

Lecture 45: Distribution-free confidence intervals for quantiles

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 44: Sample quantiles: consistency and asymptotic normality.

Today

- Distribution-Free Confidence Intervals for Population Quantiles

Exact confidence interval for a quantile

Let ξ_p be a unique p th quantile of a continuous distribution. Define

$$Y = \sum_{i=1}^n \mathbf{1}(X_i < \xi_p).$$

Then $Y \sim \text{Bin}(n, p)$. This elementary fact yields an exact, distribution-free confidence interval.

Theorem 45.1 (Exact confidence interval for a quantile)

For integers $1 \leq r < s \leq n$,

$$\mathbb{P}\{X_{(r)} \leq \xi_p \leq X_{(s)}\} = \sum_{j=r}^{s-1} \binom{n}{j} p^j (1-p)^{n-j}.$$

Hence $[X_{(r)}, X_{(s)}]$ is a distribution-free confidence interval whose coverage is the binomial probability on the right.

Proof.

The event $X_{(r)} \leq \xi_p \leq X_{(s)}$ occurs precisely when at least r observations are below ξ_p but fewer than s observations are below it; that is, when $r \leq Y \leq s - 1$. Since $Y \sim \text{Bin}(n, p)$, the stated probability follows. $\square$

For a central interval, choose r and s so that each omitted binomial tail, $\mathbb{P}(Y \leq r - 1)$ and $\mathbb{P}(Y \geq s)$, is at most $\alpha/2$ and as large as possible; the attained confidence coefficient is then at least $1 - \alpha$. Because the binomial distribution is discrete, it is usually strictly larger than $1 - \alpha$.

Example: confidence interval for a population median

Example 45.2 (Confidence interval for a population median)

For $n = 20$, find a central distribution-free confidence interval for the population median η with confidence coefficient at least 0.95.

Solution

Here $p = 1/2$ and $Y \sim \text{Bin}(20, \frac{1}{2})$ is symmetric about 10, so $s = n + 1 - r = 21 - r$. Since $\mathbb{P}(Y \leq 5) = 0.0207 \leq 0.025 < \mathbb{P}(Y \leq 6) = 0.0577$, the largest admissible r is 6, and the interval $[X_{(6)}, X_{(15)}]$ has confidence coefficient

$$\sum_{j=6}^{14} \binom{20}{j} 2^{-20} = 1 - 2(0.0207) = 0.9586.$$

Thus it is an exact distribution-free interval with at least 95% coverage for the population median.

Example: confidence interval for an upper quartile

Example 45.3 (Confidence interval for an upper quartile)

A random sample of $n = 40$ observations is taken from a continuous distribution. Find a distribution-free confidence interval for the upper quartile $\xi_{0.75}$ with confidence coefficient at least 0.95.

Solution

Now $Y = \#\{i : X_i < \xi_{0.75}\} \sim \text{Bin}(40, 0.75)$, with mean 30 and standard deviation $\sqrt{7.5} = 2.739$. This distribution is not symmetric, so the two tails must be handled separately:

k	22	23	24
$\mathbb{P}(Y \leq k)$	0.0047	0.0116	0.0262

k	35	36	37
$\mathbb{P}(Y \geq k)$	0.0433	0.0160	0.0047

Example: confidence interval for an upper quartile (cont.)

Solution (cont.)

Lower end. The largest r with $\mathbb{P}(Y \leq r - 1) \leq 0.025$ is $r = 24$, since $\mathbb{P}(Y \leq 23) = 0.0116 \leq 0.025 < \mathbb{P}(Y \leq 24) = 0.0262$.

Upper end. The smallest s with $\mathbb{P}(Y \geq s) \leq 0.025$ is $s = 36$, since $\mathbb{P}(Y \geq 36) = 0.0160 \leq 0.025 < \mathbb{P}(Y \geq 35) = 0.0433$.

Interval. By the Theorem (Exact confidence interval for a quantile), $[X_{(24)}, X_{(36)}]$ has confidence coefficient

$$\mathbb{P}(24 \leq Y \leq 35) = 1 - 0.0116 - 0.0160 = 0.9724 \quad (\geq 0.95),$$

for every continuous distribution. The point estimate is $\hat{\xi}_{0.75} = X_{(30)}$.

Example: confidence interval for an upper quartile (cont.)

Solution (cont.)

The two omitted tails (0.0116 and 0.0160) are unequal and both below 0.025: discreteness and the skewness of $\text{Bin}(40, 0.75)$ make the attained coefficient noticeably larger than 0.95.

A more extreme quantile. For $\xi_{0.90}$ with $n = 50$ the same rule gives $r = 41$ ($\mathbb{P}(Y \leq 40) = 0.0245$) and $s = 50$ ($\mathbb{P}(Y \geq 50) = 0.9^{50} = 0.0052$), i.e. $[X_{(41)}, X_{(50)}]$ with coefficient 0.9703.

The upper end is already the sample maximum: high quantiles need large samples (cf. Exercise 3.6).

Practice (Lecture Notes, Chapter 3)

- Short-answer questions (Section 3.14): Q4.
- Exercises (Section 3.15): 3.5.

Next lecture

Lecture 46: One-sided bounds; duality with the sign test; principles of nonparametric testing.