Large Sample Theory
Ferguson

Exercises, Section 13, Asymptotic Distribution of Sample Quantiles.

1. (a) Find the asymptotic joint distribution of $(X_{(np)}, X_{(n(1-p))})$ when sampling from a Cauchy distribution $\mathcal{C}(\mu, \sigma)$. You may assume $0 < p < 1/2$. See Example 13.2 and Exercise 13.3.

(b) Find the asymptotic distribution of $\hat{\mu}_n = (1/2)(X_{(np)} + X_{(n(1-p))})$.

(c) What value of p minimizes the asymptotic variance of $\hat{\mu}_n$? Compare this estimate to the sample median as an estimate of μ.

2. Let $0 < p_1 < \cdots < p_k < 1$, and let $X_{(\lceil np_i \rceil)}$ be the corresponding sample quantiles for a sample of size n from a distribution with location parameter θ having distribution function $F(x-\theta)$ and density $f(x-\theta)$. Let u_i denote the p_ith quantile of F (i.e. $F(u_i) = p_i$).

(a) Let $Z_i = X_{(\lceil np_i \rceil)} - u_i$. Let $\mathbf{Z}$ represent the vector $(Z_1, \ldots, Z_k)^T$ and $\mathbf{1}$ represent the k-vector of all 1's. Show that $\sqrt{n}(\mathbf{Z} - \theta\mathbf{1}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the symmetric matrix with components $\sigma_{ij} = \dfrac{p_i(1-p_j)}{f(u_i)f(u_j)}$ for $i \le j$.

(b) Find the asymptotic best linear unbiased estimate of θ based on $\mathbf{Z}$. That is, for $\hat{\theta} = \mathbf{a}^T\mathbf{Z}$, find $\mathbf{a}$ to minimize $\mathbf{a}^T\boldsymbol{\Sigma}\mathbf{a}$ subject to $\mathbf{1}^T\mathbf{a} = 1$ (in terms of $\boldsymbol{\Sigma}^{-1}$).

(c) In view of (b), it is comforting to know that the inverse of $\boldsymbol{\Sigma}$ has a simple form. It is a tridiagonal matrix. Find it.

(d) Find $\hat{\theta}$ of (b) explicitly, for the uniform distribution, $F(x) = x$ for $0 \le x \le 1$.

3. Suppose we have a sample, $X_1, \ldots, X_n$, from the family of distributions on the real line with density $f(x|\theta, \alpha) = c(\alpha)e^{-|x-\theta|^\alpha}$, $\alpha > 0$. We may use the sample mean, $\overline{X}_n$, or the sample median, m_n to estimate the location parameter θ.

(a) Find the constant $c(\alpha)$.

(b) What is the asymptotic distribution of $\overline{X}_n$?

(c) What is the asymptotic distribution of m_n?

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

4. Let $X_1, \ldots, X_n$ be a sample from $\mathcal{N}(\theta, \sigma^2)$ with σ^2 known. It is desired to estimate the pth quantile, $x_p = \theta + \sigma z_p$, where z_p is the pth quantile of the standard normal distribution. The maximum likelihood estimate of x_p is clearly $\hat{x}_p = \overline{X}_n + \sigma z_p$. What is the asymptotic distribution of $\sqrt{n}(\hat{x}_p - x_p)$? What is the asymptotic efficiency of $X_{(\lceil np \rceil)}$ relative to $\hat{x}_p$?

5. Let $X_1, \ldots, X_n$ be a sample from the Pareto distribution with density $f(x|\theta) = \theta/(x+\theta)^2$ for $x > 0$, and distribution function $F(x|\theta) = x/(x+\theta)$ for $x > 0$. Let $x_p(\theta)$ denote the pth quantile of the distribution and let $X_{\lceil np \rceil}$ denote the sample pth quantile.
(a) What is the asymptotic distribution of $X_{\lceil np \rceil}$ as $n \to \infty$?
(b) Find a constant $c(p)$ such that $\hat{\theta}_n = c(p)X_{\lceil np \rceil}$ is a consistent asymptotically unbiased estimate of θ. For what value of p is the asymptotic variance of $\hat{\theta}_n$ a minimum?