Basic Statistics Guide

Prepared by: Dr. Wendy Craig, Queen's University
and Michaela Young, Queen's University
March 2022

Table of Contents

Definitions

Means, Standard Deviation

Mean: The average of a data set.

- The mean is a measure of "central tendency", i.e. where the center of a data set lies.
- Compute by adding all the numbers in a data set together and dividing by the number of items in the set.
    - E.g. 10 + 10 + 50 + 30 = 100 / 4 = 25.

Standard Deviation: Shows the degree to which a dataset is spread out around the mean/average.

- Describes variability in a single sample.
- A low standard deviation means data are more clustered around the mean.
- A high standard deviation means data are more spread out from the mean.

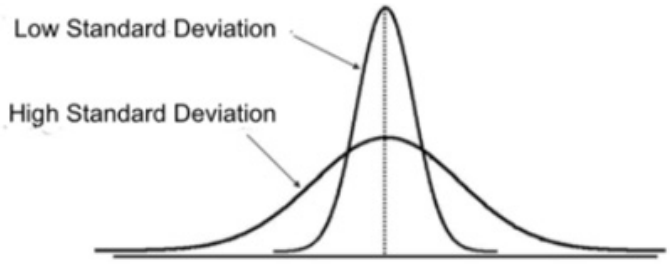

Image from https://careerfoundry.com/en/blog/data-analytics/standard-error-vs-standard-deviation

Confidence Intervals

Confidence Interval: A range of values that a true value of a parameter likely lies in.

- Measures the degree of certainty or uncertainty in a sampling method.

P-Values and Statistical Significance vs. Clinical Significance

p-value (or probability value): A number describing how likely it is that your data would have occurred by random chance.

- The smaller the p-value, the stronger the evidence that your data did **not** occur by chance.
- Most often, researchers refer to a p-value less than 0.05 ($p < .05$) as **statistically significant**, and a p-value less than 0.001 ($p < .001$) as **highly statistically significant**.

So, what is the difference between **statistical significance** and **clinical significance**?

Statistical Significance: When a result is found to be unlikely to have occurred by chance.

Clinical Significance: When a result suggests that an intervention (e.g. a treatment or youth program) has had genuine and quantifiable effects.

- The extent to which an intervention has had a meaningful impact on participants' lives (Durlak, 2009).

Effect Size

Effect Size: A statistical concept that measures the strength of the relationship between two variables on a numeric scale.

- Compares two groups, e.g. depression scores of an experimental group (who received a treatment) vs. depression scores of a control group (who did not receive a treatment)

The most commonly used value used to represent effect size is called Cohen's d. The table below shows Cohen's suggestions that 0.2 be considered a 'small' effect size, 0.5 represents a 'medium' effect size, and 0.8 a 'large' effect size.

Relative Size	Effect Size	% of control group below the mean of experimental group
small	0.2	58%
medium	0.5	69%
large	0.8	79%

Types of Variables

There are two main variable types in statistics:

1. **Numerical/Quantitative Variables:** counts, percents, or numbers, for example:
    a. Number of children in a family
    b. Percentage grade on a test
    c. Score on a questionnaire
2. **Non-numerical/Categorical Variables:** written descriptions of groups or things, for example:
    a. Gender
    b. Race/ethnicity
    c. Favourite ice cream flavour

Numerical/quantitative variables and non-numerical/categorical variables can be assigned the roles of dependent variable or independent variable, depending on the design of a research study.

1. **Independent variable (also called the predictor variable):** A variable thought to be the cause of some effect
    - What the researcher *manipulates*
2. **Dependent variable (also called the outcome variable):** A variable thought to be affected by changes in an independent variable
    - What the researcher *measures*

Confounds and Control Variables

Confounding Variable: A variable that is related to the independent variable, and therefore could have its own effect on the dependent variable – which we don't want!

- This could obscure the true relationship we are looking for between the independent and dependent variable.

So, if we know of a potential confounding variable, we should control for it.

Control Variable: A potential confound that you were able to spot.
Example: [1]

- In one study, researchers are comparing two study methods: regular, spaced study and cramming before an exam.
  - Independent variable: Study method (A. regular studying, or B. cramming)
  - Dependent variable: Scores on exam
- A potential confound could be **sleep deprivation**. Maybe people who cram the night before get less sleep, and that is why they scored worse.
- Solution: Control for sleep, by excluding people who got less than six hours of sleep.
- **Control variable = sleep**

Types of Statistical Tests

T-test: Used to determine if the mean of one group is significantly different from either a known score, or from the mean of another group/condition.

There are three different kinds of t-test:

1. **One-sample t-test:** When the mean of one sample is compared to a known population mean
   a. Example: If we know the average age of youth who experience TDV from a previous study, we can compare that mean age to the mean age of youth who experience TDV from a more recent sample.
2. **Paired-sample t-test:** When the means of two *related* samples are compared on a continuous (numerical) variable

a. Example question: *Do teens experience less dating violence after participating in a PREVNet teen dating violence prevention workshop?*
   i. Here, we would examine the average number of TDV occurrences experienced by a group of youth prior to an intervention and compare it with the average number of TDV occurrences experienced by ***the same group of youth*** after the intervention.

3. **Independent sample t-test:** When the means of two unrelated samples are compared on a continuous (numerical) variable
   a. Example question: *Do gender minority youth experience teen dating violence more often than non-gender minority youth?*
      i. Here, our two unrelated samples are the sample of gender minority youth and the sample of non-gender minority youth.

[1] *Castelhano, M. (2018). Lecture 2, lecture notes, PSYC221 Cognitive Psychology, Queen's University, delivered 6 February 2018.*

What happens if we have more than 2 groups to compare?

ANOVA (Analysis of Variance): Used to compare the means of 3 or more groups with one dependent variable

- One-way ANOVA: One independent variable with 3 or more groups, and one dependent variable
  - Example: *You want to test how three different TDV interventions affect the number of TDV occurrences experienced by teens after each intervention.*
    - You can use a one-way ANOVA to find out if there is a difference in occurrence of TDV between the three interventions.
    - Here, your independent variable is TDV intervention (where there are three different groups, or types of intervention). Your dependent variable is the # of TDV occurrences experienced after the intervention.
- Two-way ANOVA: Two independent variables each with multiple groups, and one dependent variable
  - Use a two-way ANOVA when you want to know how two independent variables affect a variable together.
  - Example: *You want to test how three different TDV interventions impact the number of TDV occurrences experienced by teens in three different age groups.*
    - Here, your independent variables are TDV intervention (where there are three different groups, or types of intervention) and age (with three age groups). Your dependent variable is the # of TDV occurrences experienced after the intervention for each age group.

MANOVA (Multivariate Analysis of Variance): Used to compare the means of 3 or more groups with more than one dependent variable

- Example: *You want to test how three different TDV interventions impact the number of TDV occurrences experienced by teens along with their depression scores.*
  - Here, your independent variable is TDV intervention (where there are three different groups, or types of intervention). Your dependent variables are the # of TDV occurrences experienced after the intervention and their depression scores after the intervention.

Correlation

Correlation: A measure of the degree to which two variables are related

- **Positive correlation:** Both variables move in the same direction. As one increases, the other also increases, or as one decreases the other also decreases.
- **Negative correlation:** When one variable increases, the other decreases.
- **Zero correlation:** When no relationship exists between two variables.

Example question: *How is time of day related to the number of occurrences of teen dating violence?*

- Potential answer: As time of day increases (i.e., as night approaches), the number of reported instances of teen dating violence increases (positive correlation).

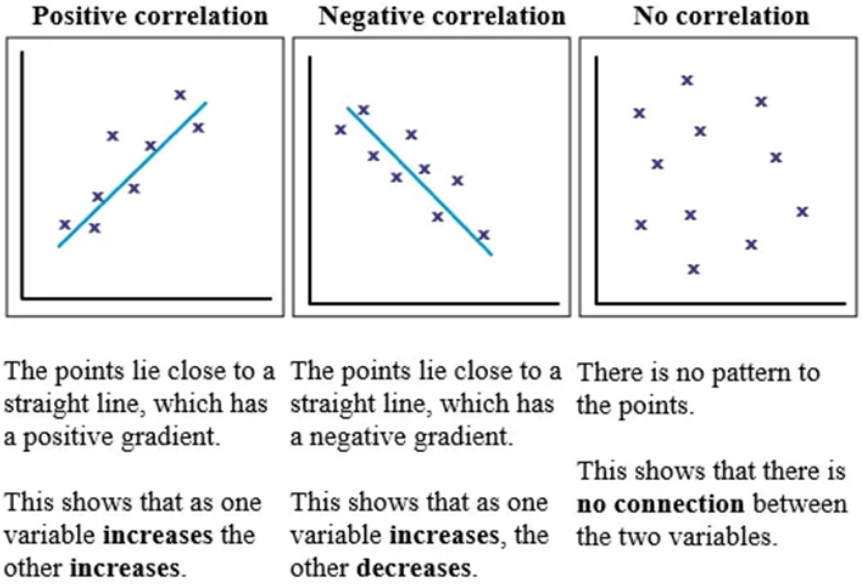

Image from https://careerfoundry.com/en/blog/data-analytics/standard-error-vs-standard-deviation

Chi-square Test

Chi-square Test of Independence: Tests whether or not distributions of categorical variables differ from each other.

- A chi-square statistic (x2) is a single value that shows you how much difference exists between your observed data counts and the counts that you would expect if there were no relationship at all in the population.
- A low chi-square value means there is a high correlation between your two variables.
- Example: Gender and type of TDV (both categorical variables)
  - *Does gender affect the types of TDV experienced by youth?*

Regression

Multiple Regression Analysis: Used to describe relationships between a set of independent variables and the (continuous) dependent variable.

- Produces a regression equation where the coefficients represent the relationship between each independent variable and the dependent variable
- Regression can also be used to make predictions!

Interpreting Regression Output - looking at regression coefficients and p-values

1. Low p-values (e.g. < .05) mean that the independent variable is statistically significant.
2. The coefficients represent the average change in the dependent variable given a one-unit change in the independent variable, while controlling for the other independent variables.

Mediators and Moderators

Mediator (or mediating variable): Explains how an independent variable and a dependent variable are related.

- E.g. Study method (an independent variable) can affect exam scores (a dependent variable) through the mediator of sleep quality
    - In this example, cramming might result in poor sleep quality, which could decrease exam scores. On the other hand, regular studying might result in better sleep quality, which could increase exam scores.

Moderator (or moderating variable): Affects the strength and direction of the relationship between an independent and dependent variable.

- E.g. Stress (independent variable) has a more noticeable impact on men's health than it does on women's health[2] (health = dependent variable).
    - In this example, gender moderates the strength of an effect between stress and health.

Repeated Measures and Between Subjects Designs

Between subjects design: When separate groups are created for each treatment in a study.

- Each participant is assigned to only one treatment group.
- E.g. Testing a new medication: One group receives the actual medication, and the other receives a placebo.

Within subjects / Repeated Measures design: When multiple measures of a dependent variable are taken on the same subjects OR matched subjects OR subjects under different conditions over two or more time periods.

- Examples:
    - Frequently used in pre-test/post-test research designs e.g. testing youth depression and anxiety scores before a school program, and again after the program is delivered.
    - When a new youth program is delivered to students, and the results are measured over time (e.g. 4 weeks after, 8 weeks after, 12 weeks after)

[2] *American Psychological Association. https://www.apa.org/news/press/releases/stress/2011/gender.pdf*

- Two main types of repeated measures designs:
  - **Cross-sectional designs:** Interview/study a new sample of people each time they are carried out
  - **Longitudinal designs:** Interview/study the same sample of people over time

Cross-sectional	Longitudinal
One point in time	Several points in time
Different samples	Same sample
Snapshot of a given point in time, change at a societal level	Change at the individual level
Eg, British Social Attitudes Survey, Labour Force Survey	Eg, British Birth Cohort Studies, Understanding Society

Image from https://learning.closer.ac.uk/wp-content/uploads/2017/07/Longitudinal-vs-cross-sectional-studies-table.pdf

Other Terms

Odds ratio: A measure of association between a certain property A and a second property B in a population.
- Tells you how the presence or absence of property A has an effect on the presence or absence of property B.
- E.g., You could use an odds ratio to figure out if a particular exposure (like vaping) is a risk factor for a particular outcome (such as lung cancer).

Logistic regression: Describes the relationship between a categorical dependent variable and a given set of independent variables.

Meta-analysis: An examination of data from a number of independent studies of the same subject, in order to determine overall trends.[3]

Post hoc Analyses: Analyses that are conducted on a data set after data collection is complete.

[3]*Oxford Languages*

Where should I look in an article for an understandable summary of the research findings?

- For a quick description of an entire research study, the **Abstract** is a great place to look. An abstract is a short 1-2 paragraph section at the beginning of a research article that succinctly describes the background, methodology, findings and implications of a research study.
- If you are looking for a more descriptive explanation of a study's results, try reading the **Discussion** section, where authors will delve into the implications of their findings.

References

https://www.statisticshowto.com/statistics-basics/

https://www.mhaonline.com/faq/clinical-vs-statistical-significance

https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/effect-size/

https://www.simplypsychology.org/effect-size.html

https://statisticsbyjim.com/regression/when-use-regression-analysis/

https://www.scribbr.com/methodology/mediator-vs-moderator/

https://www.scribbr.com/statistics/one-way-anova/

Durlak, J. A. (2009). *How to Select, Calculate, and Interpret Effect Sizes. Journal of Pediatric Psychology, 34*(9), 917–928. doi:10.1093/jpepsy/jsp004

Salkind, N. J. (2010). Encyclopedia of research design (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412961288

Public Health Agency of Canada | Agence de la santé publique du Canada

The views expressed herein do not necessarily reflect the views of the Public Health Agency of Canada