PROBABILITY GENERATING FUNCTIONS FOR DISCRETE REAL VALUED RANDOM VARIABLES

MANUEL L. ESQUÍVEL

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.

Abstract. The probability generating function is a powerful technique for studying the law of finite sums of independent discrete random variables taking integer positive values. For real valued discrete random variables, the well known elementary theory of Dirichlet series and the symbolic computation packages available nowadays, such as Mathematica 5 TM, allows us to extend to general discrete random variables this technique. Being so, the purpose of this work is twofold. Firstly we show that discrete random variables taking real values, non necessarily integer or rational, may be studied with probability generating functions. Secondly we intend to draw attention to some practical ways of performing the necessary calculations.

1. Classical probability generating functions

Generating functions are an useful and up to date tool in nowadays practical mathematics, in particular in discrete mathematics and combinatorics (see [Lando 03]) and, in the case of probability generating functions, in distributional convergence results as in [Kallenberg 02][p. 84]. Its uses in basic probability are demonstrated in the classic reference [Feller 68][p. 266]. More recently, probability generating functions for integer valued random variables have been studied intensively mainly with some applied purposes in mind. See, for instance [Dowling et al 97], [Marques et al 89], [Nakamura et al 93], [Nakamura et al 93a], [Nakamura et al 93b], [Rémillard et al 00], [Rueda et al 91] and [Rueda et al 99].

The natural extension of the definition of probability generating function to non negative real valued random variable X, as the expectation of the function t^X, is very clearly presented in the treatise [Hoffmann-Jørgensen 94][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 $\mathbb{Z}$ and consider that for all $k \in \mathbb{Z}$ we have $p_k := \mathbb{P}[X = k] \in [0, 1]$.

Date: July 19, 2004.

1991 *Mathematics Subject Classification.* 60–08, 60E10, 30B50.

Key words and phrases. Probability Generating Functions, finite sums of independent real valued discrete random variables, Dirichlet series.

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) .*

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

$$\mathbb{D}_X = \{z \in \mathbb{C} : \int_\Omega z^X d\mathbb{P} \in \mathbb{C}\} = \{z \in \mathbb{C} : \int_\Omega \mid z \mid^X d\mathbb{P} < +\infty\} .$$

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

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

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 $C(\rho_1, \rho_2) = \{0 \leq \rho_1 \leq \mid z \mid \leq \rho_2 \leq +\infty\}$ where by Hadamard's formula we have: $\rho_1 = \limsup_{n\to+\infty} \sqrt[n]{p_{-n}}$ and $\rho_2 = 1/\limsup_{n\to+\infty} \sqrt[n]{p_n}$. As 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 $\mathbb{D}_X = C(\rho_1, \rho_2)$. If for all $n < 0$ we have that $p_n = 0$ then, ψ_X is represented by a Taylor series around zero, $\{\mid z \mid < 1\} \subset \mathbb{D}_X$ and so,

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

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 as we have for some γ_r, the border of the circle of radius $r \in]\rho_1, \rho_2[$ centered at zero, that

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

The main purpose of this paper is to extend probability generating function techniques to discrete random variables taking real values non necessarily 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 $(\alpha_k)_{k\in\mathbb{Z}}$ such that for some sequence of probabilities $(p_k)_{k\in\mathbb{Z}} \in [0,1]^{\mathbb{Z}}$, thus verifying $\sum_{k=-\infty}^{+\infty} p_k = 1$, we have that $\mathbb{P}[X = \alpha_k] = p_k$ for all $k \in \mathbb{Z}$. The law of X is the probability measure μ_X is given by $\mu_X = \sum_{k=-\infty}^{+\infty} p_k \delta_{\alpha_k}$. We will constantly use that for any $t > 0$ and $x \in \mathbb{R}$, $t^x := e^{x \ln(t)}$ 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:

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

Remark 2.2. Let us observe that $\mathbb{E}[t^X]$ 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 as we have

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

A natural question is then to determine the exact description of the convergence domain of ψ_X, that is the set $\mathbb{D}_X := \{t > 0 : \psi_X(t) < +\infty\}$ where the PGF of X is, in fact, a real valued function. We will address this question in theorem 2.5 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 [Kallenberg 02, p. 84] we denote by $\hat{\mu}_X$ the **characteristic function** of X, defined for all $t \in \mathbb{R}$ by:

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

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

Definition 2.3. The **moment generating function** (MGF) $\tilde{\mu}_X$ of a real valued random variable X is defined to be

$$\tilde{\mu}_X(z) = \mathbb{E}[e^{zX}] = \int_{\mathbb{R}} e^{zx} d\mu_X(x) \ . \tag{2.1}$$

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

Remark 2.4. For any random variable X the natural domain of its MGF is never empty as $0 \in \tilde{\mathbb{D}}_X$. However, important properties depend crucially on $\tilde{\mathbb{D}}_X$ having a non empty interior. For that reason some authors (see [Resnick 01, p. 294]) consider that $\tilde{\mu}_X$ is defined only in that case. On subsection 5.2 we will deal more thoroughly with this question.

There are natural relations among these functional transforms. For all t for each the functional transforms involved are well defined we have:

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

Consider the following further convention on the notation used above, that is, X is a random variable taking as values the ordered sequence of real values $(\alpha_k)_{k\in\mathbb{Z}}$ 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 2.5. *Let X be a random variable and let ψ_X denote its PGF. We then have that:*

1. *If X takes an finite number of real values*

$$\mathbb{D}_X =]0, +\infty[\ .$$

2. *If X takes an infinite number of real values without accumulation points*

$$\exists u_0, v_0 \in]-\infty, 0] \ , \ ]e^{u_0}, e^{-v_0}[\subset \mathbb{D}_X \subset [e^{u_0}, e^{-v_0}] \ . \tag{2.3}$$

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

Proof. In the first case we have that 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 that:

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

where the sequences $(\alpha_k)_{k\in\mathbb{N}^*}$ and $(\beta_k)_{k\in\mathbb{N}^*}$ are increasing. Under the hypotheses that these sequences do not have a limit in $\mathbb{R}$, in the formula (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 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

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

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

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

The last result stated in the theorem is an immediate consequence of proposition 5.3 in the last section as we have that $\mathbb{D}_X = \{\mid e^z \mid : z \in \tilde{\mathbb{D}}_X\}$. ☐

Remark 2.6. The use of Dirichlet series or MGF is further justified as it allows to determinate 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 formulas (5.2) or (5.3). If X is discrete having an exponential decaying tail then we may get σ_0^+ and σ_0^- as defined in the proof of proposition 5.3 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 a t, the **floor function** as

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

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

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

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

Theorem 2.7. *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*

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

denote the PGF of the random variable X. Then, obviously, we have that:

$$p_{-M} = \mathbb{P}[X = \alpha_{-M}] = \lim_{t \to 0^+} \left(t^{\alpha_{-M}} \times \psi_X(t)\right)$$

and derivating enough times:

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

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

Proof. For the first result it is enough to observe that:

$$\begin{aligned} t^{\alpha_{-M}} \times \psi_X(t) = p_{-M} +& \\ & p_{-M-1} t^{\alpha_{-M} - \alpha_{-M-1}} + \cdots + p_0 t^{\alpha_{-M}} + p_1 t^{\alpha_1 + \alpha_{-M}} + \ldots \end{aligned}$$

For the second result in the statement of the theorem and, in case we have $\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor \geq 1$, just derive the preceeding formula above $\lfloor \alpha_{-M} - \alpha_{-M+1} \rfloor$ times. □

Remark 2.8. The practical interest of this theorem if 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 below.

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 2.9. *Let X, Y be random variables such that for some neighborhood V of $1 \in \mathbb{D}_X \cap \mathbb{D}_Y$ we have ψ_X and ψ_Y well defined and verifying*

$$\forall t \in V \;\; \psi_X(t) = \psi_Y(t) . \tag{2.6}$$

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

Proof. Condition (2.6) implies that $\tilde{\mu}_X$ and $\tilde{\mu}_Y$ are well defined on

$$\mathbb{D} = \{s = \sigma + it \in \mathbb{C} : \sigma \in \{\ln(u) : u \in \mathrm{Int}(V)\}\} \subset \mathbb{D}_X \cap \mathbb{D}_Y$$

and that for $s \in \mathbb{D}$ we have $\tilde{\mu}_X(s) = \tilde{\mu}_Y(s)$. As, by proposition 5.4, $\tilde{\mu}_X$ and $\tilde{\mu}_Y$ are holomorphic functions on $\mathbb{D}_X$ and $\mathbb{D}_Y$, respectively, they are certainly equal as two holomorphic functions coinciding in a set having an accumulation point coincide. In order to conclude it suffices to observe that as $\tilde{\mu}_X \equiv \tilde{\mu}_Y$ we have

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

That is, the characteristic functions of X and Y coincide . Being so it is well known that $X \overset{\mathrm{d}}{=} Y$, as wanted. □

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

Theorem 2.10. *Let $X_1, \ldots, X_n, \ldots X$ real valued random variables, and $\psi_1, \ldots \psi_n, \ldots \psi$ the correspondent PGF. Suppose that for V, some neighborhood of $1 \in \mathbb{D}_X$, we have that:*

$$\forall n \in \mathbb{N} \;\; V \subset \mathbb{D}_{X_n} \tag{2.7}$$

and

$$\forall t \in V \quad \lim_{n\to+\infty} \psi_n(t) = \psi(t) \ . \tag{2.8}$$

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

Proof. Let $\mathbb{D} := \{s = \sigma + it \in \mathbb{C} : e^\sigma \in V\}$. Then condition (2.8) implies that

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

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

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

Moreover we have $\lim_{n\to+\infty} \mid \mu_{X_n} - \mu_X \mid (e^{\sigma x}) = 0$ that is, $(\tilde{\mu}_{X_n})_{n\in\mathbb{N}}$ converges over $\mathbb{D}$ towards $\tilde{\mu}_X$. Being so, by theorem 5.6 we have that $(X_n)_{n\in\mathbb{N}}$ converges in law to X. ☐

2.3. **The PGF of a sum of iid random variables.** One important usage of PGF 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 shown in section 3. As in the case of non negative 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 2.11. *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* $\mathbb{D}_{X+Y} = \mathbb{D}_X \cap \mathbb{D}_Y$, *then*

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

Proof. The first equality is a consequence of the independence of X and Y. The second equality is simple given by the usual product of two absolutely convergent series. ☐

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

Corollary 2.12. *Let* $X_1, X_2, \ldots, X_m$ *be a sequence of independent and identically distributed with* X *a discrete real valued random variable. Then we have that for all* $t > 0$*:*

$$\psi_{X_1+X_2+\cdots+X_m}(t) = (\psi_X(t))^m$$

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

$$\psi_{X_1+X_2+\cdots+X_m}(t) = \sum_{i_1,\ldots,i_m=-\infty}^{+\infty} p_{i_1} \ldots p_{i_m} t^{\alpha_{i_1} + \cdots + \alpha_{i_m}}$$

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 independent copies of a random variable taking a finite number of real values. In he 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.

$$X_1 = \begin{cases} -1 & \text{with probability } 1/16 \\ -2/3 & \text{with probability } 3/16 \\ -1/3 & \text{with probability } 3/16 \\ 0 & \text{with probability } 1/8 \\ 1/3 & \text{with probability } 3/16 \\ 2/3 & \text{with probability } 3/16 \\ 1 & \text{with probability } 1/16\ . \end{cases} \tag{3.1}$$

The PGF of X_1 is given by

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

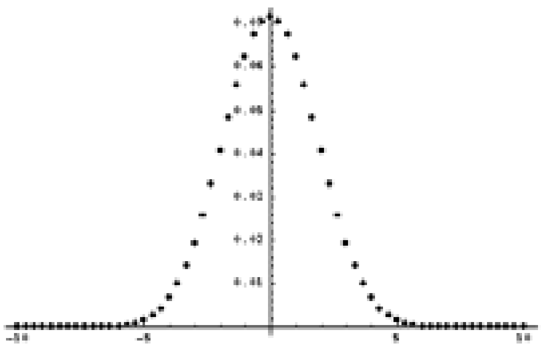
For a deeper understanding of fair marking an exam with a set of, say, ten multiple choice questions it is important to know the distribution of the sum of ten independent copies of X_1 which we denote by Y. We know that $\psi_Y = (\psi_{X_1}(t))^{10}$, With a symbolic computation package we have expanded this power in a sum of 61 terms of the form $a \times t^{\alpha}$ 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 correspondent 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. This next 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 $\alpha a t^{\alpha-1}/a$ 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 $\alpha a t^{\alpha-1}/a$ 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 correspondent 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 the following figure.



Note that a first inspection of this figure suggests the use of a normal approximation for Y.

For a second example consider a random variable X_2 taking some irrational values defined as in the following.

$$X_2 = \begin{cases} -3/4 & \text{with probability } 0.34 \\ \pi & \text{with probability } 0.33 \\ 2\pi & \text{with probability } 0.33 \, . \end{cases} \tag{3.2}$$

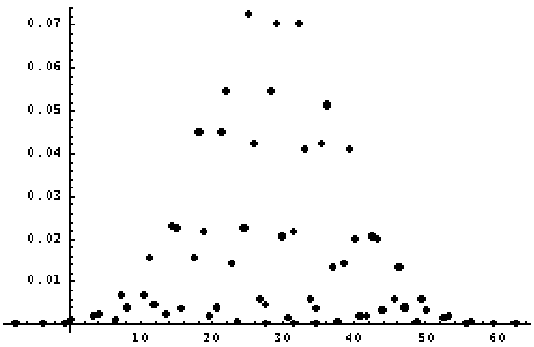
Obviously the PGF of X_2 is given by

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

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

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

Proceeding as in the first example above, we get probability distribution of Z is given in the following figure.



Obviously, using a normal approximation for Z can't be thought in this case.

4. Random variables taking an infinite number of values

In the preceeding section we showed how to effectively determine the probability distribution of a finite sum of independent identically distributed random variables taking a finite number of real values using the PGF. In this section we will show that for a sum of iid 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. Firstly we define a sequence real valued random variables $(X^{M,N})_{M,N\in\mathbb{N}}$, taking a finite number of values. It is then easy to see that the the sequence $(X^{M,N})_{M,N\in\mathbb{N}}$ 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 preceeding section, let $(\alpha_k)_{k\in\mathbb{N}}$ denote the ordered sequence of real numbers which are the values taken by X with the correspondent probabilities $(p_k)_{k\in\mathbb{N}}$. 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 $M, N \in \mathbb{N}$ the random variable $X^{M,N}$ such that:

$$\begin{cases} \tilde{p}_k := \mathbb{P}[X^{M,N} = \alpha_k] = \mathbb{P}[X = \alpha_k] = p_k & \forall k \in \{-M, \dots, -1, 1, \dots, N\} \\ \tilde{p}_0 := \mathbb{P}[X^{M,N} = 0] = p_0 + \sum_{k=-\infty}^{-M-1} p_k + \sum_{k=N+1}^{+\infty} p_k \end{cases} \tag{4.1}$$

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 $(X^{M,N})_{M,N\in\mathbb{N}}$ converges in law to X.*

Proof. Let f be a continuous bounded function of $\mathbb{R}$. As we have that $\mu_{X^{M,N}}$ 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:

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

Let K denote the bound of $\mid f \mid$. We then will have:

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

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, $X_1, \ldots, X_m$ be m independent copies of X and $X_1^{M,N}, \ldots, X_m^{M,N}$ be m independent copies of $X^{M,N}$. Then the sequence of random variables $(X_1^{M,N} + \cdots + X_m^{M,N})_{M,N \in \mathbb{N}}$ converges in law to $X_1 + \cdots + X_m$.*

Proof. It is a simple consequence of the continuity theorem of Lévy-Cramér. □

We now show that the sequence of the PGF 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 2.5. The sequence $(\psi_{X^{M,N}})_{M,N \in \mathbb{N}}$ converges uniformly for ψ_X on $]e^{u_0}, e^{-v_0}[$.*

Proof. By the definitions we simply have to observe that:

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

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 $X_1^{M,N} + \cdots + X_m^{M,N}$ converges uniformly to the PGF of $X_1 + \cdots + X_m$.

Theorem 4.4. *With the same notations used in statement and in the proof of theorem 2.5 suppose that $u_0, v_0 < 0$. Then, for every $m \geq 1$ and every $\epsilon > 0$, there exists $M_0, N_0 \in \mathbb{N}$, $u, v \in \mathbb{R}$ verifying*

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

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

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

where

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

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

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

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

$$\forall t \in]e^{u_0}, e^{-v_0}[\;\;,\;\; \mid \psi_X(t) - \psi_{X^{M,N}}(t) \mid \leq \frac{\epsilon}{2} \; . \tag{4.3}$$

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

$$\forall t \in]e^{u}, e^{-v}[\;\;,\;\; \mid \psi_X(t) \mid \leq 1 + \frac{\epsilon}{2} \; . \tag{4.4}$$

As a consequence of estimate (4.3) we then have:

$$\forall t \in]e^{u}, e^{-v}[\;\;,\;\; \mid \psi_{X^{M,N}}(t) \mid \leq 1 + \epsilon \; . \tag{4.5}$$

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

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

Defining

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

it follows from the previous estimates (4.4) and (4.5) that the 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 PGF. We suppose that the results on subsection 5.2 may have interest on its own.

5.1. **A quick review of Dirichlet series.** In this subsection we recall from [Hardy & Riesz] or [Zaks & Zygmund] some results that were needed in previous sections. A Dirichlet series is a series of the form

$$\sum_{n=1}^{+\infty} a_n e^{-\lambda_n s} \; , \tag{5.1}$$

where $(a_n)_{n\mathbb{N}^*}$ is a sequence of complex numbers and $(\lambda_n)_{n\mathbb{N}^*}$ 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

$$\{s \in \mathbb{C} :\mid \mathrm{Arg}(s - s_0) \mid \leq a < \frac{\pi}{2}\} .$$

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

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

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

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

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

$$\alpha = \limsup_{n \to +\infty} \frac{\sum_{k=1}^{n} \mid a_k \mid}{\lambda_n} . \tag{5.2}$$

and if $\alpha < 0$ then

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

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 (2.3). It is easy but somehow lengthy to show that $\tilde{\mathbb{D}}_X$ having a non empty interior happens only for random variables with exponential decaying tails.

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

$$Int(\tilde{\mathbb{D}}_X) \neq \emptyset \Leftrightarrow \exists k, c > 0 \;\; \mathbb{P}[\mid X \mid > x] \leq k e^{-cx} . \tag{5.4}$$

Proof. Suppose that the natural domain of definition of MGF has non empty interior. Let us deal with $\int_{\mathbb{R}_+} e^{\sigma x} \, d\mu_X(x)$ first. Suppose that for $\sigma^+ > 0$ we have $\int_{\mathbb{R}_+} e^{\sigma^+ x} \, d\mu_X(x) < +\infty$. We then have:

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

By Tchebytcheff inequality

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

which is equivalent by an obvious change of variable to:

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

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

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

and finally

$$\begin{aligned}\mathbb{P}[\mid X \mid> t] = \mathbb{P}[\{X^+ > t\} \cup \{X^- > t\}] \leq \\ \leq 2 \sup \left(\int_\Omega e^{\sigma^+ X^+} d\mathbb{P}, \int_\Omega e^{-\sigma^- X^-} d\mathbb{P} \right) e^{-\inf(\sigma^+, -\sigma^-)t} ,\end{aligned}$$

as wanted. Suppose that the condition on the right of (5.4) is verified. As we can write (see [Rudin 86, p. 172]) for $\sigma \geq 0$ and X a positive random variable

$$\int_\Omega e^{\sigma X} d\mathbb{P} = \sigma \int_0^{+\infty} e^{\sigma t} \mathbb{P}[\mid X \mid> t] dt$$

then by formula (5.5), and σ^+ such that $0 \leq \sigma^+ < c$

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

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

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

As a consequence we have

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

and so $\mathrm{Int}(\tilde{\mathbb{D}}_X) \neq \emptyset$ as wanted. □

Remark 5.2. 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 of definition the MGF.

Proposition 5.3. *Let X be a real valued random variable. Then, there exists $\sigma_0^- \leq 0$ and $\sigma_0^+ \geq 0$ such that:*

$$\{s = \sigma + it \in \mathbb{C} : \sigma \in]\sigma_0^-, \sigma_0^+[\} \subset \tilde{\mathbb{D}}_X \subset \{s = \sigma + it \in \mathbb{C} : \sigma \in [\sigma_0^-, \sigma_0^+]\} \tag{5.6}$$

Proof. Recall that for $\epsilon \geq 0$

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

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

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

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

$$\int_{\mathbb{R}_-} \mid e^{sx} \mid \mu_X(x) = \int_{\mathbb{R}_-} e^{\sigma x} \mu_X(x) \leq \int_{\mathbb{R}_-} e^{\sigma^- x} \mu_X(x) < +\infty \, .$$

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

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

Defining $\sigma_0^+ := \sup\{\sigma \in \mathbb{R} : \int_{\mathbb{R}_+} e^{\sigma x}\mu_X(x) \leq +\infty\}$ we have that as $\mu_X(\mathbb{R}_+) < +\infty$ we have $\sigma_0^+ \geq 0$. It is now clear that the pair (σ_0^-, σ_0^+) is bound to verify the statement above. □

The nest result shows that MGF of random variables or more generally complex Laplace transforms of probability measures are holomorphic functions whenever the natural domain of definition is non trivial.

Proposition 5.4. *Let μ be a probability measure and denote by $\tilde{\mu}$ its complex Laplace transform defined as in (2.1). Suppose that Int$(\tilde{\mathbb{D}}_\mu) \neq \emptyset$. Then $\tilde{\mu}$ is an holomorphic function on $\tilde{\mathbb{D}}_\mu$.*

Proof. 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 Lebesgue dominated convergence theorem. In fact, we can write for $s \in \mathrm{Int}(\tilde{\mathbb{D}})$ fixed that, in case $\tilde{\mu}'(s)$ exists, we have

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

Now, with

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

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

$$| \phi_{s,x}(h) | \leq | \phi_{s,x}(h) - xe^{sx} | + | xe^{sx} | \leq \epsilon + M_K \ .$$

By the Lebesgue dominated convergence theorem we will have

$$\tilde{\mu}'(s) = \int_K xe^{sx} d\mu(x) \in \mathbb{C}. \tag{5.7}$$

as 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 perfectly defined on $\tilde{\mathbb{D}}$ and so it is an holomorphic function on the interior of this set by the preceeding argument. We will now show that $(\tilde{\mu}_n)_{n\in\mathbb{N}}$ converges uniformly on compacts to $\tilde{\mu}$. Consider now an arbitrary compact set $K \subset \tilde{\mathbb{D}}$. We have the following estimates.

$$\sup_{s\in K} | \tilde{\mu}_n(s) - \tilde{\mu}(s) | \leq \sup_{s\in K}\left| \int_{\mathbb{R}_-} e^{sx}(1-\phi_n(x))d\mu(x)\right| + \\ + | \mu(\{0\}) - \mu_n(\{0\}) | + \sup_{s\in K}\left| \int_{\mathbb{R}_+} e^{sx}(1-\phi_n(x))d\mu(x)\right| \ .$$

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

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

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

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

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

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

In order to conclude let us observe that the sequence of holomorphic functions $(\tilde{\mu}_n)_{n\in\mathbb{N}}$ converges uniformly on the compact sets of $\mathrm{Int}(\tilde{\mathbb{D}})$ to $\tilde{\mu}$ and so this last function is holomorphic in $\mathrm{Int}(\tilde{\mathbb{D}})$ by a well known theorem of complex analysis (see, for instance, [Rudin 86, p. 214]). ☐

Theorem 5.5. *Suppose that $\mu_1, \mu_2, \ldots, \mu_n, \ldots \mu$ are probability measures such that $(\tilde{\mu_n})_{n\in\mathbb{N}}$ converges to $\tilde{\mu}$ over $\tilde{\mathbb{D}}_\mu$ such that $Int(\tilde{\mathbb{D}}_\mu) \neq \emptyset$. Then $(\tilde{\mu_n})_{n\in\mathbb{N}}$ is a tight sequence of measures.*

Proof. We will show that

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

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

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

In fact as $\tilde{\mu}$ is holomorphic in $\mathrm{Int}(\mathbb{D}_\mu)$, 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 formula (5.7) we have for $r=\alpha s$ with $\alpha\in[0,1]$

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

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

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

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

$$\exists n_0\in\mathbb{N}\ \ \forall n\in\mathbb{N}\ \ n\geq n_0 \Rightarrow |\ 1-\tilde{\mu}_n(0+it)\ |\leq|\ 1-\tilde{\mu}(0+it)\ |+\frac{\epsilon}{4}\ .$$

Now, by a well known tail estimate (see [Kallenberg 02, p. 85]) we have that for every $r > 0$ and $n > n_0$,

$$\mu_n(\mid x \mid > r) \le \frac{r}{2}\int_{-2/r}^{2/r}(1-\hat{\mu}_n(t))dt \le \frac{r}{2}\int_{-2/r}^{2/r}(1-\tilde{\mu}_n(0+it))dt \le$$
$$\le \frac{r}{2}\int_{-2/r}^{2/r}\Big(\mid 1-\tilde{\mu}(0+it)\mid + \frac{\epsilon}{4}\Big)\,dt \le \frac{r}{2}\int_{-2/r}^{2/r}(c\mid t\mid + \frac{\epsilon}{4})dt = \frac{4c}{r}+\frac{\epsilon}{4}\ .$$

We may now choose r_0 such that for $n > n_0$ and $r > r_0$ we have

$$\mu_n(\mid x \mid > r) \le \epsilon\ .$$

Also for $m \in \{1,\dots,n_0\}$, as we have $\mu_m(\mathbb{R}) = 1$ there exists $r_m > 0$ such that for $r > r_m$ we have $\mu_m(\mid x \mid > r) \le \epsilon$. Choosing $r_\epsilon = \max_{0\le n\le n_0} r_n$ we will have

$$\sup_{n\in\mathbb{N}^*}\mu_n(\mid x \mid > r) \le \epsilon$$

as wanted. □

In fact more is true. The next result may be deduced from an exercise stated in [Kallenberg 02, p. 101]. It is a continuity theorem for moment generating functions.

Theorem 5.6. *Suppose that $\mu_1, \mu_2, \dots, \mu_n, \dots \mu$ are probability measures such that $(\tilde{\mu_n})_{n\in\mathbb{N}}$ converges to $\tilde{\mu}$ on some set $\tilde{\mathbb{D}}_\mu = \{s = \sigma + it \in \mathbb{C} : \sigma \in [\sigma^-, \sigma^+]\}$ such that Int$(\tilde{\mathbb{D}}_\mu) \neq \emptyset$. Then $(\tilde{\mu_n})_{n\in\mathbb{N}}$ converges weakly to μ.*

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

$$\lim_{n\to+\infty}\tilde{\mu_{n_k}} = \tilde{\nu}(s)\ . \tag{5.10}$$

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

$$\mid \tilde{\mu}_{n_k}(s) - \tilde{\nu}(s) \mid = \mid \mu_{n_k}(e^{sx}) - \nu(e^{sx}) \mid \le$$
$$\le \mid \mu_{n_k}(e^{sx}) - \mu_{n_k}(\phi_r e^{sx}) \mid + \mid \mu_{n_k}(\phi_r e^{sx}) - \nu(\phi_r e^{sx}) \mid + \mid \nu(\phi_r e^{sx}) - \nu(e^{sx}) \mid\ .$$

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

$$\mid \mu_{n_k}(e^{sx}) - \mu_{n_k}(\phi_r e^{sx}) \mid + \mid \nu(\phi_r e^{sx}) - \nu(e^{sx}) \mid \le \int_{\mathbb{R}} e^{\sigma x}\,(1-\phi_r)(x)\,d(\mu_{n_k}+\nu)(x) \le$$
$$\le \left(\int_{\mathbb{R}} e^{p\sigma x} d(\mu_{n_k}+\nu)(x)\right)^{1/p}\ \left(\int_{\mathbb{R}}(1-\phi_r)^q(x)d(\mu_{n_k}+\nu)(x)\right)^{1/q} \le$$
$$\le (\tilde{\mu}_{n_k}(p\sigma) + \tilde{\nu}(p\sigma))^{1/p}\ ((\mu_{n_k}+\nu)(\{\mid x \mid \ge 2r\}))^{1/q}\ .$$

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

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

As by theorem 5.5 the sequence $(\mu_n)_{n\in\mathbb{N}}$ is tight then the sequence $(\mu_{n_k}+\nu)_{k\in\mathbb{N}}$ is also tight and so there exists $r_\epsilon > 0$ such that

$$c \times ((\mu_{n_k}+\nu)(\{\mid x \mid \ge 2r_\epsilon\}))^{1/q} \le \frac{\epsilon}{2}\ .$$

Also as $\phi_{r_\epsilon} e^{sx}$ is a continuous function with compact support then there exists $k_0 \in \mathbb{N}$ such that

$$\forall k \in \mathbb{N} \;\; k \geq k_0 \Rightarrow \mid \mu_{n_k}(\phi_{r_\epsilon} e^{sx}) - \nu(\phi_{r_\epsilon} e^{sx}) \mid \leq \frac{\epsilon}{2}$$

As a consequence we have that for all $s = \sigma + it \in \mathrm{Int}(\tilde{\mathbb{D}}_\mu)$ we will have equality (5.10) verified as wanted. As $\tilde{\mu}$ and $\tilde{\nu}$ are holomorphic functions, the hypothesis made on the sequence $(\mu_n)_{n \in \mathbb{N}}$ and equality (5.10) shows that $\tilde{\mu} \equiv \tilde{\nu}$. We now observe that we have in fact that if $(\mu_{n_k})_{k \in \mathbb{N}}$ is a subsequence of $(\mu_n)_{n \in \mathbb{N}}$ converging weakly to a probability distribution ν then then $\mu = \nu$ as the subsequence $(\mu_{n_k})_{k \in \mathbb{N}}$ is *a fortiori* vaguely convergent to ν. By a well known result [Shiryaev 96, p. 322] we finally have that $(\mu_n)_{n \in \mathbb{N}}$ being tight and such that every weakly convergent subsequence converges to the same probability measure μ, then $(\mu_n)_{n \in \mathbb{N}}$ converges weakly to μ thus ending the proof of the theorem. □

References

[Dowling et al 97] M. M. Dowling, M. Nakamura, Estimating parameters for discrete distributions via the empirical probability generating function. *Commun. Stat., Simulation Comput. 26* **1** (1997) 301–313.

[Feller 68] W. Feller, *An introduction to Probability theory and its Applications*, third edition, John Wiley & Sons 1968.

[Hardy & Riesz] G. H. Hardy, M. Riesz *The general theory of Dirichlet's series*, Cambridge University Press 1915.

[Hoffmann-Jørgensen 94] J. Hoffmann-Jørgensen, *Probability with a view towards statistics*, Volume I, Chapman & Hall 1994.

[Kahane 89] J.-P. Kahane, The last problem of Harald Bohr. *Journal of the Australian Mathematical Society* A **47** (1989), pp. 133-52

[Kallenberg 02] O. Kallenberg, *Foundations of Modern Probability*, second edition, Springer Verlag 2002.

[Lando 03] S. K. Lando, *Lectures on generating Functions*, Student Mathematical Library, American Mathematical Society 2003.

[Marques et al 89] M. S. Marques, V. Pérez-Abreu, Law of large numbers and central limit theorem for the empirical probability generating function of stationary random sequences and processes. *Caballero, María Emilia (ed.) et al., Symposium on probability and stochastic processes. Proceedings, Guanajuato, Mexico 1988. Mexico City: Sociedad Matemática Mexicana. Aportaciones Mat., Notas Invest.2* **4** (1989) 100–109.

[Nakamura et al 93] M. Nakamura, V. Pérez-Abreu, Empirical probability generating function. An overview. *Insur. Math. Econ. 12* **3** (1993) 349–366.

[Nakamura et al 93a] M. Nakamura, V. Pérez-Abreu, Exploratory data analysis for counts using the empirical probability generating function. *Commun. Stat., Theory Methods 22* **3** (1993) 827-842.

[Nakamura et al 93b] M. Nakamura, V. Pérez-Abreu, Use of an empirical probability generating function for testing a Poisson model. *Can. J. Stat. 21, No.2* **2** (1993) 149–156.

[Rémillard et al 00] B. Rémillard, R. Theodorescu, Inference based on the empirical probability generating function for mixtures of Poisson distributions. *Stat. Decis. 18* **4** (2000).

[Resnick 01] S. I. Resnick, *A Probability Path*, Birkhäuser 2001.

[Rudin 86] W. Rudin, *Real and Complex Analysis*, third edition McGraw-Hill 1986.

[Rueda et al 91] R. Rueda, V. Pérez-Abreu, F. O'Reilly, Goodness of fit for the Poisson distribution based on the probability generating function. *Commun. Stat., Theory Methods 20* **10** (1991) 3093–3110.

[Rueda et al 99] R. Rueda, F. O'Reilly, Tests of fit for discrete distributions based on the probability generating function. *Commun. Stat., Simulation Comput. 28* **1** (1999) 259–274.

[Shiryaev 96] A. N. Shiryaev, *Probability*, second edition, Springer Verlag 1996.

[Zaks & Zygmund] S. Zaks, A. Zygmund, *Analytic Functions*, Monografie Matematyczne Tom XXVIII, Warszawa 1952.

Departamento de Matemática, FCT/UNL, Quinta da Torre, 2829-516 Caparica, Portugal
E-mail address: `mlefct.unl.pt`
URL: `http://ferrari.dmat.fct.unl.pt/personal/mle`