Name: _______________________________________

1 Objective: T-tests; testing the differences between two means; paired t-tests (related samples)

The t-test assesses whether the means of two groups are statistically different from each other. The t-test is known to be relatively robust if the underlying distribution of the samples is normal. To determine whether the means of two groups are significantly different, we set up a null hypothesis (H_0). It states that the both groups would have the same mean if we repeated the experiment a large (infinite) number of times. The alternative hypotheses (H_1) to the null hypothesis are that one particular mean will be greater (or lower) than the other (called a one tailed test) or that the two means will be different (called a two-tailed test). Figure 1 illustrates the hypothesis teasting discussed above. Read CS chapter 6.

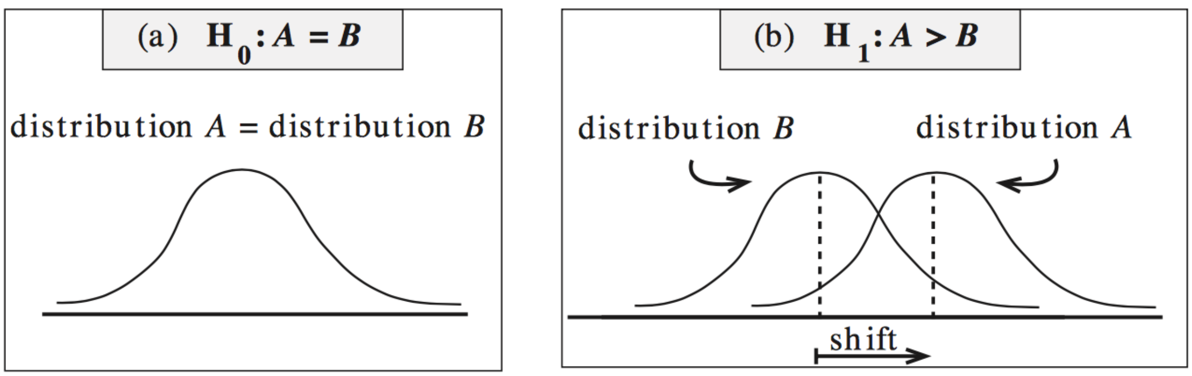

Figure 1: Testing the differences between two means

Assumptions:

Two sample t-test assumes that

1. There is one continuous dependent variable and one categorical independent variable (with 2 levels);
2. The two samples are independent;
3. The two samples follow normal distributions (check normality by QQ-Plot).

When the assumptions are not met, other methods are possible based on the two samples:

1. Two dependent samples and follow Normal distribution, use paired t-test.
2. Two independent samples and do not follow Normal distribution, use Wilcoxon Rank Sum test (or Wilcoxon Mann-Whiney U-test=WMW) test.

3. Two dependent samples and do not follow Normal distribution, use Signed Rank test (or Wilcoxon signed-rank test)

Note that SAS performs a two-sided test, meaning the hypothesis is to compare a significant difference between two groups. If one wants to test whether one group is greater(smaller) than the other, p-value can be divided by 2.

2 Key Words

`proc ttest`, independent samples, paired samples.

3 Randomly Assigned Groups

Following data steps generate headache data which contains time for subjects to feel relief from headache pain using Aspirin `"A"` and Tylenol `"T"`.

```
DATA headache;
   INPUT treat $ time @@;
DATALINES;
A 40 A 42 A 48 A 35 A 62 A 35
T 35 T 37 T 42 T 22 T 38 T 29
;
run;
```

Print data and confirm that they look ok.

Perform a t-test below. Is either product significantly faster than the other at $\alpha = 0.05$ significance level?

```
PROC ttest data=headache;
   TITLE "Headache Study";
   CLASS treat;
   VAR time;
run;
```

Use the output section of "equality of variances" to decide between the methods "Pooled" and "Satterthwaite" for your hypothesis testing:

1. When the p-value (shown under "Pr>F") is greater than 0.05, then the variances maybe equal then read the "Pooled" section of the result.

2. When the p-value (shown under "Pr>F") is no more than 0.05, then the variances are unequal then read the "Satterthwaite" section of the result

Create a box plot to visualize the means.

```
PROC boxplot data=headache;
      PLOT time*treat;
run;
```

Nonparametric tests are often more appropriate for very small samples, since assumptions concerning distributions are difficult to determine. Among many nonparametric tests, we will use Wilcoxon rank-sum test (equivalent to the Mann-Whitney U-test) for two samples. For moderate sample sizes, the Wilcoxon test is almost as powerful as t-test (parametric equivalent).

Assuming the absence of a normal distribution and having very small sample size, we can use Wilcoxon test to the `headache` data. We use this if:

1. Comparing two samples.
2. The two groups of data are independent.
3. The type of variable could be continuous or ordinal.
4. The data might not be normally distributed.

```
PROC npar1way data=headache WILCOXON;
   TITLE "Nonparametric Comparison";
   CLASS treat;
   VAR time;
   EXACT WILCOXON; /* use EXACT for exact p-value (not the asymptotic approx. usually computed) when
       sample size is small */
run;
```

Explain the output in terms of time for subjects to feel relief from headache pain using `"A"` and `"T"`.

4 Paired t-tests - Related Samples

Groups can be considered to be independent if they are both randomly assigned. Regular t-test can not be used if the groups are no longer independent. For example, there are many experiments where each subject receives a pair of data (pre and post) and therefore the groups (pre and post scores) are no longer independent. In this case, we use what's called "paired t-test". The paired t-test computes a mean and standard error of the differences and determines the probability that the absolute value of the mean difference was greater than zero by chance alone (Read CS146-148).

In this section, we will consider two types of headache medication again (CS ch.6). Each of eight subject tries each of the two drugs (for two different headaches) and the time span to pain relief is measured.

Run the following code to generate the data.

```
DATA headachepair;
     INPUT subj Atime Btime;
   DATALINES;
   1 20 18
   2 40 36
   3 30 30
   4 45 46
   5 19 15
   6 27 22
   7 32 29
   8 26 25;
run;
```

Justify why you want to use paired t-test for this data. Run the following code to compare two drugs and decide which one is acting faster.

```
PROC ttest data=headachepair;
     TITLE "Paired T-Test on headache relief time";
     PAIRED Atime * Btime;
run;
```

Assume two dependent samples and they do not follow Normal distribution, we use Signed Rank test (or Wilcoxon signed-rank test). Univariate analysis is done on the differences:

```
DATA headpairW;
       SET  headachepair;
       diff= Atime - Btime;
run;

PROC univariate data=headpairW;
         var diff;
run;
```

Create a box plot to visualize the means.

5 Turn In

Answer the following questions and report:

1. `headachepair` data dependent or independent? Why?
2. Explain the difference between paired t-test and Wilcoxon signed-rank test.
3. Choose one of the methods above and run the test.
4. Report the results with supporting charts and graphs.