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

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

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

$$\hat{SE}[\hat{\theta}_1-\hat{\theta}_2] = \sqrt{\hat{\theta}_{shared}(1-\hat{\theta}_{shared})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$ if $\hat{\theta}_{shared}$ is consistant

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

$$\Rightarrow \frac{\hat{\theta}_1-\hat{\theta}_2}{\sqrt{\frac{\Sigma x_{1i} - \Sigma x_{2i}}{n_1+n_2}\left(1-\frac{\Sigma x_{1i}+\Sigma x_{2i}}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \overset{\text{approx}}{\dot{\sim}} N(0,1)$$

(double approx)

2-Proportion Z-test (Diffrence in mean)

$H_a\ \theta_1-\theta_2 \neq 0 \quad H_0\ \theta_1-\theta_2 = 0 \quad \alpha = 5\%$

Control group: $n_1 = 81$, $\Sigma x_{1i} = 27 \Rightarrow \hat{\theta}_1 = \frac{27}{81} = .333$, $\hat{\theta}_{shared} = \frac{27+12}{81+79}$

exp group: $n_2 = 79$, $\Sigma x_{2i} = 12 \Rightarrow \hat{\theta}_2 = \frac{12}{79} = .152$

$\rightarrow \hat{\theta}_{shared} = .244$

$$(\hat{\theta}_1-\hat{\theta}_2)_{stat} = \frac{.333-.152}{\sqrt{.224(1-.244)\left(\frac{1}{81}+\frac{1}{79}\right)}} = 2.66 \notin [-1.96, 1.96]$$

reject H_0

Another (obvious) Wald test: If $x_1 \ldots x_n$ iid DGP with mean θ & var σ^2 & estimator $\hat{\theta} = \bar{x}$, then CLT implies that:

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

If σ is uknown, replace σ with any cosisitent estimator e.g (S, $\hat{\sigma}$ & $\frac{1}{n}\Sigma(Y_i-\theta)^2$

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

Scanned with CamScanner

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 ∴ I did this on Problem 11 of the midterm

$H_a: \theta > 2$, $n=30$, $\bar{x} = 2.57$, $s = 1.00$

$$\hat{\theta}_{std} = \frac{2.57-2}{\frac{1.00}{\sqrt{3}}} = 3.12 \notin (-\infty, 1.645) \text{ Reject } H_0$$

Another Wald test for two independent Samples with unknown Variance & You wish to ~~send~~ test a different in means.

$$\frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \xrightarrow{d} N(0,1) \Rightarrow \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \xrightarrow{d} N(0,1)$$

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

Lets use the asymptotic normality of the MLE thm (last class) to do Wald test. HW4 L,m has DGP: $X_1 \ldots X_n$ iid Gumbel $(\theta, 1)$ The Gumbel is a r.v model for "extreme even" think maximum rainfall per month

$$\ell'(\theta; x_1, \ldots, x_n) = n - e^{\theta} \Sigma e^{-x_i} \stackrel{set}{=} 0 \Rightarrow \hat{\theta}^{mle} = \ln\left(\frac{n}{\Sigma e^{-x_i}}\right)$$

$$\ell'(\theta; x) = 1 - e^{\theta} e^{-x} \Rightarrow \ell''(\theta; x) = -e^{\theta} e^{-x} \quad \stackrel{?}{0} \text{ (Prof made on error)}$$

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

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

Scanned with CamScanner

Ex: $x_1 = 2.15$, $x_2 = 1.91$, $\overset{x_3}{3.66}$, $\overset{x_4}{4.85}$, $\overset{x_5}{3.03}$, $\overset{x_6}{1.03}$, $\overset{x_7}{3.58}$, $n = 7$

$\hat{\theta}^{mle} = 2.26$, test $H_a: \theta > 2$ $\alpha = 5\%$

$$\hat{\theta}^{mle}_{std} = \frac{2.26 - 2}{\frac{e^2}{\sqrt{7}}} = \frac{.26}{2.79} = .09 \in [-\infty, 1.645]$$

$\Rightarrow$ Retain H_0

There are 3 goals of Statistical inference

(1) Point estimation

Goal Here is to provide a best guess, $\hat{\theta}$ of the value of θ. You dont 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 distrib" 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 specific theta "(θ)". We used hypothesis testing. What makes a "good test"? One Property is power. There are other Properties we didnt discuss.

(3) Confidence Sets

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

Definition: 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: $[w_L(x_1 \ldots x_n), w_U(x_1 \ldots x_n)]$

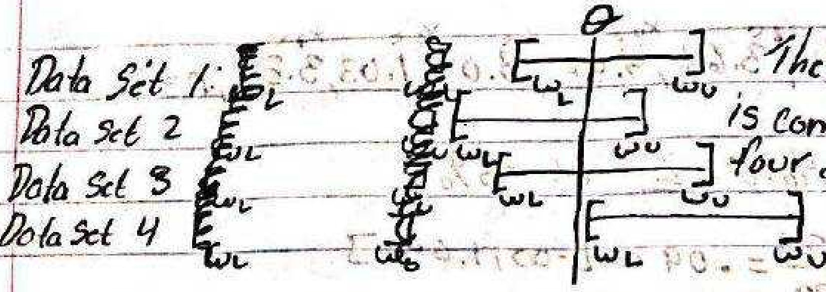
eg $[1.789, 2.463]$

and of course, the "interval estimator" $^{(r.v)}$ is: $[w_L(X_1 \ldots X_n), w_U(X_1 \ldots$

Which is a "random interval".

Defenition: An interval estimator has "Coverage Probabillity" $P(\theta \in [w_L(X_1 \ldots X_n), w_U(X_1 \ldots X_n)] \mid \theta)$. An illustration:

Do



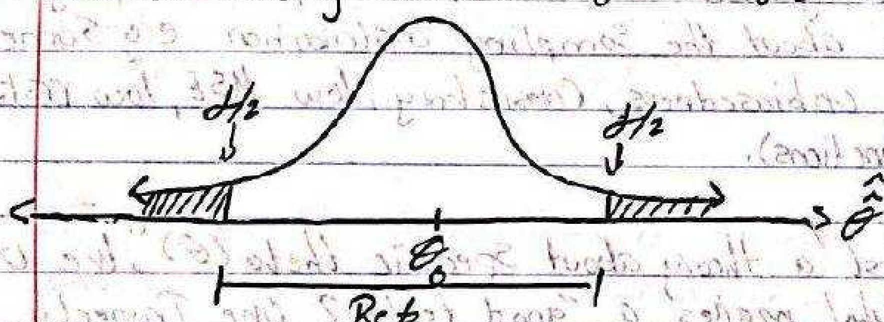
The Coverage Probability is computed over every datasets. For these four datasets, the Coverage Probability would be $3/4 = 75\%$

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

Given alpha, how do we find the confidence interval? Lets begin with the DGP iid normal mean θ, varience σ^2 & Varience known & estimator $= \bar{x}$

Consider testing $H_a: \theta = \theta_0$ vs $H_0: \theta = \theta_0$ at size α

$Z_{1-\alpha/2} : F_Z^{-1}(1-\alpha/2)$



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

$$P(\hat{\theta} \in \theta_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \theta_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}] \mid \theta = \theta_0)$$

$$= P(\hat{\theta} - \theta_0 \in [-Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, +Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}] \mid \theta = \theta_0)$$

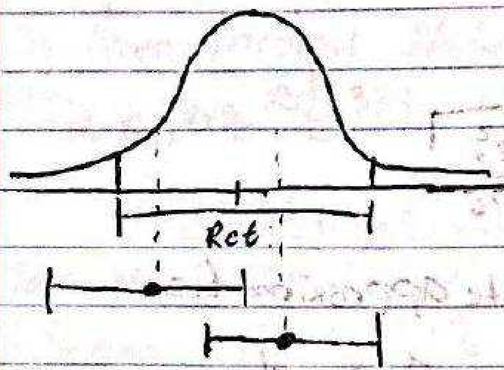
$$= P(\hat{\theta} - \theta_0 \in [-Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}, +Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}] \mid \theta = \theta_0)$$

$$= P(\theta_0 \in [\underbrace{\hat{\theta} - Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}}_{W_L}, \underbrace{\hat{\theta} + Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}}_{W_U}] \mid \theta = \theta_0)$$

$$= P(\theta_0 \in [\omega_L(X_1 \ldots X_n), \omega_U(X_1 \ldots X_n)] \mid \theta = \theta_0) \text{ since Valid } \forall \theta_0$$

$$\Rightarrow CI_{\theta, 1-\alpha} = [\hat{\theta} - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \hat{\theta} + Z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}]$$

We Constructed our first confidence interval by "inverting the test"



theta being in Confidence interval with Coverage Probability $1-\alpha$ $\hat{\theta}$ is equivalent to test at size alpha retainment.

$\hat{\theta} \in Ret_{\theta_0,\alpha} \iff \theta_0 \in CI_{\theta,1-\alpha}$

Pg421 C&B: both hypothesis testing & interval construction look for Consonance between the Sample Statistics ($\hat{\theta}$) & the Population Parameter (θ).
HyPothesis test fix Value of the Parameter θ (H_0) & ask "is the estimate" $\hat{\theta}$ in agreement? If no => Reject
Confidence set fixes the estimate ($\hat{\theta}$) & ask "which Values of the Parameter θ are in agreement?"

We inverted a 2-sided hypothesis test to get a 2-sided CI
You also have a 1-sided CI e.g
$CI_{L,\theta,1-\alpha} := [w_L(x_1 \dots x_n), \infty)$ or $CI_{R,\theta,1-\alpha} := (-\infty, w_U(x_1 \dots x_n)]$
but we wont do this in Class for the intrest of Saving time and moving on to other topics

Sometimes the sampling distribution was aproximate. Inverting that test will yield CI's with approximate coverage ie "approximate CI's. Lets build some Popular CI's!

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

$CI_{\theta,1-\alpha} = [\hat{\theta} \pm t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}]$ ← exact

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

if σ_1^2, σ_2^2 known

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

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

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

Lec 6

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

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

Scanned with CamScanner

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

$$\approx \left[(\hat{\theta}_1 - \hat{\theta}_2) \pm t_{1-\frac{\alpha}{2}, df} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right]$$

See lec 7 for Satterthwaite approximation

DGP: $\overset{iid}{\sim}$ Bern(θ), $\hat{\theta} = \bar{X}$ via the CLT $\frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \xrightarrow{d} N(0,1)$

$$\Rightarrow \frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \stackrel{\cdot}{\sim} N(0,1). \overset{\text{Via thm 5.5.4 Slutsky's}}{\Rightarrow} P\left(\frac{\hat{\theta} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \in [-z_{1-\frac{\alpha}{2}}, +z_{1-\frac{\alpha}{2}}]\right) = 1 - \alpha$$

$$\boxed{\frac{\hat{\theta} - \theta}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \xrightarrow{d} N(0,1)}$$ using this fact & following through

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

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

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

this fails, I dont know theta

We actually use

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

this is "a" CI for the binomial Proportion. It is actually a bad approximation for low n & θ near 0 or 1 there are other CI's we wont study and it's actually an area of modern Research.

Proportion

DGP: $X_{11} \ldots X_{1n_1} \overset{iid}{\sim}$ Bern(θ_1) indep of $X_{21} \ldots X_{2n_2} \overset{iid}{\sim}$ Bern(θ_2)

from lec 11, $\frac{(\hat{\theta}_1 - \hat{\theta}_2) - (\theta_1 - \theta_2)}{\sqrt{\frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}}} \xrightarrow{d} N(0,1) \overset{5.5.4 \text{ slutsky}}{\Rightarrow}$

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

Scanned with CamScanner

eg. from medical studies, $n_1 = 81, \hat{\theta}_1 = .333$, $n_2 = 79$, $\hat{\theta}_2 = .152$

$$CI_{\theta_1 - \theta_2, 95\%} \approx \left[(.333 - .152) \pm 1.96\sqrt{\frac{.333(.667)}{81} + \frac{.152 \cdot .848}{79}}\right]$$

$$= [.181 \pm 1.96(.066)] = [0.051, .311]$$

Your 95% Confident that θ the true Proportion diffrence is between 5.1% & 31.1%

by CLT $\dfrac{\hat{\theta} - \theta}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} N(0,1)$ $\Downarrow$ $\dfrac{\hat{\theta} - \theta}{\frac{s}{\sqrt{n}}} \xrightarrow{d} N(0,1)$

DGP iid some RV with mean θ, varience σ^2 unknown. $\hat{\theta} = \bar{X}$

$$CI_{\theta, 1-\alpha} \approx \left[\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right]$$

if you use the t test it wont be "so bad"

Problem 11 on midterm I: $\bar{x} = 2.57$, $s = 1.00$

$$CI_{\theta, 95\%} = [2.57 \pm 1.96 \tfrac{100}{\sqrt{30}}] = [2.212, 2.928]$$

~~DGP~~

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

from Lec 11, $\dfrac{\hat{\theta}^{mle} - \theta}{\sqrt{\frac{I(\theta)^{-1}}{n}}} \xrightarrow{d} N(0,1) \Rightarrow \dfrac{\hat{\theta}^{mle} - \theta}{\sqrt{\frac{I(\hat{\theta}^{mle})^{-1}}{n}}} \xrightarrow{d} N(0,1)$ 5.5.4

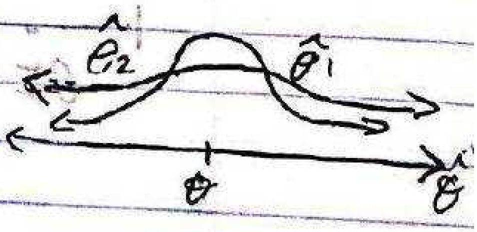
$$\Rightarrow CI_{\theta, 1-\alpha} \approx \left[\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{I(\hat{\theta}^{mle})^{-1}}{n}}\right]$$

Example from last class

DGP iid Gumbel$(\theta, 1)$ $\langle 2.15, 1.91, 3.66, 4.85, 3.03, 1.03, 3.58 \rangle$ $n = 7$ 95%

CI for θ

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



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

$$CI_{\theta, 95\%} \approx \left[2.26 \pm 1.96 \left(\frac{9.57}{\sqrt{7}}\right)\right] = [.58, 3.93]$$

Now we've been Properly introduced to Statistical inference (all three goals), lets talk about big Picture things

- for an unbiased estimator, MSE (being small) is king Why?

(1) Point ~~estimation~~ estimation

the lower the MSE, the closer $\hat{\theta}$ is to θ on average

Scanned with CamScanner

(2) HyPothesis testing
Mostly estimators we disscussed with exactly or approximately normaly distributed. Thus the retainment regior for a 2-sided test looks like:

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

(3) Confidence Intervals
for exactly or approximately normally distributed estimator,

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

or $\sqrt{\hat{MSE}}$ thrm 5.5.4

A lower MSE means a tighter/smaller CI which means youre more confident about where θ lies e.g

$$CI_{\theta, 95\%} = [.49, 5.1] \text{ vs } CI_{\theta, 95} = [.449, .5001]$$

Lets Picture all three goals