

Lecture 41: Nonparametric and distribution-free procedures; assumptions that still matter

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

Units IV–V: Nonparametric Inference

Lecture 41 begins Units IV–V: Nonparametric Inference.

Today

- Introduction to Nonparametric Inference
 - Why Nonparametric Methods are Needed
 - Basic Assumptions that Still Matter

In a parametric analysis the population distribution is assumed to belong to a specified family, such as the normal, exponential or gamma family, and the inferential problem is reduced to statements about a finite-dimensional parameter.

Such assumptions are useful when they are scientifically plausible, but a procedure derived under a parametric model may lose its stated size, coverage or efficiency when the model is seriously violated.

Nonparametric inference develops procedures that require fewer assumptions about the functional form of the population distribution.

Many of these procedures depend only on signs, ranks, order statistics or empirical distribution functions.

Their null distributions are then free of the unknown continuous population distribution, which is why they are often called *distribution-free* methods.

Standard systematic accounts include Hollander, Wolfe and Chicken (2014), Gibbons and Chakraborti (2021) and Conover (1999).

Definition 41.1 (Nonparametric procedure)

A statistical procedure is called *nonparametric* when its formulation does not require the population distribution to be specified by a finite-dimensional parameter vector. A test is called *distribution-free* under a null hypothesis when its null distribution is the same for every continuous population distribution satisfying that hypothesis.

The terms are closely related but not identical. A nonparametric estimator may have a sampling distribution that still depends on the underlying population, whereas a rank test may be exactly distribution-free under continuity and exchangeability.

Nonparametric methods are particularly useful when:

- (i) the measurement scale is ordinal rather than genuinely quantitative;
- (ii) the population is skewed, heavy-tailed or contaminated by outliers;
- (iii) the sample size is too small to assess a detailed parametric model;
- (iv) the inferential target is a median, quantile or general distribution function rather than a mean; or
- (v) robustness and exact finite-sample validity are more important than efficiency under one special model.

Advantages. Weak distributional assumptions, applicability to ranks and signs, resistance to extreme observations and the availability of many exact small-sample tests.

Limitations. Loss of information when numerical observations are reduced to signs or ranks, complications due to ties, and the fact that different procedures test different population features.

For example, the sign test concerns a median, whereas the Wilcoxon signed-rank test is naturally a test of a symmetric location distribution.

The word “nonparametric” does not mean “assumption-free.”

Depending on the procedure, one may still require

- random sampling,
- independence,
- continuity,
- identical distributions,
- exchangeability,
- symmetry, or
- a common-shape location model.

These assumptions must be stated before a test statistic is interpreted.

Practice (Lecture Notes, Chapter 3)

- Short-answer questions (Section 3.14): Q1.
- Exercises (Section 3.15): 3.1.

Next lecture

Lecture 42: Empirical distribution function; probability integral transformation.