STAT 824 sp 2025 Lec 00 slides

What is nonparametric inference?

Karl Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Nonparametric inference

What is nonparametric inference?

Methods based on ranks form a substantial body of statistical techniques that provide alternatives to the classical parametric methods...the feature of nonparametric methods mainly responsible for their great popularity (and to which they owe their name) is the weak set of assumptions required for their validity.

– E.L. Lehmann in his 1975 book *Nonparametrics* [1]

What is nonparametric inference?

The basic idea of nonparametric inference is to use data to infer an unknown quantity while making as few assumptions as possible. Usually, this means using statistical models that are infinite-dimensional.

      – Larry Wasserman in his 2006 book *All of Nonparametric Statistics* [3]

What is nonparametric inference?

The problem of nonparametric estimation consists in estimation, from the observations, of an unknown function belonging to a sufficiently large class of functions.

– Alexander Tsybakov in his 2008 book *Intro. to Nonparametric Estimation* [2].

Nonparametric inference

Nonparametric regression

Kernel density estimation
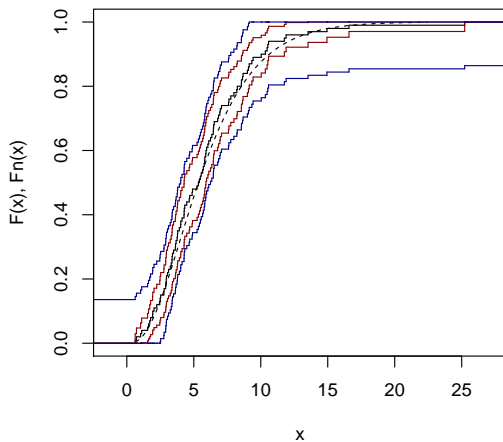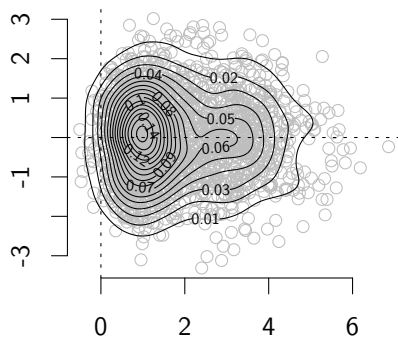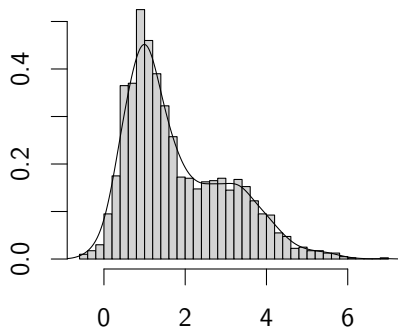
Minimax theory

Bootstrap

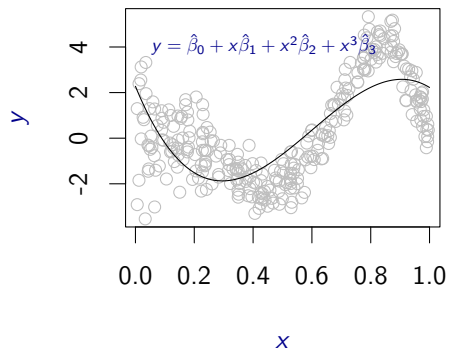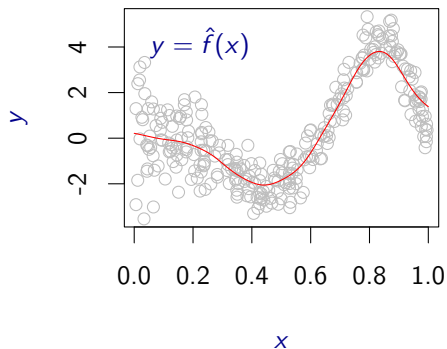cdf estimation

trad nonparm

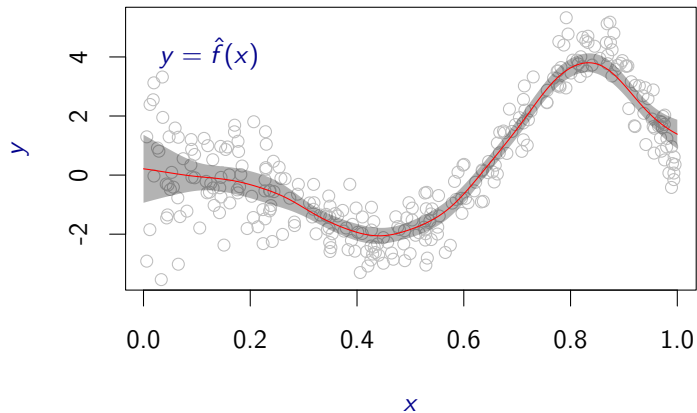# CDF estimation

Kernel density estimation

Nonparametric regression

Minimax theory

Bootstrap

Traditional nonparametrics

Lehmann (1975) says "rank-based statistics constitute methodologically the most important part" of nonparametric statistics. Some of these are:

1. Wilcoxon rank-sum test for a difference in location.
2. Siegel-Tukey test for a difference in dispersion.
3. Sign test for paired comparisons.
4. Wilcoxon signed-rank test for paired comparisons.
5. The Kruskall-Wallace test for comparing several treatments.

Wasserman (2006) classes the above as "traditional" nonparametric methods.

Overview of STAT 824 topics

1. CDF estimation: Glivenko-Cantelli, Dvoretzky-Kiefer-Wolfowitz and Hoeffding inequalities, Kolmogorov-Smirnov-Donsker, Brownian motion, Brownian bridge, KS test.

2. Density estimation: Estimating the cdf, kernel density estimation (KDE), Lipschitz and Hölder classes of densities, bounds on the mean squared error (MSE) of kernel density estimators, bounds on the mean integrated squared error (MISE) of KDEs, multivariate KDE, the "curse of dimensionality".

3. Nonparametric regression: Nadaraya-Watson estimator, MSE bounds under Lipschitz and Hölder smoothness, local polynomial estimator, least-squares and penalized splines, trend filtering, additive model, sparse additive model.

4. Minimax theory: Normal means model, James-Stein estimator, Pinsker's Theorem, minimax optimality, adaptivity.

5. Bootstrap: Bootstrap for the mean, Edgeworth expansion and second-order correctness of the bootstrap, bootstrap for statistical functionals, bootstrap for linear regression, bootstrap in nonparametric regression.

6. Traditional nonparametrics: Wilcoxon rank-sum test.

📄 E L Lehmann.
Nonparametrics.
Holden-Day, Inc., 1975.

📄 Alexandre B Tsybakov.
Introduction to nonparametric estimation.
Springer Science & Business Media, 2008.

📄 Larry Wasserman.
All of nonparametric statistics.
Springer Science & Business Media, 2006.