Variance stabilizing transformations

Recap: delta method

Suppose $\hat{\theta}$ is an estimate of $\theta \in \mathbb{R}$, such that

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

for some σ^2, and g is a continuously differentiable function with $g'(\theta) \neq 0$. Then

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

Variance stabilizing transformations

Let $\hat{\theta}$ be an estimator of θ, and suppose

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

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

g is a <u>variance-stabilizing transformation</u> if

$\sigma^2 [g'(\theta)]^2$ does not involve θ

$$P\left(\log(\hat{\theta}) - z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n}} \le \log(\theta) \le \log(\hat{\theta}) + z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n}}\right) =$$

Example

$$P\left(e^{\log(\hat{\theta}) - z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)} \le \theta \le e^{\log(\hat{\theta}) + z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)}\right)$$

Let $X_1, \ldots, X_n \overset{iid}{\sim} Exponential(\theta)$, with density $f(x|\theta) = \theta e^{-\theta x}$.

$$\hat{\theta} = \frac{1}{\overline{X}} \qquad \sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \theta^2)$$

depends on θ

(asymptotic normality of MLE)

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

if $g'(\theta) = \frac{1}{\theta}$, then $[g'(\theta)]^2 \theta^2 = 1$

$$\Rightarrow \sqrt{n}\left(g(\hat{\theta}) - g(\theta)\right) \xrightarrow{d} N(0, 1)$$

$g(\theta) = \log(\theta) \Rightarrow \sqrt{n}\left(\log(\hat{\theta}) - \log(\theta)\right) \approx N(0, 1)$

$1 - \alpha$ CI for $\log(\theta)$: $\log(\hat{\theta}) \pm z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)$

$1 - \alpha$ CI for θ: $\left[e^{\log(\hat{\theta}) - z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)}, e^{\log(\hat{\theta}) + z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)}\right]$

Also: $\hat{\theta} = \frac{1}{\bar{X}}$ asymptotic normality of MLE

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

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

Let $g(x) = \frac{1}{x} \Rightarrow g'(x) = -\frac{1}{x^2}$

$$\sqrt{n}(\hat{\theta} - \theta) = \sqrt{n}\left(g(\bar{X}) - g\left(\tfrac{1}{\theta}\right)\right)$$

$$\xrightarrow{d} N\Big(0, \underbrace{\left[g'\left(\tfrac{1}{\theta}\right)\right]^2 \frac{1}{\theta^2}}_{\theta^4 \cdot \frac{1}{\theta^2}}\Big) \quad \text{(delta method)}$$

$$= \theta^2$$

$E[X_i] = \frac{1}{\theta}$

$Var(X_i) = \frac{1}{\theta^2}$

$\Rightarrow Var(\bar{X}) = \frac{1}{n\theta^2}$

$\Rightarrow SE(\bar{X}) = \frac{1}{\theta\sqrt{n}}$

$\Rightarrow SE(\sqrt{n}\bar{X}) = \frac{1}{\theta}$

Example

Suppose that $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$.

$\hat{p} = \bar{X}$ $\qquad$ $\sqrt{n}(\hat{p} - p) \xrightarrow{d} N(0, \underbrace{p(1-p)}_{\text{depends on } p})$

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

For variance stabilizing transformation,

$$g'(p) = \frac{1}{\sqrt{p(1-p)}}$$

$g(p) = 2\arcsin(\sqrt{p})$ $\qquad$ works!

Comparison

Two approaches to Wald confidence intervals for binomial probability:

① $\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

② $g(p) = 2\arcsin(\sqrt{p})$

$g^{-1}(x) = \sin^2\left(\frac{x}{2}\right)$

$\sqrt{n}\left(g(\hat{p}) - g(p)\right) \xrightarrow{d} N(0,1)$

$\left[g^{-1}\left(g(\hat{p}) - z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)\right),\ g^{-1}\left(g(\hat{p}) + z_{\frac{\alpha}{2}}\left(\frac{1}{\sqrt{n}}\right)\right)\right]$

How could we investigate their relative performance?

① Simulate data $X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$

② Calculate both intervals for p

③ Repeat many times & compare coverage