Gulf Journal of Mathematics ISSN: 2309-4966
Vol 21, Issue 2 (2025) 149-161 https://doi.org/10.56947/gjom.v21i2.3668

ON BERNSTEIN POLYNOMIAL ESTIMATORS OF THE HAZARD RATE FUNCTION

TAYEB DJEBBOURI[1], KOUIDER DJERFI[2*] and BRAHIM REFAFA[3]

ABSTRACT. For the density and distribution functions supported on the unit interval, Bernstein polynomial estimators are known to have good asymptotic properties under appropriate regularity conditions on the function to be estimated. In particular, these estimators are asymptotically unbiased over the entire unit interval. In this paper, we consider the estimation method using Bernstein polynomials and we show that it can be naturally adapted to the hazard function. The asymptotic properties of the obtained estimators are studied. More precisely, we obtain strong consistency and asymptotic normality under appropriate assumptions.

Keywords. Bernstein polynomials; hazard function; mean squared error; asymptotic normality.
2020 Mathematics Subject Classification. Primary 46L55; Secondary 44B20

1. Introduction

The estimation of the hazard function (the conditional failure rate) $h(x)$ plays a very important role in statistics. Indeed, it is used in risk analysis and for the study of survival phenomena in many fields such as medicine, geophysics, socioeconomic studies, etc. The literature on the estimation of the hazard function is very abundant. For the case of vectorial observations, we can cite [21], [17], [4] and [16] for relatively recent references. In the functional case, we cite [18], [15], [12] and [6]. Overall, these works on the nonparametric estimation of the hazard function is based on the kernel method. We find results about almost complete convergence, asymptotic normality and uniform almost complete convergence whether in the dependent or mixing case.

The main contribution of this work is the estimation of the hazard function with an entirely different method. We follow the approach of [19] in the estimation of the density function using Bernstein polynomials. These polynomials were introduced by Sergei Bernstein in order to give a probabilistic proof of Weierstrass' Theorem (see. [3]).

Later, these polynomials are used in the approximation of functions in larger

Date: Received: Oct 7, 2025; Revised: Nov 4, 2025; Accepted: Nov 17, 2025.
* Corresponding author

functional spaces, such as spaces of integrable functions, spaces of measurable functions. Then the Bernstein polynomials themselves are generalized to a non-compact support. For more details on the analytic aspect of these polynomials, see [13].

On the other hand, for the estimation method using these polynomials, many authors agree with the idea of [19]. [1] present a method for the estimation of the density and the distribution function with a simulation of the proposed method by making a comparison with that of the kernel.

A few years later, [2] generalize Bernstein's polynomial estimation method for a $d-$dimensional model in order to define the estimator of a density whose support is a compact domain of $\mathbb{R}^d$.

A thorough study of Bernstein estimators of the density and the distribution function is mainly detailed in the works of [8], [9] and [10]. It also shows the asymptotic normality of these estimators including the fact that they are asymptotically unbiased.

In [11], the author examines the boundary properties of Bernstein estimators . Specifically, he shows that Bernstein density estimators have reduced bias, but increased variance in the boundary region. In the case of distribution function estimation, it shows that these estimators experience an advantageous boundary effect.

Therefore, this note presents an estimator of the hazard function using Bernstein polynomials. The bias and the variance of our estimator are calculated, asymptotic properties are obtained by referring to the work of [9]. The remainder of the paper is organized as follows. In Section 2 Some remarkable properties of Bernstein polynomials are presented. Section 3 is devoted to a summary of the results obtained on the Bernstein estimator of the distribution and the density function. In Section 4, we state the main theoretical result of interest for our estimator of the hazard function $h(x)$. First, we state a convergence theorem, we compute its bias, variance and mean squared error (MSE). Finally, the section ends with a result on asymptotic normality.

2. Basic results on Bernstein polynomials

In these section we establish some properties of the Bernstein polynomials. The Bernstein polynomials of degree m are defined on $[0,1]$, by

$$P_{k,m}(x) = \begin{pmatrix} m \\ k \end{pmatrix} x^k (1-x)^{m-k}$$

for $k = 0, 1, \ldots, m$, where

$$\begin{pmatrix} m \\ k \end{pmatrix} = \frac{m!}{k!(m-k)!}$$

There are $m+1$ $m-$th degree Bernstein polynomials. throughout this paper, we usually set $P_{k,m} = 0$, if $k < 0$ or $k > m$. Several important properties of these polynomials can be proved immediately from their formulas.

Proposition 2.1. [13] *and* [9]. *Bernstein polynomials have the following properties:*

(1) *Positivity:*

$$\forall k \in \{0, \dots, m\},\ \forall x \in [0,1],\ P_{k,m}(x) \geq 0,$$

and $P_{k,m}(x) = 0$ *if and only if* $x \in \{0,1\}$.

(2) *Partition of Unity:*

$$\forall m,\ \sum_{k=0}^{m} P_{k,m}(x) = 1$$

(3) *Recursive property:*

$$\forall k \in \{1, \dots, m-1\},\ P_{k,m}(x) = (1-x)P_{k,m-1}(x) + xP_{k-1,m-1}(x)$$

(4) *Derivation formula:*

$$\frac{d}{dx}P_{k,m}(x) = m(P_{k-1,m-1}(x) - P_{k,m-1}(x))$$

We also refer the reader to [13] and [9] where more detailed properties can be found.

Let f a function in $[0,1]$, the Bernstein polynomial of degree $m \geq 1$ corresponding to the given function f is defined as

$$P_m^f(x) = \sum_{k=0}^{m} f(\frac{k}{m})P_{k,m}(x).$$

Subsequently, the following Theorem is at the origin of the results obtained later for smoothing the empirical distribution function over the interval $[0,1]$.

Theorem 2.2. [5].
If f *is continuous in* $[0,1]$, *the Bernstein polynomials* $P_m^f(x)$ *tend uniformly to* $f(x)$ *as* $m \to \infty$.

3. Estimation of a distribution and density function

This section presents the main results obtained concerning the estimation method using Bernstein polynomials. Let $X_1, ..., X_n$ be independent, identically distributed random variables and f denote the probability density of X which is supported on $[0,1]$. Let F denotes the distribution function of the random variable X, we present the approach of [19] and [1], used for distribution and density estimation. These estimators are based on Bernstein polynomials of order $m > 0$ and smoothing the empirical distribution function.
Let F_n denotes the empirical distribution function obtained from the random sample $X_1, ..., X_n$. The Bernstein estimator of order $m > 0$ of the distribution F is defined as

$$\widehat{F}_{m,n}(x) = \sum_{k=0}^{m} F_n(\frac{k}{m})P_{k,m}(x). \tag{3.1}$$

On the other hand, taking the derivative of $\widehat{F}_{m,n}$ with respect to the variable x leads to

$$\widehat{f}_{m,n}(x) = \frac{d}{dx}\widehat{F}_{m,n}(x) = m\sum_{k=0}^{m-1}\left[F_n\left(\frac{k+1}{m}\right) - F_n\left(\frac{k}{m}\right)\right]P_{k,m-1}(x), \tag{3.2}$$

which is the Bernstein density estimator of order m, as it is defined by [1], [9] and many others.
First, as a consequence of Theorem 2.2, the following Theorem shows that the estimator$\widehat{F}_{m,n}$ is strongly consistent.

Theorem 3.1. [1].
Let F be a continuous probability distribution function on the interval $[0,1]$. If $m, n \longrightarrow \infty$, then

$$\|\widehat{F}_{m,n} - F\| \longrightarrow 0 \textit{ almost surely},$$

where $\|G\| = \sup_{x\in[0,1]} |G(x)|$.

A similar result of strong convergence can be established for the estimator $\widehat{F}_{m,n}$ under a regularity condition on the density f. In [1], it is shown that if f is continuous and Lipshitzian of order 1, then we have

Theorem 3.2. [1].
Assume that f is continuous and Lipschitz of order 1. For $2 \leq m \leq (n/\log n)$ and $m = o(n/\log n)$, we have

$$\|\widehat{f}_{m,n} - f\| \longrightarrow 0 \textit{ almost surely},$$

as $m, n \longrightarrow \infty$.

Later, in the main results section, we apply Theorem 3.1 and Theorem 3.2 to prove a consistency Theorem of the Bernstein estimator of the hazard function, under appropriate conditions.
We now state some other results. Let us begin by introducing the following notations, which will be used later:

- $\psi(x) = (4\pi x(1-x))^{-1/2}$,
- $\Delta_1(x) = \frac{1}{2}\left[(1-2x)f'(x) + x(1-x)f''(x)\right]$,
- $b(x) = \frac{1}{2}[x(1-x)f'(x)]$,
- $\sigma^2(x) = F(x)[1-F(x)]$,
- $V(x) = f(x)[2x(1-x)/\pi]^{1/2}$,
- $\gamma(x) = f(x)[4\pi x(1-x)]^{-1/2}$,

where f and F are defined on $[0,1]$ such that the derivatives f' and f'' exist.

On the other hand, we consider the following hypothesis:

$$\begin{aligned}&F \textit{ is continuous on } [0,1] \textit{ and admits}\\ &\textit{two continuous and bounded derivatives.}\end{aligned} \tag{3.3}$$

The following Theorem presents some results on the Bernstein estimator $\widehat{F}_{m,n}$. Specifically, it establishes the bias and variance of $\widehat{F}_{m,n}$. In particular, we find that the estimator $\widehat{F}_{m,n}$ is asymptotically unbiased and has zero variance for $x = 0$ and $x = 1$.

Theorem 3.3. [10].
Under assumption (3.3), we have for $x \in (0, 1)$ that

- $$Bias[\widehat{F}_{m,n}(x)] = \mathbb{E}[\widehat{F}_{m,n}(x)] - F(x) = m^{-1}b(x) + o(m^{-1}),$$
- $$MSE[\widehat{F}_{m,n}(x)] = n^{-1}\sigma^2(x) - m^{-1/2}n^{-1}V(x) + m^{-2}b^2(x) + o(m^{-2}) + o_x(m^{-1/2}n^{-1}),$$
- $$Var[\widehat{F}_{m,n}(x)] = n^{-1}\sigma^2(x) - m^{-1/2}n^{-1}V(x) + o_x(m^{-1/2}n^{-1}),$$

  as both $m, n \to \infty$.

On the other hand, it is well known that

$$MSE[F_n(x)] = Var[F_n(x)] = n^{-1}\sigma^2(x),$$

so that $\widehat{F}_{m,n}$ and F_n are equivalent in MSE up to the first-order. However, when considering also higher order terms, it turns out that $\widehat{F}_{m,n}$ asymptotically dominates F_n in terms of MSE when m is well chosen. Notice that the previous result shows that $\widehat{F}_{m,n}$ can asymptotically out perform the empirical distribution function at every $x \in (0, 1)$ in terms of MSE. (Both estimators achieve an MSE of zero at $x = 0$ and $x = 1$).
Let us now examine the asymptotic properties of the estimator defined by the formula (3.2). This is the Bernstein approximation of order m of the density f defined on $[0, 1]$. Obviously, $\widehat{f}_{m,n}$ is a polynomial of degree $m - 1$. In the context of density estimation, it is an approximation to f that cannot be computed.

Theorem 3.4. [19], [9].
Under assumption (3.3), we have

- *For $x \in [0, 1]$*
  $$Bias[\widehat{f}_{m,n}(x)] = \mathbb{E}[\widehat{f}_{m,n}(x)] - f(x) = \frac{\Delta_1(x)}{m} + o(m^{-1})$$
- *For $x \in (0, 1)$*
  $$MSE[\widehat{f}_{m,n}(x)] = \frac{\Delta_1^2(x)}{m^2} + \frac{m^{1/2}}{n}f(x)\psi(x) + \frac{o(1)}{m^2} + o_x(\frac{m^{1/2}}{n}),$$
  consequently
  $$Var[\widehat{f}_{m,n}(x)] = \frac{m^{1/2}}{n}f(x)\psi(x) + o_x(\frac{m^{1/2}}{n}),$$
  as both $m, n \to \infty$ in such a way that $m/n \to 0$.

- *For* $x \in \{0,1\}$
$$MSE[\widehat{f}_{m,n}(x)] = \frac{\Delta_1^2(x)}{m^2} + \frac{m}{n}f(x) + \frac{o(1)}{m^2} + o(\frac{m}{n})$$
and therefore
$$Var[\widehat{f}_{m,n}(x)] = \frac{m}{n}f(x) + o(\frac{m}{n}),$$
as both $m, n \to \infty$ *in such a way that* $m/n \to 0$.

3.1. **Asymptotic normality.** Unlike kernel estimators, the study of asymptotic normality of Bernstein estimators (whether with respect to distribution or density) has not been the subject of much research. However, the most important ones can be cited. Precisely, for the density estimator, [1] obtain strong consistency and asymptotic normality under an appropriate choice of the polynomial degree. Later, in [2], the asymptotic normality of the distribution and density function on a hypercube using Bernstein polynomials for dependent random vectors are investigated (a-mixing condition). Recently, in another context, for spectral density estimation, [7] obtains asymptotic normality for a type of generalized Bernstein estimators. More later, [10] establish the (pointwise) asymptotic normality of the distribution Bernstein estimator and show that it have highly advantageous boundary properties, including the absence of boundary bias. However, more recently, we can cite [14] and [20], where the authors established results of asymptotic normality by applying the Berry-Esseen Theorem. The next Theorems establish the asymptotic normality of the estimators defined by (3.1) and (3.2).

Theorem 3.5. [10].
Assume (3.3) holds and $m,\ n \longrightarrow \infty$. *Then, for* $x \in (0,1)$ *such that* $0 < F(x) < 1$, *if* $mn^{-1/2} \longrightarrow \infty$,
$$n^{1/2}(\widehat{F}_{m,n}(x) - F(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2(x)\right),$$
where $\sigma^2(x)$ *is defined as previously in the notations and* " $\xrightarrow{\mathcal{D}}$ " *denotes convergence in distribution.*

For the density estimator, a remarkable result on asymptotic normality is that in [1], as follows:

Theorem 3.6. [1].
Assume that f *is Lipshitzian of order* 1, *and* $f(x) > 0$, *then*
$$n^{1/2}m^{-1/4}(\widehat{f}_{m,n}(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \gamma(x)\right),$$
as $m,\ n \longrightarrow \infty$ *such that* $2 \le m \le (n/\log n)$ *and* $n^{2/3}/m \longrightarrow 0$.

Note that Theorem 3.5 is in fact a corollary in [10], the proof of which relies mainly on the relatively small distance between the empirical distribution and the estimator $\widehat{F}_{m,n}$ for an appropriate choice of the latter's degree. This argument, added to the fact that the empirical distribution itself has a well-known asymptotic normal distribution, leads to the asymptotic normality of $\widehat{F}_{m,n}$.

Let us also note that the two results cited above were chosen precisely in order to obtain a similar result on the asymptotic normality of the estimator $\widehat{h}_{m,n}$ defined below in (4.1). The interest of this choice lies in particular in the coherence existing between the hypotheses of Theorem 3.5 and Theorem 3.6. More in-depth discussions and remarks on this question will be addressed at the end of the following section.

4. Main results

Our purpose is to estimate the hazard rate $h(x)$, defined as

$$h(x) = \frac{f(x)}{1 - F(x)}, \;\; F(x) < 1$$

Therfore, we propose the estimator $\widehat{h}_{m,n}(x)$ given by

$$\widehat{h}_{m,n}(x) \;\; = \;\; \frac{m \sum_{k=0}^{m-1} \left[F_n\left(\frac{k+1}{m}\right) - F_n\left(\frac{k}{m}\right)\right] P_{k,m-1}(x)}{1 - \sum_{k=0}^{m} F_n(\frac{k}{m}) P_{k,m}(x)} \tag{4.1}$$

where $P_{k,m}(x) = \left(\begin{array}{c} m \\ k \end{array} \right) x^k (1-x)^{m-k}$ is the Bernstein polynomial of order m.

Remark 4.1. As seen in most references on the estimation of the hazard rate, the difference $\widehat{h}_{m,n}(x) - h(x)$ can be written

$$\begin{aligned}
\widehat{h}_{m,n}(x) - h(x) &= \frac{\widehat{f}_{m,n}(x)}{1 - \widehat{F}_{m,n}(x)} - \frac{f(x)}{1 - F(x)} \\
&= \frac{\widehat{f}_{m,n}(x) - F(x)\widehat{f}_{m,n}(x) - f(x) + f(x)\widehat{F}_{m,n}(x)}{(1 - \widehat{F}_{m,n}(x))(1 - F(x))} \\
&= \frac{\left(\widehat{f}_{m,n}(x) - f(x)\right)(1 - F(x)) + f(x)\left(\widehat{F}_{m,n}(x) - F(x)\right)}{\left(1 - \widehat{F}_{m,n}(x)\right)(1 - F(x))} \\
&= \frac{\widehat{f}_{m,n}(x) - f(x)}{1 - \widehat{F}_{m,n}(x)} + \frac{f(x)}{1 - F(x)} \frac{\widehat{F}_{m,n}(x) - F(x)}{1 - \widehat{F}_{m,n}(x)},
\end{aligned}$$

finally we can write

$$\widehat{h}_{m,n}(x) - h(x) \;\; = \;\; \frac{\left[(\widehat{f}_{m,n}(x) - f(x)) + h(x)(\widehat{F}_{m,n}(x) - F(x))\right]}{1 - \widehat{F}_{m,n}(x)} \tag{4.2}$$

The asymptotic properties of the estimator $\widehat{h}_{m,n}$ that we will establish in this section are based on the one hand on the results previously obtained by different authors on the estimators $\widehat{F}_{m,n}$ and $\widehat{f}_{m,n}$, and on the other hand on the following technical Lemma.

Lemma 4.2. *Let $\widehat{F}_{m,n}(x)$ be the estimator defined on $[0,1]$ by formula (3.1). We therefore have*

$$\widehat{F}_{m,n}(x) \quad = \quad x + O(1), \tag{4.3}$$

which implies for $x \neq 1$, that

$$\frac{1}{1-\widehat{F}_{m,n}(x)} \quad = \quad \frac{O(1)}{1-x} \tag{4.4}$$

Proof. According to formula (3.1), we have

$$\widehat{F}_{m,n}(x) = \sum_{k=0}^{m} F_n(\frac{k}{m})P_{k,m}(x), \ (m>0),$$

with

$$\sum_{k=0}^{m} P_{k,m}(x) = 1, \qquad P_{m,m}(x) = x^m, \qquad F_n(1) = 1.$$

We can write

$$\widehat{F}_{m,n}(x) = \sum_{k=0}^{m-1} F_n(\frac{k}{m})P_{k,m}(x) + x^m$$

Since F_n is increasing, we have for all $k \in \{0,...,m-1\}$

$$0 \leq F_n(\frac{k}{m}) \leq F_n(\frac{m-1}{m})$$

so

$$\begin{aligned} x^m \leq \widehat{F}_{m,n}(x) \quad &\leq \quad x^m + F_n(\frac{m-1}{m})\sum_{k=0}^{m-1} P_{k,m}(x) \\ x^m \leq \widehat{F}_{m,n}(x) \quad &\leq \quad x^m + (1-x^m)F_n(\frac{m-1}{m}) \end{aligned}$$

In particular we have:

$$0 \leq \widehat{F}_{m,n}(x) \leq x + F_n(\frac{m-1}{m}),$$

and we can therefore write

$$\widehat{F}_{m,n}(x) \quad = \quad x + O(1),$$

which also implies

$$\frac{1}{1-\widehat{F}_{m,n}(x)} \quad = \quad \frac{O(1)}{1-x}$$

$\square$

First, we state the following Theorem which establishes the strong consistency of the proposed estimator $\widehat{h}_{m,n}$.

Theorem 4.3. *Assume that the probability distribution F is continuous on the interval $[0,1]$ and that the corresponding density function f is Lipschitz of order 1. Let S be a compact subset of $[0,1]$ such that $1 \notin S$. For $2 \leq m \leq (n/\log n)$ and $m = o(n/\log n)$, we have*

$$\sup_{x \in S} |\widehat{h}_{m,n}(x) - h(x)| \longrightarrow 0 \text{ almost surely},$$

as $m, n \longrightarrow \infty$.

Proof. According to (4.2) and (4.4), we can write

$$\widehat{h}_{m,n}(x) - h(x) \quad = \quad \frac{\left[(\widehat{f}_{m,n}(x) - f(x)) + h(x)(\widehat{F}_{m,n}(x) - F(x))\right] O(1)}{1-x} \tag{4.5}$$

For a compact subset $S \subset [0,1]$ such that $1 \notin S$, there are two positive constants α and β such that

$$|\widehat{h}_{m,n}(x) - h(x)| \leq \alpha \|\widehat{f}_{m,n} - f\| + \beta \|\widehat{F}_{m,n} - F\|$$

for all $x \in S$.
This means that under the assumptions of Theorem 3.1 and Theorem 3.2, we obtain

$$\sup_{x \in S} |\widehat{h}_{m,n}(x) - h(x)| \longrightarrow 0 \text{ almost surely},$$

as $m, n \longrightarrow \infty$. $\square$

4.1. **Bias of the hazard Bernstein estimator.** We are now able to give a formula which establishes the bias of the estimator $\widehat{h}_{m,n}$ as a function of those of $\widehat{F}_{m,n}$ and $\widehat{f}_{m,n}$.

Theorem 4.4. *Assume that the functions F and f are continuous, and that each has a continuous and bounded second derivative on $[0,1]$. We have for $x \in (0,1)$*

$$Bias[\widehat{h}_{m,n}(x)] \quad = \quad \frac{1}{m(1-x)} \left[\Delta_1(x) + h(x)b(x)\right] O(1) \tag{4.6}$$

Proof. From the formula (4.5), we have

$$\begin{aligned}
\mathbb{E}[\widehat{h}_{m,n}(x)] - h(x) &= \frac{\left[\mathbb{E}(\widehat{f}_{m,n}(x)) - f(x)\right] O(1)}{1-x} + \frac{h(x)\left[\mathbb{E}(\widehat{F}_{m,n}(x)) - F(x)\right] O(1)}{1-x} \\
&= \frac{1}{1-x}\left[\left(\frac{\Delta_1(x)}{m} + o(m^{-1})\right) + h(x)\left(\frac{b(x)}{m} + o(m^{-1})\right)\right] O(1) \\
&= \frac{\Delta_1(x) O(1)}{m(1-x)} + \frac{o(1)}{m(1-x)} + \frac{h(x)b(x)O(1)}{m(1-x)} + \frac{h(x)o(1)}{m(1-x)} \\
&= \frac{\Delta_1(x)}{m(1-x)} O(1) + \frac{h(x)b(x)}{m(1-x)} O(1),
\end{aligned}$$

finally we have

$$Bias(\widehat{h}_{m,n}(x)) = \frac{1}{m(1-x)} \left[\Delta_1(x) + h(x)b(x)\right] O(1)$$

$\square$

4.2. **MSE and variance of the hazard Bernstein estimator.**

Theorem 4.5. *Assume that the functions F and f are continuous, and that each has a continuous and bounded second derivative on $[0,1]$. We have*

$$Var[\widehat{h}_{m,n}(x)] = \begin{cases} \left(Var[\widehat{f}_{m,n}(x)] + h^2(x)Var[\widehat{F}_{m,n}(x)]\right)\frac{O(1)}{(1-x)^2} & \textit{for } x \in (0,1) \\ \frac{m}{n}f(x) + o(\frac{m}{n}) & \textit{for } x = 0 \end{cases}$$

Proof. Firstly, we start by the formula

$$\widehat{h}_{m,n}(x) - h(x) = \frac{1}{1-x}\left[(\widehat{f}_{m,n}(x) - f(x)) + h(x)(\widehat{F}_{m,n}(x) - F(x))\right]O(1),$$

which means that

$$(\widehat{h}_{m,n}(x) - h(x))^2 = \frac{\left[(\widehat{f}_{m,n}(x) - f(x))^2 + h^2(x)(\widehat{F}_{m,n}(x) - F(x))^2 + 2h(x)A(x)\right]O(1)}{(1-x)^2},$$

where $A(x) = (\widehat{f}_{m,n}(x) - f(x))(\widehat{F}_{m,n}(x) - F(x))$. A direct calculation of this term taking into account the approximation formula (4.3), implies

$$A(x) \quad = \quad (x - F(x) + O(1))\widehat{f}_{m,n}(x) + (F(x) - x + O(1))f(x)$$

Therefore

$$2h(x)A(x)O(1) \quad = \quad h(x)[1 + x - F(x)](\widehat{f}_{m,n}(x) - f(x))O(1)$$

Consequently,

$$\begin{aligned} MSE[\widehat{h}_{m,n}] &= \frac{1}{(1-x)^2}MSE[\widehat{f}_{m,n}]O(1) + \frac{h^2(x)}{(1-x)^2}MSE[\widehat{F}_{m,n}]O(1) \\ &+ \frac{h(x)}{(1-x)^2}[1 + x - F(x)]Bias(\widehat{f}_{m,n})O(1) \end{aligned} \tag{4.7}$$

Now, to establish the formula for the variance of the estimator $\widehat{h}_{m,n}$, we first calculate the term $B^2ias(\widehat{h}_{m,n})$. Indeed, according to the proof of the Theorem 4.4, we can write

$$Bias(\widehat{h}_{m,n}) = \frac{1}{1-x}Bias(\widehat{f}_{m,n})O(1) + \frac{h(x)}{1-x}Bias(\widehat{F}_{m,n})O(1),$$

which means that

$$\begin{aligned} B^2ias(\widehat{h}_{m,n}) &= \frac{1}{(1-x)^2}B^2ias(\widehat{f}_{m,n})O(1) + \frac{h^2(x)}{(1-x)^2}B^2ias(\widehat{F}_{m,n})O(1) \\ &+ \frac{h(x)}{(1-x)^2}Bias(\widehat{F}_{m,n})Bias(\widehat{f}_{m,n})O(1), \end{aligned}$$

where the last term of the formula, taking into account the Lemma 4.2, can be written as

$$\frac{h(x)}{(1-x)^2}Bias(\widehat{F}_{m,n})Bias(\widehat{f}_{m,n})O(1) = \frac{h(x)}{(1-x)^2}(1 + x - F(x))Bias(\widehat{f}_{m,n})O(1)$$

finally, formula (4.7), leads to

$$Var[\widehat{h}_{m,n}] \quad = \quad \left(Var[\widehat{f}_{m,n}] + h^2(x)Var[\widehat{F}_{m,n}]\right)\frac{O(1)}{(1-x)^2} \tag{4.8}$$

□

4.3. **Asymptotic normality of the hazard Bernstein estimator.** As can be seen, there is a contrast between the assumptions of Theorem 3.5 and Theorem 3.6 regarding the degree m of the Bernstein polynomial used in the estimators. Indeed, for the estimator of the distribution function to be asymptotically normal, it is sufficient that $(n^{1/2})/m \longrightarrow 0$. On the other hand, in the case of density estimator, it is necessary that $(n^{2/3})/m \longrightarrow 0$. Since $n^{1/2} < n^{2/3}$, we assume that $(n^{2/3})/m \longrightarrow 0$ in order to simultaneously satisfy the conditions of Theorem 3.5 and Theorem 3.6. Therefore, the desired result follows easily.

Theorem 4.6. *Under the assumptions of Theorem 3.5 and Theorem 3.6, we have for $x \in (0,1)$ such that $f(x) > 0$ and $0 < F(x) < 1$,*

$$n^{1/2}m^{-1/4}(\widehat{h}_{m,n}(x) - h(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\gamma(x)}{1-x}\right)$$

as $m, n \longrightarrow \infty$ such that $2 \leq m \leq (n/\log n)$ and $n^{2/3}/m \longrightarrow 0$.

Proof. For $2 \leq m \leq (n/\log n)$, the formula (4.5) implies

$$\frac{n^{1/2}}{m^{1/4}}(\widehat{h}_{m,n}(x) - h(x)) \quad = \quad \frac{n^{1/2}\left[(\widehat{f}_{m,n}(x) - f(x)) + h(x)(\widehat{F}_{m,n}(x) - F(x))\right]O(1)}{m^{1/4}(1-x)}$$

According to Theorem 3.5 and Theorem 3.6, we conclude

$$\frac{n^{1/2}}{m^{1/4}}(\widehat{h}_{m,n}(x) - h(x)) \quad \xrightarrow{\mathcal{D}} \quad \mathcal{N}\left(0, \frac{\gamma(x)}{1-x}\right)$$

□

Conclusion. In this paper, Bernstein's polynomial estimation method, well known in the literature for probability distributions and density functions, is naturally extended to the hazard function h.

First, we state a strong consistency theorem for the estimator $\widehat{h}_{m,n}$ over a compact interval S contained within the unit interval. We then calculate the bias of $\widehat{h}_{m,n}$ and prove that it is asymptotically unbiased. Later, the mean squared error (MSE) and the variance of the estimator $\widehat{h}_{m,n}$ are obtained from those of the estimators $\widehat{F}_{m,n}$ and $\widehat{f}_{m,n}$. Finally, an asymptotic normality result is obtained under appropriate regularity conditions.

As observed in [19], [1], [9] and others, the Bernstein estimators possess properties comparable to those of the kernel estimators. These include order of convergence of the estimate, consistency, efficiency, and asymptotic normality. Indeed, Theorem 4.3, Theorem 4.4, Theorem 4.5 and Theorem 4.6 shows that our estimator also inherits these properties.

Acknowledgement. The authors thank the editor and the anonymous reviewers for their detailed comments and suggestions, which resulted in an improved paper.

References

1. Babu, G. J., Canty, A. J., and Chaubey, Y. P. (2002).Application of bernstein polynomials for smooth estimation of a distribution and density function.*Journal of Statistical Planning and Inference*, 105(2):377–392. https://doi.org/10.1016/S0378-3758(01)00265-8
2. Babu, G. J. and Chaubey, Y. P. (2006).Smooth estimation of a distribution and density function on a hypercube using bernstein polynomials for dependent random vectors.*Statistics & probability letters*, 76(9):959–969. https://doi.org/10.1016/j.spl.2005.10.031
3. Bernstein, S. (1912).Démonstration du théoreme de weierstrass fondée sur le calcul des probabilities.*Comm. Soc. Math. Kharkov*, 13:1–2.
4. Estévez-Pérez, G. (2002).On convergence rates for quadratic errors in kernel hazard estimation.*Statistics & probability letters*, 57(3):231–241. https://doi.org/10.1016/S0167-7152(02)00052-4
5. Feller, W. (1991). An introduction to probability theory and its applications, Volume 2. John Wiley & Sons.
6. Ferraty, F., Rahbi, A., and Vieu, P. (2008).Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle.*Revue roumaine de mathématiques pures et appliquées*, 53(1):1–18.
7. Kakizawa, Y. (2011).A note on generalized bernstein polynomial density estimators. *Statistical Methodology*, 8(2):136–153. https://doi.org/10.1016/j.stamet.2010.08.004
8. Leblanc, A. (2009).Chung–smirnov property for bernstein estimators of distribution functions.*Journal of Nonparametric Statistics*, 21(2):133–142. https://doi.org/10.1080/10485250802485676
9. Leblanc, A. (2010).A bias-reduced approach to density estimation using bernstein polynomials.*Journal of Nonparametric Statistics*, 22(4):459–475. https://doi.org/10.1080/10485250903318107
10. Leblanc, A. (2012a).On estimating distribution functions using bernstein polynomials. *Annals of the Institute of Statistical Mathematics*, 64:919–943. https://doi.org/10.1007/s10463-011-0339-4
11. Leblanc, A. (2012b).On the boundary properties of bernstein polynomial estimators of density and distribution functions.*Journal of Statistical Planning and Inference*, 142(10):2762–2778. https://doi.org/10.1016/j.jspi.2012.03.016
12. Li, J. and Tran, L. T. (2007).Hazard rate estimation on random fields. *Journal of Multivariate analysis*, 98(7):1337–1355. https://doi.org/10.1016/j.jmva.2007.03.002
13. Lorentz, G. G. (2012).*Bernstein polynomials*. American Mathematical Soc.
14. Lu, D. and Wang, L. (2021).On the rates of asymptotic normality for bernstein polynomial estimators in a triangular array.*Methodology and Computing in Applied Probability*, 23:1519–1536. https://doi.org/10.1016/j.jmaa.2022.126063
15. Nielsen, J. P. and Linton, O. B. (1995).Kernel estimation in a nonparametric marker dependent hazard model. *The Annals of Statistics*, pages 1735–1748. https://doi.org/10.1214/aos/1176324321
16. Quintela-del Rio, A. (2006).Nonparametric estimation of the maximum hazard under dependence conditions.*Statistics & probability letters*, 76(11):1117–1124. https://doi.org/10.1016/j.spl.2005.12.009
17. Roussas, G. G. (1989).Hazard rate estimation under dependence conditions. *Journal of statistical planning and inference*, 22(1):81–93. https://doi.org/10.1016/0378-3758(89)90067-0

18. Singpurwalla, N. D. and Wong, M.-Y. (1983).Estimation of the failure rate-a survey of nonparametric methods part i: Non-bayesian methods.*Communications in Statistics-Theory and Methods*, 12(5):559–588. https://doi.org/10.1080/03610928308828479
19. Vitale, R. A. (1975).A bernstein polynomial approach to density function estimation. In *Statistical inference and related topics*, pages 87–99. Elsevier. https://doi.org/10.1016/B978-0-12-568002-8.50011-2
20. Wang, L. and Lu, D. (2022).On the rates of asymptotic normality for bernstein density estimators in a triangular array.*Journal of Mathematical Analysis and Applications*, 511(1):126063. https://doi.org/10.1016/j.jmaa.2022.126063
21. Watson, G. and Leadbetter, M. (1964).Hazard analysis. i.*Biometrika*, 51(1/2):175–184. https://doi.org/10.1093/biomet/51.1-2.175

[1] Laboratory of Geometry, Analysis, Control and Applications, University of Saida, Dr. Moulay Tahar, Algeria.
Email address: tdjebbouri20@gmail.com

[2] Laboratory of Stochastic Models, Statistics and Applications, University of Saida, Dr. Moulay Tahar, Algeria.
Email address: djorfik@gmail.com

[3] Department of Economics, University of Saida, Dr. Moulay Tahar, Algeria.
Email address: rerayene@hotmail.fr