Quiz 4, Stat3

Time allowed 45 mins

Total marks 15

There is one continuous response and one categorical predictor (factor) at L levels, with $n_i, i = 1, \cdots, L$ observations at each level. Observed values of the response are $y_{ij}, j = 1, \cdots n_i, i = 1, \cdots, L$. The model is

$$y_{ij} = \beta_i + \epsilon_{ij}$$

where ϵ_{ij} are independent $\mathcal{N}(0, \sigma^2)$

1. (2) Show that the least squares estimator of β_i is $\bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i$
2. (1) Show that the sum of squares residual, also known as the sum of squares within(SSW), is

$$\sum_{i=1}^{L}\sum_{j=1}^{n_i} \hat{\epsilon}_{ij}^2 = \sum_{i=1}^{L}\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

3. (1) What is the sum of squares total(SST)?
4. (3) Show that the sum of squares between (SSB) equals

$$\sum_{i=1}^{L} n_i(\bar{y}_i - \bar{y})^2$$

   where $\bar{y} = \sum_{i=1}^{L}\sum_{j=1}^{n_i} y_{ij} / \sum_{i=1}^{L} n_i$ using SST=SSB+SSW.

5. (2+5+1) A partially completed ANOVA table is given below, where all the n_i are equal to some n.

Source	df	SS	MS	F	p-value
Between	6	18.4			0.004
Within					
Total	41	45.2			

   (a) What are n and L?

   (b) Complete the ANOVA table.

   (c) For Bonferroni test for equality of pairwise means at overall 5% level, what is the level of significance for each pairwise test?