QUANTITATIVE METHODS
Martin Huard
Winter 2005

LAB 11
Hypothesis Testing

Hypothesis Testing for the Mean (Large Samples)

Example 1

Pepsi claims that their cans contain 355ml of cola. You suspect that they underfill their cans, so you collect a sample of 36 cans and find that the mean volume is 354ml with a standard deviation of 6ml. At the 5% level of significance, is this sufficient evidence to conclude that Pepsi underfills its cans using the classical approach?

Make the usual heading (cells A1:A3) and label sheet 1 "Example 1".
Enter the following in cells A5:A18, with a right alignment.

Assumptions: n =
Ho : μ =
Ha : μ <

Left-tailed Test
α =
z critical =

Sample Mean =
Sample St. Dev =

z =

Decision:

Fill in the information in the adjacent cells of column B. You need to use formulas for z critical and for z. z critical is found using the command NORMSINV, where for probability, you enter the area up to the critical value. Since we have a left tailed test, the area is α, so you can enter B10. For the z value, we use the formula $z = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}$. You must make a reference to other cells for all of the above quantities. Once this is done, you can make a decision (Fail to reject H_o or Reject H_o). Write a conclusion below it. Your sheet should now look something like this (see next page):

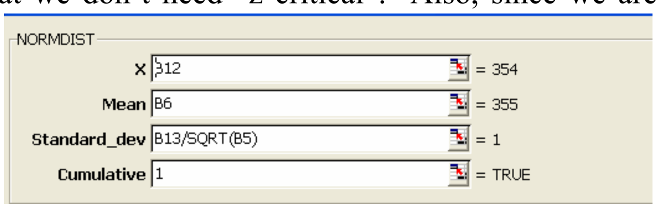
4			
5	Assumptions: n =	36	≥ 30
6	Ho : μ =	355	ml
7	Ha : μ <	355	ml
8			
9	Left-tailed Test		
10	α =	5%	
11	z critical =	-1.645	
12			
13	Sample Mean =	354	ml
14	Sample St. Dev =	6	ml
15			
16	z =	-1.00	
17		z is not in the critical region	
18	**Decision:**	Fail to Reject H0	
19			
20	Therefore there is not sufficient evidence at the 5% level of significance to conclude that Pepsi underfills its cans.		

Example 2 (same as 1 but with the *p*-value approach)

Pepsi claims that their cans contain 355ml of cola. You suspect that they underfill their cans, so you collect a sample of 36 cans and find that the mean volume is 354ml with a standard deviation of 6ml. At the 5% level of significance, is this sufficient evidence to conclude that Pepsi underfills its cans using the *p*-value approach?

Go to sheet 2 and label it "Example 2".

The procedure is the same as above, except that we don't need "z critical". Also, since we are using Excel to calculate the *p*-value, we don't need to calculate "z". To find the *p*-value, we simply use the NORMDIST function, as shown on the right. Here the X is the sample mean whereas the MEAN is the population mean. Also, the standard deviation is the standard deviation of the sampling distribution, that is $\frac{s}{\sqrt{n}}$, so you must enter this formula in the STANDARD_DEV cell.



Your results should look like this.

4			
5	Assumptions: n =	36	≥ 30
6	Ho : μ =	355	ml
7	Ha : μ <	355	ml
8			
9	Left-tailed Test		
10	α =	5%	
11			
12	Sample Mean =	354	ml
13	Sample St. Dev =	6	ml
14			
15	p-value =	0.1587	
16		*p*-value > α	
17	**Decision:**	Fail to Reject H0	
18			
19	Therefore there is not sufficient evidence at the 5% level of		
20	significance to conclude that Pepsi underfills its cans.		
21			

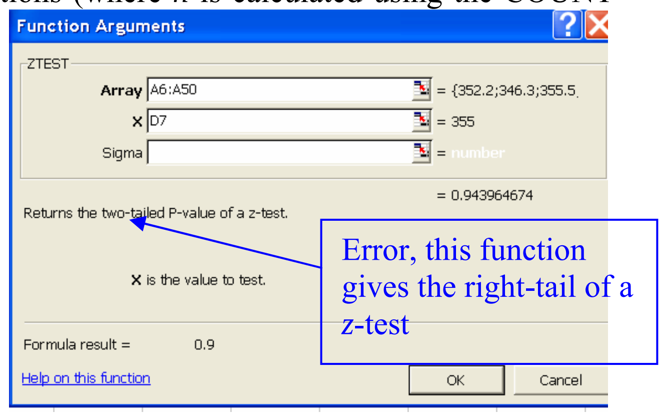
Example 3 (same as 1 but with raw data)

Pepsi claims that their cans contain 355ml of cola. You suspect that they underfill their cans, so you collect a sample cans and measure the volume. The results are given on the worksheet "Data – Volume Pepsi Cans" from my web site. At the 5% level of significance, is this sufficient evidence to conclude that Pepsi underfills its cans? Use the *p*-value approach.

Go to sheet 3 and label it "Example 3".

Copy the data in cells A6:A50. Write the assumptions (where *n* is calculated using the COUNT function), the hypothesis, the type of test and the level of significance (as shown below). In cell A14 write "*p*-value". To calculate this, we use the ZTEST command. This gives the area to the right of the *z* value. Since this is a left-tailed test, we subtract this value from 1.

With the value that you find, we compare it to the level of significance. Since the *p*-value is bigger, we fail to reject Ho.



Here is what your results should be:

	A			
4				
5	**Volume (ml)**			
6	352.2	Assumptions: n =	45	≥ 30
7	346.3	Ho : μ =	355	ml
8	355.5	Ha : μ <	355	ml
9	361.7			
10	361.2	Left-tailed Test		
11	364.4	α =	5%	
12	340.9			
13	352.6			
14	360.6	*p*-value =	0.056	
15	347.5		*p*-value > α	
16	349.9	**Decision:**	Fail to Reject H0	
17	343.9			
18	342.9	Therefore there is not sufficient evidence at the 5% level		
19	348.1	of significance to conclude that Pepsi underfills its cans.		
20	349.4			

Note: you can also use the classical approach with raw data. The method is the same as in example 1, except that you get Excel to calculate the sample mean and standard deviation.

Hypothesis Testing for the Mean (Small Samples)

Example 4

The dean at a college claims that students have, on average, 22 hours of classes per week. Sarah thinks that this is false, that students actually have more then 22 hours of classes per week. So she takes a sample of 10 students was and finds a mean of 24.3 hours of classes per week with a standard deviation of 2.3. At the 1% level of significance, can you conclude that Sarah is right? Assume that the population is normally distributed.

Go to sheet 4 and label it "Example 4".

This test is done in the same way as for hypothesis testing for the mean. The only difference is that for the critical value, we use the TINV command. This command gives the critical value for a two-tailed test. Since this is a one-tail test, we multiply the level of significance α by 2 in the dialogue box for this function.

Your results should look like this:

5	Assumptions:	Population is normally distributed	
6	n =	10	
7	Ho : μ =	22	hours
8	Ha : μ >	22	hours
9			
10	Right-tailed Test		
11	α =	5%	
12	t critical =	1.833	
13			
14	Sample Mean =	24.2	hours
15	Sample St. Dev =	2.3	hours
16			
17	t =	3.02	
18		t is in in the critical region	
19	**Decision:**	Reject H0	
20			
21		Therefore there is sufficient evidence at the 5% level	
22		of significance to conclude that Sarah is right.	