THE BASICS OF DATA LITERACY

HELPING YOUR STUDENTS *(AND YOU!)* MAKE SENSE OF DATA

MICHAEL BOWEN
ANTHONY BARTLEY

NSTApress
National Science Teachers Association
Arlington, Virginia

THE *t*-TEST

The *t*-test is used when you compare *two* means to see if they are statistically different from each other. You should *not* use a *t*-test over and over to compare many pairs of means (see the ANOVA test description for how to deal with that situation). What the *t*-test is doing is determining what the likelihood is that the difference between the two means happens because of chance or because of the variable you tested. In simple terms, it's comparing how much data scatter there is for each variable and then comparing how different the means are in relation to that data scatter so that the likelihood of the differences between the means being due to random chance can be determined.

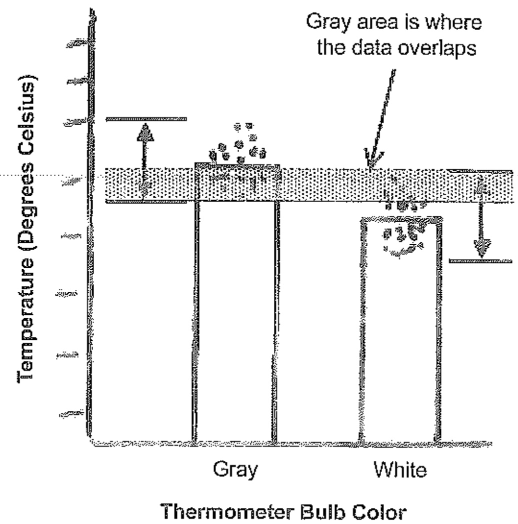
It might be a bit simpler if we looked at a graph of data (Figure 8.1 here, which you might recognize from Figure 1.7, p. 10).

A *t*-test would help you determine whether the amount of overlap in the data would be statistically significant so that you could argue that the two means are different from each other.[2]

Every test has conditions (also known as assumptions) that must be met for the results to be valid. If you meet those conditions, statistical tests are pretty good at letting you know whether there's a statistically significant difference between means, but if you violate those assumptions then the tests might not be accurate. Here are the assumptions that should be met to do the *t*-test:

FIGURE 8.1

Graph of temperature data with arrows depicting range of response and the gray area depicting where the data overlaps



1. The data scatter is reasonably the same for the two categories (in statistical terms, the *variation* is close to the same).

2. There is more data toward the middle of the circles than at the nearest and farthest points away from the middle (in statistical terms, the data has a reasonably *normal distribution).*

3. The data are *randomly* chosen (in statistical terms, this means you didn't choose data to include so that you showed what you wanted to show).

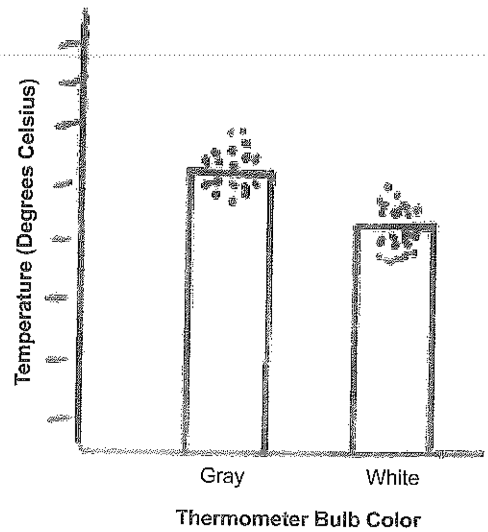
4. The replicates in the two treatments need to be *independent* of each other. For instance, the data *cannot* be before and after measures on the same individuals (there's a separate test for that called the paired *t*-test).

2. We've actually done this in Appendix IX. Go and take a look at whether the means are significantly different or not for this data set.

It might be easier to show you what this means on a graph. Let's look at Figure 1.7 (the gray and white thermometer data, p. 10) again in Figure 8.2.

FIGURE 8.2

Depiction of gray and white temperature data portraying even data scatter

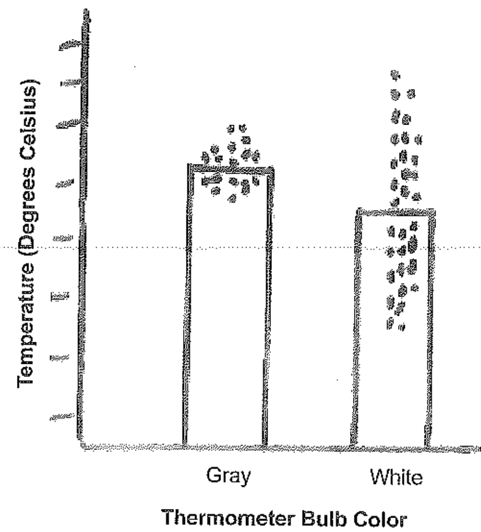


You'll notice that the data in this graph meets the assumptions listed above: The raw data depicted around the two bars is about the same distance from top to bottom, and there are more data points close to the horizontal line than far away from it.

Let's look at a couple of extreme versions of data for the same variables that do not meet those assumptions for a *t*-test (Figure 8.3).

FIGURE 8.3

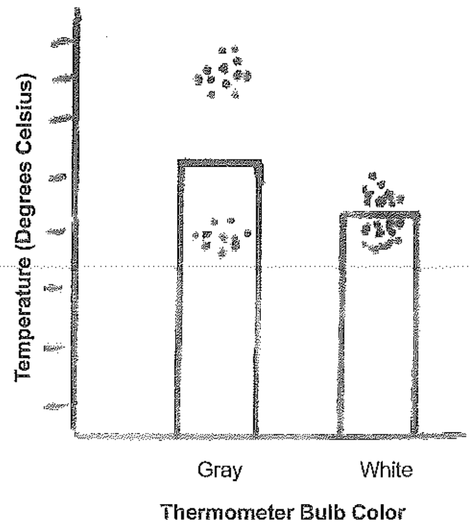
Depiction of gray and white temperature data portraying uneven data scatter



Notice that in Figure 8.3, the data on the right bar is much more scattered (has a greater variation) than the data on the left. This violates assumption 1. Now we'll look at a graph that violates assumption 2 (Figure 8.4, p. 54).

FIGURE 8.4

Depiction of gray and white temperature data portraying discontinuous data scatter



Here, you'll notice that on the left bar data are not close to the horizontal line at all; the horizontal line is at the average of two separate clusters. This condition violates assumption 2 because the data are not normally distributed (i.e., more toward the horizontal line than away from it).

The more your data looks like the last two graphs (and data could look like a combination of both of them), the less likely it is that the results of the *t*-test are reliable. In that situation you might *hedge* how you phrased your interpretation of the data analysis. For instance, if you found a significant difference (as we will describe below), you could write, "Despite finding a significant difference between the mean temperatures for the gray and white thermometer, there is still some room for doubt because of the amount of variation in the data for the white thermometer, which was much greater than that for the gray thermometer."

However, having described the problem, realize that the *t*-test is a reasonably robust test and is fairly accurate even if its assumptions are violated.

So, how do you calculate a *t*-test? The step-by-step worksheet and example in Appendix IX will show you. When you calculate your *t*-test statistic using the worksheet, compare your calculated value to the table value in Table 8.1.

TABLE 8.1

Critical values for the *t*-test statistic

5% Significance Table			
Degrees of freedom	Critical value	Degrees of freedom	Critical value
4	2.78	15	2.13
5	2.57	16	2.12
6	2.48	18	2.10
7	2.37	20	2.09
8	2.31	22	2.07
9	2.26	24	2.06
10	2.23	26	2.06
11	2.20	28	2.05
12	2.18	30	2.04
13	2.16	40	2.02
14	2.15	60	2.00
		120	1.98

If the *t*-statistic you calculated is *less than* the critical value in the table above (for the correct degrees of freedom, which you calculate on the worksheet) then the difference between the two means is not statistically significant.

If the calculated *t*-statistic is *greater than* the critical value in the table above (for the correct degrees of freedom) then the difference between the two means is statistically significant at 5%. This means we're 95% confident that the difference between the means is a real one (i.e., not due to chance).

t-TEST WORKSHEET

(Using the dog and cat data from Table 3.1, p. 19.)

TABLE 3.1

Example of how data are normally recorded in a table, using data on the number of hours dogs and cats sleep

Pet	Dogs sleeping	Cats sleeping
1	8	14
2	9	5
3	5	7
4	8	8
5	7	7
Avg.	7.4 hours/day	8.2 hours/day

	Group 1: Dogs sleeping		Group 2: Cats sleeping	
	(a)	(a) × (a)	(b)	(b) × (b)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
Sum	(c)	(d)	(e)	(f)
Count	(g)		(h)	

Step 1: Enter data in table in rows (a) and (b).

Step 2: Square (a) and put in column (a) × (a); square (b) and put in column (b) × (b).

Step 3: Sum columns (a), (a) × (a), (b) and (b) × (b) and put results on Sum row.

Step 4: Count measures in column (a) and (b) and enter them on the Count row.

Step 5: Calculate: $\frac{(c) \times (c)}{(g)}$ = ________ (i) $\frac{(e) \times (e)}{(h)}$ = ________ (j)

Step 6: Calculate: (d) – (i) = ________ (k) (f) – (j) = ________ (l)

Step 7: Calculate: (k) + (l) = ________ (m)

Step 8: Calculate: $\frac{(m)}{(g) + (h) - 2}$ = ________ (n)

Step 9: Calculate: $(n) \times \left(\frac{(1)}{(g)} + \frac{(1)}{(h)}\right)$ = (n) × (________ + ________) = ________ (o)

Step 10: Calculate: $\sqrt{(o)}$ = ________ (p)

Step 11: Calculate: $\frac{(c)}{(g)}$ = ________ (q) $\frac{(e)}{(h)}$ = ________ (r)

Step 12: Calculate: (q) – (r) = | ________ | (s) (Note: Absolute value of [s])

Step 13: Calculate: *t*-statistic = $\frac{(s)}{(p)}$ = ________

Step 14: Calculate: degrees of freedom (d.f.) = (g) + (h) – 2 = ________

5% Significance Table			
Degrees of freedom	**Critical value**	**Degrees of freedom**	**Critical value**
4	2.78	15	2.13
5	2.57	16	2.12
6	2.48	18	2.10
7	2.37	20	2.09
8	2.31	22	2.07
9	2.26	24	2.06
10	2.23	26	2.06
11	2.20	28	2.05
12	2.18	30	2.04
13	2.16	40	2.02
14	2.15	60	2.00
		120	1.98

You must now compare your calculated *t*-statistic to the appropriate value in the significance table. Find the table value beside the appropriate degrees of freedom and enter it below.

Critical value: ________

Calculated *t*-statistic: ________

If the *t*-statistic you calculated is *less than* the critical value in the table above (for the correct degrees of freedom), then the difference between the two means is *nonsignificant.* This indicates that there is statistically *no difference* between the Group 1 and Group 2 data.

If the calculated *t*-statistic is *greater than* the critical value in the table above (for the correct degrees of freedom), then the difference between the two means is statistically *significant.* This indicates that *there is a statistical difference* between the Group 1 and Group 2 data.

Check the appropriate box:

Significant ☐ Nonsignificant ☐