2013-11-06-C01

Wednesday, November 06, 2013
12:46 PM

e) (7.5) CI for $\mu_1 - \mu_2$: Ind't samples
σ_1^2, σ_2^2 known

Recall (7.1)

Single pop'n $\bar{X}$: sample mean $\quad \mu_{\bar{X}} = E(\bar{X}) = \mu, \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ normal

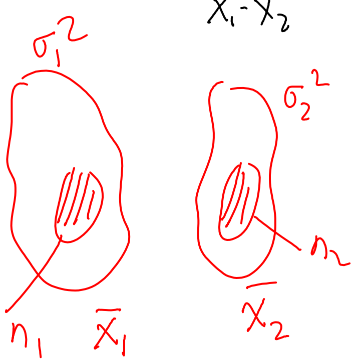
$100(1-\alpha)\%$ CI for μ : $\left(\bar{x} \mp z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$

Two pop'ns

$\bar{X}_1 - \bar{X}_2$:

$$\mu_{\bar{X}_1 - \bar{X}_2} = E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$



$100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$:

$$\left[(\bar{x}_1 - \bar{x}_2) \mp z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

Ex. Atkins vs. conventional diet

		Initial	6-month	Loss at			Initial	6-month	Loss at
	Diet	Weight (lbs)	Weight	6 Months (Atk)		Diet	Weight (lbs)	Weight	6 Months (Con)
1	Atkins	310	292.7	17.3	1	Conventional	256	259.5	-3.5
2	Atkins	309	275.1	33.9	2	Conventional	198	186.3	11.7
3	Atkins	257	217.7	39.3	3	Conventional	311	299.9	11.1
4	Atkins	227	221.1	5.9	4	Conventional	246	231.6	14.4
5	Atkins	231	204.5	26.5	5	Conventi	170	182.2	-12.2

						onal			
6	Atkins	195	148	47	6	Conventi onal	244	251.2	-7.2
7	Atkins	190	179.7	10.3	7	Conventi onal	179	180	-1
8	Atkins	164	131.1	32.9	8	Conventi onal	204	202.1	1.9
9	Atkins	190	162.7	27.3	9	Conventi onal	261	265.5	-4.5
10	Atkins	140	125.2	14.8	10	Conventi onal	271	260.3	10.7
11	Atkins	251	240.8	10.2	11	Conventi onal	265	263	2
12	Atkins	213	200.9	12.1	12	Conventi onal	262	266.8	-4.8
13	Atkins	211	210.5	0.5	13	Conventi onal	253	236.2	16.8
14	Atkins	257	241.8	15.2	14	Conventi onal	176	177.3	-1.3
15	Atkins	274	291.3	-17.3	15	Conventi onal	237	237.5	-0.5
16	Atkins	267	234	33	16	Conventi onal	170	149.1	20.9
17	Atkins	273	258.7	14.3	17	Conventi onal	169	168.9	0.1
18	Atkins	174	187.4	-13.4	18	Conventi onal	190	166.5	23.5

Pasted from <file:///C:\DOCUME~1\parlar\LOCALS~1\Temp\Atkins-vs-Conventional-Diet-Class.xls>

	6 Months (Atk)
count	33
mean	15.424
sample variance	206.528
sample standard deviation	14.371
minimum	-17.3
maximum	47
range	64.3

Pasted from <file:///C:\DOCUME~1\parlar\LOCALS~1\Temp\Atkins-vs-Conventional-Diet-Class.xls>

$n_1 = 33$

$\bar{x}_1 = 15.42$

$$\boxed{\sigma_1 = 8 \quad \text{Given} \quad \sigma_2 = 6}$$

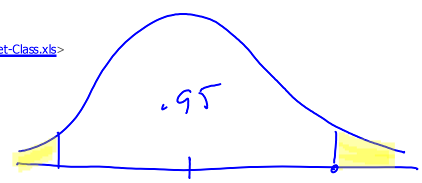
	6 Months (Con)
count	30
mean	7.007
sample variance	152.788
sample standard deviation	12.361
minimum	-12.9
maximum	36.5
range	49.4

Pasted from <file:///C:\DOCUME~1\parlar\LOCALS~1\Temp\Atkins-vs-Conventional-Diet-Class.xls>

$n_2 = 30$

$\bar{x}_2 = 7.00$

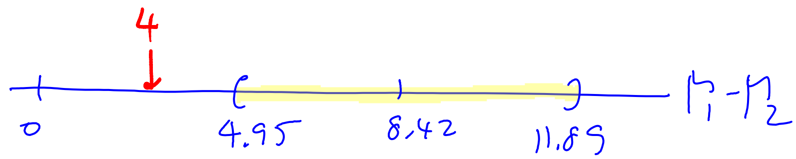
$\bar{x}_1 - \bar{x}_2 = 8.42$



$z_{.025} = 1.96$

$$\left[(\bar{x}_1 - \bar{x}_2) \mp z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

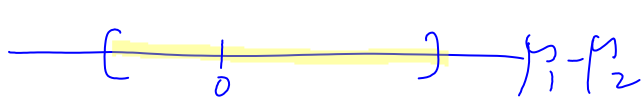
$$= \left[8.42 \mp 1.96 \sqrt{\frac{8^2}{33} + \frac{6^2}{30}}\right] = [4.95, 11.89]$$



$\Rightarrow$ Atkin's more effective: more likely that $\mu_1-\mu_2>0$

$\mu_1>\mu_2$

Suppose



maybe $\mu_1<\mu_2$ or $\mu_1>\mu_2$

f) [7.6] CI for $\mu_1-\mu_2$: Ind't sampls

σ_1^2, σ_2^2 unknown, but

equal $\rightarrow$ $\sigma_1^2=\sigma_2^2=\sigma^2$

Use t-tables

Pooled estimate for σ^2

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

CI:
$$\left[(\bar{x}_1-\bar{x}_2)\mp t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right], \quad df=n_1+n_2-2$$

Ex. Atkin's vs. conventional

A	C
$n_1=33$	$n_2=30$
$\bar{x}_1=15.42$	$\bar{x}_2=7.00$
$S_1=14.37$	$S_2=12.06$

So

So,

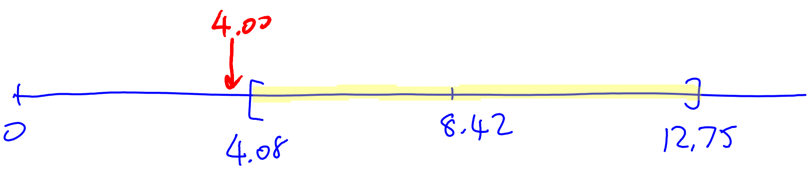
$$S_p^2 = \frac{(32)(14.37)^2 + (29)(12.36)^2}{33+30-2} = 73.72$$

$df = 33+30-2 = 61$, $\quad t_{\alpha/2} = t_{.025} = 2.00$

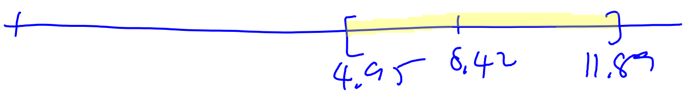
95% CI for $\mu_1 - \mu_2$

h2 $\left[8.42 \mp 2.00\sqrt{73.72\left(\frac{1}{33}+\frac{1}{30}\right)}\right] = [4.08, 12.75]$

σ_1^2, σ_2^2 unkn



σ_1^2, σ_2^2 known



Atkins appears more effective

Ch. 8 Hypothesis testing

$\underbrace{\text{Hypo}}_{\text{below}}\underbrace{\text{thesis}}_{\text{to be proven}}$

Ex. Coke vs. Pepsi experiment

	1	2	3	4	5	6	7	8	# correct
	C	C	C	C	P	P	P	P	
70	C	C	C	C	P	P	P	P	4+4=8
	C	P	C	C	P	C	P	P	3+3=6
	P	P	P	P	C	C	C	C	0+0=0
	⋮		⋮			⋮			

$$Pr(\text{all correct}) = \frac{1}{70} = 1.4\%$$

Rick got all right $\rightarrow$ reject hypothesis

Rick got all right → reject hypothesis that he's guessing

a) Null & alternative hypotheses
(z-test about μ, σ known)

Ex. legal process:
A person who's charged with a crime is innocent until proven guilty.

H_0: person innocent

H_a: " guilty

Actions. ① reject H_0
② accept (don't reject) H_0

		(innocent) H_0 true	(guilty) H_0 false
(find guilty)	Reject H_0	Type I error	Correct
(" innocent)	Accept H_0	Correct	Type II error

$$\alpha = Pr(\text{Type I error}) = Pr(\underbrace{\text{reject } H_0}_{\text{find guilty}} \mid \underbrace{H_0 \text{ true}}_{\text{innocent}})$$

$$\beta = Pr(\text{Type II }\text{ " }) = Pr(\underbrace{\text{accept } H_0}_{\text{find innocent}} \mid \underbrace{H_0 \text{ false}}_{\text{guilty}})$$

Steven Truscott & O.J. Simpson

Intuition for when to reject H_0

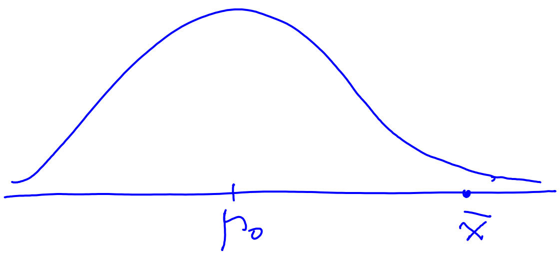
Suppose claim is $H_0: \mu \le \mu$ "°

Suppose claim is $H_0: \mu \le \mu_0 = 10$

$H_a: \mu > \mu_0 = 10$

Sample → sample mean $\bar{x}$, $\mu_{\bar{x}} = \mu_0$

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$



If $\bar{x}$ very high, reject

Ex. Tar content of cig's

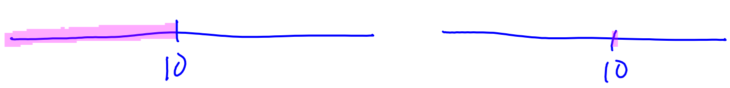
- Manuf claims $H_0: \mu \le 10$

  $H_a: \mu > 10$

  Suppose $n = 25$, and $\sigma = 3$ mg → given

  $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{25}} = .6$$

- Specify $\alpha = Pr(\text{Type I error}) = .05$
- Select a test statistic
  - Replace $H_0: \mu \le 10$ by $H_0: \mu = 10$



  - Find $\bar{x}$ (test statistic)

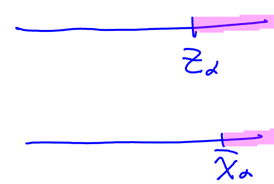
or - Use $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ (test statistic)

- Find a rejection point & region

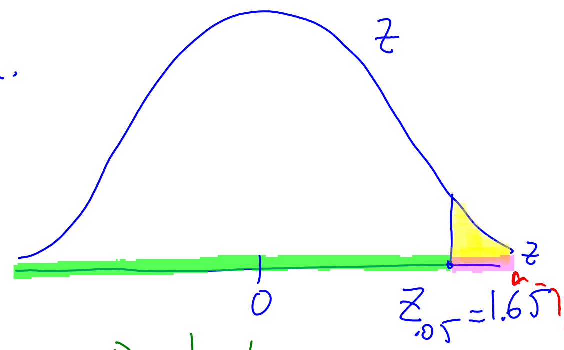
· Find a rejection point & region

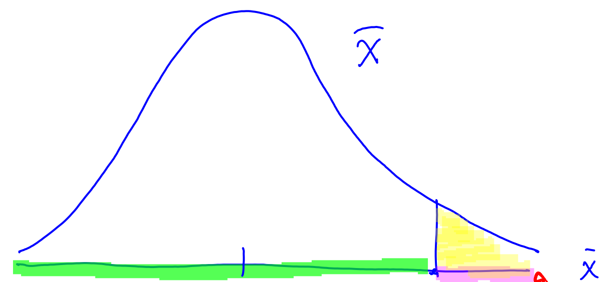
z_α

or $\bar{x}_\alpha$



Ex.





$\bar{x}_\alpha = 10.99$

$= \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$

$= 10 + 1.65(.6) = 10.99$

Reject

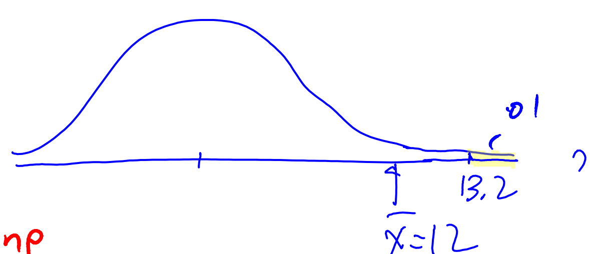
· Collect data

$n = 25, \; \boxed{\bar{x} = 12}$ Reject

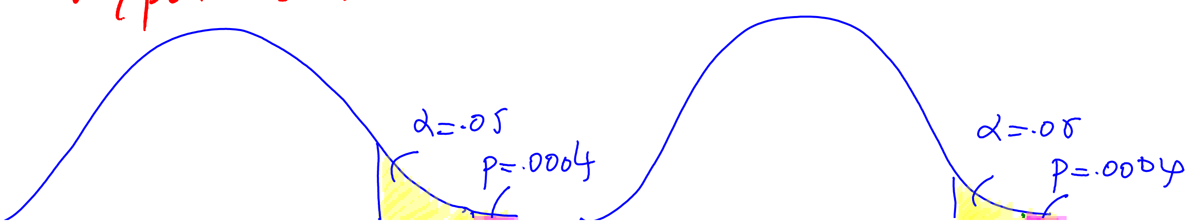
$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{12-10}{3/\sqrt{25}} = \boxed{3.33}$$

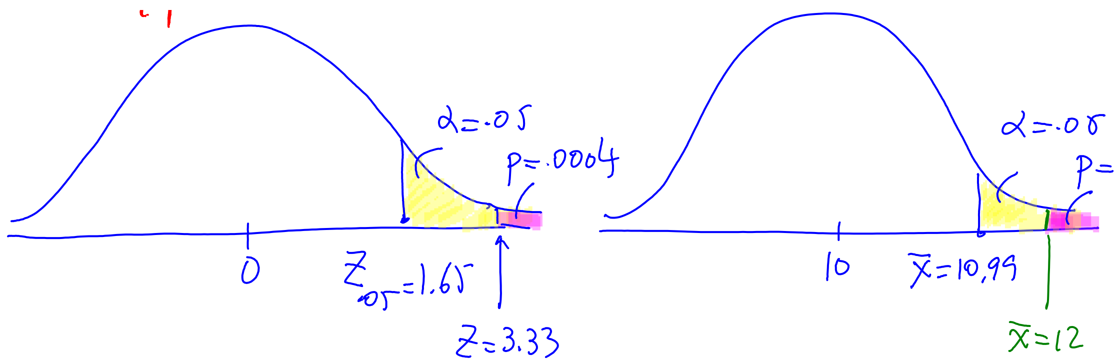
Conclusion: We reject claim that $\mu \le 10$

(If α is made smaller, rejection region gets smaller too, and this may result in not rejecting H_0)



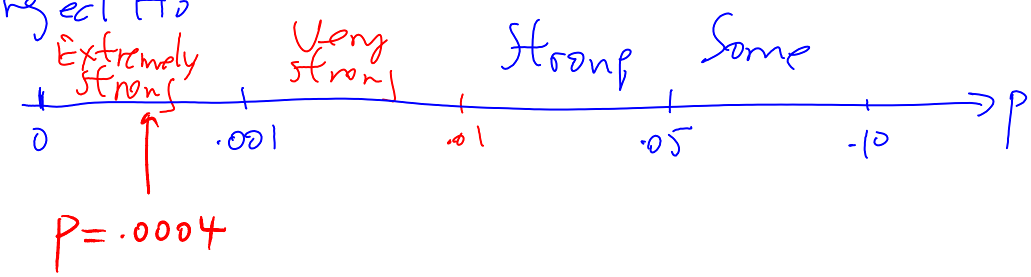
The p-value for testing hypotheses





(Small values of p, "bad news" for H_0)

Evidence to reject H_0



Summary

1. Find $\bar{x}$ and z (test stats)
2. Calculate p
3. Reject H_0 if p "small"