Chapter 7. Point Estimation

<u>Definition 7.1.1</u>. A point estimator is any function $W(X_1, \dots, X_n)$ of a sample. Any statistic is a point estimator.

If $(x_1, \dots, x_n)$ are realized values of $X_1, \dots, X_n$, $W(x_1, \dots, x_n)$ is called an estimate.

Moment estimator

$X_1, \dots, X_n$ random sample.

Let
$$\mu_1(\theta_1, \dots, \theta_k) = E[X_i], \quad i = 1, 2, \dots, n$$
$$\mu_2(\theta_1, \dots, \theta_k) = E[X_i^2], \quad i = 1, 2, \dots, n$$
$$\vdots \qquad\qquad \vdots$$
$$\mu_k(\theta_1, \dots, \theta_k) = E[X_i^k], \quad i = 1, 2, \dots, n$$

The moment estimators are found by solving the equations.
$$\mu_1(\hat{\theta}_1, \dots, \hat{\theta}_k) = \frac{1}{n} \sum_{i=1}^{n} X_i$$
$$\mu_2(\hat{\theta}_1, \dots, \hat{\theta}_k) = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$
$$\vdots$$
$$\mu_k(\hat{\theta}_1, \dots, \hat{\theta}_k) = \frac{1}{n} \sum_{i=1}^{n} X_i^k$$

<u>Example</u>. $X_1, \dots, X_n \sim iid \sim N(\mu, \sigma^2)$

$$\hat{\mu} = \bar{X}$$
$$\hat{\sigma}^2 + \hat{\mu}^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

7.2.2. Maximum Likelihood Estimators.

Let $X_1, \dots, X_n$ be iid from a population with pdf/pmf $f(x \mid \theta_1, \dots, \theta_k)$. The likelihood function is defined by: $L(\underline{\theta} \mid \underline{x}) = L(\theta_1, \dots, \theta_k \mid x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i \mid \theta_1, \dots, \theta_k)$

Definition 7.2.4

For each sample point $\underline{x}$, let $\hat{\underline{\theta}}(\underline{x})$ be a parameter value at which $L(\underline{\theta} \mid \underline{x})$ attains its maximum as a function of $\underline{\theta}$. The maximum likelihood estimator for $\underline{\theta}$ is then $\hat{\underline{\theta}}(\underline{x})$. Normally we maximize $\log L(\underline{\theta} \mid \underline{x})$

Example 7.2.5

$X_1, \dots, X_n$ iid $N(\theta, 1) \Rightarrow L(\theta \mid \underline{x}) = \prod_{i=1}^{n} (2\pi)^{-\frac{1}{2}} \cdot e^{-\frac{\sum_{i=1}^{n}(x_i - \theta)^2}{2}}$

$\Rightarrow \log L(\theta \mid \underline{x}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2$

We need to minimize $\sum_{i=1}^{n}(x_i - \theta)^2$

But $\sum_{i=1}^{n}(x_i - \theta) \geq \sum_{i=1}^{n}(x_i - \bar{x})^2$ with equality if and only if $\theta = \bar{x}$. Therefore $\hat{\theta}(\underline{x}) = \bar{x}$

Example 7.2.9

$X_1, \dots, X_n$ iid $B(k, p)$, k unknown, p known

$$L(k \mid \underline{x}, p) = \prod_{i=1}^{n} \binom{k}{x_i} p^{x_i} (1-p)^{k - x_i}$$

$k \geq \max_i x_i$ and must satisfy

$$\frac{L(k|\underline{x},p)}{L(k-1|\underline{x},p)} = \frac{\left[\prod_{i=1}^{m}\binom{k}{x_i}\right]\cdot p^{\sum x_i}(1-p)^{mk-\sum x_i}}{\left[\prod_{i=1}^{m}\binom{k-1}{x_i}\right]\cdot p^{\sum x_i}\cdot(1-p)^{m(k-1)-\sum x_i}} = \frac{k^m(1-p)^m}{\prod_{i=1}^{m}(k-x_i)} \geq 1$$

and $$\frac{L(k+1|\underline{x},p)}{L(k|\underline{x},p)} = \frac{(k+1)^m(1-p)^m}{\prod_{i=1}^{m}(k+1-x_i)} \leq 1$$

which gives $\prod_{i=1}^{m}\left(1-\frac{x_i}{k}\right) \leq (1-p)^m \leq \prod_{i=1}^{m}\left(1-\frac{x_i}{k+1}\right)$ which gives a unique $k \geq \max_i x_i$

Maximum likelihood and the Invariance principle.

$\hat{\theta}$ is the MLE of θ. What about $\tau(\theta)$?

Suppose $\eta = \tau(\theta)$ is one to one $\Rightarrow \theta = \tau^{-1}(\eta)$

For example

$$f(x|\theta) = \frac{1}{\theta} e^{-\frac{1}{\theta}x}, \quad x>0, \; \theta>0$$

Let $\eta = \tau(\theta) = \frac{1}{\theta} \Rightarrow \theta = \tau^{-1}(\eta) = \frac{1}{\eta}$

$$f(x|\eta) = f(x|\tau^{-1}(\eta)) = \eta e^{-\eta x}, \quad x>0, \; \eta>0.$$

In general we have $L(\theta|\underline{x})$ and $L'(\tau^{-1}(\eta)|\underline{x})$ with respective maximum in $\hat{\theta}$ and $\hat{\eta}$.

~~Further let $\hat{\eta}$ and $\hat{\theta}$ be the respective~~

Then we have: $L'(\hat{\eta}|\underline{x}) \geq L'(\eta|\underline{x}), \; \forall \eta$

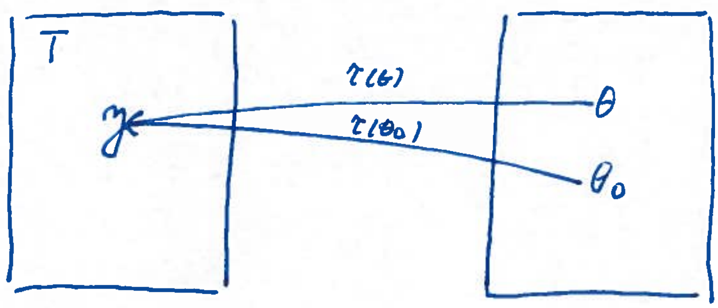
$$\Leftrightarrow f(\underline{x}|\tau^{-1}(\hat{\eta})) \geq f(\underline{x}|\tau^{-1}(\eta)), \; \forall \eta$$

$$\Leftrightarrow f(\underline{x}|\hat{\theta}) \geq f(\underline{x}|\theta), \; \forall \theta$$

$$\Leftrightarrow L(\hat{\theta}|\underline{x}) \geq L(\theta|\underline{x}), \; \forall \theta.$$

which means $\hat{\eta} = \tau(\hat{\theta})$

Suppose $\tau(\theta)$ is not one to one



$$L(\theta|\underline{x}) \overset{?}{=} L(\theta_0|\underline{x})$$

We define the induced likelihood L^* given by

$$L^*(\eta|\underline{x}) = \sup_{\{\theta:\ \tau(\theta)=\eta\}} L(\theta|\underline{x})$$

The value $\hat{\eta}$ that maximizes $L^*(\eta|\underline{x})$ is called the MLE of $\eta = \tau(\theta)$. Let $\hat{\theta}$ be the MLE of θ.

Then $$L^*(\hat{\eta}|\underline{x}) = L(\hat{\theta}|\underline{x}) = \sup_{\{\theta:\ \tau(\theta)=\tau(\hat{\theta})\}} L(\theta|x) = L^*[\tau(\hat{\theta})|\underline{x}]$$

~~Thereby~~ <u>Theorem 7.2.10</u>

Invariance properties of MLE's. If $\hat{\theta}$ is the MLE of θ, then for any function $\tau(\theta)$, the MLE of $\tau(\theta)$ is $\tau(\hat{\theta})$

<u>Consequences</u>

MLE of μ^2: $\hat{\mu}^2 = \bar{x}^2$.

MLE of $\sqrt{p(1-p)}$: $\sqrt{\hat{p}(1-\hat{p})}$

Also if $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ is the MLE of $(\theta_1, \ldots, \theta_k)$, the MLE of $\tau(\theta_1, \ldots, \theta_k) = \tau(\hat{\theta}_1, \ldots, \hat{\theta}_k)$

Example

$X_1, \dots, X_n$ iid $N(\mu, \sigma^2)$

$$\Rightarrow L(\mu, \sigma^2 | \underline{x}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}$$

and $\log L(\mu, \sigma^2 | \underline{x}) = \ell(\mu, \sigma^2 | x) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$

$$\frac{\partial}{\partial\mu}\ell(\mu, \sigma^2 | \underline{x}) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0 \qquad (1)$$

$$\frac{\partial}{\partial\sigma^2}\ell(\mu, \sigma^2 | \underline{x}) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \qquad (2)$$

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

Maximum

$\forall \sigma^2$ we have $\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{\sigma^2}} \geq \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}$

$\Rightarrow \bar{x}$ is a MLE for μ.

Now $\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{\sigma^2}} \xrightarrow{\eta = \sigma^2 \Rightarrow \sigma = \sqrt{\eta}} \frac{1}{(2\pi\eta)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{\eta}}$

$$\ell^*(\eta | \underline{x}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\eta - \frac{1}{2\eta}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\frac{\partial\ell^*(\eta | x)}{\partial\eta} = \frac{-n}{2\eta} + \frac{1}{2\eta^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 = 0 \Rightarrow \hat{\eta} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

$$\frac{\partial^2\ell^*(\eta | x)}{\partial\eta^2} = \frac{n}{2\eta^2} - \frac{1}{\eta^3}\sum_{i=1}^{n}(x_i-\bar{x})^2 \overset{\eta = \hat{\eta}}{=} \frac{n}{2\hat{\eta}^2} - \frac{n\hat{\eta}}{\hat{\eta}^3} = -\frac{n}{2\hat{\eta}^2} < 0$$

$\Rightarrow$ maximum.