Lecture 12 10/19/2020

In lec 10 $\frac{\hat{\theta}-\theta}{SE[\hat{\theta}]} \xrightarrow{d} N(0,1) \Rightarrow \frac{\hat{\theta}-\theta}{\hat{SE}[\hat{\theta}]} \xrightarrow{d.} N(0,1)$

We can use this now in our situation:

$$\frac{\hat{\theta}_1-\hat{\theta}_2}{SE[\hat{\theta}_1-\hat{\theta}_2]} \xrightarrow{d} N(0,1)$$

$$\Rightarrow \frac{\hat{\theta}_1-\hat{\theta}_2}{\hat{SE}[\hat{\theta}_1-\hat{\theta}_2]} \xrightarrow{d} N(0,1)$$

$$SE[\hat{\theta}_1-\hat{\theta}_2] = \sqrt{\theta_{shared}(1-\theta_{shared})(1/n_1 + 1/n_2)}$$

$$\hat{SE}[\hat{\theta}_1-\hat{\theta}_2] = \sqrt{\hat{\theta}_{shared}(1-\hat{\theta}_{shared})(1/n_1 + 1/n_2)}$$ if $\hat{\theta}_{shared}$ is consistent.

$\hat{\theta}_{shared}$ = avg. over both samples = $\frac{\sum X_{1i} + \sum X_{2i}}{n_1 + n_2}$

$$\Rightarrow \frac{\hat{\theta}_1-\hat{\theta}_2}{\sqrt{\frac{\sum X_{1i}+\sum X_{2i}}{n_1+n_2}\left(1-\frac{\sum X_{1i}+\sum X_{2i}}{n_1+n_2}\right)\left(1/n_1+1/n_2\right)}} \;\dot{\sim}\; N(0,1)$$

e.g. $H_a: \theta_1-\theta_2 \neq 0$, $H_0: \theta_1-\theta_2 = 0$, $\alpha = 5\%$

control $n_1 = 81$, $\sum X_{1i} = 27 \Rightarrow \hat{\theta}_1 = 27/81 = 0.333$

experimental $n_2 = 79$, $\sum X_{2i} = 12 \Rightarrow \hat{\theta}_2 = 12/79 = 0.152$

$\hat{\theta}_{shared} = \frac{27+12}{81+79} = 0.244$

0.244

$$(\hat{\theta}_1 - \hat{\theta}_2)_{STD} = \frac{0.333 - 0.152}{\sqrt{0.244(1-0.244)(1/81 + 1/99)}} = 2.66 \notin [-1.96, +1.96] \Rightarrow \text{Reject } H_0$$

Another (obvious) Wald Test: If $X_1, \ldots, X_n$ iid DGP with mean theta and variance $\sigma^2$ and the estimator $\hat{\theta}$ is $\bar{X}$, then the CLT implies that:

$$\frac{\hat{\theta} - \theta}{\sigma / \sqrt{n}} \xrightarrow{d} N(0,1) \text{ if } \sigma \text{ is known.}$$

If $\sigma$ is unknown ... I can replace $\sigma$ with any consistent estimator e.g. $S$, $\hat{\sigma}$ & $1/n \sum (X_i - \theta)^2$.

$$\Rightarrow \frac{\hat{\theta} - \theta}{S/\sqrt{n}} \xrightarrow{d} N(0,1)$$

Are you allowed to just use the T-test here? Many people just use the T-test here. Technically it's wrong because you need the DGP to be normal iid. But if you use the T-test ... it's "not so bad". (in midterm 2 Q11)

$$H_0: \theta > 2, \; n = 30, \; \bar{X} = 2.57, \; S = 1.00$$

$$\hat{\theta}_{std} = \frac{2.57 - 2}{1.00/\sqrt{30}} = 3.12 \notin (-\infty, 1.645] \Rightarrow \text{Rej } H_0.$$

Another Wald test for two independent samples with unknown variances and you wish to test a different in means.

$$\frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \xrightarrow[\text{from last class}]{d} N(0,1)$$

$$\Longrightarrow \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \xrightarrow{d} N(0,1)$$

If you use the Satterthwaite t-test, it "wouldn't be so bad" because unless your population distributions were so very skewed, it should be fine.

Let's use the asymptotic normality of the MLE thm (last class) to do a Wald Test. HW#1, m has DGP: $X_1, \ldots, X_n$ iid Gumbel $(\theta, 1)$. The Gumbel is a rv model for "extreme events" think maximum rainfall per month:

$$\ell'(\theta; X_1, \ldots X_n) = n - e^{\theta} \Sigma e^{-X_i} \overset{\text{set}}{=} 0$$

$$\Longrightarrow \hat{\theta}^{MLE} = \ln\left(n / \Sigma e^{-X_i}\right)$$

$$\ell'(\theta; X) = 1 - e^{\theta} e^{-X} \Rightarrow \ell''(\theta; X) = -e^{\theta} e^{-X} \ ?$$

$$I(\theta) = E[\ell''(\theta; X)] = E[e^{\theta} e^{-X}] = e^{\theta} E[e^{-X}]$$
$$= e^{-2\theta}.$$

$$\frac{\hat{\theta}^{MLE}-\theta}{\sqrt{\frac{I(\theta)^{-1}}{n}}} = \frac{\hat{\theta}^{MLE}-\theta}{\frac{e^{\theta}}{\sqrt{n}}} = \frac{\ln\left(n/\Sigma e^{-x}\right)-\theta}{e^{\theta}/\sqrt{n}} \xrightarrow{d} N(0,1)$$

$X_1 = 2.15$, $X_2 = 1.91$, $\underset{X_3}{3.66}$, $\underset{X_4}{4.85}$, $\underset{X_5}{3.03}$, $\underset{X_6}{1.03}$, $\underset{X_7}{3.58}$,
$n = 7$

$\hat{\theta}^{MLE} = 2.26$ Test $H_a: \theta > 2$. $\alpha = 5\%$.

$$\hat{\theta}^{MLE}_{STD} = \frac{2.26-2}{e^{2}/\sqrt{7}} = \frac{0.26}{2.79} = 0.09 \in [-\infty, 1.645]$$

$\Rightarrow$ Retain $H_0$

There are three goals of statistical inference
(1) Point Estimation
Goal here is to provide a best guess, $\hat{\theta}$ of the value of $\theta$. You don't know if your specific guess is good, is close, is bad, is far ... How do we ask the question "is it good / bad"? We imagined $\hat{\theta}$ coming from a distribution $\hat{\theta}$, the "Sampling distribution". ~~There are properties about the "sampling distribution"~~ There are properties about the sampling distribution e.g. some good properties are unbiasedness, consistency, low MSE, low risk (for general loss functions).

(2) Testing
Goal here is to test a theory about a specific $\theta$. We used hypothesis testing. What makes a good test? One property is "power". There are other properties we didn't discuss.

(3) Confidence Sets

The goal here is to create a set of values for $\theta$ that you're "confident in". The approach we use here is the "confidence interval".

Def

an "interval estimate" are two statistics:

$W_L(X_1, \ldots, X_n)$ & $W_U(X_1, \ldots X_n)$ s.t $W_L < W_U$ for all data set,

Combined in an interval: $[W_L(X_1, \ldots, X_n), W_U(X_1, \ldots X_n)]$

e.g. $[1.789, 2.463]$

and of course, the "interval estimator" is:

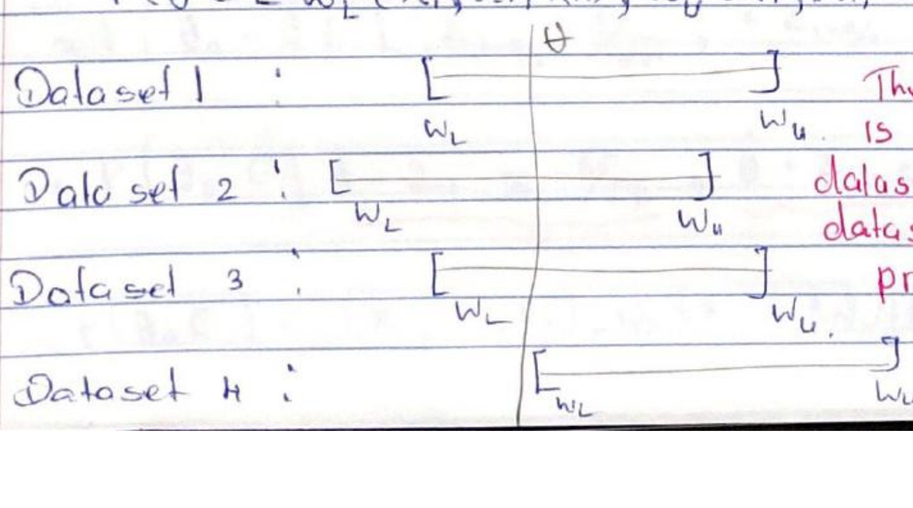
$$[W_L(X_1, \ldots X_n), W_U(X_1, \ldots, X_n)]$$

which is a "random interval".

Def

An interval estimator has "coverage probability"

$P(\theta \in [W_L(X_1, \ldots, X_n), W_U(X_1, \ldots, X_n)] \mid \theta)$ An illustration



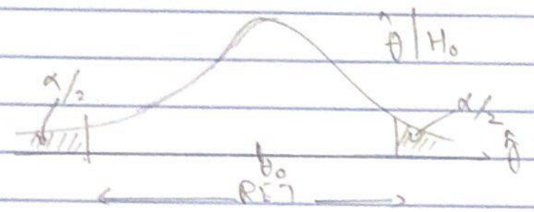

The coverage prob. is computed over every dataset. For these 4 datasets, the coverage prob. would be $3/4 = 75\%$.

We define the "confidence interval" with coverage probability 1 - alpha for parameter $\theta$ as this interval estimate and interval estimator (depending on context). Denoted $CI_{\theta, 1-\alpha}$.

Given alpha, how do we find the confidence interval? Let's begin with the DGP iid normal mean $\theta$, variance $\sigma^2$ & variance known & the estimator $= \bar{X}$.

Consider testing:

$$H_a: \theta \neq \theta_0 \quad \text{vs.} \quad H_0: \theta = \theta_0 \quad @ \text{ size } \alpha.$$




$$z_{1-\alpha/2} := F_z^{-1}(1-\alpha/2)$$

$$P(\hat{\theta} \in REJ \mid H_0) = 1-\alpha.$$

$$\parallel$$

$$P\left(\hat{\theta} \in \left[\theta_0 - z_{1-\alpha/2} \cdot \sigma/\sqrt{n}, \ \theta_0 + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}\right] \middle| \theta = \theta_0\right)$$

$$= P\left(\hat{\theta} - \theta_0 \in \left[-z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \ + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}\right] \middle| \theta = \theta_0\right)$$

$$= P\left(\theta_0 - \hat{\theta} \in \left[-z_{1-\alpha/2} \cdot \sigma/\sqrt{n}, \ + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}\right] \middle| \theta = \theta_0\right)$$

$$= P\left(\theta_0 \in \left[\underbrace{\hat{\theta} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n}}_{W_L}, \ \underbrace{\hat{\theta} + z_{1-\alpha/2} \, \sigma/\sqrt{n}}_{W_U}\right] \middle| \theta = \theta_0\right)$$

$$= P\left(\theta_0 \in \left[W_L(X_1, \ldots, X_n), W_U(X_1 \ldots X_n)\right] \middle| \theta = \theta_0\right)$$ Since valid $\forall \theta_0 \ldots$

$$\Rightarrow CI_{\theta, 1-\alpha} \doteq \left[ \hat{\theta} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \, , \, \hat{\theta} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \right]$$

We constructed our first confidence interval by "inverting the test".