STAT 824 – Nonparametric Inference

Spring 2025

Meeting times: 1:15 - 2:30 pm Tuesday and Thursday in LeConte College 205.
Instructor: Dr. Karl Gregory, LeConte College 216 C, gregorkb@stat.sc.edu.

Link to course website

Description

Nonparametric inference introduces graduate students to some important topics in nonparametric statistics such as density estimation, nonparametric regression, and the bootstrap. It assumes familiarity with topics covered in STAT 712 and STAT 713 and focuses on theory and computation.

Overview of topics

Nonparametric inference has traditionally referred to statistical inference which does not make assumptions about the distribution of the data, e.g. Gaussianity. In recent decades nonparametric inference has more and more often referred to statistical inference in settings with an infinite number of parameters, making a contrast with parametric inference, which targets a finite number of parameters. Quintessentially, nonparametric inference will target the estimation of an unknown function f of unspecified form, while parametric inference targets some parameter θ with finite dimension.

Here are some quotes, time-stamped, reflecting this history:

> *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, 1975

> *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, 2006

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, 2008

Our dance through the vast arcade of nonparametric statistics will begin at density estimation, where we will encounter all the usual "issues" that arise when one proceeds nonparametrically. From there, we will turn to nonparametric regression. The latter part of the course will focus on the bootstrap. At the very end some attention will be given to the rank-based methods mentioned in the Lehmann quote. Other topics will be inserted as time allows and as interest dictates.

Following are previews of the three main areas we will focus on:

I. Density estimation

If we observe a random sample from a distribution with unknown probability density function f, we consider how to estimate the function f. Parametric methods must assume that f belongs to a family of densities and then estimate the value of the parameters which index the densities in the family (if Gaussianity is assumed, one estimates the mean and variance; if Gamma is assumed, one estimates the shape and scale parameters, etc.). Nonparametric methods do not require any specification of what family of distribution f belongs to. Nonparametric estimators of f when we observe univariate and bivariate data may look like those shown below:

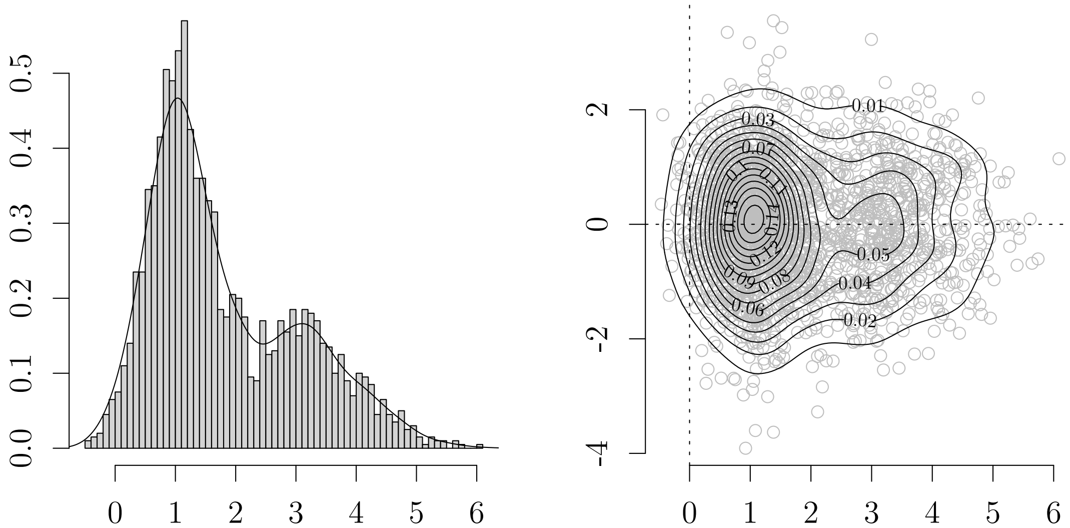

Figure 1: The left panel shows a histogram of a random sample from an arbitrary distribution with an estimate of the probability density function overlaid. The right panel shows a bivariate random sample with contours of the estimated bivariate probability density function overlaid. These are nonparametric estimators and can be computed without assuming any particular distribution for the data.

In this section of the course we will study—as a prelude to density estimation—estimation of the cumulative distribution function, with discussions of the Dvoretzky-Kiefer-Wolfowitz inequality, the Kolmogorov-Smirnov test, and Brownian motion. Then we will study kernel density estimation of a probability density function, obtain bounds on the mean squared error and mean integrated squared error of kernel density estimators, and consider optimal bandwidth selection via crossvalidation. In addition we will study multivariate kernel density estimation and encounter the so-called "curse of dimensionality".

II. Nonparametric regression

Suppose we observe data pairs $(x_1, y_1), \ldots, (x_n, y_n)$, where $y_i = f(x_i) + \varepsilon_i$ for $i = 1, \ldots, n$ for some unknown function f, and where $\varepsilon_1, \ldots, \varepsilon_n$ are noise. Now, parametric methods for estimating the unknown function f must specify a functional form for f and then estimate any unknown constants which are involved, e.g. one could assume $f(x) = \beta_0 + \beta_1 x$ and then estimate the intercept and slope parameters β_0 and β_1. See the figure below.

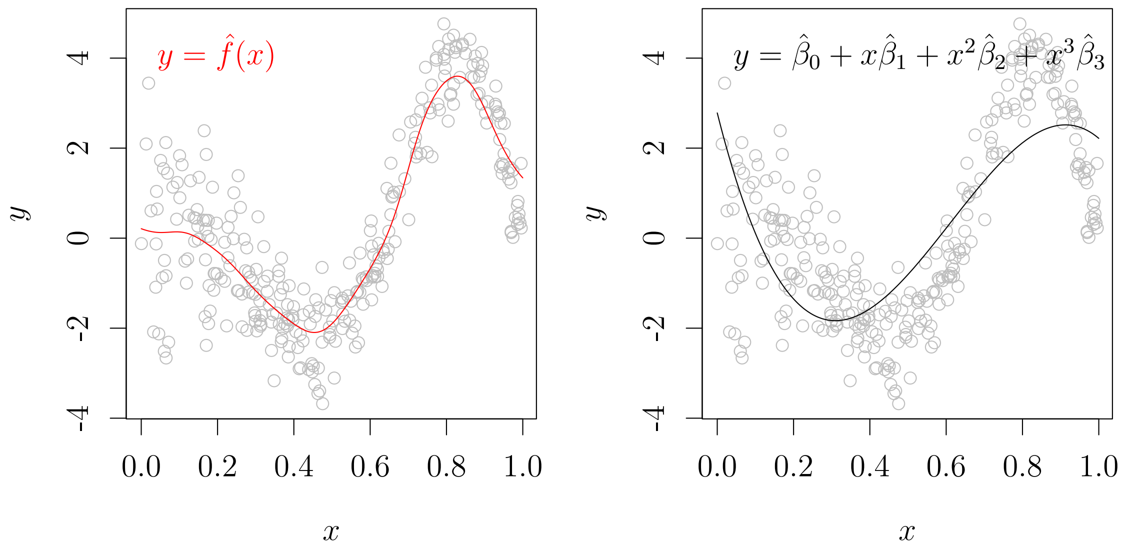

Figure 2: The left panel shows a nonparametric estimate of the function f and the right panel shows, with the same data, a parametric estimate of the function f under the assumption that it is a cubic function (it fits the data poorly).

In this section of the course we will study the Nadaraya-Watson and local polynomial estimators for the regression function f and obtain bounds on the mean squared error of these estimators under different smoothness assumptions (Lipschitz and Hölder) on f. We also study spline-based estimation and the choice of smoothness via crossvalidation. In the setting with multiple covariates we study the additive model and the sparse additive model.

III. The bootstrap

The bootstrap belongs to nonparametric inference under both the traditional and current usages of the term. It is designed for settings in which the distribution of the data is unknown and finds application in settings with a finite or an infinite number of parameters. Its goal is most often to estimate the sampling distribution of a pivotal quantity, an example of which is a sample mean centered to have zero mean and unit variance. Once one has an estimate of the sampling distribution of such a quantity one can use it to make inferences for some target parameter, for example by building confidence intervals.

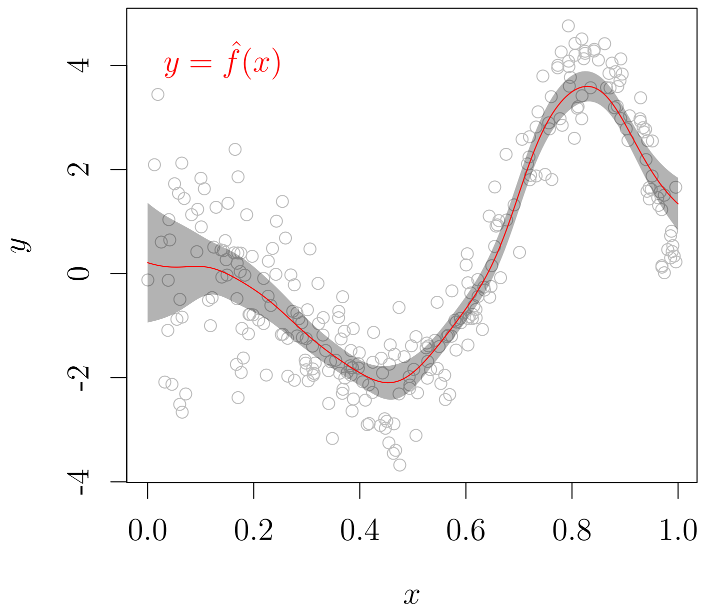

Figure 3: A bootstrap method called the wild bootstrap can be used to build a confidence band for the unknown function f.

In this section of the course we will study the bootstrap for the sample mean, proving that it provides a consistent estimator of the sampling distribution of the centered and scaled sample mean. Then we will discuss Edgeworth expansions and the so-called second-order correctness of the bootstrap for the mean. Moving beyond the mean, we will consider differentiable statistical functionals, and then introduce bootstrap methods for parametric as well as nonparametric regression estimators such as the residual bootstrap and the wild bootstrap.

Prerequisites

STAT 713 or consent of instructor.

Textbook

There is no required textbook. Complete notes will be posted at the course website. Some sources are:

1. Tsybakov, A. B. *Introduction to nonparametric estimation.* Springer Science & Business Media, 2008.
2. Wasserman, L. *All of nonparametric statistics.* Springer Science & Business Media, 2006.
3. Gyrfi, L., Kohler, M., Krzyzak, A., & Walk, H. *A distribution-free theory of nonparametric regression.* Springer Science & Business Media, 2006.
4. Ruppert, D., Wand, M. P., & Carroll, R. J. *Semiparametric regression* (No. 12). Cambridge university press, 2003.
5. Athreya, K. B., & Lahiri, S. N. *Measure theory and probability theory.* Springer Science & Business Media, 2006.
6. Hall, P. *The bootstrap and Edgeworth expansion.* Springer Science & Business Media, 2013.
7. Serfling, R. J. *Approximation theorems of mathematical statistics* (Vol. 162). John Wiley & Sons, 2009.
8. Pratt, J. W., & Gibbons, J. D. *Concepts of nonparametric theory.* Springer Science & Business Media, 2012.

Grading

The graded components of the course are the following:

> **Homework (20%)** There will be 5 homework assignments. Homework will be submitted on Blackboard. All homework assignments must be typed (LaTeX or markdown).
>
> **Project (80%)**: Will be completed in phases and will culminate in a written report and a presentation to the class.

The thresholds 90%, 87%, 80%, 77%, 70%, 67%, and 60% will be used to determine the assignment of the letter grades A, B+, B, C+, C, D+, and D, respectively. The grade of F will be assigned to those earning less than 60%.

Honor code

See the Carolinian Creed in the Carolina Community: Student Handbook & Policy Guide. Violations of the USC Honor Code may result in a 0 for the work in question, and, in accordance with University policy, other punishments up to and including expulsion from the University.

Accommodations

If you require special accommodations, they must be arranged in advance through the Office of Student Disability Services Close-Hipp, Suite 102. (803-777-6142, SADRC@mailbox.sc.edu).