Bruce M. Boghosian

Properties of Estimators

Consistency

Bruce M. Boghosian

Tufts University School of Arts and Sciences

Department of Mathematics

Tufts University

Outline

Asymptotic unbiasedness

- Recall our calculation of the sample variance of the normal distribution yielded sequence of estimators $\left\{\hat{\sigma}_n^2(\vec{Y})\right\}$, where

$$\hat{\sigma}_n^2(\vec{Y}) = \frac{1}{n}\sum_{j=1}^{n}\left(Y_j - \overline{Y}\right)^2$$

- For finite n, $\hat{\sigma}_n^2(\vec{Y})$ is *biased* since

$$E\left[\hat{\sigma}_n^2(\vec{Y})\right] = \frac{n-1}{n}\sigma^2 \neq \sigma^2.$$

- As $n \to \infty$, this approaches σ^2, so we say that the estimator is *asymptotically unbiased*.

$$E\left[\lim_{n\to\infty}\left(\hat{\sigma^2}(\vec{Y}) - \sigma^2\right)\right] = \lim_{n\to\infty}\left(\frac{n-1}{n}\sigma^2 - \sigma^2\right) = 0.$$

From asymptotic unbiasedness to consistency

- More generally, if $\hat{\theta}_n$ is a sequence of estimators for parameter θ, then
  - If $E(\hat{\theta}_n) = \theta$ for all n, then $\hat{\theta}_n$ is unbiased.
  - If $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$, then $\hat{\theta}_n$ is asymptotically unbiased.
- We may also worry about the shape of the distribution of $\hat{\theta}_n$ in the vicinity of the actual value of θ.
- The probability $P\left(\left|\hat{\theta}_n - \theta\right| < \epsilon\right)$ may not approach 1 as $n \to \infty$, even if $\hat{\theta}_n \to \theta$ in that limit.
- For that to happen, $\mathrm{Var}\left(\hat{\mu}_n\right)$ must go to zero as well.

Definition of consistency

- **Definition:** A sequence of estimators, call it $\hat{\theta}_n = h(W_1, \ldots, W_n)$ is said to be *consistent* if it *converges in probability* to θ. That is, it is consistent if, for any fixed $\epsilon > 0$, however small,

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < \epsilon\right) = 1.$$

- Consistency implies asymptotic unbiasedness, but the reverse is not true.

Example 1

- Suppose $Y_1, \ldots, Y_n$ is sampled from the uniform pdf

$$f_Y(y) = \begin{cases} 1/\theta & \text{if } 0 \le y \le \theta \\ 0 & \text{otherwise} \end{cases}$$

- The MLE estimator is $\hat{\theta}_n(\vec{Y}) = Y_{\max}$
- From order statistics, we know that

$$f_{\hat{\theta}}(y) = f_{Y_{\max}}(y) = \begin{cases} \frac{ny^{n-1}}{\theta^n} & \text{if } 0 \le y \le \theta \\ 0 & \text{otherwise} \end{cases}$$

- So $E\left(\hat{\theta}_n(\vec{Y})\right) = E\left(Y_{\max}\right) = \frac{n}{n+1}\theta$
- Hence $\hat{\theta}_n$ is biased but asymptotically unbiased.
- Is it consistent?

Example 1 (continued)

- Is $\hat{\theta}_n$ consistent?
- Note that for any fixed $\epsilon > 0$,

$$\begin{aligned} P\left(\left|\hat{\theta}_n - \theta\right| < \epsilon\right) &= P\left(\theta - \epsilon < \hat{\theta}_n < \theta\right) \\ &= \int_{\theta-\epsilon}^{\theta} dy\ \frac{ny^{n-1}}{\theta^n} \\ &= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\ &\to 1 \quad \text{as } n \to \infty \end{aligned}$$

- Hence the MLE estimator $\hat{\theta}_n$ is consistent.
- But, as we shall see in our next example, not all asymptotically unbiased estimators are consistent!

Example 2

- For $y \in \mathbb{R}$, a pdf with mean μ and unit variance is

$$f_Y(y;\mu) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|y-\mu|}$$

- Consider sequence of r.v.s $\vec{Y} = \{Y_1, \ldots, Y_n\}$
- Y_n is sampled from $f_Y\left(y;\mu_0 + \frac{1}{n}\right)$
- For r.v.s $\vec{Y}$ generated in this way, consider the estimator

$$\hat{\mu}_n(\vec{Y}) = Y_n$$

- Note that $E(\hat{\mu}_n) = E(Y_n) = \mu_0 + \frac{1}{n}$, so this is not unbiased, but it is asymptotically unbiased since

$$\lim_{n\to\infty} E(\hat{\mu}_n) = \lim_{n\to\infty}\left(\mu_0 + \frac{1}{n}\right) = \mu_0.$$

Example 2 (continued)

- We have shown the estimator for μ given by $\hat{\mu}_n(\vec{Y}) = Y_n$ is *asymptotically unbiased*. But is it *consistent*?
- From elementary integration and a bit of work,

$$P(|\hat{\mu}_n - \mu| < \epsilon) = \int_{\mu-\epsilon}^{\mu+\epsilon} dy\ f_Y(y;\mu) = \begin{cases} 1 - e^{-\sqrt{2}\epsilon}\cosh\left(\sqrt{2}/n\right) & \text{if } n\epsilon > 1 \\ e^{-\sqrt{2}/n}\sinh(\sqrt{2}\,\epsilon) & \text{otherwise} \end{cases}$$

- where we have used the *hyperbolic functions*

$$\cosh z := \frac{e^z + e^{-z}}{2} \quad \text{and} \quad \sinh z := \frac{e^z - e^{-z}}{2}$$

- Fixing $\epsilon > 0$, however small, and letting $n \to \infty$, we find

$$\lim_{n\to\infty} P(|\hat{\mu}_n - \mu| < \epsilon) = 1 - e^{-\sqrt{2}\epsilon} \neq 1,$$

so $\hat{\mu}_n$ is *not consistent*.

Markov's inequality

- **Thm. (Markov):** If X is a nonnegative r.v., and $a > 0$,

$$\boxed{P(X \geq a) \leq \frac{E(X)}{a}.}$$

- **Pf.:** From the definition of expectation and $X \geq 0$,

$$E(X) = \int_{-\infty}^{+\infty} dx\ f_X(x)x = \int_{0}^{+\infty} dx\ f_X(x)x$$

- It follows that

$$\begin{aligned} E(X) &= \int_0^a dx\ f_X(x)x + \int_a^{+\infty} dx\ f_X(x)x \\ &\geq \int_a^{+\infty} dx\ f_X(x)x \geq \int_a^{+\infty} dx\ f_X(x)a = a\ \mathrm{Prob}(x \geq a) \end{aligned}$$

- The desired result immediately follows. $\square$

Chebyshev's inequality

- **Thm. (Chebyshev):** If X is a nonnegative r.v., and $a > 0$, then

$$\boxed{P\left(|X - E(X)|\right) \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.}$$

- **Pf.:** This is really an immediate corollary of Markov's inequality. Using Markov's inequality on the r.v. $|X - E(X)|^2$, and letting $a \to a^2$ results in

$$\begin{aligned} P\left(|X - E(X)| \geq a\right) =& P\left(|X - E(X)|^2 \geq a^2\right) \\ \leq& \frac{E(|X - E(X)|^2}{a^2} = \frac{\mathrm{Var}(X)}{a^2}. \quad \square \end{aligned}$$

Consistency of mean $\hat{\mu}$ with Chebyshev's Theorem

- Let W be any random variable with mean μ and variance σ^2. For any $\epsilon > 0$,

$$P(|W-\mu| < \epsilon) \geq 1 - \frac{\sigma^2}{\epsilon^2} \quad \text{or} \quad P(|W-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

- Suppose $X_1, \ldots, X_n$ is a random sample of size n from a discrete pdf with theoretical mean μ and variance $\sigma^2 < \infty$. Let $\hat{\mu}_n = \frac{1}{n}\sum_{j=1}^n X_j$. Is $\hat{\mu}_n$ consistent?
- By Chebyshev's inequality $P(|\hat{\mu}_n - \mu| < \epsilon) > 1 - \frac{\text{Var}(\hat{\mu})}{\epsilon^2}$
- But $\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n}$, so $P(|\hat{\mu}_n - \mu| < \epsilon) > 1 - \frac{\sigma^2}{n\epsilon^2}$.
- For any ϵ and δ, an n can be found that makes $\frac{\sigma^2}{n\epsilon^2} < \delta$, so

$$\lim_{n\to\infty} P(|\hat{\mu}_n - \mu| < \epsilon) = 1 \qquad \text{(Weak law of large numbers)}$$

Summary

- We have defined and studied consistency.
- We have given two examples of asymptotically unbiased estimators – one consistent, and another not consistent.
- We have used Chebyshev's Theorem to show that the sample mean is always a consistent estimator for the mean.