Asymptotic relative efficiency, delta method

Ciaran Evans

Last time

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Three estimators of σ^2:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

biased

$\text{Bias} = -\frac{1}{n}\sigma^2$

$\text{Var} = \frac{2\sigma^4(n-1)}{n^2}$

unbiased

$\text{Var} = \frac{2\sigma^4}{n-1}$

does not attain CRLB

unbiased

$\text{Var} = \frac{2\sigma^4}{n}$

$= \text{CRLB}$

requires knowledge of μ!

Last time

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Three estimators of σ^2:

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \quad \widetilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

$$Var(\widehat{\sigma}^2) = \frac{2\sigma^4(n-1)}{n^2} < \frac{2\sigma^4}{n-1} = Var(s^2)$$

But asymptotically,

$$\frac{Var(s^2)}{Var(\hat{\sigma}^2)} = \frac{n^2}{(n-1)^2} \to 1 \quad \text{as } n \to \infty$$

So for large samples, there isn't really any difference

Asymptotic relative efficiency

Let $\theta \in \mathbb{R}$ be a parameter of interest, and let $\widehat{\theta}_{1,n}$ and $\widehat{\theta}_{2,n}$ be two estimators of θ such that

$$\sqrt{n}(\widehat{\theta}_{1,n} - \theta) \xrightarrow{d} N(0, \sigma_1^2) \qquad \sqrt{n}(\widehat{\theta}_{2,n} - \theta) \xrightarrow{d} N(0, \sigma_2^2)$$

The **asymptotic relative efficiency** (ARE) of $\widehat{\theta}_{1,n}$ compared to $\widehat{\theta}_{2,n}$ is:

$$ARE(\hat{\theta}_{1,n}, \hat{\theta}_{2,n}) = \frac{\sigma_2^2}{\sigma_1^2}$$

$ARE = 1 \iff$ estimators are equally efficient

$ARE > 1 \iff \hat{\theta}_{1,n}$ is more efficient (smaller variance)

$ARE < 1 \iff \hat{\theta}_{2,n}$ is more efficient

Example: Normal variance estimators

Let $\theta \in \mathbb{R}$ be a parameter of interest, and let $\widehat{\theta}_{1,n}$ and $\widehat{\theta}_{2,n}$ be two estimators of θ such that

$$\sqrt{n}(\widehat{\theta}_{1,n} - \theta) \xrightarrow{d} N(0, \sigma_1^2) \qquad \sqrt{n}(\widehat{\theta}_{2,n} - \theta) \xrightarrow{d} N(0, \sigma_2^2)$$

$$ARE(\widehat{\theta}_{1,n}, \widehat{\theta}_{2,n}) = \frac{\sigma_2^2}{\sigma_1^2}$$

$X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Asymptotic normality of MLE:

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} N(0, \mathcal{I}_1^{-1}(\sigma^2))$$
$$= N(0, 2\sigma^4)$$

Since $s^2 = \frac{n}{n-1}\hat{\sigma}^2$, also have

$$\sqrt{n}(s^2 - \sigma^2) \xrightarrow{d} N(0, 2\sigma^4)$$

So: $ARE(\hat{\sigma}^2, s^2) = \dfrac{2\sigma^4}{2\sigma^4} = \underline{1}$

Example: sample mean vs. sample median

Let $\theta \in \mathbb{R}$ be a parameter of interest, and let $\widehat{\theta}_{1,n}$ and $\widehat{\theta}_{2,n}$ be two estimators of θ such that

$$\sqrt{n}(\widehat{\theta}_{1,n} - \theta) \xrightarrow{d} N(0, \sigma_1^2) \qquad \sqrt{n}(\widehat{\theta}_{2,n} - \theta) \xrightarrow{d} N(0, \sigma_2^2)$$

$$ARE(\widehat{\theta}_{1,n}, \widehat{\theta}_{2,n}) = \frac{\sigma_2^2}{\sigma_1^2}$$

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ want to estimate μ

MLE: $\hat{\mu}_n = \overline{X}$

Also consider sample <u>median</u>, M_n

$$\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

$$\sqrt{n}(M_n - \mu) \xrightarrow{d} N(0, \tfrac{\pi}{2}\sigma^2)$$

$$\Rightarrow ARE(\hat{\mu}_n, M_n) = \frac{\pi}{2} \approx 1.57$$

Sample mean is more efficient than the sample median (by a factor of 1.57)

Limiting distributions of functions of estimators?

Suppose $Y_1, \ldots, Y_n \overset{iid}{\sim} Exponential(\lambda)$, with pdf $f(y) = \lambda e^{-\lambda y}$ for $y > 0$.

$E[Y_1] = \frac{1}{\lambda}$ $\quad Var(Y_1) = \frac{1}{\lambda^2}$

MoM: $\hat{\lambda} = \frac{1}{\bar{Y}}$

$E[Y_1] = \frac{1}{\lambda}$

$\Rightarrow \lambda = \frac{1}{E[Y_1]}$

$\Rightarrow \hat{\lambda} = \frac{1}{\bar{Y}}$

$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d}$ Normal?

CLT: $\sqrt{n}\left(\bar{Y} - \frac{1}{\lambda}\right) \xrightarrow{d} N\left(0, \frac{1}{\lambda^2}\right)$

$Var(\bar{Y}) = \frac{1}{n\lambda^2}$

$\Rightarrow Var\left(\sqrt{n}\left(\bar{Y} - \frac{1}{\lambda}\right)\right) = (\sqrt{n})^2\left(\frac{1}{n\lambda^2}\right) = \frac{1}{\lambda^2}$

what does this tell us about $1/\bar{Y}$?

Delta method

Let $\theta \in \mathbb{R}$, and $\hat{\theta}_n$ an estimator such that

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

for some σ^2 (could be function of θ). Let g be a continuously differentiable function, with $g'(\theta) \neq 0$

Then,

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2)$$

Pf: Taylor expansion:

$$g(\hat{\theta}_n) \approx g(\theta) + g'(\theta)(\hat{\theta}_n - \theta)$$

$$\Rightarrow \sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \approx \underbrace{g'(\theta) \underbrace{\sqrt{n}(\hat{\theta}_n - \theta)}_{\xrightarrow{d} N(0, \sigma^2)}}_{\xrightarrow{d} g'(\theta) N(0, \sigma^2) = N(0, \sigma^2 [g'(\theta)]^2)}$$

$$\Rightarrow \sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2)$$

Example: exponential

Suppose $Y_1, \ldots, Y_n \overset{iid}{\sim} Exponential(\lambda)$, with pdf $f(y) = \lambda e^{-\lambda y}$ for $y > 0$.

$\hat{\lambda} = \frac{1}{\bar{Y}}$

CLT:

$$\sqrt{n}\left(\overline{Y} - \frac{1}{\lambda}\right) \xrightarrow{d} N\left(0, \frac{1}{\lambda^2}\right)$$

$g(u) = \frac{1}{u}$ $\quad g'(u) = -\frac{1}{u^2}$

Delta method: $\sqrt{n}\left(g(\bar{Y}) - g\left(\frac{1}{\lambda}\right)\right) \xrightarrow{d} N\left(0, \frac{1}{\lambda^2}\left[g'\left(\frac{1}{\lambda}\right)\right]^2\right)$

$$\Rightarrow \sqrt{n}\left(\frac{1}{\bar{Y}} - \lambda\right) \xrightarrow{d} N\left(0, \frac{1}{\lambda^2}\left[-\lambda^2\right]^2\right)$$

$$= N(0, \lambda^2)$$

$$\Rightarrow \sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda^2)$$

Example: exponential

Suppose $Y_1, \ldots, Y_n \overset{iid}{\sim} Exponential(\lambda)$, with pdf $f(y) = \lambda e^{-\lambda y}$ for $y > 0$.

$$\widehat{\lambda} = 1/\overline{Y} \qquad \sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda^2)$$

Notice: asymptotic variance depends on the parameter λ that we are trying to estimate!

Delta method: $\sqrt{n}(g(\hat{\lambda}) - g(\lambda)) \xrightarrow{d} N(0, \lambda^2 [g'(\lambda)]^2)$

Can we find g such that $\lambda^2 [g'(\lambda)]^2$ does <u>not</u> depend on λ?

(variance stabilizing transformation)

$\Rightarrow \quad \lambda^2 [g'(\lambda)]^2 = \text{constant}$

$g'(\lambda) \propto \frac{1}{\lambda} \quad \Rightarrow \quad g(\lambda) = \log(\lambda)$

$\sqrt{n}(\log(\hat{\lambda}) - \log(\lambda)) \xrightarrow{d} N(0, \lambda^2 (\frac{1}{\lambda})^2) = N(0, 1)$

Example: Poisson

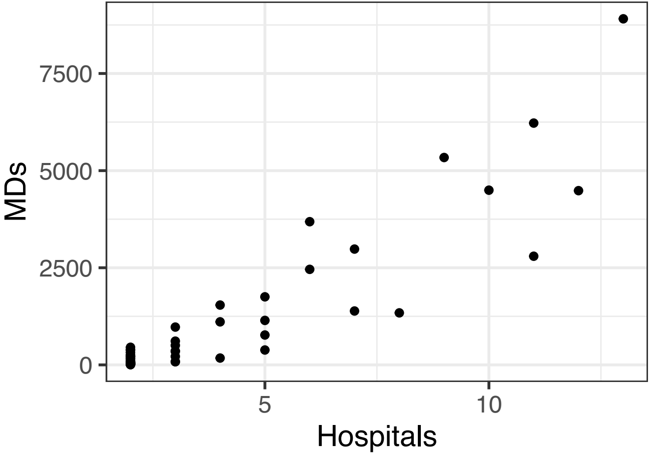
Suppose $Y_1, ..., Y_n \overset{iid}{\sim} Poisson(\lambda)$.

$$\widehat{\lambda} = \overline{Y} \qquad \sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda)$$

Want: variance-stabilizing transformation g such that asymptotic variance of $\sqrt{n}(g(\widehat{\lambda}) - g(\lambda))$ does not depend on λ

Example: non-constant variance

Example: Data on the number of hospitals and number of doctors (MDs) in US counties



Question: How do we adjust for non-constant variance in a linear model?

Example: non-constant variance

Example: Data on the number of hospitals and number of doctors (MDs) in US counties