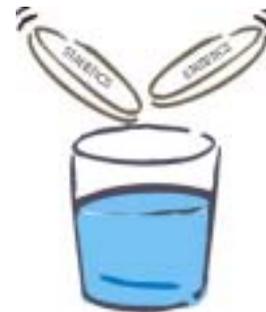


# Statistics in Divided Doses



April 2003 No 6

## Analysing small samples – the Student's *t* test

### Contents

Dealing with small samples

The *t* distribution

Using the *t* test to compare:

- a test value to a 'normal' value – (a one-sample *t* test)
- paired samples
- unpaired samples

### Dealing with small samples

In *Statistics in Divided Doses 5 (SiDD5)* we used the *z* statistic, calculated from the observations made in the *Rumbocur* study, to test the null hypothesis. We compared our test *z* value with those in the Normal probability distribution table to determine the probability of our finding having occurred if the null hypothesis were true. Using a test based on Normal distribution was acceptable in this case as our sample size was relatively large. However, such tests are less reliable for smaller samples because these are more likely to be subject to chance variation. For small samples it is preferable to use the *t* test, which is based on the *t* distribution.

### The *t* distribution

#### What is the *t* distribution?

The *t* distribution is a theoretical probability distribution. Its bell-shaped curve is similar to that of the Normal distribution but it goes 'flaky' at the tail edges (i.e. the periphery of the distribution) because of the effect of chance variation. The mathematical parameter that defines the *t* distribution is **degrees of freedom** (see later), which is closely affiliated to sample size. The *t* distribution was first described by Gosset who wrote under the pseudonym "Student". The test based on the *t* distribution is commonly referred to as the Student's *t* test or more simply, the *t* test. Gosset worked at the Guinness brewery in Dublin and a pint at this point may help you understand what follows!

#### Why is the *t* distribution important?

The *t* distribution forms the foundation of nearly all

parametric methods of analysis. (Parametric analysis, you will remember from *SiDD3*, assumes that the data are from a population with a known probability distribution.)

The *t* distribution can be used to perform hypothesis tests and calculate confidence intervals for samples that are too small for analysis using the Normal distribution.

#### Should the *t* test be used only for analysis of small samples?

No. For large samples the *t* test will provide virtually identical answers to tests based on the Normal distribution. In practice, therefore, the *t* distribution can be used routinely.

#### What do tables based on the *t* distribution look like and how do I use them?

Table 1. *t* distribution

Degrees of freedom	Two-tailed P value			
	0.10	0.05	0.01	0.001
1	6.314	12.706	63.657	636.619
2	2.920	4.303	9.925	31.599
3	2.353	3.182	5.841	12.924
4	2.132	2.776	4.604	8.610
5	2.015	2.571	4.032	6.869
6	1.943	2.447	3.707	5.959
7	1.895	2.365	3.499	5.408
8	1.860	2.306	3.355	5.041
9	1.833	2.262	3.250	4.781
18	1.734	2.101	2.878	3.922

The left hand column of the table relates to degrees of freedom. This concept was first mentioned in *SiDD2* where it was described as 'elusive'. If it helps, degrees of freedom relate to the estimated standard deviation which describes the variation within the sample around the estimated mean. If it doesn't help, all you need to know for the *t* distribution is that degrees of freedom are calculated by subtracting 1 from the sample size.

The right-hand columns of the table refer to levels of probability. Because, in general, the effect we're interested in e.g. change in blood pressure, can go up as well as down, we use a two-tailed test to allow for either eventuality. The body of the table contains

values based on the *t* distribution against which we compare our test value.

As a brief introduction to the table, let's imagine we take a sample of 6 men, measure their heights and then calculate the mean height. We now want to calculate 95% confidence intervals around this mean. If we were using an analysis based on Normal distribution, we would use a figure of 1.96 to calculate our confidence intervals (see *SiDD3*). However, as this is a small sample we use the *t* distribution. Go to the table and read across the row for five degrees of freedom (sample size 6 minus 1). The value for *t* corresponding to a probability of 0.05 is 2.571. This figure, 2.571, is the one we use to calculate 95% confidence intervals for our sample mean. Clear as mud? The examples below will hopefully clarify issues.

All examples are based on our fictitious drug *Rumbocur* and its effects on blood levels of protein X. You may need to refer back to past issues of *SiDD* for some of the calculations (*SiDD2* standard deviations; *SiDD3* standard error of the mean & confidence intervals). You'll find that our formulae for calculating standard error of the difference between means differs from that given in *SiDD5* because we are dealing with paired data and small sample sizes. Don't worry – have another pint.

**Using the *t* test to compare a test value to a 'normal' value – a one sample *t* test**

**How can I compare a test value with a 'normal' value?**

We want to find out whether *Rumbocur* affects protein X levels. Medical reference texts tell us that in the general population the average protein X level is 20ng/ml.

We measure protein X levels in a sample of 10 patients (*n*=10) on *Rumbocur* and obtain the following values:

- 8, 9, 10, 11, 11, 12, 12, 13, 14, 16 ng/ml.
- The mean level,  $\bar{x} = 11.6\text{ng/ml}$
- The standard deviation,  $s = 2.37\text{ng/ml}$
- The standard error of the mean,  $SE = 0.75\text{ng/ml}$

Our null hypothesis is that the observed difference between our sample mean and that of the population is entirely due to random variation, i.e. that our sample is from a population with a mean protein X level of 20ng/ml.

To calculate the *t* statistic, we use the formula:

$$t = \frac{\text{observed value} - \text{reference value}}{SE \text{ of observed value}}$$

therefore:

$$t = \frac{11.6 - 20}{0.75} = -11.2$$

Go to the table of *t* distribution, read across the row for 9 degrees of freedom (10 minus 1); you'll find that our *t* value of 11.2 (ignore the negative sign) is greater than any of the values given, including that column corresponding to  $P=0.001$ . We therefore have strong evidence (probability <0.001) to reject our null hypothesis.

**How should I interpret these results?**

With care, as usual! It is reasonable to assume that *Rumbocur* depresses protein X levels, but we need to consider whether other factors may have contributed to the results. Confidence intervals around the observed effect are more informative than the P value alone.

**Using the *t* test to compare paired samples**

**What are paired samples?**

In certain situations, we may be interested in determining the response of the same group of subjects to more than one treatment. To do this, we could, for example, use a clinical trial of cross-over design, where every individual gets each treatment, one after the other in random order, usually with a washout period in between. The data that we obtain are **dependent** since they are obtained from the same individuals. We refer to such data as paired data as they come from related or **paired samples**.

**How do I use the *t* test to analyse paired data?**

Instead of using general population data from a medical text as we did in the last example, we perform a trial in which we measure protein X levels in our 10 patients after 2 weeks on placebo and after two weeks on *Rumbocur* therapy. We find that the average level following placebo ( $\bar{x}_1$ ) is 19.1ng/ml, and that during treatment with *Rumbocur* ( $\bar{x}_2$ ) is 11.6ng/ml. Our mean observed difference ( $\bar{d}$ ) is therefore, 7.5ng/ml (see table 2 – but ignore the section for unpaired samples!).

Our null hypothesis is that the mean difference in protein X levels on placebo and *Rumbocur* treatment is zero i.e. that both samples in our trial are from the same population.

**Table 2: Cross over trial results**

Patient	Protein X levels ng/ml		Difference $x_1 - x_2$	$d^2$
	After placebo $x_1$	After <i>Rumbocur</i> $x_2$		
1	19	8	11	121
2	20	9	11	121
3	23	10	13	169
4	15	11	4	16
5	21	11	10	100
6	18	12	6	36
7	17	12	5	25
8	19	13	6	36
9	15	14	1	1
10	24	16	8	64
<b>Total (<math>\Sigma</math>)</b>	<b>191</b>	<b>116</b>	<b>75</b>	<b>689</b>
<b>Mean</b>	<b>19.1</b>	<b>11.6</b>	<b>7.5</b>	
<b>For our unpaired samples</b>				
<b>Standard deviation</b>	<b>S<sub>1</sub> = 3.03</b>	<b>S<sub>2</sub> = 2.37</b>		

To test our null hypothesis, our  $t$  statistic is:

$$t = \frac{\text{observed difference} - \text{hypothesised difference}(0)}{\text{standard error of the difference}}$$

To calculate the standard error of the difference, we first need to calculate the standard deviation of the difference,  $s_d$ , by substituting figures from the table into the formula:

$$S_d = \sqrt{\frac{\sum \bar{d}^2 - \frac{(\sum \bar{d})^2}{n}}{n - 1}}$$

Note: this is a short-cut (and more convenient) version of the "official" formula for the standard deviation given in *SiDD2*.

Substituting our data, we obtain;

$$S_d = \sqrt{\frac{689 - \frac{75^2}{10}}{10 - 1}}$$

$$= 3.75 \text{ ng/ml}$$

The standard error of the difference,  $SE_d$ , can then be calculated;

$$SE_d = \frac{s_d}{\sqrt{n}} = \frac{3.75}{\sqrt{10}} = 1.19 \text{ ng/ml}$$

which means that our  $t$  statistic is:

$$t = \frac{\bar{d} - 0}{SE_d} = \frac{7.5 - 0}{1.19} = 6.3$$

Entering the table of  $t$  distribution at 9 degrees of freedom we find that our  $t$  value is greater than the  $t$  value corresponding to  $P = 0.001$ . We can therefore reject the null hypothesis at the 0.001 level of probability and conclude that the effect of *Rumbocur* in lowering protein X levels is statistically significant.

### How do I calculate the confidence intervals around the mean difference in this example?

In the  $t$  distribution table, we find that with 9 degrees of freedom and a  $P$  value of 0.05 (corresponding to 95% probability),  $t = 2.262$ . In our sample, 95% of values will lie within 2.262 standard errors around the mean.

Our mean difference is 7.5 ng/ml  
 Our standard error of the difference is 1.19  
 Our 95% confidence intervals are therefore, 7.5 - (2.262 x 1.19) to 7.5 + (2.262 x 1.19)

The calculated confidence intervals (4.8 to 10.2 ng/ml) are reasonably narrow. We will discuss interpretation of confidence intervals in more detail in a later issue.

## Using the $t$ test to compare unpaired samples

### What are unpaired data?

We obtain unpaired data when we compare outcome in two separate groups of subjects. These data are **independent** as each subject provides only one result. Such data are obtained from the most common type of clinical study, the parallel trial, in which each patient receives one of two (or more) treatments that have been randomly assigned.

### How do I use the $t$ test to analyse unpaired data?

With more difficulty than for paired data. Another pint might come in handy here!

Let's assume that the values in table 2 refer to data from two independent groups of patients receiving placebo ( $x_1$ ) and *Rumbocur* ( $x_2$ ).

For the placebo group;  
 the mean protein X level,  $\bar{x}_1 = 19.1 \text{ ng/ml}$   
 the standard deviation,  $s_1 = 3.03 \text{ ng/ml}$

For the *Rumbocur* group;  
 the mean protein X level,  $\bar{x}_2 = 11.6 \text{ ng/ml}$   
 the standard deviation,  $s_2 = 2.37 \text{ ng/ml}$

Our null hypothesis is that both samples are from the same population and that the mean difference between them will be zero.

In this case, we calculate  $t$  by dividing the difference in the mean levels by the standard error of that difference. The formula for  $t$  for unpaired data is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE(\bar{x}_1 - \bar{x}_2)}$$

The difficulty arises with the standard error of the difference. With unpaired data we are unable to calculate the differences in levels since the data are independent. To get around this we first calculate the **pooled standard deviation** ( $s_p$ ) of the two samples using the formula:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Here, we're multiplying the individual degrees of freedom (9 for each sample) by the square of the standard deviation of each sample and adding the results together. We then divide the results by the degrees of freedom for the pooled samples.

$$s_p = \sqrt{\frac{(9 \times 3.03^2) + (9 \times 2.37^2)}{18}}$$

$$= \sqrt{7.38} = 2.72 \text{ ng/ml}$$

To calculate our standard error of the difference, we use the formula;

$$SE(\bar{x}_1 - \bar{x}_2) = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 2.72 \times \sqrt{(1/10 + 1/10)} = 1.22 \text{ ng/ml}$$

Substituting these values into the formula for  $t$  above:

$$t = \frac{19.1 - 11.6}{1.22} = 6.15$$

Entering the  $t$  distribution at 18 degrees of freedom we find that our  $t$  value of 6.15 is greater than any of the given values. We can, therefore, reject the null hypothesis with the usual reservations.

### **How do I calculate a 95% confidence interval for the difference in sample means?**

The  $t$  value from the distribution table corresponding to  $P = 0.05$  at 18 degrees of freedom is 2.101

The standard error of the mean is 1.22

This gives the 95% CI as  $7.5 \pm (2.101 \times 1.22)$   
i.e. 5.0 to 10.1 ng/ml

### **Are there any conditions attached to use of the $t$ test?**

Yes – the data should be from an approximately Normal distribution. For unpaired samples the standard deviations of the two samples should be similar.

If the data do not have an approximately Normal distribution, or we make no assumptions about the distribution, we must use non-parametric or 'distribution free' methods of analysis. These will be discussed in the next issue, *SiDD7*. The alcohol withdrawal programme will have to wait!