Paper 475.330 **Assignment 4** **Due: Friday 13 October**

Framingham is an industrial town located approximately 30 km from Boston. In 1948 a study was begun with the aim of identifying factors that are related to the occurrence of coronary heart disease (CHD). At the start of the study, a large proportion of the town's inhabitants were examined for the presence of CHD. Measurements were made on a number of potential risk factors. The individuals who were found to be free of CHD at that time were followed up for twelve years and those who developed CHD during that period were identified. The following dataset was extracted from that data and relates the proportions developing CHD to the initial serum cholesterol level (mg per 100 ml) of these individuals cross classified by age and sex.

		Serum Cholesterol Level			
Sex	Age	< 190	$190 - 219$	$220 - 249$	≥ 250
Male	$30 - 49$	13/340	18/408	40/421	57/362
	$50 - 62$	13/123	33/176	35/174	49/183
Female	$30 - 49$	6/542	5/552	10/412	18/357
	$50 - 62$	9/58	12/135	21/218	48/395

To access this data simply logon to the advlab computer, start Splus and type

```
> framingham.df
```

For this assignment, use binary ANOVA to investigate the relationship between the 3 regressors and CHD. The effect of serum cholesterol level on the occurrence of CHD is of particular interest.

Your assignment should consist of two parts: a report that clearly explains your findings and a statistical appendix that outlines your analysis. You are expected to use suitable diagnostics procedures and discuss any problems you identify.

This assignment should be handed in to the appropriate box in the basement of the Maths/Physics building by the Resource Centre, by 4pm on Friday, 13 October.