

## Chapter 8 Class Notes – Comparison of Paired Samples

In this chapter, we consider the analysis of paired data. To illustrate, (in the spirit of p.332 ex.8.S.5) an agronomist randomly selected six wheat plants from a field and for each plant took a batch of seeds in the central portion and another batch of seeds in the top portion of the plant. She then measured the moisture content in the batches of seeds, and her goal is to find the 95% CI for the difference of the means. Here are the (modified) data:

Plant	Center	Top	$y_C - y_T$
1	62.7	60.7	2.0
2	63.6	61.6	2.0
3	60.9	59.3	1.6
4	63.0	63.5	-0.5
5	62.7	60.6	2.1
6	63.7	62.4	1.3
Mean	62.77	61.35	1.417
SD	1.011	1.482	0.987

In Chapter 7, the measurements in the two ‘groups’ were independent, but now they are paired: here, we get a pair  $(y_C, y_T)$  for each plant.

The **correct analysis** of these paired data is to look at the differences. Let  $d = y_C - y_T$ . The average of these sample differences is denoted  $\bar{d}$ , and it turns out that  $\bar{d} = \bar{y}_C - \bar{y}_T$  (since the average of the differences is the difference of the averages). Also,  $\bar{d}$  is a point estimated of  $\mu_D = \mu_C - \mu_T$ . Defining these in this example,

- \*  $\mu_C$  is the average **center** moisture content of **all** wheat plants
- \*  $\mu_T$  is the average **top** moisture content of **all** wheat plants

The sample **SD** of the differences is denoted  $s_D$  – it is easy to compute from the above list of differences; the sample **SE** of the difference is  $SE_{\bar{D}} = \frac{s_D}{\sqrt{n_D}}$  (here,  $n_D$  is the number of differences – number of plants in the above example).

→ Important: Note that  $s_D$  is not equal to  $s_1 - s_2$ . ←

For the above Wheat data,  $\bar{d} = 1.417$ ,  $s_D = 0.987$ ,  $SE_{\bar{D}} = \frac{0.987}{\sqrt{6}} = 0.403$

**Finding the 95% CI** for  $\mu_D = \mu_C - \mu_T$  is simple using the methods of Chapter 6 – it is  $\bar{d} \pm t_{0.025} SE_{\bar{D}}$ . In this expression, the t-statistic corresponds to **df =  $n_D - 1$  degrees of freedom**, and it is simple to extend this methodology to CIs with e.g. 90% or 99% coverage.

For the Wheat data, we obtain the 95% CI for  $\mu_D$ :

$$1.417 \pm 2.571 \times 0.403 \rightarrow 1.417 \pm 1.036 \rightarrow \underline{\underline{(0.381, 2.453)}}$$

**Conclusion:** We are 95% confident that the average moisture content in the center of wheat plants exceeds the average moisture content at the top of wheat plants by at least 0.381% and at most 2.451%.

If instead of a **CI**, we were interested in a **HT** (**hypothesis test**) – and testing for a ‘difference’ and  $\alpha = 5\%$  – we would proceed as follows:

$$H_0: \mu_C = \mu_T \text{ (same as } \mu_D = 0 \text{)}$$

$$H_A: \mu_C \neq \mu_T \text{ (same as } \mu_D \neq 0 \text{)}$$

The test statistic (TS) is  $t_s = \frac{\bar{d} - 0}{SE_{\bar{D}}}$ . The ‘0’ is only added here to

‘restate’  $H_0$  and is not needed. For the Wheat study, the test statistic

is  $t_s = 1.417 / 0.403 = 3.52$ . From Table 4, it follows that the one-tailed area is between 0.005 and 0.01, so then the p-value is between 0.01 and 0.02. Since  $p < \alpha$ , we reject  $H_0$  and **conclude that this study suggests that the average moisture contents do indeed differ for the two regions of the wheat plant head.**

It is important to point out the [assumptions](#) here are that:

- the wheat plants have been randomly chosen and are therefore representative of the wheat plant population
- the differences come from a Normal distribution.

On pp.306-307, the authors point out that we only assume that the differences follow a Normal distribution – i.e., not the separate measurements ( $y_C$  and  $y_T$  in this example).

As an aside, had we performed an *incorrect* (i.e., two-independent-sample t-test) *analysis* of the Wheat data, we would obtain:

$$SE_{(\bar{Y}_1 - \bar{Y}_2)} = \sqrt{\frac{1.011^2}{6} + \frac{1.482^2}{6}} = 0.732 \text{ and } t_s = \frac{62.77 - 61.35}{0.732} = 1.94,$$

df = 8, and a non-significant p-value ( $p = 0.089$ ). The reason for this change is in the calculation of the standard error in the denominator of the test statistic (since the numerators are the same). The correct test statistic uses the denominator of **0.403** instead of the incorrect **0.732**. The reason for this 45% reduction from 0.732 to 0.403 is in taking account of the pairing, and this shows the increased efficiency in using the paired design (RCBD) over the independent group design (CRD). Of course, we are not advocating here that the incorrect analysis be performed – we are just demonstrating the increase in precision; see the discussion on p.305.

§8.3: The Paired Design is used often and in several contexts:

- RCBD's (eggplant illustration in a greenhouse)
- Observational studies (matched pair design and case-control studies)
- Repeated measurements (before and after study)
- Blocking by time (where a block = day)

One pairs to reduce bias and to increase precision.

### First Nonparametric Alternative (N/A) – the Sign Test (Section 8.4)

“A blunt but handy tool...” Examine the column of differences, ignore the ties, count the number of + signs (denoted  $N_+$ ), the number of – signs (denoted  $N_-$ ), the test statistic ( $B_s$ ) is the larger of  $N_+$  and  $N_-$ , and the p-value using the Binomial distribution (or Table 7, p.625).

To illustrate, consider the Valproate data on p.320 ex.8.4.5. This is a cross-over study involving Valproate and Placebo and the response variable is the number of minor epileptic seizures. We are asked to “test for efficacy of Valproate using the sign test at  $\alpha = 0.05$ ” which means that the alternative is one-tailed: when we form the differences (Placebo – Valproate), the alternative states that this distribution exceeds zero. The number of zeros is one; of the remaining 19,  $N_+ = 14$ ,  $N_- = 5$ , so  $B_s = 14$ , so the one-tailed p-value is  $0.064/2 = 0.032$  (from Table 7 on p.625). The exact p-value is  ${}_{19}C_{14}(\frac{1}{2})^{14}(\frac{1}{2})^5 + {}_{19}C_{15}(\frac{1}{2})^{15}(\frac{1}{2})^4 + \dots + {}_{19}C_{19}(\frac{1}{2})^{19}(\frac{1}{2})^0 = 0.031784$ . We **conclude that Valproate appears effective in reducing the rate of epileptic seizures ( $p = 0.032$ ).**

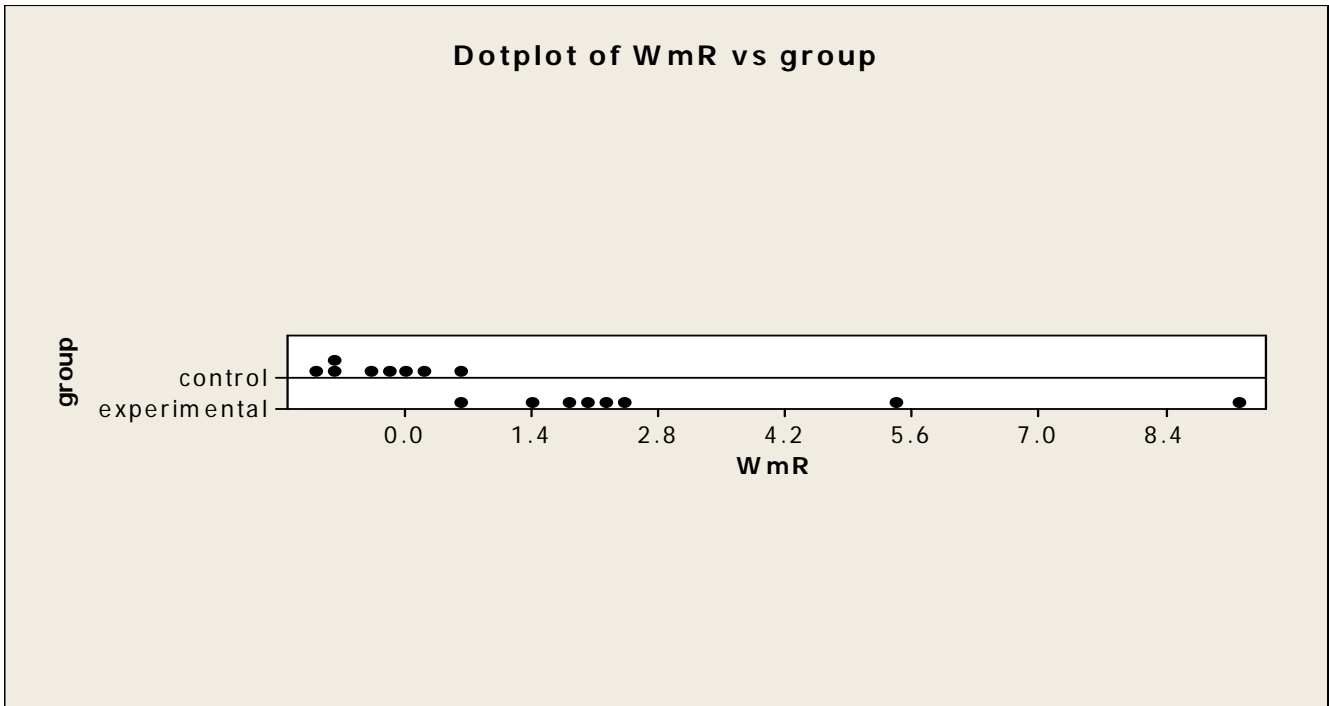
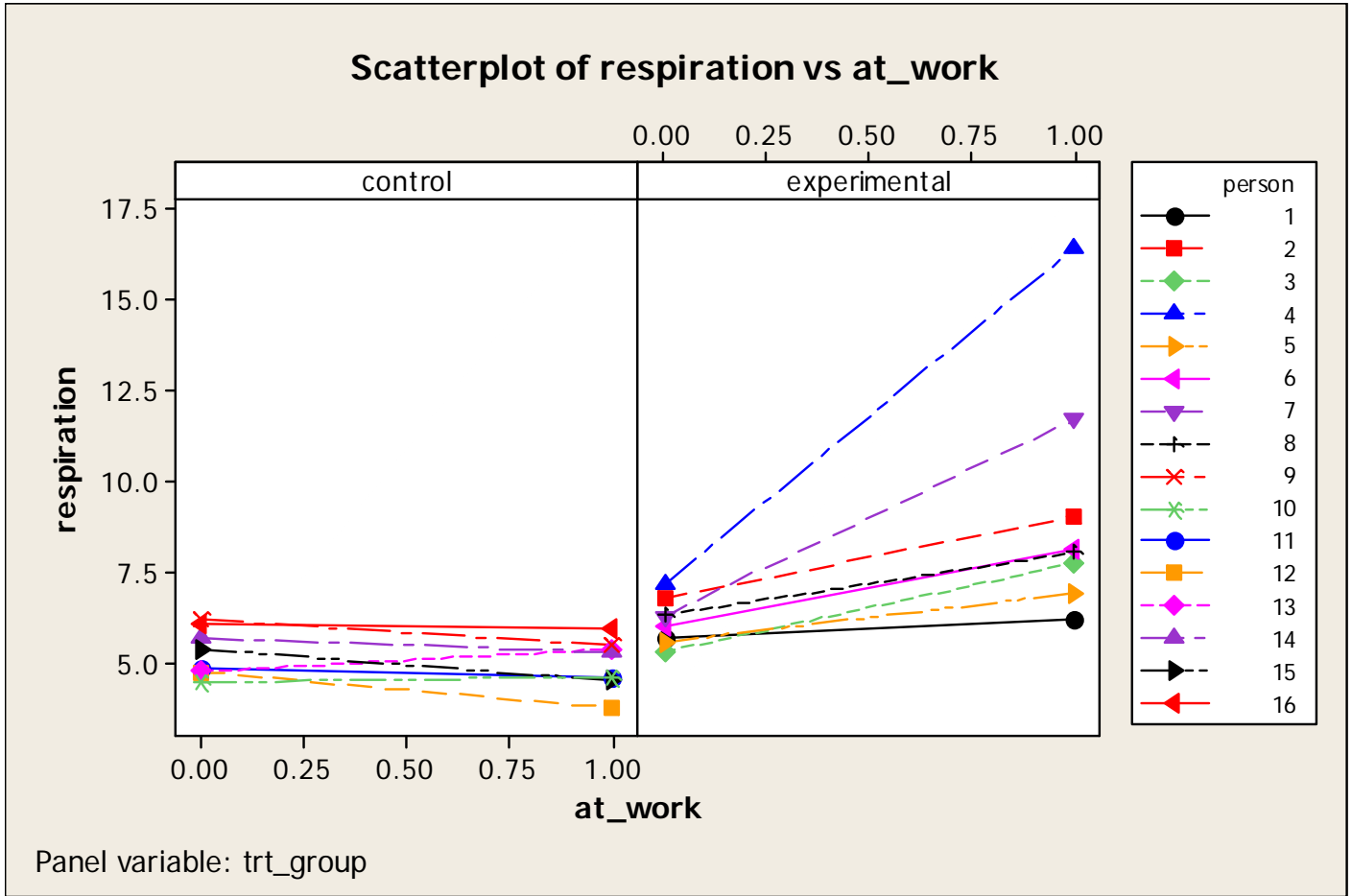
**Second N/A – the Wilcoxon Signed-Rank Test (Section 8.5)**

The Signed-Rank test is equally simple, but is more powerful than the above Sign test. We rank the absolute value of the differences and attach the sign of the difference, add the signs of the negatives ( $W_-$ ), the signs of the positives ( $W_+$ ), and the test statistic ( $W_s$ ) is the larger of these two. For the Valproate data, the test statistic is  $W_s = 166$ , and Minitab reports the p-value is 0.002; from Table 8 on p.626, the p-value is between  $0.0017/2 = 0.00085$  and  $0.0094/2 = 0.0047$ . Our conclusion here is as above. Note that the p-value here is less than for the Sign test since this Signed Rank test is more powerful.

In **Section 8.6**, the authors address comparisons of two treatments or groups in before-and-after studies, where methods of Chapter 7 are applied to the differences of before and after measurements. To illustrate, the full hypnosis/ventilation on p.334 #8.S.20:

Experimental Group					Control Group			
Subject	Rest	Work	W-R		Subject	Rest	Work	W-R
1	5.74	6.24	0.50		9	6.21	5.50	-0.71
2	6.79	9.07	2.28		10	4.50	4.64	0.14
3	5.32	7.77	2.45		11	4.86	4.61	-0.25
4	7.18	16.46	9.28		12	4.78	3.78	-1.00
5	5.60	6.95	1.35		13	4.79	5.41	0.62
6	6.06	8.14	2.08		14	5.70	5.32	-0.38
7	6.32	11.72	5.40		15	5.41	4.54	-0.87
8	6.34	8.06	1.72		16	6.08	5.98	-0.10

# Introduction to Biostatistics



### Two-Sample T-Test and CI: WmR, group

Two-sample T for WmR

group	N	Mean	StDev	SE Mean
control	8	-0.319	0.544	0.19
experimental	8	3.13	2.86	1.0

Difference = mu (control) - mu (experimental)

Estimate for difference: -3.45125

95% CI for difference: (-5.88671, -1.01579)

T-Test of difference = 0 (vs not =): T-Value = -3.35 P-Value = 0.012 DF = 7

### Mann-Whitney Test and CI: expdiff, contdiff

	N	Median
expdiff	8	2.180
contdiff	8	-0.315

Point estimate for ETA1-ETA2 is 2.495

95.9 Percent CI for ETA1-ETA2 is (1.501,5.651)

W = 99.0

Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0014

**Given the above Dotplot, we would prefer the WMW test here; our conclusion is that these data suggest a difference between the ventilation rate differences for the two groups ( $p = 0.0014$ ).**