Basic Concepts & Graphical Displays
Week 1
Prof. Patrick Bennett
Overview

• Part 1 - Introduction: What is Statistics?
• Part 2 - Basic Concepts: Variables & measurement scales
• Part 3 - Graphical displays & shapes of distributions
Part 2 Basic Concepts
Independent Variables

• Most Psychology studies collect data in designed experiments
• Experiments usually involve collection of data in various experimental conditions
  - conditions differ in terms of 1 or more independent variables
  - e.g., conditions in memory experiment defined by:
    ▸ type of items (faces vs words) being studied
    ▸ time interval between study & test phases
Dependent Variable

• The variable(s) that is(are) measured and constitute the data, or results, that will be analyzed.

• Designed experiments measure the effects of independent variables on dependent variables.
Examples of dependent variables include reaction time, response accuracy, number of items recalled in memory test, event-related brain potentials, heart rate, number of offspring, number of aggressive or affiliative behaviours, etc.
Hypothetical Recognition Memory Experiment

2 independent variables

2 levels on each independent variable

Study Items

<table>
<thead>
<tr>
<th>faces</th>
<th>words</th>
</tr>
</thead>
</table>

Retention Interval

<table>
<thead>
<tr>
<th>short (1 minute)</th>
<th>long (1 hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td># of items</td>
<td># of items</td>
</tr>
<tr>
<td>correctly</td>
<td>correctly</td>
</tr>
<tr>
<td>recognized during</td>
<td>recognized during</td>
</tr>
<tr>
<td>test</td>
<td>test</td>
</tr>
<tr>
<td># of items</td>
<td># of items</td>
</tr>
<tr>
<td>correctly</td>
<td>correctly</td>
</tr>
<tr>
<td>recognized during</td>
<td>recognized during</td>
</tr>
<tr>
<td>test</td>
<td>test</td>
</tr>
</tbody>
</table>

1 dependent variable measured in 4 experimental conditions defined by combinations of independent variables
Correlational Studies

• Correlational studies measure the association between variables
• Usually, the variables are not manipulated by investigator
  - each event/subject comes with own set of variables
  - but values on variables differ across events/subjects
• Regression measures association between predictor & criterion variables
  - e.g., measure the association between annual income (criterion variable) and parent’s income, years of education, race, gender (predictor variables)
Random Assignment

• In Psychology, designed experiments use subjects that also come with their own set of intrinsic characteristics

• These characteristics (personality, motivation, intelligence, etc.) probably affect dependent variable

• HOWEVER, in most experiments, subjects are randomly assigned to experimental conditions

• So, effects of subject differences should be UNRELATED TO EFFECTS OF INDEPENDENT VARIABLES
  - big advantage of designed experiments over correlational studies
Between-Subjects Experimental Design

2 independent variables

<table>
<thead>
<tr>
<th>Study Items</th>
<th>Retention Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>faces</td>
<td>short (1 minute)</td>
</tr>
<tr>
<td>n = 12 Ss</td>
<td>long (1 hour)</td>
</tr>
<tr>
<td>n = 12 Ss</td>
<td></td>
</tr>
<tr>
<td>words</td>
<td></td>
</tr>
<tr>
<td>n = 12 Ss</td>
<td></td>
</tr>
</tbody>
</table>

Random assignment reduces chance that observed differences among conditions are due to differences among subjects.

48 subjects randomly assigned to 1 of 4 conditions (12 Ss per condition)
Categorical vs. Numerical Variables

- **Categorical variables** typically have a few **discrete** levels:
  - sex (male or female); marital status (single, married, divorced); brain area (e.g., V1, V2, V5); profession; age (e.g., young, middle-age, elderly)
  - often, though not always, they lack numeric properties

- **Numerical variables** typically have many levels (e.g., reaction time, body weight, income, age (years)) though not always (e.g., number of offspring).
  - have obvious numeric properties: the levels are **ordered**, differences between levels are meaningful, etc.

- In truth, “categorical” and “numerical” are at ends of a continuum
  - some variables possess both categorical & numerical properties
## Between-Subjects Experimental Design

Note that ‘Retention Interval’ could be categorical or numeric independent variable.

<table>
<thead>
<tr>
<th>Study Items</th>
<th>faces</th>
<th>words</th>
</tr>
</thead>
<tbody>
<tr>
<td>recognition memory</td>
<td>recognition memory</td>
<td>accuracy recognition memory</td>
</tr>
<tr>
<td>accuracy</td>
<td>accuracy</td>
<td># of items correctly recognized is a numeric dependent variable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retention Interval</th>
<th>short (1 minute)</th>
<th>long (1 hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>recognition memory</td>
<td>recognition memory</td>
<td>accuracy recognition memory</td>
</tr>
<tr>
<td>accuracy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of items correctly recognized is a numeric dependent variable.
Measurement Scales

• nominal, ordinal, interval, & ratio
• type of scale determines the kinds of mathematical operations (addition, subtraction, multiplication, division) that can be applied to the data to yield results that make sense
• when thinking about scales, it is important to distinguish between i) actual measurements of stimuli, subjects, and behaviour; and ii) the psychological phenomena that underlie the observable, measurable quantities
  - e.g., the difference between sound intensity and loudness
  - e.g., the difference between light wavelength and colour
Nominal Scale

• Nominal (categorical) examples:
  - eye colour: blue, brown, green, etc.
  - gender: male, female
  - political party affiliation: Liberal, Conservative, NDP, None

• Levels of nominal variable have no obvious numeric properties
Ordinal Scale

• Ordinal Scale examples:
  - age: child, adolescent, young-adult, middle-age, elderly
  - musical expertise: none, novice, amateurs, experts
  - Likert scales... Rate your response to “X”:
    ‣ 1) strongly agree; 2) agree; 3) neutral; 4) disagree; 5) strongly disagree

• **Order** of levels is meaningful

• Differences between levels may not be meaningful
Saffir-Simpson Hurricane Scale (SSHS)

Why even a record-breaking hurricane can't hit Category 6

'Once you say catastrophic and there's near complete damage, why do you need a 6?'

- **Category 1 (119-153 km/h)**: Very dangerous winds will produce some damage. Well-constructed frame homes could have damage to roof, shingles, vinyl siding and gutters. Large branches of trees will snap and shallowly rooted trees may be toppled. Extensive damage to power lines and poles likely will result in power outages that could last a few to several days.
- **Category 2 (154-177 km/h)**: Extremely dangerous winds will cause extensive damage. Well-constructed frame homes could sustain major roof and siding damage. Many shallowly rooted trees will be snapped or uprooted and block numerous roads. Near-total power loss is expected with outages that could last from several days to weeks.
- **Category 3 (178-208 km/h)**: Devastating damage will occur. Well-built frame homes may incur major damage or loss of roof decking and gable ends. Many trees will be snapped or uprooted, blocking numerous roads. Electricity and water will be unavailable for several days to weeks after the storm passes.
- **Category 4 (209-251 km/h)**: Catastrophic damage will occur. Well-built frame homes can sustain severe damage with loss of most of the roof structure and/or some exterior walls. Most trees will be snapped or uprooted and power poles downed. Fallen trees and power poles will isolate residential areas. Power outages will last weeks to possibly months. Most of the area will be uninhabitable for weeks or months.
- **Category 5 (252 km/h or higher)**: Catastrophic damage will occur: A high percentage of frame homes will be destroyed, with total roof failure and wall collapse. Fallen trees and power poles will isolate residential areas. Power outages will last for weeks to possibly months. Most of the area will be uninhabitable for weeks or months.
Interval Scales

• Numeric variables usually with many levels
• but, interval scales lack a true, meaningful zero:
  - e.g., time of day [zero point is arbitrary]
  - e.g., IQ [zero doesn’t imply zero intelligence]
• consequently, differences between levels are meaningful:
  - difference btw IQs 100 & 105 is same as 115 & 120
  - ratios are not particularly meaningful:
    ‣ temperature: 30/10 deg C not equal to 86/50 deg F
• Most standardized psychological tests use interval scales
Ratio Scales

- Numeric variables:
  - reaction time (RT), proportion correct, body weight
- differences & ratios between levels are meaningful
- ratio scales have a true, non-arbitrary ZERO on the scale
- addition, subtraction, multiplication, & division are valid mathematical operations
Physical vs. Psychological Measurement

• often our dependent variables are indirect measures of psychological constructs (e.g., ratings of happiness vs. true, internal happiness)

• we may treat dependent measure as interval or ratio variable without assuming that psychological variable exists on interval or ratio scale

• but then our conclusions are about the measure, not the underlying psychological variable
  - we operationally define the psychological variable as our measure
  - a person’s “happiness” is her self-reported rating of happiness

• and this limitation probably is OK!
Part 2 Summary

• Variable types
  - independent vs dependent
  - predictor vs criterion
• Data types: continuous vs discrete
• Random Sampling & Random Assignment
• Measurement Scales: nominal, ordinal, interval, ratio
  - physical vs psychological measures (e.g., intensity vs loudness)