Lecture 16: Sample Size and Power

Chapter 4.6

Goals for Today

- More in depth discussion of
  - 10% sampling rule
  - Skew condition to check to use the normal model
- Central Limit Theorem
- How big a sample size do I need?
- Statistical Power

10% Sampling Rule

Question: Why do we have the rule that says our sample size n should be less than 10% of the population size N?

Intuition: Shouldn't we always sample as many people as we can?

Answer: Yes, if we only care about the mean. If we also care about the SE, then we need to be careful.

Explanation: Recall our example of sampling Wayne and Mario ($n = 2$) without replacement from a finite population of size $N = 4$ vs $N = 10000$ and its effect on independence.

10% Sampling Rule

The finite population correction (FPC) to the SE of point estimates accounts for the sampling without replacement:

$$SE = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

So say $N >> n$, $N = 10000$ and $n = 100$ (i.e. 1%), then

$$FPC = \sqrt{\frac{10000-100}{10000-1}} = \sqrt{\frac{9900}{9999}} = 0.995$$

but say $N = 10000$ and $n = 5000$ (i.e. 50%), then

$$FPC = \sqrt{\frac{10000-5000}{10000-1}} = \sqrt{\frac{5000}{9999}} = 0.707$$

10% Sampling Rule

The finite population correction (FPC) to the SE of point estimates accounts for the sampling without replacement:

$$SE = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

We've been ignoring the blue part in this class. So when

- $N >> n$, the correction is close to 1, so not a problem.
- As n gets bigger relative to N, factor shrinks: more and more of a problem.

Conclusion: Capping n to be less than 10% of the population, we have a rule of thumb for keeping the FPC “close” to 1.

10% Sampling Rule

We can tie the conceptual and mathematical notions of sampling:

Conceptual: If we sample everybody in our study population, then we don't need statistics because we know the true μ exactly.

and

Mathematical: If $n = N$, then the $FPC = \sqrt{\frac{N-n}{N-1}} = 0$, hence the corrected standard error is $SE = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}} = 0$.

i.e. there is no variability in our sampling procedure. If we repeat the procedure many times, we get the same value.

i.e. the sampling distribution is just one point: the true μ.

10% Sampling Rule

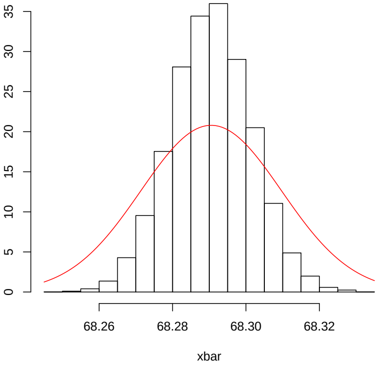
Question: Why do we care that our SE is correct?

Answer: If not, our normal model (i.e. z-score) based confidence intervals and hypothesis test p-values will be off the mark.

Example: From MATH392 where we sample $n = 40,000$ without replacement from a population of size $N = 60,000$

10% Sampling Rule

- The histogram represents the true sampling distribution
- The red curve represents the sampling distribution with the uncorrected $SE = \frac{s}{\sqrt{n}}$



Skew Condition to Check to Use Normal Model

Throughout the book, they talk about the condition for $\overline{x}$ being nearly normal and using s in place of σ in $SE = \frac{\sigma}{\sqrt{n}}$:

- On page 164: the population distribution is not strongly skewed
- On page 167: (CLT informal description) the data are not strongly skewed
- On page 168: the distribution of sample observations is not strongly skewed
- On page 185: the population data are not strongly skewed

Skew Condition to Check to Use Normal Model

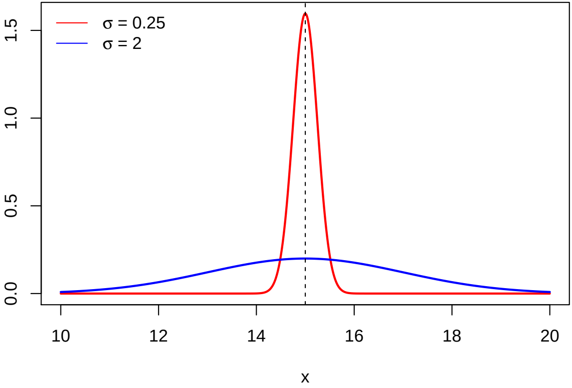
However, they all mean the same thing:

1. The true population distribution from which you are drawing your sample observations/data $x_1, \ldots, x_n$ is not too skewed.
2. The histogram of the sample observations/data $x_1, \ldots, x_n$ is not too skewed.

This skew is a problem that might affect the normality of $\overline{x}$ unless n is large.

Sample Size: Thought Experiment

Say you have two distributions with $\mu = 15$ but different σ.



Which of the two distributions do you think will require a bigger *n* to estimate μ "well"?

Margin of Error

Recall our formula for a 95% confidence interval:

$$\left[\overline{x} - 1.96\frac{s}{\sqrt{n}},\ \overline{x} + 1.96\frac{s}{\sqrt{n}}\right]$$

The margin of error is half the width of the CI. Say we knew the true standard deviation σ, then

$$\text{Margin of Error: } 1.96\frac{\sigma}{\sqrt{n}}$$

Identify n for a Desired Margin of Error

To estimate the necessary sample size n for a maximum desired margin of error m, we set

$$m \geq z^* \frac{\sigma}{\sqrt{n}}$$

where z^* is the critical value chosen to correspond to the desired confidence level. Ex. for a 95% CI, $z^* = 1.96$.

Solve for n.

Identify n for a Desired Margin of Error

Since

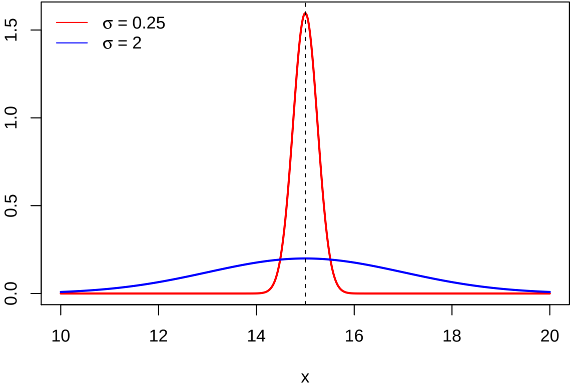
$$m \geq z^* \frac{\sigma}{\sqrt{n}}$$
$$\sqrt{n} \geq z^* \frac{\sigma}{m}$$
$$n \geq \left(z^* \frac{\sigma}{m}\right)^2$$

So

- As σ goes up, you need more n
- As z^* goes up, i.e. higher confidence level, you need more n
- As the desired margin of error goes up, you don't need as much n

Back to Thought Experiment

For the same desired maximal margin of error m and same confidence level, we need a larger n to estimate the mean of the blue curve:



Type II Error Rate and Power

The significance level α associated with a hypothesis test is the type I error rate: the rate at which we reject H_0 when it is true.

The type II error rate β is the rate at which we fail to reject H_0 when H_A is true. $1 - \beta$ is called the statistical power.

Statistical Power $=$ P(Rejecting H_0 when H_A is true)

Type II Error Rate and Power

Say we are conducting $N = A + B + C + D$ hypothesis tests.

		Test conclusion	
		do not reject H_0	reject H_0 in favor of H_A
Truth	H_0 true	A	B
	H_A true	C	D

- The Type I Error rate is the rate $\alpha = \frac{B}{A+B}$ at which B occurs given H_0 is true.
- The Type II Error is the rate $\beta = \frac{C}{C+D}$ at which C occurs given H_A is true.
- The power is the rate $1 - \beta = 1 - \frac{C}{C+D} = \frac{D}{C+D}$ at which D occurs given H_A is true.

Practical vs Statistical Significance

When rejecting the null hypothesis, we call this a statistically significant result. Note statistically significant results aren't always practically significant.

Say we suspect that using cell phones increases the risk of cancer by a multiplicative factor of θ. We have:

- $H_0 : \theta = 1$
- $H_A : \theta > 1$ (increased risk)

Say we have a large n, observe a $\widehat{\theta} = 1.000000001$, and reject the null hypothesis.

This tiny increased risk might translate to only one or two additional cases of cancer across the world i.e. not practically significancant.