

Introduction: Variables, samples, designs

Statistics and experimental design

Textbooks and websites

- Andy Field, Discovering Statistics Using SPSS
- Andrew Mayer, Introduction to Statistics and SPSS in Psychology
- Howell, Fundamental Statistics for the Behavioral Sciences
- Szokolszky Ágnes, Kutatómunka a pszichológiában. (feladatlap is van hozzá)
- <http://www.mathsisfun.com/>
- <http://students.brown.edu/seeing-theory/>
- <http://onlinestatbook.com/2/index.html>
- <http://www.socscistatistics.com/>
- <http://www.graphpad.com/guides/prism/6/statistics/>
- <http://www.statsoft.com/Textbook>
- <http://students.brown.edu/seeing-theory/>
- <http://distributom.org/V3/calc/index.html>

What we'll need

- Laptop
- Excel, Numbers or Google Sheets and SPSS or
- Calculator, (graph) paper, colour pens

Course requirements

- Course materials:
[http://www.cogsci.bme.hu/~kstkuser/
KURZUSOK/BMETE47MC38/2017_2018_1/](http://www.cogsci.bme.hu/~kstkuser/KURZUSOK/BMETE47MC38/2017_2018_1/)
- Homework exercises (send to: Markója Ádám at markoja.adam@cogsci.bme.hu by 10 pm on Wednesday) (34% of final grade)
- Format: SPSS, Google Sheets and Google Docs (or pdf)
- 2 zh (2 x 33% of final grade)

Fontos!

- Aláírás megszerzéséhez be kell adni legalább 10 házi feladatot (12 vagy 13-ból), és meg kell írni mindkét zh-t.
 - A jegybe minden házi feladat beszámít: egy elmaradt házi 0 pontot ér!
- Aki nem kap aláírást, csak egy év múlva tudja újra felvenni a kurzust.
- Vizsgázni csak annak kell, aki egyest kapott valamelyik zh-ra vagy jobb jegyet szeretne.

Schedule, Lab????

Week	Topic
1	Hypothesis testing, population & sample, variables
2	Frequency distribution, central tendency & variability,
3 szept 20	Sport nap
4	Percentiles, outliers, graphs
5	Normal curve, z-scores, Confidence intervals, effect size, power
6	Correlation
7 (okt 18)	1. zh
8	Regression
9	T-tests
10	ANOVA and GLM 1
11	ANOVA and GLM 2
12	ANOVA and GLM 3
13	Non-parametric tests
14	Ami kell, vagy 2. zh
15 (dec 13)	2. zh, ha az előző héten nem volt

Topics for today

- Statistics as an art
- Deductive and inductive reasoning
- Samples and populations
- Descriptive and inferential statistics
- Sampling error
- Dependent and independent variables, and confounds
- Hypothesis testing
- Measurement scales
- Reliability and validity, floor and ceiling effect

The problem with inductive reasoning

- Deductive reasoning:
All men are mortal. Socrates is a man.
Therefore:
- Inductive reasoning:
 1. Some mushrooms are poisonous. 1 in 10,000 people got ill after eating this kind of mushroom.
Is it safe to eat?
 2. Some mushroom are poisonous. 1 in 10 people got ill after eating this kind of mushroom.
Is it safe to eat?
 3. Some mushroom are poisonous. 100 in 1,000 people got ill after eating this kind of mushroom.
Is it safe to eat?

What is statistics?

- Analysis of numerical data:
 - A set of mathematical procedures for organising, summarising and interpreting data.
- The data comes from observations and measurement

A problem

- Does stress have an effect on cognitive performance?
 - Naturalistic observation
 - Quasi-experiment (natural groups tested)
 - Experiment (induce stress in experimental group)
- Whichever method we use, we get data from a **SAMPLE**
 - We give a quantitative description of the sample but that won't take us very far
 - So we try to generalise from the sample to the target population
 - Population of stressed people and population of relaxed people

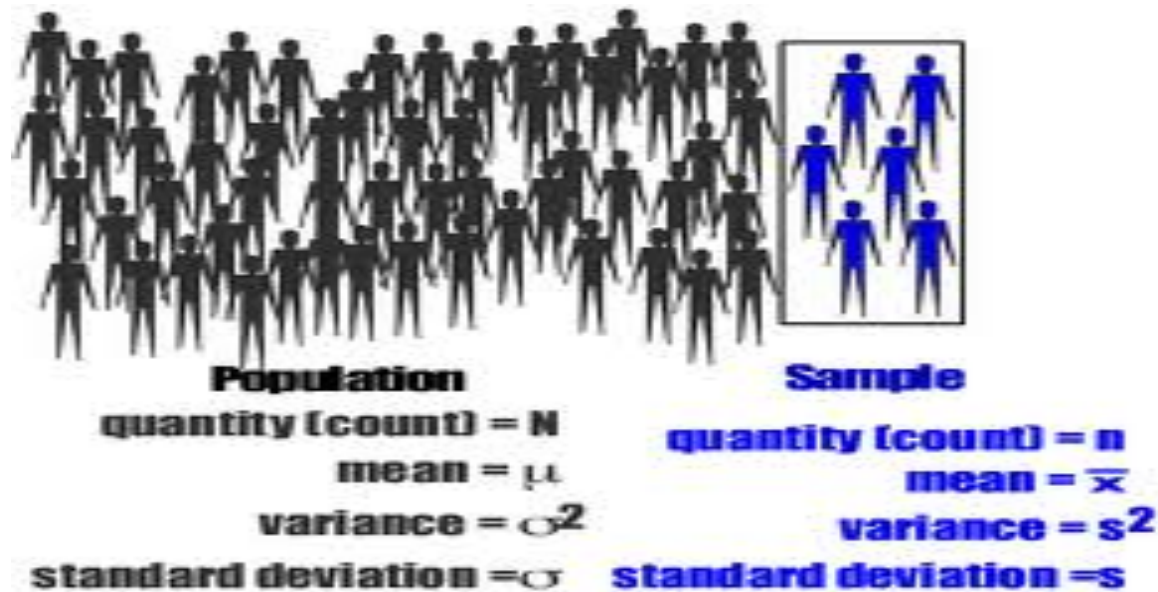
Population & Sample

Population

- All the people of interest
- Includes ALL POSSIBLE OBSERVATIONS
- Greek Letters

Sample

- A set of observations from a population
- Roman Letters



Populations and Samples

- **Target Population:**
- the people/animals/microbes/fungi of interest

Accessible population:

portion of target population that is available to the researcher

- **Sample:**
- a subset of the population that we observe

Two Branches Of Statistics

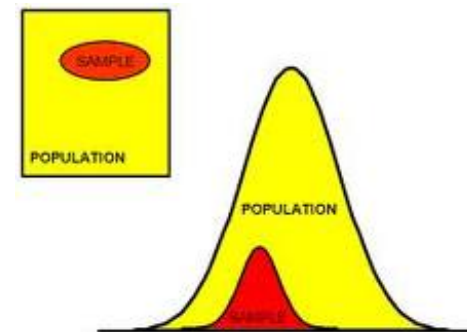
Descriptive

- Organize
- Summarize
 - Average
 - Range
 - Percentage
- Communicate

... a body of observed data
using numerical techniques
and graphical techniques

Inferential

- Using sample data to make estimates of the rest of the population (interpret data, predict behaviour)
- Can infer only from a Sample to the Population



The problem with samples

- Samples almost never represent populations perfectly (sampling error)
 - Not really an error
 - Just the natural variability that you can expect from one sample to another
 - Difference is shown by central tendency and variability

Observations to Variables

- They are called variables because they are properties that vary.
 - The outcomes of our measurements or observations vary:
Dependent Variable
 - A property can be varied by the experimenter (e.g. to compare different populations: **Independent Variable**
 - The independent variable has: values or **levels** or **conditions**
e.g. number of yawns (dependent variable) in male/female students, on Budapest campus/main campus, at 8a.m./before/after lunch/at 4 p.m.
- Confounding Variable
 - Variable that systematically varies with Independent Variable
 - Can only be controlled by good design or by limiting scope of conclusion, or statistical correction

Why do students yawn?

- Hypothesis: Statistics makes students yawn
- Dependent variable: number of yawns in a class period
- Independent variable: class type (Psycholinguistics, Neurobiology, Statistics)
- Confounding variable: Statistics is always on Thursday afternoon, Psycholinguistics on Monday morning and Neurobiology on Wednesday morning. We suspect that day of week and time of day may have a confounding effect.

Hypothesis testing

- Theory
- Hypothesis: Statement of possible relationship between variables which follow from theory
- The null hypothesis
 - A statement that there is no relationship.
- The research hypothesis
 - A statement that there is a relationship
- Hypothesis testing: a process of drawing conclusions about whether a particular relation between variables is supported by evidence.
 - Find out the probability of the null hypothesis being true

Statistical Decisions

- Either Reject or Fail to Reject the null hypothesis
 - Rejecting the null hypothesis suggests that there is a relationship between the populations sampled
 - Failing to reject suggests that no relationship exists
 - Decision is based on probability
 - **Alpha:** the statistical decision criteria used in testing the null hypothesis
 - Traditionally, alpha is set to small values (.05 or .01)
- Always a chance for error in our decision

Statistical Decision Process

	Reject Null Hypothesis	Retain Null Hypothesis
Null Hypothesis is True	Type I Error (False alarm)	Correct Decision (Correct rejection)
Null Hypothesis is False	Correct Decision (Hit)	Type II Error (Miss)

Operational definitions

An operational definition specifies the procedures used to manipulate or measure a variable

How to specify the independent variable(s)

e.g. differences between young and old people in solving math problems: who is old and who is young?

How to measure the dependent variable(s)

what differences? Success rate? Speed? Method?

MEASUREMENT SCALES

Measuring Variables

- **Measurement:** Assigning numbers to indicate the level of a variable
 - Sometimes the number assignment is intuitive (e.g., time measured in seconds)
 - Sometimes it is more arbitrary (e.g., 1 for male and 2 for female)

Scales of Measurement

- Based on how closely the measurement scale matches the real number system
- Scales of Measurement
 - Nominal
 - Ordinal
 - Interval
 - Ratio

Nominal Scales

■ Naming scale

- Each number reflects a category
- *Examples:*
diagnostic categories (*jet lag, flu, hangover*),
political affiliations (*Conservative, Liberal, Green*)

■ Produces **nominal** or **categorical** data

Ordinal Scales

- Scale indicating rank order
 - Reflects the order, but not the amount of difference between two consecutive items
 - *Examples:*
 - order of finish in a race (1st, 2nd, 3rd),
 - class rankings (1st, 2nd, 3rd, 4th, etc)
 - rating scales (always, often, sometimes, rarely)
- Produces **ordinal** data

Interval Scales

- Scale with equal intervals
 - The scale indicates amount, but with no natural zero point
 - *Examples:*
temperature on the Celsius scale (0 doesn't mean that there is no temperature),
personality tests
- Produces score or **scale** data

Ratio Scales

- Scale that fits the number system well
 - Includes equal intervals and a meaningful zero
 - *Examples:*
 - time,
 - distance,
 - cognitive tests (e.g. reaction time, math score)
- Produces score or **scale** data

Which level of measurement?

- Age
- Place of birth
- Rent paid per month
- Order of arriving at college
- Number of Facebook friends
- Favourite colour
- Eye sight (poor, fair, good, excellent)
- Answers on a Likert-scale (How happy are you?)
- IQ
- How long it takes to get to university (minutes)

RELIABILITY AND VALIDITY

Problems with the dependent variable

- Validity,
- Reliability,
- Scale attenuation effects

Reliability

- test-retest reliability: the measurements are consistent, you get the same results every time you take the measurement (under the same conditions)
- inter-rater reliability: level of agreement between two (or more) observers

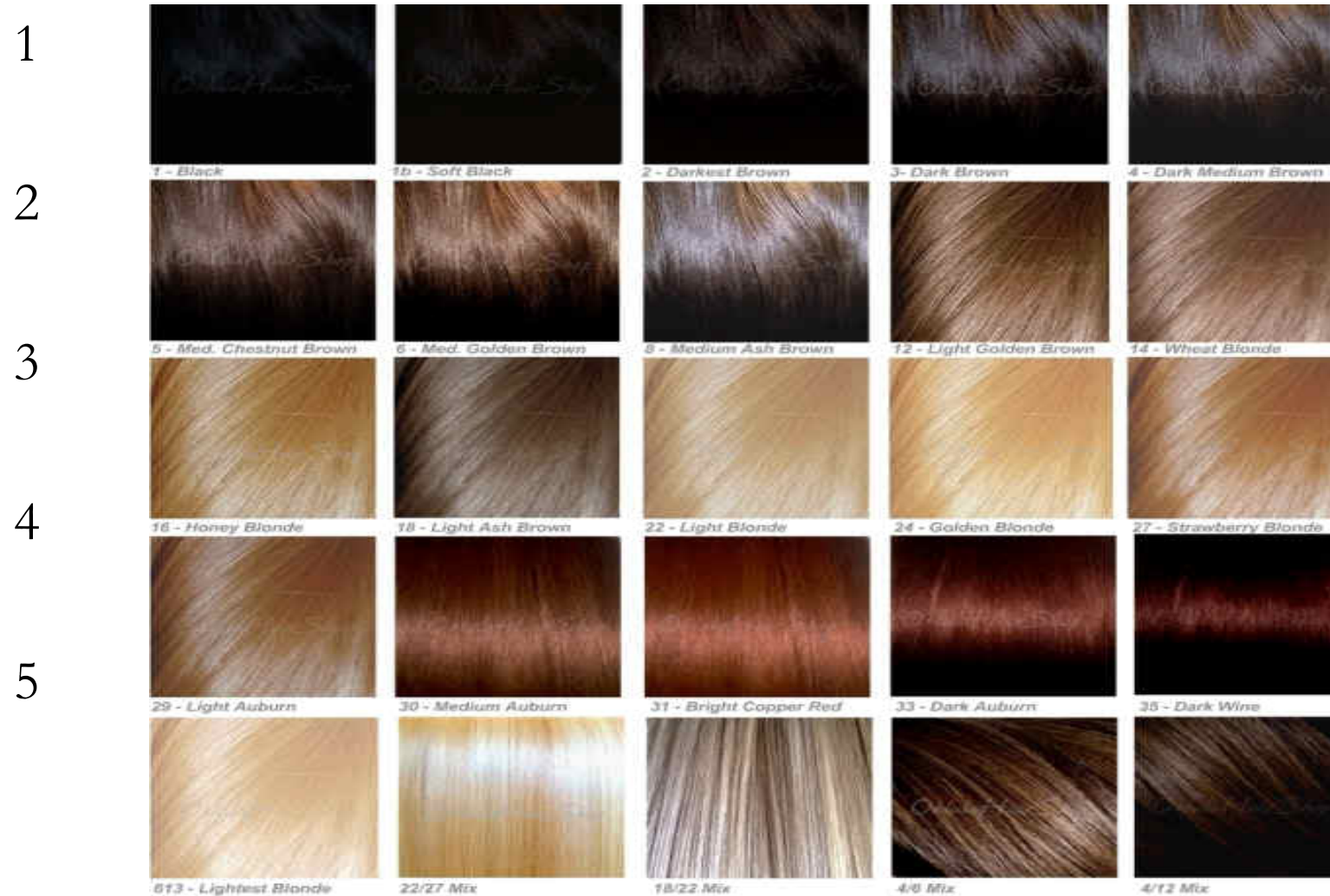
Influenced by

- how objective the measurement is
- how well-defined the measurement is

1. Black, brown, blond, red

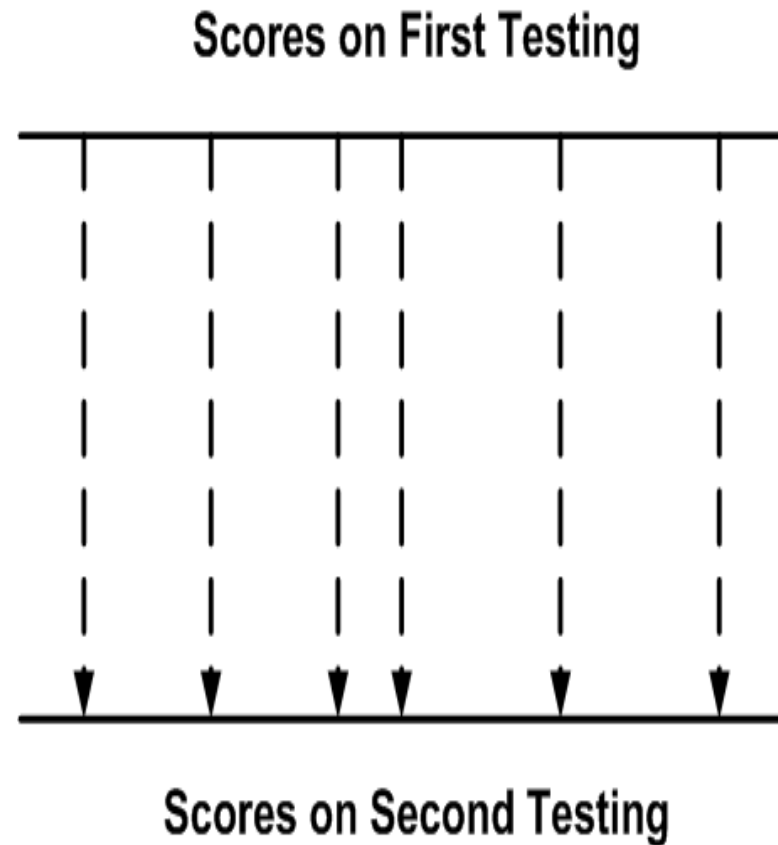
2. raven, dark brown, auburn, chestnut, golden, copper, ginger, fair, ash

• A B C D E



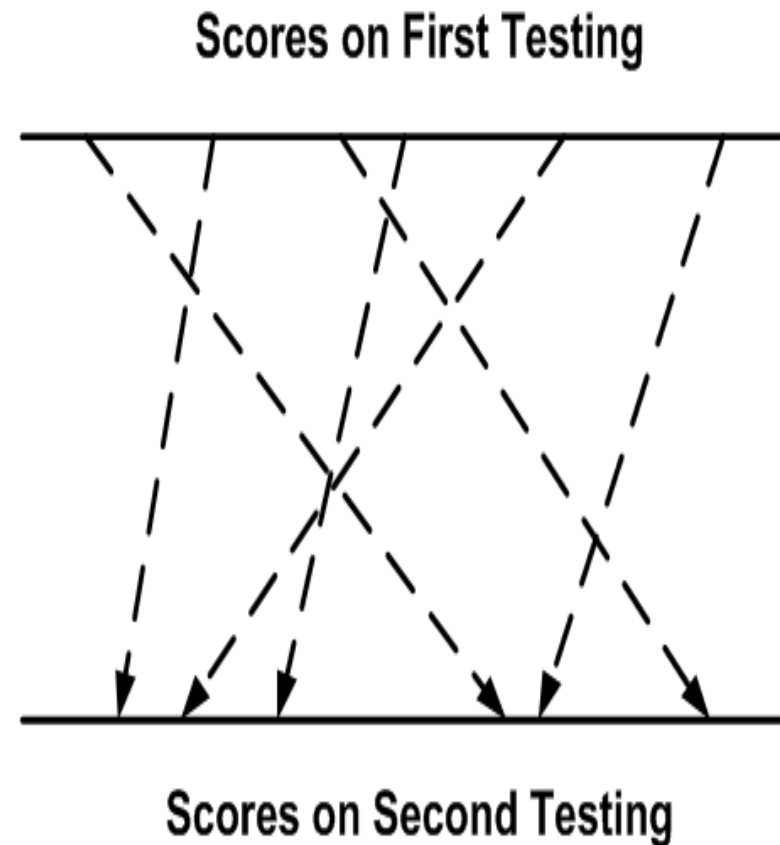
Perfect Reliability

- Reliability is a measure of consistency.
- Perfect reliability means that the scores are perfectly consistent.



Poor Reliability

- When you reach the point where the rank orderings have no relationship to one another, your reliability (i.e., consistency) is poor.



Scale Attenuation Effects

- Results from a restriction of the range of the measure
 - Therefore, people above or below the effective range are not measured accurately
- Two types of scale attenuation effects
 - Floor effects
 - Ceiling effects

English tests

Test 1

1. What's the past tense of 'give'?
2. How do you say hello in the morning?
3. List your favourite foods.

Test 2

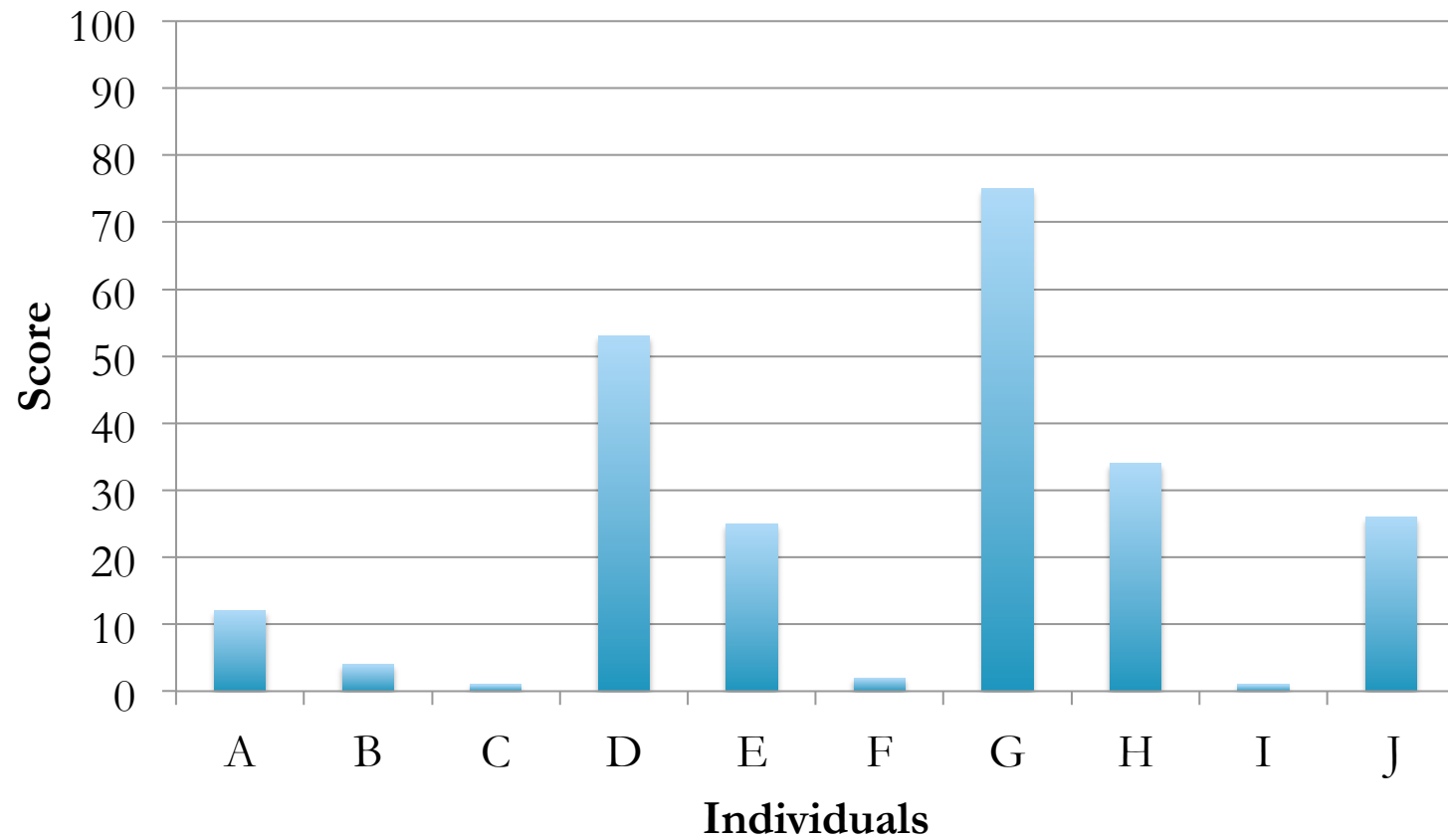
- Analyse the following passage:
- The longe love that in my thought doeth harbar :
- And in my hert doeth keep his residence :
- Into my face preseth with bolde pretence :
- And therein campeth spreding his baner.
- She that me lerneth to love and suffre :
- And willes that my trust and lustes negligence
- Be rayned by reason, shame, and reverence :
- With his hardines taketh displeasur.
- Where with all unto the hertes forrest he fleith :
- Leving his enterprise with payn and cry :
- And ther him hideth and not appereth.
- What may I do when my maister fereth ?
- But in the feld with him to lyve and dye ?
- For goode is the liff, ending faithfully.

Floor Effect

- Range on the bottom of the scale is insufficient to measure people who score near the bottom
- The test is too difficult for them

Floor Effect

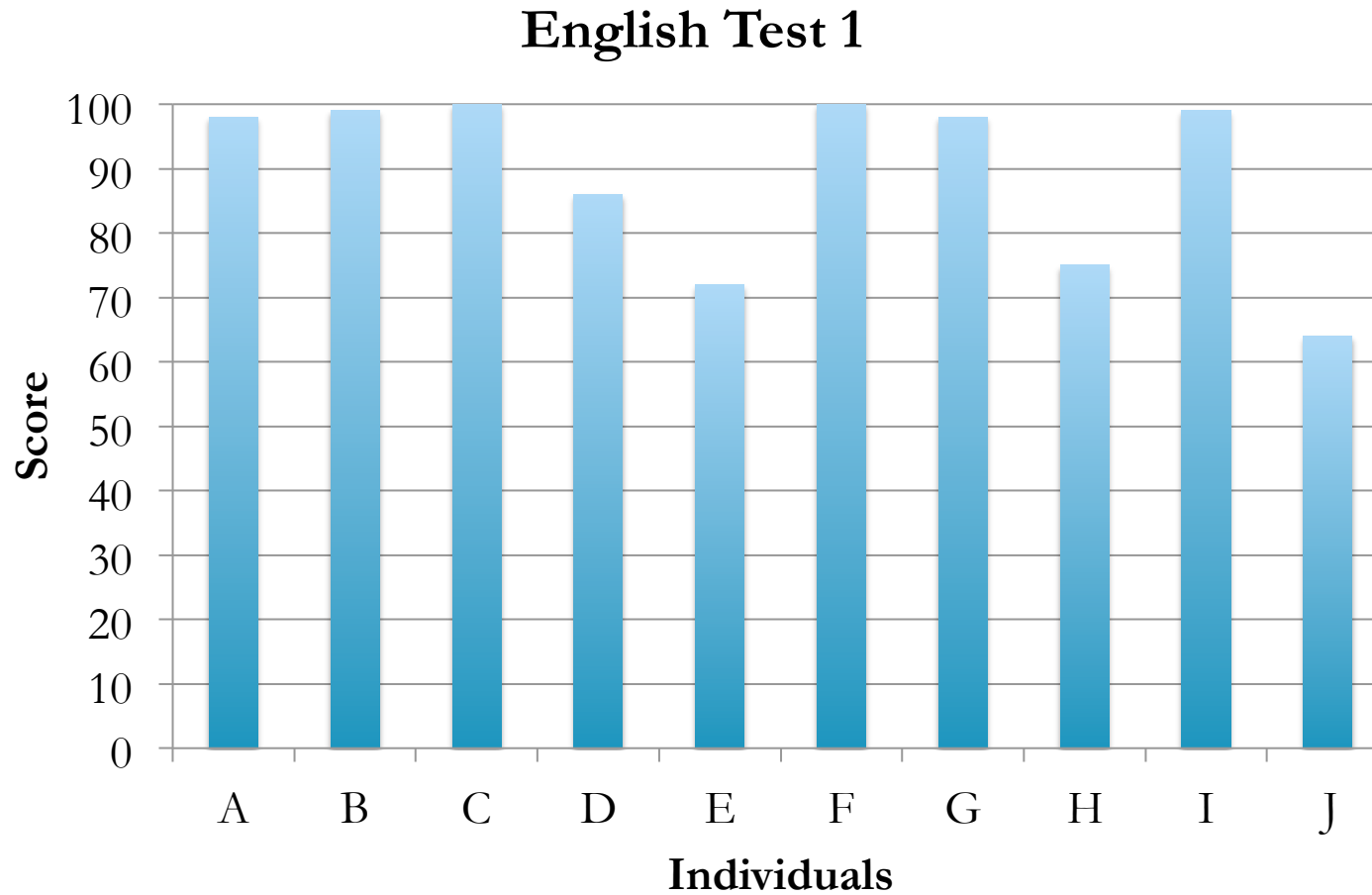
English Test 2



Ceiling Effect

- The range on the top of the scale is insufficient to measure people who score near the top
- The test is too easy for them

Ceiling Effect



Validity

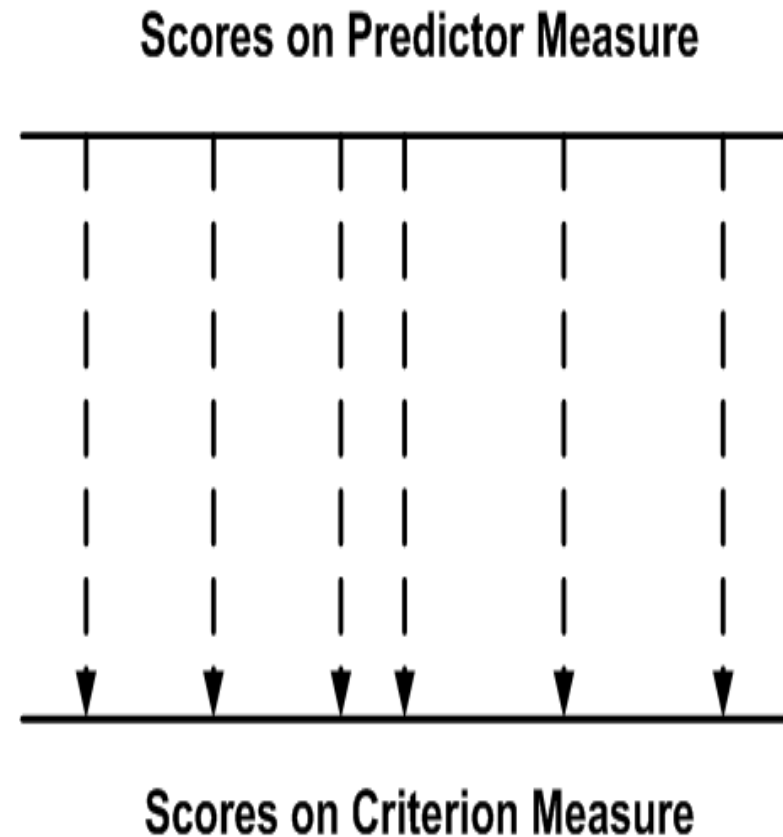
- A scale is valid if it measures what it is supposed to measure
- Validity also refers to how well a scale predicts other variables
 - *Example:* An IQ test is likely to be a valid predictor of grades in school.
 - When used this way, the scale is called the *predictor measure* and the measure predicted is called the *criterion*.

Measuring Validity

- Validity is usually measured with a correlation
 - The correlation is between
 - The measure and
 - A SPECIFIED criterion measure
 - Always list the criterion when reporting the level of validity
 - *For example*, the IQ test is a valid predictor of school grades, but not a valid predictor of athletic ability.

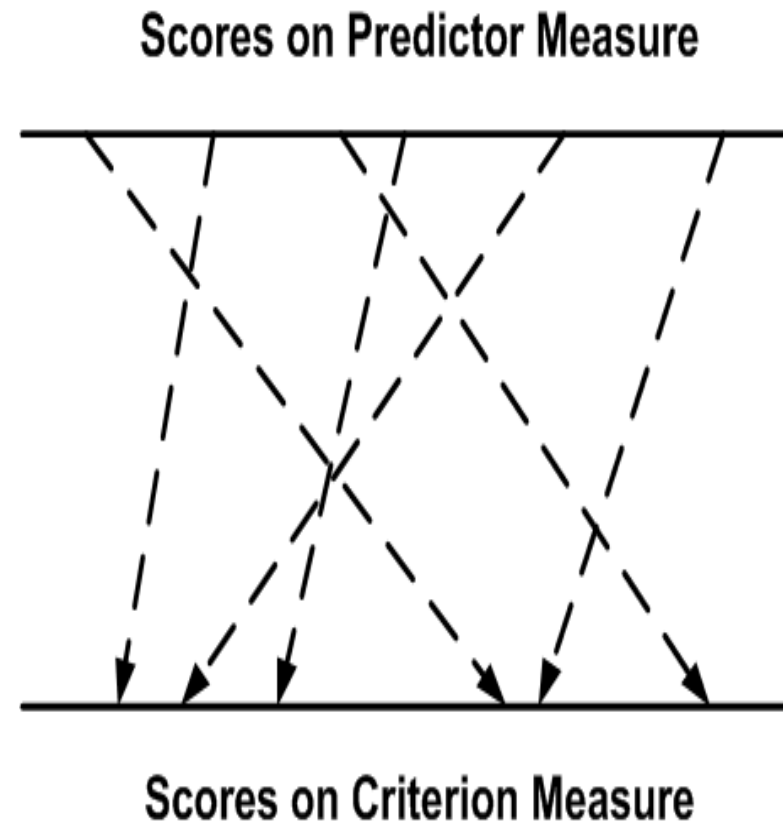
Perfect Validity

- Validity is the degree to which one measure predicts another.
- With perfect validity, the rank orderings on the predictor and criterion measures are identical.



Poor Validity

- When you reach the point that the rank orderings of predictor and criterion measures are unrelated, you have essentially zero validity.



Homework

- Make sure you have Excel, Numbers or Google Sheets
- Make sure you have SPSS
- Collect data for 3 adults (NO NAMES):
 - Gender (male/female)
 - Height (cm)
 - Shoe size
 - Hair colour
 - Eye colour
 - Number of Facebook friends
 - How many hours of sleep they got on Tuesday night.
 - How many hours of sleep they got on Saturday night.
 - Did they go out last Saturday? (yes/no)
 - How many units of alcohol they had last week (1 unit = 1 pint of beer, 1 glass of wine, 1 vodka, etc.)
- **Work with [Google Sheet](#) file Homework_1_Data2017: one row for each participant and the variables in different columns (in above order)**
- **Before next class, make sure to have the file with all the data**