

Problem 1: Mark the following statements as true or false in the blank space provided. If the statement is false, explain why it is false.

- (a) F: A computer will always give the degrees of freedom for a two-sample t procedure with sample sizes n_1 and n_2 as $\min\{n_1 - 1, n_2 - 1\}$.

Computers can use other degrees of freedom which give non-integer values.

- (b) T: Given a sample size of less than 15 people from a distribution with unknown standard deviation, the population distribution has to be normal or have no outliers/skewness to use one-sample t procedures.

This is the condition for $n \leq 15$. For $15 < n < 40$, we need no skewness. For $n \geq 40$, the procedure does not need either condition.

- (c) T: The standard error for a pooled t -procedure with two equal size samples is the same as the average of the standard deviations of the samples.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(n - 1)s_1^2 + (n - 1)s_2^2}{2n - 2} = \frac{(n - 1)(s_1^2 + s_2^2)}{2(n - 1)} = \frac{s_1^2 + s_2^2}{2}.$$

- (d) F: Given two samples from a distribution with unknown standard deviation, a two-sample t procedure must be used.

There are other options. For example, if the two samples are paired, then matched pairs is appropriate—even necessary as the samples are not independent.

- (e) F: A two-sample t statistic has a t distribution.

We approximate this with a t -distribution, but it is not exactly a t -distribution.

Problem 2: A drinks company has created a new caffeine based energy drink to compete with a rival company whose energy drinks provide a longer ‘energy boost.’ To be sure they have succeeded, the company’s scientists test 16 of their new drinks and find a mean of 11.2 hours of ‘energy boost’ with standard deviation 2.3, and test 13 of their competitor’s beverages, finding an average of 10.8 hours of ‘enhanced energy’ with standard deviation 2.5. Assuming that the ‘enhanced energy period’ for both drinks is normally distributed with approximately equal standard deviations (because they use a similar brewing recipe):

- (a) Find a 90% confidence interval for the difference in ‘extra energy’ time.

Because the population standard deviations are unknown, t -procedures are appropriate. We use a pooled t -test because the two populations have approximately the same standard deviations ($2.3/2.5 = 0.92$ and $0.5 < 0.92 < 2$). Note also that $n_1 + n_2 = 16 + 13 = 29$. We have $\bar{x}_1 = 11.2$, $s_1 = 2.3$, and $n_1 = 16$, and $\bar{x}_2 = 10.8$, $s_2 = 2.5$, and $n_2 = 13$. We have degrees of freedom $n_1 + n_2 - 2 = 29 - 2 = 27$ so that for a 90% confidence interval, $t^ = 1.703$. Then*

$$s_p^2 = \frac{(16 - 1)2.3^2 + (13 - 1)2.5^2}{16 + 13 - 2} = \frac{154.35}{27} = 5.7167 \Rightarrow s_p = \sqrt{5.7167} = 2.39$$

$$\begin{aligned} (\bar{x}_1 - \bar{x}_2) &\pm t^* s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ (11.2 - 10.8) &\pm 1.703(2.39) \sqrt{\frac{1}{16} + \frac{1}{13}} \\ 0.4 &\pm 1.52 \end{aligned}$$

which gives confidence interval $(-1.12, 1.92)$. Therefore, we are 90% certain that, on average, the company’s drink provide between 1.12 hours less to 1.92 hours more of an energy boost.

- (b) Find the t statistic and p -value for hypothesis that these scientists have designed a drink which provides a longer ‘energy boost’ than their competitor’s drink.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{11.2 - 10.8}{2.39 \sqrt{\frac{1}{16} + \frac{1}{13}}} = \frac{0.4}{0.892} = 0.448 \rightsquigarrow p > 0.25$$

Then we have test statistic $t = 0.448$ and p -value $p > 0.25$.

- (c) State your conclusion at the $\alpha = 0.05$ level.

We have

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_a : \mu_1 > \mu_2 \end{cases}$$

Because $p > 0.25 > \alpha$, there is not sufficient evidence to suggest that the company’s energy drink provides a longer energy boost than their competitor’s energy drink.