

Definitions

Population

The entire collection of subjects we wish to target

Sample

A subset of the population

Univariate

1 variable

Bivariate

2 variables from the same subject

Multivariate

2 or more variables from the same subject

Qualitative Data

Categorical, descriptive (yes/no, blue, etc)

Quantitative Data

Numeric (include discrete and continuous)

Discrete Data

Primarily count data (the number of...)

Continuous Data

Data from measurements (can take on any value w/in some interval)

Ways to Obtain a Sample

Stratified Sample helps avoid biased data (If there are 2X white people than Hispanics, then the sample should have 2X white people than Hispanics)

Convenience Sampling Stay away from this, (Choosing to only sample from one assembly line on the shop floor)

Simple Random Sampling (A name is drawn out of a hat)

Ways to Obtain Data

Experiment allows us to draw cause and effect b/c of the ways its designed (the best)

Survey A questionnaire or observation

The 2 Branches of Statistics

Descriptive Statistics

use of graphs, numeric computations to summarize the data

Inferential Statistics

Make and inference using sample statistics back to the population

Predicts

Sample Mean (\bar{x}) Population mean (μ)

Sample Median (\bar{x}) Population Median (μ)

Sample Relative Frequency (p) Population Proportion (p)

Sample Standard Deviation (s) Population Standard Deviation (σ)

Sample Variance (s^2) Population Variance (σ^2)

Things to know how to calculate:

Trimmed Mean trim a certain percentage of values from the ends of the data set, and then average whats left

Standard deviation (s) The size of a typical deviation (calculator function)

Variance (s^2) How data points vary from the mean

Symbols and Their Meanings

n sample size

N population size

Characteristics of a Graph

Center

tells us what a typical value in the data set should be (If data is fairly symmetric use mean, otherwise, use median)

Spread

The range of data

Skew

If the bell curve is shifted left (negative skew) or right (positive skew)

More on Box Plots

Box Plots They show us outliers visually, and are great for comparing multiple data sets

Quartiles Values that divide the sorted data set into 4 equal parts.

Q1 The smallest 25% of data

Q2 The median

Q3 The 75% mark

Q4 The max value

percentiles If a value is in the first quartile, then 75% of the values are greater than that, so your in the 75th percentile

Graph Types

Box Plots	Show us outliers visually and great for comparing multiple data sets
Dot Plot	Dots located above their value on the X-axis
Stem and Leaf	The stem of the number includes all but the last digit (so 38 3 would be 38.3)
Histogram	Like box graphs but there's no spaces between columns, can be used with discrete and continuous data
Histogram Shapes:	Symmetric, Right(positive) skew, Left (negative) skew, Bi-modal (2 peaks), and Multi-modal (many peaks)



By **Shelbeans** (shelbeans)
cheatography.com/shelbeans/

Published 6th September, 2024.
Last updated 28th January, 2024.
Page 2 of 2.

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
<http://crosswordcheats.com>