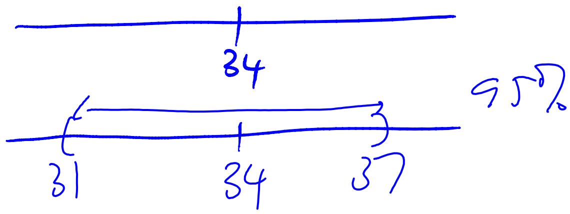
2013-10-22-EC01

Tuesday, October 22, 2013

6:18 PM

http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/LectureNotes/documents/2013-10-15-EC01.pdf

Ch. 7 Confidence Intervals (CI)



a) [7.1] z-based CI for μ (σ known)

Ex. Physicians' salaries

x (\$100,000)	$p(x)$
0	?
1	?
2	?
3	?
4	?

$\mu = ?$

$\sigma = 1.26$ (100K)

$n = 5$

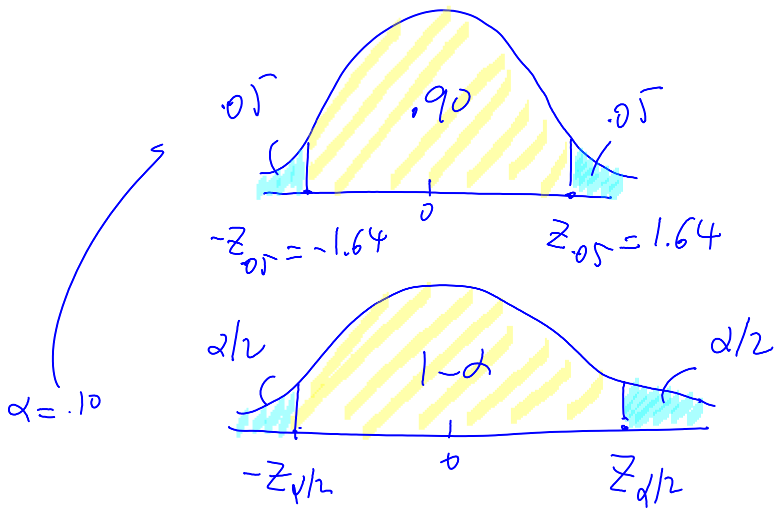
	Sample
x_1	3
x_2	0
x_3	3
x_4	0
x_5	3

$\bar{x} = 1.8$ (100K)

Recall ① Pop'n normal
or ② sample large } $\bar{X}$ normal

We want a 90% CI for μ

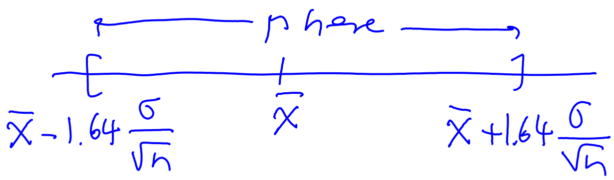
Recall 90% prob for Z



$Pr(-1.64 \le Z \le 1.64) = .90$

$Pr(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1-\alpha$

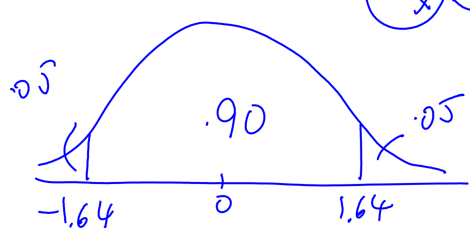
Claim: A 90% CI for μ is as follows



$1.8 - 1.64\frac{1.26}{\sqrt{5}} = .87$ $\quad = 2.72$

How? Recall $Z = \dfrac{X-\mu}{\sigma}$

$$Z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$



$$Pr(-1.64 \le Z \le 1.64) = .90$$

$$\downarrow$$

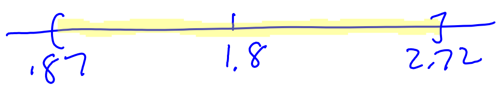
$$Pr\left(-1.64 \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le 1.64\right) = .90$$

$$Pr\left(-1.64\frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le 1.64\frac{\sigma}{\sqrt{n}}\right)=.90$$

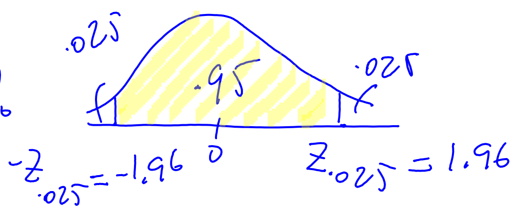
$$Pr\left(\bar{X}-1.64\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}+1.64\frac{\sigma}{\sqrt{n}}\right)=.90$$

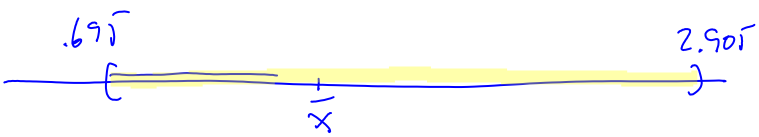
$$1.8-1.64\frac{1.26}{\sqrt{5}} \le \mu \le 1.8+1.64\frac{1.26}{\sqrt{5}}$$

$$.87 \le \mu \le 2.72$$

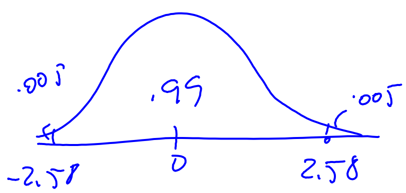


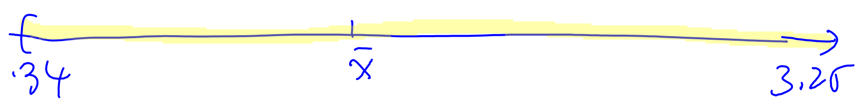
95%





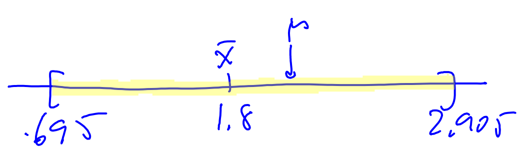
99%



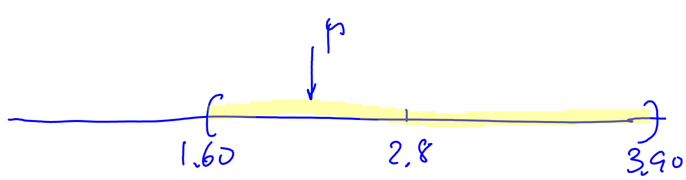


#2 Consider the 95% case & take new samples

Original $\bar{x}=1.8$
3, 0, 3, 0, 3

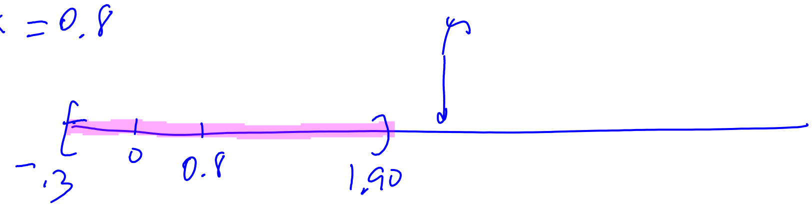


Another
4, 4, 3, 2, 1 : $\bar{x}=2.8$



Another

2, 0, 2, 0, 0 : $\bar{x} = 0.8$

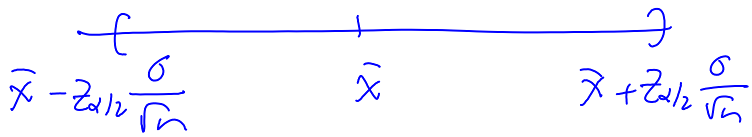


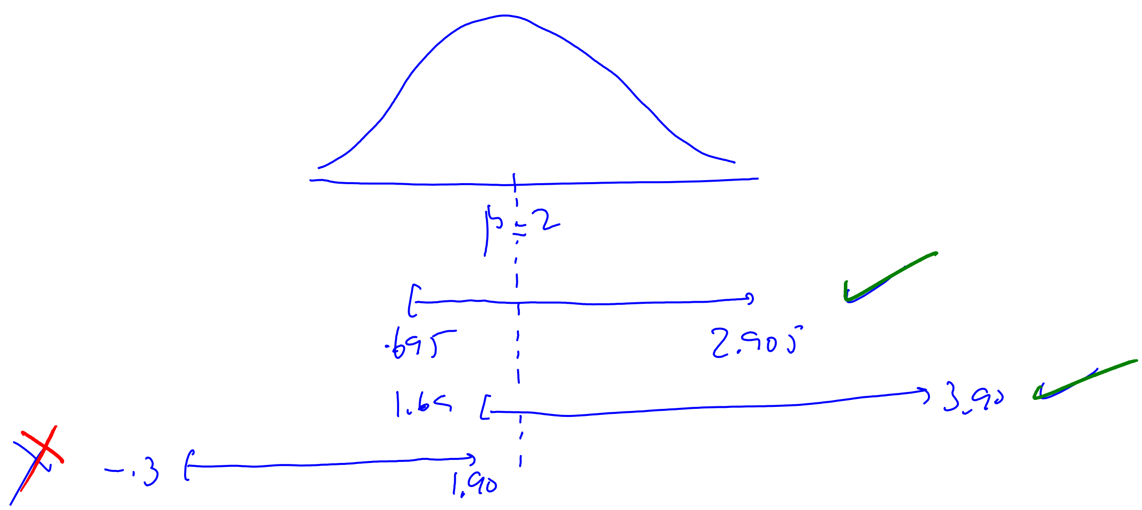
Reveal the distribution

x	$p(x)$
0	.1
1	.2
2	.4
3	.2
4	.1

$\mu = 2$

The interval $\left[\bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$ has $100(1-\alpha)\%$ chance of including the true mean μ





Ex. Visual Stat

http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/ChapterComments/CI-1.wmf

http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/ChapterComments/CI-2.wmf

(b) [7.2] t-based CI for μ (σ-unknown)

σ^2 unknown $\longrightarrow$ $S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

$$S = \sqrt{S^2}$$

σ known

$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \bar{x} \quad \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right)$: $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ Z-score

σ unknown

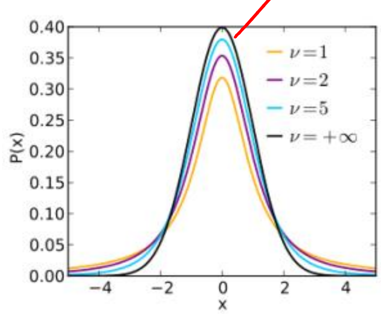
$\left(\bar{x} - (\cdot)\frac{s}{\sqrt{n}} \quad \bar{x} \quad \bar{x} + (\cdot)\frac{s}{\sqrt{n}} \right)$

t-score

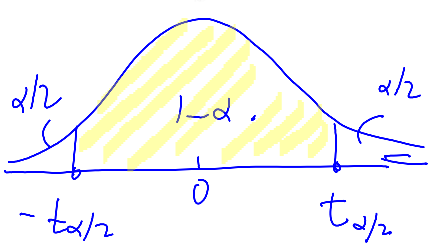
$$\cancel{Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}} \qquad t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$: t - with $n-1$ "degrees of freedom"

normal (0,1)
μ σ



$100(1-\alpha)\%$



$-t_{\alpha/2}$ $\qquad$ $t_{\alpha/2}$

$$\Big[\quad \bar{x} - (t_{\alpha/2})\frac{s}{\sqrt{n}} \qquad\qquad \bar{x} \qquad\qquad \bar{x} + (t_{\alpha/2})\frac{s}{\sqrt{n}} \quad\Big]$$

Ex. Sample of $n=6$ cars & mileage

18.8, 18.4, 19.2, 20.8, 19.4, 20.5

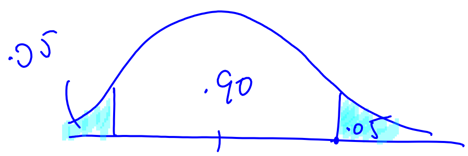
Find a 90% CI assuming pop'n normal

$\bar{x} = 19.48$

$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = 1.12$

$s = 1.06$, $\quad \frac{s}{\sqrt{n}} = \frac{1.06}{\sqrt{6}} = 0.43$

$df = n-1 = 6-1 = 5$



$-t_{.05}$ $\qquad$ $t_{.05} = 2.015$

http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/ChapterComments/documents/t-Table.pdf

df	$t_{.01}$	$t_{.05}$	$t_{.025}$	..
1				
2				
3				
4				
5		2.015		

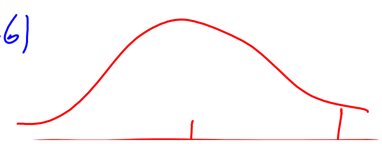
Error = E = .87

90% CI.

$$\Big[\quad \bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \qquad\qquad \bar{x} \qquad\qquad \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}} \quad\Big]$$

$19.48 - 2.015(.43) = 18.61$ $\qquad$ 19.48 $\qquad$ $19.48 + 2.015(.43) = 20.35$

If $\sigma = 1.06$



If $\sigma = 1.06$

$0 \qquad z_{.05} = 1.645$

z-results

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$[\ 18.77 \qquad 19.48 \qquad 20.19\ ]$

h3

c) [7.3] What's the best sample size n? (σ - known)

$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ (error)

$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad \bar{x} \qquad \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Fix E: Solve for n

$$n = \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

Ex. Taxable income

$n = 5, \quad \bar{x} = 1.8, \quad \sigma = 1.26$

$\alpha = .05, \quad 1-\alpha = .95, \quad z_{\alpha/2} = 1.96$

95% CI: $[\ .695 \qquad 1.8 \qquad 2.905\ ]$

$E = 1.105$ (\$110,500)

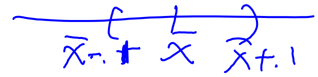
Want $E = .5$ (\$50,000)

$$n = \left(\frac{1.96\cdot 1.26}{0.5}\right)^2 = 24.39 \rightarrow 25$$

$\bar{x} - .5 \qquad \bar{x} \qquad \bar{x} + .5$

Want $E = .1$ (\$10,000)

$n = (\qquad) = 609$



Remark What if σ unknown?

- Take a preliminary sample of size m
- Estimate s
- Use $t_{\alpha/2}$ with $df = m-1$

$$n = \left(\frac{t_{\alpha/2} \cdot s}{E}\right)^2$$

Ex. Mileage $\qquad E = .87$ (before)

$m = 6$, $\quad s^2 = 1.12$

$df = 5$ $\quad s = 1.06$,

$t_{.05} = 2.015$

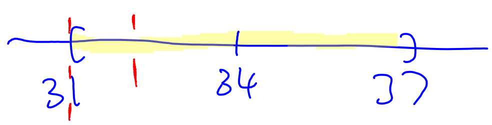
$E = .4 \qquad n = \left(\frac{t_{.05} \cdot s}{E}\right)^2 = \left(\frac{2.015 \times 1.06}{.4}\right)^2 = 28.51 \to 29$

d) [7.4] CI for a proportion p

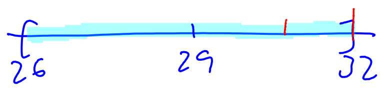
$\hat{p} = .34$ (Lib) $\qquad \mp 3\%$ margin of error

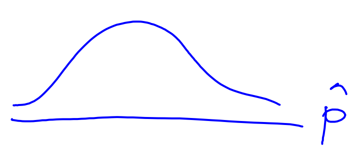
$19/20 = 95\%$

Lib



Tom





Last chapter

$\mu_{\hat{p}} = p$

$\sigma_{\hat{p}} = \sqrt{\frac{1}{n} p(1-p)} \to \sqrt{\frac{1}{n} \hat{p}(1-\hat{p})}$

$$\sigma_{\hat{p}} = \sqrt{\frac{1}{n} p(1-p)} \rightarrow \sqrt{\frac{1}{n} \hat{p}(1-\hat{p})}$$

$$\hat{p} - z_{\alpha/2}\sqrt{\tfrac{1}{n}\hat{p}(1-\hat{p})} \qquad \hat{p} \qquad \hat{p} + z_{\alpha/2}\sqrt{\tfrac{1}{n}\hat{p}(1-\hat{p})}$$

Globe (prop. of water to surface area)

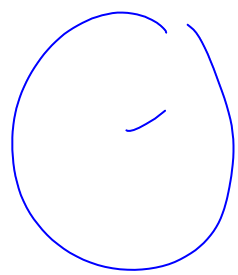
	W			L	
5	~~||||~~			~~||||~~	5
5	~~||||~~	21	8	|||	3
5	~~||||~~				
5	~~||||~~				
1	|				

$$\hat{p} = \frac{21}{29} = .72$$

Confidence interval - proportion		
	95%	confidence level
	0.72	proportion
	29	n
	1.960	z
	0.163	half-width
	0.883	upper confidence limit
	0.557	lower confidence limit

p

.557 .72 .883

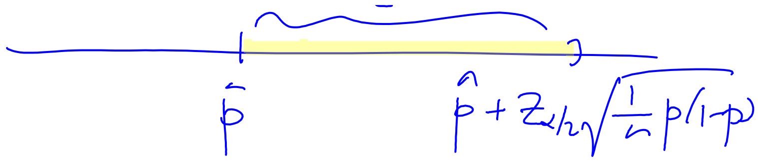


Q: Best sample size n for proportion estimation?

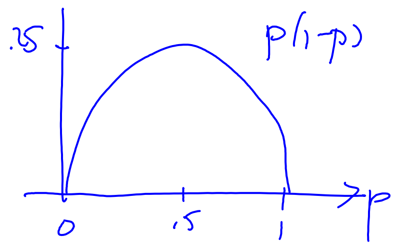
Correct interval $\left(\hat{p} \mp z_{\alpha/2}\sqrt{\frac{1}{n}p(1-p)}\right)$

E

$$\hat{p} \qquad \hat{p} + z_{\alpha/2}\sqrt{\tfrac{1}{n}p(1-p)}$$



$$E = z_{\alpha/2}\sqrt{\frac{1}{n}\underbrace{p(1-p)}_{.25}}$$



$$n = \underbrace{p(1-p)}_{.25}\left(\frac{z_{\alpha/2}}{E}\right)^2 = .25\left(\frac{z_{\alpha/2}}{E}\right)^2$$

Ex. Election polling

95% $\Rightarrow z_{\alpha/2} = 1.96$

- As of mid-June 2013, Liberals had the support of 34.2% of voters, Conservatives 29.4%, and NDP 25.3%.
- The article states that Nanos **surveyed 816 committed voters and the poll is accurate plus or minus 3.5 percentage points, 19 times out of 20.**
- So, we are 95% sure that the true proportion of Tory support is approximately somewhere between 30.7% (34.2 - 3.5%) and 39.3% (34.2 + 3.5%).
- We will see how 816 is obtained if desired "error" is plus or minus 3.5%.

Pasted from <http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/ChapterComments/ch-07.html>

$\mp 3.5\%$

$E = .035$

$$n = 0.25\left(\frac{1.96}{.035}\right)^2 = 784 \simeq 816$$

If $E = .01 \rightarrow n = 9{,}604$