Sketch solution for selected problems in Homework #9

14.50  b. If the rank sum for one treatment is much larger than the rank sum for all other treatments, it is more likely that the distributions differ in location.

14.51  At least 5 observations/measurements in any treatment group.

14.55  Based on the table, it is easy to see that $n_1 = 12$, $n_2 = n_3 = 8$. A bit calculation shows that $R_1 = 84$, $R_2 = 145$, $R_3 = 177$ which lead to $\bar{R}_1 = 7$, $\bar{R}_2 = 18.125$, $\bar{R}_3 = 22.125$, and $\bar{R} = 14.5$. Therefore,

$$H = \frac{12}{28 \times 20} \left[ (12 \times (7 - 14.5)^2 + 8 \times (18.125 - 14.5)^2 + 8 \times (22.125 - 14.5)^2 \right] = 18.40271$$

If we set $\alpha = 0.05$, the reject region is $\{ T > 5.9915 \}$ since the 95\% quantile of $\chi^2$ with $df = 2$ is 5.9915. The observed test statistic falls in the rejection region. Thus we reject the null hypothesis.

If we only consider the first two groups, we may use Wilcoxon Rank Sum Test. It is easy to see that $T_1 = 82.5$ and $T_2 = 127.5$ with $n_1 = 12$ and $n_2 = 8$. For a two-sided test, most softwares will provide a p-value around 0.000717. If you only has the textbook, you may find the $n_1 = 12$ is out of bound for Table XII. Thus you may use normal approximation

$$Z = \frac{|82.5 - 12 \times (20 + 1)/2|}{\sqrt{12 \times 8 \times 21/12}} = -3.3561$$

Therefore, the p-value for two-sided test is approximately $2 \times \Pr(Z < -3.3561) = 0.00079$.

14.63  b. For each block, rank the data across the treatments from smallest to largest.

14.65  Data were obtained from a complete randomized block design with $b = 6$ blocks and $k = 4$ treatments. The rank sum for the treatments are $R_A = 11$, $R_B = 21$, $R_C = 21$, $R_D = 7$. The null hypothesis is that the distributions of the response under four treatments are the same. The alternative is that there is difference between these distributions. The test statistic is

$$F_r = \frac{12}{6 \times 4 \times 5} \left( 11^2 + 21^2 + 21^2 + 7^2 \right) - 3 \times 6 \times 5 = 15.2,$$

which is larger than 6.251, the 90\% quantile of the $\chi^2$ with $df = 3$. Therefore, we reject the null at $\alpha = 0.10$. If use the Table VII on page 798, the 99.5\% quantile of $\chi^2$ with $df = 3$ is 12.8381 which is less than the test statistic. Therefore, the p-value should be less than 0.5\% (The actual p-value is 0.00165). There is a typo in the textbook. See the notes for detail.

14.78  All the calculation is based on Table XIV. (a) 0.01, (b) 0.01, (c) 0.975, (d) 0.05.

14.81  Let $\rho$ be the correlation coefficient between ranks at population level. We are testing $H_0 : \rho = 0$ vs. $H_a : \rho \neq 0$. Since $n = 7$, based on Table XIV, the rejection region should be $\{ |\hat{\rho}| > 0.786 \}$ at $\alpha = 0.05$. A bit calculation shows that $\hat{\rho} = 0.7451$. Therefore, we do not reject the null at $\alpha = 0.05$. Based on the Table, we can see the actual p-value should be between 0.05 and 0.10. For the above test procedure be valid, we need to assume that the pairs of observations are independent random samples from a joint continuous distribution.