Chap 7 - Confidence Intervals

centered at statistic ($\bar{x}$ or $\hat{p}$)

margin of error $= t^*$ or $z^* \cdot$ (st. dev.)

↳ ½ width of the interval

(10, 20) width = 10

ME = 5

Statistic = 15

15 ± 5

larger level of confidence ⇒ wider interval

smaller n ⇒ wider interval } n is in denom.

bigger n ⇒ narrower interval

means

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

For C.I.

z^*

= invNorm or qnorm $\left(\frac{1+CL}{2}\right)$

decimal

t^*

$df = n-1$

= invT

qT $\left(\frac{1+CL}{2}, df\right)$

Finding n

means :

$$\underline{ME} > z^* \frac{s \text{ or } \sigma}{\boxed{\sqrt{n}}}$$

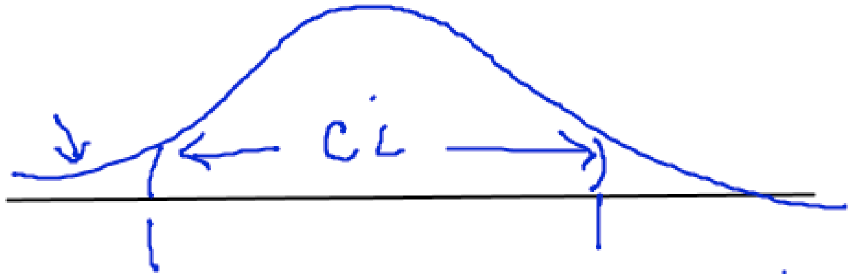
$$\sqrt{n} > \frac{z^* (s \text{ or } \sigma)}{ME}$$

$$n > \left(\frac{z^* (s \text{ or } \sigma)}{ME} \right)^2$$

round up.

$n > 56.01$

$n = 57$



z^* or t^* = invnorm
or invT of
area to left.

Hypothesis Testing

one sample – mean

$H_0: \mu = \mu_0$
$H_a: \mu \square \mu_0$

$\square$: $\neq$, $<$, $>$

proportion → $H_0: p = p_0$
$H_a: p \square p_0$.

difference (matched pairs)

↳ $H_0: \mu_D = 0$
$H_a: \mu_D \square 0$

dependent samples

two samples – means

$H_0: \mu_1 = \mu_2$
$H_a: \mu_1 \square \mu_2$

proportions

$H_0: p_1 = p_2$
$H_a: p_1 \square p_2$

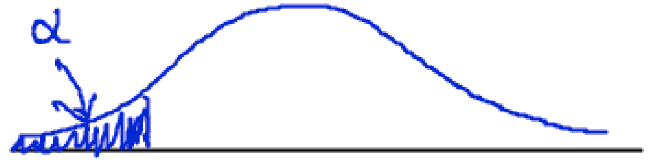
many samples – chi-square

H_0: distribution is same as claimed
H_a: distribution is different from claimed

⟨goodness of fit⟩

Alt. Hyp. | Rejection Region

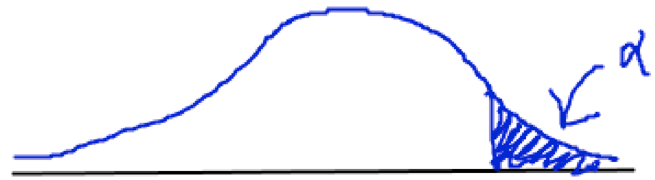
$<$



t^* or z^*

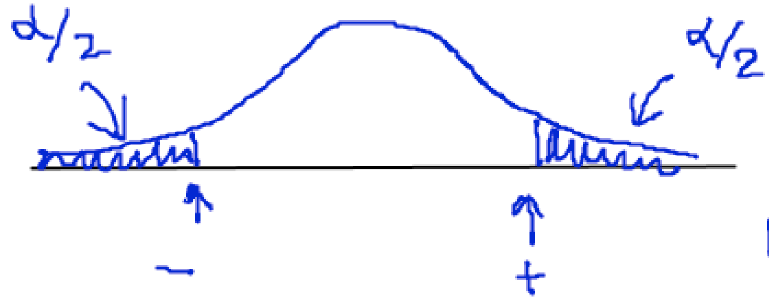
invnorm(α)

$>$



Invnorm($1-\alpha$)

$\neq$



Invnorm($\frac{\alpha}{2}$)

test statistic

Test	Null Hypothesis	Test Statistic
One-sample z-test for means	$\mu = \mu_o$	$z = \dfrac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$
One-sample t-test for means	$\mu = \mu_o$	$t = \dfrac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}}$; df = n–1
Matched Pairs t-test	$\mu_D = \mu_{D_0}$ to	$t = \dfrac{\bar{x}_D - \mu_D}{s/\sqrt{n}}$; df = n – 1
One-sample z-test for proportions	$p = p_o$	$z = \dfrac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$
Two-sample t-test for means	$\mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$	$t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$; df=min(n1,n2)-1
Two-sample z-test for proportion	$p_1 - p_2 = 0$ or $p_1 = p_2$	$z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}\right)}}$
χ^2 Goodness of fit test	no change	$\chi^2 = \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

p-value

Alt Hyp.	p value
$<$	p(z or t $<$ test statistic) normalcdf or tcdf pnorm or pt
$>$	p(z or t $>$ test statistic) normalcdf or tcdf 1 - pnorm or 1 - pt
$\neq$	2 · p(z or t $<$ if test stat is neg) 2 · p(z or t $>$ if test stat is pos)

$$p(z < -1.02) = \text{normalcdf}(-9999, -1.02)$$

$$\text{pnorm}(-1.02)$$

$$p(t > 2.6) = \text{tcdf}(2.6, 99999, df)$$

$$1 - \text{pt}(2.6, df)$$

Conclusion:

If pvalue $< \alpha \Rightarrow$ reject H_0

Based on $\underline{\alpha}$ level of significance, I will Reject the null hyp. which states __state null__ in favor of saying __state alt.__

If pvalue $> \alpha \Rightarrow$ fail to reject H_0

I will fail to reject H_0 which states ____

8.3

2 sample t-test   df = 29

2. Nitrites are often added to meat products as preservatives. In a study of the effect of these chemicals on bacteria, the rate of uptake of a radio-labeled amino acid was measured for a number of cultures of bacteria, some growing in a medium to which nitrites had been added. Here are the summary statistics from this study.

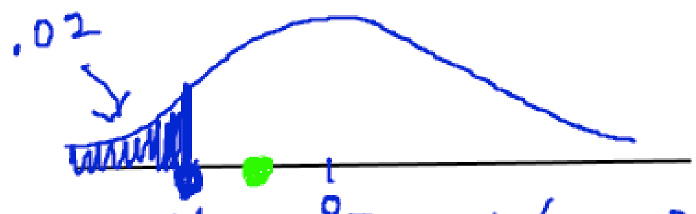
sample s.d.

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 < \mu_2$

	Group	n	$\bar{x}$	s
1	Nitrite	30	7880	1115
2	Control	30	8112	1250

Carry out a test of the research hypothesis that nitrites decrease amino acid uptake at the 2% significance level. — α



$t^* =$ invT or qt(.02, 29)

$= -2.15$

test stat:

$$t = \frac{7880 - 8112}{\sqrt{\frac{1115^2}{30} + \frac{1250^2}{30}}} = \boxed{-.7586}$$

pvalue: $p(t < -.7586) =$ tcdf(-9999, -.7586, 29)

pt(-.7586, 29)

$= .227 > \alpha$

Fail to reject H_0

4. Two methods were used to teach a high school algebra course. A sample of 75 scores was selected for method 1, and a sample of 60 scores was selected for method 2. The results are:

	Method 1	Method 2
Sample mean	85	83
Sample s.d.	3	2

Test whether method 1 was more successful than method 2 at the 1% level.

$n_1 = 75$

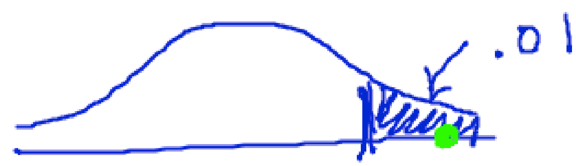
$n_2 = 60 \leftarrow \quad df = 59$

2 sample t test

$\alpha = .01$

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$



$t^* = \text{invT}(.99, 59)$
$= 2.39$

$$t = \frac{85 - 83}{\sqrt{\frac{3^2}{75} + \frac{2^2}{60}}} = \boxed{4.629}$$

pvalue: $p(t > 4.629) = \text{tcdf}(4.629, 99999, 59)$

$\boxed{p < \alpha \Rightarrow \text{Reject } H_0}$ $\quad 1 - pt(4.629, 59) = 1.03 \times 10^{-5}$

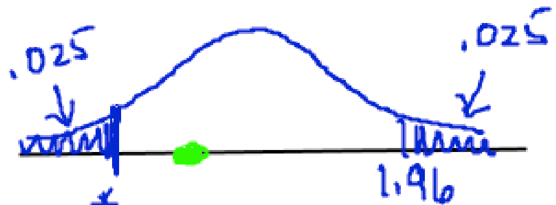
4. Data taken from a random sample of 60 students chosen from the student population of a large urban high school indicated that 36 of them planned to pursue post-secondary education. An independent random sample of 50 students taken at a neighboring large suburban high school resulted in data that indicated that 31 of those students planned to pursue post-secondary education. Do these data provide sufficient evidence at the 5% level to reject the hypothesis that these population proportions are equal?

2 prop. z-test

$\alpha = .05$

$H_0: p_1 = p_2$

$H_a: p_1 \neq p_2$



$z^* = \text{invnorm}(.025)$

-1.96

$\hat{p}_1 = 36/60 = .6$

$\hat{p}_2 = 31/50 = .62$

$$z = \frac{.6 - .62}{\sqrt{\frac{.6(.4)}{60} + \frac{.62(.38)}{50}}} = \boxed{-.214}$$

pvalue: $2p(z < -.214) = 2 \cdot \text{normalcdf}(-9999, -.214)$

$2 \cdot \text{pnorm}(-.214)$

$= .8305 > \alpha$

Fail to reject H_0

Public Page 12

2. The community hospital is studying its distribution of patients. A random sample of $\boxed{317}$ ← n patients presently in the hospital gave the following information:

OBS.

Type of patient	Old rate of occurrences of these types of patients	Present number of occurrences of these types of patients
Maternity ward	20%	65
Cardiac ward	32%	100
Burn ward	10%	29
Children's ward	15%	48
All other wards	23%	75

EXP

$.2(317) = 63.4$

$.32(317) = 101.44$

$.10(317) = 31.7$

$.15(317) = 47.55$

$.23(317) = 72.91$

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$\boxed{df = 4}$$

Using a 5% level of significance, test the claim that the distribution of patients in these wards has not changed.

$$\chi^2 = \frac{(65-63.4)^2}{63.4} + \frac{(100-101.44)^2}{101.44} + \cdots\cdots = \_\_\_$$

$$p(\chi^2 > \_\_\_) = \chi^2cdf(\_\_\_, 99999, df)$$

For questions 12-15: At Dream HS, teachers are informed that their grade distribution should model 20% A's, 30% B's, 20% C's, 15% D's, and 15% F's. Mrs. Concerned wonders if her 100 student's grades fit this pattern. Her students earned 26 A's, 34 B's, 30 C's, 6 D's, and 4 F's.

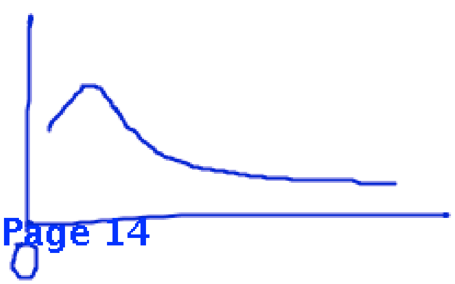
	A	B	C	D	F	total
OBS	26	34	30	6	4	= 100
EXP	20	30	20	15	15	mult by %

$$\sum \frac{(O-E)^2}{E} = \frac{(26-20)^2}{20} + \frac{(34-30)^2}{30} + \frac{(30-20)^2}{20} + \frac{(6-15)^2}{15} + \frac{(4-15)^2}{15}$$

$$\chi^2 = 20.8$$

$$p(\chi^2 > 20.8) = \left.\begin{array}{l} 1 - \text{pchisq}(20.8, 4) \\ \chi^2\text{cdf}(20.8, 99999, 4) \end{array}\right\} = .00035 < \alpha$$

Reject H_0



Ehw 13

13) A researcher wants to determine whether there is a difference between the proportions of males and females who believe in aliens. The sample information is:

	Males (1)	Females (2)
Sample size	75	100
# who said "yes"	50	45

2 prop z test

Based on the results of a significance test you should

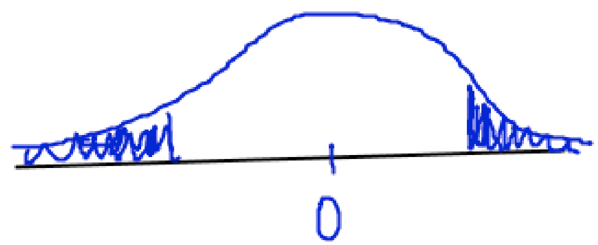
a) Reject H_0 at the 1% significance level ← if pvalue < .01

b) Fail to reject H_0 at the 1% significance level but reject H_0 at the 5% significance level ← if pvalue is between .01 and .05

c) Fail to reject H_0 at the 5% significance level ← if pvalue > .05

$H_0: p_1 = p_2$

$H_a: p_1 \neq p_2$



$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \boxed{}$$ ← Should be +

p value $2p(z > \boxed{})$

14. There are 4 TV sets in the student center of a large university. At a particular time each day, four different soap operas (1, 2, 3 and 4) are viewed on these TV sets. It is believed that the percentages of the audience captured by these shows are 25%, 30%, 25%, and 20%, respectively. 300 students are surveyed the summary of their response follows:

	1	2	3	4
Observed	80	88	79	53
Exp				

= 300

$$\sum \frac{(O-E)^2}{E} = \chi^2$$

Do these observed data fit the belief of percentages who watch each show ($\alpha = 0.05$)?

a) Yes ← fail to reject Ho
b) No ← reject Ho

$p(\chi^2 > \ )$

χ^2- goodness of fit

H_0: data fits expects

H_a: data doesnt fit

15) An association of college bookstores reported that the average amount of money spent by students on textbooks for the Fall 2010 semester was \$325.16. A random sample of 75 students at the local campus of the state university indicated an average bill for textbooks for the semester in question to be \$312.34 with a standard deviation of \$76.42. Do these data provide significant evidence that the actual average bill is different from the \$325.16 reported?

Based on the results of the significance test, you should

a) Reject H_0 at the 1% significance level

b) Fail to reject H_0 at the 1% significance level but reject H_0 at the 5% significance level

c) Fail to reject H_0 at the 5% significance level

one sample t-test

$\bar{x} = 312.34$

$s = 76.42$

$n = 75$

$H_0: \mu = 325.16$

$H_a: \mu \neq 325.16$

Question 8

You did not answer the question.

It has been observed that some persons who suffer renal failure, again suffer renal failure within one year of the first episode. This is due, in part, to damage from the first episode. The performance of a new drug designed to prevent a second episode is to be tested for its effectiveness in preventing a second episode. In order to do this two groups of people suffering a first episode are selected. There are 45 people in the first group and this group will be administered the new drug. There are 75 people in the second group and this group will be administered a placebo. After one year, 12% of the first group has a second episode and 14% of the second group has a second episode. Conduct a hypothesis test to determine, at the significance level 0.1, whether there is reason to believe that the true percentage of those in the first group who suffer a second episode is less than the true percentage of those in the second group who suffer a second episode? Select the [Rejection Region, Decision to Reject (RH_0) or Failure to Reject (FRH_0)].

a) [z > 1.28, RH0]

b) [z < -1.28, FRH0]

c) [z > -1.28 and z < 1.28, RH0]

d) [z < -1.28 ~~and~~ z > 1.28, FRH0]

e) [z < -1.28 or z > 1.28, FRH0]

$\alpha = .1$

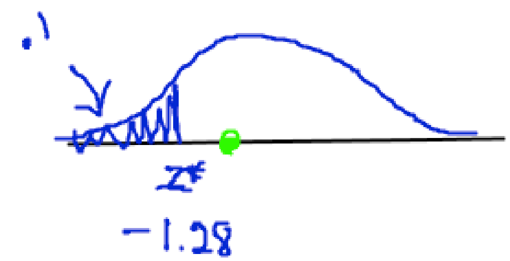
2 prop. z-test

① $n = 45$ ② $n = 75$

$\hat{p}_1 = .12$ $\hat{p}_2 = .14$

$H_0: p_1 = p_2$

$H_a: p_1 < p_2$



$$z = \frac{.12 - .14}{\sqrt{\frac{.12(.88)}{45} + \frac{.14(.86)}{75}}} = \boxed{-.318}$$

pvalue: $p(z < -.318)$

$= .375 > \alpha$