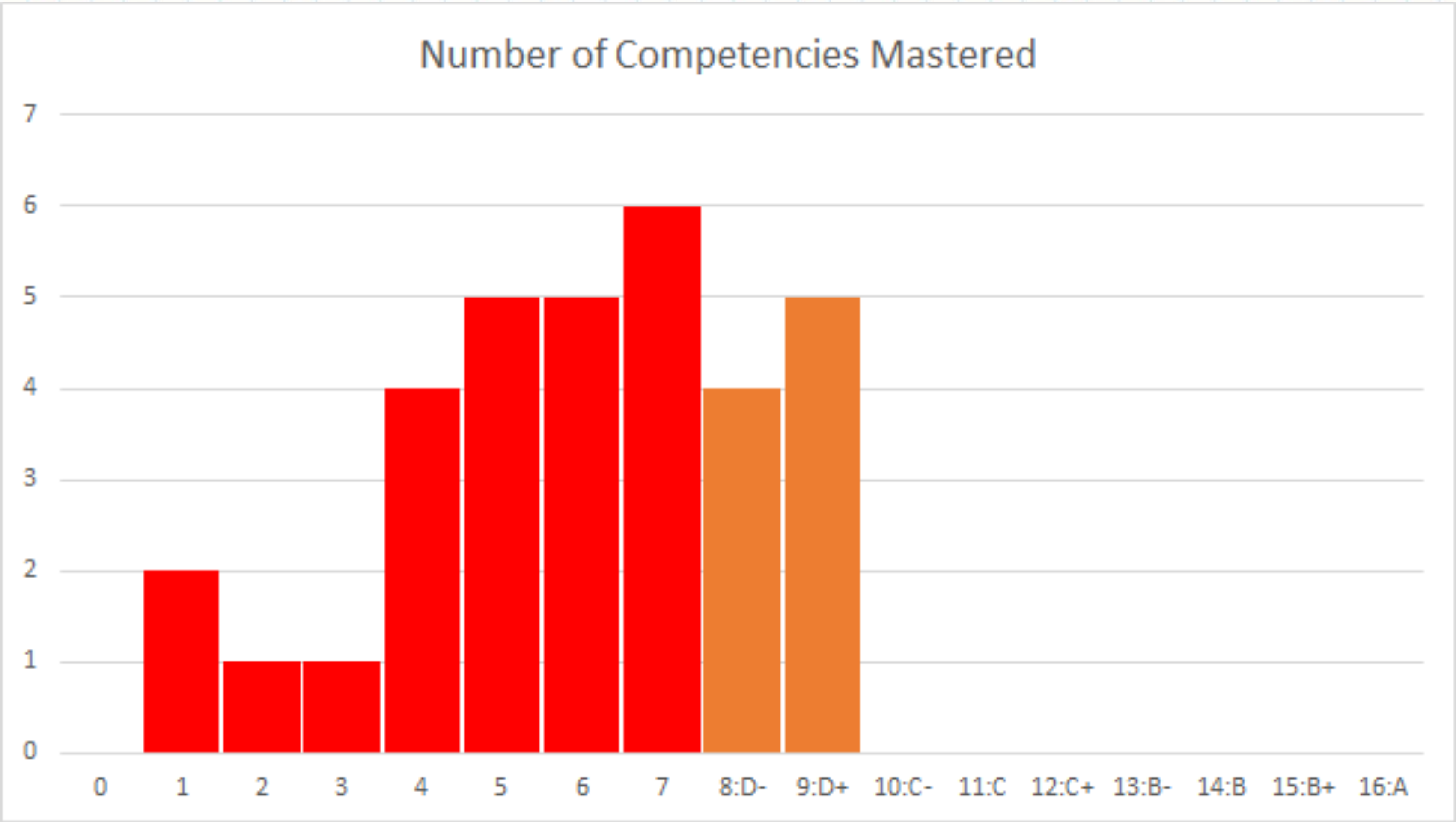
Confidence Intervals

Math 122



Next test: April 6, Competencies 11, 12, 13

Central Limit Theorem

Suppose that x is a random variable with mean μ and standard deviation σ.

Suppose that $\bar{x}$ is the mean of simple random samples of size n of values of x.

If x is normally distributed or if $n > 30$ then $\bar{x}$ is approximately normally distributed with

$$\mu_{\bar{x}} = \mu \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.$$

Central Limit Theorem
Version 2

Suppose that x is a random variable with mean μ and standard deviation σ.

Suppose that $\bar{x}$ is the mean of simple random samples of size n of values of x.

If x is normally distributed or if $n > 30$ then the distribution of

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

is approximately a standard normal distribution.

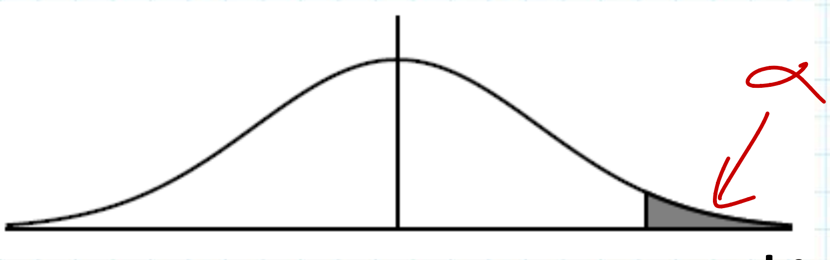
Central Limit Theorem
Other Versions

- There is a version of the CLT which allows binomial distributions to be approximated with normal distributions.
- There is a version of the CLT theorem for relating proportions to normal distributions.
- There is a version of the CLT for means which uses a Student t distribution. (Uses s instead of σ)

α Tails

- The number z_α is the unique number so that

$$P(z > z_\alpha) = \alpha$$



$z_\alpha = |\text{invnorm}(\alpha)|$

Natural Question

The best **single number estimate** or **point estimat**e of a population mean μ from a sample is the sample mean $\bar{x}$.

But how far can the sample mean $\bar{x}$ be from the population mean μ?

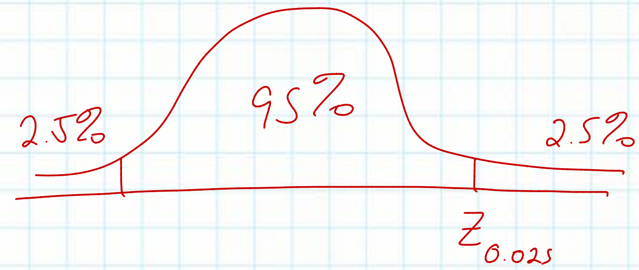
Interval Approximations

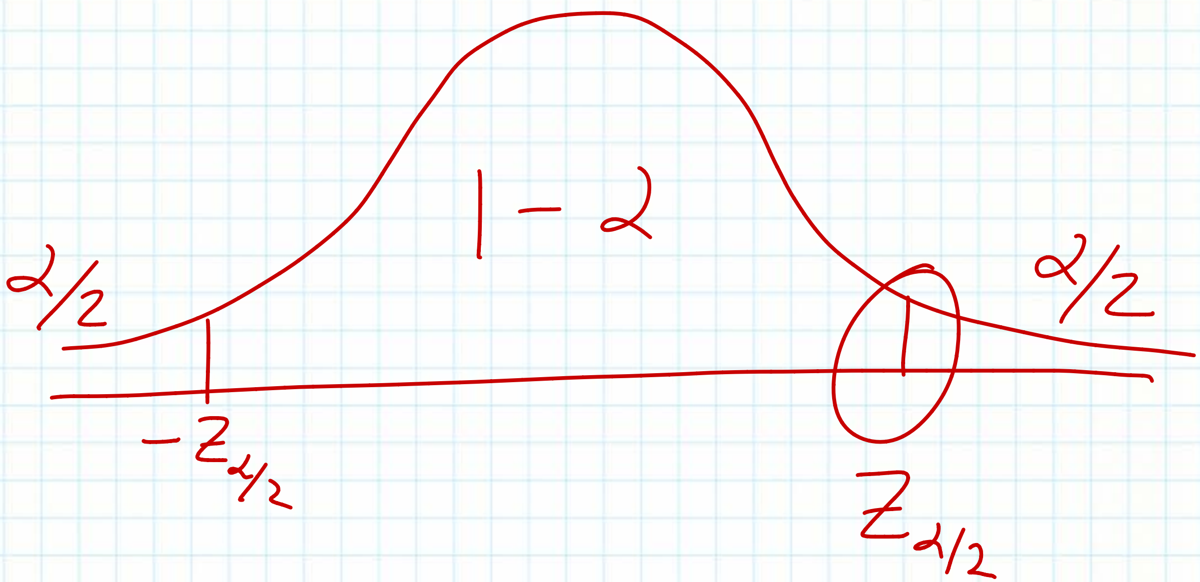
Since it is unlikely that a sample mean will be exactly equal to a population mean, we will approximate the population mean μ with an interval rather than a single number.

We want to be **reasonably confident** the actual mean is in the interval.

$$1 - \alpha = P(-z_{\alpha/2} < z < z_{\alpha/2})$$

95%





$$1-\alpha = P(-z_{\alpha/2} < z < z_{\alpha/2})$$

95%

$$1-\alpha = P\left(-z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right)$$

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

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

$$-\bar{x}$$

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

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

95%

$$1-\alpha = P\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

E – margin of Error

Confidence Interval:

$$\bar{x} - E < \mu < \bar{x} + E$$

Or

$$(\bar{x} - E, \bar{x} + E)$$

Nonsense

Instead of

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

We use

$$t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Ingredients for Confidence Intervals

- **Confidence Level**: CL (Usually 95%)
- **Significance Level**: $\alpha = 1 - CL$ (Usually 5%)
- **Critical Value**: $z_{\alpha/2} = |\text{invnorm}(\alpha/2)|$
- **Point Approximation**: $\bar{x}$ or $\hat{p}$
- **Margin of Error**: E
- **Interval**:

$$\bar{x} - E < \mu < \bar{x} + E$$

$$\hat{p} - E < p < \hat{p} + E$$

Formulas for Confidence Intervals

	Critical Value	Margin of Error	Interval
Means	$t_{\alpha/2} = -invt(\alpha/2, n-1)$	$t_{\alpha/2}\frac{s}{\sqrt{n}}$	$\bar{x} - E < \mu < \bar{x} + E$
Proportions	$z_{\alpha/2} = -invnorm(\alpha/2)$	$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$\hat{p} - E < p < \hat{p} + E$

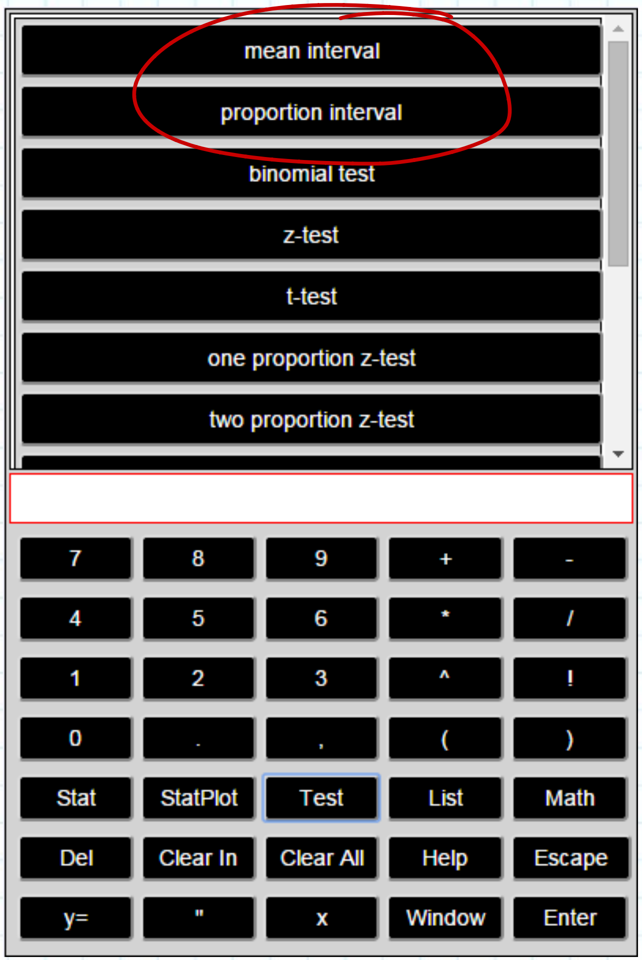
We will end with a statement such as:

We are 95% sure the population mean is between 10 and 12.

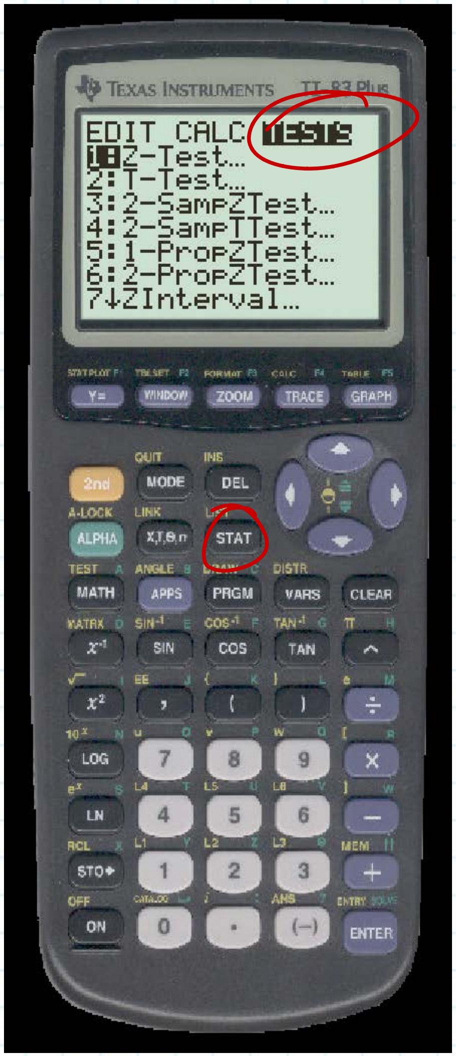
Which means…

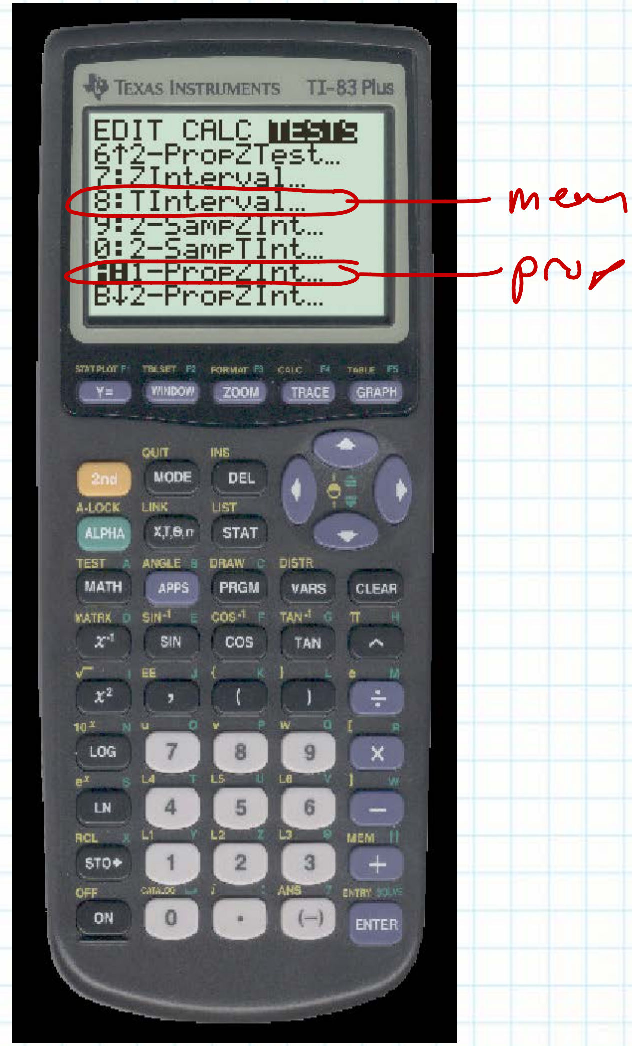
If the process that gave us this interval is repeated many times and many intervals are created, we expect 95% of the intervals to contain the actual mean.

We will use technology to calculate confidence intervals



We will use technology to calculate confidence intervals





A newscast reports that the President's approval rating is 43% with a margin of error of 4%. What is the confidence interval estimate of the proportion of voters that approve of the President's performance?

$\hat{p}$ E

Lower: $\hat{p} - E = 39\%$

Upper: $\hat{p} + E = 47\%$

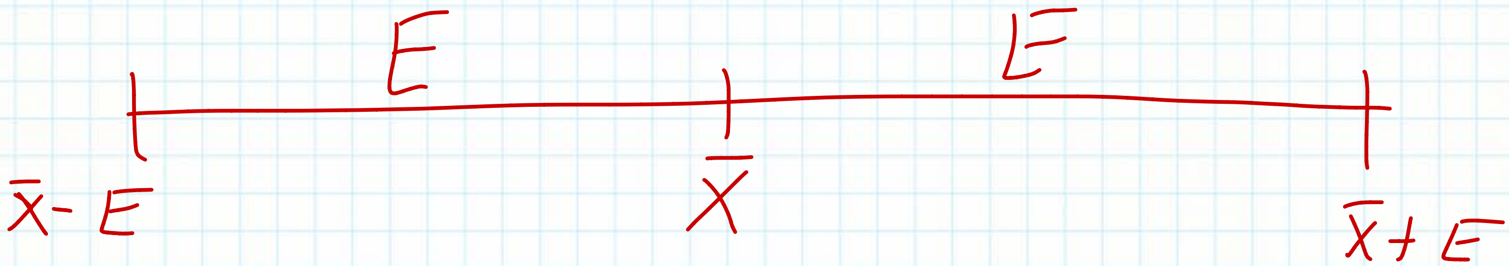
Inequalities: $39\% < p < 47\%$

Interval: $(.39, .47)$

A study of adult male heights gives a 95% confidence interval estimate of the mean adult male height of

$68.9 < \mu < 70.3.$

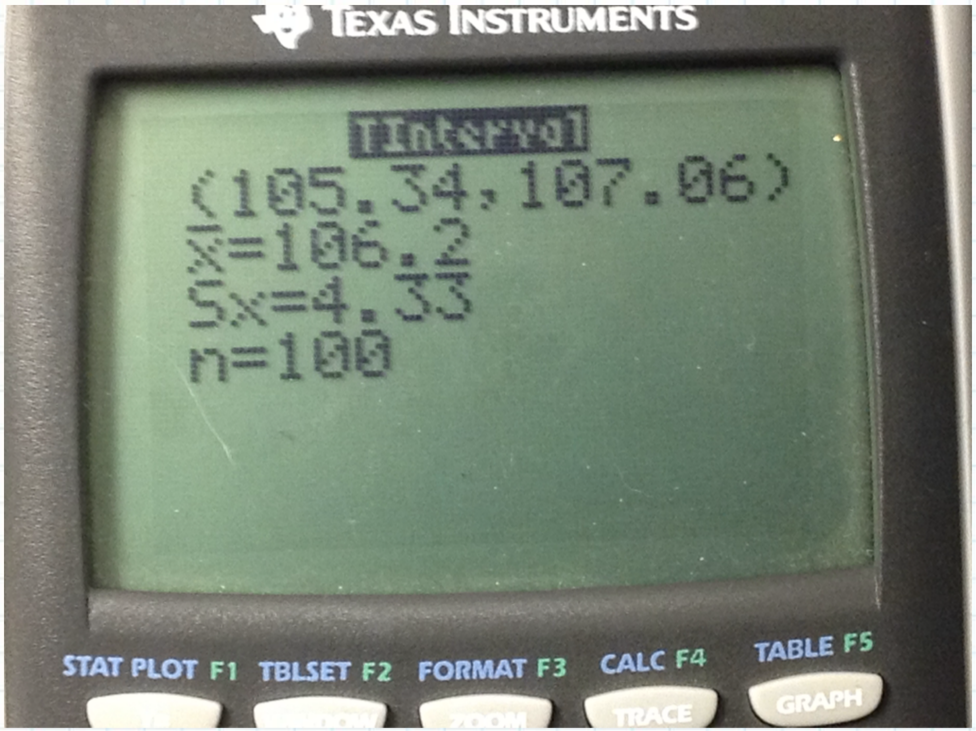
What were the margin of error and sample mean for this study?



$$\bar{x} = \frac{70.3 + 68.9}{2} = 69.6$$

$$E = 70.3 - 69.6 = 0.7$$

A sample of 100 pine needles had an average length of 106.2mm with a standard deviation of 4.33mm. Use this information to construct a 95% confidence interval estimate of the mean length of such a pine needle.



$E = 107.06 - 106.2$

$= 0.86$

A sample of 12 pine needles had these lengths (in mm):

107, 123, 99, 102, 106, 119, 115, 103, 111, 109, 102, 100

Use this data to construct a 95% confidence interval estimate of the average length of such a pine needle.

Enter in List

$\mu = 108$

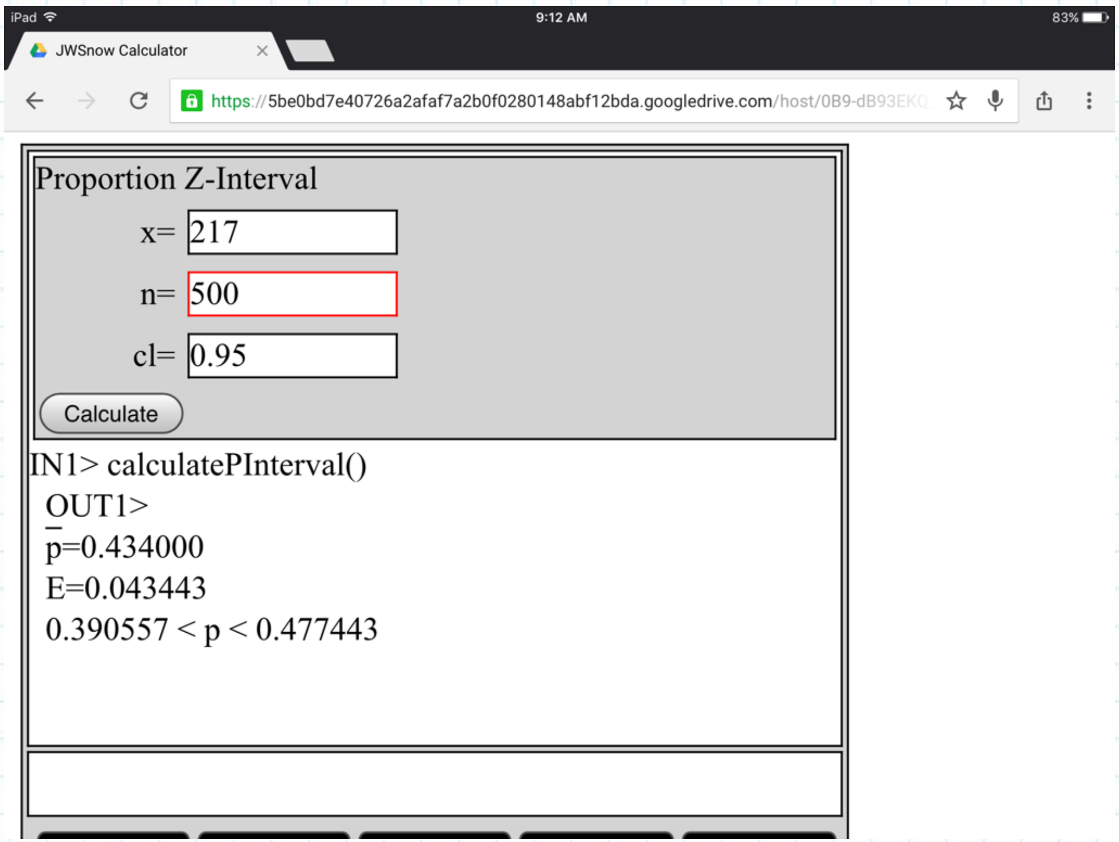
$S = 7.69$

Lower: 103.11

Upper: 112.89

$E = 4.89$

In a pre-election poll, 217 of 500 voters (or 43.4%) prefer candidate A. Use this information to construct a 95% confidence interval estimate of the proportion of voters that prefer candidate A.



In a pre-election poll, 241 of 500 voters (or 48.2%) prefer candidate B. Use this information to construct a 95% confidence interval estimate of the proportion of voters that prefer candidate B.

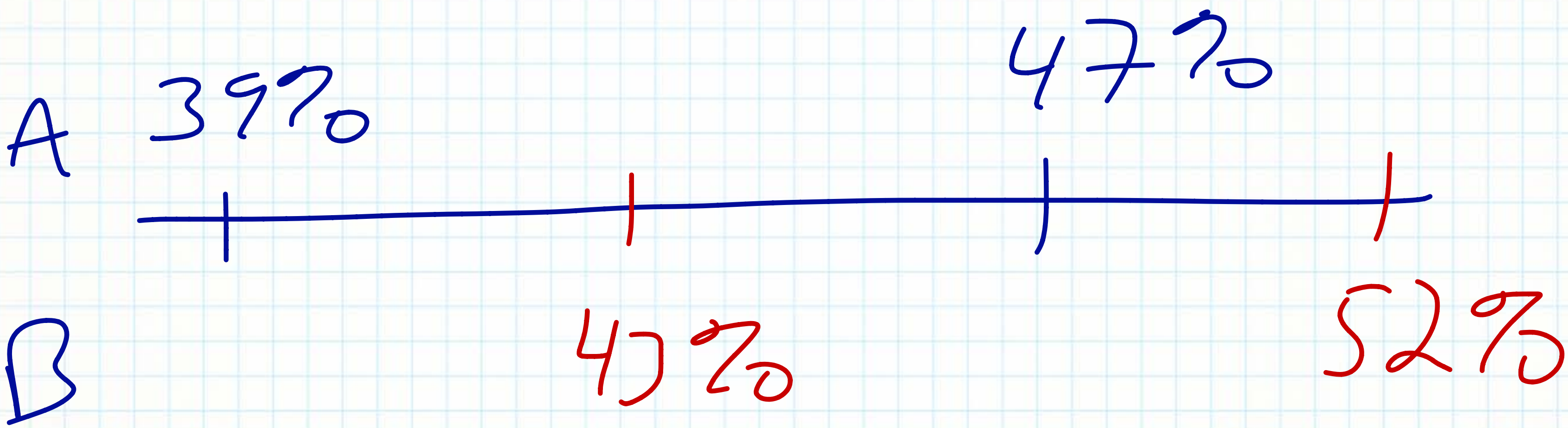
$$.438 \le p \le .526$$

$$E = .043$$

In the previous two examples, can it be said that one candidate is ahead of the other?

A: $.391 \leq p \leq .477$ 43.4%

B: $.438 \leq p \leq .526$ 48.2%



Statistical Tie

How large should a sample be to get a specified margin of error for the estimate of a proportion?

We can solve:

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

To get:

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

The largest $\hat{p}(1-\hat{p})$ can be is 1/4 so we use

$$n = \left(\frac{1}{4}\right)\left(\frac{z_{\alpha/2}}{E}\right)^2 = \left(\frac{1}{2}\right)^2\left(\frac{z_{\alpha/2}}{E}\right)^2 = \left(\frac{z_{\alpha/2}}{2E}\right)^2$$

Sufficient Sample Size

To approximate a proportion with margin of error no more than E use a sample size of

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

Know!

$\alpha = 1 - CL$

$z_{\alpha/2} = |\text{invnorm}(\alpha/2)|$

A poll is to be conducted to determine what proportion of the population prefers candidate A. If we want a margin of error of no more than 3%, then what sample size should be used for a confidence level of 90%? of 99%?

$CL = .9$

$\alpha = .1$

$E = .03$

$z_{\alpha/2} = |\text{invnorm}(\frac{.1}{2})| = 1.64$

$\left(\frac{1.64}{2 \times .03}\right)^2 = 747.1$

$n = 748$ (round up)

$CL = .99$

$\alpha = .01$

$E = .03$

$z_{\alpha/2} = |\text{invnorm}(\frac{.01}{2})| = 2.58$

$\left(\frac{2.58}{2 \times .03}\right)^2 = 1849$

$n = 1849$ (round up)