Practice Problems for Chapter 9

1. Weights (in kg) of 5 subjects were measured before and after an 8-week training program. At $\alpha = 0.05$ is there significant evidence that the program is helpful for losing weight, on average? (You may assume the data are normally distributed and choose the p-value from the list below). Follow up the test by calculating and interpreting a 95% confidence interval (selecting the critical value from the list below)

Subject	1	2	3	4	5	Mean	SD
Weight Before	55.7	54.9	59.6	62.3	74.2	61.34	7.788
Weight After	52.7	54.4	59.8	61.5	73.1	60.3	8.033
Difference (A-B)	-3.0	-0.5	0.2	-0.8	-1.1	-1.04	1.197

P-values

```
> pnorm(abs(statistic))
[1] 0.9738102
> pt(abs(statistic),df)
[1] 0.9378113
```

Critical Values

```
> qt(0.975,4)
[1] 2.776445
> qt(0.95,4)
[1] 2.131847
> qt(0.975,9)
[1] 2.262157
> qt(0.95,5)
[1] 2.015048
```

2. In each experiment, determine whether the data are obtained single sample, two independent samples, or are paired. If the data are two independent samples, indicate the two distinct populations. If the data are paired, indicate the common characteristic.

   (a) School board members believe adding a teacher's aide to each K-4 class will improve classroom management and increase instruction time. Twenty-size elementary classrooms are selected at random and the daily instruction time is recorded. A teacher's aide is then added to each classroom, and the daily instruction time is recorded again. The data will be used to determine whether there is any evidence that adding a teacher's aide increases the mean daily instruction time.

   (b) A researcher investigating home-insurance costs obtained a random sample of homes in the Northeast and another random sample of homes in the South. The yearly insurance cost for each home was recorded. The data will be used toe determine whether there is any difference in the mean yearly home-insurance costs between the Northeast and the South.

   (c) Officials at the transit authority of a large city would like to estimate the mean route time during evening rush hour. Eleven routes are selected at random and the evening route time is recorded for each. The data will be used to determine if the mean route time is longer than 45 minutes.

3. If the process of selecting individuals or objects in sample 1 has no effect on, or no relation to, the selection of individuals or objects in sample 2, the samples are:

   (a) independent   (b) dependent   (c) discrete   (d) paired

4. In a study of diverticular disease and diet among vegetarians, 23 vegetarians were surveyed. One variable of interest was total dietary fibre and it is thought that vegetarians without the disease have mean total daily dietary fibre content in their diet higher than those with the disease. The following information was obtained fort he two groups.

With	$n_1 = 5$	$\bar{X}_1 = 27.2$	$s_1 = 9.5$
Without	$n_2 = 18$	$\bar{X}_2 = 42.7$	$s_2 = 9.9$

   (a) At $\alpha = 0.05$ is there enough evidence to conclude that the mean total daily dietary fibre content is higher for those without the disease? You may assume the populations are normally distributed (Be sure to state your hypotheses, check assumptions, and perform the appropriate test)

   (b) Construct and interpret the appropriate 95% confidence interval

Critical Values

P-values

```
> pnorm(abs(statistic))
[1] 0.9993074
> pt(abs(statistic),df)
[1] 0.9835135
```

```
> qnorm(0.975,0,1)
[1] 1.959964
> qt(0.025,5)
[1] -2.570582
> qt(0.975,4)
[1] 2.776445
> qt(0.975,17)
[1] 2.109816
> qchisq(0.95,17)
[1] 27.58711
```

5. Each employee hired at an electronic parts assembly line in Edmonton, is given a general intelligence test. To determine which method of training is more effective, eight pairs of new hires were matched according to their exam scores. One set of employees was asked to read appropriate training manuals, while the other group watched interactive training videos. Each employee was then asked to assemble a part used in a locater-beacon transmitter, and the time (in minutes) to completion was recorded. Two options for R output are given below. Select the correct output and use it to perform the hypothesis test and calculate the confidence interval

Employee Pair	1	2	3	4	5	6	7	8
Written manual	4.9	4.6	5.3	4.9	4.9	5.4	5.5	5.0
Interactive Video	3.1	4.1	4.4	4.9	3.6	3.9	6.5	5.3

```
> written = c(4.9,4.6,5.3,4.9,4.9,5.4,5.5,5.0)
> video = c(3.1,4.1,4.4,4.9,3.6,3.9,6.5,5.3)
> t.test(written,video,mu=0,alternative="two.sided",conf.level=0.95)
Welch Two Sample t-test

data:  written and video
t = 1.486, df = 8.1318, p-value = 0.175
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.3216579  1.4966579
sample estimates:
mean of x mean of y
5.0625    4.4750
> difference = written - video
> t.test(difference,mu=0,alternative="two.sided",conf.level=0.95)
One Sample t-test

data:  difference
t = 1.7154, df = 7, p-value = 0.13
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-0.2223591  1.3973591
sample estimates:
mean of x
0.5875
```

(a) Is there enough evidence at $\alpha = 0.05$ to suggest that there is a difference in completion time, on average, between the two training types? (Be sure to state your hypotheses, check assumptions, and perform the appropriate test)

(b) Identify and interpret the confidence interval

6. Magnesium is used by every cell in your body, required for over 300 biochemical reactions, and helps muscles and nerves to function properly. According to the US Department of Agriculture National Nutritional Database, $\frac{1}{2}$ cup of vegetarian baked beans and one medium baked potato without the skin contain the same amount of magnesium. To check this claim, independent random samples of baked beans and potatoes were obtained and the amount of magnesium in each serving was recorded (in milligrams).

```
> t.test(beans,potatoes,mu=0,alternative="two.sided",conf.level=0.99)

Welch Two Sample t-test

data:  beans and potatoes
t = -2.9787, df = 35.76, p-value = 0.005174
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
-2.9228274 -0.1324105
sample estimates:
mean of x mean of y
38.67803  40.20565
```

(a) Is there enough evidence at $\alpha = 0.05$ to conclude that the mean magnesium content is different for the two food items? You may assume the populations are normally distributed.

(b) Identify and interpret the confidence interval

7. A university financial aid office polled a SRS of undergrads to study their summer employment. The data is shown below. Is there evidence that the proportion of male students who had summer jobs differs from the proportion of female students who had summer jobs?

	Men	Women
Employed	718	631
Not Employed	79	101
Total	797	732

```
> pnorm(abs(statistic))
[1] 0.9893042
> pt(abs(statistic),df)
[1] 0.9891639
```