Paper STATS 330 **Assignment 4** **Due: Monday 21 October 2002**

The data for this assignment comes from a study that investigated the effect of insulin on laboratory mice. The response was whether or not the mice had convulsions when given insulin. We are interested in modelling how the proportion of mice with convulsions varies with the type of preparation and with the dose applied.

Standard Preparation

dose	with convulsions	total
3.4	0	33
5.2	5	32
7.0	11	38
8.5	14	37
10.5	18	40
13.0	21	37
18.0	23	31
21.0	30	37
28.0	27	30

Test Preparation

dose	with convulsions	total
6.5	2	40
10.0	10	30
14.0	18	40
21.5	21	35
29.0	27	37

The data for this assignment is **not** on the STATS 330 webpage and is **not** on the servers in either the Advanced Laboratory or the Undergraduate Laboratory. The first thing you need to do for this assignment is to create the variables you need to analyse this data in *R*. To make it easier for the markers, please name the variable you use to indicate the dose dose and name the variable you use to specify the type of preparation prep.

For this assignment, first fit the model that just uses dose and prep as regressors. Use the diagnostics procedures we covered in class to assess the suitability of this model. Explore the possibility of improving the model by adding an interaction term and/or transforming the numeric regressor (or using a polynomial model). Identify the model that you would recommend be used for this data and use that model to explain how dose and prep affect the probability of convulsions.

As usual, your assignment should consist of two parts. The first part should be a report that contains:

1. an executive summary,
2. a description of the data,
3. a section that presents the model you selected and explains what this model indicates about how the probability of convulsions is affected by the type of preparation and the dose.

The second part of your assignment is a statistical appendix that explains to the marker what you did. Your statistical appendix should contain the following:

1. Diagnostics for the model that just contains dose and prep as regressors and a discussion of these.

2. A discussion of whether it is useful to include the interaction term and/or transform dose – include a key pieces of supporting evidence. Make sure you clearly identify which model you would recommend and justify your choice.

3. Include diagnostics for the model that you recommend (if it is different from the model from 1). Any outliers, high leverage and influential points should be identified and the impact of these observations on the fitted model should be evaluated.

This assignment should be handed in to the appropriate box in the basement of the Maths/Physics building by the SMIS Resource Centre, by 4pm on Monday, 21 October.