

PSET 13 - Due: Wednesday, August 7

1. Let μ_1 = the true average tread life for a premium brand of P205/65R15 radial tire, and let μ_2 = the true average tread life for an economy brand of the same size. Independent random samples of tires of each brand were obtained and yielded the summaries below.

	Sample Size	Sample Mean	Sample Std. Dev.
Premium	$n_1 = 45$	$\bar{x}_1 = 42500$	$s_1 = 2200$
Economy	$n_2 = 45$	$\bar{x}_2 = 36800$	$s_2 = 1500$

- Do the samples give enough evidence to conclude that μ_1 will exceed μ_2 by more than 5000 miles? Test $H_0 : \mu_1 - \mu_2 = 5000$ versus $H_A : \mu_1 - \mu_2 > 5000$ using the level of significance $\alpha = 0.05$.
- By how many miles will μ_1 exceed μ_2 ? Calculate a 90% confidence interval to estimate the true value of $\mu_1 - \mu_2$.
- Do the tread lives of the two brands need to be normally distributed for the test of hypotheses and the confidence interval to be valid? Why or why not?

a.) For a right-tailed z-test with level of significance $\alpha = 0.05$, we have $Z_{0.05} = 1.645$. Thus, assuming the null hypothesis is true, the area of rejection will be $Z > 1.645$. The test statistic is given by

$$\begin{aligned}
 Z &= \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\
 &= \frac{42500 - 36800 - 5000}{\sqrt{\frac{2200^2}{45} + \frac{1500^2}{45}}} \\
 &= 1.7635.
 \end{aligned}$$

Since $Z > 1.645$, we reject the null hypothesis. There is sufficient evidence to suggest μ_1 will exceed μ_2 by more than 5000 miles at the $\alpha = 0.05$ significance level.

b.) For a 90% CI with $Z_{\frac{\alpha}{2}} = 1.645$, we have

$$\begin{aligned}
 &(\bar{x}_1 - \bar{x}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\
 &= 5700 \pm 1.645 \cdot 396.9327 \\
 &= 5047.0457 < \mu_1 - \mu_2 < 6352.9543.
 \end{aligned}$$

c.) No, since both $n_1 > 30$ and $n_2 > 30$, we can invoke the central limit theorem and assert approximate normality for both samples.

2. Frequently, patients must wait a long time for elective surgery. Suppose the wait time for patients needing a knee replacement at two hospitals was investigated. Independent random samples of patients were obtained and the wait time for each (in weeks) was recorded. The resulting summary statistics are given in the following table.

Hospital	Sample Size	Sample Mean	Sample Variance
Hospital 1	15	17.4	34.81
Hospital 2	17	12.1	46.24

- (a) Do the samples give enough evidence to conclude that the mean wait time at hospital 1 is longer than that of hospital 2? Assume that the unknown population variances are equal and test using $\alpha = 0.05$.
- (b) Find bounds on the p value associated with the test.
- (c) How much is the mean wait time for hospital 1 longer than the mean for hospital 2? Calculate a 90% confidence interval to estimate the true value of $\mu_1 - \mu_2$.
- (d) Besides the population variances being equal, what else must be true (or what else must we assume) about the populations for parts (a) and (c) to be valid?

a.) We have

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 > 0.$$

For a right-tailed test with $\alpha = 0.05$, we have $t_{\alpha, m+n-2} = t_{0.05, 30} = 1.697$

Assuming both populations have the same variance, we have a pooled standard deviation given by

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{14 \cdot 34.81 + 16 \cdot 46.24}{15 + 17 - 2}}$$

$$= 6.3958.$$

And a test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{17.4 - 12.1 - 0}{6.3958 \sqrt{\frac{1}{15} + \frac{1}{17}}}$$

$$= 2.3392.$$

Since $t > 1.697$, reject the null hypothesis. There is sufficient evidence to suggest $\mu_1 - \mu_2 > 0$ at the $\alpha = 0.05$ significance level.

b.) We have $t = 0.3657$, the p-value is given by $p(t > 2.34)$. Using the t table, we see

$$0.025 < p < 0.01.$$

c.) For a 90% CI, we have $t_{0.05,30} = 1.697$. Thus

$$\begin{aligned} & (\bar{x}_1 - \bar{x}_2) \pm t_{0.05,30} \cdot S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ &= 5.3 \pm 1.697 \cdot 6.3958 \cdot 0.3542 \\ &= 1.4556 < \mu_1 - \mu_2 < 9.1444. \end{aligned}$$

d.) Since we are given the fact that both samples are random and Independent, we must only assume that both samples come from populations that are normally distributed

3. Many homeowners use tiki torches for outside decoration and to burn special oil to repel insects. Independent random samples of two types of oil were obtained and the burn time for 3 ounces of each was recorded (in hours). The summary statistics are given in the following table. Assume the underlying populations of burn times are normal with unknown, but equal variances.

Oil	Sample Size	Sample Mean	Sample Variance
Citronella Torch Fuel	18	6.25	1.04
Black Flag Mosquito Control	24	5.98	0.77

- (a) Do the samples give enough evidence to conclude that the mean burn time is different for the two brands? Test using $\alpha = 0.01$.
- (b) Find bounds on the p value associated with the test.

a.) We have

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 \neq 0.$$

For a two-tailed t-test with $\alpha = 0.01$, we have $t_{\frac{\alpha}{2}, m+n-2} = t_{0.005, 40} = 2.704$. Thus, we have a rejection region $|t| > 2.704$. With equal but unknown variances, we have a pooled standard deviation of

$$S_p = \sqrt{\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}}$$

$$= \sqrt{\frac{17 \cdot 1.04 + 23 \cdot 0.77}{40}}$$

$$= 0.9406.$$

And a test statistic

$$t = \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

$$= \frac{6.25 - 5.98}{0.9406 \sqrt{\frac{1}{18} + \frac{1}{24}}}$$

$$= 0.9206.$$

Since $|t| < 2.704$, we do not reject the null hypothesis. There is insufficient evidence to suggest $\mu_1 - \mu_2 \neq 0$ at the $\alpha = 0.01$ significance level.

b.) The p-value is given by $2 \cdot p(t > 0.9206)$, using the t table, we find

$$0.25 < 2p < 0.10$$

$$\implies 0.50 < p < 0.2.$$

4. A researcher investigated the lateral range of motion (in degrees) for workers with and without a history of lower back pain (LBP). The summary statistics are given in the following table.

Condition	Sample Size	Sample Mean	Sample Std. Dev.
No LBP	13	91.5	5.5
LBP	17	88.3	7.8

- (a) Calculate a 95% confidence interval to estimate the difference between the true mean extent of lateral range of motion for the populations of people with the two conditions. (Assume the underlying populations are normal with unknown, but equal variances.)
- (b) Suppose that the researcher wanted to test whether or not the means were equal (using $\alpha = 0.05$). She would want to test $H_0 : \mu_1 - \mu_2 = 0$ versus $H_A : \mu_1 - \mu_2 \neq 0$. Using the confidence interval in part (a), what would be the decision (Reject H_0 or Do not reject H_0)? Explain.

a.) For a 95% CI, we have $t_{0.025, 28} = 2.048$. Thus,

$$\begin{aligned}
 & (\bar{x} - \bar{y}) \pm 2.048 \cdot \sqrt{\left(\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2} \right) \cdot \left(\frac{1}{m} + \frac{1}{n} \right)} \\
 & 3.2 \pm 2.048 \cdot \sqrt{\left(\frac{12 \cdot 5.5^2 + 16 \cdot 7.8^2}{28} \right) \cdot \left(\frac{1}{13} + \frac{1}{17} \right)} \\
 & -2.013 < \mu_1 - \mu_2 < 8.413.
 \end{aligned}$$

b.) Using the CI in part a, we can conclude that we should not reject the null hypothesis, there seems to evidence to suggest that $\mu_1 - \mu_2 = 0$, since zero is included in our computed confidence interval.