)} - 1 \leq \theta$

① and ② $\Rightarrow X_{(n)} - 1 \leq \hat{\theta} \leq X_{(1)}$

Every ~~[illegible]~~ point in the interval $(X_{(n)} - 1 ; X_{(1)})$ is MLE.

6.3. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample from a distribution with p.d.f. $f(x;\theta) = 1, \theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2}, -\infty < \theta < \infty$, zero elsewhere. Show that every statistic $u(X_1, X_2, \ldots, X_n)$ such that

$$Y_n - \tfrac{1}{2} \leq u(X_1, X_2, \ldots, X_n) \leq Y_1 + \tfrac{1}{2}$$

is a m.l.e. of θ. In particular, $(4Y_1 + 2Y_n + 1)/6$, $(Y_1 + Y_n)/2$, and $(2Y_1 + 4Y_n - 1)/6$ are three such statistics. Thus uniqueness is not in general a property of a m.l.e.

6.3) $X \sim$ Uniform $(\theta - \frac{1}{2} ; \theta + \frac{1}{2})$

$f(x_i;\theta) = 1 \quad \theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2}$

$L(\theta) = 1.1. \ldots 1 = 1 \quad \theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2}$

$\theta - \frac{1}{2} \leq X_{(1)} \qquad X_{(n)} \leq \theta + \frac{1}{2}$

$\theta \leq X_{(1)} + \frac{1}{2}$ AND $X_{(n)} - \frac{1}{2} \leq \theta$

$$\boxed{X_{(n)} - \frac{1}{2} \leq \hat{\theta} \leq X_{(1)} + \frac{1}{2}}$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

6.4. Let X_1, X_2, and X_3 have the multinomial distribution in which $n = 25$, $k = 4$, and the unknown probabilities are θ_1, θ_2, and θ_3, respectively. Here we can, for convenience, let $X_4 = 25 - X_1 - X_2 - X_3$ and $\theta_4 = 1 - \theta_1 - \theta_2 - \theta_3$. If the observed values of the random variables are $x_1 = 4$, $x_2 = 11$, and $x_3 = 7$, find the maximum likelihood estimates of θ_1, θ_2, and θ_3.

6.5. The *Pareto distribution* is frequently used as a model in study of incomes and has the distribution function

$$F(x; \theta_1, \theta_2) = 1 - (\theta_1/x)^{\theta_2}, \qquad \theta_1 \le x, \text{ zero elsewhere,}$$

$$\text{where } \theta_1 > 0 \quad \text{and} \quad \theta_2 > 0.$$

If $X_1, X_2, \ldots, X_n$ is a random sample from this distribution, find the maximum likelihood estimators of θ_1 and θ_2.

6.4) $\underline{X} \sim Multinomial(25, \theta_1, \theta_2, \theta_3)$ whee $\underline{X} = (X_1, X_2, X_3)$

$$f(x_1, x_2, x_3) = \frac{25!}{x_1!\,.\,x_2!\,.\,x_3!\,.\,(25 - x_1 - x_2 - x_3)!} \cdot \theta_1^{x_1} \cdot \theta_2^{x_2} \cdot \theta_3^{x_3} \cdot (1 - \theta_1 - \theta_2 - \theta_3)^{25 - x_1 - x_2 - x_3}$$

$$L(\theta_1, \theta_2, \theta_3) = \frac{25!}{4!\,.\,11!\,.\,7!\,.\,3!} \cdot \theta_1^{4} \theta_2^{11} \theta_3^{7} (1 - \theta_1 - \theta_2 - \theta_3)^3$$

$$\ln L(\theta_1, \theta_2, \theta_3) = \ln\left(\frac{25!}{4!\,.\,11!\,.\,7!\,.\,3!}\right) + 4 \ln \theta_1 + 11 \ln \theta_2 + 7 \ln \theta_3 + 3 \ln (1 - \theta_1 - \theta_2 - \theta_3)$$

let $\underline{\theta} = (\theta_1, \theta_2, \theta_3)$

$$\frac{\partial \ln(\underline{\theta})}{\partial \theta_1} = \frac{4}{\theta_1} - \frac{3}{1 - \theta_1 - \theta_2 - \theta_3} = 0 \Rightarrow 3\theta_1 = 4 - 4\theta_1 - 4\theta_2 - 4\theta_3$$

$$\boxed{7\theta_1 + 4\theta_2 + 4\theta_3 = 4} \ (1)$$

$$\frac{\partial \ln(\underline{\theta})}{\partial \theta_2} = \frac{11}{\theta_2} - \frac{3}{1 - \theta_1 - \theta_2 - \theta_3} = 0 \Rightarrow 3\theta_2 = 11 - 11\theta_1 - 11\theta_2 - 11\theta_3$$

$$\boxed{11\theta_1 + 14\theta_2 + 11\theta_3 = 11} \ (2)$$

$$\frac{\partial \ln(\underline{\theta})}{\partial \theta_3} = \frac{7}{\theta_3} - \frac{3}{1 - \theta_1 - \theta_2 - \theta_3} = 0 \Rightarrow 3\theta_3 = 7 - 7\theta_1 - 7\theta_2 - 7\theta_3$$

$$\boxed{7\theta_1 + 7\theta_2 + 10\theta_3 = 7} \ (3)$$

① $7\theta_1 + 4\theta_2 + 4\theta_3 = 4 \Rightarrow$ ①' $\theta_2 + \theta_3 = 1 - \frac{7\theta_1}{4}$

② $11\theta_1 + 14\theta_2 + 11\theta_3 = 11 \Rightarrow$ ②' $\theta_1 + \theta_3 = 1 - \frac{14\theta_2}{11}$

$+$ ③ $7\theta_1 + 7\theta_2 + 10\theta_3 = 7 \Rightarrow$ ③' $\theta_1 + \theta_2 = 1 - \frac{10\theta_3}{7}$

$25\theta_1 + 25\theta_2 + 25\theta_3 = 22$

④ $\theta_1 + \theta_2 + \theta_3 = \frac{22}{25}$

④ & ①

$\theta_1 + 1 - \frac{7\theta_1}{4} = \frac{22}{25}$

$\frac{3}{25} = \frac{3\theta_1}{4}$

$\hat{\theta}_{1(MLE)} = \frac{4}{25}$

④ & ②'

$\theta_2 + 1 - \frac{14\theta_2}{11} = \frac{22}{25}$

$\frac{3}{25} = \frac{3\theta_2}{11}$

$\hat{\theta}_{2(MLE)} = \frac{11}{25}$

④ & ③'

$\theta_3 + 1 - \frac{10\theta_3}{7} = \frac{22}{25}$

$\frac{3}{25} = \frac{3\theta_3}{7}$

$\hat{\theta}_{3(MLE)} = \frac{7}{25}$

6.5) $X \sim Pareto(\theta_1 ; \theta_2)$

$$F(x ; \theta_1, \theta_2) = 1 - \left(\frac{\theta_1}{x}\right)^{\theta_2} \quad x \geq \theta_1 , \theta_1 > 0, \theta_2 > 0$$

$$f(x; \theta_1, \theta_2) = \frac{d}{dx} F(x; \theta_1, \theta_2) = \frac{\theta_1}{x^2} \cdot \theta_2 \cdot \left(\frac{\theta_1}{x}\right)^{\theta_2 - 1} = \theta_1^{\theta_2} \cdot \theta_2 \cdot \frac{1}{x^{\theta_2+1}}$$

$$L(\theta_1, \theta_2) = \theta_1^{n \cdot \theta_2} \cdot \theta_2^{n} \cdot \frac{1}{\prod x_i^{\theta_2+1}} \quad x_i \geq \theta_1$$

The only part of likelihood function regarding θ_1 is $\theta_1^{n \cdot \theta_2}$, which is maximized when θ_1 is maximum. So; $\boxed{\hat{\theta}_{1(MLE)} = X_{(1)}}$

$$\ln L(\theta_1, \theta_2) = n\theta_2 \cdot \ln\theta_1 + n \cdot \ln\theta_2 - (\theta_2 + 1) \cdot \ln \sum x_i$$

$$\frac{\partial \ln L(\theta_1, \theta_2)}{\partial \theta_2} = n \cdot \ln\theta_1 + \frac{n}{\theta_2} - \ln \sum x_i = 0$$

$$\frac{n}{\theta_2} = \ln \sum x_i - \ln \theta_1^n = \ln \frac{\sum x_i}{\theta_1^n} \Rightarrow \boxed{\hat{\theta}_{2(MLE)} = \frac{n}{\ln\left(\frac{\sum x_i}{X_{(1)}^n}\right)}}$$

Invariance Property of MLE

If $\hat{\theta}_{MLE}$ is MLE of θ, then $h(\hat{\theta}_{MLE})$ is MLE of $h(\theta)$.

Examples

(i) Let $X \sim Normal(\theta, 1)$. $\hat{\theta} = \bar{X}$ is MLE of θ.

Then; MLE of $h(\theta) = \theta^3$ is $h(\hat{\theta}) = \bar{X}^3$

(ii) Let $X \sim Bernoulli(\theta)$. $\hat{\theta} = \bar{X}$ is MLE of θ.

Then; MLE of $h(\theta) = \theta.(1-\theta)$ is $h(\hat{\theta}) = \bar{X}.(1-\bar{X})$

(iii) Let $X \sim Normal(\theta_1, \theta_2)$. We found $\hat{\theta}_1 = \bar{X}$ and $\hat{\theta}_2 = S^2$

Then; MLE of $h(\theta_1, \theta_2) = \theta_1 + 2\sqrt{\theta_2}$ is $\widehat{h(\theta_1, \theta_2)} = \bar{X} + 2S$

6.11. Let the table

x	0	1	2	3	4	5
Frequency	6	10	14	13	6	1

represent a summary of a sample of size 50 from a binomial distribution having $n = 5$. Find the m.l.e. of Pr $(X \geq 3)$.

6.11) $X \sim Binomial(n=5; \theta)$

$$P(X=x) = f(x) = \binom{5}{x}.\theta^x.(1-\theta)^{5-x}$$

We know that Binomial distribution is number of success obtained from n independent Bernoulli trials and that if $X \sim Bernoulli(\theta)$ then $\hat{\theta}_{MLE} = \frac{\sum X_i}{n} = \bar{X}$

$$\text{So; } \hat{\theta}_{MLE} = \bar{X} = \frac{\sum x_i f_i}{\sum f_i} / n = \frac{0.6 + 1.10 + \dots + 5.1}{6 + 10 + \dots + 1} / n = \frac{2{,}12}{5} = 0{,}424$$

$$\widehat{P(X \geq 3)} = \sum_{x=3}^{5} \binom{5}{x}.0{,}424^x.(1-0{,}424)^{5-x} = \underline{0{,}3597}$$

6.8. If a random sample of size n is taken from a distribution having p.d.f. $f(x;\theta) = 2x/\theta^2$, $0 < x \le \theta$, zero elsewhere, find:

(a) The m.l.e. $\hat{\theta}$ for θ.

(b) The constant c so that $E(c\hat{\theta}) = \theta$.

(c) The m.l.e. for the median of the distribution.

6.9. Let $X_1, X_2, \ldots, X_n$ be i.i.d., each with a distribution with p.d.f. $f(x;\theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty$, zero elsewhere. Find the m.l.e. of $\Pr(X \le 2)$.

6.8) a) $L(\theta) = \frac{2^n}{\theta^{2n}} \cdot \prod x_i \quad 0 < x_i \le \theta$

$L(\theta)$ is maximized when θ is minimum. So; $\boxed{\hat{\theta}_{MLE} = X_{(n)}}$

b) $f(x) = \frac{2x}{\theta^2} \quad 0 < x \le \theta$

$F(x) = \int_0^x f(w)dw = \int_0^x \frac{2w}{\theta^2} \cdot dw = \frac{2}{\theta^2} \frac{w^2}{2}\Big]_0^x = \frac{x^2}{\theta^2} \quad 0 < x \le \theta$

$g_n(y) = n \cdot [F(y)]^{n-1} \cdot f(y) = n \cdot \left(\frac{y^2}{\theta^2}\right)^{n-1} \cdot \frac{2y}{\theta^2} = \frac{2n}{\theta^{2n}} \cdot y^{2n-1} \quad 0 < y \le \theta$

$E(Y) = \int_0^\theta y \cdot g_n(y)dy = \int_0^\theta y \cdot \frac{2n}{\theta^{2n}} \cdot y^{2n-1} = \frac{2n}{\theta^{2n}} \cdot \frac{y^{2n+1}}{2n+1}\Big]_0^\theta = \frac{2n}{2n+1} \cdot \theta$

$E(\hat{\theta}) = E(X_{(n)}) = \frac{2n}{2n+1} \cdot \theta$

So; $E\left(\frac{2n+1}{2n} \cdot X_{(n)}\right) = \theta \Rightarrow c = \frac{2n+1}{2n}$

c) Median $= m \Rightarrow P(X \le m) = F(m) = \frac{1}{2}$

$F(m) = \frac{m^2}{\theta^2} = \frac{1}{2} \Rightarrow m^2 = \frac{\theta^2}{2} \Rightarrow m = \frac{\theta\sqrt{2}}{2}$

$\hat{m}_{MLE} = \frac{X_{(n)}\sqrt{2}}{2}$

6.9) $X \sim Exponential(\theta) \Rightarrow \hat{\theta}_{MLE} = \bar{X}$ as we have shown.

$f(x) = \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}}$ and $F(x) = 1 - e^{-\frac{x}{\theta}}$

$P(X \leq 2) = F(2) = 1 - e^{-\frac{2}{\theta}}$

$\widehat{F(2)}_{MLE} = 1 - e^{-\frac{2}{\bar{x}}}$ by invariance property.

METHOD of MOMENT Estimation.

Remember, given a distribution of a random variable X, the k^{th} moment is defined as $E(X^k)$

The sum $M_k = \sum_{i=1}^{n} \frac{\sum X_i^k}{n}$ is the k^{th} moment of the sample, $k = 1, 2, \ldots$

The method of moments can be described as follows:

Equate $E(X^k)$ to M_k, beginning with $k=1$ and continuing until there are enough equations to provide unique solutions for $\theta_1, \theta_2, \ldots, \theta_k$.

This can be done in an equivalent manner by equating $\mu = E(X)$ to $\bar{X}$ and $E[(X-\mu)^k]$ to $\frac{\sum(X_i - \bar{X})^k}{n}$ $k = 2, 3, \ldots$ and so on until unique solutions for $\theta_1, \theta_2, \ldots, \theta_k$ are obtained.

In most practical cases, $\hat{\theta}_{i(MME)}$ is a consistent estimator of θ_i, $i = 1, 2, \ldots, r$

6.7. For each of the distributions in Exercise 6.1, find an estimator of θ by the method of moments and show that it is consistent.

a) $X \sim \text{Poisson}(\theta)$; $E(X) = \theta$, $Var(X) = \theta$

$E(X) = M_1$

$\boxed{\hat{\theta}_{MME} = \bar{X}}$

b) $X \sim \text{Beta}(\alpha = \theta ; \beta = 1$; $E(X) = \dfrac{\alpha}{\alpha+\beta} = \dfrac{\theta}{\theta+1}$

$E(X) = M_1$

$\dfrac{\hat{\theta}}{\hat{\theta}+1} = \bar{X}$

$\hat{\theta} = \bar{X}\hat{\theta} + \bar{X}$

$\hat{\theta} - \bar{X}\hat{\theta} = \bar{X}$

$\boxed{\hat{\theta}_{MME} = \dfrac{\bar{X}}{1-\bar{X}}}$

c) $X \sim \text{Exponential}(\theta)$; $E(X) = \theta$

$E(X) = M_1$

$\boxed{\hat{\theta}_{MME} = \bar{X}}$

d) $X \sim \text{Double Exponential}(\mu = \theta ; \beta = 1)$; $E(X) = \theta$

$E(X) = M_1$

$\boxed{\hat{\theta}_{MME} = \bar{X}}$

e) $X \sim \text{Two Parameter Exponential}(\theta_1 = \theta ; \theta_2 = 1)$

$E(X) = \theta_1 + \theta_2 = \theta + 1 = M_1$

$1 + \hat{\theta} = \bar{X}$

$\boxed{\hat{\theta}_{MME} = \bar{X} - 1}$

Example: $X \sim Gamma(\theta_1; \theta_2) \Rightarrow$ Find MME of θ_1, θ_2

Answer: $E(X) = \theta_1 \cdot \theta_2$ and $Var(X) = \theta_1 \cdot \theta_2^2$

$$\left.\begin{array}{l} ① \ \theta_1 \cdot \theta_2 = \bar{X} \\ ② \ \theta_1 \cdot \theta_2^2 = S^2 \end{array}\right\} \Rightarrow \frac{②}{①} \Rightarrow \hat{\theta}_{2(MME)} = \frac{S^2}{\bar{X}}$$

$$① \Rightarrow \hat{\theta}_{1(MME)} = \frac{\bar{X}^2}{S^2}$$

Example: Note that, likewise, $\bar{X}$ and S^2 are MME estimators of $Normal(\theta_1, \theta_2)$

Example: Remember, $X \sim$ Double-Exponential $(\mu; \beta)$

$$E(X) = \mu \ ; \ Var(X) = 2\beta^2$$

Then, if $X \sim$ Double Exponential $(1; \theta)$, find MME of θ.

Answer: We have; $\mu = \bar{X} = 1$

$$Var(X) = 2\theta^2 = S^2$$

$$\theta^2 = \frac{S^2}{2} \Rightarrow \hat{\theta}_{MME} = \frac{S}{\sqrt{2}}$$

But note that, here, $S^2 = \frac{\sum (X_i - 1)^2}{n}$ since we know μ.

Same result can be reached as follows:

$$M_1 = \bar{X} = 1$$

$$M_2 = \frac{\sum X^2}{n} = E(X^2) = Var(X) + E^2(X) = 2\theta^2 + 1$$

$$\frac{\sum X^2}{n} = 2\theta^2 + 1$$

$$2\theta^2 = \frac{\sum X^2}{n} - 1 = \frac{\sum X^2 - n}{n} = \frac{\sum (X^2 - 1)}{n} \Rightarrow \hat{\theta}_{MME} = \sqrt{\frac{\sum (X^2 - 1)}{2n}}$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Example: let, X has an mgf $M_X(t) = \left(\frac{1}{1-\theta t}\right)^{\beta}$.

Find MME of θ and β.

Answer: $M_X'(t) = \beta.\theta.\left(\frac{1}{1-\theta t}\right)^{\beta+1} \Rightarrow E(X) = M_X'(0) = \beta\theta$

$M_X''(t) = \beta.(\beta+1).\theta^2.\left(\frac{1}{1-\theta t}\right)^{\beta+2} \Rightarrow E(X^2) = M_X''(0) = \beta.(\beta+1)\theta^2$

Then, ① $M_1 = \bar{X} = \beta\theta$

② $M_2 = \frac{\sum x^2}{n} = \beta(\beta+1)\theta^2$

$\frac{②}{①} \Rightarrow \frac{\sum X^2}{n\bar{X}} = \theta.(\beta+1) = \underbrace{\theta\beta}_{=M_1=\bar{X}}+\theta$

$$\hat{\theta}_{MME} = \frac{\sum X^2}{n\bar{X}} - \bar{X} = \frac{\sum X^2 - n\bar{X}^2}{n\bar{X}}$$

$$\hat{\theta}_{MME} = \frac{\frac{\sum X^2}{n} - \bar{X}^2}{\bar{X}} = \frac{S^2}{\bar{X}}$$

① $\Rightarrow \bar{X} = \hat{\beta}.\frac{S^2}{\bar{X}}$

$$\hat{\beta} = \frac{\bar{X}^2}{S^2}$$

Note that mgf is that of Gamma distribution.

Mean Squared Error (MSE)

Mean Squared Error of an estimator $\hat{\theta}$ is;

$$MSE_{\theta}(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = Var(\hat{\theta}) + [Bias_{\theta}(\hat{\theta})]^2$$

where $Bias_{\theta}(\hat{\theta}) = E(\hat{\theta}) - \theta$ and $\boxed{\text{Min. MSE} \Rightarrow \text{Better Estimator.}}$

Note that;

$$E[(\hat{\theta}-\theta)^2] = E[(\underbrace{\hat{\theta}-E(\hat{\theta})}_{=a} + \underbrace{E(\hat{\theta})-\theta}_{=b})^2]$$

$$= \underbrace{E[(\hat{\theta}-E(\hat{\theta}))^2]}_{Var(\hat{\theta})} + 2\underbrace{E[(\hat{\theta}-E(\hat{\theta})).(E(\hat{\theta})-\theta)]}_{=0} + \underbrace{E[E(\hat{\theta}-\theta)^2]}_{[Bias(\hat{\theta})]^2}$$

$$= Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Example: $X \sim Normal(\mu; \sigma^2)$. Remember, we have two estimators for σ^2, which are S^2 and S^{*2}. Which estimator is better according to MSE criterion?

Answer:

(i) $S^{*2} = \frac{1}{n-1}\sum(X_i-\bar{X})^2$ and $E(S^{*2}) = \sigma^2$; $Var(S^{*2}) = \frac{2\sigma^4}{n-1}$

(ii) $S^2 = \frac{1}{n}\sum(X_i-\bar{X})^2$ and $E(S^2) = \frac{n-1}{n}\sigma^2$; $Var(S^2) = \frac{2(n-1)\sigma^4}{n^2}$

(i) $\underset{\sigma^2}{MSE}(S^{*2}) = Var(S^{*2}) + \underbrace{\left[\underset{\sigma^2}{Bias}(S^{*2})\right]^2}_{=0 \text{ (unbiased)}} = \frac{2\sigma^4}{n-1}$

(ii) $\underset{\sigma^2}{MSE}(S^2) = Var(S^2) + \left[\underset{\sigma^2}{Bias}(S^2)\right]^2 = \frac{2(n-1)\sigma^4}{n^2} + \left[\frac{n-1}{n}\cdot\sigma^2 - \sigma^2\right]^2$

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

Since $\underset{\sigma^2}{MSE}(S^2) \leq \underset{\sigma^2}{MSE}(S^{*2})$, S^2 is a better estimator of σ^2.

Example: $X \sim Uniform(0; \theta)$

Let $\hat{\theta}_1 = X_{(n)}$: MLE of θ and $\hat{\theta}_2 = 2\bar{X}$: MME of θ. Which?

Answer: $\hat{\theta}_1$: $E(\hat{\theta}_1) = \frac{n}{n+1}\cdot\theta$; $Var(\hat{\theta}_1) = \frac{n}{(n+2)(n+1)^2}\cdot\theta^2$

$MSE(\hat{\theta}_1) = \frac{n}{(n+2)(n+1)^2}\cdot\theta^2 + \left(\frac{n}{n+1}\cdot\theta - \theta\right)^2 = \frac{2}{(n+2)(n+1)}\cdot\theta^2$ ✓ BETTER!

$\hat{\theta}_2$: $E(\hat{\theta}_2) = E(2\bar{X}) = 2E(\bar{X}) = 2\cdot\frac{\theta}{2} = \theta$

$Var(\hat{\theta}_2) = Var(2\bar{X}) = 4Var(\bar{X}) = 4\frac{Var(X)}{n} = \frac{4}{n}\cdot\frac{(\theta-0)^2}{12} = \frac{\theta^2}{3n}$

$MSE(\hat{\theta}_2) = Var(\hat{\theta}_2) + \underbrace{\left[Bias(\hat{\theta}_2)\right]^2}_{=0} = \frac{\theta^2}{3n}$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com