1 Asymptotic Properties of Estimators (“large” samples)

Let $X \sim F(x; \theta)$ where $\theta$ is an unknown parameter, $\theta \subset \Omega$.

Let $X_1, \ldots, X_n$ be a random sample from $X$, i.e., $X_1, \ldots, X_n$ are i.i.d. random variables with the CDF $F((x; \theta))$.

Let $\hat{\Theta}_n = u_n(X_1, \ldots, X_n)$ be an estimator of $\theta$ based on this random sample.

1.1 Asymptotically Unbiased Estimators

A sequence of estimators $\hat{\Theta}_n$ of $\theta$ is said to be asymptotically unbiased if $\lim_{n \to \infty} E(\hat{\Theta}_n) = \theta$ (or, in other words, if the bias $b(\hat{\Theta}_n) = E(\hat{\Theta}_n) - \theta \to 0$ as $n \to \infty$). Of course, if $\hat{\Theta}_n$ is unbiased for each $n = 1, 2, \ldots$, then it is asymptotically unbiased.

Example 1.1. Consider the empirical variance $V_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$. We know that $V_n$ is a biased estimator of the variance $\sigma^2$:

$$E(V_n) = \frac{n - 1}{n} \sigma^2.$$  

But $\frac{n-1}{n} = 1 - \frac{1}{n} \to 1$. Hence $E(V_n) \to \sigma^2$, and $V_n$ is asymptotically unbiased.

1.2 Consistent, Mean-Squared Consistent and Strongly Consistent Estimators

Definition 1.1. We say that the sequence of estimators $\{\hat{\Theta}_n\}$ of the parameter $\theta$ is consistent if for each $\varepsilon > 0$

$$\lim P\{|\theta_n - \theta| < \varepsilon\} = 1$$

for each $\theta \in \Omega$.

Definition 1.2. A sequence of estimators $\hat{\Theta}_n$ of $\theta$ is mean square error consistent (or simply mean consistent) if $\lim_{n \to \infty} E\left[ (\hat{\Theta}_n - \theta)^2 \right] = 0$ for each $\theta \in \Omega$. 
**Theorem 1.1.** \( \hat{\Theta}_n \) is mean consistent iff it is asymptotically unbiased and \( \lim_{n \to \infty} \text{Var}(\hat{\Theta}_n) = 0 \).

**Proof.** We know that

\[
E \left[ (\hat{\Theta}_n - \theta)^2 \right] = \text{Var}(\hat{\Theta}_n) + \left[ E(\hat{\Theta}_n) - \theta \right]^2.
\]

It is clear that if \( \text{Var}(\hat{\Theta}_n) \to 0 \) and \( E(\hat{\Theta}_n) \to 0 \) then \( E \left[ (\hat{\Theta}_n - \theta)^2 \right] \to 0 \). We have

\[
0 \leq \text{Var}(\hat{\Theta}_n) \leq E \left[ (\hat{\Theta}_n - \theta)^2 \right] \quad \text{and} \quad 0 \leq \left[ E(\hat{\Theta}_n) - \theta \right]^2 \leq E \left[ (\hat{\Theta}_n - \theta)^2 \right].
\]

Hence Mean Consistency implies \( E(\hat{\Theta}_n) \to \theta \) and \( \text{Var}(\hat{\Theta}_n) \to 0 \).

\[\square\]

**Definition 1.3.** A sequence of estimators \( \hat{\Theta}_n \) of \( \theta \) is said to be strongly consistent if \( \lim_{n \to \infty} \hat{\Theta}_n = \theta \) with probability 1.

**Theorem 1.2.** Mean consistence implies consistence.

**Proof.** The statement follows from the Chebyshev inequality

\[
P \left\{ \left| \hat{\Theta}_n - \theta \right| > \varepsilon \right\} \leq \frac{1}{\varepsilon^2} \text{Var}(\hat{\Theta}_n).
\]

If \( \text{Var}(\hat{\Theta}_n) \to 0 \) then \( P \left\{ \left| \hat{\Theta}_n - \theta \right| > \varepsilon \right\} \to 0 \) and

\[
P \left\{ \left| \hat{\Theta}_n - \theta \right| < \varepsilon \right\} = 1 - P \left\{ \left| \hat{\Theta}_n - \theta \right| > \varepsilon \right\} \to 1
\]

\[\square\]
Theorem 1.3. Strong consistency implies consistence.

Example 1.2. $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, $n = 1, 2, \ldots$ is a strongly and mean consistent sequence of estimators of $\mu$. This follows from the Strong and Mean Laws of Large Numbers.

Example 1.3. Let $E|X|^r < \infty$. Let us consider the empirical moments $M_n^{(r)} = \frac{1}{n} \sum_{i=1}^{n} X_i^r$ as estimators of the moment of order $r$, $\mu^{(r)} = E(X^r)$. $\{M_n^{(r)}\}$ is strongly consistent. Indeed, $E(X^r) = \mu^r$, so it remains to apply the SLLN with $X_i^r$ instead of $X_i$.

Example 1.4. Assume $E(X^2) < \infty$. The empirical variance

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

By Example 1.2, $\bar{X}_n^2 \rightarrow \mu^2$ with probability 1, by Example 1.3 (with $r = 2$) $\frac{1}{n} \sum_{i=1}^{n} X_i^2 \rightarrow E(X^2)$ with probability 1. Hence, with probability 1 $\lim_{n \to \infty} V_n = E(X^2) = \mu^2 = Var(X)$. Hence the sequence of estimators $\{V_n\}$ is strongly consistent for $Var(X) = \sigma^2$. $S_n^2 = \frac{n}{n-1} V_n$. Hence $\lim_{n \to \infty} S_n^2 = \lim_{n \to \infty} \frac{n}{n-1} V_n = Var(X)$ since $\lim_{n \to \infty} \frac{n}{n-1} = \lim_{n \to \infty} \frac{1}{1 - \frac{1}{n}} = 1$. Thus $\{S_n^2\}$ is also strongly consistent for $Var(X)$.

$$S_n = \sqrt{S_n^2} \rightarrow \sqrt{Var(X)} = \sigma$$ with probability 1.

This shows that $\{S_n\}$ is strongly consistent for $\sigma$.

1.3 Asymptotically Normal Estimators

Definition 1.4. Let $Z_n = \frac{\hat{\Theta}_n - \theta}{\sqrt{\gamma_n}}$. We say that $\{\hat{\Theta}_n\}$ is asymptotically normal with mean $\theta$ and asymptotic variance $\gamma_n^2$ (notation: $(AN(\theta, \gamma_n^2))$) if for each $x$ the CDFs of $Z_n$, $F_{Z_n}(x) \xrightarrow{n \to \infty} \Phi(x)$ (the standard normal CDF).

Example 1.5. Consider $\bar{X}_n$ as an estimate of $\mu = E(X)$.

We have $Z_n = \frac{\bar{X}_n - \mu}{\sqrt{\gamma_n}} = \frac{\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)}{\sigma / \sqrt{n}} = \frac{\sum_{i=1}^{n} (X_i - \mu)}{\sigma \sqrt{n}}$ where $\sigma^2 = Var(X)$.

By the CLT $F_{Z_n}(x) \rightarrow \Phi(x)$.

Example 1.6. Let $A$ be a random event with $P(A) = p$. Assume that $n$ Bernoulli trials have been performed and $\hat{P}$ is the relative frequency of occurrences of $A$ in these trials. Let

$$Z_n = \frac{\hat{P}_n - p}{\sqrt{p(1-p)}}.$$ 

Then by the De Moivre-Laplace Theorem $F_{Z_n}(x) \rightarrow \Phi(X)$.

We’ll have more examples later.
1.4 Strong consistency of the Method of Moments estimators

Let the distribution of $X$ depend on a parameter $\theta$, $\mu = E(X)$. Then $\mu = v(\theta)$; we solve this equation for $\theta$ and obtain $\theta = u(\mu)$. The Method of Moments estimator (MME) is calculated by putting the estimator $\bar{X}_n$ instead of the parameter $\mu$: $\hat{\Theta}_n = u(\bar{X}_n)$.

**Theorem 1.4.** If the function $u(\cdot)$ is continuous then $\hat{\Theta}_n$ is strongly consistent.

**Proof.** $\bar{X}_n \to \mu$ with probability 1. Since $u(\cdot)$ is continuous

$$\hat{\Theta}_n = u(\bar{X}_n) \to u(\mu) = \theta$$

with probability 1. \qed

**Example 1.7.** $f(x, \theta) = \theta x^{\theta - 1}, 0 \leq x \leq 1, (0 < \theta < \infty)$.

$$\mu = \int_0^1 x \cdot \theta x^{\theta - 1} dx = \theta \int_0^1 x \cdot \theta x^{\theta - 1} d\theta =$$

$$= \frac{\theta}{\theta + 1} \text{ (this is our function } \mu = v(\theta)).$$

We solve the equation for $\theta$ and find $\theta \equiv \frac{\mu}{1 - \mu}$. Hence the MME $\hat{\Theta}_n = \frac{\bar{X}_n}{1 - \bar{X}_n}$. Since $\bar{X}_n \to \mu$ with probability 1 we have $\hat{\Theta}_n \to \frac{\mu}{1 - \mu} = \theta$ with probability 1. Thus $\{\hat{\Theta}_n\}$ is strongly consistent.
1.5 Asymptotic Properties of the MLEs

Under some “regularity” conditions the MLEs $\tilde{\Theta}_n$ of a parameter $\theta$ have the following properties:

1. $\{\tilde{\Theta}_n\}$ is consistent $\theta$.

2. $\{\tilde{\Theta}_n\}$ is asymptotically normal.

3. $\{\tilde{\Theta}_n\}$ has asymptotically the least square error, i.e $re(\tilde{\Theta}_n : \hat{\Theta}_n) := \frac{MSE(\hat{\Theta}_n)}{MSE(\tilde{\Theta}_n)} \geq 1$

   for each other estimators $\hat{\Theta}_n$, if $n$ is large enough. We can also say that for each
   other estimators $\hat{\Theta}_n$ we have $MSE(\tilde{\Theta}_n) < MSE(\hat{\Theta}_n)$ if $n$ is large enough,

*Example* 1.8. $\{V_n\}$ has properties (1-3) an MLE of $\sigma^2$. This implies that $\{S_n^2\}$ has these properties too.