Research Article

Accuracy of Approximation for Discrete Distributions

Tamás F. Móri

Department of Probability Theory and Statistics, Eötvös Loránd University, Pázmány P. s. 1/C, Budapest 1117, Hungary

Correspondence should be addressed to Tamás F. Móri; mori@math.elte.hu

Received 28 July 2016; Accepted 13 October 2016

Academic Editor: Ramón M. Rodríguez-Dagnino

Copyright © 2016 Tamás F. Móri. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper is a contribution to the problem of estimating the deviation of two discrete probability distributions in terms of the supremum distance between their generating functions over the interval $[0, 1]$. Deviation can be measured by the difference of the k th terms or by total variation distance. Our new bounds have better order of magnitude than those proved previously, and they are even sharp in certain cases.

1. Introduction

Dealing with random combinatorial structures needs to estimate the deviation of two discrete probability distributions in terms of the maximal distance between their generating functions over $[0, 1]$. This is often the case when given a couple of not necessarily independent random events, one needs to estimate the number N of those that occur. Among many popular methods of Poisson approximation, sieve methods with generalized Bonferroni bounds, such as the graph sieve [1], are at hand. They provide estimates not only for the probability that none of the events occur, but also for the difference between the generating function of N and that of the corresponding Poisson distribution over the interval $[0, 1]$ (see [2] for more details). This raises the following problem.

Let $\mathbf{p} = (p_0, p_1, \dots)$ and $\mathbf{q} = (q_0, q_1, \dots)$ be discrete probability distributions with generating functions $g_{\mathbf{p}}(z)$ and $g_{\mathbf{q}}(z)$, respectively. Given $\Delta = \max_{0 \leq z \leq 1} |g_{\mathbf{p}}(z) - g_{\mathbf{q}}(z)|$, estimate the distance $|p_k - q_k|$ of k th terms, or the total variation distance between $\mathbf{p}$ and $\mathbf{q}$ defined as $\|\mathbf{p} - \mathbf{q}\| = \sum_{k=0}^{\infty} |p_k - q_k|$.

The difficulty lies in the constraint that the difference of the generating functions is only available over the real interval $[0, 1]$, and not over the whole complex unit disc, which would make it possible to apply standard methods of characteristic functions.

Several positive and negative results were achieved in the last three decades, beginning with [3]; see Section 2.

Lower and upper estimates got closer and closer, but the final answer is still ahead. The aim of the present note is to provide new bounds that have better order of magnitude than those proved previously. They are even sharp in certain cases.

The paper is organized as follows. In Section 2 we introduce the necessary notions and notations and cite some earlier results. Section 3 is devoted to the case where $|p_k - q_k|$ is estimated in terms of Δ , while in Section 4 the total variation distance is treated.

2. Preliminaries

The difference of two probability generating function belongs to the following class. Let $\mathcal{F}$ denote the set of all functions $f: [0, 1] \rightarrow \mathbb{R}$ such that

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad \sum_{k=0}^{\infty} |a_k| \leq 2, \quad f(1) = 0. \quad (1)$$

On the other hand, every $f \in \mathcal{F}$ can be obtained as the difference of two probability generating functions, namely, $f(z) = g_1(z) - g_2(z)$, where g_1 and g_2 can be constructed in the following way. Let $0 \leq b_k \leq 1 - |a_k|$, $k = 0, 1, \dots$ (note that $|a_k| \leq 1$ for $f \in \mathcal{F}$) such that

$$\sum_{k=0}^{\infty} b_k = 1 - \frac{1}{2} \sum_{k=0}^{\infty} |a_k| = 1 - \sum_{k=0}^{\infty} a_k^+ = 1 - \sum_{k=0}^{\infty} a_k^-, \quad (2)$$

where $x^+ = \max\{x, 0\}$, $x^- = (-a_k)^+$. Then set

$$\begin{aligned} g_1(z) &= \sum_{k=0}^{\infty} (a_k^+ + b_k) z^k, \\ g_2(z) &= \sum_{k=0}^{\infty} (a_k^- + b_k) z^k. \end{aligned} \quad (3)$$

For $f \in \mathcal{F}$ let $\Delta = \Delta(f) = \max_{0 \leq z \leq 1} |f(z)|$, and $L = \log(1/\Delta)$. For $d \in (0, 1]$ and $k = 0, 1, \dots$ define

$$\varphi_k(d) = \sup \{|a_k| : f \in \mathcal{F}, \Delta(f) = d\}. \quad (4)$$

In [2] it was shown that, for every $k = 0, 1, \dots$,

$$C_k \Delta L^{2k} \leq \varphi_k(\Delta) \leq \Delta \exp(2(kL)^{4/5}) \quad (5)$$

holds with a suitable positive constant C_k , if Δ is sufficiently small. In fact, the upper estimate is valid for every $\Delta \in (0, 1]$, but it becomes trivial for $\Delta \geq \exp(-32k^4)$ by elementary calculus. Though in both the upper and lower bounds the multiplier of Δ is slowly varying as $\Delta \rightarrow 0$, they are not of the same order of magnitude.

It is easy to see that $\sum_{k=0}^{\infty} |a_k|$ cannot be estimated in a nontrivial way, because for arbitrary $\Delta \in (0, 1]$ there exists a function $f \in \mathcal{F}$, such that $\Delta(f) = \Delta$ and $\sum_{k=0}^{\infty} |a_k| = 2$, maximal. Indeed, let $n > 1/\Delta$ and

$$\begin{aligned} f(z) &= \Delta - \Delta z + (1 - \Delta) z^n - (1 - \Delta) z^{n+1} \\ &= \Delta - \Delta z + z(1 - \Delta)(z^{n-1} - z^n). \end{aligned} \quad (6)$$

Then $0 \leq f(z) \leq \Delta$ for $z \in [0, 1]$, because

$$z^{n-1} - z^n \leq \left(\frac{n-1}{n}\right)^{n-1} \frac{1}{n} < \frac{1}{n-1} < \frac{\Delta}{1-\Delta} \quad (7)$$

by the choice of n , provided $\Delta < 1$. For $\Delta = 1$ the estimation obviously holds. Hence $\Delta(f) = f(0) = \Delta$.

However, if $f = g_{\mathbf{p}} - g_{\mathbf{q}}$, and one of the distributions, say $\mathbf{p}$, is fixed (as in the case of Poisson approximation), then the class of feasible functions is smaller; thus the bounds for $|a_k|$ may decrease, and even the total variation distance can be estimated nontrivially. The following results can be found in [2].

Let $\mathbf{p}$ be a fixed discrete probability distribution such that $p_k > 0$ for every $k = 0, 1, \dots$, and

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \log \frac{1}{p_k} < \infty. \quad (8)$$

Let k be a positive integer and C a sufficiently small positive constant. Then for every sufficiently small positive Δ there exists a discrete probability distribution $\mathbf{q}$, such that $\Delta(g_{\mathbf{p}} - g_{\mathbf{q}}) = \Delta$, and

$$|p_k - q_k| > C \Delta L^{2k}. \quad (9)$$

If the tail of $\mathbf{p}$ is lighter than exponential, the lower estimate decreases.

Instead of (8), suppose that

$$\limsup_{k \rightarrow \infty} \frac{1}{h(k)} \log \frac{1}{p_k} \quad (10)$$

is positive and finite, where h is a positive, continuous, and increasing function, regularly varying at infinity, and $\lim_{k \rightarrow \infty} h(k)/k = \infty$. Let k be a positive integer and C a sufficiently small positive constant. Then for every sufficiently small positive Δ there exists a discrete probability distribution $\mathbf{q}$, such that $\Delta(g_{\mathbf{p}} - g_{\mathbf{q}}) = \Delta$, and

$$|p_k - q_k| > C \Delta (h^{-1}(L))^{2k}. \quad (11)$$

Particularly, when $\mathbf{p}$ is the Poisson distribution with parameter λ , it follows that for every sufficiently small positive Δ there exists a discrete probability distribution $\mathbf{q}$, such that $\Delta(g_{\mathbf{p}} - g_{\mathbf{q}}) = \Delta$, and

$$|p_k - q_k| > C_k \Delta \left(\frac{L}{\log L} \right)^{2k}. \quad (12)$$

(The constant C_k does not depend on λ and $\mathbf{q}$. The parameter λ only appears in the bounds implicit in the phrase “sufficiently small.”)

Let us turn to the case of total variation. In [2] for every fixed $\mathbf{p}$ an increasing function $\psi : [0, 1] \rightarrow \mathbb{R}$ was constructed in such a way that $\lim_{t \rightarrow 0} \psi(t) = 0$ and $\|\mathbf{p} - \mathbf{q}\| \leq \psi(\Delta)$ for arbitrary $\mathbf{q}$. However, apart from the case where the tail of $\mathbf{p}$ was extremely light, the function ψ proved to be slowly varying at 0, which is just a little bit better than nothing. For example, in the case of Poissonian $\mathbf{p}$, the following inequality was obtained:

$$\|\mathbf{p} - \mathbf{q}\| = o(\exp(-L^{1/4})), \quad (13)$$

as $\mathbf{q}$ varies in such a way that $\Delta \rightarrow 0$.

Since $\|\mathbf{p} - \mathbf{q}\| \geq |p_k - q_k|$, every lower estimate obtained for fixed k will do for the total variation. However, if p_k does not decrease faster than exponential, that is, condition (8) is fulfilled, there is a lower estimate of the form

$$\|\mathbf{p} - \mathbf{q}\| \geq \Delta^C \quad (14)$$

with an exponent $C \in (0, 1)$ depending on $\mathbf{p}$.

When the tail of $\mathbf{p}$ is lighter than exponential, namely, condition (10) holds, then for every sufficiently small positive Δ the following lower estimate is valid.

$$\|\mathbf{p} - \mathbf{q}\| \geq \Delta \exp(Ch^{-1}(L)), \quad (15)$$

with a constant C depending on $\mathbf{p}$. Particularly, in the case where $\mathbf{p}$ is Poisson,

$$\|\mathbf{p} - \mathbf{q}\| > \Delta \exp\left(1.76 \frac{L}{\log L}\right) \quad (16)$$

was proved.

3. Estimation for the Difference of k th Terms

The following important result can be traced back to Markoff, 1892 [4], who dealt with the extremal properties of Chebyshev polynomials over the interval $[-1, 1]$; see Chapter 2 of [5]. The proof can be found in [6] or in [7].

Theorem 1. *Let f be a polynomial of degree less than or equal to n , and $0 \leq k \leq n$. Then*

$$|a_k| \leq \frac{(2n)^{2k}}{(2k)!} \Delta. \quad (17)$$

Using this result an upper bound can be proved without any restriction on f , which is of the same order as the lower bound in the left-hand side of (5)

Theorem 2. *Let $0 \leq k \leq 2L$. Then*

$$|a_k| \leq \frac{e^{2k+1}}{(2k)!} \Delta L^{2k} \leq \frac{e^3}{2} \Delta L^{2k}. \quad (18)$$

Proof. Suppose first that $\Delta \leq e^{-2k}$. Let $c = e^{-2}$ and $n = \lfloor L/2 \rfloor$. Then $n \geq k$, and

$$\begin{aligned} \max_{0 \leq y \leq 1} \left| \sum_{j=0}^n a_j c^j y^j \right| &= \max_{0 \leq t \leq c} \left| \sum_{j=0}^n a_j t^j \right| \\ &\leq \max_{0 \leq t \leq 1} \left| \sum_{j=0}^{\infty} a_j t^j \right| + \max_{0 \leq t \leq c} \left| \sum_{j=n+1}^{\infty} a_j t^j \right| \\ &\leq \Delta + \sum_{j=n+1}^{\infty} c^j = \Delta + \frac{1}{1-c} e^{-2(n+1)}. \end{aligned} \quad (19)$$

Here $e^{-2(n+1)} < e^{-L} = \Delta$; therefore the right-hand side is less than $e\Delta$. By Theorem 1 we then have

$$|a_k c^k| \leq \frac{(2n)^{2k}}{(2k)!} e\Delta \leq \frac{e}{(2k)!} \Delta L^{2k}, \quad (20)$$

as claimed.

If $e^{-2k} < \Delta \leq e^{-k/2}$, then the upper bound, being greater than 1, is trivial. Indeed, in that interval, ΔL^{2k} is decreasing, hence the upper bound attains its minimum at $\Delta = e^{-k/2}$. Then its value is

$$\frac{e^{2k+1}}{(2k)!} \Delta L^{2k} = \frac{1}{(2k)!} \left(\frac{k}{2} \right)^{2k} e^{1+3k/2}. \quad (21)$$

Here the right-hand side is equal to e for $k = 0$ and to $(1/8)e^{5/2} > 1$ for $k = 1$. Stepping further from k to $k+1$, the right-hand side gets multiplied by

$$\begin{aligned} &\frac{1}{(2k+1)(2k+2)} \cdot \frac{(k+1)^{2(k+1)}}{k^{2k}} \cdot \frac{e^{3/2}}{4} \\ &> \left(\frac{k+1}{k} \right)^{2k} \frac{e^{3/2}}{16} > \frac{e^{3/2}}{4} > 1, \end{aligned} \quad (22)$$

where we used the fact that $((k+1)/k)^k$ is increasing; hence it is not less than 2.

Finally, it is easy to see that $e^{2k+1}/(2k)!$ is decreasing for $k \geq 1$, from which the second inequality follows. $\square$

If $\mathbf{p}$ satisfies (8), that is, p_k cannot decrease faster than exponential, then (9) implies that the estimate of Theorem 2 is sharp in the order of magnitude.

Theorem 3. *Let $a_k = p_i - q_k$, where $\mathbf{p} = (p_k)_{k \geq 0}$ and $\mathbf{q} = (q_k)_{k \geq 0}$ are discrete probability distributions. Suppose that*

$$\sum_{k=n}^{\infty} p_k \leq F(n), \quad (23)$$

where F is a positive, continuous, and strictly decreasing function tending to zero. Then

$$\begin{aligned} |p_k - q_k| &\leq 3 \frac{(8e+4)^k}{(2k)!} \Delta \left(F^{-1}(\Delta) \right)^{2k} \\ &\leq 2(2e+1)^2 \Delta \left(F^{-1}(\Delta) \right)^{2k} \end{aligned} \quad (24)$$

for every $k < F^{-1}(\Delta)$.

Proof. Let $n = \lceil F^{-1}(\Delta) \rceil$, and introduce

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} (q_k - p_k) z^k, \\ g(z) &= \sum_{k=0}^{\infty} q_k z^k - \sum_{k=0}^{n-1} q_k z^k, \\ h(z) &= \sum_{k=0}^{\infty} q_k z^k, \end{aligned} \quad (25)$$

$$0 \leq z \leq 1.$$

Then

$$|f(z) - g(z)| \leq \sum_{k=n}^{\infty} p_k \leq F(n) \leq \Delta. \quad (26)$$

Hence

$$\sup_{0 \leq z \leq 1} |g(z)| \leq 2\Delta. \quad (27)$$

Let $c = 2e/(2e+1)$, and $z_i = (1-c) + c(i/n)$, $i = 0, \dots, n$. Let us approximate $g(z)$ with its Lagrange interpolation polynomial $p_n(z)$ over the uniform grid $\{z_i : 0 \leq i \leq n\}$. That is,

$$p_n(z) = \sum_{i=0}^n g(z_i) \prod_{j \neq i} \frac{z - z_j}{z_i - z_j}. \quad (28)$$

Then $g(z) - p_n(z)$ has $n+1$ zeros in the interval $[1-c, 1]$; hence, there exists a $\xi \in [1-c, 1]$ such that

$$\begin{aligned} g^{(n)}(\xi) &= p_n^{(n)}(\xi) = \sum_{i=0}^n g(z_i) \binom{n}{i} \left(\frac{n}{c} \right)^n (-1)^{n-i} \\ &\leq 2\Delta \left(\frac{2n}{c} \right)^n < \Delta \left(\frac{2e}{c} \right)^n n! = \Delta (1-c)^{-n} n!. \end{aligned} \quad (29)$$

Since $h^{(n)}$ is increasing in the interval $[0; 1]$, and it is equal to $g^{(n)}$, we have that $0 \leq h^{(n)}(z) \leq \Delta(1-c)^{-n}n!$ for all $z \in [0; 1-c]$. Moreover, $h(0) = h'(0) = \dots = h^{(n-1)}(0) = 0$; thus

$$0 \leq h(z) \leq \Delta \left(\frac{z}{1-c} \right)^n, \quad 0 \leq z \leq 1-c. \quad (30)$$

From all these it follows that

$$\left| \sum_{k=0}^{n-1} (p_k - q_k) z^k \right| \leq 3\Delta \quad (31)$$

for $0 \leq z \leq 1-c$. Applying Theorem 1 in the same way as in the proof of Theorem 2 we get that

$$\begin{aligned} |p_k - q_k| &\leq \frac{3\Delta}{(1-c)^k} \frac{(2(n-1))^{2k}}{(2k)!} \\ &\leq 3 \frac{(8e+4)^k}{(2k)!} \Delta \left(F^{-1}(\Delta) \right)^{2k}, \end{aligned} \quad (32)$$

if $k < F^{-1}(\Delta)$. Investigating the ratio of consecutive terms one can easily see that

$$3 \frac{(8e+4)^k}{(2k)!} \leq 3 \frac{(8e+4)^2}{4!} = 2(2e+1)^2. \quad (33)$$

□

Remark 4. Let us write F in exponential form: $F(n) = e^{-h(n)}$, where h is nonnegative, continuous, and tends increasingly to infinity. Then $F^{-1}(\Delta) = h^{-1}(L)$. The estimate of Theorem 3 is better than that of Theorem 2 in the order of magnitude only if the distribution $\mathbf{p}$ has a tail lighter than exponential; that is, $h(n)/n \rightarrow \infty$.

Remark 5. If $1/h(k) \log(1/p_k)$ is bounded away from zero and infinity as well, then both the upper estimate of Theorem 3 and the lower estimate of (11) are applicable and they are of the same order of magnitude. Use the fact that for regularly varying functions h of order $\alpha > 0$ we have $(ch)^{-1}(z) \sim c^{-1/\alpha} h^{-1}(z)$ as $z \rightarrow \infty$.

Remark 6. Reference [6] proved similar bounds with different constants, but of the same order of magnitude. However, they imposed conditions similar to (10) on the sequence $a_k = p_k - q_k$, rather than on $\mathbf{p}$, which is less useful for applications in probability. Besides, for the estimate of Theorem 2, which is true without any restriction on the coefficients, they needed exponential decay of the sequence (a_k) .

Remark 7. If h is linear, that is, $\sum_{k \geq n} p_k \leq e^{-\nu n}$, and it holds with a sufficiently large ν ($\nu \geq 2$ will do), the upper bound of Theorem 3 is better for $k > 0$ than that of Theorem 2. (For $k = 0$ the bound $|p_0 - q_0| \leq \Delta$ is obviously the best possible.)

Particularly, let $\mathbf{p}$ be the Poisson distribution with mean λ . Then, for $n > \lambda$, we have

$$\sum_{k=n}^{\infty} p_k \leq e^{-\lambda} \left(\frac{\lambda e}{n} \right)^n \quad (34)$$

(see Theorem A.15 in [8]); hence $h(x) = x \log x + O(x)$, and

$$h^{-1}(x) = \frac{x}{\log x - \log \log x + O(1)} \sim \frac{x}{\log x} \quad (35)$$

as $x \rightarrow \infty$. In addition, $h^{-1}(x) > x/\log x$ eventually. Let us plug this back into Theorem 3 to get the following estimate.

Corollary 8. *If $\mathbf{p}$ is Poisson, then, uniformly in $\mathbf{q}$, one has*

$$|p_k - q_k| \leq 4 \frac{(8e+4)^k}{(2k)!} \Delta \left(\frac{L}{\log L} \right)^{2k} \quad (36)$$

for every $k < L/\log L$, if Δ is sufficiently small.

Note that the order of this upper bound is the same as that of the lower bound (12).

4. Estimation for the Total Variation Distance

Let again $a_k = p_k - q_k$, where $\mathbf{p} = (p_k)_{k \geq 0}$ and $\mathbf{q} = (q_k)_{k \geq 0}$ are discrete probability distributions. As we have seen, if nothing is known about $\mathbf{p}$ and $\mathbf{q}$, it is impossible to give a nontrivial upper bound for the total variation distance

$$\|\mathbf{p} - \mathbf{q}\| = \sum_{k=0}^{\infty} |p_k - q_k|. \quad (37)$$

However, when $\mathbf{p}$ is fixed, the situation is completely different.

Theorem 9. *Let $a_k = p_k - q_k$, where $\mathbf{p} = (p_k)$ satisfies condition (23). Then*

$$\|\mathbf{p} - \mathbf{q}\| < 8\Delta \exp(\kappa F^{-1}(\Delta)), \quad (38)$$

where $\kappa = 2 + \log(2e+1) = 3.861\dots$.

The right-hand side tends to 0 as $\Delta \rightarrow 0$ only if the tail of $\mathbf{p}$ is not too heavy; namely, $F(x) = o(e^{-\kappa x})$.

The method of proof will be applied a couple of times in the sequel with different parameters. Therefore we formulate its essence in a separate lemma as a master inequality.

Lemma 10. *Let A , B , and C be positive real numbers; $C \leq B$. Suppose*

$$|p_k - q_k| \leq A \frac{(2B)^{2k}}{(2k)!}, \quad 0 \leq k < C. \quad (39)$$

Then

$$\|\mathbf{p} - \mathbf{q}\| \leq 2A \left(\frac{eB}{C} \right)^{2C} + 2F(C). \quad (40)$$

Proof. Let $n = \lceil C \rceil$, then $n-1 < C \leq n$. Clearly,

$$\begin{aligned} \|\mathbf{p} - \mathbf{q}\| &= \sum_{k=0}^{\infty} |p_k - q_k| \\ &= 2 \sum_{k=0}^{\infty} (p_k - q_k)^+ + \sum_{k=0}^{\infty} (p_k - q_k) \\ &= 2 \sum_{k=0}^{\infty} (p_k - q_k)^+ \leq 2 \sum_{k=0}^{n-1} |p_k - q_k| + 2 \sum_{k=n}^{\infty} p_k. \end{aligned} \quad (41)$$

By supposition we have

$$(2B)^{2k} = (2C)^{2k} \left(\frac{B}{C}\right)^{2k} \leq (2C)^{2k} \left(\frac{B}{C}\right)^{2C}. \quad (42)$$

Hence

$$\begin{aligned} \|\mathbf{p} - \mathbf{q}\| &\leq 2A \sum_{k=0}^{n-1} \frac{(2B)^k}{(2k)!} + 2F(n) \\ &\leq 2A \left(\frac{B}{C}\right)^{2C} \sum_{k=0}^{n-1} \frac{(2C)^{2k}}{(2k)!} + 2F(n) \\ &\leq 2A \left(\frac{eB}{C}\right)^{2C} + 2F(C). \end{aligned} \quad (43)$$

□

Proof of Theorem 9. Starting from Theorem 3, let us apply Lemma 10 with $A = 3\Delta$, $B = (2e + 1)^{1/2} F^{-1}(\Delta)$, and $C = F^{-1}(\Delta)$. We get that

$$\begin{aligned} \|\mathbf{p} - \mathbf{q}\| &\leq 6\Delta \left(e(2e + 1)^{1/2}\right)^{2F^{-1}(\Delta)} + 2\Delta \\ &\leq 8\Delta \exp(\kappa F^{-1}(\Delta)). \end{aligned} \quad (44)$$

□

Particularly, let $\mathbf{p}$ be the Poisson distribution with mean λ . As we have seen in (35),

$$F^{-1}(\Delta) = h^{-1}(L) = (1 + o(1)) \frac{L}{\log L}. \quad (45)$$

Let us plug this into Theorem 9. Writing 4 in place of κ in the exponent we can get rid of the term $1 + o(1)$, and even the multiplier 8 gets eventually absorbed. Thus we obtain the following estimate.

Corollary 11. *Let $\mathbf{p}$ be Poisson; then, uniformly in $\mathbf{q}$, one has*

$$\|\mathbf{p} - \mathbf{q}\| \leq \Delta \exp\left(\frac{4L}{\log L}\right), \quad (46)$$

if Δ is sufficiently small.

This is already similar to the lower bound (16), and is much better than (13).

If the tail of $\mathbf{p}$ is subexponential, that is, $\sum_{k \geq n} p_k = \exp(-o(n))$, then the estimate of Theorem 3 is useless: it tends to infinity as $\Delta \rightarrow 0$. However, with suitably chosen parameters in the master inequality, a reasonable upper bound can be obtained: not really sharp, but at least not trivial.

Theorem 12. *Suppose $F(n) = e^{-o(n)}$. Then*

$$\begin{aligned} \sup \left\{ \|\mathbf{p} - \mathbf{q}\| : \max_{0 \leq z \leq 1} |g_{\mathbf{p}}(z) - g_{\mathbf{q}}(z)| = \Delta \right\} \\ = O\left(F\left(\frac{L}{8}\right)\right), \quad \text{as } \Delta \rightarrow 0. \end{aligned} \quad (47)$$

Proof. This time we use Theorem 2. Let $A = e\Delta$, $B = eL/2$, and $C = L/8$. Then Lemma 10 gives that

$$\|\mathbf{p} - \mathbf{q}\| \leq 2e\Delta \left(4e^2\right)^{L/4} + 2F\left(\frac{L}{8}\right). \quad (48)$$

Since $(4e^2)^{1/4} = 2.331 \dots < e$, the first term in the right-hand side can be estimated by a positive power of Δ , while the second term decreases slower than any positive power of Δ , as $\Delta \rightarrow 0$; thus it will eventually dominate. □

Corollary 13. *If the α th moment of $\mathbf{p}$ is finite, then $\sum_{k=n}^{\infty} p_k = o(n^{-\alpha})$; thus*

$$\begin{aligned} \sup \left\{ \|\mathbf{p} - \mathbf{q}\| : \max_{0 \leq z \leq 1} |g_{\mathbf{p}}(z) - g_{\mathbf{q}}(z)| = \Delta \right\} &= o(L^{-\alpha}), \\ &\text{as } \Delta \rightarrow 0. \end{aligned} \quad (49)$$

Finally, let us deal with the case of exponentially decaying tails. Let $F(n) = e^{-vn}$, with $v \geq 0$. Then the total variation $\|\mathbf{p} - \mathbf{q}\|$ can be estimated by a positive power of Δ . This follows from Theorem 9, but only for $v > \kappa$. The following theorem is valid for all positive v .

Theorem 14. *Suppose $F(n) \leq e^{-vn}$, with $v > 0$. Then*

$$\|\mathbf{p} - \mathbf{q}\| \leq 8\Delta^{v/(v+7)}. \quad (50)$$

As not sharp, the concrete form of the exponent holds no interest; what really matters is that it is positive and tends to 1 as $v \rightarrow \infty$. Note that if p_k tends to 0 at an exponential rate, (14) provides a lower bound which is also a positive power of Δ (from Theorem 2.4 of [2] it follows that $C = (v+1/4)/(v+2)$ can be used for exponent).

Proof. Firstly, suppose $v \leq 1.9$. Based on Theorem 2, apply Lemma 10 with $A = e\Delta$, $B = eL/2$, and $C = L/(v+7)$:

$$\begin{aligned} \|\mathbf{p} - \mathbf{q}\| &\leq 2e\Delta \left(\frac{e^2(v+7)}{2}\right)^{2L/(v+7)} \\ &\quad + 2 \exp\left(-\frac{vL}{v+7}\right). \end{aligned} \quad (51)$$

Let us deal with the first term on the right-hand side

$$\begin{aligned} &\Delta \left(\frac{e^2(v+7)}{2}\right)^{2L/(v+7)} \\ &= \exp\left(-L \left[1 - \frac{2}{v+7} \left(2 + \log \frac{v+7}{2}\right)\right]\right) \\ &= \exp\left(-\frac{L}{v+7} \left(v+3 - 2 \log \frac{v+7}{2}\right)\right). \end{aligned} \quad (52)$$

Since $3 - 2 \log((v+7)/2) \geq 3 - 2 \log 4.45 > 0$, it follows that

$$\|\mathbf{p} - \mathbf{q}\| \leq 2(e+1) \Delta^{v/(v+7)}. \quad (53)$$

Secondly, let $\nu > 1.9$. Now we apply Theorem 3. Note that $F^{-1}(\Delta) = L/\nu$, and set $A = 3\Delta$, $B = (2e + 1)^{1/2}L/\nu$, and $C = L/(\nu + 7)$. Then we obtain that

$$\begin{aligned} \|\mathbf{p} - \mathbf{q}\| &\leq 6\Delta \left(\frac{e(2e + 1)^{1/2}(\nu + 7)}{\nu} \right)^{2L/(\nu+7)} \\ &\quad + 2 \exp \left(-\frac{\nu L}{\nu + 7} \right). \end{aligned} \quad (54)$$

Again, the first term can be estimated by the second one, because

$$\begin{aligned} \Delta \left(\frac{e(2e + 1)^{1/2}(\nu + 7)}{\nu} \right)^{2L/(\nu+7)} &= \exp \left(-L \left[1 - \frac{2}{\nu + 7} \left(\frac{\kappa}{2} + \log \frac{\nu + 7}{\nu} \right) \right] \right) \\ &= \exp \left(-\frac{L}{\nu + 7} \left(\nu + 7 - \kappa - 2 \log \frac{\nu + 7}{\nu} \right) \right), \end{aligned} \quad (55)$$

and $7 - \kappa - 2 \log((\nu + 7)/\nu) > 7 - 3.862 - 2 \log 4.685 > 0$. Thus, it follows that

$$\|\mathbf{p} - \mathbf{q}\| \leq 8\Delta^{\nu/(\nu+7)} \quad (56)$$

for $\nu > 1.9$, as needed. $\square$

Competing Interests

The author declares that there are no competing interests regarding the publication of this paper.

Acknowledgments

This research has been supported by the Hungarian Scientific Research Fund OTKA, Grant no. K 108615. The idea of the proof of Theorem 3 is due to Gábor Halász, whom the author is indebted to for his invaluable help.

References

- [1] A. Rényi, “A general method for proving theorems in probability theory and some of its applications,” in *Selected Papers of A. Rényi*, vol. 2, pp. 581–602, Akadémiai Kiadó, Budapest, Hungary, 1976.
- [2] T. F. Móri, “Deviation of discrete distributions—positive and negative results,” *Statistics & Probability Letters*, vol. 79, no. 8, pp. 1089–1096, 2009.
- [3] T. F. Móri, “Bonferroni inequalities and deviations of discrete distributions,” *Journal of Applied Probability*, vol. 33, no. 1, pp. 115–121, 1996.
- [4] W. Markoff and J. Grossmann, “Über Polynome, die in einem gegebenen Intervalle möglichst wenig von Null abweichen,” *Mathematische Annalen*, vol. 77, no. 2, pp. 213–258, 1916.
- [5] P. Borwein and T. Erdélyi, *Polynomials and Polynomial Inequalities*, vol. 161 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 1995.
- [6] L. Antal and T. F. Móri, “How large can the coefficients of a power series be?” *Annales Universitatis Scientiarum Budapestensis de Rolando Eötvös Nominatae. Sectio Computatorica*, vol. 34, pp. 25–32, 2011.
- [7] K. S. Berenhaut, J. V. Baxley, and R. G. Lyday, “Deviations of discrete distributions and a question of Móri,” *Statistics & Probability Letters*, vol. 81, no. 12, pp. 1940–1944, 2011.
- [8] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons, New York, NY, USA, 1989.