Two-Sample Inferential Statistics

Greg C Elvers

Two-Sample Inferential Statistics

- In an *experiment* there are two or more conditions
  - One condition is often called the *control condition* in which the treatment is *not* administered
  - The other condition is often called either the *treatment condition* or the *experimental condition*; the treatment is administered

Example of an Experiment

- Some people see a brief description of the subject of a passage
- Other people see nothing
- Both groups then hear a passage
- The groups then rate their comprehension of the passage and then recall it
- What is the control group?
- What is the experimental group?

Question Asked in an Experiment

- The basic question asked in such an experiment is whether the treatment caused an effect
- That is, are the values of the dependent variables (the measured variables) different in the treatment and control conditions?
- E.g., did having a context increase comprehension and recall of the passage relative to the control condition?

Two-Sample t-Tests

- Inferential statistics are used to answer the primary question in an experiment
  - The particular inferential statistic that you use depends on the experimental design of the study; more about that in Experimental Psychology
  - When there are just two conditions (control and experimental), you often want to use a *two-sample t-test*

4

Two-Sample t-Tests

- The two-sample t-test is not fundamentally different from the single-sample t-test that we have already discussed
- There are two primary differences:
  - You have two sample means instead of a single sample mean and a population mean
  - The population standard deviation is unknown, *and* you have *two* estimates of it
    - You have one estimate of the population standard deviation from each of the two sample standard deviations

6

Two-Sample t-Tests

$$t = \frac{\overline{X}_{exp} - \overline{X}_{control}}{s_{\overline{X}_{exp} - \overline{X}_{control}}}$$

$$s_{\overline{X}_{exp} - \overline{X}_{control}} = \sqrt{s^2_{\overline{X}_1} + s^2_{\overline{X}_2}}$$

$$df = n_1 + n_2 - 2$$

- The t-test formula says to take the difference of the means and then divide that by the standard error of the difference of the means
- Note: This formula is applicable only if $n_1 = n_2$

7

Standard Error of the Difference of the Means

- We want our estimate of the population standard deviation to be as accurate as possible
- To make it as accurate as possible, we should base it on as large of a sample as possible
- Under H_0, the two samples come from the same population, so we should use the data from both samples when we estimate the population standard deviation

8

Standard Error of the Difference of the Means

- We cannot simply add the two sample standard deviations together
- Rather, we should convert them to variances (by squaring them), sum the variances, and then convert the sum back to a standard deviation (by taking the square root)
- The above procedure produces the *standard error of the difference of the means*

An Example

- Two groups of students were asked to perform 5 simple tasks at specified times during the next hour
- One group tied a string (the external memory cue) around their finger to remind them that they had tasks to perform
- The other group had no external memory cue

An Example

- The number of tasks completed was recorded for each group

$$\overline{X}_{no\ memory} = 1.5, \quad s^2_{\overline{X}_{no\ memory}} = 1.75, \quad n_{no\ memory} = 12$$

$$\overline{X}_{memory} = 3.5, \quad s^2_{\overline{X}_{memory}} = 1.25, \quad n_{memory} = 12$$

Steps

- Write the hypotheses:
  - H_0: $\propto_{no\ memory\ cue} \geq \propto_{memory\ cue}$
  - H_1: $\propto_{no\ memory\ cue} < \propto_{memory\ cue}$
- Is it a one-tailed or two-tailed test?
  - One-tailed
- Specify the α level
  - α = .05