

Approaches to Nonparametric Statistical Analysis

Methods, Probability Distributions, and Hypothesis Tests

1. Introduction

Nonparametric statistics refers to a broad family of methods that make few or no assumptions about the parameters or functional form of the population distribution from which the data were sampled. Where classical (“parametric”) procedures such as the t-test or ANOVA assume the underlying data follow a known distribution — typically the normal distribution — characterized by a fixed set of parameters (mean μ , variance σ^2), nonparametric procedures instead work directly with the ordering, ranks, signs, or empirical distribution of the observations. This makes them especially valuable when sample sizes are small, when data are ordinal or measured on a scale that does not support arithmetic means (for example, Likert-type ratings), when the data contain outliers, or when the true distributional shape is unknown or clearly non-normal.

This summary surveys three things: (1) the major estimation and modeling approaches used in nonparametric analysis, (2) the probability distributions that underlie nonparametric inference, and (3) the specific hypothesis tests most commonly applied, together with the formulas that govern them.

2. Core Approaches and Methods

2.1 Rank-Based Methods

Rank-based methods replace raw observations with their ranks within the combined sample. Because ranks are invariant to any monotonic transformation of the data, tests built on ranks are robust to outliers and skewness. The rank transform is the foundation of the Wilcoxon, Mann-Whitney, Kruskal-Wallis, Friedman, and Spearman procedures discussed in Section 4.

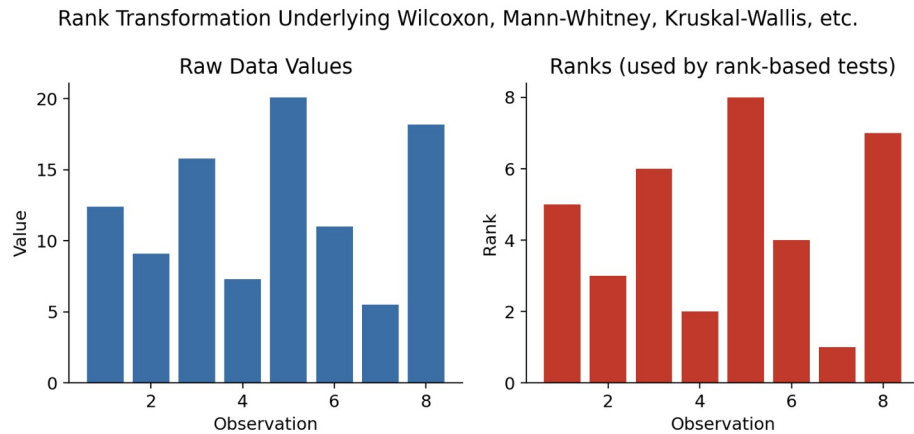


Figure 1. Raw observations are converted to ranks before rank-based tests are applied.

2.2 Empirical Distribution Function (EDF) Methods

The empirical distribution function estimates the population CDF directly from the sample without assuming a parametric family:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x) \quad (1)$$

where $\mathbb{1}(\cdot)$ is the indicator function. By the Glivenko–Cantelli theorem, $F_n(x)$ converges uniformly to the true CDF $F(x)$ as $n \rightarrow \infty$. The EDF underlies goodness-of-fit tests such as the Kolmogorov–Smirnov and Cramér–von Mises tests.

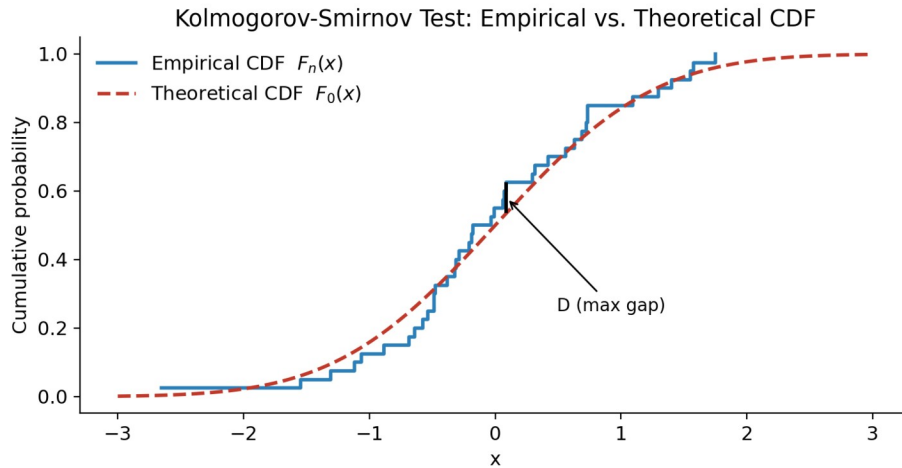


Figure 2. The empirical CDF (solid) approximates the theoretical CDF (dashed); the largest vertical gap D is the Kolmogorov–Smirnov statistic.

2.3 Kernel Density Estimation (KDE)

KDE is a nonparametric way to estimate a continuous probability density without assuming a parametric shape. Each observation contributes a small “bump” (kernel), and the estimate is the sum of these bumps:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (2)$$

Here $K(\cdot)$ is the kernel function (commonly Gaussian) and $h > 0$ is the bandwidth, which controls the smoothness of the estimate. A small h produces a noisy, highly detailed estimate; a large h over-smooths and can obscure real structure.

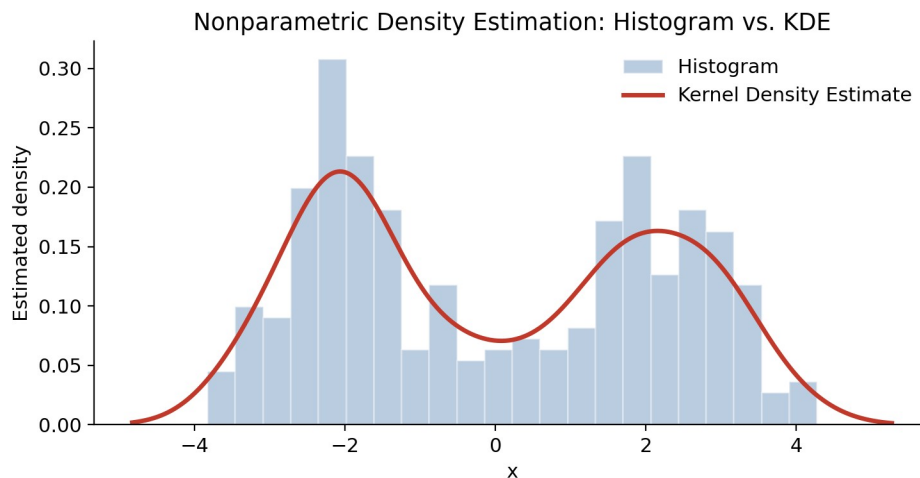


Figure 3. A histogram (bars) and a kernel density estimate (red curve) of the same bimodal sample.

2.4 Resampling Methods: Bootstrap and Permutation

Resampling methods build the sampling distribution of a statistic empirically rather than from theoretical assumptions.

- Bootstrap: repeatedly resamples the observed data with replacement to approximate the sampling distribution of a statistic (mean, median, correlation, etc.) and to construct confidence intervals. The bootstrap standard error of an estimator $\hat{\theta}$ over B resamples is:

$$\widehat{SE}_{\text{boot}} = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}^*)^2} \quad (3)$$

- Permutation (randomization) tests: repeatedly reshuffle group labels to build the null distribution of a test statistic directly from the data; this is exact under the null hypothesis of exchangeability.

Both approaches are considered nonparametric because they do not require specifying a parametric family for the data-generating process.

2.5 Smoothing and Nonparametric Regression

Nonparametric regression estimates a mean function without assuming a linear or otherwise parametric form, using local averaging techniques such as:

- Moving averages and local regression (LOESS / LOWESS)
- Nadaraya–Watson kernel regression
- Smoothing splines and regression splines

These methods let the data determine the shape of a relationship between variables rather than imposing it in advance.

3. Probability Distributions Used in Nonparametric Inference

Although nonparametric methods avoid assuming a distribution for the raw data, the test statistics they produce (ranks, sums of ranks, counts, maximum deviations) often converge to well-known reference distributions under the null hypothesis. These reference distributions are used to obtain p-values.

3.1 Chi-Square (χ^2) Distribution

The chi-square distribution with k degrees of freedom arises as the null distribution of the Kruskal-Wallis test, the Friedman test, and chi-square goodness-of-fit and independence tests:

$$f(x; k) = \frac{1}{2^{k/2} \Gamma(k/2)} x^{k/2-1} e^{-x/2}, \quad x > 0 \quad (4)$$

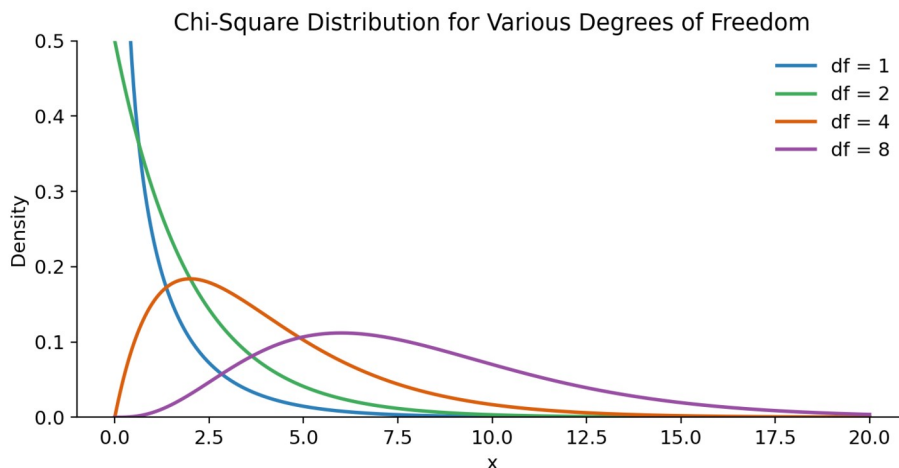


Figure 4. Chi-square probability density for several degrees of freedom.

3.2 F-Distribution

The F-distribution is the ratio of two independent chi-square variables, each divided by its own degrees of freedom. It appears in some rank-transformed ANOVA-type nonparametric procedures and in Levene's test for equality of variances:

$$F = \frac{\chi_1^2/d_1}{\chi_2^2/d_2} \quad (5)$$

3.3 Standard Normal Distribution (Large-Sample Approximation)

For moderately large samples, many rank-sum statistics (Mann-Whitney U, Wilcoxon signed-rank W) are approximately normally distributed, which allows a z-test approximation instead of exact tables. See formula (8) in Section 4.1.

3.4 Discrete Uniform Distribution of Ranks

Under the null hypothesis of no difference between groups, the ranks assigned to the combined sample are, by construction, uniformly distributed over the integers 1 through n — this is the combinatorial basis for exact rank-test p-values:

$$R_i \sim \text{DiscreteUniform}\{1, 2, \dots, n\} \text{ under } H_0$$

3.5 Binomial Distribution

The binomial distribution governs sign-based procedures, most notably the Sign test, where each observation is reduced to a plus or minus sign and the count of positive signs follows:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n - k} \quad (6)$$

4. Hypothesis Tests for Nonparametric Data

The table below summarizes the most widely used nonparametric hypothesis tests, followed by the formula and typical use case for each.

Test	Purpose	Parametric Analogue
Mann-Whitney U (Wilcoxon rank-sum)	Compare two independent groups	Independent-samples t-test
Wilcoxon signed-rank	Compare two paired / matched samples	Paired-samples t-test
Kruskal-Wallis H	Compare three or more independent groups	One-way ANOVA
Friedman test	Compare three or more repeated measures	Repeated-measures ANOVA
Sign test	Compare paired samples using signs only	Paired-samples t-test
Spearman rank correlation	Association between two ranked variables	Pearson correlation
Kolmogorov-Smirnov (K-S)	Goodness-of-fit / compare two distributions	—
Chi-square goodness-of-fit / independence	Compare observed vs. expected category counts	—
Runs test	Test randomness / independence of a sequence	—

4.1 Mann-Whitney U Test (Wilcoxon Rank-Sum Test)

Used to compare two independent groups when the outcome is ordinal or continuous but not normally distributed. All observations from both groups are ranked together, and R_1 is the sum of ranks in group 1:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad (7)$$

For larger samples (typically $n_1, n_2 > 10$), U is approximately normal, giving a z-test:

$$z = \frac{U - \mu_U}{\sigma_U}, \quad \mu_U = \frac{n_1 n_2}{2}, \quad \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \quad (8)$$

4.2 Wilcoxon Signed-Rank Test

Used for paired data. Differences $d_i = x_i - y_i$ are computed, ranked by absolute value, and the ranks are split by sign; the test statistic is the smaller of the two rank sums:

$$W_+ = \sum_{i: d_i > 0} R_i, \quad W_- = \sum_{i: d_i < 0} R_i, \quad T = \min(W_+, W_-) \quad (9)$$

4.3 Kruskal-Wallis H Test

An extension of the Mann-Whitney test to $k \geq 3$ independent groups. N is the total sample size, n_j is the size of group j , and R_j is the sum of ranks in group j :

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1) \quad (10)$$

Under H_0 , H follows an approximate chi-square distribution with $k - 1$ degrees of freedom (see Figure 4).

4.4 Friedman Test

The rank-based analogue of repeated-measures ANOVA for k related conditions measured on n subjects; R_j is the rank sum for condition j :

$$\chi_F^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1) \quad (11)$$

Under H_0 , χ_F^2 is approximately chi-square distributed with $k - 1$ degrees of freedom.

4.5 Sign Test

The simplest paired-sample test: only the sign of each difference d_i is used, ignoring magnitude entirely. Under H_0 (median difference = 0), the number of positive signs follows:

$$X \sim \text{Binomial}\left(n, \frac{1}{2}\right), \quad n = \text{number of non-zero } d_i$$

4.6 Spearman Rank Correlation

Measures monotonic association between two variables using the ranks of the observations rather than raw values. Here d_i is the difference between the paired ranks of the two variables for observation i :

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (12)$$

4.7 Kolmogorov-Smirnov (K-S) Test

Tests whether a sample comes from a specified distribution (one-sample) or whether two samples come from the same distribution (two-sample), using the largest vertical distance between CDFs:

$$D_n = \sup_x |F_n(x) - F_0(x)| \quad (13)$$

See Figure 2 for a visual illustration of D.

4.8 Chi-Square Goodness-of-Fit and Independence Tests

Compares observed category counts O_i to counts E_i expected under a null model, across k categories or cells:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (14)$$

The statistic follows a chi-square distribution with degrees of freedom depending on the number of categories (goodness-of-fit) or the table dimensions (independence).

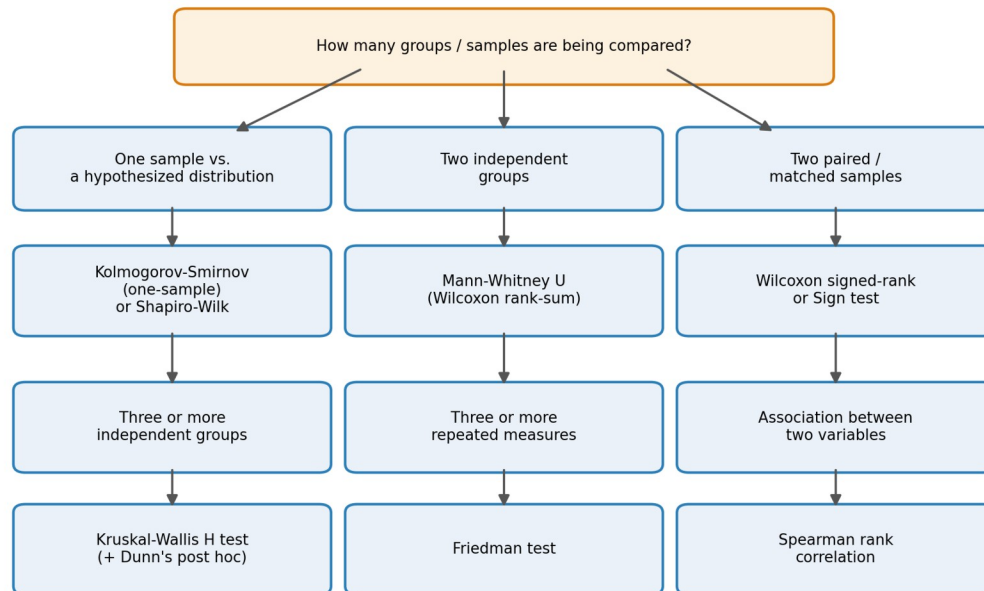
4.9 Runs Test (Wald–Wolfowitz)

Tests whether a binary sequence is random by counting the number of “runs” (consecutive identical values). With n_1 and n_2 the counts of each symbol, the expected number of runs and its standard deviation are:

$$\mu_R = \frac{2n_1n_2}{n_1 + n_2} + 1, \quad \sigma_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \quad (15)$$

5. Choosing the Right Test

Test selection depends mainly on the number of groups being compared, whether samples are independent or paired, and whether the goal is to compare distributions, locations, or associations. The flowchart below summarizes this decision process for the tests covered in Section 4.



Selecting a Nonparametric Hypothesis Test

Figure 5. A decision flowchart for selecting an appropriate nonparametric hypothesis test.

6. Conclusion

Nonparametric methods provide a flexible and robust toolkit for analyzing data that do not meet the distributional assumptions required by classical parametric procedures. The approaches surveyed here — rank

transformation, empirical distribution functions, kernel density estimation, resampling, and nonparametric regression — share a common philosophy: let the data reveal their own structure rather than imposing a fixed parametric form. The associated hypothesis tests (Mann-Whitney, Wilcoxon signed-rank, Kruskal-Wallis, Friedman, Sign, Spearman, Kolmogorov-Smirnov, chi-square, and Runs) extend the logic of their parametric counterparts to ranks, signs, and empirical distributions, relying on reference distributions such as the chi-square, F, standard normal, uniform, and binomial distributions to determine statistical significance. Because they trade some statistical power (when parametric assumptions genuinely hold) for robustness and fewer assumptions, nonparametric methods are generally preferred whenever sample sizes are small, data are ordinal, distributions are heavily skewed, or outliers are present.

Quick Reference: Formula Index

#	Formula	Used In
1	Empirical distribution function $F_n(x)$	§2.2 EDF Methods
2	Kernel density estimate $\hat{f}_h(x)$	§2.3 KDE
3	Bootstrap standard error	§2.4 Resampling
4	Chi-square PDF	§3.1 / multiple tests
5	F statistic	§3.2 F-distribution
6	Binomial PMF	§3.5 / Sign test
7	Mann-Whitney U_1	§4.1
8	Mann-Whitney z-approximation	§4.1
9	Wilcoxon signed-rank W^+ , W^- , T	§4.2
10	Kruskal-Wallis H	§4.3
11	Friedman χ^2_F	§4.4
12	Spearman ρ_s	§4.6
13	Kolmogorov-Smirnov D_n	§4.7
14	Chi-square goodness-of-fit	§4.8
15	Runs test mean / SD	§4.9