$\mu = PopulationMean$
$\bar{x} = SampleMean$
$\mu - \bar{x} = SampleError$

1 Lesson 1 Notes

independent variable is the predictor dependent variable is the outcome

Correlation does not prove Causation

Show Relationships then use Observational Studies / Surveys Show Causation requires a Controlled Experiment

Benefits of Surveys Easy way to get information on a population Relatively inexpensive Can be conducted remotely Anyone can access and analyze results

Downsides of Surveys Untruthful responses Biased responses Respondents not understanding the question (response bias) Respondents refusing to answer (non-response bias)

Controlled Experiment Blinding - participant not told which treatment (actual or palcebo) they are receiving control group receives placebo Double Blind - researchers do not know which treatment participants received

2 Lesson 3 Notes

population mean = $\sum_{i=1}^{N} \frac{x_i}{N}$ sample mean = $\sum_{i=1}^{n} \frac{x_i}{n}$ where N = population size and n = sample size

3 Lesson 4 Notes

a data point is an outlier if it is less than Q1 - 1.5(IQR) or greater than Q3 + 1.5(IQR) where IQR is the inter-quartile range

variance $\Sigma \frac{(x-\bar{x})^2}{n}$

4 Lesson 5 and 6 Notes

z score for a population $z = \frac{x-\mu}{\sigma}$

5 Lesson 7 Notes

Standard Error $\frac{\sigma}{\sqrt{n}}$ where n = sample size

6 Lesson 8 Notes

z score for a sample mean
$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

95% of sample means will fall within $\frac{1.96\sigma}{\sqrt{n}}$
$\bar{x} \pm \frac{1.96\sigma}{\sqrt{n}}$ is a 95% confidence interval around the sample mean

7 Lesson 9 Notes

alpha levels $\alpha = .05$
$\alpha = .01$
$\alpha = .001$

$H_0 = NullHypothesis : \mu = \mu_I$
$H_a = AlternativeHypothesis : \mu \neq \mu_I or \mu < \mu_I or \mu > \mu_I$
The Null Hypothesis that the new population mean will not be statistically different than the original population mean after a treatment (i.e. there was not a statistically significant impact from the treatment). The Alternative Hypothesis assumes there is an effect.

Statistical Desicion Error
Type I Error - Reject the Null when it's TRUE
Type II Error - Retain the Null when it's FALSE

8 Lesson 10 Notes

Standard Error of the Mean $\frac{S}{\sqrt{n}}$ Degrees of Freedom is the number of choices we can make before the remaining requirements are forced. For example, in an n by n table where the rows and columns must sum to a value we have $(n-1)^2$ degrees of freedom.

t score for a sample mean
$t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$

Dependent t tests for paired samples within-subject designs two conditions pre-test, post-test growth over time (longitudinal study)

Cohen's D is the Mean of the differences divided by the standard deviation of the differences

9 Lesson 10b Notes

Difference Measures Mean Difference Standardized Difference Cohen's D

Correlation Measures r2 - Proportion of variation in one variable that is related to ("explained by") another variable

Meaningfulness of Results What was measured? do the results have practical, social, or theoretical importance Effect Size Can we rule out random chance Can we rule out alternative explanations (i.e. lurking variables)

r2 - coefficient of determination ranges from 0.00 to 1.00

$r^2 = \frac{t^2}{t^2+df}$ where df is the degrees of freedom and t2 is our computed t value not the critical t value

Results Section 1. Descriptive Statistics (M, SD) a. in text, in graphs, in tables 2. Inferential Statistics Hypothesis test - kind of test (e.g. one sample t test) - test statistic - degress of freedom - p-value - direction of the text (e.g. one-tailed)

10 Lesson 11 Notes

Independent Samples

Standard Deviation for Independent Samples
$N(\mu_1, \sigma_1) - N(\mu_2, \sigma_2) = N(\mu_1 - \mu_2, \sqrt{\sigma_1 + \sigma_2})$

Standard Error for Independent Samples
$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

Degrees of Freedom for Independent Samples
$df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$

Pooled Variance
$\frac{\sum_{i=1}^{n_x}(x_i - \bar{x})^2 + \sum_{i=1}^{n_y}(y_i - \bar{y})^2}{df_x + df_y}$