

Test 1 Review

Data

Sampling Techniques

1. **Simple Random** – Picked randomly from a population. Everyone has the same chance.
2. **Convenience Sampling** – Easy to contact. Send them an email.
3. **Stratified Sampling** – Divide them into ‘strata’ or group.
4. **Cluster Sampling** – Randomly divide students into groups.

Data types

1. **Qualitative** - Non-numeric, categorical data.
2. **Quantitative** - Numeric data.
 - a. **Discrete** - A specific count.
 - b. **Continuous** - Measurement on some scale, such as length, time duration, distance, weight)

Data Description

1. Histograms
2. Boxplot
3. Shapes:
 - a. Symmetric
 - b. Left-skewed
 - c. Unimodal
 - d. Bimodal

4. **Mean** - a data set is the measure of center found by adding the sample values and dividing the total by the number of values.

$$y = \sum(y_i)/n$$

5. **Median** - A measure of center that overcomes the influence of extreme values is the median.
6. **Range** - Difference between the maximum value and minimum value.

7. **Variance** -

$$s^2 = \sum (y_i - \text{mean})^2 / n - 1$$

$$\text{Variance Coefficient: } CV = \text{Standard Deviation} / |\text{mean}|$$

8. **Standard Deviation** - A measure of variation of all values from the mean. Can vary greatly when including outliers.
9. **IQR**
10. **Five Number Summary**
11. **Empirical Rule** - 68% of the data is in the middle. 95% is one standard deviation away. 99.7% are two standard deviations away.
- Formula: $(y \pm s), (y \pm 2s), (y \pm 3s)$
12. **Outliers** - Unusual or extreme values or possible errors
- Mild
 - Extremes

Basic Probability

- Addition Rule
- Complement Rule
- Conditional Probability
- Multiplication Rule
 - Probability Tree
- Independent Events

Discrete Random Variables

1. General probability distribution:
 - a. probabilities
 - b. mean
 - c. standard deviation
2. Binomial distribution:
 - a. calculate probabilities
 - b. mean
 - c. standard deviation