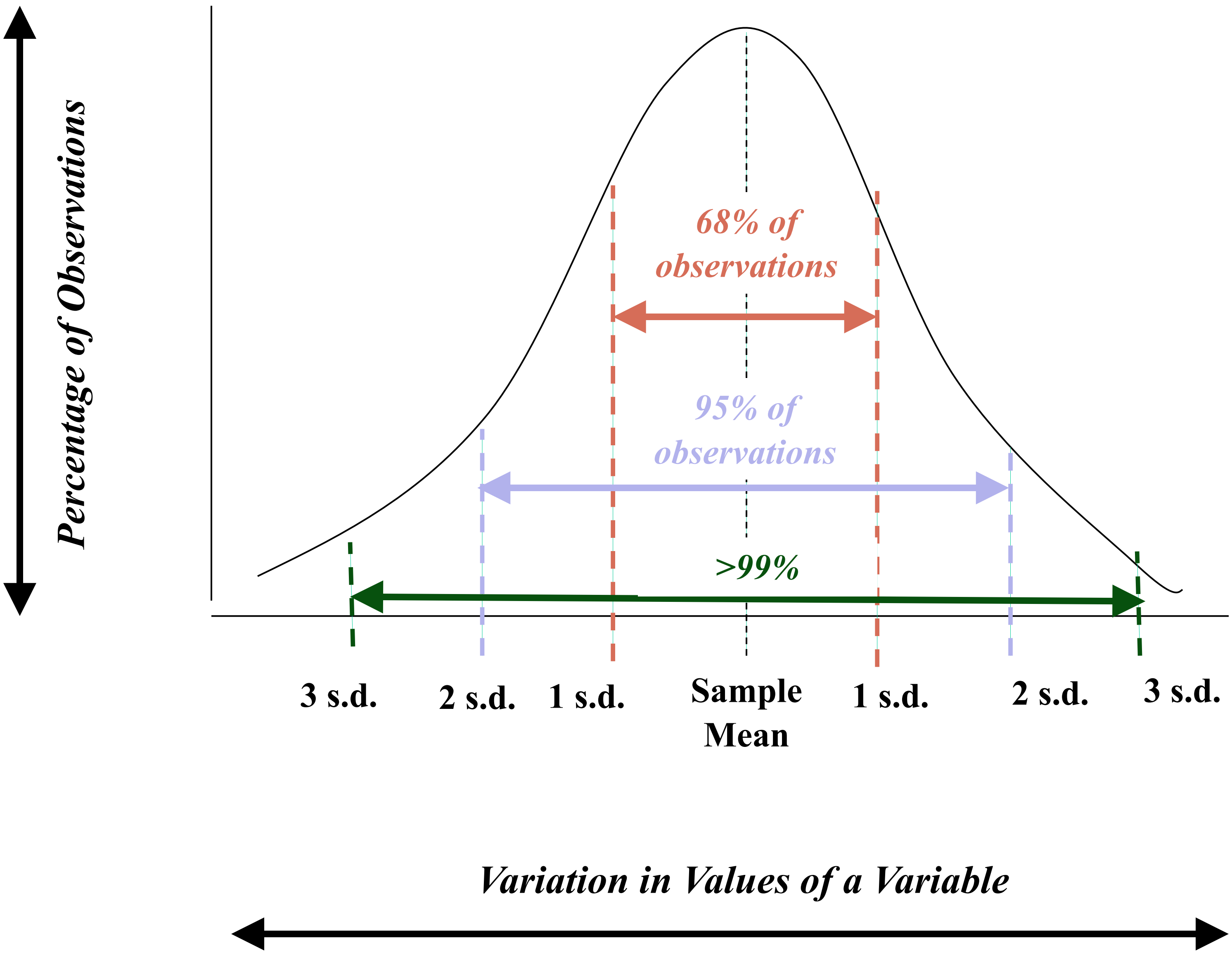
Inference

Outline

- Inferential statistics
- Sample statistics and population parameters – infinite samples, sampling distribution, uncertainty
- Central limit theorem
- Standard errors and confidence intervals
- Hypothesis testing
- Applications to proportions (Ch. 5, OpenIntro) and other areas.

Standard Deviation and the Normal Distribution



Standard Deviation

- Standard deviation measures the variation in the *sample* – how dispersed the values of the variable are in the sample.

- $s.d. = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N-1}}$

- Use properties of the normal distribution to understand inference, sampling, and *sampling* distributions.

Dispersion and Inference

- These properties are also the heart of **inferential statistics**.
  - *Using information from sample ("known facts") to make an inference about the population ("unknown facts").*

- In inferential statistics, we make inferences based on the notion of and theory behind *repeated sampling*.

- If we were to take an infinite number of random samples, what would the distribution of the *statistics* across those samples look like?

- In other words, if we estimated a mean of each of the infinite samples, what would the distribution of those means look like?

Central Limit Theorem

- The **central limit theorem** *is the foundation of inferential statistics*. ***If we were to take an infinite number of samples of the same sample size, the means (and other statistics) from these samples would be normally distributed.***

- The distribution of a *statistic* (e.g., $\bar{x},\ \hat{p}, or\ \hat{\beta}$) based on a hypothetical infinite sampling process, is called a ***sampling distribution***.

- The standard deviation of a sampling distribution is called a ***standard error***.

- Importantly: The mean of the sampling distribution is the true population parameter, μ.

- Illustration….

Illustration of CLT and Sampling Distribution

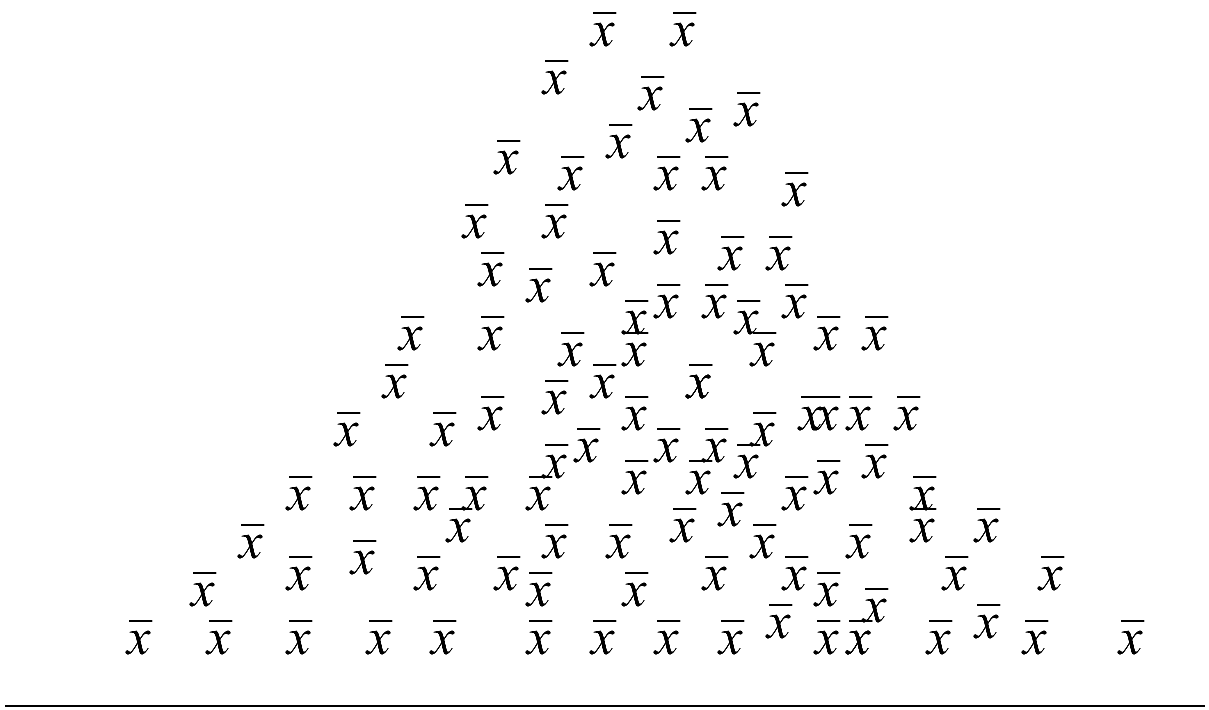
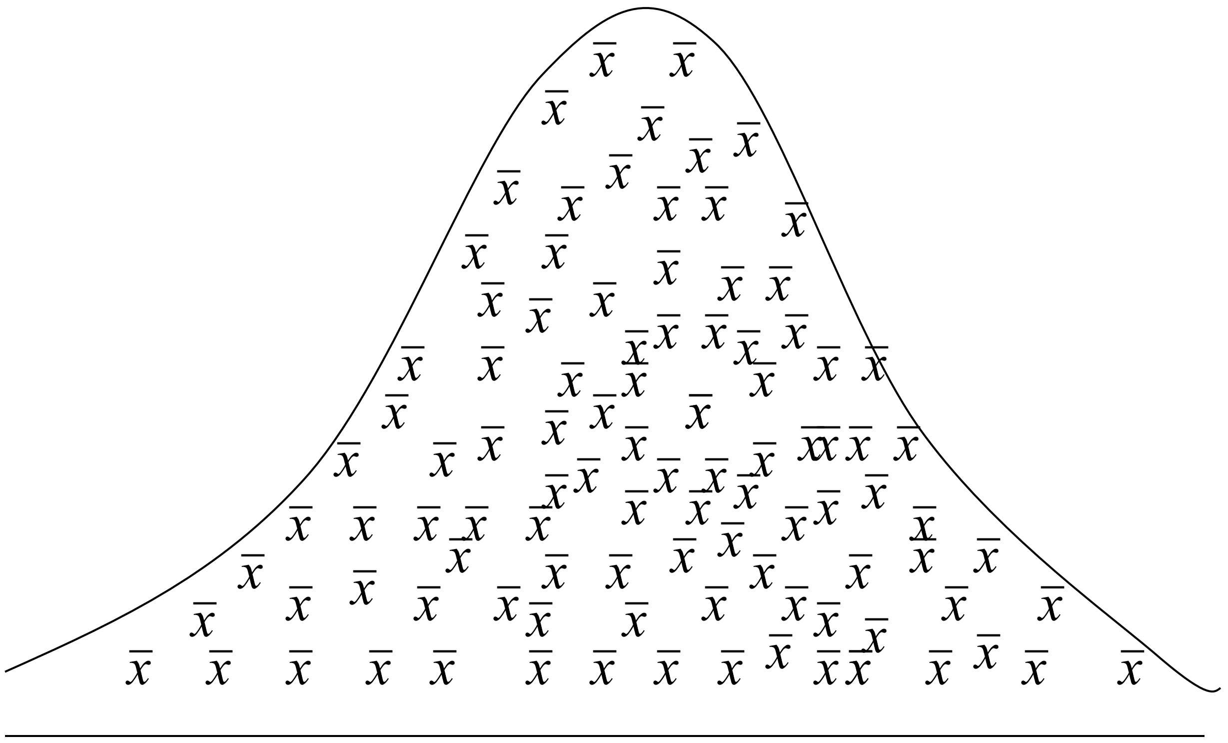
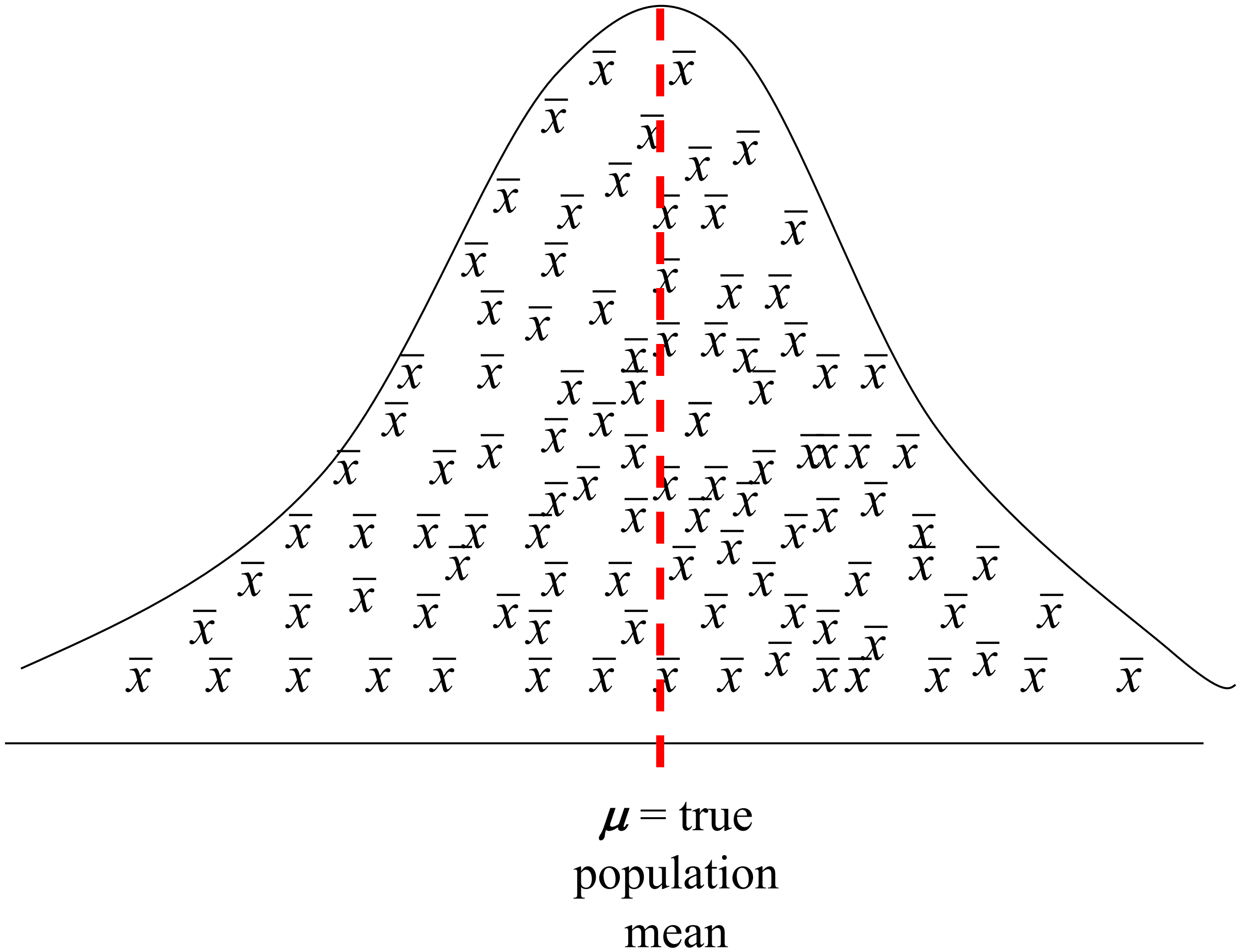


Illustration of CLT and Sampling Distribution



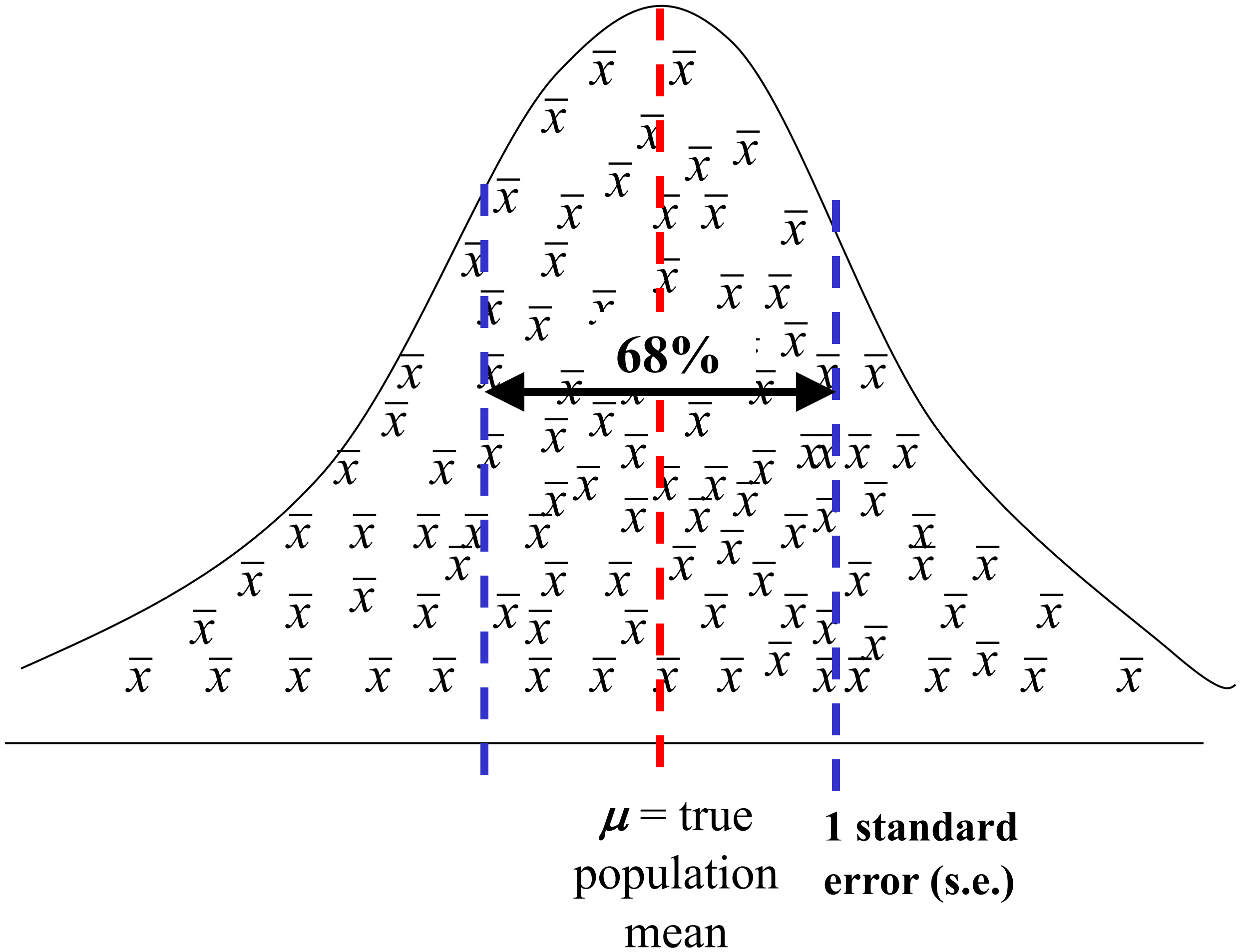
Central Limit Theorem: The distribution of sample statistics (here, sample means) from an infinite number of samples is a ***normal distribution***.

Illustration of CLT and Sampling Distribution



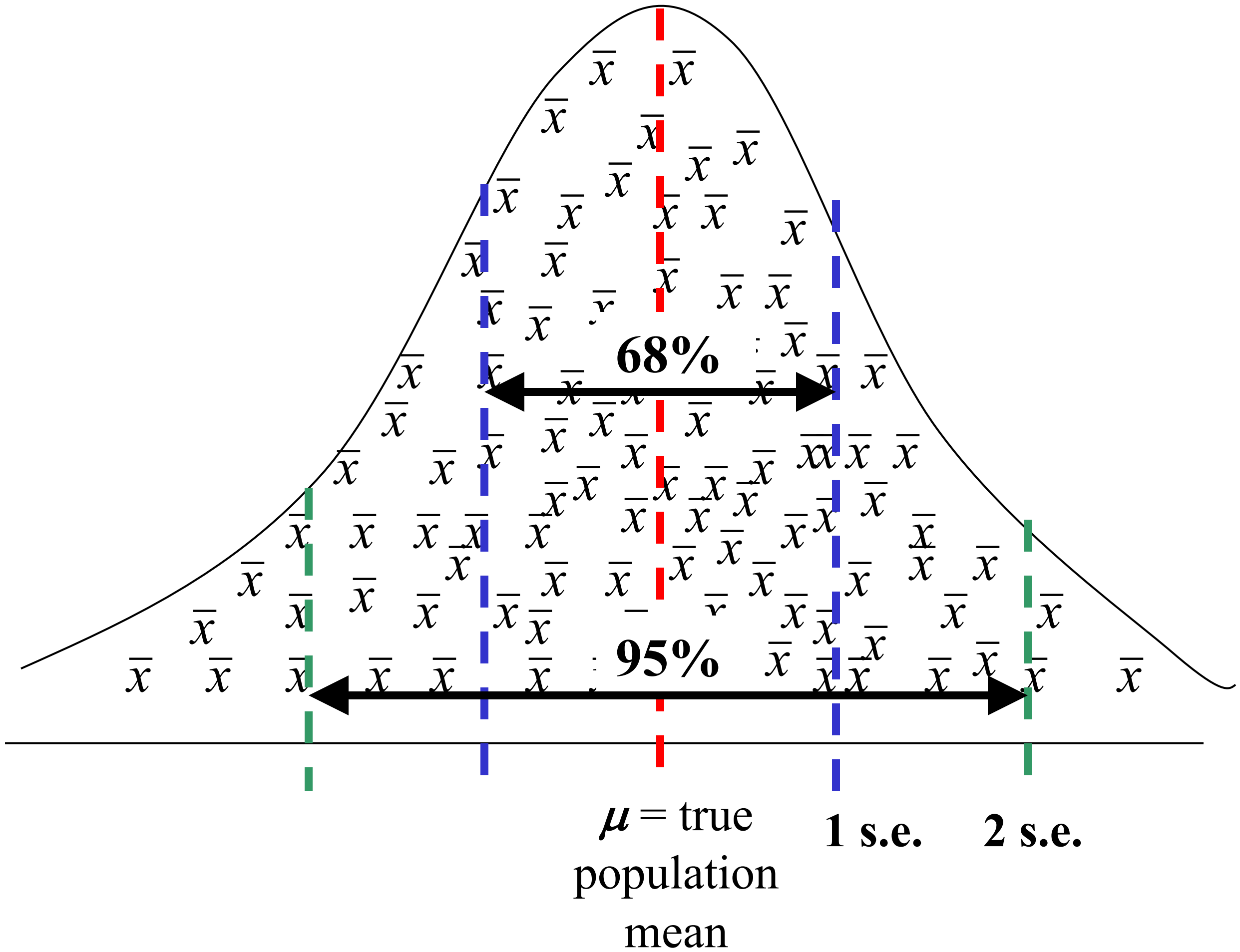
What does dispersion mean in a sampling distribution?

Illustration of CLT and Sampling Distribution



What does dispersion mean in a sampling distribution?

Illustration of CLT and Sampling Distribution



What does dispersion mean in a sampling distribution?

Central Limit Theorem

- Now, we can use our sample statistic in conjunction with our standard error to make certain inferential claims.

- For instance, if we had a sample of 100 students, the mean=2.8 and the standard error is .04, we can say that 95% of all random samples of 100 will produce means of 2.8 $\pm$ .08 (between 2.72 and 2.88).

Probability and Sampling Distributions

- Use CLT and the normal distribution → *degree of uncertainty* associated with our parameter estimates.
  - We make probability statements about how close our estimate reflects the true parameter.
  - Related to "margin of error."

- Our degree of uncertainty about a population parameter is *random sampling error*.

Population parameter = sample statistic + random sampling error

Random Sampling Error

- *Random sampling error* is the degree to which a sample statistic differs from a population parameter due to random sampling error (or "by chance").

- What affects the magnitude of random sampling error?
  - Sample size
  - Standard deviation

- *Random sampling error =*

  (variation component) / (sample size component)

- Random sampling error is synonymous with what we call the *standard error*.

Applied Inferential Statistics

BOTTOM LINE:

For ALL inferential statistics problems, we need two pieces of information:

1. ***Sample statistic*** (e.g., sample mean, sample percentage, estimated causal effect)

2. ***Standard error***

Standard Error

- A ***standard error*** is a standard deviation of a sampling distribution.

- A ***standard error*** communicates the amount of random sampling error associated with a sample statistic (e.g., mean, estimated causal effect).
  - It has the same properties as a standard deviation.

Calculating a Standard Error for a Sample Mean or Proportion

- A standard error estimate for a *sample mean* is calculated as:

$$s.e. = \frac{\sigma}{\sqrt{N}}$$

- σ = *standard deviation* from the sample

- How about for a proportion?

95% Confidence Interval

- A common form of communicating uncertainty in our sample statistic is by using the **95% confidence interval**.

- Recall: 95% of the cases in a normal distribution are within about 2 standard deviations (or standard errors).

- What's the exact number? Go to R.

- 1.96. So 95% of the time, the population mean falls within 1.96 standard errors of the sample mean.

  95% CI = sample mean $\pm$ (1.96 x s.e.)

- *Margin of error* = 1.96 x s.e.

- Example 5.3, p. 173 (OpenIntro).

Interpretation of 95% Confidence Intervals

- Three related ways to interpret the 95% CI (application to mean, but can apply to any sample statistic).
  - If we were to infinitely draw samples (of size N), the population mean (true mean) would fall within our calculated interval 95% of the time.
  - There's a 95% chance that the true mean would fall within our 95% confidence interval, and only a 5% chance that it would fall outside of our 95% confidence interval.
  - We're 95% confident that the population mean falls between [lower bound of the 95% CI] and [upper bound of the 95% CI].

Statistical Significance and Hypothesis Testing

- "Statistical significance" is directly related to our construction and interpretation of a 95% confidence interval.

- Now, switch gears a little bit and construct explicit hypothesis tests.

- Overview – We go through **5 steps** conduct a hypothesis test:
  1. A null and alternative hypothesis
  2. An alpha level – i.e., a confidence level.
  3. Test statistic
  4. Decision rule for rejecting or failing to reject the null hypothesis.
  5. Assess the evidence – make decision, render verdict on null hypothesis.

1. Null and Alternative Hypotheses

- Consider this hypothesis:
  - *In comparing individuals, those with higher levels of education are more likely to participate in politics than those with lower levels*.

- We act as **skeptics**. We need evidence "beyond a reasonable doubt" to conclude there's a relationship between two variables.

- The *null hypothesis* acts the role of skeptic: Presumption of **no relationship.**
  - It states the skeptical counterargument to our hypothesis.

- What is the null hypothesis for the example above?

- Our *alternative hypothesis* is the one we formulated above; suggests that there *is* a relationship.

1. Null and Alternative Hypotheses

- Null hypothesis: H_0.
- Alternative hypothesis: H_A.
- *Important*: In hypothesis testing, we always make decisions with respect to the null hypothesis. We either *reject* it or *fail to reject* it.
- Relation to murder trials

2. Alpha Level

- Back to confidence: Typically 95%. This is our standard by which we will reject or fail to reject the null hypothesis.

- Our confidence level is determined by our *alpha level*. For 95% confidence, alpha = 1 - .95 = **0.05**).

- If we can reject the null hypothesis, we say that the relationship is *statistically significant* at the alpha=.05 level (i.e., at the 95% confidence level).

2. Alpha Level

- ***Type 1 error***:
  - Concluding there is a relationship in the population when in fact there isn't one; i.e., we rejected the null hypothesis, but we shouldn't have.

- ***Type 2 error***:
  - Concluding there is no relationship in the population when in fact there really is one; i.e., we failed to reject the null hypothesis, but we should have rejected it.

- Back to murder trials and "beyond a reasonable doubt."

- With an alpha level of .05, this means that if we reject the null, we'll commit a type 1 error less than 5% of the time if we took an infinite number of samples.

2. Alpha Level

- Just as we want to guard against letting an innocent person go to prison, we want to set a fairly high threshold for rejecting the null hypothesis.

- The *alpha* level a researcher decides on is the threshold for rejecting the null hypothesis. Recall: alpha=1 – confidence level. We typically use the .05 alpha level.

- Using confidence intervals to test the null hypothesis.

3. *Test Statistics*

- Like confidence intervals, a test statistic can give us a formal test for rejecting or failing to reject a null hypothesis.

- In general a test statistic is calculated as:

  Test statistic = (H_A value – H_0 value) / (standard error)

- This is converted to a z-score. We can make inferences based on the normal distribution (via the central limit theorem).

- Proportions example

- Then, we can do this two ways (which will give us the same answer):
  - 95% confidence interval.
  - Find the *p-value* associated with our z-score. Use R and pnorm.

4. Decision Rule

- Decision rule:
  - If the null hypothesis value (0 in a diff. of means test) is not inside our 95% confidence interval, ***reject the null***. If it is inside, ***fail to reject the null***.
  - Recall we want to minimize type-1 error. So we want *p* to be small. ***If p < alpha, reject the null***. ***If p>alpha, fail to reject the null.***
- If we reject the null hypothesis, we conclude that *the value or the effect is significantly different from zero*.
  - The effect is "statistically significant" at the alpha level we've chosen.

5. Assess the Evidence - Making a Decision

- Communicating **statistical significance** or **statistical insignificance**.
- Statistical significance: Do we have enough statistical evidence to reject the null hypothesis value?
- If we reject the null, we conclude that the difference between means is statistically significant; i.e., ***the difference is significantly different from zero***.
- If we fail to reject the null, we conclude that the difference between the means is not statistically significant.

Example

- Let's go through Example 5.32, p. 197.

Estimating Causal Effects: Difference of Means

- Example: Voter mobilization experiment

- Do voter mobilization efforts by parties cause an increase in turnout intentions?

- Individuals are randomly assigned to receive either:
  - **Treatment**: Receive a phone call from a party worker strongly encouraging them to vote.
  - **Control**: Does not receive such a phone call

- **Dependent variable:** intention to vote; self-reported probability that individual will vote.

Estimating Causal Effects: Difference of Means

- Hypothetical results I:
  - **Mean probability of vote intention for Treatment Group:** 0.66
  - **Mean for Control Group:** 0.52
- ***Estimated causal effect:***
  - 0.66 – 0.52 = **0.14**
- Standard error = 0.02
- 95% CI = 0.14 $\pm$ (1.96 x 0.02) = **[0.10, 0.18]**

Estimating Causal Effects: Difference of Means

- Hypothetical results II:
  - **Mean for Treatment Group:** 0.58
  - **Mean for Control Group:** 0.55
- ***Estimated causal effect:*** **0.03**
- Standard error = 0.02
- 95% CI = 0.03 + (1.96 x 0.02) = **[-0.01, 0.07]**