Supplemental Examples (Chapter 11)

Online S520

1. In a genetic inheritance study discussed by Margolin (1988), blood was collected from samples of individuals from several ethnic groups, and mean sister chromatid exchange (MSCE) was measured for each individual. We wish to use a t-test to compare the average MSCE for the groups labeled "Native American" and "Caucasian" :

   - Native American: 8.50, 9.48, 8.65, 8.16, 8.83, 7.76, 8.63
   - Caucasian: 8.27, 8.20, 8.25, 8.14, 9.00, 8.10, 7.20, 8.32, 7.70

   The Native American individuals had mean MSCE 8.57 with standard deviation 0.54. The Caucasian individuals had mean MSCE 8.13 with standard deviation 0.49.

   (a) What is the experimental unit? What measurements are taken on the experimental units? Is this a problem with one or two independent samples?

   (b) Carefully state the assumptions of Welch's two-sample t-test. (The single word "normal" is not sufficient.)

   (c) Give null and alternative hypotheses for an appropriate t-test, and calculate the t-statistic.

   (d) What is the significance probability? If we adopt a significance level of $\alpha = 0.1$, should we reject the null hypothesis?

   Solutions:

   (a) Experimental unit: Individual (Native American or Caucasian). Measurement: MSCE. Two independent samples.

   (b) Two independent samples from normally distributed populations with unknown variances.

   (c) Let Δ be the population mean MSCE in Native Americans minus the population mean MSCE in Caucasians. $H_0 : \Delta = 0, H_1 : \Delta \neq 0$.

   $$T_w = \frac{\hat{\Delta} - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{8.57 - 8.13}{\sqrt{\frac{0.54^2}{7} + \frac{0.49^2}{9}}} = 1.68.$$

   (It's 1.70 if you use the raw data.)

   (d)
   ```
   Delta.hat = mean(NativeAmerican) - mean(Caucasian)
   std.error = sqrt(var(NativeAmerican)/7 + var(Caucasian)/9)
   Tw = Delta.hat / std.error
   df = (var(NativeAmerican)/7+var(Caucasian)/9)^2 / ((var(NativeAmerican)/7)^2/6 +
       (var(Caucasian)/9)^2/8)
   2*(1 - pt(Tw, df=df))
   # or just do
   # t.test(NativeAmerican, Caucasian)
   ```
   We get a p-value 0.114, which is greater than α. So we don't have sufficient evidence to reject the null hypothesis.

2. The heights of men and women are approximately normally distributed. Suppose we wish to estimate the average difference in heights between men and women attending a certain university. A random sample of seven men has average height 68.5 inches with standard deviation 3.0 inches, while a random sample of seven women has average height 65.5 inches with standard deviation 2.5 inches.

   (a) Explain why we would use the t-distribution and not the standard normal distribution to calculate P-values and confidence intervals in this case.

   (b) According to Welch's approximation, the number of degrees of freedom is 11.6. Find an approximate 95% confidence interval for the average difference in heights between men and women at the university.

   (c) The confidence interval you calculated in part (b) contains the value 0. Should you conclude that there is no difference between the average height of men and the average height of women at the university? Explain why or why not.

   Solutions:

   (a) We have small samples, so the CLT doesn't help us. Since the samples are small, we need to account for the extra variation that occurs because we estimate the standard deviation using s. This is what the t-distribution is for.

   (b) The R code `qt(.975,df=11.62)` gives the value 2.187, we go up and down 2.187 standard errors from the observed difference in means.

$$\hat{\Delta} \pm q_t \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = (68.5 - 65.5) \pm 2.187 \times \sqrt{\frac{3^2}{7} + \frac{2.5^2}{7}}$$

   or -0.2 inches to 6.2 inches.

   (c) No. Based on the data alone, it's plausible that men and women at the university are the same height on average, but it's also plausible that men are on average six inches taller. Based on everything you have ever seen in real life, it's not plausible that men and women at the university average the same height.

3. A randomized experiment is carried out to test two diets for patients with diverticulosis. Diet A is given to 15 patients, while Diet B is given to 12 patients. The outcome studied is the transit time through the alimentary canal. The patients on Diet A had a mean transit time of 68.4 hours, with standard deviation 16.47 hours. The patients on Diet B had a mean transit time of 83.42 hours, with standard deviation 17.63 hours. Normal probability plots (qqnorm) of each sample give approximately straight lines.

   (a) Which type of test would you use to test for a difference between the two diets? Explain your choice.

   (b) Calculate the test statistic for the test you chose in part (a), and give R code to produce the P-value (significance probability) for a two-tailed test.

   (c) Calculate an approximate 95% confidence interval for the difference in mean transit time between Treatment A and Treatment B.

   Solutions:

(a) Welch's two-sample t-test is a safe choice. The samples are plausibly from normal populations with unknown variances, and there's no a priori reason to be sure that the variances are equal. (Student's two-sample t-test is also justifiable here.)

(b) The observed average difference (Diet B minus Diet A) is 15.02 hours, with a standard error of 6.63 hours. The t-statistic is 2.265. The degrees of freedom calculation comes out to be 22.94. The P-value is `2*(1-pt(2.265, df=22.94))` (which is about 0.03).

(c) The R command `qt(0.975, df)`, where df is the correct number of degrees of freedom, gives the output 2.07.

$$\hat{\Delta} \pm q_t \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = 15.02 \pm 2.07 \times 6.63$$

or about 1.3 hours to 28.7 hours longer for Diet B.