MA4206: Linear Models

Satvik Saha, 19MS154 March 1, 2023

1 Introduction

There are 90 observed abrasion resistance scores y_{ijl} for jeans subjected to three types of denim treatments (indexed by i) and laundry cycles (indexed by j).

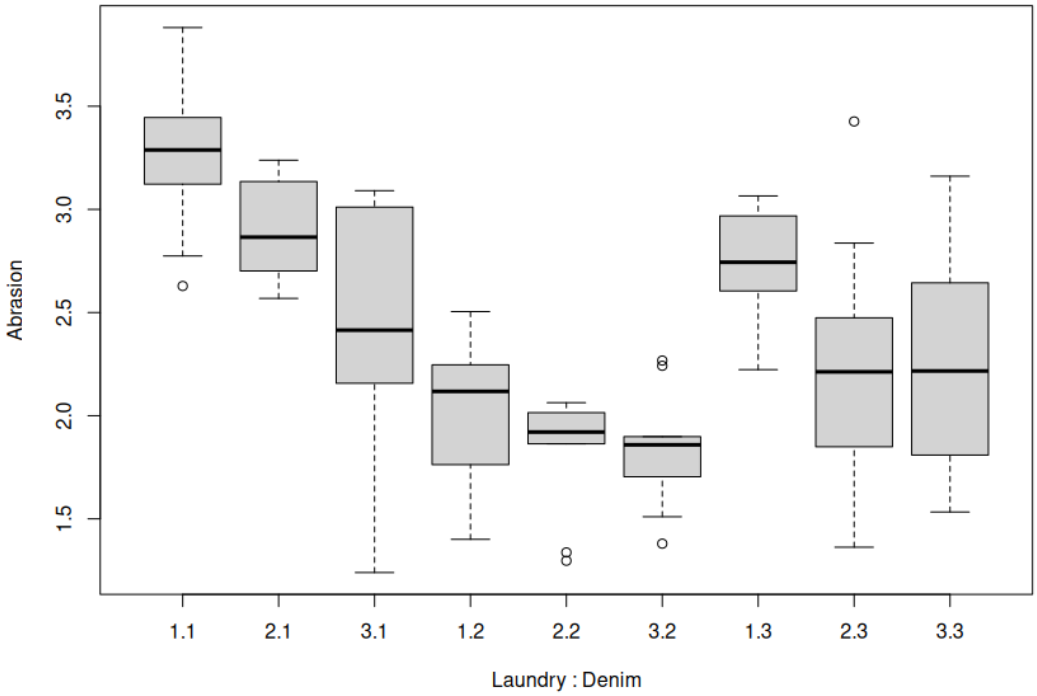

Figure 1: Abrasion scores for each of the 9 groups of treatment. Lower scores are worse.

These are fitted against the linear model

$$y_{ijl} = \mu + \tau_i + \beta_j + \epsilon_{ijl}.$$

By collecting the observations y_{ijl} into a vector $\boldsymbol{y}_{90\times 1}$, our model looks like

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where

$$\boldsymbol{X}_{90\times 7} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}, \qquad \boldsymbol{\beta}_{7\times 1} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \qquad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I}_{90}).$$

In the expression for $\boldsymbol{X}$, all $\mathbf{1}, \mathbf{0}$ are 10×1 filled vectors.

Now, we can calculate the least square estimates

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{y}, \qquad \hat{\boldsymbol{y}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{y}.$$

Note that $\text{rank}(\boldsymbol{X}) = 5$, so $\hat{\boldsymbol{\beta}}$ is not uniquely determined; however, $\hat{\boldsymbol{y}}$ is! Indeed, we can check that for any $1 \le l, l' \le 10$, we have $\hat{y}_{ijl} = \hat{y}_{ijl'}$, so it is enough to present the group estimates

$$\begin{bmatrix} \hat{y}_{11\cdot} \\ \hat{y}_{12\cdot} \\ \hat{y}_{13\cdot} \\ \hat{y}_{21\cdot} \\ \hat{y}_{22\cdot} \\ \hat{y}_{23\cdot} \\ \hat{y}_{31\cdot} \\ \hat{y}_{32\cdot} \\ \hat{y}_{33\cdot} \end{bmatrix} = \begin{bmatrix} 3.148348 \\ 2.792664 \\ 2.625998 \\ 2.203681 \\ 1.847998 \\ 1.681331 \\ 2.715011 \\ 2.359328 \\ 2.192661 \end{bmatrix}. \tag{$*$}$$

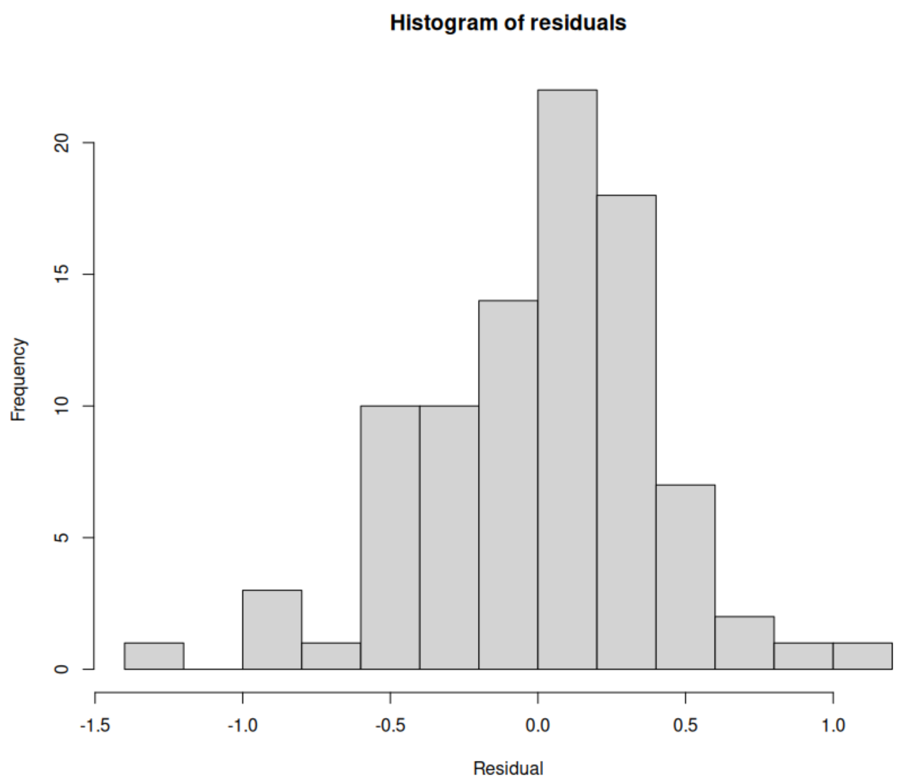

Figure 2: The distribution of residuals from the linear model.

2 The constrained model

We now impose the conditions $\sum_i \tau_i = \sum_j \beta_j = 0$. To do so, we replace $\tau_3 = -\tau_1 - \tau_2$, $\beta_3 = -\beta_1 - \beta_2$, whence our new model looks like

$$\boldsymbol{y} = \boldsymbol{X}^* \boldsymbol{\beta}^* + \boldsymbol{\epsilon},$$

where

$$\boldsymbol{X}^*_{90\times 5} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & -\mathbf{1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & -\mathbf{1} & -\mathbf{1} \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{1} & -\mathbf{1} & -\mathbf{1} & -\mathbf{1} & -\mathbf{1} \end{bmatrix}, \qquad \boldsymbol{\beta}^*_{5\times 1} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \qquad \boldsymbol{\epsilon} \sim N(0, \sigma^2 \boldsymbol{I}_{90}).$$

With this, $\boldsymbol{X}^*$ has full rank 5. The group estimates $\hat{y}_{ij\cdot}$, hence the residuals, remain *unchanged* from $(*)$. The least square estimates of the parameters are

$$\hat{\boldsymbol{\beta}}^* = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 2.39633556 \\ 0.45933444 \\ -0.48533222 \\ 0.29267778 \\ -0.06300556 \end{bmatrix}.$$

3 Estimating linear parametric functions

Consider the LPF

$$\boldsymbol{A}\boldsymbol{\beta} = \begin{bmatrix} \tau_1 - \tau_2 \\ \tau_1 - \tau_3 \\ \tau_2 - \tau_3 \end{bmatrix},$$

where

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

Here, we can calculate $\rho = \operatorname{rank}(A) = 2$; the first two rows of A are linearly independent, the third is their difference. Importantly, each row of A belongs to $\operatorname{col}(X^\top)$. For example, the first row of A can be expressed as the difference of rows 1 and 31 of $\boldsymbol{X}$. Thus, $\boldsymbol{A}\boldsymbol{\beta}$ has a BLUE

$$\widehat{\boldsymbol{A}\boldsymbol{\beta}} = \boldsymbol{A}\hat{\boldsymbol{\beta}} = \boldsymbol{A}(\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{X}^\top \boldsymbol{y}.$$

We calculate

$$\boldsymbol{A}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 0.9446667 \\ 0.4333367 \\ -0.5113300 \end{bmatrix}.$$

4 Confidence intervals/regions for estimates

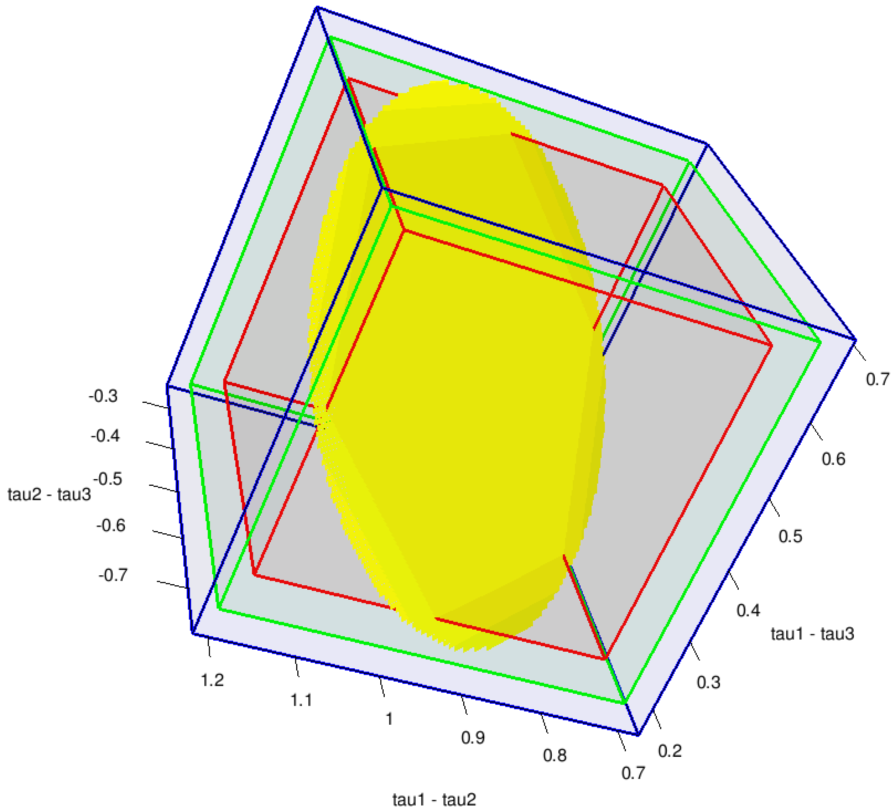

Figure 3: Confidence regions for $\boldsymbol{A\beta}$, with $\alpha = 0.05$. The joint confidence region is indicated by the *yellow* ellipse, the individual confidence intervals are the sides of the *red* cube, the Bonferroni confidence region is indicated by the *green* cube, and the Scheffé confidence region is indicated by the *blue* cube.

Note that $(\tau_1 - \tau_2) - (\tau_1 - \tau_3) + (\tau_2 - \tau_3) = 0$.

4.1 Individual confidence intervals

Denoting each row of $\boldsymbol{A}$ by $\boldsymbol{a}_j^\top$, we can calculate the individual $(1-\alpha)$ confidence intervals for $\boldsymbol{a}_j^\top \boldsymbol{\beta}$ as

$$\left[\boldsymbol{a}_j^\top \hat{\boldsymbol{\beta}} - \sqrt{\hat{\sigma}^2 \boldsymbol{a}_j^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{a}_j}\, t_{85;\alpha}, \quad \boldsymbol{a}_j^\top \hat{\boldsymbol{\beta}} + \sqrt{\hat{\sigma}^2 \boldsymbol{a}_j^\top (\boldsymbol{X}^\top \boldsymbol{X})^- \boldsymbol{a}_j}\, t_{85;\alpha}\right]$$

Here, $\hat{\sigma}^2 = R_0^2/(n-r)$, where $n = 90$, $r = \text{rank}(\boldsymbol{X}) = 5$, and R_0^2 is the sum of squares of residuals $\boldsymbol{y} - \hat{\boldsymbol{y}}$.

The calculated confidence intervals for $\alpha = 0.05$ are

$$\tau_1 - \tau_2 : [0.727, 1.162], \qquad \tau_1 - \tau_3 : [0.216, 0.651], \qquad \tau_2 - \tau_3 : [-0.729, -0.294]$$

All of them have half-width 0.217.

4.2 Joint confidence region

The joint $(1-\alpha)$ confidence region for $\boldsymbol{A\beta}$ is given by

$$\{\boldsymbol{A\beta}\colon (\boldsymbol{A\beta}-\boldsymbol{A}\hat{\boldsymbol{\beta}})^\top(\boldsymbol{A}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{A}^\top)^-(\boldsymbol{A\beta}-\boldsymbol{A}\hat{\boldsymbol{\beta}}) \le 2\hat{\sigma}^2 F_{2,85;1-\alpha}, \quad \boldsymbol{A\beta}-\boldsymbol{A}\hat{\boldsymbol{\beta}} \in \mathrm{col}(\boldsymbol{A}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{A}^\top)\}.$$

It is easy to see that the latter condition is equivalent to the restriction that the sum of the first and third components of $\boldsymbol{A\beta}$ must be equal to the second. Thus, this region is a planar section of an ellipsoid, and looks like a (filled) ellipse in $\mathbb{R}^3$.

Indeed,

$$\boldsymbol{A}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{A}^\top = \frac{1}{30}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

4.3 Bonferroni simultaneous intervals

Setting $\gamma = \alpha/\rho$, we have the Bonferroni simultaneous intervals

$$\left[\boldsymbol{a}_j^\top\hat{\boldsymbol{\beta}} - \sqrt{\hat{\sigma}^2\boldsymbol{a}_j^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{a}_j}\; t_{85;\gamma}, \quad \boldsymbol{a}_j^\top\hat{\boldsymbol{\beta}} + \sqrt{\hat{\sigma}^2\boldsymbol{a}_j^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{a}_j}\; t_{85;\gamma}\right]$$

These together define a cuboidal region in $\mathbb{R}^3$.

The calculated confidence intervals for $\alpha = 0.05$ are

$$\tau_1-\tau_2 : [0.695, 1.194], \qquad \tau_1-\tau_3 : [0.184, 0.683], \qquad \tau_2-\tau_3 : [-0.761, -0.262],$$

All of them have half-width 0.250.

4.4 Scheffé simultaneous intervals

We have the Scheffé simultaneous intervals

$$\left[\boldsymbol{a}_j^\top\hat{\boldsymbol{\beta}} - \sqrt{2\hat{\sigma}^2\boldsymbol{a}_j^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{a}_j\; F_{2,85;\alpha}}, \quad \boldsymbol{a}_j^\top\hat{\boldsymbol{\beta}} + \sqrt{\hat{\sigma}^2\boldsymbol{a}_j^\top(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{a}_j\; F_{2,85;\alpha}}\right]$$

These together define a cuboidal region in $\mathbb{R}^3$, larger than the Bonferroni intervals and enclosing the ellipsoidal joint confidence region.

The calculated confidence intervals for $\alpha = 0.05$ are

$$\tau_1-\tau_2 : [0.672, 1.217], \qquad \tau_1-\tau_3 : [0.161, 0.706], \qquad \tau_2-\tau_3 : [-0.784, -0.239],$$

All of them have half-width 0.273.