

# 140.754 Homework 1

Due date: Apr 10 (Tuesday) 11:59pm

1. **Number of Moons** The file Ex0206.csv contains data that show some characteristics of terrestrial planets, gas giants, and dwarf planets in our solar system, including the known number of moons. Apparently, larger planets have more moons, but is it the volume (as indicated by diameter) or mass that are more relevant, or is it both? Answer this question and arrive at a model for describing the mean number of moons as a function of planet size (Hint: you may need to consider overdispersion in the data). Use diagnostic statistics to check whether there are isolated influential points.

2. **Endometrial Cancer** In order to determine the effects of potential risk factors for developing endometrial cancer, 63 case-control matched sets were analyzed. Each set includes one case and four controls matched by age, marital status, and the time of entering an identical community when diagnosed. **Primary risk factors** considered include the history of gall-bladder disease, presence of hypertension, estrogen use, non-estrogen drug use, and obesity measure.

Number	Name	Description	Codes/Range
1	SET	Matched set indicator	1-63
2	CASE	Case-control indicator	1 = Case; 0 = control
3	AGE	Age at diagnosis	55-84
4	AGEG	Age group indicator	1 = 55-64; 2 = 65-74; 3 = 75+
5	EST	Estrogen usage	0 = No; 1 = Yes
6	GALL	Gallbladder disease	0 = No; 1 = Yes
7	HYP	Hypertension	0 = No; 1 = Yes
8	OB	Obesity	0 = No; 1 = Yes
9	NON	Non-estrogen drug	0 = No; 1 = Yes
10	num	Counts Subjects within Set	1-5

The data can be read into R with command

```
data <- read.table("endometrial.txt", header = TRUE)
```

Analyze the data to answer the following questions.

(1) Which primary risk factors are significantly associated with endometrial cancer occurrence? Which factors are not?

(2) Compare a model with all primary risk factors with a model containing just those factors that are significantly associated with cancer. Perform statistical test to see whether the extra factors add any predictive information?

(3) Is there an interaction between estrogen use and history of gall bladder disease?