Additional note 6

Lecture 15

1. **Revision questions of L14**

   - What constraints are imposed to solve the singularity problem of matrix $\boldsymbol{X}^\top \boldsymbol{X}$?

     Answer: Impose constraint to reduce 1 df:

     Sum constraint: $\sum_i n_i \alpha_i = 0$ which implies $\hat{\mu} = \bar{Y}_{\bullet\bullet}$ and $\hat{\alpha}_i = \bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet}$ or

     Treatment constraint: $\alpha_1 = 0$ which implies $\hat{\mu} = \bar{Y}_{1\bullet}$ and $\hat{\alpha}_i = \bar{Y}_{i\bullet} - \bar{Y}_{1\bullet}$.

   - In comparing H_0 with null model with H_1 with treatment model, why $RSS_{H_0} = ToSS$?

     Answer: Null model has no regression parameter α_i apart from μ which is estimated by the overall mean $\bar{Y}_{\bullet\bullet}$. As the Total SS also allows for the overall mean, they are the same.

   - If the null model is rejected, that is, not all treatments are equal, how to test which treatments are different?

     Answer: Contrast or comparison is expressed as $c = \sum_i c_i \alpha_i$ such that $\sum_i c_i = 0$ and $\hat{c} = \sum_i c_i \bar{Y}_{i\bullet}$ for both sum and treatment constraints.

2. TSS with $t-1$ df can be partitioned into $t-1$ independent contrasts of any general form with $\sum_i c_i = 0$ and $\sum_i \frac{c_i d_i}{n_i}$ (for dependent contrasts) similar to the $t-1$ pairwise contrasts $\alpha_i - \alpha_1, i = 2, \ldots t$ comparing each treatment with the baseline treatment. Note that we don't have t independent contrasts.

Lecture 16

1. **Revision questions of L15**

   - Is WSS/BSS the TSS or RSS?

     Answer: WSS looks at the variability within a treatment, still not allowed for by each treatment. Hence WSS is RSS. BSS looks at variability across treatments and so this is the variability allowed for having treatments. Otherwise, there will be just one overall mean. Hence BSS is TSS.

   - For a data set with t treatments, how many independent contrasts are there? What is the condition for independence between two contrasts?

     Answer: There are $t-1$ independent contrasts. For two contrasts $c = \sum_i c_i \alpha_i$ and $d = \sum_i d_i \alpha_i$ with $\sum_i c_i = \sum_i d_i = 0$, the independence condition between c and d is $\sum_i c_i d_i = 0$.

   - What are the ranks of $\mathbf{P}_{V_0}$, $\mathbf{P}_{V_T}$, $\mathbf{P}_{W_T} = \mathbf{P}_{V_T} - \mathbf{P}_{V_0}$, $\mathbf{P}_{Res}$ and $\mathbf{P}_{Tot}$?

     Answer:

     $\mathbf{P}_{V_0}$ is the matrix of all $1/n$ and has rank 1 for SS of μ.

     $\mathbf{P}_{V_T}$ has block diagonal matrices of $1/n_i$ and 0 outside the block diagonal matrices. It has rank t for SS of $\alpha_1, \ldots, \alpha_t$.

     $\mathbf{P}_{W_T}$ has block diagonal matrices of $1/n_i - 1/n$ and $-1/n$ the block diagonal matrices. It has rank $t-1$ for TSS for the SS of treatments adjusting for the overall mean.

     $\mathbf{P}_{Res}$ has block diagonal matrices of $1 - 1/n_i$ along the diagonal, $-1/n_i$ off-diagonal and 0 outside the block diagonal matrices. It has rank $n-t$ for RSS.

     $\mathbf{P}_{Tot}$ has diagonal $1 - 1/n$ and $-1/n$ off-diagonal and has rank $n-1$ for total SS adjusting for the overall mean.

2. When the H_0 of null model is accepted, all α_i are equal and so they all equal to 0. Check previous additional note. When H_0 is rejected, α_i are not all equal and we look for which pairs of treatments will give the difference. P-values give the significance of each treatment but to look for which pairs of treatments that are different, we construct pairwise CI and check if it contains 0. However the usual CI works for each pair individually.

   To allow for other treatments, we look for simultaneous CI for all pairwise comparisons. To ensure simultaneous coverage of the true parameter for all CIs at level $1-\alpha$, each CI should have a wider coverage. Bonferroni adjustment suggests each CI to cover $1-\alpha/m$ in proportion where the number of pairs $m = \binom{t}{2}$ so that the simultaneous coverage of all CIs will have a lower bound of $1-\alpha$. Then the multiplier for each CI is $t_{n-t}(1-\alpha/2m)$. It provides more conservative (wider) CIs. Tukey proposed exact simultaneous CI with multiplier based on the distribution of studentized range called HSD with two parameters, number of treatments t and the df for $\hat{\sigma}^2$. Scheffe simultaneous CI uses the results that $c_i = n_i(\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet})$ provides the most significant sample contrast and the multiplier makes use of $F_{t-1,n-p}$.

Lecture 17

1. **Revision questions of L16**

   - In multiple testing with t treatments, what should be the level of each CI to achieve a simultaneous CI coverage of 95% with Bonferroni adjustment? What limitation is Bonferroni adjustment?

     Answer: The level for each CI is $1-0.05/t$ but the 95% is not an exact coverage but a lower bound for the coverage. The multiplier is quantile of $t_{n-t}(1-\alpha/t)$. The actual coverage may be 97% say so that it is wider than it should be.

   - Which adjustment tries to provide exact coverage? How?

     Answer: Tukey adjustment which is based on the studentised range with a distribution HSD. The multiplier is given by the quantile of $HSD_{t,n-t}(1-\alpha)/\sqrt{2}$.

   - Which adjustment uses instead the idea of the most significant sample contrast? What is the limitation?

     Answer: Scheffe adjustment which is based on the most significant contrast with coefficients $c_i = n_i(\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet})$ and its test statistic which involves F random variable. Then the multiplier is $\sqrt{(t-1) \times \text{quantile of } F_{t-1,n-t}(1-\alpha)}$ but it may provide the widest CI partly due to the under-estimation of the SE of the test statistic.

2. The most important idea is to link polynomial regression with one-way ANOVA if the predictor can be treated as both numerical and (ordinal) categorical. If there are t treatments, one can show the $t-1$ order polynomial and one-way ANOVA provide the same TSS and RSS. Using contrast with coefficients $c_i = n_i(\bar{x}_i - \bar{x})$, one can show that $\hat{c} = \sum_{i=1}^t c_i \bar{Y}_{i\bullet} = \hat{\beta}_1$ with $\sum_{i=1}^t c_i = 0$ and the 1 df component of TSS corresponds to $\frac{(\sum_{i=1}^t c_i \bar{Y}_{i\bullet})}{\sum_{i=1}^t \frac{c_i^2}{n_i}} = \hat{\beta}_1^2 S_{xx} = \text{LinRegSS}$.

3. The Bartlett test tests for the equality of variance across treatment groups.

---

last adjustments: April 20, 2021 by JC