Large Sample Theory
Ferguson

Exercises, Section 13, Asymptotic Distribution of Sample Quantiles.

1. (a) Find the asymptotic joint distribution of \((X_{(n)p}, X_{(n(1-p))})\) when sampling from a Cauchy distribution \(C(\mu, \sigma)\). You may assume \(0 < p < 1/2\). See Example 13.2 and Exercise 13.3.
   
   (b) Find the asymptotic distribution of \(\hat{\mu}_n = (1/2)(X_{(n)p} + X_{(n(1-p))})\).
   
   (c) What value of \(p\) minimizes the asymptotic variance of \(\hat{\mu}_n\)? Compare this estimate to the sample median as an estimate of \(\mu\).

2. Let \(0 < p_1 < \cdots < p_k < 1\), and let \(X_{([np_i])}\) be the corresponding sample quantiles for a sample of size \(n\) from a distribution with location parameter \(\theta\) having distribution function \(F(x - \theta)\) and density \(f(x - \theta)\). Let \(u_i\) denote the \(p_i\)th quantile of \(F\) (i.e. \(F(u_i) = p_i\)).
   
   (a) Let \(Z_i = X_{([np_i])} - u_i\). Let \(Z\) represent the vector \((Z_1, \ldots, Z_k)^T\) and \(1\) represent the \(k\)-vector of all 1’s. Show that \(\sqrt{n}(Z - \theta 1) \xrightarrow{d} N(0, \Psi)\), where \(\Psi\) is the symmetric matrix with components \(\sigma_{ij} = p_i(1-p_j)/f(u_i)f(u_j)\) for \(i \leq j\).
   
   (b) Find the asymptotic best linear unbiased estimate of \(\theta\) based on \(Z\). That is, for \(\hat{\theta} = a^T Z\), find \(a\) to minimize \(a^T \Psi a\) subject to \(1^T a = 1\) (in terms of \(\Psi^{-1}\)).
   
   (c) In view of (b), it is comforting to know that the inverse of \(\Psi\) 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 \(0 \leq x \leq 1\).

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 \(\theta\).
   
   (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 \(\alpha\) 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 \(N(\theta, \sigma^2)\) with \(\sigma^2\) known. It is desired to estimate the \(p\)th quantile, \(x_p = \theta + \sigma z_p\), where \(z_p\) is the \(p\)th 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 \(X_{([np])}\) 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 \(p\)th quantile of the distribution and let \(X_{([np])}\) denote the sample \(p\)th quantile.
   
   (a) What is the asymptotic distribution of \(X_{([np])}\) as \(n \to \infty\)?
   
   (b) Find a constant \(c(p)\) such that \(\hat{\theta}_n = c(p)X_{([np])}\) is a consistent asymptotically unbiased estimate of \(\theta\). For what value of \(p\) is the asymptotic variance of \(\hat{\theta}_n\) a minimum?