Statistics 582, Problem Set 2

Wellner; 1/21/98

Reading: Lehmann, TPE, Chapter 6, sections 6.1 and 6.2. Chapter 4, sections 3 - 6.
Due: Wednesday, January 14, 1998.

1. This is a further continuation of problem 10.4 from 581:
   A. With the given data, carry out the Wald, likelihood ratio, and Rao tests of the (composite) null hypothesis $H_0 : \beta = 1$ versus the alternative hypothesis $H_1 : \beta \neq 1$. Briefly compare the results.
   B. Find what the statistics you used in part A, appropriately normalized, converge to in probability. [That is, prove analogues for this special case of the general results proved in theorem 4.2.1 for the simple null hypothesis case.]

2. Consider the following parametric model: suppose that
   $$N = (N_{ij} : 1 \le i \le r,\ 1 \le j \le s) \sim \text{Multinomial}_{rs}(n, \underline{p} = (p_{ij} : 1 \le i \le r,\ 1 \le j \le s)).$$
   Here $\sum_{i=1}^r \sum_{j=1}^s p_{ij} = 1$.
   A. What is the maximum likelihood estimator $\underline{\widehat{p}}$ of $\underline{p}$? What is the limiting distribution of $\sqrt{n}(\widehat{p} - p)$?
   B. Suppose we consider estimation of $\underline{p}$ in the restricted model in which the marginal probabilities are known: $p_{i\cdot} \equiv \sum_{j=1}^s p_{ij} = p_{i\cdot}^0$, $i = 1, \ldots, r$ and $p_{\cdot j} \equiv \sum_{i=1}^r p_{ij} = p_{\cdot j}^0$, $j = 1, \ldots, s$ where the $p_{i\cdot}^0$'s and $p_{\cdot j}^0$'s are known. Show that the maximum likelihood estimator $\widehat{p}$ has the form $\widehat{p}_{ij} = (N_{ij}/n)/(a_i + b_j)$. What can you say about the limiting distribution of $\sqrt{n}(\widehat{p} - p)$ in this (sub)model?

3. Suppose, as in Example 4.3.10, that $\underline{X}_1, \ldots, \underline{X}_n$ are i.i.d. $\text{Mult}_k(1, \underline{p})$, so that $\underline{N}_n \equiv \sum_{j=1}^n \underline{X}_j \sim \text{Mult}_k(n, \underline{p})$. Use Jensen's inequality to show that the log-likelihood
   $$l_n(\underline{p}|\underline{X}) = \sum_{j=1}^k N_j \log p_j + \sum_{i=1}^n \log\left(\frac{1!}{X_{i1}! \cdots X_{ik}!}\right)$$
   is maximized by $\underline{\widehat{p}} = \underline{N}_n/n$. [Hint: write the first term of $l_n(\underline{p}|\underline{X})$ as $n \sum_{j=1}^k \widehat{p}_j \log p_j$.]

4. (Optional bonus problem). Suggest and investigate other approaches to the same problem as in #2; e.g. minimum chi-square, modified mimimum chi-square, minimum Kullback-Leibler distance, or minimum Hellinger distance estimation. Does your modified approach yield an (asymptotically) efficient estimator of $\underline{p}$ in the constrained model?