Sampling Distributions

5-2

Means

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}} \right)$$

5-4

Proportions

$$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}} \right)$$

- Mean
- Standard Deviation
- Normality - ?
- Calculate Probabilities

Remember that this just means that it has a normal distribution with mean μ

and standard deviation $\frac{\sigma}{\sqrt{n}}$

We will now focus on Sampling Distributions for Proportions

Sampling Distributions for Sample Proportions

- The **mean** of the sampling distribution is the same as the proportion of the population:

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

- The **standard deviation** of the sampling distribution gets smaller according to this equation:

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

- We assume **Normality** through three checkpoints:

  1. $np \geq 10$
  2. $n(1-p) \geq 10$
  3. SSSRTP (Sample Size Small Relative To Population) a.k.a. the 10% rule

- Calculate **probabilities** using *normalcdf*

Parameter - Greek alphabet

$$\mu, \sigma, p$$

…except for population proportion which is our alphabet

Statistic - Our alphabet

$$\bar{x}, s, \hat{p}$$

We will use statistics to approximate parameters.

			Sampling Distribution		
Variable	**Parameter**	**Statistic**	**Center**	**Spread**	**Shape**
Categorical (example: left-handed or not)	p = population proportion 5-5	$\hat{p}$ = sample proportion	p	$\sqrt{\frac{p(1-p)}{n}}$	Normal if np ≥ 10 and n(1 - p) ≥ 10
Quantitative (example: age)	μ = population mean, σ = population standard deviation 5-2	$\bar{x}$ = sample mean	μ	$\frac{\sigma}{\sqrt{n}}$	Normal if n > 30 (always normal if population is normal)

ALWAYS, I MEAN ALWAYS, CHECK YOUR ASSUMPTIONS!!!

WHAT DO I MEAN BY THIS?

A random sample of 100 students is taken from the population of all part-time students in the US, for which the overall proportion of females is 0.6 ← There's our p value

a) There is a 95% chance that the sample proportion ($\hat{p}$) falls between what two values?

$$0.6 - 2(0.049) < \hat{p} < 0.6 + 2(0.049)$$

$$0.502 < \hat{p} < 0.698$$

HERE WE ARE CHECKING OUR ASSUMPTIONS

$np = 100(0.6) = 60 \geq 10$ ✓

$n(1-p) = 100(1-0.6) = 40 \geq 10$ ✓

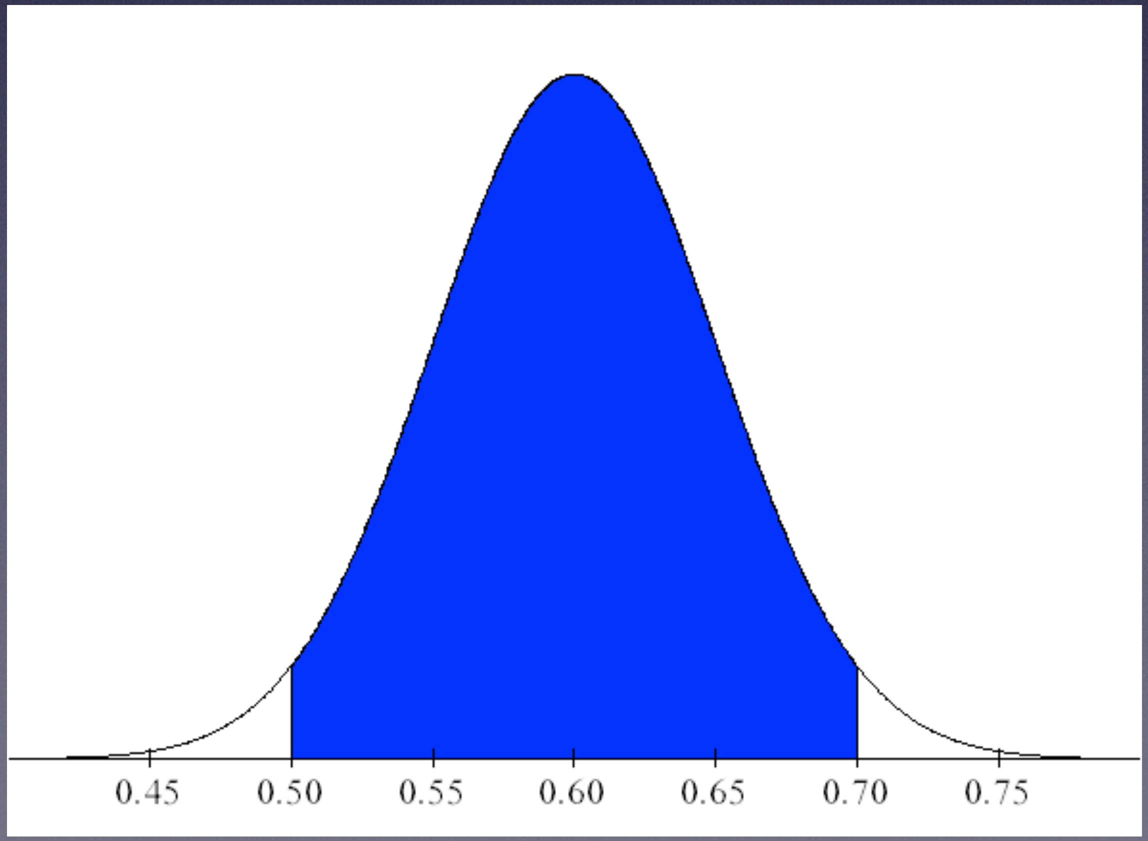
SSSRTP (Sample Size Small Relative To Population) a.k.a. the 10% rule

It's safe to say that 100 is less than 10% of the population ✓

So there is a 95% chance that between 50.2% and 69.8% of the sample will be women

Recalling The Empirical Rule (68%, 95%, 99.7%) we want to measure two standard deviations from the mean so

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.6)(1-0.6)}{100}} = 0.049$$



A random sample of 100 students is taken from the population of all part-time students in the US, for which the overall proportion of females is 0.6

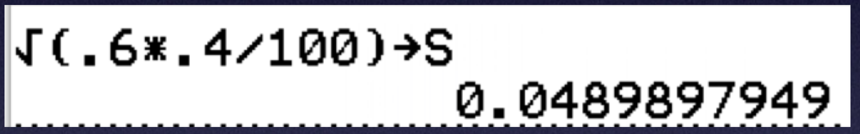
a) There is a 95% chance that the sample proportion ($\hat{p}$) falls between what two values?

b) What is the probability that the sample proportion is less than 0.56?

Since we are assuming the sample to have a normal distribution: $p = 0.6 \quad \sigma_{\hat{p}} = \sqrt{\dfrac{(0.6)(0.4)}{100}}$

$P(\hat{p} < 0.56)$ meaning what is the probability that the sample will be less than 56% women?

$$P(\hat{p} < 0.56) = \mathit{normalcdf}(-1E99, 0.56, 0.6, S)$$
$$= 0.207$$

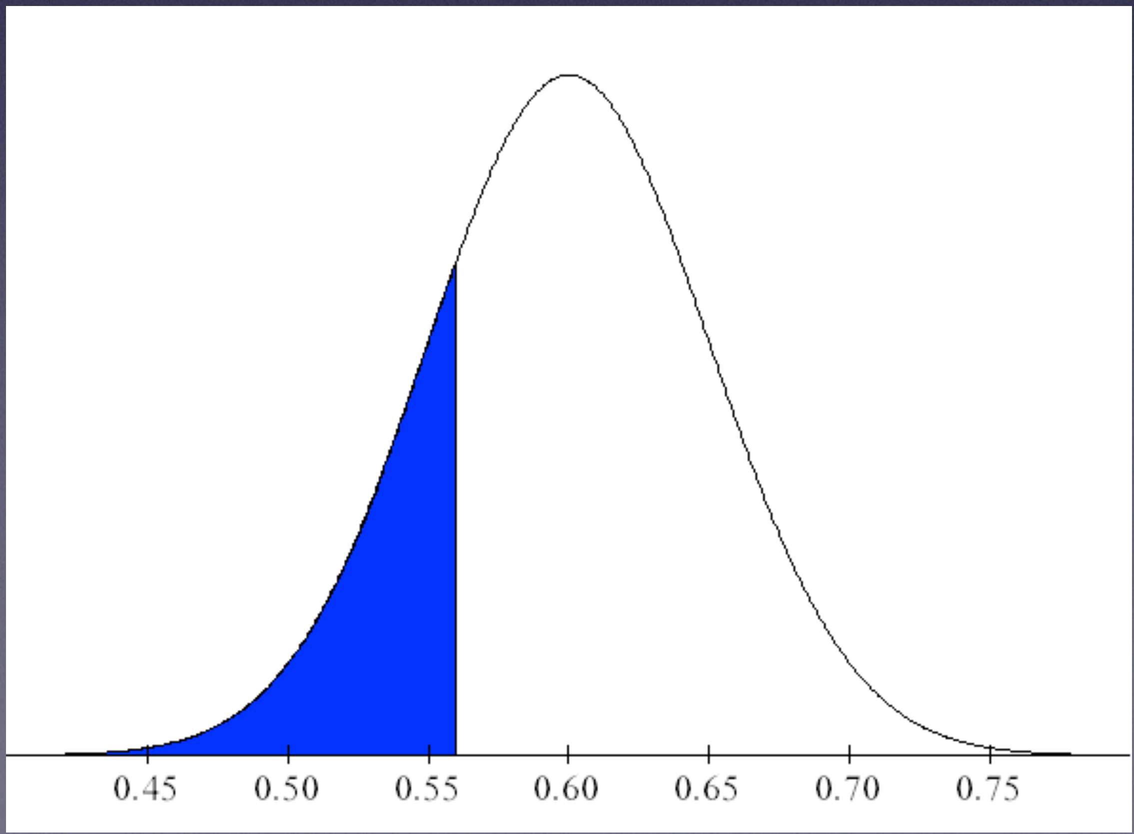


Either store the SD on the calculator (which I recommend) or at the very least round it to 0.049

This can also be done with *z*-scores and the standard normal curve

$$z_{0.56} = \frac{0.56 - 0.6}{0.049} = -0.816$$

$$\mathit{normalcdf}(-1E99, -0.816) = 0.207$$



A random sample of 2500 students is taken from the population of all part-time students in the US, for which the overall proportion of females is 0.6

a) There is a 95% chance that the sample proportion ($\hat{p}$) falls between what two values?

Bring your solutions to these questions to our next class

$$0.580 < \hat{p} < 0.620$$

b) What is the probability that the sample proportion is less than 0.56?

0.0000223