

Independence of Linear Statistics with Random Coefficients and Characterizations of Geometric and Poisson Distributions

Lev B. Klebanov^a

Abstract

There is given a characterization of the geometric distribution by the independence of linear forms with random coefficients. The result is a discrete analog of the corresponding theorem on exponential distribution. The property of linear statistics independence is also a characterization of Poisson law.

Key words: geometric distribution; exponential distribution; Poisson distribution; linear forms; random coefficients.

1 Introduction

Many different characterizations of the distribution are known (see [2]). An essential part of them is connected to the characterizations property of independence of statistics, especially by that of linear forms. The main result here is the linear forms are independent for Gaussian distribution only. The property allows explaining the appearance of Maxwell distribution in Physics. Recently, there were published facts on the independence of linear forms with random coefficients (see [3, 4]). They lead to the characterization of exponential and hyperbolic secant distributions. Unfortunately, it was not clear if it is possible to use the independence property to characterize some discrete distributions. Here we modify previous methods to obtain characterizations

^aDepartment of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic. e-mail: lev.klebanov@mff.cuni.cz

of geometric and Poisson distribution. We hope the methods will lead to characterizations of other discrete distributions. As far as we know, there are only a few characterizations of the distributions of positive integer-valued random variables. Among them let us note the characteristic properties of geometric distribution by the independence of linear forms of order statistics [1]. However, the structure of the forms used here is absolutely different.

2 Main results

In the paper, [4] there was given a characterization of exponential distribution by the independence of linear forms

$$S_1 = (1 - p)aX + \varepsilon(p)aY \quad \text{and} \quad S_2 = pbX + (1 - \varepsilon(p))bY.$$

Here X and Y are independent identically distributed (i.i.d) positive random variables, $p \in (0, 1)$, and $\varepsilon(p)$ is independent of the pair (X, Y) Bernoulli random variable with the parameter p . Below we give a similar result for a geometric random variable. However, the multiplication of X by p or by $1 - p$ leads out the class of integer-valued random variables. Therefore, we have to use a thinning operator instead of multiplication.

Let us give a precise formulation. Suppose X and Y are i.i.d. random variables taking non-negative integer values. Suppose that $p \in (0, 1)$ is fixed, and ε_p is independent of the pair (X, Y) Bernoulli random variable with the parameter p . Let $\{\varepsilon_j(1 - p), j = 1, 2, \dots\}$ be a sequence of i.i.d. Bernoulli random variables with parameter $1 - p$ and independent on $X, Y, \varepsilon(p)$. Define the forms

$$L_1 = \tilde{X}_{1-p} + \varepsilon(p)Y \quad \text{and} \quad L_2 = \tilde{X}_p + (1 - \varepsilon(p))Y. \quad (2.1)$$

Here $\tilde{X}_{1-p} = \sum_{j=1}^X \varepsilon_j(1 - p)$ and $\tilde{X}_p = \sum_{j=1}^X (1 - \varepsilon_j(1 - p))$.

Theorem 2.1. *The forms L_1 and L_2 defined above are stochastical independent if and only if X has a geometric distribution.*

Proof. Consider joint probability generating function of L_1 and L_2 .

$$\begin{aligned} \mathbb{E}(u^{L_1}v^{L_2}) &= \mathbb{E}(u^{\tilde{X}_{1-p} + \varepsilon(p)Y} v^{\tilde{X}_p + (1 - \varepsilon(p))Y}) = \\ &= \mathcal{P}((1 - p)u + pv) \left(p\mathcal{P}(u) + (1 - p)\mathcal{P}(v) \right), \end{aligned} \quad (2.2)$$

Forms L_1 , and L_2 are independent if and only if the right-hand side of (2.2) may be written as a product of two probability generating functions depending on u and v separately.

Let us verify the forms are independent in the case of exponentially distributed X . Really, in this case, we have

$$\mathcal{P}(z) = \frac{a-1}{a-z}$$

for any $a > 1$. Substitution of this function on the right-hand side of (2.2) gives us the product

$$\frac{(a-1)^2}{(a-u)(a-v)}.$$

It means the forms are independent for exponentially distributed X .

Let us prove the inverse statement. Independence of L_1 and L_2 holds if and only if their joint probability generating function is a product of corresponding marginal generating functions:

$$\begin{aligned} & \mathcal{P}((1-p)u + pv) \left(p\mathcal{P}(u) + (1-p)\mathcal{P}(v) \right) = \\ & \mathcal{P}((1-p)u + p) \left(p\mathcal{P}(u) + (1-p) \right) \mathcal{P}((1-p) + pv) \left(p + (1-p)\mathcal{P}(v) \right). \end{aligned} \quad (2.3)$$

The relation (2.3) holds for all u and v such that $|u| \leq 1$, $|v| \leq 1$.

Setting here $u = v = 0$ we obtain

$$\mathcal{P}^2(0) = \mathcal{P}(p) \left(\mathcal{P}(0)p + 1 - p \right) \mathcal{P}(1-p) \left(p + (1-p)\mathcal{P}(0) \right).$$

However, $p \in (0, 1)$ and from the definition of probability generating function we see $\mathcal{P}(p) > 0$ and $\mathcal{P}(1-p) > 0$. Therefore $\mathcal{P}(0) > 0$. Because any probability generating function is analytic in the disc $|z| \leq 1$ on the complex plain $\mathcal{P}(z)$ is uniquely defined by the sequence of its derivatives at point $z = 0$.

Rewrite (2.3) in the form

$$\mathcal{P}((1-p)u + pv) \left(p\mathcal{P}(u) + (1-p)\mathcal{P}(v) \right) = H_1(u)H_2(v). \quad (2.4)$$

Taking the logarithm and differentiating with respect to v both sides of (2.4), we obtain at the point $v = 0$ that

$$p \frac{\mathcal{P}'((1-p)u)}{\mathcal{P}((1-p)u)} + \frac{(1-p)\mathcal{P}'(0)}{p\mathcal{P}(u) + (1-p)\mathcal{P}(0)} = C, \quad (2.5)$$

where $C = H_2'(0)/H_2(0)$. From (2.5) follow two facts:

1. The value $\mathcal{P}'(0)$ may be arbitrary;
2. For any $k > 1$ the value of $\mathcal{P}^k(0)$ is uniquely determined by previous values $\mathcal{P}^{k-1}(0), \mathcal{P}^{k-2}(0), \dots, \mathcal{P}(0)$ that is by two parameters $\mathcal{P}(0)$ and $\mathcal{P}'(0)$.

However, one of these parameters is fixed in view of the condition

$$\sum_{k=0}^{\infty} \frac{\mathcal{P}^k(0)}{k!} = 1$$

and the solution of (2.2) may depend on one parameter only. Such solution is $\mathcal{P}(z) = (a-1)/(a-z)$. \square

Let us now proceed to the characterization of Poisson distribution. Let X, Y are i.i.d. non-negative integer-valued random variables. Suppose that $\{\varepsilon_j(p), j = 1, 2, \dots\}$ and $\{\tilde{\varepsilon}_j(q), j = 1, 2, \dots\}$ are independent with each other and with the pair (X, Y) sequences of Bernoulli random variables with parameters $p, q \in (0, 1)$. Consider linear forms K_1 and K_2 of X and Y :

$$\begin{aligned} \tilde{X}_{1-p} &= \sum_{j=1}^X \varepsilon_j(1-p), & \tilde{X}_q &= \sum_{j=1}^X (1 - \tilde{\varepsilon}_j(1-q)) \\ \tilde{Y}_{1-q} &= \sum_{j=1}^Y \tilde{\varepsilon}_j(1-q), & \tilde{Y}_p &= \sum_{j=1}^Y (1 - \varepsilon_j(1-p)) \\ K_1 &= \tilde{X}_{1-p} + \tilde{Y}_{1-q}, & K_2 &= \tilde{X}_q + \tilde{Y}_p. \end{aligned} \quad (2.6)$$

Theorem 2.2. *Forms (2.6) are independent if and only if X has Poisson distribution.*

Proof. The joint probability generating function of K_1 and K_2 is

$$\begin{aligned} \mathbb{E}(u^{K_1} v^{K_2}) &= \mathbb{E}\left(u^{\tilde{X}_{1-p}} v^{\tilde{Y}_{1-q}} u^{\tilde{X}_q} v^{\tilde{Y}_p}\right) = \\ &= \mathcal{P}((1-p)u + pv) \mathcal{P}((1-q)u + qv). \end{aligned} \quad (2.7)$$

Probability generating function of Poisson distribution has form

$$\mathcal{P}_o(z) = \exp\{\lambda(z-1)\}.$$

Substituting this function on the right-hand side of (2.7) leads to a product of functions depending on u and v separately.

The rest of the proof is similar to that of the Theorem 2.1; we omit it. \square

3 Conclusions

Characterization of distributions by the property of independent statistics is an interesting part of Probability. Basically, it is purely theoretical interest. However, characterizations of distributions can also be useful in statistical hypotheses testing (see [5]). Obtained here characterizations use rather simple statistics and there is a hope they may also be used for the construction of statistical tests.

ACKNOWLEDGEMENT

The work was partially supported by Grant GAČR 19-28231X.

References

- [1] Janos Galambos, Samuel Kotz (1978) *Characterizations of Probability Distributions*, Springer-Verlag, Berlin, Heidelberg, New York.
- [2] A.M. Kagan, Yu.V. Linnik and C.R. Rao (1973) *Characterization problems in mathematical statistics*. John Wiley & Sons, New York, London, Sydney, Toronto.
- [3] Lev B. Klebanov (2019) *Linear Statistics with Random Coefficients and Characterization of Hyperbolic Secant Distribution*, arXiv: 1905.09910v1, 1-6.
- [4] Lev B. Klebanov and Zeev E. Volkovich (2019) *Some more results on characterization of the exponential distribution*, arXiv: 1911.11193v1, 1-11.
- [5] Yakov Nikitin (2017) *Tests based on characterizations, and their efficiencies: a survey*, arXiv:1707.01522, 1-19.