Problem Set 11 (An extended problem set, I may add a few more problems before my last lecture)

Sun Gang will discuss some of these problems on Monday Nov 12, 2007. I will go through the remaining others during my final lecture.

Rice 3rd ed.

Chapter 4: Problems (Section 4.7) – 79, 90.

Chapter 5: Problems (Section 5.4) – 14 (recall exercise 13 from our previous practice problem set).

Chapter 8: Problems (Section 8.10) – 6, 18, 50, 71, 73. (I don't know how difficult problem 18 is, I haven't tried it myself. Try it out yourself, if it turns out to be too difficult, I would go through it during my last lecture)

Chapter 9: Problems (Section 9.11) – 5,6 (I may add a few more, depending on how much I can cover during this week's lecture).

I am adding an extra problem to the existing problem set 11.

The following question came in the last year's final exam. However, I would ask you to do a generalized likelihood ratio test (instead of Pearson Chi-square test, which I did not cover in class).

The following data are about horse racing. A certain racetrack ("travbane") contains 8 starting gates ("startporter") where the horses are kept just before the start of a race. It is claimed that starting from gate no. 1 – 4 represents an advantage in the sense that it adds to the probability of winning the race. Results from n = 144 races are given in table 1. The table shows for example that 19 of the races were won by the horse starting from gate 3. The horses starting from a given gate varies from race to race.

Table 1

Start gate	1	2	3	4	5	6	7	8	Sum
Winners	32	21	19	20	16	11	14	11	144

Let Y_j denote the number of winning horses among those starting from gate j ($j = 1, 2, \ldots, 8$). Assume that $(Y_1, Y_2, \ldots, Y_8)$ is multinomially distributed with parameters $(n, p_1, p_2, \ldots, p_8)$, where n = 144, $p_1 + p_2 + \cdots + p_8 = 1$, and $Y_1 + Y_2 + \cdots + Y_8 = n$. Here p_j is the probability that a horse starting from gate j wins given that nothing else is known about the horse except that it starts from gate j.

Consider the following special case where the first 4 gates have the same winning probability and the same is true for the last 4 gates:

Model 1 $p_1 = p_2 = p_3 = p_4 = \theta$ and $p_5 = p_6 = p_7 = p_8 = \eta$ where $4\theta + 4\eta = 1$ (which is equivalent to $\eta = \frac{1}{4} - \theta$), and θ, η otherwise unknown.

Assuming model 1 to be true, show that the mle's for θ, η are given by

$$\hat{\theta} = \frac{S_1}{4n} \quad \text{where} \quad S_1 = Y_1 + Y_2 + Y_3 + Y_4$$

$$\hat{\eta} = \frac{1}{4} - \hat{\theta}$$

Calculate the mle estimates from table 1.

(i) Perform a generalized likelihood ratio test at the level of significance 10% for the hypothesis that model 1 is true.