Types of Questions for the Mid-term

- (1) Small questions – short answer
- (2) Interpretation of graphs, results, design
- (3) Provide a critic of a design/experiment
- (4) Open question: design an experiment to best answer the question

Experimental summary:

Joe wanted to see how well people were able to predict the temperature tomorrow, based on their experience of the weather prior to that day. On 10 random days of the calendar, Joe asked subjects to predict the temperature for the next day, and he recorded their answers and the actual temperature that following day. Joe reasoned that he could do a paired (aka dependant, matched) t-test to determine whether the predictions were accurate. If the t-test had a low p-value ($p < 0.05$), then there was a statistical difference between the actual temperatures and the predicted temperatures, and therefore, the predictions were inaccurate. If the t-test had a high p-value, then there was no statistical difference between the two conditions, therefore the predictions were accurate. Here is what he got:

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>69</td>
</tr>
<tr>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>30</td>
<td>85</td>
</tr>
<tr>
<td>40</td>
<td>53</td>
</tr>
<tr>
<td>50</td>
<td>21</td>
</tr>
<tr>
<td>60</td>
<td>43</td>
</tr>
<tr>
<td>70</td>
<td>101</td>
</tr>
<tr>
<td>80</td>
<td>9</td>
</tr>
<tr>
<td>90</td>
<td>29</td>
</tr>
<tr>
<td>100</td>
<td>94</td>
</tr>
</tbody>
</table>

$p > 0.90$

With this whopping huge p-value, Joe concluded that subjects were quite accurate in their predictions, since there is no statistical difference between Actual and Predicted temperatures.

Perception and Memory

"Yes-No" paradigms

- A research domain where SDT has been successfully applied is in the study of memory. Typically in memory experiments, participants are shown a list of words and later asked to make a "yes" or "no" statement as to whether they remember seeing an item before. Alternatively, participants make "old" or "new" responses. The results of the experiment can be portrayed in what is called a decision matrix.

- The hit rate is defined as the proportion of "old" responses given for items that are Old and the false alarm rate is the proportion of "old" responses given to items that are New.

What does $d'$ measure?

- $d'$ is a measure of sensitivity. Sensitivity (discrimination) $d' = |z(H) - z(F)|$ (z score difference between HIT and FA)

- How do you interpret $d'$ value?
  - The larger the $d'$ value, the better your performance. A $d'$ value of zero means that you cannot distinguish trials with the target from trials without the target. A $d'$ of 4.6 indicates a nearly perfect ability to distinguish between trials that included the target and trials that did not include the target.

- How do you interpret response bias in a memory experiment?
  - Response bias $c = -0.5[z(H) + z(F)]$
  - $C$ is a measure of response bias. A value greater than 0 indicates a conservative bias (a tendency to say "absent" more than "present") and a value less than 0 indicates a liberal bias. Values close to 0 indicate neutral bias.
Consider two participants in a recognition memory test. Participant A has a hit rate of .70 and participant B has a hit rate of .75. Can we be sure that B has a better memory than A? What if participant A had a false alarm rate of .05 and participant B had a false alarm rate of .24? This indicates that B is more willing to say "old." Would you still think that B is doing better than A? If we look at only the hit rate, we overlook the fact that participant B might be doing better just because he is willing to say "old" more often, and hence, he is getting more hits at the expense of more false alarms.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Hit Rate</th>
<th>False Alarm Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>.70</td>
<td>.05</td>
</tr>
<tr>
<td>B</td>
<td>.75</td>
<td>.24</td>
</tr>
<tr>
<td>C</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Participant C is an extreme case, with a hit rate of 1.00. He attains this rate by saying "old" to every item. Great Job! Or is it? This performance wouldn’t really say anything about his memory ability, however, because his false alarm rate would also be 1.00. This performance wouldn’t demonstrate any discrimination between Old and New items.

**Interaction interpretation**

- Subjects were presented with 36 non words trigrams (NTG), one at a time.
- Factor 1: Modality of presentation
  - 1) Auditory + Visual
  - 2) Auditory only
  - 3) Visual only
- Factor 2: Retention interval (before starting reporting the trigrams)
  - 0,3,6,9,12 or 18 seconds (6 levels)
- Task: a free recall task of the trigram
- Method: 1 trigram is presented; 1 second delay; subjects heard a 3 digits number (except in the zero interval condition); subjects were asked to count backgrounds per threes aloud; after the retention interval, participants recall the trigram.
- **Dependent variable**: proportion of correct report

**Issue with interaction interpretation**

- Problem: performance at the 0 second retention interval is very nearly perfect in all conditions. When performance is perfect (ceiling effect) it is impossible to tell whether there are any real differences among conditions because of the scale attenuation.
- **General Rule about interaction**: Extreme caution should be used in interpreting interactions where performance on the dependent variable is at either the floor or the ceiling at some level of one of the independent variable

**Question**

- Standing (1971) tested the visual memory of people by presenting 10,000 images. At the test, the participants were presented with old and new images, one image at a time, and ask to say if the picture was in the test. Overall participants’ accuracy (for old pictures) was 92%. Standing concluded that memory for pictures was outstanding.
- You are working for an advertising company and the company asks you to put together a memory test, to test the visual memory of their ads.
- Describe an experimental design which will test the memory of visual ads.

**Design**

- Independent variable(s)?
- Dependent variable(s)?
- Within or between groups? Why?
- Control conditions/groups?
- Confounded factors?
- How many subjects / group?
- Who are the subjects?
- What are the stimuli?
- Hypothesis
- Task
- Predictions of the results
- Which statistics do you run?
- Interpretation of the results