

Homework 4

1. In a study of the effect of glucose on insulin release, identical specimens of pancreatic tissue were equally and randomly assigned to three different levels of glucose concentration (low, medium, high). The amount of insulin produced by each tissue after treatment was recorded. The data set named **insulin.dat** needs to be downloaded. In `insulindata`, the first column contains the amounts at the low concentration level, the second column the amounts at the medium concentration level, and the third column the amounts at the high concentration level. To read a data set like this, you can use the following SAS code in the data step to create a data set suitable for the `glm` procedure.

```
data insulin;
infile 'H:\dataset\insulin.dat';
input t1 t2 t3;
y=t1; trt=1; output;
y=t2; trt=2; output;
y=t3; trt=3; output;
drop t1 t2 t3;
```

This creates a treatment variable (`trt`) and a response variable (y). Or otherwise you can re-organize the data in a text editor. If you use SAS OnDemand, please refer to the YouTube video we showed in class to specify the path on the cloud service to your `insulin.dat`.

- a). Test the hypothesis that there is no difference between the treatments in the amount of insulin produced (use $\alpha = 0.01$).
- b). Assess whether the model assumptions are valid?
- c). Construct 99% CIs for the average insulin amounts at the low, medium and high glucose concentrations separately.

2. A factor with three levels was studied in an experiment. The data is given as follows, in which the first column includes the treatments and the second column includes the responses. You can download the data named **hw4pr3.dat**.

```
1  2.23
1  3.04
.  ...
2  3.65
.. .....
3  8.53
.  .....
3  8.12
```

- Test the hypothesis that there is no difference across the treatments (use $\alpha = .05$).
- Use proper plots to check whether the constant variance assumption is valid. Can you use a formal test to support your conclusion?
- The Box-Cox procedure suggests a log transformation of the response y . Using `logresp`, repeat a) and b) for the transformed data. You can use some SAS functions in the data step to generate the new `logresp`.

3. Four different designs for a digital computer circuit are being studied to compare the amount of defects. The following data have been obtained

```
design defect
1      7
1      2
1      4
1      7
```

1	2
2	10
2	6
2	9
2	7
2	5
3	16
3	13
3	11
3	13
3	13
4	5
4	5
4	2
4	2
4	7

a). Is the amount of defects present the same for all four designs? Use $\alpha = 0.05$ in standard ANOVA.

b). Check the model assumptions. In particular, how do you think about the normality assumption? Show the QQ plot and the test results for constant variance assumption