Solutions to Chapter 9 Problems

1. **Parameter**: μ_d = mean change in weight after an 8-week training program (After - Before)

   Hypotheses: $H_o : \mu_d = 0$ $H_a : \mu_d < 0$

   Conditions: assumed normally distributed

   Statistic:

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-1.04 - 0}{1.197/\sqrt{5}} = -1.94$$

$$\nu = 4$$

   p-value: $p - value = P(T < -1.943) = 0.0621$

   Conclusion: $p - value > \alpha$
   There is not enough evidence to suggest that the training program decreases weight, on average.

   Confidence interval using critical value $t* = 2.776$

   Interval:

$$\bar{d} \pm t^* \frac{s_d}{\sqrt{n}} = -1.04 \pm 2.776(1.197/\sqrt{5}) = -1.04 \pm 1.486 = (-2.526, +0.446)$$

   Interpretation: With 95% confidence, we expect the weight change to be between a 2.5 kg decrease to 0.4 kg increase, on average. Since this interval includes 0, the differences observed are not significant and the training program appears to not be effective.

2. (a) Paired means - since it is the same classroom measured twice (with and without an aide).

   (b) Independent samples - the two samples are from different regions and have no reason to believe they are dependent.

   (c) One Sample - we are comparing our data to the fixed value of 45 minutes.

3. (a)

4. (a) **Parameter**: $\mu_1 - \mu_2$ = difference between the true mean dietary fibre of vegetarians with (1) and without (2) diverticular disease.

   Hypotheses: $H_o : \mu_1 - \mu_2 = 0$ $H_a : \mu_1 - \mu_2 < 0$

   Conditions: normally distributed populations, independent

Statistic:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{27.2 - 42.7 - 0}{\sqrt{\frac{9.5^2}{5} + \frac{9.9^2}{18}}} = -3.1977$$

$$\nu = \min(5 - 1.18 - 1) = 4$$

p-value: p-value = 1- 0.9835 = 0.0165

```
> pt(abs(statistic),df)
[1] 0.9835135
```

Conclusion: $p - value < \alpha = 0.05$. There is significant evidence to suggest that the true mean dietary fibre content is higher for those without diverticular disease.

(b) **Interval:** use $df = 4$

$$\bar{X}_1 - \bar{X}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 27.2 - 42.7 \pm 2.7764\sqrt{\frac{9.5^2}{5} + \frac{9.9^2}{18}} = -15.5 \pm 13.46 = (-28.96, -2.04$$

Interpretation: We are 95% confident that the true mean dietary fibre content for vegetarians without diverticular disease is between 2.04 and 28.96 higher than the mean of those with the disease.

5. (a) **Parameter**: μ_d = true mean difference in assembly time for written manual and interactive video training.
Hypotheses: $H_o : \mu_d = 0$ $H_a : \mu_d \neq 0$
Conditions: check normal probability plots using software to confirm normality.
Statistic: t = 1.7154 $\nu = 7$ **p-value**: 0.13 **Conclusion**:
$p - value > \alpha = 0.05$
There is not enough evidence to reject the null hypothesis. It is reasonable to assume that there is no difference, on average, in assembly time for those trained with the written manual or the interactive video.

(b) **Interval:** (-0.222, 1.397)
Interpretation: We are 95% confident that the true mean difference in assembly time for the two training types is between -0.22 and 1.40 minutes. Since this interval includes zero, there is not a significant difference, on average.

6. (a) **Parameter**: $\mu_1 - \mu_2$ = difference between the true mean magnesium content of beans (1) and potatoes (2)
Hypotheses: $H_o : \mu_1 - \mu_2 = 0$ $H_a : \mu_1 - \mu_2 \neq 0$
Conditions: normally distributed populations, independent, both σ unknown (use t procedure).
Statistic: t = -2.9787, $\nu = 35.76$ **p-value**: 0.005174
Conclusion: $p - value < \alpha = 0.05$. There is significant evidence to suggest that the true mean magnesium content differs for beans and potatoes

(b) **Interval:** (-2.923,-0.1324)

Interpretation: We are 99% confident that the true mean magnesium content of beans is between 0.14 and 2.92 mg lower than the true mean content in potatoes.

7. **Parameter**: $p_1 - p_2$ = difference between the true proportion of employed men (1) and women (2)

Hypotheses: $H_o : p_1 - p_2 = 0$ $H_a : p_1 - p_2 \neq 0$

Conditions: First we need the pooled sample proportion $\hat{p}_{pool} = \dfrac{X_1 + x_2}{n_1 + n_2} = \dfrac{718 + 631}{797 + 732} = \dfrac{1339}{1529} = 0.8757$

$n_1 \hat{p}_{pool} = 697.9$, $n_1(1 - \hat{p}_{pool}) = 99.1$, $n_2 \hat{p}_{pool} = 641$, $n_2(1 - \hat{p}_{pool}) = 91$ are all more than 5.

Statistic:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \delta_0}{\sqrt{\hat{p}_{pool}(1 - \hat{p}_c)(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(\frac{718}{797} - \frac{631}{732}) - 0}{\sqrt{\frac{1339}{1529}(1 - \frac{1339}{1529})(\frac{1}{797} + \frac{1}{732})}} = 2.301$$

p-value: $2 \cdot P(Z > 2.301) = 2 \times (1 - 0.9893) = 0.0214$

Using empirical rule: Z between 2 and 3 means one-sided p-value between 0.0015 and 0.025, so two sided p-value between 0.003 and 0.05

```
> pnorm(abs(statistic))
[1] 0.9893042
```

Conclusion: $p - value < \alpha = 0.01$. There is significant evidence to suggest that there is a difference between the true proportions of male and female students that get summer jobs.