(2) **Comparing Two Populations 2**

Suppose you wish to compare a new method of teaching reading to slow learners with the current standard method. You decide to base your comparison on the results of a reading test given at the end of a learning period of six months. Of a random sample of 22 slow learners, 10 are taught by the new method and 12 are taught by the standard method. All 22 children are taught by qualified instructors under similar conditions for the designated six-month period. The results of the reading test at the end of this period are given below.

New Method: 80, 76, 70, 80, 66, 85, 79, 71, 81, 76.

Standard Method: 79, 73, 72, 62, 76, 68, 70, 86, 75, 68, 73, 66.

(a) Use the data in the table to estimate the true mean difference between the test scores for the new method and the standard method. Use a 95% confidence interval.

(b) Interpret the interval you found in the previous part.

(c) What assumptions must be made in order that the estimate be valid? Are they reasonably satisfied?

a) Since the sample sizes $n_1 = 10 < 30$ and $n_2 = 12 < 30$ are small, we use the t-test from slides 33/34 of lecture 11.

The variances are $s_1^2 = 34.04$ and $s_2^2 = 40.24$, hence we rather don't assume their equality. The 95% confidence interval is given by $[-1.36, 9.49)$

```
> meth1 <- c(80, 76, 70, 80, 66, 85, 79, 71, 81, 76)
> meth2 <- c(79, 73, 72, 62, 76, 68, 70, 86, 75, 68, 73, 66)
>
> t.test(meth1, meth2)

        Welch Two Sample t-test

data:  meth1 and meth2
t = 1.5643, df = 19.769, p-value = 0.1336
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.360091  9.493424
sample estimates:
mean of x mean of y
 76.40000  72.33333
```

b) Since the interval contains 0, we fail to reject the hypothesis $H_0: \mu_1 - \mu_2 = 0$, when compared against $H_1: \mu_1 - \mu_2 \neq 0$ if we choose $\alpha = 1/20$

We can interpret the confidence interval as follows. If we take k samples like the one we just analysed, then we expect that the true parameter (which we estimated here by the mean) lies within the confidence interval 0.95 * k times.

c) We assumed that the samples were independent and that they are normally distributed. Looking at their histograms, it seems fairly reasonable to assume normal distribution.