Assignment 18: Two-group mean comparisons

Scenario: Charles Darwin, the father of evolutionary theory, was the first person to document the operation of natural selection. Simply stated, natural selection is a process by which organisms best suited to their environment become the ones most likely to survive and leave descendants.

The central argument of this theory starts from Darwin's observation of variations among animals. Experience with animal and plant breeding demonstrates that variations can be developed that are "useful to man," such as developing new breeds of dogs or domesticating farm animals. So, reasoned Darwin, variations must occur in nature that are favorable or useful in some way to the organism itself in its own struggle for existence. Favorable variations are ones that increase chances for survival and procreation. Those advantageous variations are preserved and multiplied from generation to generation at the expense of less advantageous ones. This is the process known as natural selection. The outcome of the process is an organism that is well adapted to its environment.

In February of 1898, there was a severe winter storm with rain, sleet and snow near Providence, RI. Fifty-nine English sparrows were found freezing and brought to the laboratory of Dr. Herman Bumpus at Brown University. Of those 59 sparrows, 35 survived and 24 died. Bumpus analyzed 9 characteristics of the birds to see if the surviving sparrows were markedly different from the birds that died (which would support natural selection). We will analyze one particular characteristic of the sparrows – the humerus length (length of the arm bone):

Group	n	mean	std. dev.	Min	Q25	Median	Q75	Max
Survivors	35	738.00	19.839	687.00	728.00	736.00	752.00	780.00
Dead	24	727.92	23.543	659.00	714.75	733.50	743.75	765.00

1) Write out the null and alternate hypotheses in this study. Briefly explain the consequences of making a type I or type II error in this study.

Null hypothesis: ______________________________

Alternate hypothesis: ______________________________

Type I error consequence: ______________________________

Type II error consequence: ______________________________

2) Sketch the sampling distribution we would get if we repeatedly took samples of 59 sparrows and calculated the difference in means between the survivors and the dead sparrows. Label the mean and standard error of your sampling distribution.

3) Are you concerned about any of the assumptions needed to run an independent samples t-test? List the assumptions and briefly explain how you might determine if each assumption is satisfied in this scenario.

4) Conduct an independent samples t-test. Report the p-value and state any conclusions you can make.

5) Calculate a 99% confidence interval for the difference in means. Interpret this interval.

Scenario: Can you text more quickly on your own phone than you could on another phone?

15 teenagers were asked to text "The quick brown fox jumps over the lazy dog" on two phones. Some were asked to first send the message on their own phone; then on another phone. Others were asked to use the other phone first.

Here's the time, in seconds, it took those teens to send the message on each phone:

Own Phone	Other Phone	Difference
37.60	58.80	-21.20
54.87	72.96	-18.09
27.27	37.19	- 9.92
65.58	68.20	- 2.62
38.35	41.02	- 2.67
35.65	42.18	- 6.53
38.60	53.80	-15.20
50.00	61.94	-11.94
21.20	31.80	-10.60
36.55	62.00	-25.45
43.96	91.10	-47.14
39.17	44.82	- 5.65
57.31	56.31	+ 1.00
30.00	61.25	-31.25
42.30	51.11	- 8.81

6) Write out the null and alternate hypotheses in this scenario.

Null: ______________________ Alternate: ______________________

7) Would it be appropriate to conduct an independent samples t-test to compare the mean time for "own phone" against the mean time for "other phone?" Explain.

8) Conduct a dependent-samples t-test and report your p-value. How many degrees-of-freedom did your test have? Briefly write out any conclusions you can make.

9) If you were to construct a 95% confidence interval for the mean difference in texting times, would that interval contain zero? Explain how you know.

10) Conduct a sign test (like we did in activity 23) and report your p-value.

Scenario: Robert Martin turned 55 in 1991. Earlier that year, the Westvaco Corporation, which makes paper products, decided to downsize. They laid off roughly half of the 50 employees in the engineering department, including Martin. Later that year, Martin went to court, claiming that he had been fired because of his age. A major piece of evidence in Martin's case was based on a statistical analysis of the relationship between the ages of the workers and whether they lost their jobs.

Part of the data analysis presented at his trial concerned the ten hourly workers who were at risk of layoff in the second of five rounds of reductions. At the beginning of Round 2, there were ten employees in this group. Their ages were 25, 33, 35, 38, 48, 55, 55, 55, 56, 64. Three were chosen for layoff: the two 55-year-olds (including Martin) and the 64-year old

What to make of these data requires balancing two points of view, one that favors Martin, and the other that favors Westvaco. To understand the two views, imagine a dialog between two people, one representing Martin, the other, Westvaco:

Martin: The pattern in the data is very striking. Of the five people under age 50, all five kept their jobs. Of the five over age 50, only two kept their jobs. The average age of those chosen to lose their jobs is 58 years; that's way above the average for the whole group. The pattern is clear evidence of discrimination

Westvaco: Not so fast! Your sample is way too small to be evidence of anything. There are only ten people in all, and only three in the group that got fired. How can you expect me to take your patterns seriously when just a small change will destroy the pattern? Look how different things would be if we just switch the 64-year old and the 25-year old

Actual data: 25 33 35 38 48 **55** **55** 55 56 **64**
What-if data: **25** 33 35 38 48 **55** **55** 55 56 64

Martin: But look at what you did: You deliberately choose the oldest one fired and switched him with the youngest one not fired. Of all the possible choices, you picked the most extreme. Why not compare what actually happened with all the possible choices

Westvaco: What do you mean?

Martin: Start with the ten workers, and treat them all alike. Let random chance decide which three get chosen for layoff. Repeat the same process over and over, to see what typically happens. Then compare the actual result with what typically happens

11) Write out the null and alternate hypotheses.

Null: ______________________________ Alternate: ______________________________

12) I'd argue that we should **not** conduct an independent samples t-test on this data. Explain why.

13) Even though I shouldn't conduct a t-test, I can have a computer do so. Here's the output from Stata. What conclusions can we make?

```
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
not fired |       7    41.42857    4.455372    11.78781    30.52667    52.33047
   fired |       3          58           3    5.196152    45.09204    70.90796
---------+--------------------------------------------------------------------
combined |      10        46.4    4.033747    12.75583    37.27503    55.52497
---------+--------------------------------------------------------------------
    diff |           -16.57143    7.269122               -33.33405     .191197
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -2.2797
Ho: diff = 0                                     degrees of freedom =        8

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0260         Pr(|T| > |t|) = 0.0521          Pr(T > t) = 0.9740
```

14) Use one of the following websites to conduct a randomization-based test of the means between the fired and not-fired groups. Enter your data, run the simulation, report your estimated p-value, and state your conclusion.

Website 1: http://lock5stat.com/statkey/randomization_1_quant_1_cat/randomization_1_quant_1_cat.html

Website 2: http://www.bradthiessen.com/html5/stats/m300/2grouprandomization.swf

Scenario: With the proliferation of the internet and 24-hour cable news, it has become much easier for people to hear much more information, much more quickly. However, this has led to speculation that news organizations attempt to convey information before it has been properly verified in an effort to feed our impatience. In 2004, the USA Today reported that newspapers appear to be losing credibility over time. They cited a nationwide sample of 1,002 adults interviewed via telephone (under the direction of the Princeton Survey Research Associates) during the period May 6-16, 2002. One of the survey questions asked subjects to rate the credibility of "the daily newspaper you are most familiar with."

Of the 932 respondents to this question, 587 (or 63%) rated their daily newspaper as being "largely believable."

When the same question was asked four years earlier, 618 of 922 (or 67%) rated their daily newspaper as being "largely believable."

We want to test the hypothesis to see if the proportion of Americans rating their newspapers as being believable has actually dropped over the past four years.

15) Write out the null and alternate hypotheses. These should be in terms of proportions.

Null: ______________________ Alternate: ______________________

16) Use your calculator to construct a 95% confidence interval for the difference in proportions. Report that interval.

17) Under the null hypothesis, we can find the weighted average proportion across the two groups:

$$\hat{p}_{pooled} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{932(.63) + 922(.67)}{932 + 922} = 0.65$$

We can then use that weighted average to calculate the standard error for the difference in proportions:

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\hat{p}_{pooled}(1 - \hat{p}_{pooled})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{.65(1 - .65)\left(\frac{1}{932} + \frac{1}{922}\right)} = 0.0222$$

Conduct a t-test using this information and report a p-value. Remember that a t-test, in general, looks like:

$$t_{df} = \frac{\text{observed} - \text{hypothesized}}{\text{standard error}} = \frac{(\text{proportion from 1st group} - \text{proportion from 2nd group}) - (\text{hypothesized difference})}{\text{standard error of the difference in proportions}}$$