One Sample t-tests

In this exercise, we will learn how to perform a one sample t-test using software. We will use the dataset `LDLdiet.xls` posted on Blackboard. This dataset describes a crossover study that was conducted to investigate whether oat bran cereal helps to lower serum cholesterol levels in $n = 14$ hypercholesterolemic males.

Each subject was randomly placed on a diet that included either oat bran or cornflakes. After two weeks, LDL was recorded, and then each man was switched to the alternative diet. After another two weeks, LDL was recorded again for comparison.

The SAS code in this exercise can be found in the file Lab7.sas, which is posted on Blackboard.

1 SAS

1.1 Read In The Data

As usual, our first step is to read the data into SAS. We do this in the data step below using the `CARDS` statement.

```
DATA Ldl ;
  INPUT Subject CornFlakes OatBran ;
  Difference = CornFlakes - OatBran ;
  CARDS ;
   1   4.61  3.84
   2   6.42  5.57
   3   5.40  5.85
   4   4.54  4.80
   5   3.98  3.68
   6   3.82  2.96
   7   5.01  4.41
   8   4.34  3.72
   9   3.80  3.49
  10   4.56  3.84
  11   5.35  5.26
  12   3.89  3.73
  13   2.25  1.84
  14   4.24  3.14
;
RUN ;
```

1.2 Descriptive Statistics

Our first step is to explore the data using descriptive statistics on the variables `cornflakes`, `oatbran` and `difference`. We use the `MEANS` procedure for this.

```
PROC MEANS DATA=ldl N MEAN STD MIN MAX MAXDEC=2;
  VAR cornflakes oatbran difference;
RUN ;
```

1.3 One Sample T-test

1.3.1 Null Value: 0

Suppose we would like to perform a 2-sided t-test on the variable `cornflakes` with $\mu_0 = 0$ at significance level $\alpha = 0.05$. That is, we would like to test whether the LDL of men on the corn flakes diet is significantly different from 0, or

$$H_0 : \mu = 0$$

$$H_A : \mu \neq 0$$

To carry out a t-test, we use the `TTEST` procedure. We implement the above hypothesis test in SAS by submitting the following three lines of code.

```
PROC TTEST DATA=ldl ;
  VAR cornflakes ;
RUN ;
```

Assuming everything runs correctly, SAS will perform the desired hypothesis test and generate the results. The output is shown below.

```
                    The TTEST Procedure

                    Variable:  CornFlakes

 N        Mean     Std Dev     Std Err     Minimum     Maximum
14      4.4436      0.9688      0.2589      2.2500      6.4200


   Mean       95% CL Mean         Std Dev      95% CL Std Dev
 4.4436     3.8842   5.0030       0.9688      0.7024   1.5608


                   DF    t Value    Pr > |t|
                   13      17.16      <.0001
```

SAS output often includes lots of things you don't need. The most important things that you'll need from this output are

1. First two tables
   - Descriptive statistics on variable `cornflakes` (N, Mean, Std Deviation & Error).
   - Two-sided 95% CI for μ (95% CL Mean).
2. Third table
   - Degrees of freedom (DF).
   - Test statistic t.
   - Corresponding p-value associated with t (`Pr > |t|`).

1.3.2 Null Value: Not 0

The Mayo Clinic recommends that the LDL levels of an average person should be below 5.2 mmol/L. In §1.3.1, we performed a t-test with $\mu_0 = 0$, which doesn't make much sense for this data. Suppose instead, we would like to test whether the men who ate cornflakes had LDL levels different than the level recommended by the Mayo Clinic. In that case, we would test

$$H_0 : \mu = 5.2$$
$$H_A : \mu \neq 5.2$$

To carry out this hypothesis test in SAS, we add only one modification to the code from §1.3.1. In the `PROC` statement, we add the SAS keyword `H0=`, along with whatever we have chosen to be our value of μ_0. For this example, we would submit the following code.

```
PROC TTEST DATA=ldl H0=5.2;
  VAR cornflakes ;
RUN ;
```

1.3.3 Varying Significance Level

Both of the t-tests above were conducted at the significance level $\alpha = 0.05$, which is the default level. To change this, we add the SAS Keyword option `alpha=` to the `PROC` statement. For instance, to conduct the previous test at the $\alpha = 0.01$ significance level, we would use the following code.

```
PROC TTEST DATA=ldl H0=5.2 alpha=0.01;
  VAR cornflakes ;
RUN ;
```

1.4 Paired T-test

Since this is a crossover study, we are interested in how a man's LDL changes between the two diets, Corn Flakes and Oat Bran. Therefore, we would like to perform a paired t-test on the set of $n = 14$ differences at the $\alpha = 0.05$ significance level. That is, we would like to test

$$H_0 : \mu_\Delta = 0$$
$$H_A : \mu_\Delta \neq 0$$

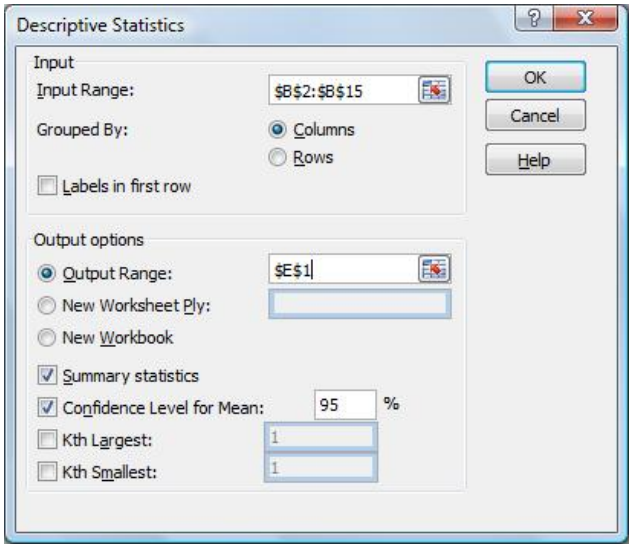
To carry out this test, we simply replace the `VAR` statement in our earlier SAS code with a `PAIRED` statement. The following code performs the paired t-test we want on the set of LDL differences between the diets.

```
PROC TTEST DATA=ldl ;
  PAIRED cornflakes*oatbran ;
RUN ;
```

2 Excel

2.1 Descriptive Statistics

We use the Descriptive Statistics tool from the Analysis ToolPak in order to compute summary statistics on the CornFlakes variable. Below is a picture of the parameters that we will use with this tool. In order to follow the rest of the exercise instructions exactly, please use these parameters.



2.2 One Sample T Test

Unfortunately, there is no tool for the one sample t-test in the Analysis ToolPak. However, we can use built-in Excel functions in order to come up with the test statistic and corresponding p-value.

2.2.1 Null Value: 0

As before, we would like to test $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$ at the $\alpha = 0.05$ significance level. To perform this test, we must calculate the test statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{\bar{x} - 0}{s/\sqrt{n}}$$

We plug this **test statistic** formula into Excel manually.

1. Select the desired cell J4
2. Click in the Formula Bar (above spreadsheet) to activate it
3. Type "=(F3-0)/(F7/SQRT(F15))"
4. Press Enter

Finally, we calculate the ***p*-value** manually using the TDIST function.

1. Select the desired cell J5
2. Click in the Formula Bar (above spreadsheet) to activate it
3. Type "=TDIST(J4,F15-1,2)" (Format "=TDIST(positive test stat, DF, 2-tails)")
4. Press Enter

2.2.2 Null Value: Not 0

Now we would like to test $H_0 : \mu = 5.2$ vs $H_A : \mu \neq 5.2$ at the $\alpha = 0.05$ significance level. To perform this test, we must calculate the test statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{\bar{x} - 5.2}{s/\sqrt{n}}$$

We plug this **test statistic** formula into Excel manually.

1. Select the desired cell K4
2. Click in the Formula Bar (above spreadsheet) to activate it

3. Type "=(F3-5.2)/(F7/SQRT(F15))"
4. Press Enter

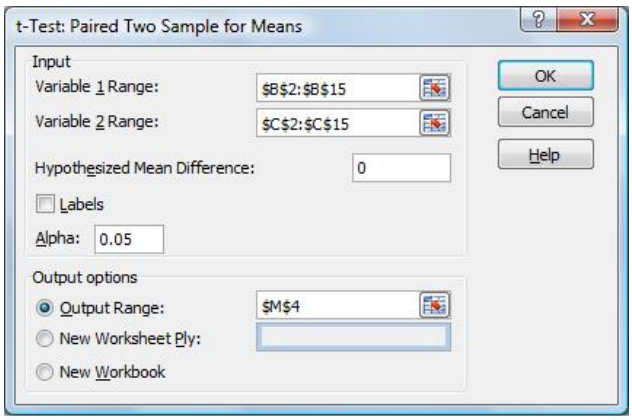
Finally, we calculate the ***p*-value** manually using the TDIST function.

1. Select the desired cell K5
2. Click in the Formula Bar (above spreadsheet) to activate it
3. Type "=TDIST(-K4,F15-1,2)" (Format "=TDIST(positive test stat, DF, 2-tails)")
4. Press Enter

2.3 Paired T-test

We can use the Paired t-test tool from the Analysis ToolPak in order to conduct the desired paired t-test. These are the steps we shall use to carry out the paired t-test on the set of LDL differences. Below is a picture of the parameters chosen.

1. Open the Data Analysis dialog box from the Data ribbon.
2. Highlight "t-Test: Paired Two Sample for Means" and press OK.
3. Set CornFlake data as Variable 1, and OatBran data as Variable 2.
4. Set Hypothesized Mean to be 0 (or whatever value you want).
5. Set Alpha to 0.05 (or whatever value you want).
6. Decide where you want the output to print.



3 Exercises

1. Perform a 1-sample t-test on the variable CornFlakes at $\mu_0 = 4$ with $\alpha = 0.01$.

   $t =$ $p =$

2. Perform a 1-sample t-test on the variable OatBran at $\mu_0 = 0$ with $\alpha = 0.01$.

   $t =$ $p =$

3. Perform a 1-sample t-test on the variable OatBran at $\mu_0 = 5.2$ with $\alpha = 0.05$.

   $t =$ $p =$

4. Perform a paired t-test between variables Air1 & Air2 with $\alpha = 0.05$ in dataset CAD.xls.

   $t =$ $p =$

5. Perform a paired t-test between variables 2co1 & 2co2 with $\alpha = 0.10$ in dataset CAD.xls.

   $t =$ $p =$

6. Perform a paired t-test between variables 4co1 & 4co2 with $\alpha = 0.01$ in dataset CAD.xls.

   $t =$ $p =$