Chapter 11
Comparing Two Populations or Treatments

11.0 Introduction

Our discussion of the confidence intervals and hypothesis tests for two treatments will seem abbreviated after our earlier discussions. Actually, our discussion will in fact be abbreviated because of our earlier discussions. As we have mentioned previously one of the strengths of the TI-83 is that the learning builds as you go. The screens, the requested information, and the output will all be familiar after performing the procedures for one population.

11.1 Inferences Concerning the Difference between two means

Example 11.2: Oral Contraceptive and Bone Mineral Density – a hypothesis test

To assess the impact of oral contraceptive use on bone mineral density (BMD), researchers in Canada carried out a study comparing BMD for women who had used oral contraceptives for at least three months to BMD for women who had never used oral Data consistent with summary quantities given in the paper appear in the accompanying table.

	Bone mineral density (g/cm)									
Never used oral contra-captives	0.82	0.94	0.96	1.31	0.94	1.21	1.26	1.09	1.13	1.14
Used oral Contra-ceptives	0.94	1.09	0.97	0.98	1.14	0.85	1.30	0.89	0.87	1.01

Enter these data in your calculator – we once again are using List1 and List2. An obvious point, but one we should mention is that you need to keep it straight which data is in which list. It will become an issue if your alternative hypothesis is one-tailed.

After entering the data, you will need to check once again the plausibility of the assumptions of normal populations. This time you will need to check both samples. After you have satisfied yourself of this plausibility we'll proceed to our hypothesis testing. Here are the keystrokes:

STAT > `TESTS` > `2-SampTTest…` >

And we're off and running!

```
2-SampTTest
 Input:Data Stats
 List1:L1
 List2:L2
 Freq1:1
 Freq2:1
 μ1:≠μ2  <μ2  >μ2
 Pooled: No Yes
 Calculate Draw
```

The requested information for the two-sample hypothesis test is the requested information for the one-sample hypothesis test, except for the "pooling." As explained in the text, pooling is a procedure that is not used for the two-sample *t* procedures, because of its unstable behavior if the population variances are unequal. The appropriate choice is to "Just Say No" to pooling in the above screen. Arrow down to **`Calculate`** and press **ENTER**. The calculator takes a short bit of time – must be that horrible degrees-of-freedom calculation that slows it down – and we get the following information:

```
2-SampTTest
μ1>μ2
t=1.134848135
p=.1358017443
df=17.66529895
x̄1=1.08
x̄2=1.004
Sx1=.159721981
Sx2=.139060339
n1=10
n2=10
```

The reading of this screen is very similar to the reading of the one-sample hypothesis testing screen – notice that the descriptive information is duplicated for both samples, and the degrees of freedom have been calculated for you.

Example 11.4: Effect of Talking on Blood Pressure – a Confidence Interval

Does talking elevate blood pressure, contributing to the tendency for blood pressure to be higher when measured in a doctor's office than when measured in a less stressful environment (called the "white coat" effect)? A study in which patients with high blood pressure were randomly assigned to one of two groups produced the following data. Those in the first group (the talking group) were asked questions about their medical history and about the sources of stress in their lives in the minutes prior to measuring blood pressure. Those in the second group (the counting group) were asked to count aloud from 1 to 100 four times prior to having blood pressure measured. The data values for diastolic blood pressure (mmHg) are as follows:

Talking	104	110	107	112	108	103	108	118
Counting	110	96	103	98	100	109	97	105

Once again enter the data into the calculator, and be sure to keep it straight which treatment is in which TI List! We will use List3 and List4 this time. When the data is entered, the keystrokes for a 2-sample confidence interval for means are decently predictable:

STAT > **`TESTS`** > **`2-SampTInt`**… >

The Lists and choices presented will typically reflect what you chose the last time you constructed a confidence interval, so fill in the information with the appropriate Lists, choose a 95% confidence level, and of course Just Say No to pooling. Arrow down to **`Calculate`** and press **ENTER.**

```
2-SampTInt
 Inpt:Data Stats
 List1:L3
 List2:L4
 Freq1:1
 Freq2:1
 C-Level: .95
 Pooled:No Yes
 Calculate
```

After pressing **ENTER** above, your calculations are soon at hand:

```
2-SampTInt
 (1.0461,11.954)
 df=13.77631629
```
$\bar{x}_1 = 108.75$

$\bar{x}_2 = 102.25$

$Sx_1 = 4.74341649$

$Sx_2 = 5.3917927$

$n_1=8$

$n_2=8$

Once again – we are in agreement with POD. We are getting to be great with this calculator.

11.3 Inferences Concerning the Paired Difference between means

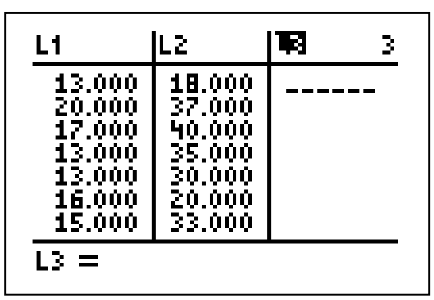
Example 11.8: Lactic Acid in the Blood After Exercise

The effect of exercise on the amount of lactic acid in the blood was examined in the article “A Descriptive Analysis of Elite-Level Racquetball” (*Research Quarterly for Exercise and Sport* (1991): 109–114). Eight males were selected at random from those attending a weeklong training camp. Blood lactate levels were measured before and after playing three games of racquetball, as shown in the accompanying table. We will use this data to estimate the mean change in blood lactate level using a 95% confidence interval.

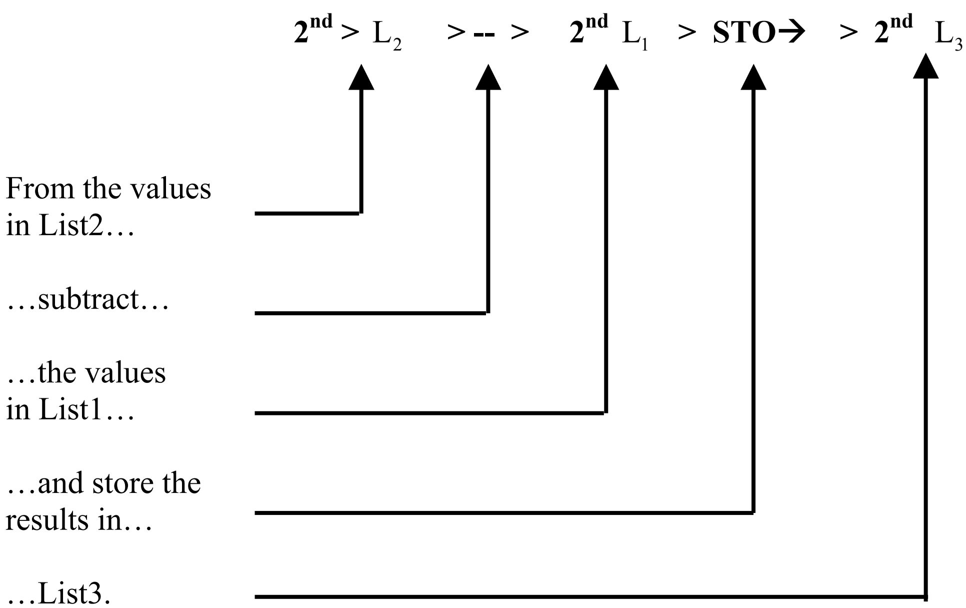
Player	Before	After	Difference
1	13	18	-5
2	20	37	-17
3	17	40	-23
4	13	35	-22
5	13	30	-17
6	16	20	-4
7	15	33	-18
8	16	19	-3

The analysis of paired data is, as you know, an analysis performed on the differences between the values of the variables for each unit of analysis. In this case, the differences between the levels of blood lactate before and after the three games of racquetball. To analyze these data, enter the Before scores in List1, and the After scores in List2. You may, of course, use whichever Lists you wish, but if you use different lists be sure to follow the discussion carefully, since it makes a difference for the analysis which variable is subtracted from which. Do not enter the calculated differences. The calculator will do all that for us.

After entering the data into these lists your Edit screen should look like this:



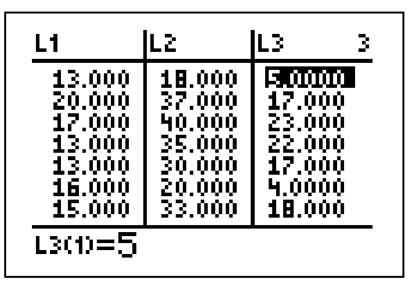
Now that the data is entered, press **2nd > QUIT** to clear the screen. The process now before us consists of two steps: subtract to find the differences, and then analyze those differences. We can't just subtract the differences and throw them away, we must store the differences in a separate list. The keystrokes to accomplish this are new to us, and a little cryptic, so we will do a little annotating after we show you the steps:



Here are the keystrokes to enter for subtracting and then storing the differences:

2^{nd} > L_2 > **--** > **2^{nd}** L_1 > **STO→** > **2^{nd}** L_3 > **ENTER**

Now you should see the following in the Edit screen:



L1	L2	L3 3
13.000	18.000	5.0000
20.000	37.000	17.000
17.000	40.000	23.000
13.000	35.000	22.000
13.000	30.000	17.000
16.000	20.000	4.0000
15.000	33.000	18.000

L3(1)=5

Now our task is to perform the hypothesis test or find the confidence interval for the data in List3. We have already discussed these procedures earlier in Chapters 9 and 10, so we refer to there for further detail.

11.4 Large Scale inferences for the Difference Between Proportions

Example 11.10: AIDS and housing availability

The authors of the article “Accommodating Persons with AIDS: Acceptance and Rejection in Rental Situations” (*J. Applied Social Psychology* (1999): 261–270) state that even though landlords participating in a telephone survey indicated that they would generally be willing to rent to persons with AIDS, they wondered whether this was true in actual practice. To investigate, two random samples of 80 advertisements for rooms for rent were independently selected from newspaper advertisements in three large cities. An adult male caller responded to each ad in the first sample of 80 and inquired about the availability of the room and was told that the room was still available in 61 of these calls. The same caller also responded to each ad in the second sample. In these calls, the caller indicated that he was currently receiving some treatment for AIDS and was about to be released from the hospital and would require a place to live. The caller was told that a room was available in 32 of these calls. Based on this information, the authors concluded that “reference to AIDS substantially decreased the likelihood of a room being described as available.”

Here are the raw data:

No AIDS Reference	$n_1 = 80$	$p_1 = 61/80 = .763$
AIDS Reference	$n_2 = 80$	$p_2 = 32/80 = .400$

Once again, the actual requirements for inference in the two-sample situation are very much like the requirements for the one-sample inference procedures. We will both perform the hypothesis test for the equality of the population proportions and construct a confidence interval for the difference in proportions using these data.

The first steps on the TI-83 for both of these procedures is the same:

STAT > `TESTS`

After that, the choices differ. However, when we discuss the two different procedures please note how similar the data entry tasks are for both hypothesis tests and confidence intervals.

The hypothesis test of the equality of the two population proportions

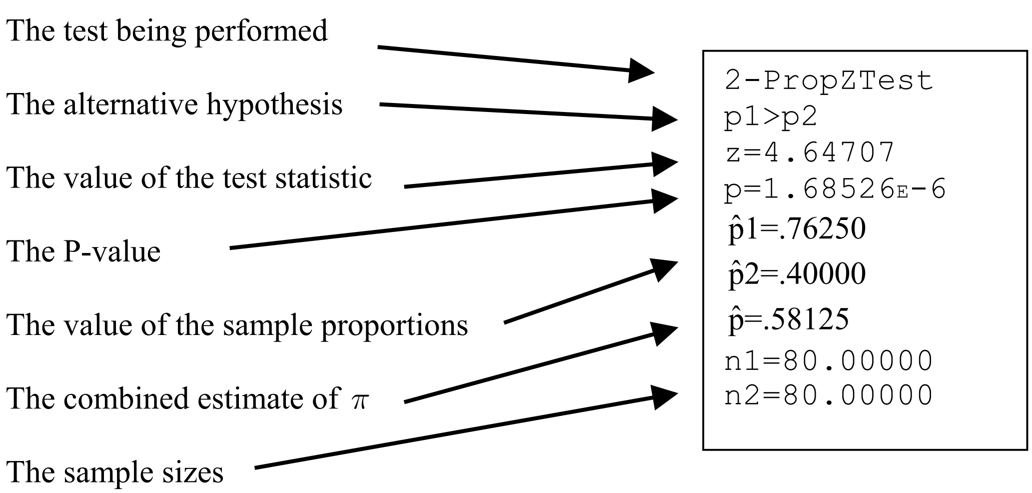
The choice for the hypothesis test is the **`2-PropZTest`**... and the results of this choice appear below:

```
2-PropZTest
x1:0
n1:0
x2:0
n2:0
p1:≠p2 <p2 >p2
Calculate Draw
```

Once again we see that the entry here is basically the entry for the single sample test of a population proportion, writ twice. Recall our warning that you need to be very careful to keep track of which proportion is which, especially since in this hypothesis test we have a one-tailed alternative. Filling in the numbers from above we should see:

```
2-PropZTest
x1:61
n1:80
x2:32
n2:80
p1:≠p2 <p2 >p2
Calculate Draw
```

The alternative hypothesis is p1>p2. Arrow down to **`Calculate`** and press **ENTER**. The notation differences between the text and the TI-83 are pretty significant here, so let's point out what this information is conveying:



The confidence interval for two population proportions

The choice for the hypothesis test is the **`2-PropZInt`**... and the results of this choice appear below:

```
2-PropZTest
x1:61
n1:80
x2:32
n2:80
C-Level:.95
Calculate
```

Notice that the calculator has "remembered" the numbers from the last 2-proportion procedure, which in our case was the hypothesis test we just performed. We might as well take advantage of this, so just ▼▼…▼ right on down to **`Calculate`** and **ENTER.**

```
2-PropZInt
 (.22030, .50470)
 p̂1=.76250
 p̂2=.40000
 n1=80.0000
 n2=80.0000
```

Once again, the confidence interval in plus/minus form can be found with a little arithmetic.