

Stat 705, Spring 2015: Homework 1

Rehabilitation therapy Consider the data of problem 16.9. Note that parts 4 and 5 can be carried out in `proc glm` via `lsmeans` or `estimate`, and in `proc glimmix` via `lsmestimate`.

1. Fit the ANOVA model

$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad \epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2),$$

as described in class and report the ANOVA table including the p-value for testing $H_0 : \mu_1 = \mu_2 = \mu_3$. Do you reject at the 5% level?

2. Include a graph with a boxplot for each group (part of the default graphics). Comment on the assumption of equal variance across groups.
3. Qualitatively describe the relationship between physical fitness status and duration of required physical therapy.
4. Problem 17.10 parts b, c, d, and f.
5. Problem 17.15 parts a, b, c.
6. Report and comment on the following plots:
 - (a) e_{ij} vs. \hat{Y}_{ij} ; does constant variance look reasonable?
 - (b) t_{ij} vs. \hat{Y}_{ij} , are there any outliers at the 5% level? Note that $t(1 - \frac{0.05}{48}; 20) = 3.53$.
 - (c) the normal probability plot and histogram of the e_{ij} . Is normality reasonable?

Premium distribution Consider the data of problem 16.12.

1. Fit the ANOVA model

$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad \epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2),$$

as described in class and report the ANOVA table including the p-value for testing $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$. Do you reject at the 5% level?

2. Include a graph with a boxplot for each group (part of the default graphics). Comment on the assumption of equal variance across groups.
3. Qualitatively describe the relationship between agent and time lapse in days.
4. Problem 17.13 parts b, c, d.
5. Problem 17.18 parts a and b.
6. Report and comment on the following plots:
 - (a) e_{ij} vs. \hat{Y}_{ij} ; does constant variance look reasonable?
 - (b) t_{ij} vs. \hat{Y}_{ij} , are there any outliers at the 5% level? Note that $t(1 - \frac{0.05}{200}; 94) = 3.61$.
 - (c) the normal probability plot and histogram of the e_{ij} . Is normality reasonable?

Helicopter service: Consider the data of problem 18.15.

1. Report side-by-side boxplots using `proc boxplot` with the `boxstyle=schematic` option. Does constant variance seem likely?

2. Fit the ANOVA model

$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad \epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2),$$

as described in class. Do you reject $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 5% level?

3. Report the standard diagnostic panel. Comment on the residuals versus the predicted values.
4. Formally test $H_0 : \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4$ using the default Levene's test using something like `means shift / hovtest`; Do you reject at the 5% level? Does this coincide with your observations from parts 1 and 3?
5. Perform a nonparametric rank-based test of $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 5% level using `proc npar1way`. Do you reject at the 5% level? Does this coincide with part 2? Follow up the test with multiple comparisons using the `dscf` option and discuss.
6. Test $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ in the model with separate group variances

$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad \epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_i^2),$$

using `proc mixed` as shown in class. Compare to parts 2 and 5.

7. Perform a Box-Cox analysis on the response $Y_{ijk}^* = Y_{ijk} + 1$ in `proc transreg`. Which λ is picked?
8. Refit the model (with constant variance) suggested by part 7. Do you reject $H_0 : \mu_1^* = \mu_2^* = \mu_3^* = \mu_4^*$ at the 5% level? How does the plot of the residuals versus the fitted values look now? What does μ_i^* represent?