Comparing Two Population Means: Independent Sampling

(*See pp.352-364*)

Many experiments involve a comparison of two or more population means. Recall our earlier comparison of two methods of teaching reading. We compared a new method of teaching reading to the current standard method. Our comparison was based on the results of a reading test given at the end of a learning period. Of a random sample of 22 learners, 10 were taught by the new method and 12 were taught by the standard method. All 22 learners were taught by qualified instructors under similar conditions. The results we considered are shown in the table below. (Figure 1) (The data is in the CD file **READING**.) We also considered a box plot of the data. (Figure 2)

C3	C4
NewScore	StandardScore
80	79
76	73
70	72
80	62
66	76
85	68
79	70
71	86
81	75
76	68
	73
	66

Figure 1

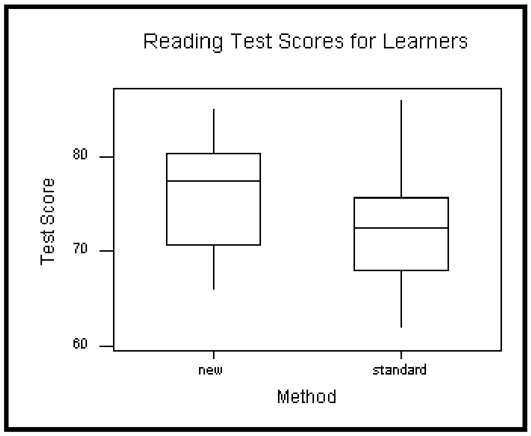

Figure 2

We found that the mean score on the post test for those taught by the new method was μ_1 = 76.4 and the mean score on the post test for those taught by the standard method was μ_2 = 72.33. We wonder if the difference in the two means is sufficient to allow us to make a valid comparison of the relative effectiveness of the two methods.

In this case we are considering two population means with small (n < 30) samples; so we will use the t-distribution. ***To use the t-distribution, both sampled populations must be approximately normal with equal variances, and the random samples must be selected independently of each other.***

Our hypotheses are as follows:

H_0: $\mu_1 - \mu_2 = 0$

H_a: $\mu_1 - \mu_2 \neq 0$

First we use our data to estimate the true mean difference between the test scores of the two methods. We construct a 95% **(α = 0.05)** confidence interval for ($\mu_1 - \mu_2$). MINITAB output is shown below. (p. 428)

Two sample T for SCORE

METHOD	N	Mean	StDev	SE Mean
NEW	10	76.40	5.83	1.8
STD	12	72.33	6.34	1.8

95% CI for mu (NEW) - mu (STD): (-1.4, 9.5)

T-Test mu (NEW) = mu (STD) (vs not =): T = 1.55 P = 0.14 DF = 20 ($n_1 + n_2 - 2$)

Both use Pooled StDev = 6.12

We can estimate with (with 95% confidence) the mean test score for the new method is anywhere between 1.4 points less than, to 9.5 points more than, the mean test score for the standard method. There is insufficient evidence to indicate that $\mu_1 - \mu_2 \neq 0$ because the 95% CI includes 0 as a possible value. That is, we cannot reject the null hypothesis that $\mu_1 - \mu_2 = 0$ because the 95% CI includes 0 as a possible value.

Our test of hypothesis proceeds as follows. We use a two-tailed test in this case.

Our rejection region: $t < -t_{\alpha/2} = -t_{0.025} = -2.086$ or $t > t_{\alpha/2} = t_{0.025} = 2.086$ based on 20 df.

Since the observed value of t (1.55) is not in the rejection region we do not reject H_0. As above, we cannot reject the premise that the two populations have the same mean. (At the 0.05 level, we cannot conclude that the new method is better.)

TI Graphing Calculator Hypothesis Test for (μ_1 - μ_2) (pp. 362-363)

STAT, **EDIT**, Place data in **L_1** and **L_2**. **STAT**, **TESTS**, **2-Samp TTest**. (Figure 3) Arrow down n to "**Calculate**" and press **ENTER**. (Figure 4)

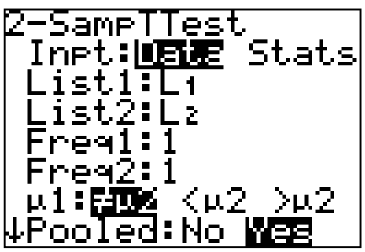

Figure 3

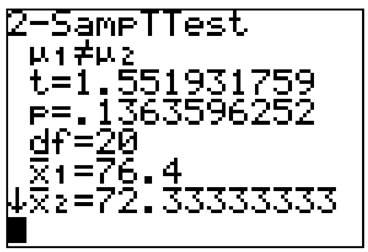

Figure 4

Recall that *to use the t-distribution, both sampled populations must be approximately normal with equal variances, and the random samples must be selected independently of each other.* Are these conditions satisfied?

In observing the box plots, the variances appear to be similar. An examination of the normal probability plots suggests that the distributions are approximately normal.

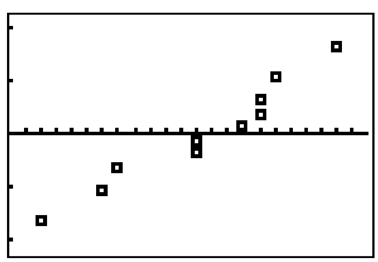

Figure 2. New Method

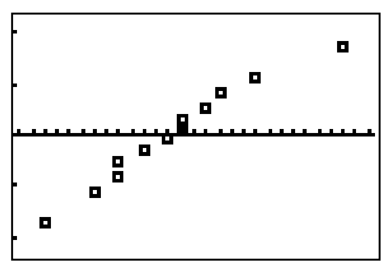

Figure 3. Standard Method