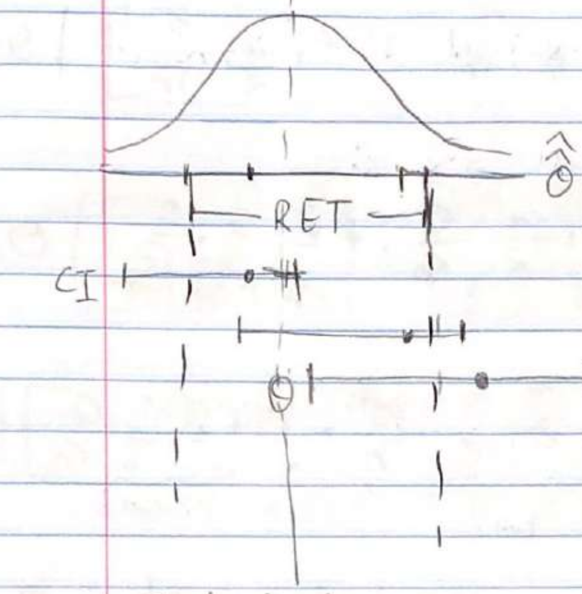
lecture 13:




theta0 ($\theta_0$) being in the confindence interval with coverage probability $1-\alpha$, is equivalent to the test at size alpha ~~retaining~~ retaining.

$$\hat{\theta} \in RET_{\theta_0,\alpha}$$

$$\Updownarrow$$

$$\hat{\theta}_0 \in CI_{\hat{\theta},1-\alpha}$$

P421 C & B: Both hypothesis testing and interval construction look for consonance b/t the sample statistic ($\hat{\theta}$) and the Population parameter ($\theta$).

Hypothesis tests fix the value of the Parameter theta ($H_0$) and ask "is the estimate $\hat{\theta}$ is agreement?" if no $\Rightarrow$ Reject.

Confidence sets fixes the estimate ($\hat{\theta}$) and ask "which values of the Parameter ($\theta$) are in agreement"?

We invented a 2 sided hypothesis test to get a 2 sided CI. You can also have a 1 sided CI eg:

$$CI_{L,\theta,1-\alpha} := [\,w_L(x_1,..,x_n), \infty) \text{ or } CI_{R,\theta,1-\alpha} := (-\infty, w_u(x_1, \ldots x_n)]$$

But we won't do this in class only for the interest of saving time and moving on to other topics.

Sometimes the Sampling distribution was approximate. Inverting that test will yield CI's with approximate coverage i.e "approximate CI's". Let's build some popular CI's!

DGP: $\overset{iid}{\sim} N(\theta, \sigma^2)$ with $\sigma^2$ unknown

$\hat{\theta} = \bar{X}$

$$CI_{\theta,\,1-\alpha} = \left[\hat{\theta} \pm \underbrace{t_{1-\alpha/2,\,n-1}\,\frac{S}{\sqrt{n}}}_{\text{Margin of error}}\right]$$

$$= \left[\longleftarrow \; \hat{\theta} \; \longrightarrow\right]$$

DGP:

$X_{11}, \dots, X_{1n} \overset{iid}{\sim} N(\theta_1, \sigma_1^2)$ indep of $X_{21}, \dots, X_{2n_2} \overset{iid}{\sim} N(\theta_2, \sigma_2^2)$

$\hat{\theta}_1 - \hat{\theta}_2 = \bar{X}_1 - \bar{X}_2$

if $\sigma_1^2, \sigma_2^2$ known

$$CI_{\theta_1-\theta_2,\,1-\alpha} \overset{\uparrow}{=} \left[(\hat{\theta}_1 - \hat{\theta}_2) \pm z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

if $\sigma_1^2 = \sigma_2^2 = \sigma^2$ known

$$\overset{\uparrow}{=} \left[(\hat{\theta}_1 - \hat{\theta}_2) \pm z_{1-\frac{\alpha}{2}}\,\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

if $\sigma_1^2 = \sigma_2^2$ but unknown

see lec 6

$$\overset{\uparrow}{=} \left[(\hat{\theta}_1 - \hat{\theta}_2) \pm t_{1-\frac{\alpha}{2},\,n_1+n_2-2}\, S_{\text{pooled}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$$

if $\sigma_1^2 \neq \sigma_2^2$ and unknown

$$\overset{\uparrow}{\approx} \left[(\hat{\theta}_1 - \hat{\theta}_2) \pm t_{1-\frac{\alpha}{2},\,df}\sqrt{\frac{S_1}{n_1} + \frac{S_2}{n_2}}\right]$$

See lec 7 for the Satterthwaite approximation

DGP: iid Bern($\theta$), $\hat{\theta} = \bar{X}$ via the CLT

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

$$\Rightarrow \frac{\hat{\theta}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \stackrel{\cdot}{\sim} N(0,1) \xRightarrow{\text{Via thms 5.5, 5.4, 5.8}} \frac{\hat{\theta}-\theta}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \xrightarrow{d} N(0,1)$$

using this fact and following [illegible]

$$\Rightarrow P\left(\frac{\hat{\theta}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \in \left[-z_{1-\frac{\alpha}{2}}, +z_{1-\frac{\alpha}{2}}\right]\right) \approx 1-\alpha$$

$$\Rightarrow P\left(\frac{\theta-\hat{\theta}}{\sqrt{\frac{\theta(1-\theta)}{n}}} \in \left[-z_{1-\frac{\alpha}{2}}, +z_{1+\frac{\alpha}{2}}\right]\right) \approx 1-\alpha$$

$$\Rightarrow P\left(\theta \in \left[\hat{\theta} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta(1-\theta)}{n}}, \hat{\theta} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta(1-\theta)}{n}}\right]\right) \approx 1-\alpha$$

$$\Rightarrow CI_{\theta,1-\alpha} \approx \left[\hat{\theta} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta(1-\theta)}{n}}, \hat{\theta} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta(1-\theta)}{n}}\right]$$

this is a fail... I don't know theta!

$$\Rightarrow CI_{\theta,1-\alpha} \approx \left[\hat{\theta} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}, \hat{\theta} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}\right]$$

this is a CI for binomial Proportion. It is actually a lead approximation for low n and theta near 0 or 1. There are other CI'S we won't study and it is actually an area of modern research.

DGP; $X_{11}, \ldots X_{1n} \overset{iid}{\sim} Bern(\theta_1)$ indep of

$X_{21}, \ldots X_{2n_2} \overset{iid}{\sim} Bern(\theta_2)$

From lec 11, $\dfrac{(\hat{\theta}_1 - \hat{\theta}_2)\ (\theta_1 - \theta_2)}{\sqrt{\dfrac{\theta_1(1-\theta_1)}{n_1} + \dfrac{\theta_2(1-\theta_2)}{n_2}}} \xrightarrow{d} N(0,1)$,

$$\xrightarrow[\text{Slutsky's}]{\text{Thm 5.5.4-8}} \frac{(\hat{\theta}_1 - \hat{\theta}_2) - (\theta_1 - \theta_2)}{\sqrt{\dfrac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \dfrac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}} \xrightarrow{d} N(0,1)$$

$$\Rightarrow CI_{\theta_1 - \theta_2, 1-\alpha} \approx \left[ (\hat{\theta}_1 - \hat{\theta}_2) \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}} \right]$$

E.g. from the medical study

$n_1 = 81$, $\hat{\hat{\theta}}_1 = 0.333$, $n_2 = 79$, $\hat{\hat{\theta}} = 0.152$

$$CI_{\theta_1-\theta_2, 95\%} \approx \left[(0.333-0.152) \pm 1.96\sqrt{\frac{0.333\cdot 0.667}{81} + \frac{.152 \times .848}{79}}\right]$$

$$= [.181 \pm 1.96 \cdot 0.066]$$

$$= [0.051, 0.311]$$

You are 95% confident that the true proportion difference is between 5.1% and 31.1%.

DGP iid some rv with mean $\theta$, variance $\sigma^2$ unknown.

$\cancel{CI_{\theta, 1-\alpha}} =$ $\hat{\theta}_2 \longrightarrow \bar{X}$ By CLT $\frac{\hat{\theta}-\theta}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} N(0,1)$

$$\Downarrow$$

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

$$CI_{\theta, 1-\alpha} \approx \left[\hat{\hat{\theta}} \pm z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right]$$ if you use the $t$ it won't be "so bad"

Prob 11 on Midterm I, $\bar{X} = 2.57$, $S = 1.00$

$$CI_{\theta, 95\%} \approx \left[2.57 \pm 1.96 \frac{1.00}{\sqrt{30}}\right]$$

$$= [2.212, 2.228]$$

* DGP iid $f(\theta)$ where $\hat{\theta} = \hat{\theta}^{MLE}$

from lec 11;

$$\frac{\hat{\theta}^{MLE} - \theta}{\sqrt{\frac{I(\theta)^{-1}}{n}}} \xrightarrow{d} N(0,1)$$

$$\overset{\text{Richard}}{\Rightarrow} \frac{\hat{\theta}^{MLE} - \theta}{\sqrt{\frac{I(\hat{\theta}^{MLE})^{-1}}{n}}} \xrightarrow{d} N(0,1)$$

$$\Rightarrow CI_{\theta, 1-\alpha} \approx \left[\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{I(\hat{\theta}^{MLE})^{-1}}{n}}\right]$$

example from last class.

DGP: iid Gumbel$(\theta, 1)$ and the

data is $\langle 2.15, 1.91, 3.66, 4.85, 3.03,$

1.03, 3.58 > and $n = 7$. Find a 95% CI for $\theta$.

$$\hat{\theta}^{MLE} = \ln\left(\frac{n}{\sum e^{-x_i}}\right), \quad \hat{\theta}^{MLE} = 2.26$$

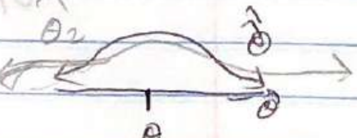
$$I(\theta)^{-1} = e^{\theta} \Rightarrow I(\hat{\theta}^{MLE})^{-1} = 9.57$$

$$CI_{\theta, 95\%} \approx \left[2.26 \pm 1.96 \cdot \frac{2.57}{\sqrt{7}}\right]$$

$$= [0.58, 3.93]$$

NOW that we've been properly introduced to statistical inference (all three goals) let's talk about some big picture things.

For an unbiased estimator, MSE (being small) is KING. Why?



(1) point estimation: The lower the MSE, the closer $\hat{\theta}$ is to $\theta$ on average.

(2) Hypothesis testing: Most estimators we discussed with exactly or approximately normally distributed. Thus the retainment region for a ~~$\theta$~~ 2-sided test looks like:

$$RET = \left[\theta_0 \pm z_{1-\frac{\alpha}{2}} \underbrace{\sqrt{\widehat{MSE}}}_{= SE[\hat{\theta}]}\right]$$

with a smaller MSE ⇒ Smaller RET ⇒ Higher power!

(3) Confidence Intervals:

For exactly or approximately normally distributed estimators,

$$CI_{\theta, 1-\alpha} \underset{\approx}{\overset{=}{\text{or}}} \left[\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{MSE}\right]$$

or $\sqrt{\widehat{MSE}}$

via Richard Man

A lower MSE means a tighter / smaller CI which means you are more confidence about where $\theta$ lies e.g.

$$CI_{\hat{\theta}, 95\%} = [0.49, 5.1]$$

vs

$$CI_{\theta, 95\%} = [.449, .5001]$$

Let's picture all these goals.