Statistics 581/582, Winter Quarter 2008
Problem Set 10

Reading: Ferguson, Sections 12–15.

Problem 35 (van der Waerden test, 2 points). The van der Waerden test is a competitor of the two-sample randomization and rank-sum tests, in which the value of each observation is replaced by $\Phi^{-1}(r/(N+1))$, where Φ is the standard normal distribution function, and r is the rank of the observation among the pooled sample of size $N = m + n$. Thus, under the hypothesis of no treatment effect the van der Waerden statistic is of the form

$$S_N = \sum_{i=1}^{N} a_{\sigma(i)} z_i, \qquad \text{where} \qquad a_i = \begin{cases} 1, & 1 \le i \le m, \\ 0, & m+1 \le i \le N, \end{cases}$$

and $z_i = \Phi^{-1}(i/(N+1))$ for $i = 1, \ldots, N$.

(a) Show that $\bar{z} = 0$ and $\sigma_z^2 \to 1$ as $N \to \infty$.

(b) Suppose that $n/N \to r$ as $N \to \infty$, where $0 < r < 1$. Find a nondegenerate limit distribution for S_N.

Problem 36 (estimation for a Cauchy sample, 2 points). Let $X_1, X_2, \ldots$ be a sample from the Cauchy distribution $\mathcal{C}(\mu, \sigma)$ with Lebesgue density

$$f(x) = \frac{1}{\pi\sigma} \frac{1}{1 + ((x-\mu)/\sigma)^2}.$$

Suppose that $0 < p < 1/2$.

(a) Find the asymptotic distribution of the estimator

$$\hat{\mu}_n = \frac{1}{2}(X_{([np])} + X_{([n(1-p)])}).$$

(b) Study the asymptotic variance of the estimator $\hat{\mu}_n$ in its dependence on p, and compare to the sample median as an estimator of the location parameter μ.

Problem 37 (estimation for an exponential power sample, 6 points). Suppose that we have a sample $X_1, X_2, \ldots$ from the exponential power distribution with Lebesgue density

$$f(x) = c(\alpha)\, e^{-|x-\theta|^\alpha}$$

where $\alpha > 0$ and $\theta \in \mathbb{R}$.

(a) Find the value of the normalizing constant $c(\alpha)$.

(b) Find an asymptotic distribution for the sample mean $\bar{X}_n$ as an estimator of the location parameter θ.

(c) Find an asymptotic distribution for the sample median Med_n as an estimator of the location parameter θ.

(d) For what values of α is the asymptotic variance of Med_n smaller than the asymptotic variance of $\bar{X}_n$?

(e) Design and implement an interesting simulation study that supports the theoretical findings in parts (b), (c) and (d).

Problem 38 (asymptotic distribution for the maximum of a sample, 6 points). Let $X_1, X_2, \ldots$ be a sample from a distribution with cumulative distribution function F. For $n = 1, 2, \ldots$, let $M_n = \max\{X_1, \ldots, X_n\}$ be the maximum of the first n observations. For each of the following distributions, find a normalization $(M_n - a_n)/b_n$, if any exists, that has a nondegenerate limiting distribution as $n \to \infty$.

(a) The logistic distribution with Lebesgue density $f(x) = e^{-x}/(1 + e^{-x})^2$.

(b) The standard normal distribution.

(c) The distribution with Lebesgue density $f(x) = 2/x^3$ for $x > 1$.

(d) The cosine distribution with Lebesgue density $f(x) = (1/2)\cos x$ for $-\pi/2 < x < \pi/2$.

Tilmann Gneiting, January 11, 2008. Solutions are due Friday, January 25 at the beginning of the class session.