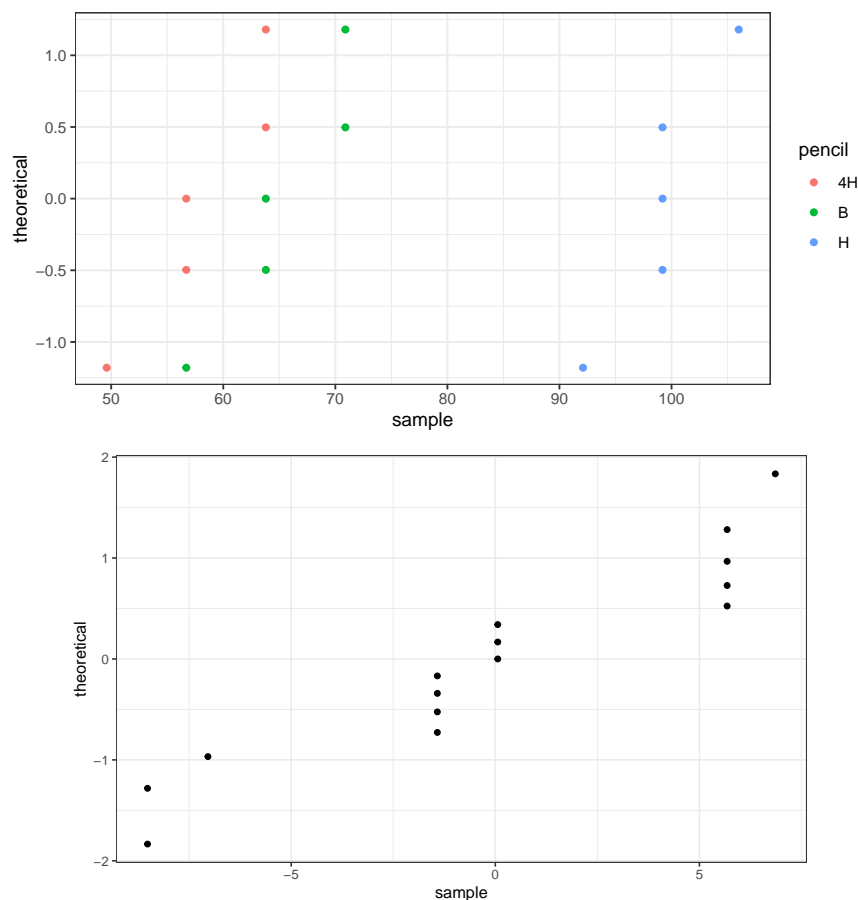# Homework 13

Due April 30, 2020 at 11:59 PM

1. P. 533: 1 (ignore (f), (g)) (14 points for (d), 6 points for others), dataset: `pencil.jmp`

   (a) Data within each sample must be iid normal, the samples must be independent, and the three distributions must have the same standard deviation.

   The three normal plots of the data are roughly linear with no outliers, providing no evidence against the normal part of the assumption. The slopes are also similar, providing no evidence against the common standard deviation assumption. The normal plot of the residuals is roughly linear (considering the number of ties). This also provides nor evidence against the normal part of the assumption.

   (b) Using equation (7.7) in the textbook, $s_P^2 = 31.57733$, and so $s_P = \sqrt{31.57733} = 5.619$, with $n - r = 15 - 3 = 12$ degrees of freedom associated with it. This measures the magnitude of baseline variation within any of the 3 conditions, assuming it is the same for all 3 conditions.

(c) Use equation (7.14). The $\pm$ part (margin of error) is the same for all three intervals because all three sample size are the same. For 95% confidence, the quantile used is $t_{n-r,1-\alpha/2} = t_{12,0.975} = 2.179$. So the margin of error is

$$t_{12,0.975}s_P\sqrt{\frac{1}{n_i}} = 2.179\frac{5.619}{\sqrt{5}} = 5.475956$$

The confidence intervals are

$$\hat{\mu}_i \pm t_{12,0.975}s_P\sqrt{\frac{1}{n_i}}, i = 4H, H, B$$

Plugging in $\hat{\mu}_i$ in the JMP output and the margin of error, we have the confidence intervals for μ_{4H}, μ_H, μ_B are

$$(52.64, 63.60), (93.66, 104.62), (59.74, 70.70)$$

respectively.

(d) Use equation (7.15). The margin of error is the same for all 3 intervals because all three sample sizes are the same. The quantile used is the same as in (c). The margin of error is

$$t_{12,0.975}s_P\sqrt{\frac{1}{n_i} + \frac{1}{n_j}} = 2.179(5.619)\sqrt{\frac{1}{5} + \frac{1}{5}}$$

Using the $\hat{\mu}_i$ in the JMP output, the resulting confidence intervals for $\mu_{4H}-\mu_H$, $\mu_{4H}-\mu_B$, $\mu_H-\mu_B$ are

$$(-48.76, -33.28), (-14.83, 0.64), (26.18, 41.66)$$

respectively.

(e) Label the 4H lead Sample 1, the H lead Sample 2 and the B lead Sample 3. Use equation (7.20) with $c_1 = c_2 = 1/2$, $c_3 = -1$. The quantile used is the same $t_{12,0.975}$ as in (c). The resulting interval is

$$\frac{\hat{\mu}_1 + \hat{\mu}_2}{2} - \hat{\mu}_3 \pm t_{12,0.975}s_P\sqrt{\frac{(1/2)^2}{n_1} + \frac{(1/2)^2}{n_2} + \frac{(-1)^2}{n_3}}$$

$$= \frac{58.12 + 99.14}{2} - 65.2213 \pm 2.179(5.619)\sqrt{\frac{1/4 + 1/4 + 1}{5}}$$

$$= 13.41 \pm 6.706649$$

$$= (6.70, 20.12)$$

(h) 1. $H_0 : \mu_1 = \mu_2 = \mu_3$, $H_a :$ not all μ_1, μ_2, μ_3 are the same

   2. The test statistic is

   $$F = \frac{SSR/(r-1)}{SSE/(n-r)}.$$

   Assuming H_0 is true and the model $Y_{ij} = \mu_i + \epsilon_{ij}$, $\epsilon_{ij} \sim N(0, \sigma^2)$ is valid. The test statistic $F \sim F_{2,12}$.

3. The observed F in the JMP output is 76.10. The p-value is

$$P(F_{2,12} > 76.10) < 0.0001$$

4. The p-value is small, so we reject H_0 and conclude H_a.
5. There is overwhelming evidence that all 3 leads do not produce the same mean strength.

(i) The ANOVA table is given below.

Source	SS	df	MS	F
Treatment	4806.0	2	2403.0	76.10
Error	378.9	12	31.6	
Total	5185.0	14		