Exercise 27.42 of PSLS3e

Data: Vaccine antibody responses in rhesus monkeys against a particular virus after vaccination with four different types of vaccine and placebo (sham). The outcome measured is given as V2-specific antibody productions measured by surface plasmon resonance. The trial included eight monkeys per vaccine group.

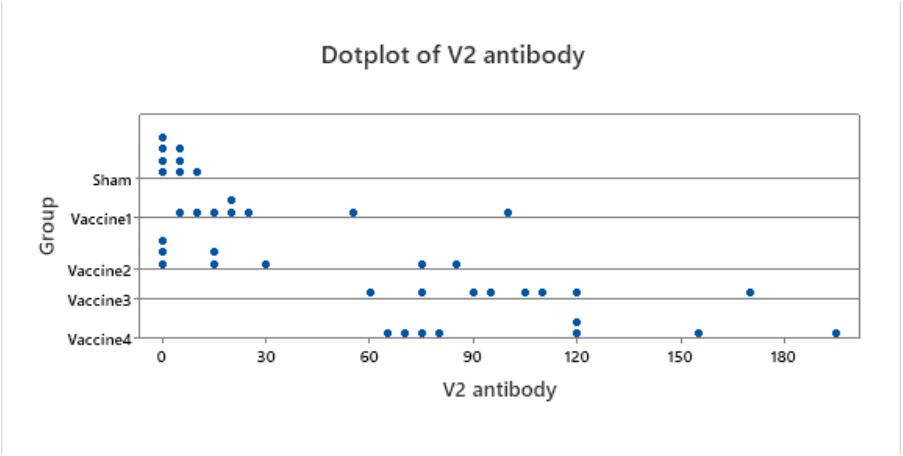
Model: If we denote by X_{ij} the V2 antibody response of the jth animal in vaccine group i (where $i = 1, \ldots, 5$ and $j = 1, \ldots, 8$), the statistical model is that the X_{ij}'s for each group i constitute a simple random sample from a suitable distribution (alternatively, we could say that the 8 values for each group are i.i.d.). More discussion of model assumptions will follow below.

(a) For descriptive statistics, refer to the Minitab table and the dotplots below.

```
Describe 'V2 antibody';
 By 'Group'.
Dotplot ( 'V2 antibody' ) * 'Group'.
```

Statistics

Variable	Group	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Skewness
V2 antibody	Sham	8	3.83	3.73	0.10	1.07	2.85	5.92	11.40	1.29
	Vaccine1	8	31.1	31.6	4.3	11.7	20.0	46.8	100.1	1.84
	Vaccine2	8	27.6	33.1	0.6	0.8	14.9	62.8	83.5	1.08
	Vaccine3	8	103.0	33.5	60.0	79.9	98.3	117.0	171.6	1.15
	Vaccine4	8	109.5	47.2	65.5	69.9	98.7	145.8	196.1	0.93



Comments:
The distributions appear as very different, not only in their center but also in their spread, with the Sham group tightly clustered close to zero and the other distributions more spread out. Vaccines 3 and 4 appear to have triggered the strongest antibody responses. All distributions show some right-skewness, with skewness values ranging from 0.93 to 1.84. Normality tests are not too informative here with the small group sizes, because a non-significant test can reflect the low sample size rather than a distribution being close to normal. Due to the right-skewness, most or all of the distributions appear as non-normal. In response to the question, the data violate the assumptions of ANOVA in two ways: the non-normality within groups and the different within-group standard deviations (very clearly violating the guidelines of ratios being less than 2). Two obvious ways to proceed are by a non-parametric analysis or by attempting to transform the data to a better scale for ANOVA.

(b) For Kruskal-Wallis test we have the option of assuming the distributions to be of equal shape (the "Δ-assumption"), in order to phrase our inference in terms of the medians. That assumption is however not suitable here because the distributions have very different spread; the "Δ-assumption" requires the distributions to differ only in their position. Without that assumption, the hypotheses are simply,

H_0 : equal distributions in the five vaccine groups,

H_a : not H_0, that is, some differences exist between the distributions.

The alternative hypothesis can be worded a bit stronger as: "some distributions are systematically larger than others", however this interpretation is most helpful when comparing only two distributions. The Kruskal-Wallis results are obtained from Minitab, as shown below.

```
Kruskal-Wallis 'V2 antibody' 'Group'.
```

EX27_042.MTW

Kruskal-Wallis Test: V2 antibody versus Group

Descriptive Statistics

Group	N	Median	Mean Rank	Z-Value
Sham	8	2.85	7.4	-3.53
Vaccine1	8	20.00	17.8	-0.74
Vaccine2	8	14.90	14.4	-1.64
Vaccine3	8	98.35	31.4	2.94
Vaccine4	8	98.65	31.5	2.98
Overall	40		20.5	

Test

Null hypothesis H_0: All medians are equal
Alternative hypothesis H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	4	26.59	0.000
Adjusted for ties	4	26.59	0.000

Comments:
The Kruskal-Wallis test is strongly significant ($P<0.0005$), so we conclude that the five distributions are not equal. In view of the large differences we saw between the distributions descriptively, that is not surprising and indeed not too helpful. It also does not answer the question asked about whether some treatments (i.e., vaccines) have systematically larger antibody responses than others because the result could be due to the Sham group only. This naturally leads to a wish to proceed with pairwise comparisons.

Before doing that we may want to consider the role of the Sham group in the experiment. It may not be real interest to compare the vaccine groups with this group, so that the only reason for including it is to provide evidence that the measurements were not flawed in some way. Having already noted that the Sham group gives the lowest values of all groups, we can probably conclude that there was no such flaw. If at that point there is no further interest in the Sham group, it could be considered to exclude it from further analysis. In an ordinary ANOVA, all groups contribute to the pooled standard deviation, and therefore such a decision will also affect the analysis for the other groups; however, in a rank-based analysis there is no loss of power for the remaining groups by omitting a group. The Kruskal-Wallis test for the reduced

dataset with only the four vaccines included gives: $H = 17.59$ and $P = 0.001$. We conclude that there are some differences between the four vaccine variants in their antibody response.

The Kruskal-Wallis test does not have as well-established procedures for pairwise group comparisons as ordinary ANOVA. One reason for this is that we are comparing entire distributions instead of single parameters (for the ANOVA, the means). One sensible method to carry out pairwise comparisons between all (or a selected subset of) group pairs by Wilcoxon-Mann-Whitney two-sample tests and use a Bonferroni correction for multiple comparisons; this approach is described in Exercise 15.47 in Chapter 15 of IPS (available at Moodle).

In order to carry out this analysis in Minitab, one needs to unstack the data for the vaccine groups into separate columns and then use the Mann-Whitney menu repeatedly. Below the results are shown for one of those comparisons, and all results are summarized in tabular form. For two groups of size 8, the total rank sum of their combined sample is $16 \cdot 17/2 = 136$, and hence the expected rank sum for any of the groups under the null hypothesis equals $136/2 = 68$; the W-values are assessed in terms of how far are away from this value.

```
Mann-Whitney 95.0 'V2 antibody_Vaccine1' 'V2 antibody_Vaccine2';
 Alternative 0.
```

GROUP LESS THAN VACCINE1 EXCLUDED

Mann-Whitney: V2 antibody_Vaccine1, V2 antibody_Vaccine2

Method

η_1: median of V2 antibody_Vaccine1
η_2: median of V2 antibody_Vaccine2
Difference: $\eta_1 - \eta_2$

Descriptive Statistics

	Sample N	Median
V2 antibody_Vaccine1	8	20.0
V2 antibody_Vaccine2	8	14.9

Estimation for Difference

Difference	CI for Difference	Achieved Confidence
7.2	(-29.6, 24.7)	95.94%

Test

Null hypothesis H_0: $\eta_1 - \eta_2 = 0$
Alternative hypothesis H_1: $\eta_1 - \eta_2 \neq 0$

Method	W-Value	P-Value
Not adjusted for ties	74.00	0.564
Adjusted for ties	74.00	0.563

P / W	Vaccine 1	Vaccine 2	Vaccine 3	Vaccine 4
Vaccine 1	-	74	40	40
Vaccine 2	0.563	-	97	94
Vaccine 3	0.004	0.003	-	66
Vaccine 4	0.004	0.007	0.875	-

Bonferroni correction for carrying out 6 tests corresponds to assessing each with a significance level of $0.05/6 = 0.0083$ or to multiplying each of the P-values in the table by 6 and comparing the resulting values with 0.05. The conclusion is clear: Vaccines 1 and 2 show no significant difference in their antibody response, and nor do Vaccines 3 and 4. However, any of the first two vaccines are significantly different than any of the last two vaccines. The directions we have seen in the desciptive statistics tell us that we can phrase the significant results by saying that Vaccines 3 and 4 give sysstematically larger antibody responses than Vaccines 1 and 2.

Another option for analysis of these data would be to try transform to a suitable scale for ANOVA. Square-root transformation reduces the right-skewness within the groups quite substantially, and while the Sham group still has the clearly lowest spread, the other groups are now reasonably similar in spread (one ratio of standard deviations slightly exceeds 2, but tests for equal variances are clearly non-significant). An ANOVA for the four vaccine variants yields essentially the same conclusions as the analysis above; the details are left for the interested student.