JEAN-MARC DESHOUILLERS
GREGORY A. FREIMAN
WILLIAM MORAN

On series of discrete random variables, 1 : real trinomial distributions with fixed probabilities

Astérisque, tome 258 (1999), p. 411-423
<http://www.numdam.org/item?id=AST_1999__258__411_0>

Astérisque **258**, 1999, p. 411–423

ON SERIES OF DISCRETE RANDOM VARIABLES, 1: REAL TRINOMIAL DISTRIBUTIONS WITH FIXED PROBABILITIES

by

Jean-Marc Deshouillers, Gregory A. Freiman & William Moran

Abstract. — This paper begins the study of the local limit behaviour of triangular arrays of independent random variables $(\zeta_{n,k})_{1\le k\le n}$ where the law of $\zeta_{n,k}$ depends on on n. We consider the case when $\zeta_{n,1}$ takes three integral values $0 < a_1(n) < a_2(n)$ with respective probabilities p_0, p_1, p_2 which do not depend on n. We show three types of limit behaviours for the sequence of r. v. $\eta_n = \zeta_{n,1} + \cdots + \zeta_{n,n}$, according as $a_2(n)/\gcd(a_1(n), a_2(n))$ tends to infinity slower, quicker or at the same speed as $\sqrt{n}$.

These notes are a first step in the description of the local behaviour of *series* of discrete random variables, that is to say sequences $(\eta_1, \dots, \eta_n, \dots)$ of random variables such that η_n is the sum of n independent discrete random variables $(\xi_{n,k})_{1\le k\le n}$ following a same law that may depend on n. We are restricting ourselves to the case when the $\xi_{n,k}$'s take three integer values $a_0 = 0 < a_1(n) < a_2(n)$, where a_1 and a_2 are coprime, with fixed positive probabilities p_0, p_1 and p_2 respectively.

When the values $a_1(n)$ and $a_2(n)$ do not depend on n, it follows from a result of Gnedenko that we have a *local limit* result, namely

$$P\{\eta_n = N\} = \frac{1}{\sigma\sqrt{2\pi n}}\left(\exp\left(-\frac{(n\mu - N)^2}{2n\sigma^2}\right) + o(1)\right) \qquad \text{as } n \to \infty,$$

uniformly in N, where μ and σ^2 are the expectation and the variance of the $\xi_{n,k}$'s.

Our aim is to give a complete description of the case when a_1 and a_2 depend on n, showing that there exist three different behaviors according as $a_2(n)$ is bounded or tends to infinity slower than $\sqrt{n}$, tends to infinity at the same speed as $\sqrt{n}$, or tends to infinity quicker than $\sqrt{n}$. In the first case, we get a *local limit* result similar to

1991 ***Mathematics Subject Classification.*** — 60F05, 60E10, 11P55, 11Z05.
Key words and phrases. — Sums of discrete random variables, local limit theorems, characteristic function, number theoretic methods, circle method.

J.-M. D: Cette recherche a bénéficié du soutien du CNRS (UMR 9936, Université Bordeaux 1) et de l'Université Victor Segalen Bordeaux 2.

the one we just quoted. The second case leads to a result of a similar structure with a limiting law which is no more normal. The third case can be seen as *isomorphic* to a two-dimensional series with a fixed law; this notion, which may be of future importance, will be presented in the last section. This notion will also be useful to explain what happens when the coprimality condition of the a_i's is removed, without having to rewrite the statement of the Theorem in a heavier form where a_1 and a_2 would be replaced by $a_1/\gcd(a_1, a_2)$ and $a_2/\gcd(a_1, a_2)$, and $\{\eta_n = N\}$ by $\{\eta_n = N\gcd(a_1, a_2)\}\ldots$

We now state our main result.

Theorem. — *Let p_0, p_1, p_2 be three positive numbers with sum 1, and $a_0 = 0 < a_1(n) < a_2(n)$ be three coprime integers. Let further*

$$\begin{aligned} \mu &= \mu_n = p_1 a_1(n) + p_2 a_2(n), \\ \sigma^2 &= \sigma_n^2 = p_1 a_1^2(n) + p_2 a_2^2(n) - \mu_n^2. \end{aligned}$$

We consider n independent random variables $(\xi_{n,k})_{1\le k\le n}$, each of which takes the values $a_0, a_1(n), a_2(n)$ with probabilities p_0, p_1, p_2 respectively, and we denote by η_n the sum $\xi_{n,1} + \cdots + \xi_{n,n}$.

When $a_2(n) = o(\sqrt{n})$ as n tends to infinity, we have, uniformly with respect to the integer N

$$P\{\eta_n = N\} = \frac{1}{\sigma_n\sqrt{2\pi n}}\left(\exp\left(-\frac{(n\mu_n - N)^2}{2n\sigma_n^2}\right) + o(1)\right) \quad \text{as } n \to \infty.$$

When $a_2(n)/\sqrt{n}$ tends to infinity with n, we have, uniformly with respect to the integer N

$$P\{\eta_n = N\} = \frac{n!}{k_0!k_1!k_2!}p_0^{k_0}p_1^{k_1}p_2^{k_2} + o(\frac{1}{n}) \quad \text{as } n \to \infty,$$

where the integral triple (k_0, k_1, k_2) is defined by

$$-a_2(n)/2 < k_1 - np_1 \le a_2(n)/2,\ k_1a_1(n) + k_2a_2(n) = N,\ k_0 + k_1 + k_2 = n.$$

When $a_2(n)/\sqrt{n}$ tends to a finite positive limit c when n tends to infinity, we have, uniformly with respect to c and to the integer N

$$P\{\eta_n = N\} = \frac{1}{2\pi n\sqrt{p_0p_1p_2}}\sum_{(k_0,k_1,k_2)}\exp(Q(k_0, k_1, k_2)) + o(\frac{1}{n}) \quad \text{as } n \to \infty,$$

where the sum is extended to integral triples satisfying $k_0 + k_1 + k_2 = n$, $a_1(n)k_1 + a_2(n)k_2 = N$, and the quadratic form Q is defined by

$$Q(k_0, k_1, k_2) = -\frac{1}{2}\sum_{i=1}^{3}\frac{1}{np_i}(k_i - np_i)^2.$$

1. The case when $a_2(n) = o(\sqrt{n})$

Due to the arithmetical nature of our problem (the a_i's are integers), we shall use the Fourier kernel $\exp(2\pi i \cdot x)$ and define by

$$\Psi(t) = \Psi_n(t) = p_0 + p_1 \exp(2\pi i t a_1) + p_2 \exp(2\pi i t a_2)$$

the characteristic function of $\xi_{n,k}$ so that we have

$$P\{\eta_n = N\} = \int \Psi^n(t) \exp(-2\pi i t N)\, dt\ ,$$

where the integral is performed over any interval of length 1.

We shall divide the range of summation into a *major arc*, when t is close to 0, and a *minor arc* when t is far from 0. Let ε be a positive real number (that will be specified later to be $1/(a_2(n)n^{2/5})$), and let

$$\mathfrak{M} = [-\varepsilon, \varepsilon] \qquad \text{and} \qquad \mathfrak{m} =]\varepsilon; 1-\varepsilon].$$

1.1. Contribution of the minor arc. — The following lemma will be used to get an upper bound for Ψ on the minor arc, playing either with the term $\exp(2\pi i a_1 t)$ or with $\exp(2\pi i a_2 t)$.

Lemma 1. — *Let $C = 8\min\left(\frac{p_0p_1}{p_0+p_1}, \frac{p_0p_2}{p_0+p_2}\right)$. We have*

$$|p_0 + p_1 \exp(2\pi i u_1) + p_2 \exp(2\pi i u_2)| \le \exp(-C\max(||u_1||^2, ||u_2||^2))\ ,$$

where $||u||$ denotes the distance from u to the nearest integer.

Proof. — It is of course enough to prove the inequality

$$|p_0 + p_1 \exp(2\pi i u_1) + p_2 \exp(2\pi i u_2)| \le \exp\left(\frac{-8p_0p_1}{p_0+p_1}||u_1||^2\right).$$

We have

$$\begin{aligned}
|p_0 + p_1 \exp(2\pi i u_1)|^2 &= p_0^2 + p_1^2 + 2p_0p_1 \cos 2\pi u_1 \\
&= (p_0+p_1)^2 - 4p_0p_1 \sin^2 \pi u_1 \\
&\le (p_0+p_1)^2 - 16p_0p_1||u_1||^2 \\
&\le (p_0+p_1)^2\Big(1 - \frac{8p_0p_1}{(p_0+p_1)^2}||u_1||^2\Big)^2\ .
\end{aligned}$$

This implies

$$\begin{aligned}
|p_0 + p_1 \exp(2\pi i u_1) + p_2(2\pi i u_2)| &\le p_0 + p_1 + p_2 - \frac{8p_0p_1}{p_0+p_1}||u_1||^2 \\
&\le 1 - \frac{8p_0p_1}{p_0+p_1}||u_1||^2 \\
&\le \exp\left(-\frac{8p_0p_1}{p_0+p_1}||u_1||^2\right)\ ,
\end{aligned}$$

which is the inequality we looked for. □

We now present the dissection of the minor arc. For any integer r, we shall denote by $\overline{r}$ the integer in $[0, a_2/2]$ such that r is congruent either to $\overline{r}$ or to $-\overline{r}$ modulo a_2. The reader will easily check that $\mathfrak{m}$ is the disjoint union of the following intervals:

$$\begin{aligned}
\mathfrak{m}_1(r) &= \left]\frac{r}{a_2} - \frac{\overline{ra_1}}{2a_2^2}, \frac{r}{a_2} + \frac{\overline{ra_1}}{2a_2^2}\right[, \quad \text{for } r = 1, 2, \ldots, a_1 - 1,\\
\mathfrak{m}_2(0) &= \left]\varepsilon, \frac{1}{a_2} - \frac{\overline{1a_1}}{2a_2^2}\right],\\
\mathfrak{m}_2(a_2 - 1) &= \left[\frac{a_2 - 1}{a_2} + \frac{\overline{(a_2 - 1)a_1}}{2a_2^2}, 1 - \varepsilon\right[\ ,\\
\mathfrak{m}_2(r) &= \left[\frac{r}{a_2} + \frac{\overline{ra_1}}{2a_2^2}, \frac{r+1}{a_2} - \frac{\overline{(r+1)a_2}}{2a_2^2}\right], \quad \text{for } r = 1, 2, \ldots, a_2 - 2\ .
\end{aligned}$$

The intervals of type $\mathfrak{m}_2$ stay away from rationals with denominator a_2, so that $||a_2 t||$ is rather large when t is in such an interval. More precisely, if we consider $\mathfrak{m}_2^-(r) = [\frac{r}{a_2} + \frac{\overline{ra_1}}{2a_2^2}, \frac{2r+1}{2a_2}]$, in the case when $1 \le r \le a_2/2$, we have $t \in \mathfrak{m}_2^-(r) \Rightarrow ||ta_2|| = ||ta_2 - r|| = (ta_2 - r)$, so that, by Lemma 1, we have

$$\begin{aligned}
\int_{\mathfrak{m}_2^-(r)} |\Psi(t)|^n dt &\le \int_{\mathfrak{m}_2^-(r)} \exp(-Cn(ta_2 - r)^2)dt\\
&\le \frac{1}{a_2\sqrt{n}} \int_{\frac{\overline{ra_1}}{2a_2}\sqrt{n}}^{\infty} \exp(-Cu^2)du\ .
\end{aligned}$$

In a similar way, the contribution of each of $]\varepsilon, \frac{1}{2a_2}]$ and $[\frac{2a_2-1}{a_2}, 1-\varepsilon[$ is at most

$$\frac{1}{a_2\sqrt{n}} \int_{\varepsilon a_2\sqrt{n}}^{\infty} \exp(-Cu^2)du.$$

The coprimality of a_1 and a_2 implies that $\overline{ra_1}$ is different from 0 for $r = 1, 2, \ldots, a_1 - 1$, and that any integer s in $[1, a_2/2]$ is equal to some $\overline{ra_1}$ for at most two values of r. If we denote by $\mathfrak{m}_2$ the union of the $\mathfrak{m}_2(r)$ for $r = 0, 1, \ldots, a_2 - 1$, we get

$$\int_{\mathfrak{m}_2} |\Psi(t)|^n dt \le \frac{2}{a_2\sqrt{n}} \int_{\varepsilon a_2\sqrt{n}}^{\infty} \exp(-Cu^2)du + \frac{4}{a_2\sqrt{n}} \sum_{s=1}^{\infty} \int_{\frac{s\sqrt{n}}{2a_2}}^{\infty} \exp(-Cu^2)du\ ,$$

so that the condition $a_2 = o(\sqrt{n})$ implies

$$\int_{\mathfrak{m}_2}^{\infty} |\Psi(t)|^n dt \le \frac{2}{a_2\sqrt{n}} \int_{\varepsilon a_2\sqrt{n}}^{\infty} \exp(-Cu^2)du + o(\frac{1}{a_2\sqrt{n}}). \tag{1}$$