

MATHEMATICS 201-510-LW

Business Statistics

Martin Huard

Fall 2008

Inferences for μ with Excel

Confidence intervals for the mean


Example

The height (in m) of 10 students chosen at random at SLC was measured. Here are the results.

1.55 1.68 1.70 1.72 1.74 1.79 1.80 1.82 1.83 1.94

Find a 99% confidence interval for the mean height of students at SLC, assuming that the height of SLC students are normally distributed.

Make the usual heading in cells A1:A4. In cell A7 write "Height (m)". Enter each of the above numbers below this (cells A8:A17).

To construct the interval, go to DATA – DATA ANALYSIS – DESCRIPTIVE STATISTICS to access the following dialogue box. If you cannot find DATA ANALYSIS, go to  - EXCEL OPTIONS – ADD-INS and include the ANALYSIS TOOLPAK.



For the INPUT RANGE, we use the heights in cells A7:A17. Click on LABELS IN FIRST ROW and SUMMARY STATISTICS. Use C7 for the OUTPUT RANGE, and make sure that you put 99% for the CONFIDENCE LEVEL FOR MEAN. Click OK, and then adjust the columns widths accordingly.

To construct the confidence interval, go through the five steps. Your results should look like this, where the shaded cells indicate that the result was obtained with a cell reference or with a formula. For the mean, it is cell D9, for E it is cell D22, and the interval itself is given by H10-H11 and H10+H11.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Business Statistics													
2	Inferences for μ													
3	By: Martin Huard													
4	August 18, 2008													
5														
6														
7	Height (m)		Height (m)											
8	1.55													
9	1.68	Mean		1.757										
10	1.70	Standard Error		0.033235										
11	1.72	Median		1.705										
12	1.74	Mode		#N/A										
13	1.79	Standard Deviation		0.105098										
14	1.80	Sample Variance		0.011045										
15	1.82	Kurtosis		1.123976										
16	1.83	Skewness		-0.32113										
17	1.94	Range		0.39										
18		Minimum		1.55										
19		Maximum		1.94										
20		Sum		17.57										
21		Count		10										
22		Confidence Level(99.0%)		0.108008										
23														

Step 1	Assumptions: Population normally Distributed
Step 2	Test statistic: t
	1 - α = 99%
Step 3	Mean = 1.75 m
Step 4	E = 0.11
Step 5	The 95% confidence interval for the mean height of SLC students is
	1.65 m to 1.87 m

Note: This method gives the confidence interval using the t statistic, no matter what the sample size is. For a confidence interval with the z statistic (when $n \geq 30$) then use the CONFIDENCE command to find E. Note that it asks for α , not the level of significance $1 - \alpha$, so for a 99% confidence interval, you would enter 0.01 (or 1-H9). The standard deviation and the size can be found using the values calculated in Descriptive Statistics above (cells D13 and D21).

Hypothesis testing for the mean

Example

The height (in m) of 10 students chosen at random at SLC was measured. Here are the results.

1.55 1.68 1.70 1.72 1.74 1.79 1.80 1.82 1.83 1.94

Test the claim that the average height of an SLC student is 1.7 meters at the 1% level of significance, assuming that the height of SLC students are normally distributed.

For this, we start in the same way as before, that is, we use DESCRIPTIVE STATISTICS to obtain the mean and standard deviation of our sample. Next, go through the five steps. For the Greek letters μ and α and the \neq symbol, you can find them by going to INSERT – SYMBOL and going to SYMBOL in FONT. The whole Greek alphabet should be there.

To find the value of t , we use the formula $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$. Using cell references, this becomes (D9-I8)/D10 where D10 is the standard error $\frac{s}{\sqrt{n}}$. For the t critical, we use the function TINV, where for probability we have α (H12) and for the degrees of freedom $n - 1$ (D21-1). If using the p -value approach, we use the function TDIST to find the p -value, where we have the t value just calculated for x (H14), the degrees of freedom are the same as before, and for tails, we will write 2 since this is a two tailed test. Note that this last command will only work if t is positive. If it is not, then you can enter -H15 so that it will become positive. Your results should look like this (if using the classical approach):

6					
7	Height (m)	Height (m)		Step 1	Assumptions: Population normally Distributed
8	1.55			Step 2	H ₀ : $\mu = 1.7$ m
9	1.68	Mean	1.757		H _a : $\mu \neq 1.7$ m
10	1.70	Standard Error	0.033235	Step 3	Test statistic: t
11	1.72	Median	1.765		Right-tailed Test
12	1.74	Mode	#N/A		$\alpha = 5\%$
13	1.79	Standard Deviation	0.105098		t critical = 2.262
14	1.80	Sample Variance	0.011046	Step 4	$t = 1.72$
15	1.82	Kurtosis	1.123978	Step 5	t is not in the critical region
16	1.83	Skewness	-0.32113		Fail to reject H ₀
17	1.94	Range	0.39		There is not sufficient evidence, at the 5% level
18		Minimum	1.55		of significance to conclude that student heights
19		Maximum	1.94		are different than 1.7 m.
20		Sum	17.57		
21		Count	10		
22		Confidence Level(99.0%)	0.108008		
23					

Note: this test works in the same way for a large sample. We simply replace the t by z and use the commands NORMSINV and NORMSDIST instead of TINV and TDIST.