1. A multicenter study of survival rates for breast cancer patients is available on the course web page. The variables are:

- # survived
- # cases = (# survived + # not survived)
- Diagnostic Center (1=Tokyo, 2=Boston, 3=Glamorgan)
- Age (1=under 50, 2=50-69, 3=over 70)
- Inflammed (1=minimal inflammation, 2=greater inflammation)
- Appearance (1=malignant appearance, 2=benign appearance)

(a) What are the main effects of Center, Age, Inflammed, and Appearance upon survival probability? Which factors have a significant effect and what are the natures of the significant effects? In particular, which centers, if any, have better survival probabilities than the others? Also, how do age and the two histologic variables affect survival probability?

(b) Are there any signs of interactions? If so, what are their natures?

(c) Are there any data points that seem outlying or highly influential?

2. Suppose that $X_1, \ldots, X_n$ are iid from a distribution with expectation $\mu$, variance $\sigma^2$, and third central moment $\kappa_3 := E(X_i - \mu)^3$. Justify formally that

$$E\{h(\overline{X}) - E[h(\overline{X})]\}^3 = n^{-2}\left\{h'(\mu)^3\kappa_3 + 3h''(\mu)[h'(\mu)]^2\sigma^4\right\} + O(n^{-3}).$$

To be continued.