# MAST30025: Linear Statistical Models

## Week 10 Lab

1. Consider again the data from Question 5 from the Week 8 lab. In a manufacturing plant, filters are used to remove pollutants. We are interested in comparing the lifespan of 5 different types of filters. Six filters of each type are tested, and the time to failure in hours is given in the dataset `filters` (on the website, in `csv` format).

   (a) Calculate $s^2$.

   (b) Calculate a 95% confidence interval for the difference in lifespan between filter types 3 and 4.

   (c) Show that the hypothesis that the filters all have the same lifespan is testable.

   (d) Test this hypothesis, using matrix theory.

   (e) Test the same hypothesis using the `linearHypothesis` function from the `car` package.

2. Derive a formula for a prediction interval for $\mathbf{t}^T\boldsymbol{\beta}$ in the less than full rank model.

3. Consider a one-way classification model with three levels. To test $\tau_1 = \tau_2 = \tau_3$ we could use either of the following matrices:

$$C_1 = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \; C_2 = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

   Show that the test statistics formed using these two matrices are the same.

   Hint: find $A$ such that $AC_1 = C_2$.

4. An industrial psychologist is investigating absenteeism among production-line workers, based on different types of work hours: (1) 4-day week with a 10-hour day, (2) 5-day week with a flexible 8-hour day, and (3) 5-day week with a structured 8-hour day. A study is conducted and the following data obtained of the average number of days missed:

| | Work plan | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Mean | 9 | 6.2 | 10.1 |
| Number | 100 | 85 | 90 |

   They also find $s^2 = 110.15$.

(a) Test the hypothesis that the work plan has no effect on the absenteeism.

(b) Test the hypothesis that work plans 1 and 3 have the same rate of absenteeism.