The *t* Test

In biology you often want to compare two sets of replicated measurements to see if they are the same or different. For example are plants treated with fertiliser taller than those without? If the means of the two sets are very different, then it is easy to decide, but often the means are quite close and it is difficult to judge whether the two sets are the same or are significantly different. The t test compares two sets of data and tells you *the probability (P) that the two sets are basically the same.*

P varies from 0 (not likely) to 1 (certain). The higher the probability, the more likely it is that the two sets are the same, and that any differences are just due to random chance. The lower the probability, the more likely it is that that the two sets are significantly different, and that the differences are real. Where do you draw the line between these two conclusions? In biology the critical probability is usually taken as 0.05 (or 5%). This may seem very low, but it reflects the facts that biology experiments are expected to produce quite varied results. So if $P > 0.05$ then the two sets are the same, and if $P < 0.05$ then the two sets are different. For the t test to work, the number of repeats should be as large as possible, and certainly > 5.

(T.TEST for Excel 2008 and later)

In Excel the t test is performed using the formula: =TTEST (range1, range2, tails, type) . For the examples you'll use in biology, tails is always 2 (for a "two-tailed" test), and type can be 1 or 2 depending on the circumstances.

	A	B	C
1	lymphocyte count	infected patients	unaffected patients
2		150	165
3		155	170
4		152	151
5		146	164
6		152	160
7	mean	151	162
8	conf. limit	2.91	6.23
9	t-test *P*	0.014	
10			
11			

=TTEST(B2:B6,C2:C6,2,2)

The usual form of the t test is for "unmatched pairs" (type=2), where the two sets of data are from different individuals. For example the number of white blood cells of 5 patients infected with a parasite were compared with 5 unaffected individuals. The results are shown in the spreadsheet on the right, and the means and confidence limits have been worked out as usual. The infected patients certainly have a smaller mean white blood cell count, but is it significantly smaller? Cell B9 has the t-test probability, which is 0.014. This probability is smaller than the critical value of 0.05, so the two groups are significantly different, and the infected patients really do have significantly fewer white cells than normal.

	A	B	C
1	**pulse rate**	**before eating**	**after eating**
2		105	109
3		79	87
4		79	86
5		103	109
6		87	100
7		74	82
8		73	80
9		82	90
10	**mean**	85.25	92.875
11	**conf. limit**	8.58	8.05
12	**t-test *P***	**0.000065**	
13			
14			

=TTEST(B2:B9,C2:C9,2,1)

The other form of the t test is for "matched pairs" (type=1), where the two sets of data are from identical individuals. A good example of this is a "before and after" test. For example the pulse rate of 8 individuals was measured before and after eating a large meal, with the results shown in the left. The mean pulse rate is certainly higher after eating, but is it significantly higher? Cell B12 has the t-test probability, which is a tiny 0.000065, and indicates that the difference between the before and after results is <u>very highly</u> significant. If we had used the normal unmatched pairs t-test, we would have obtained a *P* of 0.225, which is higher than 0.05, so indicates that the apparent increase in pulse rate with eating is <u>not</u> significant! This shows the importance of choosing the right test.