Empirical Process Proof of the Asymptotic Distribution of Sample Quantiles

**Definition:** Given $\theta \in (0, 1)$, the $\theta^{th}$ quantile of a random variable $X$ with CDF $F$ is defined by:

$$\mu_\theta \equiv F^{-1}(\theta) = \inf\{x \mid F(x) \geq \theta\}.$$  

Note that $\mu_{0.5}$ is the median, $\mu_{0.25}$ is the 25th percentile, etc. Further if we define the $\theta^{th}$ quantile as $\mu_\theta = \lim_{\theta \to 0} \mu_\theta$ and define $\mu_1$ similarly, it is easy to see that these are the lower and upper points in the support of $X$ (i.e. the minimum and maximum possible values of $X$ which might be $-\infty$ and $+\infty$ if $X$ has unbounded support). Note also that if $F$ is strictly increasing in a neighborhood of $\mu_\theta$, then $\mu_\theta = F^{-1}(\theta)$ is the usual inverse of the CDF $F$. If $F$ happens to have "flat" sections, say an interval of points $x$ satisfying $F(x) = \theta$, then $\mu_\theta$ is the smallest $x$ in this interval. The following lemma, a slightly modified version of a lemma from R. J. Serfling, (1980) *Approximation Theorems of Mathematical Statistics* Wiley, New York, provides some basic properties of the quantile function $F^{-1}(\theta)$:

**Lemma 1:** Let $F$ be a CDF. The quantile function $F^{-1}(\theta)$, $\theta \in (0, 1)$ is non-decreasing and left continuous, and satisfies:

1. $F^{-1}(F(x)) \leq x$, $-\infty < x < \infty$
2. $F(F^{-1}(\theta)) \geq \theta$, $0 < \theta < 1$
3. If $F$ is strictly increasing in a neighborhood of $\mu_\theta = F^{-1}(\theta)$ we have: $F(F^{-1}(\theta)) = \theta$ and $F^{-1}(F(\mu_\theta)) = \mu_\theta$.
4. $F(x) \geq \theta$ if and only if $x \geq F^{-1}(\theta)$.

**Definition:** Let $(\tilde{X}_1, \ldots, \tilde{X}_N)$ be a random sample of size $N$ from a CDF $F$. Then the sample quantile $\hat{\mu}_\theta$, $\theta \in (0, 1)$ is defined by:

$$\hat{\mu}_\theta = F_N^{-1}(\theta),$$

where $F_N$ is the empirical CDF defined by:

$$F_N(x) = \frac{1}{N} \sum_{i=1}^{N} I\{\tilde{X}_i \leq x\}.$$  

Thus $\hat{\mu}_{0.5} = F_N^{-1}(.5)$ is the sample median $F_N^{-1}(.5) = \text{med}(\tilde{X}_1, \ldots, \tilde{X}_N)$, and $\hat{\mu}_0 = F_N^{-1}(0)$ is the sample minimum, $F_N^{-1}(0) = \min(\tilde{X}_1, \ldots, \tilde{X}_N)$, and $\hat{\mu}_1 = F_N^{-1}(1)$ is the sample maximum, $F_N^{-1}(1) = \max(\tilde{X}_1, \ldots, \tilde{X}_N)$. Since empirical CDF's have jumps of size $1/N$ (unless more than one of the $\{\tilde{X}_i\}$'s take the same value), then we can bound the maximum difference between $\theta$ and $F(F^{-1}(\theta))$ in Lemma 1-2 as follows:
**Lemma 2:** Let \((\bar{X}_1, \ldots, \bar{X}_N)\) be a random sample from a CDF \(F\) and suppose that in this sample each \(\bar{X}_i\) happens to be distinct, so that by reindexing we have \(\bar{X}_1 < \bar{X}_2 < \cdots < \bar{X}_{N-1} < \bar{X}_N\). Then for all \(\theta \in (0, 1)\) we have:

\[
|F_N(F_N^{-1}(\theta)) - \theta| \leq \frac{1}{N}.
\]

The following theorem shows that the asymptotic distribution of the sample quantiles \(\hat{\mu}_\theta\) for \(\theta \in (0, 1)\) are normally distributed. It is important to note that we exclude the two cases \(\theta = 0\) and \(\theta = 1\) in this theorem since the asymptotic distribution of these *extreme value statistics* is very different and generally non-normal.

**Theorem:** Let \((\bar{X}_1, \ldots, \bar{X}_N)\) be IID draws from a CDF \(F\) with continuous density \(f\). Then if \(f(\mu_0) > 0\), we have:

\[
\sqrt{N}(\hat{\mu}_\theta - \mu_0) = \sqrt{N}(F_N^{-1}(\theta) - F^{-1}(\theta)) \implies N(0, \sigma^2),
\]

where:

\[
\sigma^2 = \frac{\theta(1 - \theta)}{f(\mu_0)^2}.
\]

**Proof:** The Central Limit Theorem for IID random variables implies that for any \(x\) in the support of \(F\) we have:

\[
\sqrt{N}(F_N(x) - F(x)) \implies N(0, \gamma^2),
\]

where \(\gamma^2 = F(x)[1 - F(x)]\). Letting \(x = \mu_0 = F^{-1}(\theta)\) and using Lemma 1-3 we have:

\[
\sqrt{N}(F_N(\mu_0) - F(\mu_0)) = \sqrt{N}(F_N(F_N^{-1}(\theta)) - F(F^{-1}(\theta))) \implies N(0, \theta(1 - \theta)).
\]

Furthermore, the property of *stochastic equicontinuity* from the theory of *empirical processes* (see D. Andrews, (1996) *Handbook of Econometrics* (vol. 4) for an accessible introduction and definition of stochastic equicontinuity), we have that the result given above is unaffected if we replace \(\mu_0\) by a consistent estimate \(\hat{\mu}_0\):

\[
\sqrt{N}(F_N(\hat{\mu}_0) - F(\hat{\mu}_0)) = \sqrt{N}(F_N(F_N^{-1}(\theta)) - F(F^{-1}(\theta))) \implies N(0, \theta(1 - \theta)).
\]

Now note that Lemma 1-2 implies that:

\[
\sqrt{N}(F_N(F_N^{-1}(\theta)) - F(F_N^{-1}(\theta))) \geq \sqrt{N}(\theta - F(F_N^{-1}(\theta))).
\]

However since the true CDF \(F\) has a density, the probability of observing duplicate \(\{\bar{X}_i\}\)'s is zero, so Lemma 2 implies that with probability 1 we have:

\[
\sqrt{N}(F_N(F_N^{-1}(\theta)) - F(F_N^{-1}(\theta))) = \sqrt{N}(\theta - F(F_N^{-1}(\theta))) + O_p(1/\sqrt{N}),
\]

which implies that:

\[
\sqrt{N}(\theta - F(F_N^{-1}(\theta))) \implies N(0, \theta(1 - \theta)).
\]

Now we apply the Delta theorem, i.e. we do a Taylor series expansion of \(F(F_N^{-1}(\theta))\) about the limiting point \(\mu_0 = F^{-1}(\theta)\) to get:

\[
F(F_N^{-1}(\theta)) = F(F^{-1}(\theta)) + f(\hat{\mu}_0)[F_N^{-1}(\theta) - F^{-1}(\theta)],
\]
where $\tilde{\mu}_\theta$ is a point on the line segment between $\hat{\mu}_\theta$ and $\mu_\theta$. Using the result above and Lemma 1-3 we have:

$$\sqrt{N} \left( F^{-1}_N(\theta) - F^{-1}(\theta) \right) = \sqrt{N} \left( \frac{\theta - F(F^{-1}_N(\theta))}{f(\tilde{\mu}_\theta)} \right) \Rightarrow N(0, \sigma^2),$$

where we have used Slutsky’s Theorem and the fact that $\tilde{\mu}_\theta \to \mu_\theta$ since $\hat{\mu}_\theta \to \mu_\theta$ with probability 1.