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

4.1 - 4.3

		Total
5 m/c	8 pts	= 40 pts
3 f/R	15, 20, 25	= 60 pts
		100 pts

also quiz

+ .10 (practice test 2)

Written hw:

0 in grade calculator

→ check gradebook or Assignments

if blank ⇒ not graded

7.1 – Margins of Error and Estimates

What is estimation?

good guess

A **point estimate** is a single value that has been calculated from sample data to estimate the unknown population parameter.

Population Parameter	**Sample Statistic**
p - Population Proportion	$\hat{p}$ - Sample Proportion
μ - Population Mean	$\bar{x}$ - Sample Mean
σ - Population Standard Deviation	s – Sample Standard Deviation

Suppose we would like to make an estimate of a population parameter based on a sample statistic. A **confidence interval** is a range of possible values that is likely to contain the unknown population parameter that we are seeking.

First, we must have a **level of confidence**. Then, based on this level, we will compute a **margin of error** (we will discuss how to compute this in the next sections). Last, we can say that we are --% confident that the true population parameter falls within our confidence interval.

Formula for a confidence interval is: sample statistic ± margin of error

Example (problem 11 from text):

Suppose the heights of the population of basketball players at a certain college are in question. A sample of size 16 is randomly selected from this population of basketball players and their heights are measured. The average height is found to be 6.2 feet and the margin of error is found to be ±0.4 feet. If this margin of error was determined with a 95% confidence level, find and interpret the confidence interval.

$n = 16$ $\quad \bar{X} = 6.2$ $\quad ME = \pm .4$

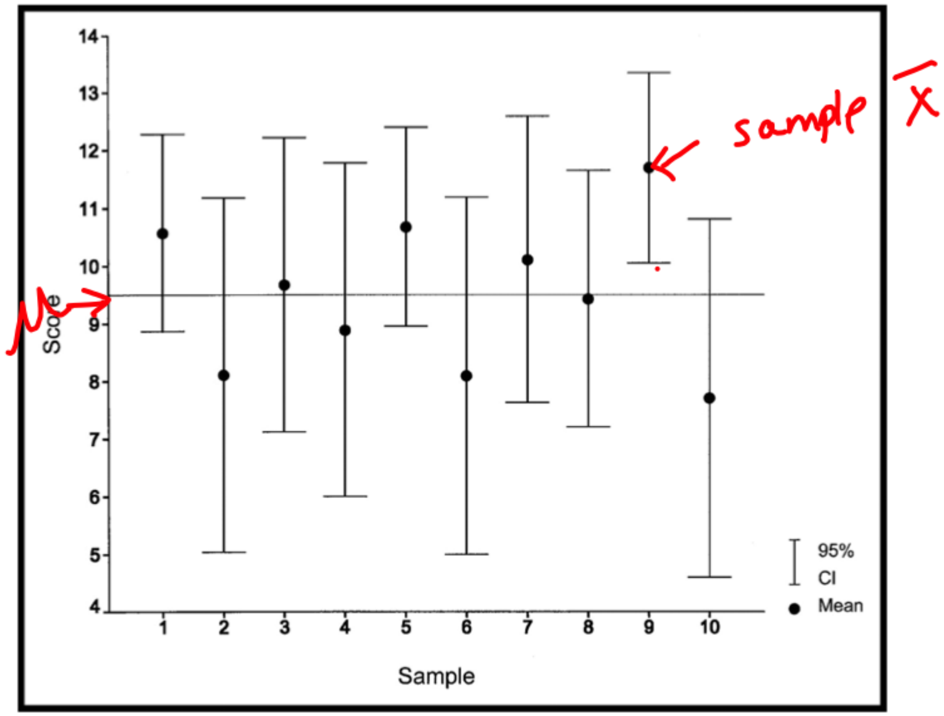
$6.2 \pm .4 = [5.8, 6.6]$

I am 95% conf. mean height of the population of bball players at a certain college is between 5.8 ft and 6.6 ft.

When we interpret the confidence level, what does this interpretation really mean?

What does it mean to be 95% conf?

If I repeated this procedure (taking sample and finding mean height) say 100 times, 95 of the intervals will contain the true population mean.



Popper 16

1. Which of the following is not a valid confidence interval centered at 15?

a. [10, 20]

10 15 20

15 ± 5

b. [2, 22]

c. [0, 30]

d. [11, 19]

e. none of these

7.2 - Confidence Interval for a Proportion

$\hat{p}$ - sample proportion

Before any inferences can be made about a proportion, certain conditions must be satisfied:

1. The sample must be an SRS from the population of interest.
2. The population must be at least 10 times the size of the sample.
3. The number of successes and the number of failures must each be at least 10 (both $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$).

The sample statistic for a population proportion is $\hat{p}$, so based on the formula for a CI, we have $\hat{p} \pm \textit{margin of error}$

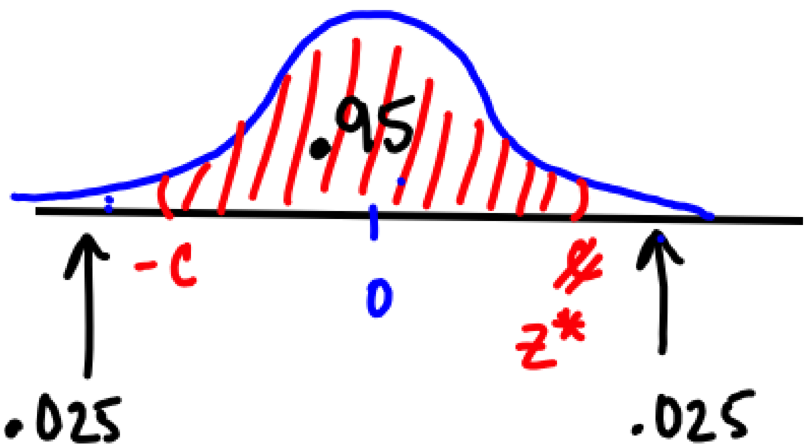
based on level of confidence

How do we find the $\boxed{\text{margin of error}}$ if it is not given to us?
The margin of error is equal to the **critical value** (a number based on our level of confidence) and the **standard deviation** (or **standard error** when needed) of the statistic.

Critical Value: When the distribution is assumed to be normal, our critical value is found from the z table (or using invNorm on calculator or qnorm in R). If it is not normal, we will use the t distribution (discussed later).

Standard Deviation/Error: When working with proportions, the standard deviation of the statistic $\hat{p}$ is $\sqrt{p(1-p)/n}$. Since p is unknown, we will use the standard error. To calculate the standard error of $\hat{p}$, use the formula $\boxed{\sqrt{\hat{p}(1-\hat{p})/n}}$.

95% confidence level



$P(-c < Z < c) = .95$

$.025 + .95 + .025 = 1$ ✓

$\text{invNorm}\left(\frac{1.95}{2}\right) = 1.96$

z^* is critical value $\quad z^* = \text{invNorm}\left(\frac{1+CL}{2}\right)$

$$\frac{2(.95)}{2} + \frac{1-.95}{2} = \frac{2(.95)+1-.95}{2} = \frac{.95+1}{2}$$

1.96

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

Examples:

In the first eight games of this year's basketball season, Lenny made 25 free throws in 40 attempts. $\leftarrow n$

a. What is $\hat{p}$, Lenny's sample proportion of successes? $25/40 = .625$

b. Find and interpret the 90% confidence interval for Lenny's proportion of free-throw success.

$$z^*_{.90} = 1.645$$

invNorm(.95)

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

$$.625 \pm 1.645 \sqrt{\frac{.625(1-.625)}{40}}$$

$$\text{St.dev} = \sqrt{\frac{p(1-p)}{n}}$$

$$.625 \pm .1259 = [.4991, .7509]$$

86% of ppl. like Coke from a SRS of 100 students at UH

find 92% conf. interval for all college students who like coke

Ck assumptions: SRS ✓

pop. > 10 times sample

$np = 100(.86) = 86 \geq 10$

$n(1-p) = 100(.14) = 14 \geq 10$ ✓

Calculate: $z^* = \text{invNorm}\left(\frac{1.92}{2}\right) = 1.75$

$s = \sqrt{\frac{.86(.14)}{100}} = .0347$

mult.

ME .0607

$.86 \pm .0607 = [.7993, .9207]$

Sometimes we are asked to find the minimum sample size needed to produce a particular margin of error given a certain confidence level. When working with a one-sample proportion, we can use the formula:

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

$$\therefore \overset{\text{max}}{ME} \geq z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

if $\hat{p}$ is unknown, use est. of p, if thats unknown use .5

Example:
It is believed that 35% of all voters favor a particular candidate. How large of a simple random sample is required so that the margin of error of the estimate of the percentage of all voters in favor is no more than 3% at the 95% confidence level?

est. p = .35

find n

ME = .03

$z^* = \text{invNorm}\left(\frac{1.95}{2}\right) = 1.96$

$$.03 \geq 1.96 \sqrt{\frac{.35(1-.35)}{n}}$$

$$\left(\frac{.03}{1.96}\right)^2 \geq \left(\sqrt{\frac{.35(.65)}{n}}\right)^2$$

$$n(.000234) \geq \frac{.2275}{n} \cdot n$$

$$.000234n \geq .2275$$

$$n \geq \frac{.2275}{.000234}$$

$$n \geq 972.2$$

$$\boxed{n = 973}$$

$$n \geq \frac{\overset{\text{or } \hat{p} \text{ or } .5}{\overbrace{p(1-p)}}}{(ME/z^*)^2}$$

$$n \geq \frac{.35(1-.35)}{(.03/1.96)^2}$$

$$n \geq 971.07$$

$$n = 972$$

Popper 16

2. What is the z* for a 90% confidence interval?

a. 1.96

b. 2.10

c. 1.645

d. 1.045

e. none of these

7.3 - Confidence Interval for the Difference of Two Proportions

The assumptions that need to be satisfied for a two-sample proportion are slightly different than those for a one-sample.

1. Both samples must be independent SRSs from the populations of interest.
2. The population sizes are both at least ten times the sizes of the samples.
3. The number of successes and failures in both samples must all be ≥ 10.

To make the comparison, we will need to find the difference of the two proportions, $\hat{p}_1 - \hat{p}_2$. The standard error for this difference is $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$. So our formula for the confidence interval is:

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

Monday

Example:

The National Research Council of the Philippines reported that 210 of 361 members in biology are women, but only 34 of 86 members in mathematics are women. Establish a 96% confidence interval estimate of the difference in proportions of women in biology and mathematics in the Philippines. Interpret your results.

Popper 16

$\hat{p}$ $n = 200$ $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

3. In an opinion poll, 35% of 200 people sampled said they were strongly opposed to the state lottery. The standard error of the sample proportion is approximately

a. .03 b. .25 c. .00094 d. 6.12 e. .06

4. A bag of M&Ms was randomly selected from the grocery store shelf, and the color counts were:

Brown	16	Red	11	Yellow	19
Orange	5	Green	7	Blue	3

Construct a 95% confidence interval for the probability of a brown M&M.

$n = 61$

$\hat{p} \pm z^*(s)$

$\hat{p} = 16/61 = .2623$

$z^* = 1.96$ $s = \sqrt{\frac{.2623(1-.2623)}{\cancel{200}\ 61}}$

a. (.20, .32)
b. (.13, .39)
c. (.37, .63)
d. (.15, .37)
e. none of these

5. The width of a confidence interval is dependent on the level of confidence.

a. True
b. False

PT 2 # 7

$p = .67$ $n = 256$

In a large population, 67% of the households have cable tv. A simple random sample of 256 households is to be contacted and the sample proportion computed. What is the mean and standard deviation of the sampling distribution of the sample proportions?

a) [0.67, 0.0512]
b) [0.1072, 0.0294]
c) [0.67, 0.0294]
d) [67, 0.0294]
e) [67, 0.0512]
f) None of the above

$$\mu_{\hat{p}} = p = .67$$

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

$$\mu = \mu_{\bar{x}} \qquad \sigma_{\bar{x}} = \sigma/\sqrt{n}$$