

Using Excel for statistical analysis

➤ Spreadsheet formulas: Enter these formulas, as they appear, into a cell in a worksheet. Where a formula specifies “*data range*,” enter the addresses of the beginning and ending cells of the range that contains the data. If some other argument is needed, enter either the argument or a cell address that contains the value for the argument. (Note: an “argument” here means a value that must be entered into the formula and that will change what the formula calculates. Examples appear below.)

=AVERAGE(*data range*)

arithmetic mean of the data in specified range

EX: =average(a4:a18) gives the mean of the data in column A, rows 4 through 18

=STDEV(*data range*)

sample standard deviation of data in specified range

=STDEVP(*data range*)

population standard deviation

=MEDIAN(*data range*)

=QUARTILE(*data range, number*)

“number” is the number of quartile desired: 1, 2, 3, 4. “Number” is an argument in the quartile formula.

EX: =quartile(a4:a18,3) gives the third quartile of the data in range a4:a18

=MAXIMUM(*data range*)

=MINIMUM(*data range*)

=SMALL(*data range, number*)

the nth smallest value in the data range, where n is specified as “number”

EX: =small(a4:a18,2) gives the second smallest number in the range a4:a18

NOTE: When inserting formulas into a cell, the equals sign “=” must appear as the first character in the cell, but thereafter one need not enter “=” again to indicate a formula.

EX: The expressions 4-13 and average(a4:a18) will be treated as labels and show as entered

EX: =maximum(a4:a18)-minimum(a4:a18) will give the range of the data in a4:a18

General solution?

- NORMSDIST(z value): proportion of distribution \leq given z value
- NORMSINV(probability): value z_0 such that $P(z \leq z_0)$ has the probability entered

EXAMPLES: =normsdist(1.96) will give the result 0.975
=normsdist(-1.96) yields 0.025
=normsinv(.05) yields -1.64485
=normsinv(.95) yields 1.64485

- TDIST(t value, degrees of freedom, tails)
Here t value is calculated from the data, degrees of freedom are specified by the nature of the problem, and the argument tails has values 1 or 2. If tails = 1, the formula gives the area in one tail of a t distribution beyond the specified t value; if tails = 2, the formula gives the area in both tails. The tdist function is used to find the p-value in hypothesis tests; it gives the probability of a t value as large or larger than the one actually obtained.
- TINV(probability, degrees of freedom) gives the two-tailed t value for the specified probability.
The tinv function is used for finding t values to be used in calculating confidence intervals and for finding critical t values in hypothesis tests.

EXAMPLES: =tdist(2,9,2) will yield 0.076553
=tdist(2,9,1) will yield 0.038276
note that $0.076553 = 2 \times 0.038276$
=tinv(0.05,9) will show 2.262159, the value that would be used in constructing a 95% confidence interval
=tinv(0.1,9) will show 1.833114, the critical t-value for an upper one tail test at 10% significance level