Point Estimation

For a population with distribution $P_{\underline{\theta}}$, $\underline{\theta} \in \mathbb{R}^k$, we take a random sample $X_1, \ldots, X_n$. We are interested in estimating the parametric function $g(\underline{\theta})$.

An estimator $T(\underline{X})$ is simply a function of random sample.

<u>Criteria for Good Estimators</u>

1. <u>Unbiasedness</u> : An estimator $T(\underline{X})$ of $g(\underline{\theta})$ is said to be unbiased if $E\,T(\underline{X}) = g(\underline{\theta})$ for all $\underline{\theta}$.

Examples: 1. Let $X_1, \ldots X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$

$E(\bar{X}) = \mu$. So $\bar{X}$ is unbiased estimator of μ.

$$S^2 = \boxed{\frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2} \to \text{sample variance}$$

$$W = \frac{(n-1) S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$E\left(\frac{\cancel{(n-1)}S^2}{\sigma^2}\right) = \cancel{(n-1)}$$

$$\Rightarrow \quad E(S^2) = \sigma^2$$

So S^2 is unbiased estimator of σ^2.

$$\eta = \frac{\mu}{\sigma} \rightarrow \text{standardized mean}$$

We want an unbiased estimator

of η.

We note that $\bar{X}$ & W are independently distributed.

$$E\frac{1}{(W^{1/2})} = \int_0^\infty \frac{1}{w^{1/2}} \cdot \frac{1}{2^{\frac{n-1}{2}}\Gamma\frac{n-1}{2}} e^{-w/2} w^{\frac{n-1}{2}-1} dw$$

$$= \frac{1}{2^{\frac{n-1}{2}}\Gamma\frac{n-1}{2}} \int_0^\infty e^{-w/2} w^{\frac{n}{2}-2} dw$$

$$= \frac{2^{\frac{n}{2}-1}}{2^{\frac{n-1}{2}}\,\Gamma\frac{n-1}{2}} \quad \frac{\Gamma\frac{n}{2}-1}{} = \frac{\Gamma\left(\frac{n-2}{2}\right)}{\sqrt{2}\,\Gamma\frac{n-1}{2}} \quad k$$

$$E\left(\frac{\sigma}{\sqrt{n-1}\,S}\right) = k \Rightarrow E\left(\frac{1}{k\sqrt{n-1}\,S}\right) = \frac{1}{\sigma}$$

So $T = \dfrac{\bar{X}}{k\sqrt{n-1}\,S}$ is unbiased estimator of μ/σ.

2. $X \sim Bin(n, p)$, n is known

$$E(X) = np \Rightarrow E\left(\frac{X}{n}\right) = p$$

So sample proportion $\frac{X}{n}$ is an unbiased estimator of pop^n proportion p.

3. $X_1 \ldots X_n \sim P(\lambda)$, $\lambda > 0$

$$E(\bar{X}) = \lambda$$

$\bar{X}$ is unbiased for λ.

4. $X_1 \dots X_n \overset{i.i.d.}{\sim} \lambda e^{-\lambda x}$, $\lambda > 0$, $x > 0$

$$E(X_1) = \frac{1}{\lambda} \qquad E(\bar{X}) = \left(\frac{1}{\lambda}\right)$$

So $\bar{X}$ is unbiased for $1/\lambda$.

$$Y = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \lambda)$$

$$E\left(\frac{1}{Y}\right) = \int_0^\infty \frac{1}{y} \cdot \frac{\lambda^n}{\Gamma n} e^{-\lambda y} y^{n-1} \, dy$$

$$= \frac{\lambda^n}{\Gamma n} \cdot \int_0^\infty e^{-\lambda y} y^{n-2} \, dy$$

$$= \frac{\lambda^n}{\Gamma n} \cdot \frac{\Gamma n-1}{\lambda^{n-1}} = \frac{\lambda}{(n-1)}$$

$$E\left(\frac{n-1}{Y}\right) = \lambda \Rightarrow E\left(\frac{n-1}{n\overline{X}}\right) = \lambda$$

So $\frac{(n-1)}{n\overline{X}}$ is unbiased estimator of λ.

<u>Consistency</u>: $T_n = T(X_1, \ldots X_n)$ is said to be consistent estimator of $g(\underline{\theta})$ if for any $k > 0$

$$P\left(|T_n - g(\underline{\theta})| > k\right) \longrightarrow 0$$

as $n \to \infty$

or

$$P\left(|T_n - g(\theta)| \leq k\right) \longrightarrow 1$$

as $n \to \infty$

$$T_n \xrightarrow{P} g(\theta)$$

<u>**Example**</u>: Let $X_1, X_2 \ldots$ be a sequence of i.i.d. random

variables with mean μ and variance σ^2. For estimating μ let us consider $T_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

$$E(\bar{X}) = \mu, \quad V(\bar{X}) = \frac{\sigma^2}{n}$$

Using Chebyshev's inequality,

$$P\left(|\bar{X} - \mu| \geq k \right) \leq \frac{\sigma^2}{nk^2} \longrightarrow 0$$

as $n \to \infty$

So the sample mean is a consistent estimator of the population mean (if variance exists)

2. Let $X_1 \ldots \ldots X_n \sim U(0, \theta)$

$T_1 = 2\bar{X}$, $\quad E(\bar{X}) = \frac{\theta}{2}$

$E(T_1) = \theta \quad 2\bar{X}$ is unbiased

T_1 is also consistent for θ.

$T_2 = \max(X_1, \ldots, X_n) = X_{(n)}$

$$f_{X_{(n)}}(x) = \frac{n x^{n-1}}{\theta^n}, \quad 0 < x < \theta$$

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

$T_3 = \left(\frac{n+1}{n}\right) X_{(n)}$ is unbiased for θ

$$P(|X_{(n)} - \theta| > k)$$

$$= P(\theta - X_{(n)} > k)$$

$$= P(X_{(n)} < \theta - k) = \left(\frac{\theta - k}{\theta}\right)^n$$

$$\longrightarrow 0 \text{ as } n \to \infty$$

So $T_2 = X_{(n)}$ is consistent but biased estimator of θ

$T_3 = \left(\frac{n+1}{n}\right) X_{(n)}$ is unbiased and consistent for θ

If T_n is consistent for $g(\underline{\theta})$

then $a_n T_n + b_n$ is consistent for $g(\underline{\theta})$ if

$a_n \to 1$ & $b_n \to 0$ as $n \to \infty$.

Mean Squared Error Criteria

We define M.S.E. of an estimator $T(\underline{X})$ for $g(\underline{\theta})$ as

$$MSE(T) = E\left(T - g(\underline{\theta})\right)^2$$

We say that an estimator T_1 is better than T_2 if

$$MSE(T_1) \leq MSE(T_2) \ \forall \ \underline{\theta}$$

with strict inequality for at least some $\underline{\theta}$.

Let us consider $U(0, \theta)$ example.

$$T_1 = 2\bar{X}, \quad MSE(T_1) = E(2\bar{X} - \theta)^2$$

$$= V(2\bar{X})$$

$$= 4V(\bar{X}) = \frac{4\theta^2}{12n}$$

$$= \frac{\theta^2}{3n}.$$

$$T_2 = X_{(n)}, \quad MSE(T_2) = E(X_{(n)} - \theta)^2$$

$$= EX_{(n)}^2 - 2\theta E(X_{(n)}) + \theta^2$$

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

$$E X_{(n)} = \int_0^\theta \frac{n \; x^n}{\theta^n} dx = \frac{n}{n+1} \theta$$

$$E\left(X_{(n)}^2\right) = \int_0^\theta \frac{n \, x^{n+1}}{\theta^n} \, dx = \frac{n}{n+2} \cdot \theta^2$$

$$MSE(T_1) - MSE(T_2) = \theta^2 \left(\frac{1}{3n} - \frac{1}{(n+1)(n+2)} \right)$$

$$= \frac{(n^2+2)}{3n(n+1)(n+2)} \theta^2 \quad > 0$$

T_2 is better than T_1

<u>Ex.</u> find $MSE(T_3)$ and compare with $MSE(T_2)$ & $MSE(T_1)$