When to Use a Particular Statistical Test

Univariate Descriptive

Central Tendency

Mode

- the most commonly occurring value

ex: 6 people with ages 21, 22, 21, 23, 19, 21 - mode = 21

Median

- the center value
- the formula is $\frac{N+1}{2}$

ex: 6 people with ages 21, 22, 24, 23, 19, 21
line them up in order form lowest to highest
19, 21, 21, 22, 23, 24
and take the center value - mode =21.5

Mean

- the mathematical average
- the formula is $\sum X/N$

ex: mean age = age of person one + age of person two + age of person three, etc./number of people

Variance

- a measure of how spread out a distribution is
- it is computed as the average squared deviation of each number from its mean

Standard Deviation

- how much scores deviate from the mean
- it is the square root of the variance
- it is the most commonly used measure of spread

Bi- and Multivariate Inferential Statistical Tests

Differences of Groups

Chi Square

- compares observed frequencies to expected frequencies

ex: Is the distribution of sex and voting behavior due to chance or is there a difference between the sexes on voting behavior?

t-Test

- looks at differences between two groups on some variable of interest
- the IV must have only two groups (male/female, undergrad/grad)

ex: Do males and females differ in the amount of hours they spend shopping in a given month?

ANOVA

- tests the significance of group differences between two or more groups
- the IV has two or more categories
- only determines that there is a difference between groups, but doesn't tell which is different

ex: Do SAT scores differ for low-, middle-, and high-income students?

ANCOVA

- same as ANOVA, but adds control of one or more covariates that may influence the DV

ex: Do SAT scores differ for low-, middle-, and high-income students after controlling for single/dual parenting?

MANOVA

- same as ANOVA, but you can study two or more related DVs while controlling for the correlation between the DV
- if the DVs are not correlated, then separate ANOVAs are appropriate

ex: Does ethnicity affect reading achievement, math achievement, and overall scholastic achievement among 6^{th} graders?

MANCOVA

- same as MANOVA, but adds control of one or more covariates that may influence the DV

ex: Does ethnicity affect reading achievement, math achievement, and overall scholastic achievement among 6th graders after controlling for social class?

Relationships

Correlation

- used with two variables to determine a relationship/association
- do two variables covary?
- does not distinguish between independent and dependent variables

ex: Amount of damage to a house on fire and number of firefighters at the fire

Multiple Regression

- used with several independent variables and one dependent variable
- used for prediction
- it identifies the best set of predictor variables
- you can enter many IVs and it tells you which are best predictors by looking at all of them at the same time
- in sequential regression the computer adds the variables one at a time based on the amount of variance they account for

ex: IVs drug use, alcohol use, child abuse
DV. suicidal tendencies

Path Analysis

- looks at direct and indirect effects of predictor variables
- used for relationships/causality

ex: Child abuse causes drug use which leads to suicidal tendencies.

Group Membership

Logistic Regression

- like multiple regression, but the DV is a dichotomous variable
- logistic regression estimates the odds probability of the DV occurring as the values of the IVs change

ex: What are the odds of a suicide occurring at various levels of alcohol use?