Bio Factsheet

Number 85

The Paired t-Test & When To Use It

The t-test is widely used within project work. This Factsheet will tell you when and how to use the paired t-test. Questions using the t-test may occur on exam papers, but you will not have to remember the formula - it will be given to you.

Hypotheses and 1 or 2-tailed tests.

The t-test is used to compare two *means* (the mean is the normal "average"). As with all other statistical tests, this involves choosing between two *hypotheses*. For the t-test, the *null hypothesis* (H_0) is **always** that the two means are equal - for example "mean length of leaves at the top of the tree = mean length of leaves at the bottom of the tree". It can **never** be that the means are different, or that one is bigger/smaller than the other. You **always** start out assuming that the null hypothesis is true, and only change your mind if the evidence is good enough.

The other hypothesis - *alternative hypothesis* (H_1) - can take one of three forms (Table 1); the precise form of it tells you whether you are carrying out a *one-tailed* or a *two-tailed* test. A one-tailed test means that you are only considering one alternative - for example that the leaves are bigger at the top of the tree. A two-tailed test means you are considering either alternative - that the leaves could be either bigger or smaller at the top of the tree.

Table 1. Alternative hypotheses for the t-test

Alternative Hypothesis	Example	1-tailed or 2-tailed	Use it if you are expecting....
mean 1 > mean 2	mean leaf length at top > mean leaf length at bottom	1-tailed	leaves to be longer at the top
mean 1 < mean 2	mean leaf length at top < mean leaf length at bottom	1 tailed	leaves to be shorter at the top
mean 1≠ mean 1	mean leaf length at top ≠ mean leaf length at bottom	2-tailed	not sure - could be either!

Paired and Unpaired Tests

The t-test may be **paired** or **unpaired**. The calculations involved are different in each case. This Factsheet concentrates on the method for the (commoner) paired test; a later Factsheet will deal with the unpaired test.

Paired tests are used when the data occur in "natural" pairs. Examples include:
- the same organism reacting to two different stimuli.
- comparisons of plant growth between points in the same position on different sides of a hedge.
- measurements made on identical twins.

Note that the pairs must be matched in a natural way - you cannot just decide to pair results up.

The idea of pairing is to increase the chance that any differences really are due to what you are investigating - e.g. which side of the hedge the point is - rather than due to other variations - e.g. soil quality.

If a paired test is not applicable, then an unpaired test must be used. Table 2 shows some examples of investigations requiring paired and unpaired tests.

Table 2. t-test investigations

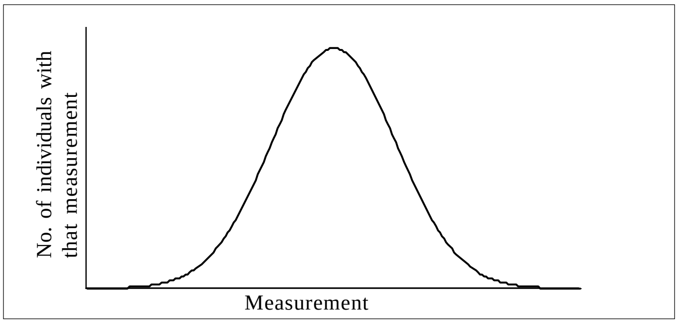
Investigation	Paired or Unpaired?
Comparison of plant growth on two sides of a hedgerow	Paired
Effect of different soil types on the yield of a crop plant	Unpaired
Effect of pollution on vegetation	Unpaired

Exam Hint: - If you are unsure, it is better to use an unpaired test. Even if a paired test would have been suitable, it is not wrong to use an unpaired one (although it is harder to get a significant result). But using a paired test where it is not appropriate is ***definitely wrong*** *and will not gain you marks!*

When can the t-test be used?

You can only use the t-test when the population your observations come from is *normally distributed*. That means that if you were to collect the same data from every member of the population and draw a graph of your results, you would get a symmetrical, bell-shaped graph (Fig 1).

Fig 1. Normal distribution curve



What sorts of things are normally distributed?

Normally distributed variables must be *continuous* - this means that they can take any value (like lengths), rather than only set values (like numbers of organisms - you can't have 2.85 spiders!)

Generally speaking, anything that is **measured** will be normally distributed. This includes length, width, height, weight and velocity.

Anything that is **counted** (like numbers of organisms) or **calculated** (like a diversity index) will **not** be normally distributed. In these cases, you should use the Mann-Whitney U-test instead.

Sample size

You cannot use the t-test unless you have at least 6 measurements in each sample for the paired test.

However, you will find it much easier to get a significant result if you have larger numbers than this - aim for at least 10 in each sample.

Worked Example - paired t-test

Eight adults who normally ate a "junk-food" diet agreed to follow a "healthy" diet for two months. The following data gives their weights in kilograms immediately before and immediately after the two-month period.

Person	A	B	C	D	E	F	G	H
"Before" weight	76.4	85.9	50.6	62.4	90.5	55.7	78.3	67.4
"After" weight	74.5	80.1	50.8	62.3	79.4	55.7	73.2	65.0

Step 1: Write down the **hypotheses**, deciding which form of the alternative hypothesis you want to (use and so whether it is 1-tailed or 2-tailed).

Since we would expect a healthy diet to reduce weight compared to a junk-food diet, we have:
H_0: Mean weight "before" = Mean weight "after"
H_1: Mean weight "before" > Mean weight "after" So it is a 1-tailed test

Step 2: Decide whether it is **paired** or **unpaired**

This is paired, because we are considering the same people "before" and "after"

Step 3: For a paired test, work out "1st value - 2nd value" for each individual

Person	A	B	C	D	E	F	G	H
"Before - "After"	1.9	5.8	-0.2	0.1	11.1	0	5.1	2.4

(76.4 - 74.5)

Step 4: Work out the mean of these differences by adding them up and dividing by how many there are.

(1.9 + 5.8 + -0.2 + 0.1 + 11.1 + 0 + 5.1 + 2.4) ÷ 8 = **3.275**

Step 5: Calculate the standard deviation **s**, using

$$s = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$$

where Σx^2 means square each difference, then add up
n = no. of values
$\bar{x}$ = mean of the differences
Σ = add up

$\Sigma x^2 = 1.9^2 + 5.8^2 + (-0.2)^2 + 0.1^2 + 11.1^2 + 0^2 + 5.1^2 + 2.4^2 = 192.28$
n = 8

$$s = \sqrt{\frac{192.28}{8} - 3.275^2} = 3.648$$

> ***Exam Hint:*** *- When you are working out Σx^2, you **must** square each value then add them - you **cannot** add then square afterwards.*

Step 6: Work out the **test statistic,** using

$$\text{t-test} = \frac{\bar{x}\sqrt{n-1}}{s}$$

$$\text{t-test} = \frac{3.275\sqrt{(8-1)}}{3.648} = 2.375$$

Step 7: Work out the **degrees of freedom**, using degrees of freedom = n - 1

degrees of freedom = 8 - 1 = 7

> ***Exam Hint*** *- Do not worry about what this means, but remember to use n - 1*

Step 8: Get a t-test table and **look up the value** for the significance level (usually 5%), the degrees of freedom and whether it is a 1-tailed or 2-tailed test.

We are carrying out a 1-tailed test;
we will use a 5% significance level.
We need to use 7 degrees of freedom.

So our tables value is **1.895**

Table 3. t-table

	Significance level				
df	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055

Step 9: **Make a decision** - if your t-value is **bigger** than the one from the tables, you can **reject** the null hypothesis. Otherwise you have to accept it.

Our test-value (2.539) is bigger than the tables value.
So we reject the null hypothesis.
So the mean weight before is significantly greater than the mean weight after.

Practice Question

The plant known as Indian Mustard (Brassica juncea) will accumulate gold from gold salts in the soil in which it grows. It can be used to extract gold from impoverished mine wastes which still contain residual gold which is uneconomical to extract by traditional methods. Addition of ammonium thiocyanate to the mine waste growth substrate is thought to enhance the accumulation of gold by the plants. An investigation was carried out to test whether this was so. The results of the investigation are shown in the table below.

plant	gold accumulated/ mg gold Kg^{-1} dry mass		x = b - a	x^2
	no thiocyanate(a)	with thiocyanate(b)		
1	1.01	0.99		
2	1.09	1.09		
3	0.98	1.25		
4	0.71	1.34		
5	1.15	1.22		
6	1.21	1.18		
7	1.18	1.23		
8	0.89	1.34		
9	1.26	1.21		
10	1.07	0.97		
11	1.30	1.18		
12	1.17	1.23		

The null hypothesis was proposed that:

the mean mass of gold accumulated without thiocyanate = mean mass of gold accumulated with thiocyanate

The alternate hypothesis considered was:

mean mass of gold accumulated without thiocyanate ≠ mean mass of gold accumulated with thiocyanate

(a) Carry out a t-test to test the validity of the null hypothesis.

The formula for this t-test is: $\mathbf{t = \dfrac{\bar{x}\sqrt{(n-1)}}{s}}$

where $\bar{x}$ is the mean of the weight differences, n is the number of readings in each group and s is found from the formula:

$$\mathbf{s^2 = \frac{\sum x^2}{n} - \bar{x}^2}$$

(i) Calculate the value of $\bar{x}$. 3
(ii) Calculate the value of $\sum x^2$. Show your working. 2
(iii) Calculate the value of $(\bar{x})^2$. Show your working. 2
(iv) Calculate the value of s. Show your working. 2
(v) Calculate the value of t. Show your working. 2

(b) At 11 degrees of freedom the critical values for t at various probabilities are shown in the table.

Probability	0.1	0.05	0.02	0.01	0.001
Critical value	**1.80**	**2.20**	**2.72**	**3.11**	**4.44**

(i) Does the calculated value for t enable you to accept or reject the null hypothesis? Explain your answer. 2
(ii) Suggest two precautions which should have been taken to ensure that the investigation yielded valid results. 2

Total 14

Answer

(a) (i)

plant	gold accumulated/ mg gold Kg^{-1} dry mass		x = b - a	x^2
	no thiocyanate(a)	with thiocyanate(b)		
1	1.01	0.99	-0.02	0.0004
2	1.09	1.09	0	0.0000
3	0.98	1.25	0.27	0.0729
4	0.71	1.34	0.63	0.3969
5	1.15	1.22	0.07	0.0049
6	1.21	1.18	-0.03	0.0009
7	1.18	1.23	0.05	0.0025
8	0.89	1.34	0.45	0.2025
9	1.26	1.21	-0.05	0.0025
10	1.07	0.97	-0.1	0.0100
11	1.30	1.18	-0.12	0.0144
12	1.17	1.23	0.06	0.0036

x values;
$\sum x = 1.21$;
$\bar{\mathbf{x}} = \dfrac{1.21}{12} = 0.101$; 3

(ii) $\mathbf{x^2}$ values;
$\sum \mathbf{x^2} = 1.464$; 2

(iii) $\mathbf{(\bar{x})^2} = 0.010(2)$; 1

(iv) $\mathbf{s^2} = \dfrac{1.464}{12} - 0.010 = 0.112$;
$\mathbf{s} = 0.335$; **2**

(v) $\mathbf{t} = \dfrac{0.101\sqrt{12-1}}{0.335}$; $= 1.00$; (allow 0.999) 2

(b) (i) accept the null hypothesis;
calculated value is less than the critical value (at 0.05 probability level/at all probability levels shown); 2

(ii) use same strain/seed batch/similar Brassica seedlings;
mine waste material should have identical gold contents to begin with;
use a standard concentration of thiocyanate;
grow under similar conditions of temperature/humidity/water availability; max 2

Total 14

Acknowledgments: This Factsheet was researched and written by Cath Brown. *Curriculum Press, Unit 305B The Big Peg, 120 Vyse Street, Birmingham B18 6NF*
 ISSN 1351-5136