

Assignment 3 Answer Keys

1. In the experiment, $a = 5$ and $n_1 = n_2 = \dots = n_a = 13$. Only the summary statistics, sample means (\bar{y}_i) and sample standard deviations (s_i), are given. Though the original observations are not reported, the summary statistics are enough to calculate the sum of squares for ANOVA. Recall that $SS_{tr} = \sum_i n_i (\bar{y}_i - \bar{y}_{..})^2$ and $SS_E = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2$. The grand mean $\bar{y}_{..}$ is in fact the average of sample means, i.e.,

$$\bar{y}_{..} = \frac{\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5}{5} = 11.336$$

So, $SS_{tr} = 2.28176$.

To calculate SS_E , note that, for any fixed i , $\sum_j (y_{ij} - \bar{y}_i)^2 = (n_i - 1)s_i^2$, hence

$$SS_E = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_5 - 1)s_5^2 = 7.9488$$

Based on SS_{tr} and SS_E , one can construct the ANOVA table. $F_0 = 4.036$, and P-value=.004. The conclusion is that the treatment means are not all equal.

2. Use proc glm to get ANOVA table. Residual plots and QQ plot are generated to check model assumptions. Two formal tests for constant assumption are Levene's test and Bartlett's test. For normality, Kolmogorov-Smirnov test and Anderson-Darling test can be used. Construct 99% confidence interval for each of the treatment means. They are not simultaneous confidence intervals, hence the overall confidence level is less than 99%. If 99% simultaneous confidence intervals are requested, Bonferroni procedure can be used.

Assignment 4 Answer Keys

1.

a) Usual ANOVA shows that $F_0 = 21.31$ and P-value is less than 0.001. We reject null hypothesis and conclude that difference exists among the treatment effects.

b) Residual plots indicate that constant variance assumption might be invalid. Both Levene's test and Bartlett's test report P-value less than 5%, which implies that some remedy is in order.

c) Using the SAS file for approximate Box-Cox transformation, one has

$$\log s_i^2 = -.714 + .835 \log \bar{y}_i,$$

and $\hat{\beta} = .835$, so the possible (variance stabilizing) transformation is

$$Y' = Y^{1-\hat{\beta}} = Y^{.165}$$

Since $\hat{\beta}$ is approximately 1.00, a more meaningful transformation should be

$$Y' = \log(Y).$$

d) Use the SAS file for the exact Box-Cox transformation, SS_E is minimized at $\lambda = .25$. The transformation is

$$Y' = Y^{1/4}$$

In fact, $\lambda = 1/4$ and $\lambda = 0$ might not be different statistically, so both transformations can be used.

e) Apply ANOVA to the transformed responses. Residual plots and formal tests show that the violation of constant variance assumption has been corrected.

2. Apply usual ANOVA first to the data. QQ plot reveals some departure from normality, formal tests (Shapiro-Wilk's test) report p-values less than 5%. This implies that the

normality assumption is not valid and the result from ANOVA are questionable. Hence a nonparametric procedure is called for. Use `proc npar1way` to perform the Kruskal-Wallis test. The conclusion from Kruskal-Wallis is consistent with that from ANOVA in this problem.