

Statistics Notes

Weighted comparison of means

J Martin Bland, Sally M Kerry

In a recent Statistics Note¹ we referred to a weighted two sample *t* test. Here we describe how it is done. The data were the percentage of requests from general practitioners for *x* ray examinations which were judged appropriate (table 1), where general practitioners had been randomised to intervention or control groups.²

If we compare the two sets of percentages by the usual two sample *t* method, each observation (practice) has an equal impact on the result. As some practices contributed fewer requests than others, we wish these practices to have a lesser effect on the estimate of the difference. We can do this by weighting the practices by the number of requests.

To calculate the mean percentage in each group, we simply add the observations together and divide by the number of observations. To calculate the weighted mean, we multiply each observation by the weight, add, then divide by the sum of the weights:

$$\frac{20 \times 100 + 7 \times 100 + 16 \times 94 + \dots + 18 \times 56}{20 + 7 + 16 + \dots + 18} = \frac{34\,105}{429} = 79.50$$

If the weights are all the same this gives the usual, unweighted, mean. Note that the weighted mean, 79.50, is not the same as the unweighted mean in the table, 81.6. There is a slight (but not significant) tendency for general practitioners who make more referrals to have a lower proportion conforming to the guidelines, which explains this. For the second group the weighted mean is $51050/704 = 72.51$.

The weighted standard deviation is found in a similar way. Firstly, we need a weighted sum of squares. To calculate an unweighted sum of squares about the mean, we square and add the observations, then we subtract a correction term, the number of observations times the mean squared. Here we calculate the weighted sum of the observations squared, then subtract the number of observations times the weighted mean squared. For the first group, the weighted sum of the observations squared is:

$$\frac{20 \times 100^2 + 7 \times 100^2 + 16 \times 94^2 + \dots + 18 \times 56^2}{(20 + 7 + 16 + \dots + 18)/17} \\ = \frac{2755709}{429/17} = 109\,200.59$$

To get the weighted mean we divided by the sum of the weights; to get a weighted sum we divide by the mean of the weights.) To get the sum of squares about the mean we subtract the correction term, $17 \times 79.50^2 = 107\,444.25$, giving $109\,200.59 - 107\,444.25 = 1756.34$. Dividing this by the degrees of freedom, $17 - 1 = 16$, gives the weighted estimate of the variance, $1756.34/16 = 109.77$, and the square root is the standard deviation, $\sqrt{109.77} = 10.48$. For the second group, the weighted sum of the observations squared is $3\,751\,934/(704/17) = 90\,600.68$ and the sum of squares about the mean is $90\,600.68 - 17 \times 72.51^2 =$

Table 1 Number of requests conforming to guidelines for each practice in the intervention and control groups

Practices	Intervention group		Control group	
	No of requests	% Conforming	No of requests	% Conforming
1	20	100	7	100
2	7	100	37	89
3	16	94	38	84
4	31	90	28	82
5	20	90	20	80
6	24	88	19	79
7	7	86	9	78
8	6	83	25	76
9	30	83	120	75
10	66	80	88	73
11	5	80	22	68
12	43	77	76	68
13	43	74	21	67
14	23	70	126	66
15	64	69	22	64
16	6	67	34	62
17	18	56	10	40
Total	429		702	
Mean (SD)		81.6 (11.9)		73.6 (13.1)

1219.78. Hence the estimated variance is $1219.78/(17-1) = 76.24$ and the standard deviation $\sqrt{76.24} = 8.73$.

We find the pooled sum of squares by adding the sums of squares within the two groups, $1756.34 + 1219.78 = 2976.12$ and the common variance estimate for the two groups by dividing by the combined degrees of freedom, $2976.12/(17 + 17 - 2) = 93.00$. We can now use the weighted estimates of the means and common variance in the usual two sample *t* formulas. The standard error of the difference between the means is $\sqrt{93.00 \times (1/17 + 1/17)} = 3.31$ and the difference is $79.50 - 72.51 = 6.99$. The 95% confidence interval is therefore $6.99 - 2.04 \times 3.31 = 0.2$ to $6.99 + 2.04 \times 3.31 = 13.7$. The test of significance is $t = 6.99/3.31 = 2.11$. With 32 degrees of freedom this gives $P = 0.04$.

For comparison, the unweighted difference is 8.00 and the pooled variance estimate is 157.81. The standard error of the difference is $\sqrt{157.81 \times (1/17 + 1/17)} = 4.3$. The 95% confidence interval is -1 to 17 percentage points. In this example the number of requests varies greatly between practices. This must lead to some deviation from the assumptions of the *t* test, as the variance will not be uniform. The weighted analysis meets the assumptions better and produces a worthwhile reduction in the size of the confidence interval. Some statistical software will do these calculations very simply. The same basic principle is used in meta-analysis to combine studies of varying size.

- 1 Kerry SM, Bland JM. Statistics Notes: analysis of a trial randomised in clusters. *BMJ* 1997;316:54.
- 2 Oakshott P, Kerry SM, Williams JE. Randomised controlled trial of the effect of the Royal College of Radiologists' guidelines on general practitioners' referral for radiographic examination. *Br J Gen Pract* 1994;44:197-200.

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
J Martin Bland, professor of medical statistics

Division of General Practice and Primary Care, St George's Hospital Medical School, London SW17 0RE
Sally M Kerry, lecturer in medical statistics

Correspondence to: Professor Bland.

BMJ 1998;316:129