Math 2311

Bekki George – bekki@math.uh.edu

Office Hours: MW 11am to 12:45pm in 639 PGH

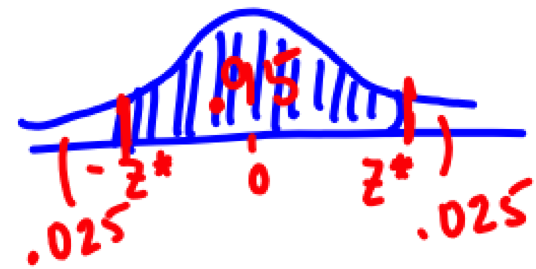
Online Thursdays 4-5:30pm

And by appointment

Class webpage: http://www.math.uh.edu/~bekki/Math2311.html

Confidence Intervals

proportions - $\hat{p}$ or $\hat{p}_1 - \hat{p}_2$



critical value (conf. level) = z^* (invNorm, qnorm)

tied to our level of confidence

CL

$\text{invNorm}\left(\frac{1+CL}{2}\right)$

ex: 95% conf. interval

$z^* = \text{invNorm}\left(\frac{1.95}{2}\right)$

bottom of t-table

1000			
z^*:	1.645	1.96	
	90%	95%	. . . . .

larger level of confidence means wider interval

$$\text{Statistic} \pm \underbrace{\text{margin of error}}$$

proportion(s): $\hat{p}$ or $\hat{p}_1 - \hat{p}_2$

or

mean(s): $\bar{x}$ or $\bar{x}_1 - \bar{x}_2$

$$= z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{or} \quad \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= z^* \cdot \sigma/\sqrt{n}$$ (know σ, the pop. std. dev)

or

$$\underbrace{t^*}_{\text{critical value}} \cdot \underbrace{s/\sqrt{n}}_{\text{standard error}}$$ sample std dev.

bigger $n \rightarrow$ narrower interval

Finding n given margin of erron

for proportions:

$$ME \geq z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{solve for } n$$

p or $\hat{p}$ or $.5$

Ex: Suppose we want the width of a conf. interval to be .3 w/ 95% level of confidence. How many should we sample?

$z^* = 1.96$

Use $p = .5$ $ME = .15$ $\left(\frac{.3}{2}\right)$

$$.15 \geq 1.96\sqrt{\frac{.5(1-.5)}{n}}$$

$$\frac{.15}{1.96} \geq \sqrt{\frac{.5(.5)}{n}} \quad \frac{\sqrt{.25}}{\sqrt{n}}$$

$$\frac{.15}{1.96}\sqrt{n} \geq .5$$

$$\sqrt{n} \geq \frac{.5(1.96)}{.15}$$

$\boxed{} \pm ME$

12 ± 5

$(7, 17)$

width $= 10$

$$n \geq \left(\frac{.5(1.96)}{.15}\right)^2$$

$n \geq 42.68$ need $n = 43$

$$n \geq \left(\frac{\sqrt{p(1-p)} \cdot z^*}{ME} \right)^2$$

☆ always round up

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

Popper 18

1. What does a 95% confidence interval tell us?
    a. 95% of the sample means have values between those endpoints.
    b. 95% of the population means have values between those endpoints.
    c. If the procedure were repeated many times, 95% of the resulting confidence intervals would contain the population mean.
    d. The probability that the population mean [redacted] is between the endpoints is .95

7.4 - Confidence Interval for a Population Mean

3. The effect of exercise on the amount of lactic acid in the blood was examined in an article for an exercise and sport magazine. Eight males were selected at random from those attending a week-long training camp. Blood lactate levels were measured before and after playing three games of racquetball, as shown in the accompanying table. Use this data to estimate the mean increase in blood lactate level using a 95% confidence interval.

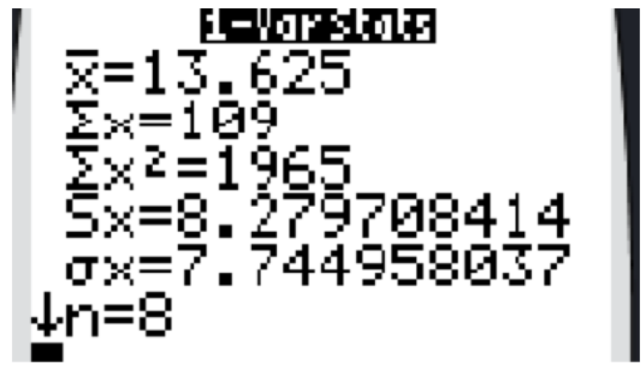
Player	1	2	3	4	5	6	7	8
Before	13	20	17	13	13	16	15	16
After	18	37	40	35	30	20	33	19

$n = 8$

$df = 7$

lact. incr 5, 17, 23, 22, 17, 4, 18, 3

assume normality



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

$$t^* = \text{invt}\left(\frac{1.95}{2}, 7\right) = 2.365$$

$$13.625 \pm 2.365 \cdot \frac{8.2797}{\sqrt{8}}$$

[6.70 , 20.545]

4. A 95% confidence interval for the mean of a population is to be constructed and must be accurate to within 0.3 unit. A preliminary sample standard deviation is 2.9. Find the he smallest sample size n that provides the desired accuracy.

$s \rightarrow$ have to use z^*

☆ since we dont have df to find t^*

$$ME = z^* \frac{s}{\sqrt{n}}$$

$$.3 \geq 1.96 \frac{2.9}{\sqrt{n}}$$

$$\sqrt{n} \geq \frac{1.96(2.9)}{.3}$$

$$n \geq \left(\frac{1.96(2.9)}{.3}\right)^2$$

$$n \geq 358.98 \qquad \boxed{n = 359}$$

7.5 - Confidence Interval for the Difference of Two Means

A confidence interval for two population means is used when you have two independent random samples and you wish to make a comparison of the difference ($\mu_1 - \mu_2$).

The assumptions that need to be satisfied are:

1. Both samples must be independent SRSs from the populations of interest.
2. Both sets of data must come from normally distributed populations. If this is not the case or if we are unsure whether the population is normally distributed, the sampling distributions of $\bar{x}_1$ and $\bar{x}_2$ must be normally distributed. (Recall from section 4.4 that we can assume that the sampling distribution of $\bar{x}$ is normal for values of *n* greater than 30.)

When the population standard deviations are known, we use the formula $\boxed{(\bar{x}_1 - \bar{x}_2) \pm z^* \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

and when it is unknown, we will need to find the sample standard deviations, s_1 and s_2, and use the formula $\boxed{(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ where t^* is the *t*-critical value based on the smaller of $n_1 - 1$ or $n_2 - 1$ degrees of freedom.

$1 = \text{men}$

Examples: $\sigma_1 = 3.6$ $\sigma_2 = 2.9$ $2 = \text{women}$

1. The height (in inches) of men at UH is assumed to have a normal distribution with a standard deviation of 3.6 inches. The height (in inches) of women at UH is also assumed to have a normal distribution with a standard deviation of 2.9 inches. A random sample of 49 men and 38 women yielded respective means of 68.3 inches and 64.6 inches. Find the 90% confidence interval for the difference in the heights of men at UH and women at UH.

$\hookrightarrow z^* = 1.645$

$n_1 = 49$ $n_2 = 38$

$\bar{x}_1 = 68.3$ $\bar{x}_2 = 64.6$

$$(\bar{x}_1 - \bar{x}_2) \pm z^* \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(68.3 - 64.6) \pm 1.645 \sqrt{\frac{3.6^2}{49} + \frac{2.9^2}{38}}$$

$$3.7 \pm 1.645(.696975505)$$

$$3.7 \pm 1.14656$$

$$[2.55, 4.85]$$

2. A researcher wants to see if birds that build larger nests lay larger eggs. He selects two random samples of nests: one of small nests and the other of large nests. He weighs one egg from each nest. The data are summarized below:

	Small nests	Large nests
Sample size	60 n_1	159 n_2
Sample mean (g)	37.2 $\bar{x}_1$	35.6 $\bar{x}_2$
Sample variance	24.7 $s_1 = 4.97$	39.0 $s_2 = 6.24$

Find the 95% confidence interval for the difference between the average mass of eggs in small and large nests.

t^* 95% CL w/ $df = 60 - 1 = 59$

$t^* = 2.00$

large - small

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

$$(35.6 - 37.2) \pm 2.00 \sqrt{\frac{24.7}{60} + \frac{39.0}{159}}$$

$[-3.22, 0.02]$

-1.6 ± 1.62

Popper 18

Suppose we compare the class averages for two classes on the same exams and get the following data:

Class	n	$\bar{x}$	s
A 1	25	88.4	4.3
B 2	36	86.7	1.9

t^*

2. What degrees of freedom will we use?
   a. 25
   b. 36
   c. 24
   d. 35
   e. none of these

$t^* = 2.0639$

3. Find the margin of error for a 95% CI
   a. 1.89
   b. 1.57
   c. 2.10
   ~~d.~~ 2.06
   e. none of these

$$\boxed{t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm$$

4. Find the 95% CI for the difference of these two means.
   a. [0.13, 3.27]
   b. [-0.19, 3.59]
   c. [0, 3.4]
   d. [-1.7, 1.7]
   e. none of these

5, A