

Lecture 44: Sample quantiles: consistency and asymptotic normality

Statistical Inference–II ■ Units IV–V: Nonparametric Inference

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

Previous lecture

Lecture 43: Order statistics; the distribution of $X_{(r)}$.

Today

- Order Statistics and Estimation of Quantiles
 - Consistency and Asymptotic Normality

Theorem 44.1 (Consistency of a sample quantile)

Suppose that F is continuous and strictly increasing in a neighbourhood of ξ_p . If $r_n/n \rightarrow p$, then

$$X_{(r_n)} \xrightarrow{P} \xi_p.$$

Proof.

For real x let $N_n(x) = \sum_{i=1}^n \mathbf{1}(X_i \leq x) \sim \text{Bin}\{n, F(x)\}$, so that $\{X_{(r_n)} \leq x\} = \{N_n(x) \geq r_n\}$. Fix $\varepsilon > 0$ small enough that F is strictly increasing on $[\xi_p - \varepsilon, \xi_p + \varepsilon]$. By continuity $F(\xi_p) = p$, hence

$$F(\xi_p - \varepsilon) < p < F(\xi_p + \varepsilon).$$

Proof (cont.)

By the weak law of large numbers $N_n(x)/n \xrightarrow{P} F(x)$ for each fixed x , while $r_n/n \rightarrow p$. Therefore

$$\mathbb{P}\{X_{(r_n)} > \xi_p + \varepsilon\} = \mathbb{P}\left\{\frac{N_n(\xi_p + \varepsilon)}{n} < \frac{r_n}{n}\right\} \rightarrow 0,$$

$$\mathbb{P}\{X_{(r_n)} \leq \xi_p - \varepsilon\} = \mathbb{P}\left\{\frac{N_n(\xi_p - \varepsilon)}{n} \geq \frac{r_n}{n}\right\} \rightarrow 0,$$

since in the first line the left side converges to $F(\xi_p + \varepsilon) > p$ and in the second to $F(\xi_p - \varepsilon) < p$. Hence $\mathbb{P}\{|X_{(r_n)} - \xi_p| > \varepsilon\} \rightarrow 0$. □

Theorem 44.2 (Asymptotic distribution of a sample quantile)

Suppose F has a positive and continuous density f at ξ_p . Then

$$\sqrt{n} \left(\hat{\xi}_p - \xi_p \right) \xrightarrow{d} N \left(0, \frac{p(1-p)}{f(\xi_p)^2} \right).$$

Consequently an approximate standard error is

$$\text{se}(\hat{\xi}_p) \approx \frac{\sqrt{p(1-p)}}{\sqrt{n} f(\xi_p)}.$$

Heuristic explanation.

Put $r = \lceil np \rceil$. By the distribution of $F\{X_{(r)}\}$ found above, $F(\hat{\xi}_p) = U_{(r)} \sim \text{Beta}(r, n - r + 1)$, whose mean is $r/(n + 1) \approx p$ and whose variance is

$$\frac{r(n - r + 1)}{(n + 1)^2(n + 2)} \approx \frac{p(1 - p)}{n}.$$

Since F has derivative $f(\xi_p) > 0$ at ξ_p , a first-order Taylor expansion gives $F(\hat{\xi}_p) - p \approx f(\xi_p)(\hat{\xi}_p - \xi_p)$, that is,

$$\hat{\xi}_p - \xi_p \approx \frac{U_{(r)} - p}{f(\xi_p)}, \quad \text{so} \quad \text{Var}(\hat{\xi}_p) \approx \frac{p(1 - p)}{n f(\xi_p)^2}.$$

Asymptotic normality of $U_{(r)}$ completes the argument. (Rigorous proof: Lecture Notes, Appendix A.) □

The formula

$$\text{se}(\hat{\xi}_p) \approx \frac{\sqrt{p(1-p)}}{\sqrt{n} f(\xi_p)}$$

shows that quantile estimation is more difficult when the density is small near the target quantile.

Since $f(\xi_p)$ is unknown, a large-sample interval requires a density estimate.

The exact order-statistic interval in the next section avoids this nuisance quantity entirely.

Example 44.3 (Asymptotic variance of the sample median)

For $p = 1/2$,

$$\sqrt{n}(\hat{\xi}_{1/2} - \xi_{1/2}) \xrightarrow{d} N\left(0, \frac{1}{4f(\xi_{1/2})^2}\right).$$

If $X \sim N(\mu, \sigma^2)$, then $f(\mu) = 1/(\sigma\sqrt{2\pi})$, so the asymptotic variance of the sample median is $\pi\sigma^2/(2n)$. The sample mean has variance σ^2/n ; therefore the asymptotic relative efficiency of the median to the mean under normality is $2/\pi \approx 0.637$.

Practice (Lecture Notes, Chapter 3)

- Exercises (Section 3.15): 3.7.

Next lecture

Lecture 45: Distribution-free confidence intervals for quantiles.