On Bernoulli Sums and Bernstein Polynomials

Jacek Cichoń[†] and Zbigniew Gołębiewski[‡]

Faculty of Fundamental Problems of Technology, Wrocław University of Technology, Poland

---

In the paper we discuss a technology based on Bernstein polynomials of asymptotic analysis of a class of binomial sums that arise in information theory. Our method gives a quick derivation of required sums and can be generalized to multinomial distributions. As an example we derive a formula for the entropy of multinomial distributions. Our method simplifies previous work of Jacquet, Szpankowski and Flajolet from 1999.

Keywords: asymptotic analysis, Bernoulli sums, Bernstein polynomials, entropy, binomial distribution, multinomial distribution

---

1 Introduction

Let $f = \{f_k\}$ be a sequence of real numbers and let $p \in [0, 1]$. The Bernoulli transform of the sequence f with parameter p is defined as

$$S_n^f(p) = \sum_{k=0}^{n} f_k \binom{n}{k} p^k (1-p)^{n-k} \ .$$

The problem of asymptotic estimation of S_n^f (when n tends to infinity) has been investigated in many papers. Jacquet and Szpankowski (1999) used the technology of exponential generating function, saddle point estimates and "analytic depoissonization". They showed that for "smooth" sequences $\{f_k\}$ asymptotically holds $S_n^f \sim f(np)$. Besides, they showed a method of deriving asymptotic expansions of required precision. Flajolet (1999) proposed another approach. It is based on the identity

$$\sum_n S_n z^n = \frac{1}{1-qz} f\left(\frac{pz}{1-qz}\right)$$

where $S_n = S_n^f(p)$, $f(z) = \sum_n f_n z^n$ and $q = 1 - p$.

In this paper we propose an alternative method of investigating Bernoulli sums which is based on Bernstein polynomials. Our method gives in many cases a simple and quick derivation of required formulas.

[†]Email: `Jacek.Cichon@pwr.wroc.pl`
[‡]Email: `Zbigniew.Golebiewski@pwr.wroc.pl`

Moreover, our method can be easily automated. Most of the formula derived in this paper were previously obtained in Jacquet and Szpankowski (1999), Flajolet (1999) and probably in other papers. However, the formula for entropy of multinomial distribution (see Section 4) applied even only to the binomial case seems to have a better precision than any other formula known today. Probably the formulas calculated in the last section for multivariate sums are new.

$\mathbb{N}$ and $\mathbb{R}$ denote the sets of natural and real numbers, respectively, while $\log$ denotes the natural logarithm.

1.1 Bernstein Polynomials

The Bernstein approximation B_n^f to a function $f : [0;1] \to \mathbb{R}$ (see e.g. Lorentz (1986)) is the polynomial

$$B_n^f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \; .$$

It is clear that $B_n^f(x)$ is the expected value of the random variable $f\left(\frac{X_n}{n}\right)$ where X_n is a random variable with binomial distribution with parameters n an x. These polynomials were introduced in Bernstein (1912) in order to give a constructive proof of the Weierstrass Approximation Theorem (see Weierstrass (1885)). Namely, for each function $f : [0,1] \to \mathbb{R}$ and a point of continuity x of f we have

$$\lim_{n\to\infty} B_n^f(x) = f(x)$$

and the convergence is uniform if f is continuous. E. Voronovskaya (see Voronovskaya (1932)) showed that if the second derivative $f''(x)$ of the function f exists, then

$$\lim_{n\to\infty} n(B_n^f(x) - f(x)) = \frac{1}{2}x(1-x)f''(x) \; .$$

Therefore $B_n^f(x) = f(x) + \mathrm{O}\left(\frac{1}{n}\right)$. This result has been later improved by S. Bernstein in the same year (see Bernstein (1932) and also Lorentz (1986) Chapt. 1.6). Namely, he showed that if the $2k$th order derivative $f^{(2k)}$ is bounded in the interval $(0,1)$ then for each $x \in [0,1]$ we have

$$B_n^f(x) = f(x) + \sum_{a=2}^{2k-1} \frac{f^{(a)}(x)}{a!n^a} T_{n,a}(x) + \mathrm{O}\left(\frac{1}{n^k}\right) \tag{1}$$

where

$$T_{n,a}(x) = \sum_k (k-nx)^a \binom{n}{k} x^k(1-x)^{n-k} \; .$$

Remark Notice that $T_{n,a}(x)$ is the ath central moment of a random variable with a binomial distribution with parameters n and x. Clearly, $T_{n,0} = 1$, $T_{n,1} = 0$. It is well known that the sequence $\{T_{n,a}(x)\}$ satisfies the following recurrence

$$T_{n,a+1}(x) = x(1-x)\left(T'_{n,a}(x) + naT_{n,a-1}(x)\right) \; .$$

Therefore, the values of $T_{n,a}$ can be easily computed (see Appendix 6).

If we apply Equation (1) to $k = 1, 2, 3$ then we get

$$B_n^f(x) = f(x) + \mathrm{O}\left(\frac{1}{n}\right) \tag{2}$$

$$B_n^f(x) = f(x) + \frac{x(1-x)f''(x)}{2n} + \mathrm{O}\left(\frac{1}{n^2}\right) \tag{3}$$

$$B_n^f(x) = f(x) + \frac{(1-x)xf''(x)}{2n} + \frac{x(1-x)(4(1-2x)f^{(3)}(x) + 3(1-x)xf^{(4)}(x))}{24n^2} + \mathrm{O}\left(\frac{1}{n^3}\right) \tag{4}$$

and higher level approximations can be computed, too. In the next two subsections we will show how to use this formulas for asymptotics of binomial sums.

2 Simple Applications

There exists a direct connection between Bernoulli sums and Bernstein polynomials for some class of functions. Let us consider, for example, the function $f(x) = x^a$ where $a > 0$. From Equation 3 we get

$$\sum_{k=0}^{n} \left(\frac{k}{n}\right)^a \binom{n}{k} x^k (1-x)^{n-k} = x^a + \frac{a(a-1)x^{a-1}(1-x)}{2n} + \mathrm{O}\left(\frac{1}{n^2}\right) ,$$

therefore

$$\sum_{k=0}^{n} k^a \binom{n}{k} x^k (1-x)^{n-k} = (nx)^a \left(1 + \frac{a(a-1)(1-x)}{2nx} + \mathrm{O}\left(\frac{1}{n^2}\right)\right) .$$

In further considerations we will use the abbreviation $\pi_{n,k}(x) = \binom{n}{k}x^k(1-x)^{n-k}$. We will also use the following consequence of the strong asymptotic concentration of the binomial distribution around its mean which follows from the Chernoff bounds (see Mitzenmacher and Upfal (2005), Corollary 4.6):

Lemma 1 *Suppose that $f : \mathbb{R} \to \mathbb{R}$, $x \in (0, 1)$ and $\delta > 0$. Then*

$$|S_n^f(x) - \sum_{|k-nx|<\delta nx} f(k)\pi_{n,k}(x)| \leq \max_{0\leq k\leq n} |f(k)| \cdot \frac{2}{e^{nx\delta^2/3}} .$$

2.1 Function f(k) = log(k)

Our current tactical goal is to find asymptotic of the Bernoulli sum of the function $f(k) = \log k$. Let us observe that we cannot apply Bernstein polynomials directly to this function, since it cannot be extended to a continuous function on the interval $[0, 1]$. But, let us fix $x \in (0, 1)$. We shall replace the function f by a function $g : [0, 1] \to \mathbb{R}$ such that $g^{(6)}$ is bounded in $[0, 1]$ and $g(t) = \log(t)$ for $t > \frac{1}{2}x$. Then

$$B_n^g(x) = \sum_{k\leq nx/2} g\left(\frac{k}{n}\right)\pi_{n,k}(x) + \sum_{k>nx/2} \log\left(\frac{k}{n}\right)\pi_{n,k}(x).$$

From Lemma 1 applied for $\delta = \frac{1}{2}$ we get

$$|\sum_{k\leq nx/2} g\left(\frac{k}{n}\right)\pi_{n,k}(x)| \leq \sup\{|g(t)| : 0 \leq t \leq \frac{x}{2}\} \sum_{k\leq nx/2} \pi_{n,k}(x) = \mathrm{O}\left(\frac{1}{e^{nx/12}}\right) .$$

Hence

$$B_n^g(x) = \mathrm{O}\left(\frac{1}{e^{nx/12}}\right) + \sum_{k>nx/2} \log\left(\frac{k}{n}\right)\pi_{n,k}(x) =$$

$$\mathrm{O}\left(\frac{1}{e^{nx/12}}\right) + \sum_{k>nx/2} \log(k)\pi_{n,k}(x) - \log(n)\sum_{k>nx/2} \pi_{n,k}(x) =$$

$$\mathrm{O}\left(\frac{1}{e^{nx/12}}\right) + \sum_{k>nx/2} \log(k)\pi_{n,k}(x) - \log(n) + \mathrm{O}\left(\frac{\log n}{e^{nx/12}}\right) .$$

Let us observe that

$$0 \leq \sum_{1\leq k\leq nx/2} \log(k)\pi_{n,k}(x) \leq \log(\frac{nx}{2}) \sum_{1\leq k\leq nx/2} \pi_{n,k}(x) = \mathrm{O}\left(\frac{\log n}{e^{nx/12}}\right) .$$

Therefore,

$$\sum_{k>0} \log(k)\pi_{n,k}(x) = B_n^g(x) + \log(n) + \mathrm{O}\left(\frac{\log n}{e^{nx/12}}\right).$$

Using the third order approximation of B_n^g from Equation 4 we get

$$\sum_{k>0} \log(k)\pi_{n,k}(x) = \log(x) - \frac{1-x}{2xn} - \frac{5-6x+x^2}{12x^2n^2} + \mathrm{O}\left(\frac{1}{n^3}\right) + \log(n) + \mathrm{O}\left(\frac{\log n}{e^{nx/612}}\right)$$

and we derive finally the formula

$$\sum_{k\geq 1} \log(k)\pi_{n,k}(x) = \log(nx) - \frac{1-x}{2xn} - \frac{5-6x+x^2}{12x^2n^2} + \mathrm{O}\left(\frac{1}{n^3}\right) ,$$

so we get precisely the same formula as in Flajolet (1999).

2.2 Function f(k) = k log(k)

In a similar way we are able to find an approximation of Bernoulli sums of the function $f(k) = k\log(k)$. It is sufficient to observe that

$$\sum_{k\geq 1} \frac{k}{n}\log\left(\frac{k}{n}\right)\pi_{n,k}(x) = \frac{1}{n}\left(\sum_{k\geq 1} k\log(k)\pi_{n,k}(x) - \log(n)\sum_{k\geq 1} k\pi_{n,k}(x)\right) .$$

Hence

$$\sum_{k\geq 1} k\log(k)\pi_{n,k}(x) = nB_n^f(x) + xn\log(n) ,$$

so

$$\sum_{k\geq 1} k\log(k)\pi_{n,k}(x) = n\left(\frac{1-x}{2n} + \frac{1-x^2}{12n^2x} + x\log(x) + \mathrm{O}\left(\frac{1}{n^3}\right)\right) + xn\log(n)\ ,$$

therefore we finally get

$$\sum_{k\geq 1} k\log(k)\pi_{n,k}(x) = nx\log(nx) + \frac{1-x}{2} + \frac{1-x^2}{12nx} + \mathrm{O}\left(\frac{1}{n^2}\right)\ .$$

2.3 Formulas From Other Papers

Here is a summary of some formulas which we checked using the method described above:

1. $\sum_{k=1}^{n} \log(k)\pi_{n,k} = \log(nx) - \frac{1-x}{2xn} - \frac{5-6x+x^2}{12x^2n^2} + \mathrm{O}\left(\frac{1}{n^3}\right)$,
2. $\sum_{k=1}^{n} k\log(k)\pi_{n,k}(x) = nx\log(nx) + \frac{1-x}{2} + \frac{1-x^2}{12nx} + \mathrm{O}\left(\frac{1}{n^2}\right)$,
3. $\sum_{k=1}^{n} \frac{1}{\sqrt{k}}\pi_{n,k}(x) = \frac{1}{\sqrt{nx}}\left(1 + \frac{3(1-x)}{8nx} + \frac{5(1-x)(13-5x)}{128n^2x^2}\right) + \mathrm{O}\left(\frac{1}{n^3}\right)$,
4. $\sum_{k=1}^{n} \frac{1}{k}\pi_{n,k}(x) = \frac{1}{nx} + \frac{1-x}{x^2n^2} + \mathrm{O}\left(\frac{1}{n^3}\right)$,
5. If $r > 0$ then
$\sum_{k=1}^{n} \frac{1}{k^r}\pi_{n,k}(x) = \frac{1}{(nx)^r}\left(1 + \frac{r(r+1)(1-x)}{2nx} + \frac{r(r+1)(r+2)(1-x)(4+3r(1-x)-x)}{24(nx)^2} + \mathrm{O}\left(\frac{1}{n^3}\right)\right)$.

Formulas (1), (3) and (4) may be found in Flajolet (1999) and formula (5) may be found in Žnidarič (2009).

2.4 Transfer Lemma

In general, the method described in this section works with the class of functions of the form $f(x) = x^a(\log(x))^b$ where a is an arbitrary complex number and b is a positive integer. The following simple consequence of the binomial formula explains how to recover Bernoulli sums of functions of this form from corresponding Bernstein polynomials:

Lemma 2 (Transfer lemma) *Let $n > 0$ be an integer and let a be a fixed real number and let b be a natural number. For an arbitrary natural number c we put $f_c(x) = x^a(\log x)^c$. Let $B_c(x) = \sum_{k=1}^{n} f_c\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}$. Then*

$$\sum_{k=1}^{n} k^a\left(\log k\right)^b\binom{n}{k}x^k(1-x)^{n-k} = n^a\sum_{c=0}^{b}\binom{b}{c}B_c(x)(\log n)^{b-c}\ .$$

Proof: Observe that $(\log k)^b = (\log\frac{k}{n} + \log n)^b$. Hence

$$k^a(\log k)^b = n^a\sum_{c=0}^{b}\binom{b}{c}\left(\frac{k}{n}\right)^a\left(\log\frac{k}{n}\right)^c(\log n)^{b-c}\ ,$$

from which the thesis follows immediately. □

If we apply the second order approximation of Bernstein polynomials to the Transfer Lemma we obtain the following result:

Corollary 1 *Let a be a fixed real number and let b be a natural number. Then*

$$\sum_{k=1}^{n} k^a(\log k)^b \binom{n}{k} x^k(1-x)^{n-k} = (nx)^a \left(\log(nx)\right)^b +$$
$$\frac{(nx)^{a-1}(1-x)}{2}\Big((b-1)b(\log(nx))^{b-2} + (2a-1)b(\log(nx))^{b-1} + (a-1)a(\log(nx))^b\Big) +$$
$$\mathrm{O}\left(n^{a-2}(\log(n))^b\right) .$$

From this equality we can automatically derive second order approximations of all sums considered in Section 2.3. In a similar way we can derive a formula which gives approximations of Bernoulli sums of a function of the form $x^a(\log x)^b$ of any required order. We do not quote these formulas in this paper because they are too long. However, they can be easily derived and manipulated by symbolic computation packages.

3 More advanced formula

We will show that we can extend the class of functions which can be analyzed by the method described in the previous section. We show that on one example from Flajolet (1999). Namely, let $f(k) = \log(k!)$. This function does not belong to the class described in the previous section. But

$$f(k) = k\log(k) - k + \frac{1}{2}\log(k) + \frac{1}{2}\log(2\pi) + \frac{1}{12k} + \mathrm{O}\left(\frac{1}{k^3}\right) ,$$

so we may split the function f into functions $f_1(k) = k\log(k)$, $f_2(k) = -k$, $f_3(k) = \frac{1}{2}\log(k)$, $f_4(k) = \frac{1}{2}\log(2\pi)$, $f_5(k) = \frac{1}{12k}$ and $f_6(k) = \mathrm{O}\left(\frac{1}{k^3}\right)$ For functions f_1, f_3 and f_5 we may use formulas (1), (2) and (4) from Section 2.3. Notice that $S_n^{f_2}(x) = -nx$ and that $S_n^{f_4}(x) = \frac{1}{2}\log(2\pi)$. For the function f_6 we may use the formula (5) from Section 2.3 and deduce that $S_n^{f_6}(x) = \mathrm{O}\left(\frac{1}{n^3}\right)$. Finally, adding up all of the factors, with a little help of symbolic computation packages we derive the formula

$$\sum_k \log(k!)\pi_{n,k}(x) = \log\Big((nx)^{nx}\mathrm{e}^{-nx}\sqrt{2\pi nx}\Big) + \frac{1-x}{2} - \frac{1-3x+x^2}{12xn} + \mathrm{O}\left(\frac{1}{n^2}\right) \qquad (5)$$

from Flajolet (1999). Let us remark that in Flajolet (1999) the derivation of this formula is based on Mellin's transform and complex analysis and that our approach is elementary.

4 Entropy of Multinomial Distribution

The entropy of a discrete random variable X with possible values in a set A is defined by the formula

$$H(X) = -\sum_{a\in A} \log(\Pr(X=a))\Pr(X=a)$$

(in the case of $\Pr(X = a) = 0$ for some $a \in A$, the value of the corresponding summand $0\log(0)$ is taken to be 0).

For a sequence $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$ of natural numbers we put $|\vec{\alpha}| = \alpha_1 + \ldots + \alpha_m$ and $\vec{\alpha}! = (\alpha_1!)\cdots(\alpha_m!)$. If $\vec{\alpha} \in \mathbb{N}^m$ and $n = |\vec{\alpha}|$ then we put $\binom{n}{\vec{\alpha}} = \frac{n!}{\vec{\alpha}!}$. Moreover, if $\vec{\alpha} \in \mathbb{N}^m$ and $\vec{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$ then we put $\vec{x}^{\vec{\alpha}} = x_1^{\alpha_1} \cdots x_m^{\alpha_m}$.

Let us fix a vector $\vec{p} = (p_1, \ldots, p_m)$ such that $p_1, \ldots, p_m > 0$ and $p_1 + \cdots + p_m = 1$. Let us recall that a random vector $\vec{X} = (X_1, \ldots, X_m)$ has multinomial distribution with parameters n and $\vec{p}$ if

$$\Pr[\vec{X} = \vec{\alpha}] = \binom{n}{\vec{\alpha}} \vec{p}^{\,\vec{\alpha}}$$

for such sequences on non negative integers $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$ that $|\vec{\alpha}| = n$. Notice that for $m = 2$ the multinomial distribution with parameters n, (p_1, p_2) coincides with binomial distribution with parameters n and p_1 (after projection on first axis). It was shown in Jacquet and Szpankowski (1999) that the entropy of a random variable with the binomial distribution with parameters n and p equals $\frac{1}{2}\log(2\pi e n p(1-p)) + \mathrm{o}(1)$ (in fact they showed a full asymptotic expansion with complicated expressions). The next theorem formulates this result for accuracy of order $\mathrm{O}(n^{-2})$ in a clearer form.

Theorem 1 *The entropy of a random vector with multinomial distribution with parameters n and $p_1 > 0, \ldots, p_m > 0$ is equal to*

$$\frac{1}{2}\log\left((2\pi n \mathrm{e})^{m-1} p_1 \cdots p_m\right) + \frac{1}{12n}\left(3m - 2 - \sum_{j=1}^{m} \frac{1}{p_j}\right) + \mathrm{O}\left(\frac{1}{n^2}\right) . \tag{6}$$

Proof: Let us fix $n \in \mathbb{N}$ and $\vec{p} \in (0,1]^m$ such that $|p| = 1$. For $\vec{\alpha} \in \mathbb{N}^m$ we put $\pi_{\vec{\alpha}} = \binom{n}{\vec{\alpha}} p^{\vec{\alpha}}$. Notice that

$$\sum_{|\vec{\alpha}|=n} \log(\pi_{\vec{\alpha}})\pi_{\vec{\alpha}} = \log(n!) - A + B$$

where $A = \sum_{k=1}^{m} \sum_{|\vec{\alpha}|=n} \log(\alpha_k!)\pi_{\vec{\alpha}}$ and $B = \sum_{|\alpha|=n} \sum_{k=1}^{m} \alpha_k \log(p_k)\pi_{\vec{\alpha}}$. We will also use the entropy function $H(\vec{x}) = -\sum_{k=1}^{m} x_i \log(x_i)$. Note that

$$B = \sum_{k=1}^{m} \log(p_k) \sum_{|\alpha|=n} \alpha_k \pi_\alpha = \sum_{k=1}^{m} \log(p_k) n p_k = -nH(\vec{p}).$$

We transform formula (5) into a form more convenient for our purposes, namely,

$$\sum_a \log(a!)\pi_{n,a}(p) =$$
$$nx\log n + nx\log x - x\left(n + \frac{1}{2} + \frac{1}{12n}\right) + \log(\sqrt{x}) + \log(\sqrt{2\pi n \mathrm{e}}) + \frac{3}{12n} - \frac{1}{12nx} + \mathrm{O}\left(\frac{1}{n^2}\right)$$

and we get

$$\sum_{k=1}^{m}\sum_{|\vec{\alpha}|=n}\log(\alpha_k!)\pi_{\vec{\alpha}} + \mathrm{O}\left(\frac{1}{n^2}\right) =$$

$$n\log n - nH(p) - \left(n+\frac{1}{2}+\frac{1}{12n}\right) + \log(\sqrt{p_1\cdots p_m}) + m\log(\sqrt{2\pi n\mathrm{e}}) + \frac{3m}{12n} - \frac{1}{12n}\sum_{k=1}^{m}\frac{1}{p_i}\ .$$

We use finally the Stirling formula for approximation of $\log(n!)$ and we get required formula. □

The second term of the expression (6) is in a closed form. The same term from Jacquet and Szpankowski (1999) (Theorem 2) for the entropy of binomial distribution contains a complicated infinite sum containing coefficients of the expansion of the function $\exp(x\ln(1+\lambda(\mathrm{e}^y-1)) - \lambda xy)$ as an infinite double series of variables x and y. Flajolet (1999) calculated the entropy of binomial distribution with accuracy of order $\mathrm{O}(n^{-1})$. Let us finally mention that in Knessl (1998) a similar formula for the entropy of binomial distribution was derived.

5 Application of Bernstein Polynomial Defined on a Simplex

Let us fix an integer $m > 0$. Let $\sigma_m = \{(x_1,\ldots,x_m) \in [0,1]^m : x_1+\ldots+x_m \leq 1\}$ and $\Sigma_m = \{(\lambda_0,\ldots,\lambda_m) \in [0,1]^{m+1} : \lambda_0+\ldots+\lambda_m = 1\}$. We put $\lambda(\vec{x}) = (1-|\vec{x}|, x_1,\ldots,x_m)$ for $\vec{x}\in\sigma_m$. Observe that $\lambda : \sigma_n \to \Sigma_m$ is a bijection. Notice that $\lambda(\vec{x})$ are the barycentric coordinates of the point $\vec{x}$ in the simplex generated by points $\vec{e}_0 = (0,0,\ldots,0)$, $\vec{e}_1 = (1,0,\ldots,0)$, $\vec{e}_2 = (0,1,\ldots,0)$, $\ldots$, $\vec{e}_m = (0,0,\ldots,0,1)$.

Let us recall that for $\vec{\lambda} = (\lambda_0,\ldots,\lambda_m) \in \mathbb{R}^{m+1}$ and $\vec{\alpha} = (\alpha_0,\ldots,\alpha_m) \in \mathbb{N}^{m+1}$ we put $|\vec{\alpha}| = \alpha_0+\ldots+\alpha_m$, $\vec{\alpha}! = \alpha_0!\cdots\alpha_m!$, $\binom{n}{\vec{\alpha}} = \frac{n!}{\vec{\alpha}!}$ (where $n = |\vec{\alpha}|$) and $\vec{\lambda}^{\,\vec{\alpha}} = \prod_{i=0}^{m}\lambda_i^{\alpha_i}$.

For a vector $\beta = (\beta_1,\ldots,\beta_m)$ of natural numbers we put $D_\beta = \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1}\cdots\partial x_m^{\beta_m}}$. Finally, by $C^k(\sigma_m)$ we denote the class of all real functions on the simplex σ_m such that $D_\beta(f)$ is continuous for each vector β such that $|\beta| = k$.

The Bernstein polynomial of degree n on a simplex σ_m of a function $f : \sigma_m \to \mathbb{R}$ (see Jian and Wu (1988)) is defined as

$$B_n(f,\vec{x}) = \sum_{|\alpha|=n} f(x_\alpha)\binom{n}{\alpha}\vec{\lambda}^{\vec{\alpha}},$$

where the sum is taken over all $\vec{\alpha} = (\alpha_0,\ldots,\alpha_m) \in \mathbb{N}^{m+1}$, $\vec{\lambda} = \lambda(\vec{x})$ and $x_\alpha = \left(\frac{\alpha_1}{n},\ldots,\frac{\alpha_m}{n}\right)$. For any continuous function f defined on the simplex σ_n the multivariate Bernstein polynomials $B_n(f,x)$ converges uniformly to $f(x)$ (see e.g. Lai (1992)). Moreover (see Feng and Kozak (1992)), if $f \in C^{2s}(\sigma_m)$ then

$$B_n(f,\vec{x}) = f(\vec{x}) + \sum_{k=1}^{2s}\sum_{|\beta|=k}\frac{1}{n^k}\frac{T_{n,\beta}(\vec{x})}{k!}D_\beta f(\vec{x}) + \mathrm{O}\left(\frac{1}{n^s}\right) \tag{7}$$

where $T_{n,\beta}(\vec{x}) = \sum_{|\vec{\alpha}|=n}(\vec{\alpha}-n\vec{x})^\beta\binom{n}{\vec{\alpha}}\lambda(\vec{x})^{\vec{\alpha}}$. Observe that $T_{n,\beta}(\vec{x})$ is the βth moment of a random vector with multinomial distribution with parameters n and $\lambda(\vec{x})$. Hence, we have $T_{n,\beta}(\vec{x}) = 0$ if $|\beta| = 1$.

Other βth moments for moderate $|\beta|$ can be calculated from the central moment generating function for the multinomial distribution, i.e. from the function

$$m(t_0, \ldots, t_m) = \left(\frac{\mathrm{e}^{t_0}\lambda_0 + \ldots + \mathrm{e}^{t_m}\lambda_m}{\exp(\lambda_0 t_0 + \ldots + \lambda_m t_m)} \right)^n .$$

Let

$$A_1(f)(\vec{x}) = \frac{1}{2} \left(\sum_{j=1}^{m} x_j \frac{\partial^2 f}{\partial x_j^2} - \sum_{i,j=1}^{m} x_i x_j \frac{\partial^2 f}{\partial x_i \partial x_j} \right) \tag{8}$$

If we apply formula (7) to $s = 2$ then we get the following approximation for multinomial Bernstein polynomials

$$B_n(f, \vec{x}) = f(\vec{x}) + \frac{1}{n} A_1(f)(\vec{x}) + \mathrm{O}\left(\frac{1}{n^2}\right) \tag{9}$$

for functions $f \in C^4(\sigma_m)$.

Let us fix $\vec{x} \in \sigma_m$. Let $\vec{\lambda} = \lambda(\vec{x})$. Let $\pi_{\vec{\alpha}} = \binom{|\vec{\alpha}|}{\vec{\alpha}} \vec{\lambda}^{\vec{\alpha}}$. We shall formulate a theorem which allows us to automatically translate multinomial Bernstein approximation formulas for corresponding multinomial sums.

Lemma 3 (Multinomial Transfer Lemma) *Let us fix integers $n, m > 0$, two sequences $a_1, \ldots, a_m$ and $b_1, \ldots, b_m$ of reals such that $b_1 > 0, \ldots, b_m > 0$. For an arbitrary natural number c we define*

$$f_c(x_1, \ldots x_m) = \left(\prod_{i=1}^{m} x_i^{a_i} \right) \left(\log \left(\sum_{i=1}^{m} b_i x_i \right) \right)^c .$$

Let

$$B_c(\vec{x}) = \sum_{|\vec{\gamma}|=n} f_c\left(\frac{\gamma_1}{n}, \ldots, \frac{\gamma_m}{n}\right) \pi_{\vec{\gamma}} , \quad S_c(\vec{x}) = \sum_{|\vec{\gamma}|=n} f_c(\gamma_1, \ldots, \gamma_m) \pi_{\vec{\gamma}} .$$

Then for each natural number c we have

$$S_c(\vec{x}) = n^{a_1 + \ldots + a_m} \sum_{d=0}^{c} \binom{c}{d} B_d(\vec{x}) (\log n)^{c-d}.$$

Proof: The proof mimics the proof of Lemma 2 and is based on the following observation:

$$\left(\prod_{i=1}^{m} \left(\frac{x_i}{n}\right)^{a_i} \right) \left(\log \left(\sum_{i=1}^{m} b_i \frac{x_i}{n} \right) \right)^c = \frac{1}{n^{a_1 + \ldots + a_m}} \left(\prod_{i=1}^{m} x_i^{a_i} \right) \left(\log \left(\sum_{i=1}^{m} b_i x_i \right) - \log(n) \right)^c$$

for fixed integers $n > 0$, $m > 0$, $a_1, \ldots, a_m$ and $b_1, \ldots, b_m$ and an arbitrary natural number c. $\square$

We finish this section with three examples of direct application of methods discussed above. The first two examples are artificial and the third example is from from Jacquet and Szpankowski (1999).

Example 1 Let $f(x,y)=\sqrt{xy}$. After application of equation (9) we get

$$B_n(f,x,y)=\sqrt{xy}-\frac{x+y}{8n\sqrt{xy}}+\mathrm{O}\left(\frac{1}{n^2}\right)\ ,$$

from which we conclude that

$$\sum_{a+b+c=n}\sqrt{ab}\binom{n}{a\,b\,c}x^a y^b(1-x-y)^c=n\sqrt{xy}-\frac{x+y}{8\sqrt{xy}}+\mathrm{O}\left(\frac{1}{n}\right)\ .$$

Example 2 Let $g(x,y)=\log(x)y^2$. Then from (9) we get

$$B_n(g,x,y)=\log(x)y^2-\frac{y}{2nx}\big(y+x(-2\log x+(3+2\log x)y)\big)$$

and after some transformations, using strong concentration of the multinomial distribution near its mean, in a similar way as in Section 2, we obtain

$$\sum_{\substack{a+b+c=n\\ a\geq 1}}\log(a)b^2\binom{n}{a\,b\,c}x^a y^b(1-x-y)^c=$$
$$\log(nx)(ny)^2+n\log(n)y(1-y)+n\left(\log(x)y(1-y)-y^2\left(\frac{3}{2}+\frac{1}{2x}\right)\right)+\mathrm{O}\left(1\right)\ .$$

Example 3 Let $f(x,y)=\log(x)^3y^2$. From Equation (9) and Theorem 3, after some transformations, we get approximation formula for the function mentioned in Concluding Remarks of Jacquet and Szpankowski (1999):

$$\sum_{a+b+c=n}(\log a)^3b^2\binom{n}{a\,b\,c}x^a y^b(1-x-y)^c=(\log(nx))^3(ny)^2+$$
$$\frac{\log(nx)ny}{x}\left(x(1-y)(\log(nx))^2-\frac{3}{2}(1+3x)y\log(nx)+3(1-x)y\right)+\mathrm{O}\left((\log(n))^3\right)\ .$$

6 Codes for Mathematica

The following Mathematica code may be used to compute polynomials $T_{n,a}(x)$ (i.e. ath moment of the Binomial distribution with parameters n and x) defined in Section 1.1:

```
T[n,0,x]=1;
T[n,1,x]=0;
T[n_,a_,x_]:= T[n,a,x]= x (1-x)(D[T[n,a-1,x],x]+ n (a-1)T[n,a-2,x])
```

The next Mathematica code may be used to calculate approximation formula 1 of the Bernstein polynomial of a function f with precision $\mathrm{O}\left(n^{-k}\right)$

```
BA[n_,f_,x_,k_] := Collect[
  f[x] + Sum[D[f[x], {x,a}]/(n^a a!) T[n,a,x], {a,2,2 k - 1}],
  n,
  Simplify]
```

7 Summary

We proposed a simple method for extracting asymptotics of Bernoulli sums of functions which can be approximated by linear combinations of functions of the form $x^a(\log x)^b$, where $a \in \mathbb{R}$ and $b \in \mathbb{N}$. Our method works with all examples from Flajolet (1999), can be extended onto the class of multinomial sums and can give approximation of any required precision. We have also shown that our method can be automated to a large extent.

Acknowledgements

The paper was supported by grant N N206 369739 of the Polish Ministry of Science and Higher Education.

References

S. Bernstein. Démonstration du théoreme de Weierstrass, fondée sur le calcul des probabilités. *Communications of the Kharkov Mathematical Society*, pages 1–2, 1912.

S. Bernstein. Complement a l'article de E. Voronovskaya. *C. R. Acad. URSS*, pages 86–92, 1932.

Y. Y. Feng and J. Kozak. Asymptotic expansion formula for bernstein polynomials defined on a simplex. *Constructive Approximation*, 8:49–58, 1992. ISSN 0176-4276. URL `http://dx.doi.org/10.1007/BF01208905`. 10.1007/BF01208905.

P. Flajolet. Singularity analysis and asymptotics of bernoulli sums. *Theor. Comput. Sci.*, 215(1-2):371–381, 1999.

P. Jacquet and W. Szpankowski. Entropy computations via analytic depoissonization. *IEEE Transactions on Information Theory*, 45(4):1072–1081, 1999.

R. Jian and Z. Wu. Bernstein polynomials defined on a simplex. *Acta Mathematica Sinica*, pages 510–522, 1988.

C. Knessl. Integral representations and asymptotic expansions for shannon and renyi entropies. *Applied Math Letters*, pages 69–74, 1998.

M.-J. Lai. Asymptotic formulae of multivariate bernstein approximation. *J. Approx. Theory*, 70:229–242, August 1992. ISSN 0021-9045. doi: 10.1016/0021-9045(92)90086-4. URL `http://dl.acm.org/citation.cfm?id=160278.160285`.

G. Lorentz. *Bernstein polynomials*. Chelsea Publishing Series. Chelsea Pub. Co., 1986. ISBN 9780828403238. URL `http://books.google.pl/books?id=nXiRAKzpdjUC`.

M. Mitzenmacher and E. Upfal. *Probability and Computing. Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

E. Voronovskaya. Determination de la forme asymptotique d'approximation des fonctions par les polynomes de M. Bernstein. *Doklady Akademii Nauk SSSR*, pages 79–85, 1932.

K. Weierstrass. Über die analytische darstellbarkeit sogenannter willkürlicher functionen einer reellen veränderlichen. *Sitzungsberichte der Königlich PreuSSischen Akademie der Wissenschaften zu Berlin*, 1885.

M. Žnidarič. Asymptotic expansion for inverse moments of binomial and poisson distributions. *The Open Statistics & Probability Journal*, 1:7–10, 2009.