Statistical Methodology 24 (2015) 37–51

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Approximating moments of continuous functions of random variables using Bernstein polynomials

A.I. Khuri [a], S. Mukhopadhyay [b,*], M.A. Khuri [c]

[a] Department of Statistics, University of Florida, P.O. Box 118545, Gainesville, FL 32611-8545, USA

[b] Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

[c] Department of Mathematics, State University of New York at Stony Brook, Stony Brook, NY, USA

ARTICLE INFO

Article history:
Received 11 September 2013
Received in revised form 19 September 2014
Accepted 18 November 2014
Available online 24 December 2014

Keywords:
Balanced and unbalanced data
Delta method
Heritability function
Jensen's inequality
Polynomial approximation
Tchebycheff polynomials
Uniform convergence
Weierstrass approximation theorem

ABSTRACT

Bernstein polynomials have many interesting properties. In statistics, they were mainly used to estimate density functions and regression relationships. The main objective of this paper is to promote further use of Bernstein polynomials in statistics. This includes (1) providing a high-level approximation of the moments of a continuous function $g(X)$ of a random variable X, and (2) proving *Jensen's inequality* concerning a convex function without requiring second differentiability of the function. The approximation in (1) is demonstrated to be quite superior to the *delta method*, which is used to approximate the variance of $g(X)$ with the added assumption of differentiability of the function. Two numerical examples are given to illustrate the application of the proposed methodology in (1).

1. Introduction

Bernstein [3] demonstrated that the well-known *Weierstrass approximation theorem* (concerning the uniform approximation of a continuous function by polynomials) could be easily derived by using the *law of large numbers*. This led to the introduction of the so-called *Bernstein polynomials*, which later became widely used in many scientific and computer engineering applications. In statistics, Bernstein polynomials were mainly used in estimating density functions, quantile functions, and regression

* Corresponding author. Tel.: +91 2225767495.
E-mail address: siuli@math.iitb.ac.in (S. Mukhopadhyay).

http://dx.doi.org/10.1016/j.stamet.2014.11.004

relationships. In the area of density estimation, the statistical literature is very rich. Vitale [28] was the first to use Bernstein polynomials to obtain smooth estimators for density functions on a closed interval. Babu et al. [1] considered the application of Bernstein polynomials for approximating continuous density functions. The authors also investigated the asymptotic properties of the resulting estimators of these functions. A multivariate extension of this work was considered by Babu and Chaubey [2]. Turnbull and Ghosh [26] presented a unimodal density estimation method on the basis of Bernstein polynomials. The density estimate was obtained using quadratic programming techniques to minimize a scaled squared distance between the Bernstein distribution function estimate (constrained to unimodality) and the empirical cumulative distribution function of the data. Leblanc [12] examined the boundary properties of Bernstein estimators of density and distribution functions. Petrone [17] presented the general framework for Bayesian nonparametric density estimation using a prior based on Bernstein polynomials.

In the area of regression estimation, Stadtmüller [22] used Bernstein polynomials to approximate an unknown regression function. This method was generalized by Tenbusch [24] to regression models with several predictors. Curtis and Ghosh [6] took a Bayesian approach to estimate a regression function using Bernstein polynomials. They established a connection between monotonic regression and variable selection. Osman and Ghosh [16] presented a nonparametric regression model for the conditional hazard rate using Bernstein polynomials. Wang and Ghosh [29] developed an estimator of the regression function subject to various shape constraints. Their use of Bernstein polynomials made it possible to obtain an estimator that could be adapted to accommodate shape constraints such as nonnegativity, monotonicity, and convexity.

In this article, we present further applications of Bernstein polynomials which include the approximation of moments of functions of random variables, and the derivation of a proof of *Jensen's inequality* concerning a convex function $\phi(X)$ of a random variable X. The proof does not require that the function be twice differentiable. In the area of moment approximation, the proposed procedure is contrasted with the *delta method* which is used to approximate the mean and variance of a differentiable function of a random variable. The benefits of using Bernstein polynomials in the approximation of moments are outlined in Section 2.2. Two examples are presented to demonstrate how to use these polynomials to approximate the moments of a continuous function $g(X)$ of a random variable X.

1.1. Bernstein polynomials

A Bernstein polynomial of degree n of a function $g(t)$ defined on $[0, 1]$ is given by

$$B_n^*(g; t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} g\left(\frac{i}{n}\right), \quad n = 1, 2, \ldots. \tag{1.1}$$

Note that $B_n^*(g; t)$ represents the expected value of $g(n^{-1}\sum_{i=1}^{n} Y_i)$, where $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed Bernoulli random variables with success probability t. Thus, $\sum_{i=1}^{n} Y_i$ has the binomial distribution $\mathrm{Bin}(n, t)$. If the function g is defined on $[a, b]$, then by making the transformation $t = (x-a)/(b-a)$, the corresponding Bernstein polynomial of the function $g(x),\ a \le x \le b$, can be expressed as

$$B_n(g; x) = \frac{1}{(b-a)^n} \sum_{i=0}^{n} \binom{n}{i} (x-a)^i (b-x)^{n-i} g\left[a + (b-a)\frac{i}{n}\right], \quad a \le x \le b. \tag{1.2}$$

The following theorems are known about Bernstein polynomials (see [14]; [7, Chapter 6]; [18, Chapter 7]):

Theorem 1.1. *If $g(x)$ is continuous on $[a, b]$, then for every $\epsilon > 0$, there exists an integer $N(\epsilon)$ such that $|g(x) - B_n(g; x)| < \epsilon$ for $n \ge N$ and for all x in $[a, b]$.*

Bernstein [3] noted that this well-known approximation theorem of Weierstrass could be easily proved using the *law of large numbers* and the fact that $B_n^*(g; t)$ in (1.1) is the expected value of $g(n^{-1}\sum_{i=1}^{n} Y_i)$, as was mentioned earlier (see [13]).