Norms & Basic Statistics

Psych 427
Introduction to Psychological Testing
Gary S. Katz, Ph.D.

Why Statistics?

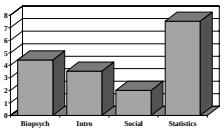
- Organize data.
- Extract ***meaning*** from ***data***.

Two Kinds of Statistics

- Descriptive
  - Describe a set of data.
  - Reporting measures (observations).
- Inferential
  - Make predictions based upon a collection of observations

Descriptive Statistics

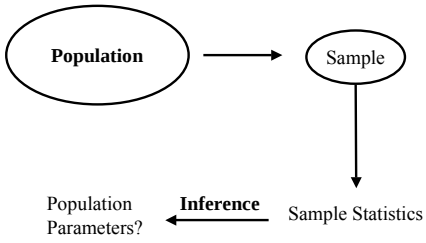
Course	Average Rating
Biopsych	4.42
Intro	3.53
Social	2.00
Statistics	7.50



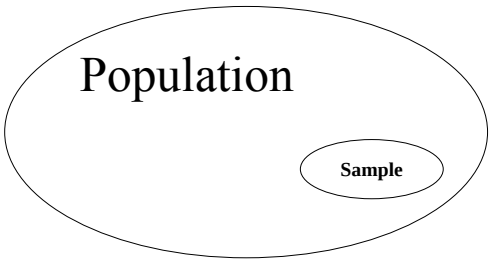
Inferential Statistics

- What is the rating that this class will likely give psychological testing at the end of the semester?
- Do students really like psychological testing better than social psychology?

Inferential Statistics



Samples from Populations



Populations & Samples

Scales of Measurement

- Nominal
- Ordinal
- Interval
- Ratio

Nominal

- Labels
- Think: Nominal, Names
- Identify different categories
- No concept of more, less; no magnitude

- Gender, Ice Cream Flavors, Fruits

Ordinal

- Ordered set of observations.
- Think: Ordinal, ORDER
- Different categories, but the categories have a logical order; magnitude
- Distances between categories varies.

- Class rank, Derby Results, Sports Rankings

Interval

- Different categories, logical order, distance between categories is constant.
- Think: Interval has constant intervals.
- Distances between categories are meaningful.
- No meaningful zero point - no ratio comparisons.

- Temperature in F, Pain sensitivity?

Ratio

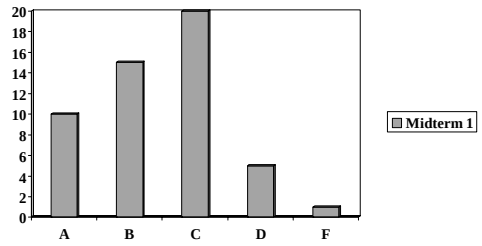
- Different categories, logical order, constant distances, meaningful zero.
- Think: Ratio allows ratio comparisons (or "Ra-she" 0)
- "Twice as much as" works here.
- Zero - absence of quantity being measured.
- Height, temperature in K

Scales of Measurement & Their Properties

Three things we want to know about a set of test data:

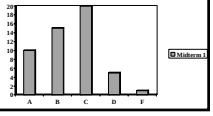
- Shape
- "Typical" value
  - measure of central tendency
- "Spread" of scores
  - measure of variability

Types of Graphs: Bar Graph

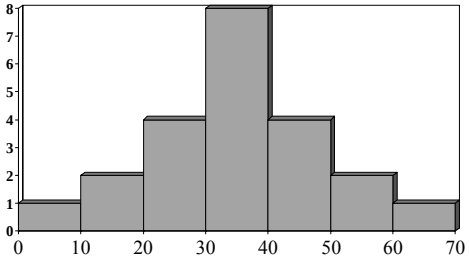


Things to Remember: Bar Graphs

- Height of bar = frequency of that category.
- Width of bar is irrelevant, but should be constant.
- If Width is irrelevant, so is Area.
- ***Bars Don't Touch!***
- Nominal or Ordinal Data



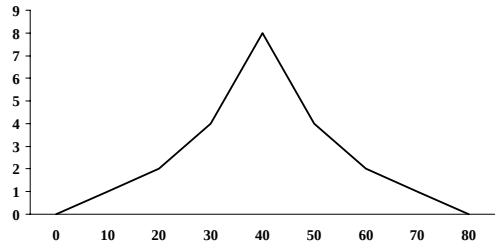
Types of Graphs: Histogram



Things to Remember: Histograms

- Height of bar = number of observations in "that range."
- Width of bar = width of "range."
- Height is meaningful, Width is meaningful, therefore, Area is meaningful.
- Area = total number of observations in a given range.
- ***Bars touch!***
- Interval or Ratio Data

Types of Graphs: Frequency Polygon



Percentile Ranks

- What percent of the scores fall below a particular score?

- Percentile ranks are RANKS, not scores.
- *Height example.*

Height Example - Percentiles

Percentiles

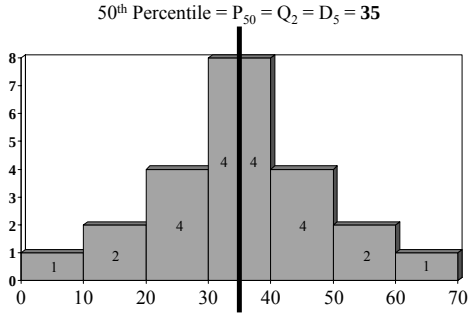
- **Scores** which divide distributions into specific proportions.
  - Percentiles = hundredths
    - P_1, P_2, P_3, … P_{97}, P_{98}, P_{99}
  - Quartiles = quarters
    - Q_1, Q_2, Q_3
  - Deciles = tenths
    - D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9
- **Percentiles are SCORES.**

Things to Remember: Histograms

- Height of bar = number of observations in "that range."
- Width of bar = width of "range."
- Height is meaningful, Width is meaningful, therefore, Area is meaningful.
- Area = total number of observations in a given range.
- ***Bars touch!***
- Interval or Ratio Data

Percentiles

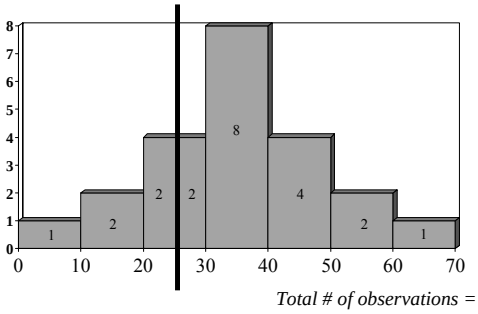
50th Percentile = $P_{50} = Q_2 = D_5 =$ **35**



Total # of observations = 22

Percentiles

What percentile is a raw score of 25?



Total # of observations = 22

Percentiles

Percentiles & Percentile Ranks

- Percentiles and percentile ranks are used to compare scores of individuals with groups of individuals.
- Suppose you earned 40 points on the midterm.
  - Fruit magnet score?
- Suppose you earned a 95th percentile score on the midterm?
  - Fruit magnet score?

Three things we want to know about a set of test data:

- Shape
- "Typical" value
  - measure of central tendency
- "Spread" of scores
  - measure of variability

Mean

- Arithmetic mean:
- Mean of a sample:
- Mean of a population: μ

Height Example - Mean

Three things we want to know about a set of test data:

- Shape
- "Typical" value
  - measure of central tendency
- "Spread" of scores
  - measure of variability

Variance

- The *average squared* distance of each score from the mean.
- Also known as: **mean square**

- Variance of a sample: s^2
- Variance of a population: σ^2

Variance

Height Example - Variance

What's a squared unit?

Standard Deviation

Height Example - Standard Deviation

Standard Deviation, Variance, Samples & Populations

- To estimate the population mean, we use the sample mean.
- To estimate the population variance or standard deviation, we need to "fix" the sample formula.

Estimating Population Variance & Standard Deviations

Z-Scores

Indicate how many standard deviations a score is away from the mean.

Two components:

- **Sign:** positive (above the mean) or negative (below the mean).
- **Magnitude:** how far from the mean the score falls.

Z-Scores: Formulae

- Raw score -> Z-Score:

$$z = \frac{x - \bar{x}}{s} = \frac{\text{score} - \text{mean}}{\text{standard deviation}}$$

- Z-Score -> Raw score

$$x = \bar{x} + z(s) = \text{mean} + z(\text{standard deviation})$$

Properties of Z-Scores

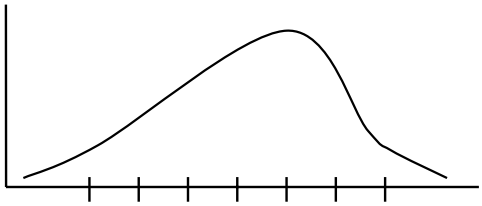
- Z-score indicates how many SD's a score falls above or below the mean.
  - Positive z-scores are above the mean.
  - Negative z-scores are below the mean.
- Mean of distribution of z-scores = 0
- SD of distribution of z-scores = 1

Properties of Z-Scores

- Most z-scores fall between -3 and +3 because scores beyond 3SD from the mean are very rare.
- Z-scores can be used to compare scores within a distribution and across different distributions, provided that the different distributions have the same shape.

Properties of Z-Scores

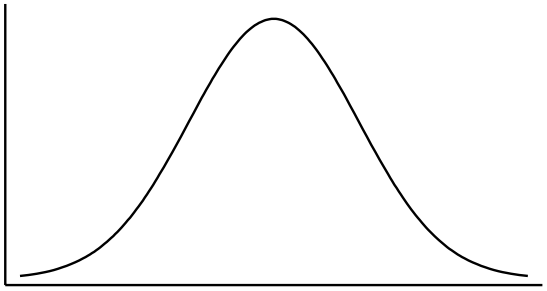
- If you transform a set of raw scores into z-scores, the shape of the distribution does not change!



Raw Scores	20	30	40	50	60	70	80
Z-scores	-3	-2	-1	0	+1	+2	+3

Height Example - Z Scores

The Normal Curve



Normal Curve: Characteristics

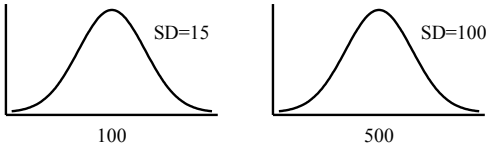
- Theoretical distribution; generated by a mathematical formula:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{(x-\mu)^2/(2\sigma^2)}$$

  - No real distribution is perfectly normal.
  - Many actual distributions are approximately normal, so normal curve statistics apply.
  - Normal curve statistics underlie procedures in inferential statistics.

Normal Curve: Characteristics

- Family of distributions
  - there is a normal curve for every possible combination of mean and standard deviation
    - IQ: mean 100; sd 15
    - SAT: mean 500; sd 100



Normal Curve: Characteristics

- Symmetrical, unimodal
- Mean = median = mode
- Continuous
  - for every x, there is a corresponding y
- Infinite range, asymptotic
  - tails never touch the abscissa

Normal Curve: Characteristics

In any normal curve:

68.26% of the area falls within 1SD of the mean

95.44% of the area falls within 2SD of the mean

99.74% of the area falls within 3SD of the mean

Guidelines:

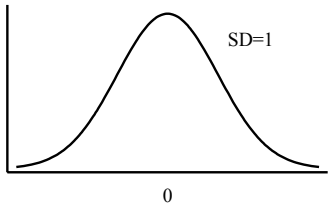
2/3 of the area falls within 1SD of the mean

95% of the area falls within 2SD of the mean

99% of the area falls within 3SD of the mean

Normal Curve: Characteristics

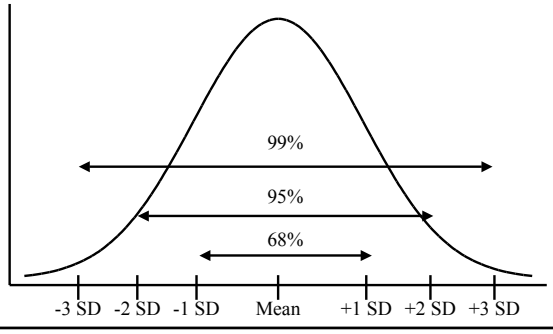
- A normal curve expressed as z-scores is called a **standard normal curve.**



Normal Curve - Area

- In any distribution, the percentage of the area in a given portion is equal to the percent of scores in that portion.
  - Since 68% of the area falls between the mean and ±1 SD of a normal curve,
  - 68% of the scores in a normal curve fall between ±1 SD of the mean.

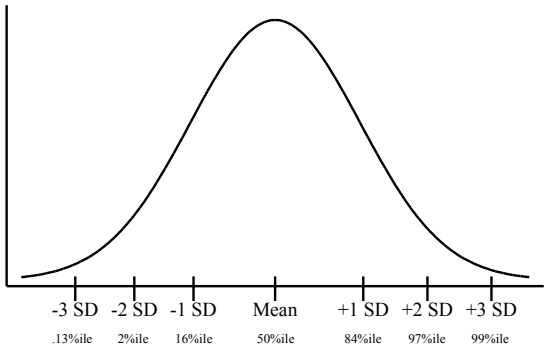
The Normal Curve



Other "Easy" Percentages

- 50% of the area lies above or below the mean
- 34% of the area lies between the mean and ±1SD
- 47.5% of the area lies between the mean and ±2SD
- 16% of the area lies beyond ±1SD

The Normal Curve



Other Percentages?

- For other percentages, we use a tabled chart -- see appendix I and appendix II

Other Standard Scores

- McCall's T scores (or just T Scores)
  - Mean 50, SD 10

$$T = 10Z + 50$$

- Stanines (standard nines)
  - Mean 5, standard deviation ~ 2