STAT3015/4030/7030 Generalised Linear Modelling

Tutorial 3

1. Reconsider the one-way ANOVA example `coagulation` from class. Refit the one-way ANOVA but this time treat diet effect as random. Write down the structure of your random effects model using mathematical notation. Interpret the output of your model fit. In particular, report the intra-class correlation coefficient, test whether the diet effect is significant, and provide estimates of the random effects for each diet. Also provide 95% confidence interval estimates for the effect of each diet on blood coagulation time.

2. Recall the blood group data from Tutorial 2:

Blood Type	Responses			
$O-$	9	11		
$O+$	20	19	23	19
$A-$	12	10		
$A+$	17	18	21	20
$B-$	16			
$B+$	24	28	25	
$AB-$	15			
$AB+$	25			

   (a) Recall that we determined that the A-antigen was not significantly related to the response, so that we could reduce the number of factor levels down to 4. Now, create two categorical variables, one for whether the individual can create the B-antigen and one for whether the individual has the Rh factor, and fit a two-way ANOVA model

$$Y_{ijk} = \mu + \tau_i + \alpha_j + \epsilon_{ijk}$$

   to the data. Do both factors appear to be significant?

   (b) Construct appropriate indicator variables for the two categorical predictors and refit the model as a linear regression to verify that the same results are obtained.

   (c) What are the estimators for τ_2 and α_2 in the two-way ANOVA model (assuming we have used the constraints $\tau_1 = \alpha_1 = 0$). Compare these estimates with the linear contrasts calculated in part (c) of Question 2 on Tutorial 2. What do you notice?

(d) Multiply the two indicators together to arrive at the indicator variable for the two-factor interaction, and test for additivity in the model.

3. This question was adapted from Ramsey and Schafer (2013). A 1989 study investigated the effect of heredity and environment on intelligence. From adoption registers in France, researchers selected samples of adopted children whose biological parents and adoptive parents came from either the very highest or the very lowest socio-economic status (SES) categories (based on years of education and occupation). They attempted to obtain samples of size 10 from each combination: (1) high adoptive SES and high biological SES, (2) high adoptive SES and low biological SES, (3) low adoptive SES and high biological SES, and (4) low SES for both parents. It turned out, however, only eight children belonged to combination three. The 38 children were given IQ tests. The scores are in the data file `SES.csv`.

   (a) Does the difference in mean scores for those with high and low SES biological parents depend on whether the adoptive parents were high or low SES?

   (b) As you should know by now, the answer to part (a) is NO. Now find out how much the mean IQ score is affected by the SES of adoptive parents, and how much it is affected by the SES of the biological parents? Is one of these effects larger than the other?

References

F. L. Ramsey and D. W. Schafer. The statistical sleuth: a course in methods of data analysis. Brooks/Cole, 2013.