Two-Sample Inferential Statistics

Greg C Elvers

1

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

2

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?

3

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?

4

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*

5

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

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

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

Steps

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

Steps

- Calculate the appropriate statistic
  - There are two samples with ratio scaled data, or the two-sample t-test is appropriate

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

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

$$df = n_1 + n_2 - 2$$

Plug and Chug

$$s_{\overline{X}_{exp} - \overline{X}_{control}} = \sqrt{s_{\overline{X}_1}^2 + s_{\overline{X}_2}^2} = \sqrt{1.75 + 1.25} = 1.732$$

$$t = \frac{\overline{X}_{exp} - \overline{X}_{control}}{s_{\overline{X}_{exp} - \overline{X}_{control}}} = \frac{3.5 - 1.5}{1.732} = 1.15$$

$$df = n_1 + n_2 - 2 = 12 + 12 - 2 = 22$$

Steps

- Find the critical value in a table of critical t values
  - Find the column with the correct α level (.05) and the correct number of tails (1)
  - Find the row with the correct degrees of freedom (22)
    - If the row with the correct degrees of freedom does not exist, use the next smallest degrees of freedom
  - The critical t is at the intersection of the row and column
  - Critical $t(22)_{\alpha=.05,\text{ 1-tailed}} = 1.717$

Steps

- Decide whether we can reject H_0 or not
- If the calculated t value is greater than or equal to the critical t value, then you can reject H_0
- 1.15 (calculated t) is not greater than or equal to 1.717 (critical t)
- We fail to reject H_0
  - We cannot conclude that the treatment caused an effect

Using a Computer

- Using a computer to perform t-tests is easy
- The computer will print out the value of t, its degrees of freedom, and a *p value*
- The p value is the probability of observing a difference between the means this large due to chance
- When the p value is less than or equal to the specified α level, you can reject H_0

17

Within-Subjects t-Tests

- Simple experiments can be either *between-subjects* or *within-subjects designs*
- A between-subject design occurs when different people participate in the control and experimental condition
  - Thus, the two sample standard deviations should be independent of each other

18

Within-Subjects t-Tests

- A within-subjects design has the same people in both the control and experimental condition
  - The two sample standard deviations should not be independent of each other because each individual provides two scores, one in each condition
  - It is inappropriate to use the preceding formula for the standard error of the difference of the means when the sample standard deviations are not independent of each other

19

Within-Subjects t-Tests

- When using a within-subjects design, the formula for the standard error of the difference between the means becomes:

$$s_{\bar{X}_{cont}-\bar{X}_{exp}} = \sqrt{s^2_{\bar{X}_{cont}} + s^2_{\bar{X}_{exp}} - 2r\, s_{\bar{X}_{cont}}\, s_{\bar{X}_{exp}}}$$

20

Within-Subjects t-Tests

- The formula simply corrects for the fact that when the standard errors of the mean are correlated, they convey less new information than when they are not correlated

Within-Subjects t-Tests

- The degrees of freedom are reduce in the within-subjects design experiment
- For the within-subjects t-test, the degrees of freedom is the number of pairs of scores minus1

Between- vs Within-Subjects Designs

- In general, within-subjects designs are more powerful
  - That is, you are more likely to reject H_0 when H_0 is false with a within-subject design compared to a between-subjects design
- This arises because the standard error of the difference of the means tends to be smaller in a within-subjects design compared to a between-subjects design

Between- Vs Within-Subjects Designs

- However, because the within-subjects design has fewer degrees of freedom than does the corresponding between-subjects design, the critical t value will often be larger in the within-subjects design
- In general, the smaller size of the standard error of the difference of the means will more than compensate for the larger critical t that is due to the smaller degrees of freedom

Within-Subjects t-Tests

- When using the computer to analyze the results of a study, always be sure to use the correct statistical procedure
- SPSS calls within-subjects t-tests a *paired samples t-test*
- SPSS calls between-subjects t-tests an *independent samples t-test*

25

Assumptions of Student's t

- Student's t test makes several assumptions that should be verified prior to accepting the results of a t-test
- If the assumptions are violated, then you may not be safe in making any conclusions based on the t-value

26

Homogeneity of Variance

- The assumption of *homogeneity of variance* states that the variance in each condition (i.e. sample) should be equal
- It is possible to perform a statistical test to determine if the variances are probably different from each other
  - If the variances are found to be different, then a modified form of the t-test should be used
  - Most statistics programs automatically calculate both the normal t-test and the heterogeneous variance t-test

27

Normal Distributions

- The data in each condition should be normally distributed
  - This is often the case in the social sciences
  - Unless the distribution is very non-normal, the t-test will still give good results
    - That is, the t-test is *robust* -- it will give good results even when some of its assumptions are violated

28

Samples are Independent

- The data in the two samples should be independent of each other
  - This assumption is violated in within-subjects designs
  - Again, the t-test is robust to violations of this assumption