Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Improved Estimates for Binomial Tails

José A. Adell[1], **Daniel Cárdenas-Morales**[2*]

[1] Department of Statistical Methods, University of Zaragoza, Zaragoza 50009, Spain
[2] Department of Mathematics, University of Jaén, Jaén 23071, Spain
E-mail: cardenas@ujaen.es

Received: 27 August 2025; **Revised:** 15 December 2025; **Accepted:** 22 December 2025

Abstract: In a recent paper, Ferrante gave these estimates for binomial tails, $P(S_n(x) \geq na) \leq F_n(a, x)\exp(-nH\ (\alpha, x))$, in terms of the Kullback-Leibler divergence $H(\alpha, x)$ between Bernoulli random variables. Although the upper bound is asymptotically sharp in n for fixed values of a and x, the term $F_n(a, x)$ goes to infinity when a approaches to x. In this note, we replace $F_n(a, x)$ by a bounded term making use of the Berry-Esseen inequality. As an application, we show that Bernstein polynomials acting on locally affine functions exhibit an exponential rate of pointwise convergence, in contrast with the usual polynomial rates.

***Keywords*:** binomial tails, Kullback-Leibler divergence, Berry-Esseen bounds, Bernstein polynomials

MSC: 60E05, 41A25

1. Introduction

This note deals with the large deviation theory specialized to the binomial distribution. It considers the approximation of the probability of k or more successes in n independent trials when the fraction of successes belongs to the real interval $[x, 1]$, x being the probability of success of each trial. Specifically, given a random variable $S_n(x)$ following the binomial distribution, that is,

$$P(S_n(x) = j) = \binom{n}{j} x^j (1-x)^{n-j}, \quad j = 0, 1, \ldots, n, \tag{1}$$

we present new bounds on the probabilities of right-tail events

$$\{S_n(x) \geq na\}, \quad a \in [x, 1].$$

The estimates will be given in terms of the bivariate function

Copyright ©2026 Daniel Cárdenas-Morales, et al.
DOI: https://doi.org/10.37256/cm.7320268403
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

$$H(u, v) = u \log \frac{u}{v} + (1-u) \log \frac{1-u}{1-v}, \quad u, v \in (0, 1), \tag{2}$$

which is called the Kullback-Leibler divergence between Bernoulli random variables with success probabilities u and v.

It is well-known (see Massart [1], Arratia and Gordon [2], or Carter and Pollard [3], among others) that

$$P(S_n(x) \geq na) \leq e^{-nH(a, x)}, \quad 0 < x \leq a < 1. \tag{3}$$

A very short proof of this inequality is given in Lemma 1 below.

Denote by $\lceil \cdot \rceil$ the ceiling function, and let φ be the function defined by

$$\varphi(\theta) = \sqrt{\theta(1-\theta)}, \quad \theta \in [0, 1].$$

Recently, Ferrante [4] proved that, whenever

$$0 < x < a < 1, \qquad 1 \leq na \leq n-1, \tag{4}$$

one has

$$P(S_n(x) \geq na) \leq F_n(a, x) e^{-nH(\alpha, x)}, \tag{5}$$

where $\alpha = \lceil an \rceil / n$ and

$$F_n(a, x) = \frac{1}{1-r} \frac{1}{\varphi(\alpha)} \frac{1}{\sqrt{2\pi n}}, \qquad r = \frac{x(1-\alpha)}{\alpha(1-x)}, \tag{6}$$

r being called the odds ratio.

For fixed values of x and a satisfying (4), Ferrante [4] showed that estimate (5) is asymptotically sharp, as $n \to \infty$. However, for a fixed value of n, the quantity $F_n(a, x)$ in (6) goes to infinity, as $a \to x$, since in such a case, the odds ratio $r \to 1$. This is the main drawback in inequality (5). On the other hand, it follows from the central limit theorem (see also Lemma 3 below) that $P(S_n(x) \geq nx) \to 1/2$, as $n \to \infty$. Since $H(x, x) = 0$, this means that estimate (3) is not sharp for $x = a$.

The purpose of this note is to refine estimates (3) and (5), having in mind the previous comments. To this end, we assume that

$$0 < x \leq a < 1, \qquad 1 \leq na \leq n-1, \qquad n \geq 2, \tag{7}$$

and consider α as above.

Observe that if the second condition in (7) is not satisfied, the probability $P(S_n(x) \geq na)$ can be directly computed. On the other hand, concerning the left-tail events, we have

$$P(S_n(x) \leq na) = P(S_n(1-x) \geq n(1-a)), \qquad a \in [0, x], \tag{8}$$

since the random variables $S_n(x)$ and $n - S_n(1-x)$ have the same distribution.

The main result of this note and a subsequent corollary are stated as follows.

Theorem 1 Under assumption (7), we have

$$P(S_n(x) \geq na) \leq G_n(a) e^{-nH(\alpha, x)},$$

where

$$G_n(a) = \frac{1}{2} + \left(\frac{1}{\sqrt{2\pi}} + 0.469 \left(\alpha^2 + (1-\alpha)^2 \right) \right) \frac{1}{\varphi(\alpha)\sqrt{n}}. \tag{9}$$

Corollary 1 Under assumption (7), we have

$$P(S_n(x) \geq na) \leq C_n(a, x) e^{-nH(\alpha, x)},$$

where

$$C_n(a, x) = \min\left(F_n(a, x), G_n(a)\right).$$

Obviously, Corollary 1 improves estimate (5). On the other hand, fix $a \in (0, 1)$ satisfying (4). For moderate or large values of n, we see from (9) that $G_n(a) \sim 1/2$. Also, observe that

$$H(\alpha, x) \geq H(a, x), \tag{10}$$

since $\alpha \geq a$. Therefore, Corollary 1 provides us with a refinement of estimate (3) for moderate or large values of n.

Denote by $x_{n,a}$ the solution to the following equation with unknown x:

$$F_n(a, x) = G_n(a).$$

It is readily seen from (6) and (9) that

$$x_{n,a} = \frac{\alpha R_n(a) - 1}{R_n(a) - 1}, \qquad R_n(a) := \frac{G_n(a)\varphi(\alpha)\sqrt{2\pi n}}{\alpha}.$$

Figure 1 shows the graphs of $F_n(a, x)$ and $G_n(a)$ for $0 < x \leq a$, in the particular case $a = 0.4$, $n = 50$. The improvement in Corollary 1 with respect to estimate (5) is achieved on the interval $[x_{n,a}, a] = [0.3567\cdots, 0.4]$.

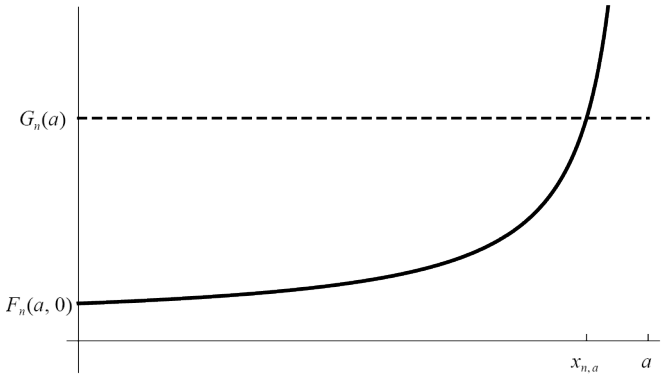


Figure 1. Graphs of $F_n(a, x)$ (solid line) and $G_n(a)$ (dashed line) for $0 < x \leq a$ in the particular case $a = 0.4$, $n = 50$. In the figure, $x_{n,a} = 0.3567\cdots$, $F_n(a, 0) = 0.1151\cdots$, $G_n(a) = 0.6855\cdots$

2. Proof of Theorem 1

The following auxiliary result is interesting by itself.

Lemma 1 Under assumption (7), we have

$$P(S_n(x) \geq na) \leq P(S_n(\alpha) \geq n\alpha))\, e^{-nH(\alpha,x)}.$$

Proof. Let $k \in \{0, 1, \ldots, n(1-\alpha)\}$. Following the ideas in the proof of [4, Theorem 1], we have from (1), (2), and (6)

$$P(S_n(x) = n\alpha + k) = P(S_n(\alpha) = n\alpha + k)\, \frac{x^{n\alpha+k}}{\alpha^{n\alpha+k}}\, \frac{(1-x)^{n(1-\alpha)-k}}{(1-\alpha)^{n(1-\alpha)-k}}$$

$$= P(S_n(\alpha) = n\alpha + k)\, r^k \left(\frac{x}{\alpha}\right)^{n\alpha} \left(\frac{1-x}{1-\alpha}\right)^{n(1-\alpha)}$$

$$= P(S_n(\alpha) = n\alpha + k)\, r^k e^{-nH(\alpha,x)} \leq P(S_n(\alpha) = n\alpha + k)\, e^{-nH(\alpha,x)},$$

since $r \leq 1$. We thus have

$$P(S_n(x) \geq na) = P(S_n(x) \geq n\alpha) = \sum_{k=0}^{n(1-\alpha)} P(S_n(x) = n\alpha + k) \leq P(S_n(\alpha) \geq n\alpha) e^{-nH(\alpha,x)}.$$

The proof is complete. □

By (10), Lemma 1 not only gives us an immediate proof of estimate (3), but also improves it.

Lemma 2 For $k = 1, \ldots, n-1$, we have

$$P(S_n(k/n) = k) \leq \frac{1}{\varphi(k/n)\sqrt{2\pi n}}.$$

Proof. Stirling's approximation (see Robbins [5]) states that

$$\sqrt{2\pi k}\exp\left(\frac{1}{12k+1}\right) \leq \frac{k!e^k}{k^k} \leq \sqrt{2\pi k}\exp\left(\frac{1}{12k}\right), \quad k = 1, 2, \ldots \tag{11}$$

On the other hand, it can be checked that

$$\frac{1}{12n} - \frac{1}{12k+1} - \frac{1}{12(n-k)+1} < 0, \quad k = 1, 2, \ldots, n-1.$$

Hence, we have from (1) and (11)

$$P(S_n(k/n) = k) = \binom{n}{k}\left(\frac{k}{n}\right)^k\left(\frac{n-k}{n}\right)^{n-k} = \frac{n!}{n^n}\frac{k^k}{k!}\frac{(n-k)^{n-k}}{(n-k)!}$$

$$\leq \frac{\sqrt{2\pi n}}{\sqrt{2\pi k}\sqrt{2\pi(n-k)}}\exp\left(\frac{1}{12n} - \frac{1}{12k+1} - \frac{1}{12(n-k)+1}\right) \leq \frac{1}{\varphi(k/n)\sqrt{2\pi n}}.$$

This completes the proof. □

Let Z be a random variable having the standard normal density and let $(Y_k)_{k\geq 1}$ be a sequence of independent copies of a random variable Y such that $\mathbb{E}Y = \mu$, $0 < \mathbb{E}(Y-\mu)^2 = \sigma^2$, and $\mathbb{E}|Y-\mu|^3 = \gamma < \infty$, where $\mathbb{E}$ denotes mathematical expectation. The Berry-Esseen bounds establish that

$$\sup_{y\in\mathbb{R}}\left|P\left(\frac{Y_1 + \cdots + Y_n - n\mu}{\sigma\sqrt{n}} > y\right) - P(Z > y)\right| \leq C\frac{\gamma}{\sigma^3}\frac{1}{\sqrt{n}}, \tag{12}$$

for some positive constant C. According to Esseen [6], C cannot be less than $(3+\sqrt{10})/(6\sqrt{2\pi}) = 0.4097\ldots$ Shevtsova [7] (see also Korolev and Shevtsova [8]) showed that $C \leq 0.4690$.

The application of this result to the case at hand reads as follows.

Lemma 3 Let $0 < x < 1$. Then

$$\sup_{y\in\mathbb{R}}\left|P\left(\frac{S_n(x) - nx}{\varphi(x)\sqrt{n}} > y\right) - P(Z > y)\right| \leq 0.469\frac{x^2 + (1-x)^2}{\varphi(x)\sqrt{n}}.$$

Proof. Let $0 < x < 1$. Note that $S_n(x) = Y_1 + \cdots + Y_n$, where $(Y_k)_{k\geq 1}$ is a sequence of independent copies of a random variable Y having the Bernoulli law with parameter x, that is,

$$P(Y=1)=1-P(Y=0)=x.$$

Thus, $\mathbb{E}Y=x$ and

$$\mathbb{E}|Y-x|^k=(1-x)^k x+x^k(1-x)=x(1-x)(x^{k-1}+(1-x)^{k-1}),\quad k=1,2,\ldots,$$

thus implying that

$$\mu=x,\qquad \sigma^2=\varphi^2(x),\quad \text{and}\quad \gamma=\varphi^2(x)\left(x^2+(1-x)^2\right).$$

Therefore, the result follows from (12). □

2.1 *Proof of Theorem 1*

In view of Lemma 1, it will be enough to show that

$$P(S_n(\alpha)\geq n\alpha)\leq G_n(a), \tag{13}$$

where $G_n(a)$ is defined in (9). Indeed, we have from Lemmas 2 and 3

$$P(S_n(\alpha)\geq n\alpha)\leq \frac{1}{2}+P(S_n(\alpha)=n\alpha)+\left|P(S_n(\alpha)>n\alpha)-\frac{1}{2}\right|\leq \frac{1}{2}+\frac{1}{\varphi(\alpha)\sqrt{2\pi n}}+0.469\frac{\alpha^2+(1-\alpha)^2}{\varphi(\alpha)\sqrt{n}}.$$

This shows (13) and completes the proof.

3. Applications

Ferrante [4] gave several interesting applications to coding theory of estimate (5). Of course, Corollary 1 can be applied to all the examples considered in [4].

Here, we give an application to approximation theory by means of Bernstein polynomials. Such polynomials constitute the paradigmatic example of positive linear operators approximating functions (see, for instance, the monograph by Bustamante [9]). If f is a real function defined on $[0,1]$, the nth Bernstein polynomial of f is defined as

$$B_n f(x)=\sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j}x^j(1-x)^{n-j}=\mathbb{E}f\left(\frac{S_n(x)}{n}\right),\quad 0\leq x\leq 1.$$

Many different results on pointwise convergence have been obtained. We mention the following (cf. Păltănea [10]). If f is continuous, then

$$|B_n f(x) - f(x)| \leq \frac{3}{2}\omega_2\left(f; \frac{\varphi(x)}{\sqrt{n}}\right), \qquad 0 \leq x \leq 1, \tag{14}$$

where $\omega_2(f;\cdot)$ is the second modulus of smoothness of f given by

$$\omega_2(f;\delta) = \sup\{|f(x-\varepsilon) - 2f(x) + f(x+\varepsilon)| : x \pm \varepsilon \in [0,1],\ 0 \leq \varepsilon \leq \delta\}, \qquad \delta \geq 0.$$

It turns out that the best order of magnitude of the upper bound in (14) is n^{-1}. However, if f is locally affine, such a rate of convergence becomes exponential, as it was proved in [11]. Here, as a consequence of Corollary 1, we prove a result that improves [11, Theorem 1]. With this purpose, let ℓ be an affine function, denote by $\|\cdot\|$ the usual supremum norm in $[0,1]$, and denote by 1_A the indicator function of the set A.

Theorem 2 Let $1/n \leq c < d \leq 1 - 1/n$. Let $f : [0,1] \to \mathbb{R}$ be an arbitrary function such that $f(t) = \ell(t)$, $t \in (c,d)$. For any $x \in (c,d)$, we have

$$|B_n f(x) - f(x)| \leq \|f - \ell\| \left(C_n(d,x)e^{-nH(\lceil dn\rceil/n,\,x)} + C_n(1-c,\,1-x)e^{-nH(\lceil (1-c)n\rceil/n,\,x)}\right).$$

Proof. Let $x \in (c,d)$. Define the function $g(t) = f(t) - \ell(t)$, $0 \leq t \leq 1$. Since $g(x) = 0$ and $B_n\ell(x) = \ell(x)$, we have from (8)

$$\begin{aligned}
|B_n f(x) - f(x)| &= |B_n(g+\ell)(x) - (g+\ell)(x)| = |B_n g(x)| \\
&= \left|\mathbb{E}g\left(S_n(x)/n\right)1_{\{S_n(x)/n \leq c\}} + \mathbb{E}g\left(S_n(x)/n\right)1_{\{S_n(x)/n \geq d\}}\right| \\
&\leq \|f-\ell\|\left(P\left(S_n(x) \leq nc\right) + P\left(S_n(x) \geq nd\right)\right) \\
&= \|f-\ell\|\left(P\left(S_n(1-x) \geq n(1-c)\right) + P\left(S_n(x) \geq nd\right)\right).
\end{aligned}$$

Therefore, the result follows from Corollary 1. □

4. Conclusions

The main contribution of this note are the following. In first place, the explicit upper bound for binomial tails given in Corollary 1, namely,

$$P(S_n(x) \geq na) \leq C_n(a,x)e^{-nH(\alpha,\,x)},$$

where the term $C_n(a,x)$, having the order of $n^{-1/2}$, is bounded. This improves Ferrante's bound when a and x are close.

In second place, we give in Lemma 1 a short simple proof of the inequality

$$P(S_n(x) \geq na) \leq P(S_n(\alpha) \geq n\alpha)e^{-nH(\alpha,x)},$$

thus improving classical estimates.

Finally, as an application, we show in Theorem 2 that Bernstein polynomials acting on a function $f : [0, 1] \to \mathbb{R}$ which coincides with an affine one ℓ on (c, d), satisfy the following exponential rate of pointwise convergence

$$|B_n f(x) - f(x)| \leq \|f - \ell\| \left(C_n(d, x)e^{-nH(\lceil dn \rceil/n, x)} + C_n(1-c, 1-x)e^{-nH(\lceil (1-c)n \rceil/n, x)} \right), \quad x \in (c, d),$$

in contrast with the usual polynomial rates.

Acknowledgement

The authors would like to thank the referees for their comments and suggestions, which greatly improved the final presentation.

Funding

Diputación General de Aragón (Spain), through Research Project DGA (E48_23R) supports the first author, and Junta de Andalucía (Spain), through Research Group FQM-0178, supports the second author. In addition, both authors are supported by the Spanish Grant PID2024-155364NB-I00.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Massart P. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Annals of Probability*. 1990; 18(3): 1269–1283. Available from: https://doi.org/10.1214/aop/1176990746.

[2] Arratia R, Gordon L. Tutorial on large deviations for the binomial distribution. *Bulletin of Mathematical Biology*. 1989; 51(1): 125–131. Available from: https://doi.org/10.1007/BF02458840.

[3] Carter A, Pollard D. Tusnády's inequality revisited. *The Annals of Statistics*. 2004; 32(6): 2731–2741. Available from: https://doi.org/10.1214/009053604000000733.

[4] Ferrante GC. Bounds on binomial tails with applications. *IEEE Transactions on Information Theory*. 2021; 67(12): 8273–8279. Available from: https://doi.org/10.1109/TIT.2021.3115044.

[5] Robbins H. A remark on Stirling's formula. *The American Mathematical Monthly*. 1955; 62(1): 26–29. Available from: https://doi.org/10.2307/2308012.

[6] Esseen CG. A moment inequality with an application to the central limit theorem. *Skandinavisk Aktuarietidskrift*. 1956; 39: 160–170. Available from: https://doi.org/10.1080/03461238.1956.10414946.

[7] Shevtsova IG. On absolute constants in inequalities of Berry-Esseen type. *Doklady Mathematics*. 2014; 89(3): 378–381. Available from: https://doi.org/10.1134/S1064562414030338.

[8] Korolev VY, Shevtsova IG. On the upper bound for the absolute constant in the Berry-Esseen inequality. *Theory of Probability and Its Applications*. 2010; 54(4): 638–658. Available from: https://doi.org/10.1137/S0040585X97984449.

[9] Bustamante J. *Bernstein Operators and Their Properties*. Basel: Birkhäuser/Springer; 2017. Available from: https://doi.org/10.1007/978-3-319-55402-0.

[10] Păltănea R. On some constants in approximation by Bernstein operators. *General Mathematics*. 2008; 16(4): 137–148.

[11] Adell JA, Cárdenas-Morales D, López-Moreno AJ. On the rates of pointwise convergence for Bernstein polynomials. *Results in Mathematics*. 2025; 80: 85. Available from: https://doi.org/10.1007/s00025-025-02397-3.