Asymptotic Properties of Estimators

- The <u>asymptotic behavior</u> ~~of~~ of an estimator refers to its behavior as the sample size n increases toward ∞.
- Since an estimator $\hat{\theta}$ usually depends on n, we may denote a sequence of such estimators by $\hat{\theta}_n$.

9.3 Consistency

- Intuitively, we would like our estimator to "get closer" to the target parameter as $n \longrightarrow \infty$.
- <u>Definition</u>: An estimator $\hat{\theta}_n$ is a <u>consistent</u> estimator of θ if, for any $\varepsilon > 0$,

Unbiasedness and Consistency

- **Theorem:** If $\hat{\theta}_n$ is unbiased for θ and if $\lim_{n\to\infty} \text{var}(\hat{\theta}_n) = 0$, then $\hat{\theta}_n$ is a consistent estimator of θ.

- **Proof:**

- If $\hat{\theta}_n$ is consistent for θ, we may write $\hat{\theta}_n$ "converges in probability" to θ, or $\hat{\theta}_n \xrightarrow{P} \theta$.

Example 1 (Law of Large Numbers): Show that if $Y_1, \ldots, Y_n$ are iid with $E(Y_i) = \mu$ and $var(Y_i) = \sigma^2 < \infty$, then $\bar{Y}_n$ is consistent for μ.

Theorem: Suppose $\hat{\theta}_n \xrightarrow{p} \theta$ and $\hat{\theta}_n^* \xrightarrow{p} \theta^*$.

Then: (a) $\hat{\theta}_n + \hat{\theta}_n^* \xrightarrow{p} \theta + \theta^*$

(b) $\hat{\theta}_n \hat{\theta}_n^* \xrightarrow{p} \theta \theta^*$

(c) $\hat{\theta}_n / \hat{\theta}_n^* \xrightarrow{p} \theta / \theta^*$ as long as $\theta^* \neq 0$.

(d) If $g(\cdot)$ is continuous at θ, then $g(\hat{\theta}_n) \xrightarrow{p} g(\theta)$.

Example 2: Let $Y_1, \ldots, Y_n$ be iid with $E(Y_i) = \mu$ and $E(Y_i^2) < \infty$, and $E(Y_i^4) < \infty$. Show that the sample variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

is a consistent estimator of $var(Y_i) = \sigma^2$.

Proof: It can be easily shown that

- The following result justifies the use of s^2 in place of σ^2 in the large-sample CIs about μ.
- Slutsky's Theorem (Special case):

If $U_n \xrightarrow{d} N(0,1)$ as $n \to \infty$, and $W_n \xrightarrow{P} 1$, then $\frac{U_n}{W_n} \xrightarrow{d} N(0,1)$.

Example 3: Let $Y_1, \ldots, Y_n$ be iid with $E(Y_i) = \mu$ and $var(Y_i) = \sigma^2 < \infty$. Show that

$$\frac{\overline{Y} - \mu}{S/\sqrt{n}} \xrightarrow{d} N(0,1)$$

Proof:

- A similar argument will show that

- Exercise: If $Y_1, \ldots, Y_n \overset{iid}{\sim} Unif(0, \theta)$, show $Y_{(n)}$ is a consistent estimator of θ. [Hint: Use the definition of consistency directly.]

The cdf of $Y_{(n)}$ is:

For every $\varepsilon > 0$,

So for $0 < \varepsilon \leq \theta$,

9.8 Large-sample properties of MLEs

- If $Y_1, \ldots, Y_n \overset{iid}{\sim} f_Y(y;\theta)$ and $\hat{\theta}$ is the MLE of θ, then, assuming certain regularity conditions:

$$\hat{\theta} \xrightarrow{P} \theta \quad \text{as } n \to \infty$$

and

for large n, $\hat{\theta} \,\dot{\sim}\, N(\theta, \sigma^2_{\hat{\theta}})$ where

- This implies that assuming regularity, MLEs are <u>consistent</u>, <u>asymptotically normal</u>, and <u>asymptotically efficient</u>.

- We can plug in $\hat{\theta}$ in place of θ in the expression for $\sigma^2_{\hat{\theta}}$ and use Slutsky's Theorem to show that:

Example 4: $Y_1, \ldots, Y_n \overset{iid}{\sim} \text{Pois}(\lambda)$. It is easy to show that the MLE of λ is $\hat{\lambda} = \bar{Y}$. Now,

- Note that a large-sample $100(1-\alpha)\%$ CI for λ is thus:

Example: $Y_1, \ldots, Y_{100} \overset{iid}{\sim} \text{Pois}(\lambda)$ with $\bar{Y} = 26.03$:

Delta Method

- If $\hat{\theta}$ is an asymptotically normal MLE of θ, the delta method allows us to derive the asymptotic distribution of $g(\hat{\theta})$, which is the MLE of $g(\theta)$ (any differentiable function of θ).

Theorem (Delta Method): Let $Y_1, \ldots, Y_n$ be iid r.v.'s with pdf $f_Y(y; \theta)$. If $\hat{\theta}$ is the MLE of θ, then (assuming regularity conditions), for large n:

where:

Example: $Y_1, \ldots, Y_n \overset{iid}{\sim}$ Bernoulli (p). Define the "log-odds" as $g(p) = \ln\left(\frac{p}{1-p}\right)$. Find the MLE of the log-odds and its asymptotic distribution.

Proof: We know $\hat{p} = \overline{Y}$ is the MLE for p, so:

- **Exercise 1:** Write a large-sample $(1-\alpha)100\%$ CI for the log-odds.

- **Exercise 2:** If $Y_1, \ldots, Y_n \stackrel{iid}{\sim} \text{Pois}(\lambda)$, write a large-sample $(1-\alpha)100\%$ CI for $P(Y=0) = e^{-\lambda}$.