1/25/22

Tuesday, January 25, 2022 4:14 PM

Standard Error & the δ-method

Usual setup; $X_1, \cdots, X_n$ iid $f(x;\vartheta)$

$T = T(X_1, \cdots, X_n)$ is an estimator of ϑ

$Bias(T;\vartheta) = \mathbb{E}[T] - \vartheta$

T is <u>unbiased</u> if $Bias(T;\vartheta) = 0$

$MSE(T;\vartheta) = \mathbb{E}[(T-\vartheta)^2]$

if T is unbiased for ϑ

then $MSE(T;\vartheta) = \mathbb{V}[T] \; (\sigma_T^2) = \mathbb{E}[T^2] - \mathbb{E}[T]^2$

Standard Error (T) = $SE(T) = \sqrt{\mathbb{V}[T]} = \sigma_T$

If $\mathbb{E}[X_i] = \mu$; $\mathbb{V}[X_i] = \sigma^2$ then $SE(\bar{X}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$

If we have n iid $X_i \sim N(\mu, \sigma^2)$

then $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, $[\mu = n-1, \sigma^2 = 2(n-1)]$

$$\mathbb{V}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1) = \frac{(n-1)^2}{\sigma^4}\mathbb{V}(S^2) \; ; \; \mathbb{V}(S^2) = \frac{2\sigma^4}{(n-1)}$$

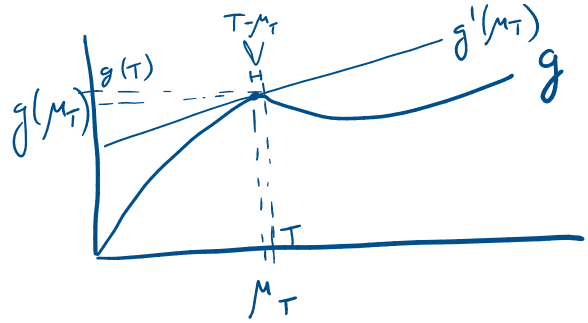
so $SE(S^2) = \frac{\sigma^2\sqrt{2}}{\sqrt{n-1}}$

Introduction to the δ-method

Given $SE(T)$ and a function g:

Consider $g(T)$ .... & $g(T) \approx g(\mu_T) + g'(\mu_T)(T-\mu_T)$

To Estimate $SE(g(T))$



$$SE(g(T)) \approx |g'(\mu_T)| \cdot SE(T)$$

Example: $SE(S^2) = \dfrac{\sigma^2\sqrt{2}}{\sqrt{n-1}}$ — from above

estimate $SE(g(S^2))$ where $g(x) = \sqrt{x}$ & $g(S^2) = S$

$$g'(x) = \frac{1}{2\sqrt{x}}$$

So $SE(g(S^2)) = \dfrac{1}{2\sqrt{\mathbb{E}(S^2)}} \cdot SE(S^2)$

$$= \frac{1}{2\sqrt{\sigma^2}} \cdot \frac{\sigma^2\sqrt{2}}{\sqrt{n-1}} = \frac{\sigma}{\sqrt{2(n-1)}}$$

The notion of Sufficient Statistics

Consider an estimator T of ϑ with $T(X_1 \ldots X_n)$ w/ $X_i \overset{iid}{\sim} f(x_i;\vartheta)$

Consider an ...

$iid \sim f(x_i; \vartheta)$

We say that T is <u>sufficient</u> for $\vartheta \in \Theta$

if $f_{(X_1, \cdots X_n | T)}(x_1, \cdots x_n | T)$ does not depend on ϑ.

Example: let $f(X_i; \vartheta) \sim \text{Binomial}(k=1, p)$

$[\, \mathbb{P}(X_i = 1) = p, \ \mathbb{P}(X_i = 0) = 1 - p \,]$

Consider $T = \sum_{i=1}^{n} X_i$, is it sufficient for $\vartheta = p$?

the pmf

$$f_{(X_1, \cdots, X_n | T)}\left(x_1, \cdots x_n \,\middle|\, T = \sum_{i=1}^{n} x_i\right) = \frac{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)}{\mathbb{P}\left(T = \sum_{i=1}^{n} x_i\right)}$$

$$= \frac{p^{x_1}(1-p)^{(1-x_n)} \cdots p^{x_n}(1-p)^{(1-x_n)}}{\binom{n}{x_1 + \cdots + x_n} p^{x_1 + \cdots + x_n} (1-p)^{n - (x_1 + \cdots + x_n)}}$$

$$= \frac{1}{\binom{n}{x_1 + \cdots + x_n}}$$

* which does not depend on p *

Thus T is sufficient for p.

Note Here*

We can only compute the conditional pmf/pdf if our statistic is sufficient...

otherwise, we would need to know ϑ & our analysis would be circular.

$\boxed{\text{As a result, the conditional pmf/pdf is a measure of the accuracy of our estimation}}$

Likelihood function $L(\vartheta \mid x_1, \ldots, x_n)$

generally cannot compute as we generally do not know ϑ

$$\underbrace{L(\vartheta \mid x_1, \ldots, x_n)}_{\text{statistical notation; } x_1, \ldots x_n \text{ "data" given}} = \underbrace{f(x_1, \ldots, x_n ; \vartheta)}_{\text{probability notation; think } \vartheta \text{ given}}$$

this is purely notational or psychological

note, not a conditional, think of as "such that", or "given"