# MAST30027: Modern Applied Statistics

## Assignment 2, 2023.

Due: 5pm Monday September 4th

---

- This assignment is worth 7% of your total mark.
- To get full marks, show your working including 1) R commands and outputs you use, 2) mathematics derivation, and 3) rigorous explanation why you reach conclusions or answers. If you just provide final answers, you will get zero mark.
- The assignment you hand in must be typed (except for math formulas), and be submitted using LMS as a single PDF document only (no other formats allowed). For math formulas, you can take a picture of them. Your answers must be clearly numbered and in the same order as the assignment questions.
- The LMS will not accept late submissions. It is your responsibility to ensure that your assignments are submitted correctly and on time, and problems with online submissions are not a valid excuse for submitting a late or incorrect version of an assignment.
- We will mark a selected set of problems. We will select problems worth $\geq 50\%$ of the full marks listed.
- If you need an extension, please contact the lecturer before the due date with appropriate justification and supporting documents. Late assignments will only be accepted if you have obtained an extension from the lecturer before the due date. To ensure that the lecturer responds to your extension request email before the due date, please contact 24h before the due date. Under no circumstances an assignment will be marked if solutions for it have been released.
- Also, please read the "Assessments" section in "Subject Overview" page of the LMS.

---

**Note: There is no unique answer for this problem. The report for this problem should be typed and include R code used for your analysis. Hand-written report or screen-captured R codes/figures won't be marked. An example report written by a student previous year has been posted on LMS.**

**Experiment Design:** Components are attached to an electronic circuit card assembly by a wave-soldering process. The soldering process involves baking and preheating the circuit card and then passing it through a solder wave by conveyor. Defects arise during the process, and an experiment was run to try and determine the effect on the number of defects of various aspects of the process.

**Data:** The data is taken from Condra, Lloyd, Reliability Improvement with Design of Experiment. CRC Press, 2001. Full wavesolder data has 48 observations, each of which has the number of defects and seven predictor variables. In this assignment, we will consider only the number of defects (response variable), and four predictor variables, prebake, flux, cooling, temp. The data can be found in the file `assignment2_2023.txt`. The dataset has 48 rows representing 48 observations. Each row has entries for:

- numDefects: number of defects
- prebake: prebake condition - a factor with levels 1 2

- flux: flux density - a factor with levels 1 2
- cooling: cooling time - a factor with levels 1 2
- temp: solder temperature - facctor with levels 1 2

You can read the data using the following command.

```
> data <- read.table(file ="assignment2_2023.txt", header=TRUE)
> dim(data)
[1] 48  5
> names(data)
[1] "numDefects" "prebake"    "flux"       "cooling"    "temp"
> data$prebake <- factor(data$prebake)
> data$flux <- factor(data$flux)
> data$cooling <- factor(data$cooling)
> data$temp <- factor(data$temp)
```

**Problem:** We want to determine which factors (prebake, flux, cooling, temp) and two-way interactions are related to the number of defects. Write a report on the analysis that should summarise the substantive conclusions and include the highlights of your analysis: for example, data visualisation, choice of model (e.g., Poisson, binomial, gamma, etc), model fitting and model selection (e.g., using AIC), diagnostic, check for overdispersion if necessary, and summary/interpretation of your final model.

At each step of you analysis, you should write why you do that and your interpretation/conclusion. For example, "I make an interaction plot to see whether there are interactions between X and Y", show a plot, and "It seems that there are some interaction between X and Y".