556: Mathematical Statistics I

ASYMPTOTIC DISTRIBUTION OF SAMPLE QUANTILES

Suppose $X_1, \ldots, X_n$ are i.i.d. continuous random variables from distribution with cdf F_X. Let $Y_n(x)$ be a random variable defined for fixed $x \in \mathbb{R}$ by

$$Y_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \leq x\} = \frac{1}{n}\sum_{i=1}^{n} Z_i$$

where $Z_i(x) = I\{X_i \geq x\} = 1$ if $X \leq x$, and zero otherwise. Then Z_i has expectation $\mu(x) = F_X(x)$ and variance $\sigma^2(x) = F_X(x)\{1 - F_X(x)\}$, and by the Central Limit Theorem

$$\sqrt{n}(Y_n(x) - F_X(x)) \xrightarrow{d} X \sim N(0, F_X(x)\{1 - F_X(x)\}).$$

Now consider the transformation through function $g(t)$ defined for $0 < t < 1$ by $g(t) = F_X^{-1}(t)$. We have the first derivative of g as

$$g^{(1)}(t) = \frac{d}{dt}\{F_X^{-1}(t)\} = \frac{1}{f_X(F_X^{-1}(t))}$$

as

$$y = F_X^{-1}(t) \quad \Longleftrightarrow \quad F_X(y) = t \quad \Longrightarrow \quad f_X(y)dy = dt \quad \Longrightarrow \quad \frac{dy}{dt} = \frac{1}{f_X(y)} = \frac{1}{f_X(F_X^{-1}(t))}$$

Thus, using the Delta method

$$\sqrt{n}(F_X^{-1}(Y_n(x)) - F_X^{-1}(F_X(x))) \xrightarrow{d} X \sim N\left(0, \frac{F_X(x)\{1 - F_X(x)\}}{\{f_X(F_X^{-1}(F_X(x)))\}^2}\right).$$

and writing $p = F_X(x)$, we have

$$\sqrt{n}(F_X^{-1}(Y_n(x)) - x) \xrightarrow{d} X \sim N\left(0, \frac{p(1-p)}{\{f_X(x)\}^2}\right).$$

Now $F_X^{-1}(Y_n(x))$ is a random variable that lies between the $(p-1)$st and pth sample quantile, that can be written using via order statistic notation as $X_{([np])}$. In fact,

$$|X_{([np])} - F_X^{-1}(Y_n(x))| \xrightarrow{a.s.} 0.$$

It follows that

$$\sqrt{n}(X_{([np])} - x) \xrightarrow{d} X \sim N\left(0, \frac{p(1-p)}{\{f_X(x)\}^2}\right).$$

EXAMPLE: For the sample median, $\tilde{X}_n$, from a symmetric distribution with location θ, where the distribution median is θ, we consider $x = \theta$ and $p = F_X(\theta) = 1/2$, so

$$\sqrt{n}(\tilde{X}_n - \theta) \xrightarrow{d} X \sim N\left(0, \frac{1}{4\{f_X(\theta)\}^2}\right).$$