Paper STATS 330 **Assignment 4** **Due: Friday, 12 October, 2001**

In this assignment you will investigate data from a survey of workers in the US cotton industry. The data records whether the workers were suffering from a lung disease called byssinosis. Also recorded were the values for five categorical explanatory variables: the race, sex and smoking status of the worker, the length of employment and the amount of dust in the workplace. The amount of dust in the workplace is thought to be a major factor in the occurrence of byssinosis but the other measured variables may also have an impact.

The data has been put into a dataframe called byssinosis.df which can accessed from S-plus in the advanced lab. It has also been put on the webpage. This dataframe contains 7 variables: dust, race, sex, smoking, employ, yes and total. The first five of these are the categorical explanatory variables which have the following codings for their levels:

dust	amount of dust in workplace: 1-high, 2-medium, 3-low.
race	ethnic origin of worker: 1-white, 2-other.
sex	sex of worker: 1-male, 2-female.
smoking	smoking status: 1-smoker, 2-nonsmoker.
employ	years of employment: 1-less than 10, 2-10 to 20, 3-more than 20.

For each combination of the explanatory variables, total represents the total number of workers and yes represent the number that suffer from byssinosis.

For this assignment, use binary ANOVA to investigate the relationship between the explanatory variables and the occurrence of byssinosis. Primary interest is in how the incidence of byssinosis is affected by the amount of dust in the workplace but the other variables need to be taken into account as well.

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 SMIS Resource Centre, by 4pm on Friday, 12 October.