

Math-Net.Ru

All Russian mathematical portal

M. L. Esquível, Probability generating functions for discrete real-valued random variables, *Teor. Veroyatnost. i Primenen.*, 2007, Volume 52, Issue 1, 129–149

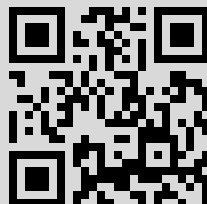
DOI: <https://doi.org/10.4213/tvp8>

Use of the all-Russian mathematical portal Math-Net.Ru implies that you have read and agreed to these terms of use
<http://www.mathnet.ru/eng/agreement>

Download details:

IP: 106.51.226.7

August 4, 2022, 21:05:10



© 2007 г.

ESQUÍVEL M. L.*

PROBABILITY GENERATING FUNCTIONS FOR
DISCRETE REAL-VALUED RANDOM VARIABLES¹⁾

*This work is dedicated to José de Sam Lazaro,
my friend and my teacher in General Mathematics,
Analytic Functions and Probability Theory,
as a token of respect and gratitude*

Вероятностные производящие функции являются мощным инструментом изучения распределений конечных сумм независимых дискретных случайных величин, принимающих целые положительные значения. Для вещественнозначных дискретных случайных величин хорошо известная элементарная теория рядов Дирихле и доступные сейчас программы символьных вычислений, такие как Mathematica 5TM, позволяют применить этот инструмент и к общим дискретным случайным величинам. Ввиду этого, настоящая статья имеет двоякую цель. Во-первых, мы показываем, что дискретные случайные величины, принимающие вещественные (не обязательно целые или рациональные) значения, можно изучать с помощью вероятностных производящих функций. Во-вторых, мы стремимся привлечь внимание к некоторым практическим путям выполнения необходимых вычислений.

Ключевые слова и фразы: вероятностные производящие функции, конечные суммы независимых вещественнозначных дискретных случайных величин, ряды Дирихле.

1. Classical probability generating functions

Generating functions are a useful and up-to-date tool in nowadays practical mathematics, in particular, in discrete mathematics and combinatorics (see [7]) and, in the case of probability generating functions, in distributional convergence results as in [6, p. 84]. Their uses in basic probability are demonstrated in the classic reference [2, Chap. XI]. More recently, probability generating functions for integer-valued random variables have been

* Departamento de Matemática, FCT/UNL, Quinta da Torre, 2829-516 Caparica, Portugal; e-mail: mle@fct.unl.pt; URL: <http://ferrari.dmat.fct.unl.pt/personal/mle>

¹⁾ Part of this work was done while visiting the Stockholm School of Economics; we hereby express our thanks to Professor Tomas Björk for the support and warm hospitality. This work was done with partial support from the FCT (Fundação para a Ciência e Tecnologia), program POCTI (Portugal/FEDER-EU).

studied intensively mainly with some applied purposes in mind. See, for instance, [1], [8]–[12], [15], and [16].

The natural extension of the definition of probability generating function to a nonnegative real-valued random variable X , as the expectation of the function t^X , is very clearly presented in the treatise [4, p. 288], where some of the consequences, drawn directly from this definition, are stated.

Let us briefly formulate some classical results on probability generating functions for integer-valued random variables recalling the useful connection between the topics of probability generating functions and of analytic function theory. Let X be a random variable taking values in \mathbf{Z} and let $p_k := \mathbf{P}\{X = k\} \in [0, 1]$ for all $k \in \mathbf{Z}$.

The probability generating function (PGF) of X , denoted by ψ_X , is given by $\psi_X(z) = \mathbf{E}[z^X] = \int_{\Omega} z^X d\mathbf{P}$ for all z in the set \mathbf{D}_X in which it is well defined, that is,

$$\mathbf{D}_X = \left\{ z \in \mathbf{C}: \int_{\Omega} z^X d\mathbf{P} \in \mathbf{C} \right\} = \left\{ z \in \mathbf{C}: \int_{\Omega} |z|^X d\mathbf{P} < +\infty \right\}.$$

Since the law of X is the probability measure μ_X given by $\mu_X = \sum_{n=-\infty}^{+\infty} p_n \delta_n$, where δ_a represents the Dirac measure with support in $\{a\}$, we can conclude that, by the standard result on the integration with respect to the law of X ,

$$\forall z \in \mathbf{D}_X \quad \psi_X(z) = \sum_{n=-\infty}^{n=+\infty} p_n z^n. \quad (1.1)$$

That means that the PGF of X is given by a Laurent series around zero in its domain of existence as a complex function. The domain of simple convergence of such a series is a set of the form $C(\rho_1, \rho_2) = \{0 \leq \rho_1 \leq |z| \leq \rho_2 \leq +\infty\}$, where by Hadamard's formula we have: $\rho_1 = \limsup_{n \rightarrow +\infty} \sqrt[n]{p_n}$ and $\rho_2 = 1 / \limsup_{n \rightarrow +\infty} \sqrt[n]{p_n}$. Since the series in (1.1) is absolutely (and uniformly) convergent in the closure of $C(r_1, r_2)$ for every $\rho_1 < r_1 < r_2 < \rho_2$ we have that $\mathbf{D}_X = C(\rho_1, \rho_2)$. If for all $n < 0$ we have that $p_n = 0$, then ψ_X is represented by a Taylor series around zero, $\{|z| < 1\} \subset \mathbf{D}_X$ and so,

$$\forall n \in \mathbf{N} \quad p_n = \frac{\psi_X^{(n)}(0)}{n!},$$

thus showing that ψ_X generates the probabilities in a very nice way that, in some cases, is useful in practice. In the general case, one can still, in a sense, generate the probabilities from the PGF since we have for some γ_r , the border of the circle of radius $r \in]\rho_1, \rho_2[$ centered at zero, that

$$\forall n \in \mathbf{Z} \quad p_n = \frac{1}{2\pi i} \int_{\gamma_r} \frac{\psi_X(\xi)}{\xi^{n+1}} d\xi.$$

The main purpose of this paper is to extend probability generating function techniques to discrete random variables taking real values nonnecessarily integer.

2. Probability generating functions for real-valued random variables

Consider, from now on, a discrete random variable X taking the sequence of real values $(\alpha_k)_{k \in \mathbf{Z}}$ such that for some sequence of probabilities $(p_k)_{k \in \mathbf{Z}} \in [0, 1]^{\mathbf{Z}}$, thus verifying $\sum_{k=-\infty}^{+\infty} p_k = 1$, we have $\mathbf{P}\{X = \alpha_k\} = p_k$ for all $k \in \mathbf{Z}$. The law of X is the probability measure μ_X which is given by $\mu_X = \sum_{k=-\infty}^{+\infty} p_k \delta_{\alpha_k}$. We will constantly use that for any $t > 0$ and $x \in \mathbf{R}$, $t^x := e^{x \ln t}$ is well defined. The formal definition follows naturally.

Definition 2.1. The *probability generating function* (PGF) of X is defined for all $t > 0$ by

$$\psi_X(t) = \mathbf{E}[t^X] = \int_{\mathbf{R}} t^x d\mu_X = \sum_{k=-\infty}^{+\infty} p_k t^{\alpha_k}.$$

Remark 2.1. Let us observe that $\mathbf{E}[t^X]$ is always well defined as a Lebesgue integral of a well defined positive function although, possibly, equal to $+\infty$, and that ψ_X takes at least a real value since

$$\psi_X(1) = \sum_{k=-\infty}^{+\infty} p_k = 1.$$

A natural question is then to determine the exact description of the convergence domain of ψ_X , that is, the set $\mathbf{D}_X := \{t > 0: \psi_X(t) < +\infty\}$, where the PGF of X is, in fact, a real-valued function. We will address this question in Theorem 2.1 below referring to Section 5 for some of the results on Dirichlet series that we use in the proof.

It is convenient to notice that PGF is one among other very important functional transforms, namely the characteristic function and the moment generating function. For future reference let us define precisely these notions. Let X be a real-valued random variable with law μ_X . Following [6, p. 84] we denote by $\widehat{\mu}_X$ the *characteristic function* of X , defined for all $t \in \mathbf{R}$ by

$$\widehat{\mu}_X(t) = \mathbf{E}[e^{itX}] = \int_{\mathbf{R}} e^{itx} d\mu_X(x).$$

For $s \in \mathbf{C}$ write $s = \sigma + it$. Another functional transformation of the law of a random variable gives us the moment generating function.

Definition 2.2. The *moment generating function* (MGF) $\widetilde{\mu}_X$ of a real-valued random variable X is defined to be

$$\widetilde{\mu}_X(z) = \mathbf{E}[e^{zX}] = \int_{\mathbf{R}} e^{zx} d\mu_X(x) \quad (2.1)$$

for all z in the set $\widetilde{\mathbf{D}}_X = \{z \in \mathbf{C}: \int_{\mathbf{R}} |e^{zx}| d\mu_X(x) < +\infty\}$, that is, such that the integral on the right exists.

R e m a r k 2.2. For any random variable X the natural domain of its MGF is never empty since $0 \in \tilde{\mathbf{D}}_X$. However, important properties depend crucially on \mathbf{D}_X having a nonempty interior. For that reason some authors (see [13, p. 294]) assume that $\tilde{\mu}_X$ is defined only in that case. In subsection 5.2 we will deal more thoroughly with this question.

There are natural relations between these functional transforms. For all t for which the functional transforms involved are well defined we have

$$\tilde{\mu}_X(it) = \int_{\mathbf{R}} e^{itz} d\mu(x) = \hat{\mu}(t) \quad \text{and} \quad \tilde{\mu}_X(\ln t) = \psi_X(t). \quad (2.2)$$

Consider the following further convention on the notation used above, namely, X is a random variable taking as values the ordered sequence of real values $(\alpha_k)_{k \in \mathbf{Z}}$ each one with the corresponding probability p_k and suppose that for $k < 0$ we have $\alpha_k < 0$, $\alpha_0 = 0$ and for $k > 0$ we have $\alpha_k > 0$.

Theorem 2.1. *Let X be a random variable and let ψ_X denote its PGF. Then*

1) *If X takes a finite number of real values, then*

$$\mathbf{D}_X =]0, +\infty[;$$

2) *if X takes an infinite number of real values without accumulation points, then*

$$\text{there exist } u_0, v_0 \in]-\infty, 0] \text{ such that }]e^{u_0}, e^{-v_0}[\subset \mathbf{D}_X \subset [e^{u_0}, e^{-v_0}]; \quad (2.3)$$

3) *if X is a discrete random variable with exponentially decaying tails, that is, if for some $k, c > 0$ we have that $\mathbf{P}\{|X| > x\} \leq ke^{-cx}$ then we get also the condition expressed by (2.3).*

P r o o f. In the first case the PGF takes the form

$$\psi_X(t) = \sum_{k=-M}^{+N} p_k t^{\alpha_k} = \frac{p_{-M}}{t^{\alpha_{-M}}} + \dots + \frac{p_{-1}}{t^{\alpha_{-1}}} + p_0 + p_1 t^{\alpha_1} + \dots + p_N t^{\alpha_N},$$

for some integers M and N and the result announced is obviously true. For the second result, defining

$$\begin{cases} q_k = p_{-k}, \\ \beta_k = -\alpha_{-k}, \end{cases}$$

we have

$$\psi_X(t) = \sum_{k=1}^{+\infty} q_k e^{-\beta_k \ln t} + p_0 + \sum_{k=1}^{+\infty} p_k e^{-\alpha_k \ln(1/t)}, \quad (2.4)$$

where the sequences $(\alpha_k)_{k \in \mathbf{N}^*}$ and $(\beta_k)_{k \in \mathbf{N}^*}$ are increasing with the notation $\mathbf{N}^* = \mathbf{N} \setminus \{0\}$. Under the hypotheses that these sequences do not have a limit in \mathbf{R} , in relation (2.4) above we have expressed ψ_X as a sum of a constant p_0

and two Dirichlet series taken at $\ln t$ and at $\ln(1/t)$. The description of the set of convergence of ψ_X can then be tied up with the description of the convergence sets of two Dirichlet series, one for the positive values, other for the negative values of the random variable. Consider the Dirichlet series defined by

$$\sum_{k=1}^{+\infty} q_k e^{-\beta_k s} \quad \text{and} \quad \sum_{k=1}^{+\infty} p_k e^{-\alpha_k s}$$

and apply the results on absolutely convergent series, recalled in subsection 5.1, to obtain u_0 and v_0 , the abscissas of absolute convergence of the series on the left and on the right, respectively. Since $\sum_{k=-\infty}^{+\infty} p_k = 1$ we have $v_0, u_0 \leq 0$. This now obviously implies the result because

$$\ln t > u_0 \Leftrightarrow t > e^{u_0} \quad \text{and} \quad \ln \frac{1}{t} > v_0 \Leftrightarrow t < e^{-v_0}.$$

The last result stated in the theorem is an immediate consequence of Proposition 5.2 in the last section since $\mathbf{D}_X = \{|e^z|: z \in \tilde{\mathbf{D}}_X\}$. Theorem 2.1 is proved.

R e m a r k 2.3. The use of Dirichlet series or MGF is further justified since it allows one to determine the parameters u_0, v_0 in the theorem above. In fact, if X takes an infinite number of values without accumulation points, then u_0, v_0 are given by (5.2) or (5.3). If X is discrete and have an exponential decaying tail, then we may get σ_0^+ and σ_0^- as defined in the proof of Proposition 5.2 and then it is clear that $u_0 = \sigma_0^-$ and $v_0 = -\sigma_0^+$.

2.1. Generating the probabilities. One reason for the denomination used in the definition is the following result. We will recall first some notation. Let $t > 0$ and define, for such t , the *floor function* as

$$\lfloor t \rfloor = \sup\{n \in \mathbf{N}: n \leq t\},$$

that is, the greatest integer less or equal to t and the *fractional part* of t as

$$\text{frac}(t) = t - \lfloor t \rfloor.$$

Also, as a notation, let us say that $\prod_{l=0}^{-1} a_l = 1$.

Theorem 2.2. *Let X be a random variable taking only a finite number of values $(\alpha_k)_{-M \leq k \leq N}$, suppose that the values in this sequence are ordered as an increasing sequence, and let*

$$\psi_X(t) = \sum_{k=-M}^{+N} p_k t^{\alpha_k} = \frac{p_{-M}}{t^{\alpha_{-M}}} + \dots + \frac{p_{-1}}{t^{-\alpha_1}} + p_0 + p_1 t^{\alpha_1} + \dots + p_N t^{\alpha_N} \quad (2.5)$$

denote the PGF of the random variable X . Then, obviously,

$$p_{-M} = \mathbf{P}\{X = \alpha_{-M}\} = \lim_{t \rightarrow 0^+} (t^{\alpha_{-M}} \psi_X(t))$$

and derivating enough times, we have

$$p_{-M+1} = \lim_{t \rightarrow 0^+} \frac{1}{t^{\text{frac}(\alpha_{-M} - \alpha_{-M+1})}} \frac{(t^{\alpha_{-M}} \psi_X(t))^{\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor}}{\prod_{l=0}^{\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor - 1} (\alpha_{-M} - \alpha_{-M+1} - l)}.$$

By induction we can get the formulas for the remaining values of X .

P r o o f. For the first result it is enough to observe that

$$\begin{aligned} t^{\alpha_{-M}} \psi_X(t) &= p_{-M} + p_{-M-1} t^{\alpha_{-M} - \alpha_{-M-1}} + \dots + p_0 t^{\alpha_{-M}} \\ &\quad + p_1 t^{\alpha_1 + \alpha_{-M}} + \dots \end{aligned} \quad (2.6)$$

For the second result in the statement of the theorem, in case $\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor \geq 1$, just derivate (2.6) $\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor$ times. Theorem 2.2 is proved.

R e m a r k 2.4. The practical interest of this theorem is reduced by the fact that with the software allowing symbolic calculus it is easy to extract the coefficient for a given exponent of t . We will show this in a couple of examples in what follows (see Section 3).

2.2. Fundamental properties of PGF. The next result shows that the PGF, whenever properly defined, characterizes the MGF of a random variable and consequently characterizes the distribution of this random variable.

Theorem 2.3. *Let X, Y be random variables such that for some neighborhood V of $1 \in \mathbf{D}_X \cap \mathbf{D}_Y$ their PGFs ψ_X and ψ_Y are well defined and verify*

$$\forall t \in V \quad \psi_X(t) = \psi_Y(t). \quad (2.7)$$

Then $\tilde{\mu}_X \equiv \tilde{\mu}_Y$ and consequently $X \stackrel{d}{=} Y$.

P r o o f. Condition (2.7) implies that $\tilde{\mu}_X$ and $\tilde{\mu}_Y$ are well defined on

$$\mathbf{D} = \left\{ s = \sigma + it \in \mathbf{C} : \sigma \in \{ \ln u : u \in \text{Int}(V) \} \right\} \subset \mathbf{D}_X \cap \mathbf{D}_Y$$

and that for $s \in \mathbf{D}$ we have $\tilde{\mu}_X(s) = \tilde{\mu}_Y(s)$. Since, by Proposition 5.3, $\tilde{\mu}_X$ and $\tilde{\mu}_Y$ are holomorphic functions on \mathbf{D}_X and \mathbf{D}_Y , respectively, they are certainly equal as two holomorphic functions coinciding in a set having an accumulation point. In order to conclude it suffices to observe that since $\tilde{\mu}_X \equiv \tilde{\mu}_Y$ we have

$$\forall t \in \mathbf{R} \quad \hat{\mu}_X(t) = \tilde{\mu}_X(0 + it) = \tilde{\mu}_Y(0 + it) = \hat{\mu}_Y(t).$$

That is, the characteristic functions of X and Y coincide. Being so it is well known that $X \stackrel{d}{=} Y$, as wanted. Theorem 2.3 is proved.

The next result shows that given a sequence of random variables $(X_n)_{n \in \mathbf{N}}$, the convergence of the corresponding sequence of PGFs to a PGF of a random variable X on a set having a nonempty interior is enough to ensure that the sequence $(X_n)_{n \in \mathbf{N}}$ converges in distribution to X .

Theorem 2.4. *Let X, X_1, X_2, \dots be real-valued random variables and $\psi, \psi_1, \psi_2, \dots$ the corresponding PGFs. Suppose that for V , some neighborhood of $1 \in \mathbf{D}_X$, we have*

$$\forall n \in \mathbf{N} \quad V \subset \mathbf{D}_{X_n} \tag{2.8}$$

and

$$\forall t \in V \quad \lim_{n \rightarrow +\infty} \psi_n(t) = \psi(t). \tag{2.9}$$

Then $(X_n)_{n \in \mathbf{N}}$ converges in distribution to X .

P r o o f. Let $\mathbf{D} := \{s = \sigma + it \in \mathbf{C} : e^\sigma \in V\}$. Then condition (2.9) implies that

$$\forall s = \sigma + it \in \mathbf{D} \quad \lim_{n \rightarrow +\infty} \tilde{\mu}_{X_n}(\sigma) = \lim_{n \rightarrow +\infty} \mu_{X_n}(e^\sigma) = \mu_X(e^\sigma) = \tilde{\mu}_X(\sigma).$$

Consider now $s = \sigma + it \in \mathbf{D}$ arbitrary. Considering the complex measure $\mu_{X_n} - \mu_X$ and its total variation $|\mu_{X_n} - \mu_X|$ it is clear that

$$|\tilde{\mu}_{X_n}(s) - \tilde{\mu}_X(s)| \leq \int_{\mathbf{R}} |e^{sx}| d|\mu_{X_n} - \mu_X|(x) = |\mu_{X_n} - \mu_X|(e^{\sigma x}) < +\infty.$$

Moreover, we have $\lim_{n \rightarrow +\infty} |\mu_{X_n} - \mu_X|(e^{\sigma x}) = 0$, that is, $(\tilde{\mu}_{X_n})_{n \in \mathbf{N}}$ converges over \mathbf{D} towards $\tilde{\mu}_X$. Being so, by Theorem 5.2 we have that $(X_n)_{n \in \mathbf{N}}$ converges in law to X . Theorem 2.4 is proved.

2.3. The PGF of a sum of independent and identically distributed random variables. One important usage of PGFs is the determination of the law of a sum of independent random variables when the laws of the terms are known. Examples of this usage will be given in Section 3. As in the case of nonnegative integer-valued random variables, the following simple result on the PGF of a sum of independent random variables is obtained as a consequence of elementary facts from probability theory.

Theorem 2.5. *Let X and Y be two independent discrete real-valued random variables. Then*

$$\psi_{X+Y}(t) = \psi_X(t) \times \psi_Y(t)$$

and if $\mu_X = \sum_{k=-\infty}^{+\infty} p_k \delta_{\alpha_k}$ and $\mu_Y = \sum_{l=-\infty}^{+\infty} q_l \delta_{\beta_l}$, and $\mathbf{D}_{X+Y} = \mathbf{D}_X \cap \mathbf{D}_Y$, then

$$\forall t \in \mathbf{D}_{X+Y} \quad \psi_{X+Y}(t) = \sum_{k,l=-\infty}^{+\infty} p_k q_l t^{\alpha_k + \beta_l}.$$

P r o o f. The first equality is a consequence of the independence of X and Y . The second equality is simply given by the usual product of two absolutely convergent series. Theorem 2.5 is proved.

Using this result it is now possible to obtain, in a very simple way, the PGF of a finite sum of independent and identically distributed random variables.

Corollary 2.1. Let X_1, X_2, \dots, X_m be a sequence of random variables which are independent and identically distributed with a discrete real-valued random variable X . Then for all $t > 0$

$$\psi_{X_1+\dots+X_m}(t) = (\psi_X(t))^m,$$

and if $\mu_X = \sum_{k=-\infty}^{+\infty} p_k \delta_{\alpha_k}$, then for every $t \in \mathbf{D}_X$

$$\psi_{X_1+\dots+X_m}(t) = \sum_{i_1, \dots, i_m = -\infty}^{+\infty} p_{i_1} \cdots p_{i_m} t^{\alpha_{i_1} + \dots + \alpha_{i_m}}.$$

Proof. The first equality is a consequence of the theorem and the second one is a consequence of the product formula for absolutely convergent series.

3. Two calculation examples

The next examples show how to take advantage of the symbolic calculation capabilities of usual software in order to obtain the distribution function of a sum of a finite number of independent copies of a random variable taking a finite number of real values. In the first example the random variable takes rational positive and negative values. In the second example the random variable takes irrational values.

The discrete random variable taking rational values X_1 defined below appears naturally in the context of fair marking multiple choice questions:

$$X_1 = \begin{cases} -1 & \text{with probability } \frac{1}{16}, \\ -\frac{2}{3} & \text{with probability } \frac{3}{16}, \\ -\frac{1}{3} & \text{with probability } \frac{3}{16}, \\ 0 & \text{with probability } \frac{1}{8}, \\ \frac{1}{3} & \text{with probability } \frac{3}{16}, \\ \frac{2}{3} & \text{with probability } \frac{3}{16}, \\ 1 & \text{with probability } \frac{1}{16}. \end{cases} \quad (3.1)$$

The PGF of X_1 is given by

$$\psi_{X_1}(t) = \frac{1}{16} t^{-1} + \frac{3}{16} t^{-2/3} + \frac{3}{16} t^{-1/3} + \frac{2}{16} + \frac{3}{16} t^{1/3} + \frac{3}{16} t^{2/3} + \frac{1}{16} t.$$

For a deeper understanding of fair marking an exam with a set of, say, ten multiple choice questions it is important to know the distribution of the

sum of ten independent copies of X_1 which we denote by Y . We know that $\psi_Y = (\psi_{X_1}(t))^{10}$. With a symbolic computation package we have expanded this power in a sum of 61 terms of the form at^α and extracted each term of the sum. From each one of these terms we extracted the coefficient a and the power α of t thus obtaining the probabilities and the corresponding values taken by Y .

In order to fully demonstrate the usefulness of our approach we present next the commented lines of a very crude program for Mathematica 5 TM used to produce the probability distribution of Y , and the corresponding graphic representation.

1. This first command defines ψ_{X_1} as a function of the variable t :

```
PGF[t_] := ((1/16)*t^(-1)) + ((3/16)*t^(-2/3))
+ ((3/16)*t^(-1/3)) + ((2/16)*t^(0)) + ((3/16)*t^(1/3))
+ ((3/16)*t^(2/3)) + ((1/16)*t^(1))
```

2. Here the full expansion of $(\psi_{X_1})^{10}$ as a sum of terms of the form at^α is defined as a function of the variable t :

```
GPGF10[t_] := Expand[PGF[t]^10]
```

3. The following command just counts the number of terms of the form at^α in the expansion:

```
numTer=Count[GPGF10[t],_*t_^_]
```

4. A list, which is function of the variable t , is made of the terms of the expansion:

```
TabProb[t_] = Table[Part[GPGF10[t], k], {k, 1, numTer}]
```

5. This command calculates $\alpha at^{\alpha-1}/a$ the derivative of the term at^α divided by a in order to get the exponent α :

```
derTabProb[n_Integer, t_]
:= D[TabProb[t][[n]], t]/TabProb[1][[n]]
```

6. The exponent of a term at^α is just the value of $\alpha at^{\alpha-1}/a$ when $t = 1$:

```
Exponents[n_Integer] := derTabProb[n, t] /. t -> 1
```

7. The list ProbExpon is just the probability distribution of $(\psi_{X_1})^{10}$ given by the pairs having as a first element the value taken by the random variable and, as a second term, the corresponding probability:

```
ProbExpon = Table[Exponents[k], TabProb[1][[k]],
{ k, 1, numTer}]
```

8. The probability distribution is sorted with the lexicographic order so that smaller values of the random variable come first:

```
SortProbExpon = Sort[ProbExpon]
```

9. This last command draws the graphic:

```
DistFunc = ListPlot[SortProbExpon,
PlotStyle -> PointSize[.012]]
```

The graphic representation of the probability distribution of Y is given in Fig. 1. Note that a first inspection of this figure suggests the use of a normal approximation for Y .

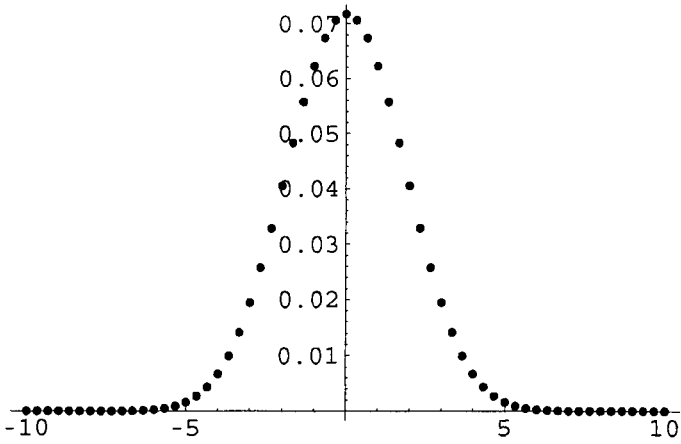


Fig. 1

For a second example consider a random variable X_2 taking some irrational values defined as follows:

$$X_2 = \begin{cases} -\frac{3}{4} & \text{with probability } 0.34, \\ \pi & \text{with probability } 0.33, \\ 2\pi & \text{with probability } 0.33. \end{cases} \quad (3.2)$$

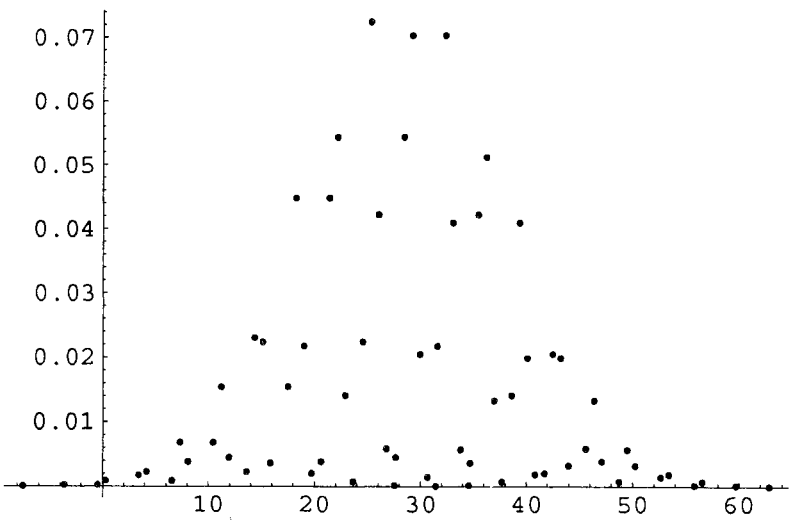


Fig. 2

Obviously the PGF of X_2 is given by

$$\psi_{X_2}(t) = \frac{0.34}{t^{3/4}} + 0.33 t^\pi + 0.33 t^{2\pi}.$$

As above, we are interested in the law of Z , the sum of ten independent and identically distributed copies of X_2 . We know that

$$\psi_Z(t) = \left(\frac{0.34}{t^{3/4}} + 0.33 t^\pi + 0.33 t^{2\pi} \right)^{10}.$$

Proceeding as in the first example above, we get probability distribution of Z given Fig. 2. Obviously, using a normal approximation for Z cannot be thought in this case.

4. Random variables taking an infinite number of values

In the preceding section we showed how, using the PGF, to determine effectively the probability distribution of a finite sum of independent identically distributed random variables taking a finite number of real values. In this section we will show that for a sum of independent and identically distributed random variables taking an infinite number of real values with no accumulation points, the same procedure can be used up to an approximation error, under some mild restrictive hypotheses.

Let X be a discrete real-valued random variable taking an infinite number of values. The method we propose is as follows. First we define a sequence of real-valued random variables $(X^{M,N})_{M,N \in \mathbf{N}}$, taking a finite number of values. It is then easy to see that the sequence $(X^{M,N})_{M,N \in \mathbf{N}}$ converges in law to X . We may then use a sum of independent copies of $X^{M,N}$, for M and N large enough, to approach the sum of independent copies of X .

As in the preceding section, let $(\alpha_k)_{k \in \mathbf{N}}$ denote the ordered sequence of real numbers which are the values taken by X with the corresponding probabilities $(p_k)_{k \in \mathbf{N}}$. Suppose that $\alpha_0 = 0$ and that for $k < 0$ we have $\alpha_k < 0$ and for $k > 0$ we have $\alpha_k > 0$. Consider for each $M, N \in \mathbf{N}$ the random variable $X^{M,N}$ such that

$$\begin{cases} \tilde{p}_k := \mathbf{P}\{X^{M,N} = \alpha_k\} = \mathbf{P}\{X = \alpha_k\} = p_k \\ \quad \forall k \in \{-M, \dots, -1, 1, \dots, N\}, \\ \tilde{p}_0 := \mathbf{P}\{X^{M,N} = 0\} = p_0 + \sum_{k=-\infty}^{-M-1} p_k + \sum_{k=N+1}^{+\infty} p_k. \end{cases} \quad (4.1)$$

This random variable $X^{M,N}$ takes the values $\{\alpha_{-M}, \dots, \alpha_{-1}, \alpha_1, \dots, \alpha_{N-1}\}$ with same probabilities as X and takes the value $\alpha_0 = 0$ with a probability equal to the sum of p_0 plus the sum of remaining probabilities for the other negative values taken by X , plus the sum of remaining probabilities for the other positive values taken by X .

Theorem 4.1. *The sequence of random variables $(X^{M,N})_{M,N \in \mathbf{N}}$ converges in law to X .*

P r o o f. Let f be a continuous bounded function on \mathbf{R} . Since $\mu_{X^{M,N}}$ and μ_X are given, respectively, by

$$\mu_{X^{M,N}} = \sum_{k=-M}^N p_k \delta_{\alpha_k} + \left(\sum_{k=-\infty}^{-M-1} p_k + \sum_{k=N+1}^{+\infty} p_k \right) \delta_0 \quad \text{and} \quad \mu_X = \sum_{k=-\infty}^{+\infty} p_k \delta_{\alpha_k},$$

we then have that

$$\begin{aligned} & |\mu_X(f) - \mu_{X^{M,N}}(f)| \\ &= \left| \sum_{k=-\infty}^{-M-1} p_k f(\alpha_k) - \left(\sum_{k=-\infty}^{-M-1} p_k \right) f(0) \right. \\ &\quad \left. + \sum_{k=N+1}^{+\infty} p_k f(\alpha_k) - \left(\sum_{k=N+1}^{+\infty} p_k \right) f(0) \right| \\ &= \left| \sum_{k=-\infty}^{-M-1} p_k (f(\alpha_k) - f(0)) + \sum_{k=N+1}^{+\infty} p_k (f(\alpha_k) - f(0)) \right|. \end{aligned}$$

Let K denote the bound of $|f|$. We then will have

$$|\mu_X(f) - \mu_{X^{M,N}}(f)| \leq 2K \left(\sum_{k=-\infty}^{-M-1} p_k + \sum_{k=N+1}^{+\infty} p_k \right).$$

As we have $\sum_{k=-\infty}^{+\infty} p_k = 1$, the theorem is proved.

We may now proceed to the second step of our approximation procedure.

Theorem 4.2. *Let $m \geq 1$ be a fixed integer, let X_1, \dots, X_m be m independent copies of X , and let $X_1^{M,N}, \dots, X_m^{M,N}$ be m independent copies of $X^{M,N}$. Then the sequence of random variables $(X_1^{M,N} + \dots + X_m^{M,N})_{M,N \in \mathbf{N}}$ converges in law to $X_1 + \dots + X_m$.*

P r o o f. It is a simple consequence of the continuity theorem of Lévy-Cramér.

We now show that the sequence of the PGFs of the random variables $X^{M,N}$ converges uniformly to the PGF of X .

Theorem 4.3. *Let u_0 and v_0 be as in Theorem 2.1. The sequence $(\psi_{X^{M,N}})_{M,N \in \mathbf{N}}$ converges uniformly for ψ_X on $]e^{u_0}, e^{-v_0}[$.*

P r o o f. By the definitions we simply have to observe that

$$\begin{aligned} |\psi_X(t) - \psi_{X^{M,N}}(t)| &\leq \left| \sum_{k=-\infty}^{-M-1} p_k t^{\alpha_k} \right| + \left(\sum_{k=-\infty}^{-M-1} p_k + \sum_{k=N+1}^{+\infty} p_k \right) \\ &\quad + \left| \sum_{k=N+1}^{+\infty} p_k t^{\alpha_k} \right| \end{aligned}$$

and use the fact that on $]e^{u_0}, e^{-v_0}[$ the terms $\sum_{k=-\infty}^{-M-1} p_k t^{\alpha_k}$ and $\sum_{k=N+1}^{+\infty} p_k t^{\alpha_k}$ are the remaining terms of two uniformly convergent Dirichlet series.

In order to finish we will show that the sequence of PGF of $X_1^{M,N} + \dots + X_m^{M,N}$ converges uniformly to the PGFs of $X_1 + \dots + X_m$.

Theorem 4.4. *With the same notation as used in the statement and in the proof of Theorem 2.1 suppose that $u_0, v_0 < 0$. Then, for every $m \geq 1$ and every $\varepsilon > 0$, there exist $M_0, N_0 \in \mathbf{N}$, and $u, v \in \mathbf{R}$ verifying*

$$1 \in]e^u, e^{-v}[\subseteq]e^{u_0}, e^{-v_0}[,$$

such that for all $M \geq M_0$ and $N \geq N_0$,

$$\forall t \in \mathbf{D}_X \quad (\psi_X(t))^m = (\psi_{X^{M,N}}(t))^m + R_\varepsilon^{M,N}(t),$$

where

$$\forall M \geq M_0, N \geq N_0, \text{ and } t \in]e^u, e^{-v}[, \quad |R_\varepsilon^{M,N}(t)| \leq \varepsilon(1 + \varepsilon)^{m-1}m. \quad (4.2)$$

Proof. As a consequence of $u_0, v_0 < 0$ we have $1 \in]e^{u_0}, e^{-v_0}[$. By the results on Dirichlet series, the series defining ψ_X converges uniformly on $]e^{u_0}, e^{-v_0}[$ and so

$$\lim_{M,N \rightarrow +\infty} \sup_{t \in]e^{u_0}, e^{-v_0}[} |\psi_X(t) - \psi_{X^{M,N}}(t)| = 0.$$

It is then possible to choose $M_0, N_0 \in \mathbf{N}$ such that for all $M \geq M_0$ and $N \geq N_0$,

$$\forall t \in]e^{u_0}, e^{-v_0}[\quad |\psi_X(t) - \psi_{X^{M,N}}(t)| \leq \frac{\varepsilon}{2}. \quad (4.3)$$

Since $\psi_{X^{M,N}}$ is continuous for $t > 0$ and the sequence $(\psi_{X^{M,N}})_{M,N \in \mathbf{N}}$ converges uniformly to ψ_X on $]e^{u_0}, e^{-v_0}[$, the PGF ψ_X is continuous at the point $t = 1$, where $\psi_X(1) = 1$. We may then choose u, v such that $1 \in]e^u, e^{-v}[\subseteq]e^{u_0}, e^{-v_0}[$ and

$$\forall t \in]e^u, e^{-v}[\quad |\psi_X(t)| \leq 1 + \frac{\varepsilon}{2}. \quad (4.4)$$

As a consequence of estimate (4.3) we then have

$$\forall t \in]e^u, e^{-v}[\quad |\psi_{X^{M,N}}(t)| \leq 1 + \varepsilon. \quad (4.5)$$

A very well-known formula tells us that for all $t \in \mathbf{D}_X$

$$(\psi_X(t))^m = (\psi_{X^{M,N}}(t))^m + (\psi_X(t) - \psi_{X^{M,N}}(t)) \sum_{i=0}^{m-1} (\psi_X(t))^{m-1-i} (\psi_{X^{M,N}}(t))^i.$$

Defining

$$R_\varepsilon^{M,N}(t) = (\psi_X(t) - \psi_{X^{M,N}}(t)) \sum_{i=0}^{m-1} (\psi_X(t))^{m-1-i} (\psi_{X^{M,N}}(t))^i,$$

it follows from the previous estimates (4.4) and (4.5) that estimate (4.2) in the statement of the theorem is valid, thus finishing the proof of this theorem.

5. Auxiliary results

For the reader's convenience we present in this section some technical results on Dirichlet series and moment generating functions that were essential to prove some fundamental results on PGFs. We suppose that the results of subsection 5.2 may have interest on its own.

5.1. A quick review of Dirichlet series. In this subsection we recall from [3] or [18] some results that were needed in previous sections. A Dirichlet series is a series of the form

$$\sum_{n=1}^{+\infty} a_n e^{-\lambda_n s}, \quad (5.1)$$

where $(a_n)_{n \in \mathbf{N}^*}$ is a sequence of complex numbers and $(\lambda_n)_{n \in \mathbf{N}^*}$ is an unbounded increasing sequence of positive real numbers. Let us observe first that if the series in (5.1) converges absolutely for $s_0 = \sigma_0 + it_0$, then the series converges absolutely and uniformly for every $s = \sigma + it$ such that $\sigma \geq \sigma_0$, as a consequence of Weierstrass criteria. This result implies the existence of α , named the *abscissa of absolute convergence*, such that for $s = \sigma + it$ such that $\sigma > \alpha$ the series converges absolutely and if $\sigma < \alpha$, then the series does not converge absolutely. On the line $\sigma = \alpha$ more analysis is needed to decide on the absolute convergence of the series.

In what concerns simple convergence, and as an easy consequence of Abel's lemma on series summation, we get that if the series in (5.1) converges for $s_0 = \sigma_0 + it_0$, then

- (a) The series converges for every $s = \sigma + it$ such that $\sigma > \sigma_0$;
- (b) the series converges uniformly in

$$\left\{ s \in \mathbf{C}: |\text{Arg}(s - s_0)| \leq a < \frac{\pi}{2} \right\}.$$

Once again, this result implies the existence of β , named the *abscissa of convergence*, such that for $s = \sigma + it$ such that $\sigma > \beta$ the series converges and if $\sigma < \beta$, then the series diverges. Also in this case, on the line $\sigma = \beta$ more analysis is needed to decide the simple convergence of the series.

Moreover, another application of the same lemma shows that if $\beta > 0$ or if $\beta = 0$ but $\sum_{n=1}^{+\infty} a_n \neq 0$, then

$$\beta = \limsup_{n \rightarrow +\infty} \frac{|\sum_{k=1}^n a_k|}{\lambda_n}.$$

It can also be shown that if $\beta < 0$, then

$$\beta = \limsup_{n \rightarrow +\infty} \frac{|\sum_{k=n+1}^{+\infty} a_k|}{\lambda_{n+1}}.$$

As a consequence we also have that if $\alpha > 0$ or if $\alpha = 0$ but $\sum_{n=1}^{+\infty} |a_n| \neq 0$, then

$$\alpha = \limsup_{n \rightarrow +\infty} \frac{\sum_{k=1}^n |a_k|}{\lambda_n} \tag{5.2}$$

and if $\alpha < 0$, then

$$\alpha = \limsup_{n \rightarrow +\infty} \frac{\sum_{k=n+1}^{+\infty} |a_k|}{\lambda_{n+1}}. \tag{5.3}$$

5.2. On the moment generating function Recall the definition of the moment generating function and of its natural domain of existence given in Definition 2.2. It is easy but somehow lengthy to show that $\tilde{\mathbf{D}}_X$ having a nonempty interior happens only for random variables with exponential decaying tails.

Proposition 5.1. *Let X be a real-valued random variable. Then*

$$\text{Int}(\tilde{\mathbf{D}}_X) \neq \emptyset \iff \exists k, c > 0 \quad \mathbf{P}\{|X| > x\} \leq ke^{-cx}. \tag{5.4}$$

P r o o f. Suppose that the natural domain of definition of MGF has nonempty interior. Let us deal with $\int_{\mathbf{R}_+} e^{\sigma x} d\mu_X(x)$ first. Suppose that for $\sigma^+ > 0$ we have $\int_{\mathbf{R}_+} e^{\sigma^+ x} d\mu_X(x) < +\infty$. Let X^+ and X^- denote, respectively, the positive and the negative parts of X . We then have

$$\int_{\Omega} e^{\sigma^+ X^+} d\mathbf{P} = \int_{\mathbf{R}_+} e^{\sigma^+ x} d\mu_X(x) + \mu_X(\{0\}) < +\infty. \tag{5.5}$$

By Chebyshev’s inequality

$$\mathbf{P}\{e^{\sigma^+ X^+} > u\} \leq \frac{1}{u} \int_{\Omega} e^{\sigma^+ X^+} d\mathbf{P},$$

which is equivalent by an obvious change of variable to

$$\mathbf{P}\{X^+ > t\} \leq e^{\sigma^+ t} \int_{\Omega} e^{\sigma^+ X^+} d\mathbf{P}.$$

In the same way if for $\sigma^- < 0$ we have $\int_{\mathbf{R}_+} e^{\sigma^- x} \mu_X(x) < +\infty$ we may conclude

$$\mathbf{P}\{X^- > t\} \leq e^{\sigma^- t} \int_{\Omega} e^{-\sigma^- X^-} d\mathbf{P},$$

and finally

$$\begin{aligned} \mathbf{P}\{|X| > t\} &= \mathbf{P}(\{X^+ > t\} \cup \{X^- > t\}) \\ &\leq 2 \max \left(\int_{\Omega} e^{\sigma^+ X^+} d\mathbf{P}, \int_{\Omega} e^{-\sigma^- X^-} d\mathbf{P} \right) e^{-\min(\sigma^+, -\sigma^-) t}, \end{aligned}$$

as wanted. Suppose that the condition on the right of (5.4) is verified. For $\sigma \geq 0$ and X a positive random variable we can write (see [14, p. 172])

$$\int_{\Omega} e^{\sigma X} d\mathbf{P} = \sigma \int_0^{+\infty} e^{\sigma t} \mathbf{P}\{|X| > t\} dt.$$

Then by (5.5), for σ^+ such that $0 \leq \sigma^+ < c$

$$\begin{aligned} \int_{\mathbf{R}_+} e^{\sigma^+ x} d\mu_X(x) &\leq \int_{\Omega} e^{\sigma^+ X^+} d\mathbf{P} = \sigma^+ \int_0^{+\infty} e^{\sigma^+ t} \mathbf{P}\{X^+ > t\} dt \\ &\leq \sigma^+ k \int_0^{+\infty} e^{(\sigma^+ - c)t} dt < +\infty. \end{aligned}$$

A similar argument shows that for σ^- such that $-c < \sigma^- \leq 0$

$$\begin{aligned} \int_{\mathbf{R}_-} e^{\sigma^- x} d\mu_X(x) &\leq \int_{\Omega} e^{(\sigma^-)X^-} d\mathbf{P} = -\sigma^- \int_0^{+\infty} e^{(\sigma^-)t} \mathbf{P}\{X^- > t\} dt \\ &\leq -\sigma^- k \int_0^{+\infty} e^{(\sigma^- - c)t} dt < +\infty. \end{aligned}$$

As a consequence we have

$$\{s = \sigma + it \in \mathbf{C}: \sigma \in]-c, +c[\} \subset \tilde{\mathbf{D}}_X$$

and so $\text{Int}(\tilde{\mathbf{D}}_X) \neq \emptyset$ as wanted. Proposition 5.1 is proved.

R e m a r k 5.1. It is this limitation that forces the use of characteristic functions, which are always well defined for general random variables.

The next result clarifies the general form of the natural domain of definition of the MGF.

Proposition 5.2. *Let X be a real-valued random variable. Then there exist $\sigma_0^- \leq 0$ and $\sigma_0^+ \geq 0$ such that*

$$\begin{aligned} \{s = \sigma + it \in \mathbf{C}: \sigma \in]\sigma_0^-, \sigma_0^+[\} &\subset \tilde{\mathbf{D}}_X \\ &\subset \{s = \sigma + it \in \mathbf{C}: \sigma \in [\sigma_0^-, \sigma_0^+]\}. \end{aligned} \quad (5.6)$$

P r o o f. Recall that for $\varepsilon \geq 0$

$$e^{\sigma x} \begin{cases} \leq e^{(\sigma+\varepsilon)x}, & x \geq 0, \\ \geq e^{(\sigma+\varepsilon)x}, & x \leq 0, \end{cases}$$

and that by definition we have that for $s \in \tilde{\mathbf{D}}_X$

$$\tilde{\mu}_X(s) = \int_{\mathbf{R}_-} e^{sx} d\mu_X(x) + \mu_X(\{0\}) + \int_{\mathbf{R}_+} e^{sx} d\mu_X(x).$$

Now, if for σ^- we have $\int_{\mathbf{R}_-} e^{\sigma^- x} d\mu_X(x) < +\infty$, then for $s = \sigma + it$ such that $\sigma \geq \sigma^-$

$$\int_{\mathbf{R}_-} |e^{sx}| d\mu_X(x) = \int_{\mathbf{R}_-} e^{\sigma x} d\mu_X(x) \leq \int_{\mathbf{R}_-} e^{\sigma^- x} d\mu_X(x) < +\infty.$$

Let $\sigma_0^- := \inf\{\sigma \in \mathbf{R}: \int_{\mathbf{R}_-} e^{\sigma x} d\mu_X(x) \leq +\infty\}$ and observe that since $d\mu_X(\mathbf{R}_-) < +\infty$ we have $\sigma_0^- \leq 0$. Similarly if for σ^+ we have $\int_{\mathbf{R}_+} e^{\sigma^+ x} d\mu_X(x) < +\infty$, then for $s = \sigma + it$ such that $\sigma \leq \sigma^+$

$$\int_{\mathbf{R}_+} |e^{sx}| d\mu_X(x) = \int_{\mathbf{R}_+} e^{\sigma x} d\mu_X(x) \leq \int_{\mathbf{R}_+} e^{\sigma^+ x} d\mu_X(x) < +\infty.$$

Define $\sigma_0^+ := \sup\{\sigma \in \mathbf{R}: \int_{\mathbf{R}_+} e^{s x} d\mu_X(x) \leq +\infty\}$. Since $\mu_X(\mathbf{R}_+) < +\infty$ we have $\sigma_0^+ \geq 0$. It is now clear that the pair (σ_0^-, σ_0^+) is bound to verify the statement above. Proposition 5.2 is proved.

The next result shows that MGFs of random variables or more generally complex Laplace transforms of probability measures are holomorphic functions whenever the natural domain of definition is nontrivial.

Proposition 5.3. *Let μ be a probability measure. Denote by $\tilde{\mu}$ its complex Laplace transform defined as in (2.1). Suppose that $\text{Int}(\tilde{\mathbf{D}}_\mu) \neq \emptyset$. Then $\tilde{\mu}$ is an holomorphic function on $\tilde{\mathbf{D}}_\mu$.*

P r o o f. Suppose first that μ is a positive finite measure with compact support denoted by K . In this case the result is a simple consequence of the Lebesgue dominated convergence theorem. In fact, in case $\tilde{\mu}'(s)$ exists, we can write for $s \in \text{Int}(\tilde{\mathbf{D}})$ fixed that

$$\tilde{\mu}'(s) = \lim_{h \rightarrow 0} \frac{\tilde{\mu}(s+h) - \tilde{\mu}(s)}{h} = \lim_{h \rightarrow 0} \int_K e^{s x} \frac{e^{h x} - 1}{h} d\mu(x).$$

Now, with

$$\phi_{s,x}(h) := e^{s x} \frac{e^{h x} - 1}{h}$$

we have $\lim_{h \rightarrow 0} \phi_{s,x}(h) = x e^{s x}$ which is a bounded function for $x \in K$, say by a constant M_K . For $x \in K$ and h small enough

$$|\phi_{s,x}(h)| \leq |\phi_{s,x}(h) - x e^{s x}| + |x e^{s x}| \leq \varepsilon + M_K.$$

By the Lebesgue dominated convergence theorem we will have

$$\tilde{\mu}'(s) = \int_K x e^{s x} d\mu(x) \in \mathbf{C}, \tag{5.7}$$

since μ is a finite measure. For a general μ we will consider an approximation by a sequence of measures with compact support. Consider ϕ_n a continuous function with compact support such that $0 \leq \phi_n \leq 1$, $\phi_n \equiv 1$ over $[-n, +n]$ and the support of ϕ_n is a subset of $[-2n, +2n]$. Let $\mu_n := \phi_n \mu$. Then μ_n is a finite measure with compact support such that $\tilde{\mu}_n$ is well defined on $\tilde{\mathbf{D}}$ and so it is a holomorphic function on the interior of this set by the preceding argument. We will now show that $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ converges uniformly on compacts to $\tilde{\mu}$. Consider now an arbitrary compact set $K \subset \tilde{\mathbf{D}}$. We have the following estimates:

$$\begin{aligned} \sup_{s \in K} |\tilde{\mu}_n(s) - \tilde{\mu}(s)| &\leq \sup_{s \in K} \left| \int_{\mathbf{R}_-} e^{s x} (1 - \phi_n(x)) d\mu(x) \right| + |\mu(\{0\}) - \mu_n(\{0\})| \\ &\quad + \sup_{s \in K} \left| \int_{\mathbf{R}_+} e^{s x} (1 - \phi_n(x)) d\mu(x) \right|. \end{aligned}$$

For a start it is obvious that $\lim_{n \rightarrow +\infty} |\mu(\{0\}) - \mu_n(\{0\})| = 0$. Observe also that since for all $s \in K$

$$|e^{s x} (1 - \phi_n(x))| \leq e^{\sigma_0^- x} (1 - \phi_n(x)) \leq 2 e^{\sigma_0^- x}$$

with the function on the right being μ integrable we may apply the Lebesgue dominated convergence theorem to have

$$\lim_{n \rightarrow +\infty} \sup_{s \in K} \left| \int_{\mathbf{R}_-} e^{sx} (1 - \phi_n(x)) d\mu(x) \right| \leq \lim_{n \rightarrow +\infty} \int_{\mathbf{R}_-} e^{\sigma_0^- x} (1 - \phi_n(x)) d\mu(x) = 0.$$

The same reasoning applies to the integral over \mathbf{R}_+ and as a consequence we also have

$$\lim_{n \rightarrow +\infty} \sup_{s \in K} \left| \int_{\mathbf{R}_+} e^{sx} (1 - \phi_n(x)) d\mu(x) \right| = 0.$$

In order to conclude let us observe that the sequence of holomorphic functions $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ converges uniformly on the compact sets of $\text{Int}(\tilde{\mathbf{D}})$ to $\tilde{\mu}$ and so this last function is holomorphic in $\text{Int}(\tilde{\mathbf{D}})$ by a well-known theorem of complex analysis (see, for instance, [14, p. 214]).

Theorem 5.1. *Suppose that μ, μ_1, μ_2, \dots are probability measures such that $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ converges to $\tilde{\mu}$ over $\tilde{\mathbf{D}}_\mu$ such that $\text{Int}(\tilde{\mathbf{D}}_\mu) \neq \emptyset$. Then $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ is a tight sequence of measures.*

P r o o f. Let us prove that

$$\forall \varepsilon > 0 \quad \exists r_\varepsilon > 0 \quad \sup_{n \in \mathbf{N}} \int_{|x| \geq r_\varepsilon} d\mu_n(x) \leq \varepsilon. \tag{5.8}$$

Consider σ_1^-, σ_1^+ such that $0 \in [\sigma_1^-, \sigma_1^+] \subset]\sigma_0^-, \sigma_0^+[$. We will show first that

$$\exists c > 0 \quad \forall s \in \{z = \sigma + it \in \mathbf{C} : \sigma \in [\sigma_1^-, \sigma_1^+]\} \quad |1 - \tilde{\mu}(s)| \leq c|s|. \tag{5.9}$$

In fact, since $\tilde{\mu}$ is holomorphic in $\text{Int}(\mathbf{D}_\mu)$, for $\sigma + it$ such that $\sigma \in [\sigma_1^-, \sigma_1^+]$ we will have

$$|1 - \tilde{\mu}(s)| \leq |s| \sup_{r \in]0, s|] } |\tilde{\mu}'(r)|,$$

where $]0, s|]$ is the convex envelope in the plane of the set $\{0, s\}$. By (5.7) we have for $r = \alpha s$ with $\alpha \in [0, 1]$

$$\begin{aligned} |\tilde{\mu}'(r)| &= \left| \int_{\mathbf{R}} x e^{rx} d\mu(x) \right| \leq \int_{\mathbf{R}_-} (-x) e^{\sigma_1^- x} d\mu(x) + \mu(\{0\}) + \int_{\mathbf{R}_+} x e^{\sigma_1^+ x} d\mu(x) \\ &= \int_{\mathbf{R}_-} (-x) e^{(\sigma_1^- - \sigma_0^-)x} e^{\sigma_0^- x} d\mu(x) + \mu(\{0\}) \\ &\quad + \int_{\mathbf{R}_+} x e^{(\sigma_1^+ - \sigma_0^+)x} e^{\sigma_0^+ x} d\mu(x). \end{aligned}$$

Now, since $\sigma_1^- > \sigma_0^-$, we have $\lim_{x \rightarrow +\infty} -x e^{(\sigma_1^- - \sigma_0^-)x} = 0$ and so this function is bounded by some constant, say M_- , in \mathbf{R}_- . By a similar argument the function $x e^{(\sigma_1^+ - \sigma_0^+)x}$ is bounded by some constant M_+ in \mathbf{R}_+ . As a consequence,

$$|\tilde{\mu}'(r)| \leq M_- \int_{\mathbf{R}_-} e^{\sigma_0^- x} d\mu(x) + \mu(\{0\}) + M_+ \int_{\mathbf{R}_+} e^{\sigma_0^+ x} d\mu(x) = C < +\infty,$$

and so (5.9) is proved. We will now prove (5.8). Consider a given $\varepsilon > 0$. As a consequence of the hypotheses made on the sequence $(\mu_n)_{n \in \mathbf{N}}$ we have that

$$\exists n_0 \in \mathbf{N} \quad \forall n \in \mathbf{N} \quad n \geq n_0 \implies |1 - \tilde{\mu}_n(0 + it)| \leq |1 - \tilde{\mu}(0 + it)| + \frac{\varepsilon}{4}.$$

Now, by a well-known tail estimate (see [6, p. 85]), for every $r > 0$ and $n > n_0$,

$$\begin{aligned} \mu_n(|x| > r) &\leq \frac{r}{2} \int_{-2/r}^{2/r} (1 - \hat{\mu}_n(t)) dt \leq \frac{r}{2} \int_{-2/r}^{2/r} (1 - \tilde{\mu}_n(0 + it)) dt \\ &\leq \frac{r}{2} \int_{-2/r}^{2/r} \left(|1 - \tilde{\mu}(0 + it)| + \frac{\varepsilon}{4} \right) dt \leq \frac{r}{2} \int_{-2/r}^{2/r} \left(c|t| + \frac{\varepsilon}{4} \right) dt \\ &= \frac{4c}{r} + \frac{\varepsilon}{4}. \end{aligned}$$

We can now choose r_0 such that for $n > n_0$ and $r > r_0$

$$\mu_n(|x| > r) \leq \varepsilon.$$

Also for $m \in \{1, \dots, n_0\}$, since $\mu_m(\mathbf{R}) = 1$, there exists $r_m > 0$ such that for $r > r_m$ we have $\mu_m(|x| > r) \leq \varepsilon$. Choosing $r_\varepsilon = \max_{0 \leq n \leq n_0} r_n$ we will have

$$\sup_{n \in \mathbf{N}^*} \mu_n(|x| > r) \leq \varepsilon$$

as wanted. Theorem 5.1 is proved.

In fact more is true. The following result can be deduced from an exercise stated in [6, p. 101]. It is a continuity theorem for moment generating functions.

Theorem 5.2. *Suppose that μ, μ_1, μ_2, \dots are probability measures such that $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ converges to $\tilde{\mu}$ on some set $\tilde{\mathbf{D}}_\mu = \{s = \sigma + it \in \mathbf{C} : \sigma \in [\sigma^-, \sigma^+]\}$ such that $\text{Int}(\tilde{\mathbf{D}}_\mu) \neq \emptyset$. Then $(\tilde{\mu}_n)_{n \in \mathbf{N}}$ converges weakly to μ .*

P r o o f. We show first that if $(\mu_{n_k})_{k \in \mathbf{N}}$ is a subsequence of $(\mu_n)_{n \in \mathbf{N}}$ converging vaguely to some probability distribution ν , then for $s \in \text{Int}(\tilde{\mathbf{D}}_\mu)$ we have

$$\lim_{n \rightarrow +\infty} \tilde{\mu}_{n_k} = \tilde{\nu}(s). \tag{5.10}$$

Let $\varepsilon > 0$ be given. Let $s = \sigma + it \in \text{Int}(\tilde{\mathbf{D}}_\mu)$ and consider for all $r > 0$, ϕ_r a continuous function such that $0 \leq \phi_r \leq 1$, $\phi_r \equiv 1$ on $[-r, +r]$ and such that the support of ϕ_r is contained in $[-2r, +2r]$. We can write

$$\begin{aligned} |\tilde{\mu}_{n_k}(s) - \tilde{\nu}(s)| &= |\mu_{n_k}(e^{sx}) - \nu(e^{sx})| \leq |\mu_{n_k}(e^{sx}) - \mu_{n_k}(\phi_r e^{sx})| \\ &\quad + |\mu_{n_k}(\phi_r e^{sx}) - \nu(\phi_r e^{sx})| + |\nu(\phi_r e^{sx}) - \nu(e^{sx})|. \end{aligned}$$

By Hölder's inequality, for $p \geq 1$ such that $p\sigma \in]\sigma^-, \sigma^+[$ we have

$$\begin{aligned} & |\mu_{n_k}(e^{s_x}) - \mu_{n_k}(\phi_r e^{s_x})| + |\nu(\phi_r e^{s_x}) - \nu(e^{s_x})| \\ & \leq \int_{\mathbf{R}} e^{\sigma x} (1 - \phi_r)(x) d(\mu_{n_k} + \nu)(x) \\ & \leq \left(\int_{\mathbf{R}} e^{p\sigma x} d(\mu_{n_k} + \nu)(x) \right)^{1/p} \left(\int_{\mathbf{R}} (1 - \phi_r)^q(x) d(\mu_{n_k} + \nu)(x) \right)^{1/q} \\ & \leq (\tilde{\mu}_{n_k}(p\sigma) + \tilde{\nu}(p\sigma))^{1/p} ((\mu_{n_k} + \nu)(\{|x| \geq 2r\}))^{1/q}. \end{aligned}$$

Because $\lim_{k \rightarrow +\infty} \tilde{\mu}_{n_k}(p\sigma) = \tilde{\mu}(p\sigma)$ we have that for some constant $c > 0$

$$\forall k \in \mathbf{N} \quad (\tilde{\mu}_{n_k}(p\sigma) + \tilde{\nu}(p\sigma))^{1/p} \leq c.$$

Since by Theorem 5.1 the sequence $(\mu_n)_{n \in \mathbf{N}}$ is tight it follows that the sequence $(\mu_{n_k} + \nu)_{k \in \mathbf{N}}$ is also tight and so there exists $r_\varepsilon > 0$ such that

$$c \left((\mu_{n_k} + \nu)(\{|x| \geq 2r_\varepsilon\}) \right)^{1/q} \leq \frac{\varepsilon}{2}.$$

Also since $\phi_{r_\varepsilon} e^{s_x}$ is a continuous function with compact support, there exists $k_0 \in \mathbf{N}$ such that

$$\forall k \in \mathbf{N} \quad k \geq k_0 \implies |\mu_{n_k}(\phi_{r_\varepsilon} e^{s_x}) - \nu(\phi_{r_\varepsilon} e^{s_x})| \leq \frac{\varepsilon}{2}.$$

As a consequence we have that for all $s = \sigma + it \in \text{Int}(\tilde{\mathbf{D}}_\mu)$ equality (5.10) holds as wanted. Since $\tilde{\mu}$ and $\tilde{\nu}$ are holomorphic functions, the hypothesis made on the sequence $(\mu_n)_{n \in \mathbf{N}}$ and equality (5.10) shows that $\tilde{\mu} \equiv \tilde{\nu}$. We now observe that, in fact, if $(\mu_{n_k})_{k \in \mathbf{N}}$ is a subsequence of $(\mu_n)_{n \in \mathbf{N}}$ converging weakly to a probability distribution ν then $\mu = \nu$ since the subsequence $(\mu_{n_k})_{k \in \mathbf{N}}$ is *a fortiori* vaguely convergent to ν . By a well-known result [17, Chap. III, § 3, Lemma 1, p. 414] we finally have that $(\mu_n)_{n \in \mathbf{N}}$ being tight and such that every weakly convergent subsequence converges to the same probability measure μ , $(\mu_n)_{n \in \mathbf{N}}$ converges weakly to μ thus ending the proof of the theorem.

СПИСОК ЛИТЕРАТУРЫ

1. Dowling M. M., Nakamura M. Estimating parameters for discrete distributions via the empirical probability generating function. — Comm. Statist. Simulation Comput., 1997, v. 26, № 1, p. 301–313.
2. Феллер В. Введение в теорию вероятностей и ее приложения, т. 1. М.: Мир, 1984, 527 с., 752 с.
3. Hardy G. H., Riesz M. The General Theory of Dirichlet's Series. Cambridge Univ. Press, 1915.
4. Hoffmann-Jørgensen J. Probability with a View towards Statistics. V. I. New York: Chapman & Hall, 1994, 589 p.
5. Kahane J.-P. The last problem of Harald Bohr. — J. Aust. Math. Soc., 1989, v. A47, p. 133–152.

6. *Kallenberg O.* Foundations of Modern Probability. New York: Springer-Verlag, 2002, 638 p.
7. *Ландо С. К.* Лекции о производящих функциях. М.: МИИМО, 2002, 143 с.
8. *Marques M. S., Pérez-Abreu V.* Law of large numbers and central limit theorem for the empirical probability generating function of stationary random sequences and processes. — Symposium on Probability Theory and Stochastic Processes (Guanajuato, 1988). Ed. by M. E. Caballero et al. México: Soc. Mat. Mexicana, 1989, p. 100–109. (Aportaciones Mat. Notas Investigación, v. 4.)
9. *Nakamura M., Pérez-Abreu V.* Empirical probability generating function: An overview. — Insurance Math. Econom., 1993, v. 12, № 3, p. 287–295.
10. *Nakamura M., Pérez-Abreu V.* Exploratory data analysis for counts using the empirical probability generating function. — Comm. Statist. Theory Methods, 1993, v. 22, № 3, p. 827–842.
11. *Nakamura M., Pérez-Abreu V.* Use of an empirical probability generating function for testing a Poisson model. — Canad. J. Statist., 1993, v. 21, № 2, p. 149–156.
12. *Rémillard B., Theodorescu R.* Inference based on the empirical probability generating function for mixtures of Poisson distributions. — Statist. Decisions, 2000, v. 18, № 4, p. 349–366.
13. *Resnick S. I.* A Probability Path. Boston: Birkhäuser, 1999, 453 p.
14. *Rudin W.* Real and Complex Analysis. New York: McGraw-Hill, 1987, 416 p.
15. *Rueda R., Pérez-Abreu V., O'Reilly F.* Goodness of fit for the Poisson distribution based on the probability generating function. — Comm. Statist. Theory Methods, 1991, v. 20, № 10, p. 3093–3110.
16. *Rueda R., O'Reilly F.* Tests of fit for discrete distributions based on the probability generating function. — Comm. Statist. Simulation Comput., 1999, v. 28, № 1, p. 259–274.
17. *Ширяев А. Н.* Вероятность, т. 1, 2. М.: МИИМО, 2004, 927 с.
18. *Saks S., Zygmund A.* Analytic Functions. Warszawa–Wrocław: Polskie Towarzystwo Matematyczne, 1952, 451 p. (Monografie Matematyczne, t. XXVIII.)

Поступила в редакцию
16.VII.2004