Problem Set #6: Independent Samples t-tests

1. Research suggests that people are more likely to gamble if they are in a good mood. Professor Keno and Professor Roulette want to replicate this finding. They agree to manipulate mood by showing their subjects hilarious cat videos on youtube. They also agree to assess willingness to gamble using a standard measure in the field. However, Professor Keno uses two groups of subjects: one group watches a hilarious cat video before completing the gambling measure; the other does not. Professor Roulette has her subjects complete the gambling measure twice: once after watching a hilarious cat video and once after watching an emotionally neutral do-it-yourself video about home composting. Which professor – Keno or Roulette – constructed an experiment that should be analyzed as a paired sample t-test? Explain.

Roulette should analyze her data with a paired t-test because she measured her subjects on the DV (willingness to gamble) twice, once before and once after the cat video. Keno needs to use an independent samples t-test because she had different subjects in her two groups.

1. You and Biff are playing a heated game of Jacks, when the conversation invariably turns to who is the superior player.  You both know enough statistics to know that one game won't settle the matter completely.  So, you each play six games, you pick up 5, 4, 8, 4, 5, and 6 jacks.  Biff picks up 4, 4, 5, 3, 4, and 5 jacks.
	1. Conduct a hypothesis test (alpha = .05) to determine who is the better jackster (Steps 1 through 8).  Be sure to interpret your results and to report the test statistic correctly. Set alpha to .05

This is an independent samples t-test because we have two groups with different “subjects” -in this case games. We must use t because of the small sample size and unknown population SD. The first step is to find the standard deviation for each sample, which we must use to compute SE for the sampling distribution. I would use the short-cut formula to do this (the math is simpler)

|  |  |
| --- | --- |
| You | Biff |
| 5, 4, 8, 4, 5, 6 | 4, 4, 5, 3, 4, 5 |
| Σ(x) = 32 | Σ(x) = 25 |
| Σ (x2) = 182 | Σ (x2) = 107 |
| Mean = 5.33 | Mean = 4.17 |
| s = 1.51 | s = .75 |

Step 1: Decide whether you are conducting a one- or a two-tailed test.

You should do a two-tailed

Step 2: Specify the ***NULL*** hypothesis (HO)

Ho: μy = μb

Step 3: Specify the ***ALTERNATIVE*** hypothesis (HA)

Ho: μy ≠ μb

Step 4: Designate the rejection region by selecting α.

α= .05

Step 5: Determine the critical value of your test statistic

tcrit = ±2.228

Step 6: Use sample statistics to calculate test statistic.

(**you don’t need to pool the variance** because we have equal sample sizes, but if you did, you would do it this way)

s2p = n1-1 (s21) + n2-1 (s22) / n1 + n2 - 2

 = (6-1)(1.512) + (6-1)(.752) / 6 + 6- 2

 = [(5) (2.267) + (5) (.567)] / 10

 = 11.33 + 2.83 / 10

 = 14.167 / 10 = 1.42

Or this way

s2p = SS1 + SS2 / df1 + df2

 = 11.33 + 2.83 / 5 + 5

= 14.167 / 10 = 1.42

SE = √s2p / n1 + s2p / n1

 = √1.42/ 6+ 1.42 / 6

 = √.472 = .688

If you don’t pool the variance first and just use the original standard deviations **you get the same number (with rounding errors).** Make sure to think about whether you are working with standard deviation (s) or variance (s2). The standard error formula for independent samples t-tests uses variance- so if you have a standard deviation you will need to square it.

SE = √s21 / n1 + s22 / n1

 = √1.512 / 6+ .752/ 6

 = √.380+.094

 = √.474 = .688

tobs = (Meany - Meanb) / (SE)

 = 5.33 - 4.17 / .688

 = 1.16 / .688

 = 1.69

Step 7: Compare *observed* value with *critical* value

Because tobs does not fall in the rejection region, we would fail to reject the null: t(10) = 1.69, p > .05.

Step 8: Interpret your decision regarding the null

There is not a significant difference between your jack skills (M = 5.33) and Biff’s jack skills (M = 4.17), t(10) = 1.69, p > .05.

* 1. Calculate and report a measure of effect size

If you haven’t already, you need to pool the variance across the two samples to obtain the average variance (see above for calculations): s2p = 1.42

Cohen’s d =$\frac{│M1-M2│}{\sqrt{s\_{p}^{2}}}$ = $\frac{│1.16│}{\sqrt{1.42}}$ = 1.16/1.19 = .97, this is a large effect.

Why did we get a large effect size but fail to reject the null? Our sample size is super small!