

General Method: Difference of Means

1. Calculate $\bar{x}_1$, $\bar{x}_2$, SE_1 , SE_2 .
2. Combined $SE = \sqrt{SE_1^2 + SE_2^2}$.
ASSUMES INDEPENDENT SAMPLES.
3. Calculate df : either Welch-Satterthwaite formula or simpler $df = \min(n_1, n_2) - 1$.
4. For a hypothesis test, $t = \frac{\bar{x}_1 - \bar{x}_2}{SE}$; convert to a P-value using table of t statistics
5. For a confidence interval, calculate critical value t^* corresponding to desired confidence level (e.g. in our example we had $df = 13$, confidence level 95%, led to $t^* = 2.16$). Then the confidence interval is

$$\bar{x}_2 - \bar{x}_1 \pm t^* \times SE.$$

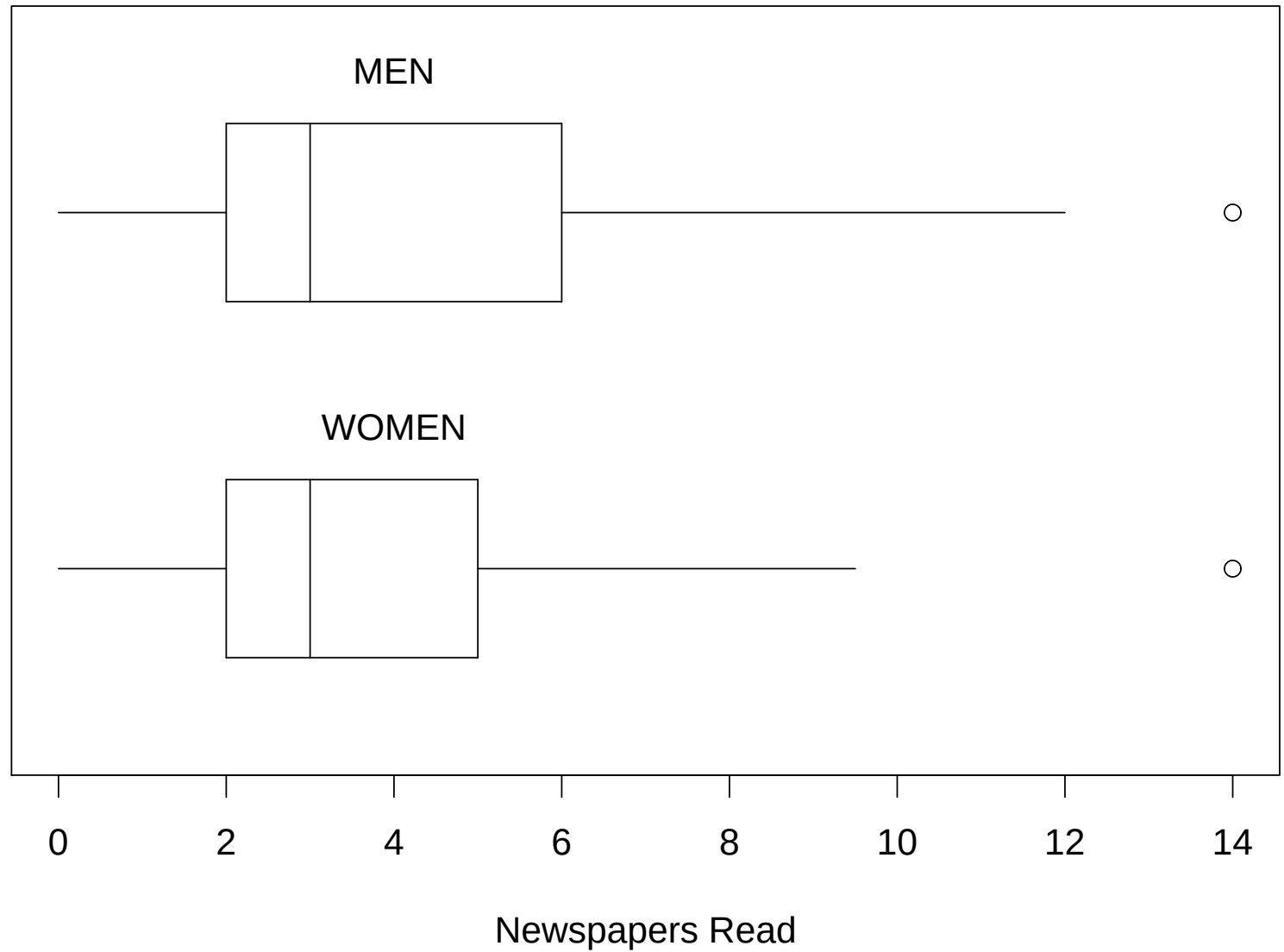
Example (Question 10.28, page 493)

Following are the numbers of newspapers read by a sample of women and of men

Women: 5,3,6,3,7,1,1,3,0,4,7,2,2,7,3,0,5,0,4,4,5,14,3,1,2,1,7,2,5,3,7

Men: 0,3,7,4,3,2,1,12,1,6,2,2,7,7,5,3,14,3,7,6,5,5,2,3,5,5,2,3,3

- (a) Construct and interpret a plot comparing responses of males and females
- (b) Construct and interpret a 95% confidence intervals comparing populations means
- (c) Show the five steps of a significance test comparing populations means
- (d) State and check the assumptions



Box plot for number of newspapers read by women and by men.

- (a) See boxplots; male and female distributions look very similar
- (b) With suffix 1 indicating women, 2 for men: $\bar{x}_1 = 3.774$, $\bar{x}_2 = 4.414$, $s_1 = 2.929$, $s_2 = 3.100$, $n_1 = 31$, $n_2 = 29$, $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 0.78$. We have $df = 57.1$ according to Welch-Satterthwaite formula, $df = 28$ by simpler formula. Based on $df = 28$, the critical value of t for a 95% confidence interval is $t^* = 2.048$, so the 95% confidence interval for $\mu_2 - \mu_1$ is $4.414 - 3.774 \pm 2.048 \times 0.78 = (-0.957, 2.237)$.
- (c) The t statistic for a hypothesis test is $\frac{\bar{x}_2 - \bar{x}_1}{SE} = 0.82$; this is well below the critical value for a test at significance level .05 (the critical value is $t^* = 2.048$, as in part 2) so we DO NOT REJECT the null hypothesis that the means for men and women are equal.

(d) The assumptions require independence of the two samples (probably more or less correct); randomness of the two samples (depends on how the samples were obtained); and approximately normal distributions for the samples themselves (probably true to a reasonable approximation).

Paired Comparison Tests

Consider the following dataset, based on the midterm and final exam scores of a recent course of mine (not STOR 151!):

Student	1	2	3	4	5	6	7	8	9	10
Midterm	86	100	90	90	94	73	76	76	95	87
Final	84	95	77	83	70	76	54	81	90	84
Difference	2	5	13	7	24	-3	22	-5	5	3

Mean midterm score=86.7; Mean final score=79.4; Difference 7.3

Is this significant evidence of a difference?

We could apply the same test as previously, which would lead to

$$\bar{x}_1 = 86.7,$$

$$\bar{x}_2 = 79.4,$$

$$s_1 = 9.06,$$

$$s_2 = 11.37,$$

$$SE = \sqrt{\frac{9.06^2}{10} + \frac{11.37^2}{10}} = 4.60,$$

$$t = \frac{86.7 - 79.4}{4.60} = 1.59,$$

$$df = 9.$$

The two-sided P-value is 0.15 — this is greater than 0.05, therefore the result is *not significant*.

However, there is an error in this calculation....

The two samples are *not independent* since they represent scores from the same students.

Instead, we apply a *matched pairs* test:

1. Calculate the difference in scores for each student.
2. Carry out a single-sample test for the null hypothesis that the mean difference is 0.

Details

The 10 differences 2, 5, 13, 7, 24, -3, 22, -5, 5, 3 have a mean $\bar{x} = 7.3$ and a standard deviation $s = 9.67$.

The standard error is $\frac{9.67}{\sqrt{10}} = 3.06$.

The t statistic is $\frac{7.3}{3.06} = 2.39$.

The corresponding two-sided P-value is 0.04.

Therefore, the result is statistically significant

Message of this example:

The standard two-sample test is valid only when the two samples are *independent* (along with other assumptions: quantitative variables, random sampling, approximately normal distributions)

When the observations are directly paired (e.g. two exam scores from the same student, two medical results from the same patient, etc.) it is possible to apply a *paired comparison test* instead.

The main difference is a different method of computing the standard error. This has implications for both hypothesis testing and confidence interval calculations.

Example: Problem 10.47, page 511/512.

In a study to determine whether exercise reduced blood pressure, a sample of three patients was tested, with the following results:

Subject	Before	After
1	150	130
2	165	140
3	135	120

- (a) Explain why the three “before” observations and the three “after” observations are dependent samples.
- (b) Find the sample mean of the before scores, sample mean of the after scores, and the sample mean of $d = \text{before} - \text{after}$.
- (c) Find a 95% confidence interval for the mean difference, and interpret it.

- (a) They are from the same subjects; it is a “matched pairs” design.
- (b) $\frac{150+165+135}{3} = 150$; $\frac{130+140+120}{3} = 130$; difference 20.
- (c) The three differences 20, 25, 15 have mean 20 and standard deviation $\sqrt{\frac{5^2+0+5^2}{2}} = 5$. The standard error is $\frac{5}{\sqrt{3}} = 2.887$. Based on the t table with $df = 2$, the critical value of t^* for a 95% confidence interval is 4.303. Therefore, a 95% confidence interval for the difference is $20 \pm 4.303 \times 2.887 = (7.58, 32.42)$. This does not contain 0, so despite the very small sample size, the difference is statistically significant.