J. Appl. Prob. **17**, 1138–1144 (1980)
Printed in Israel
0021-9002/80/041138-07 $00.95

A CHARACTERIZATION OF THE GAMMA DISTRIBUTION BY THE NEGATIVE BINOMIAL DISTRIBUTION

JAN ENGEL* AND
MYNT ZIJLSTRA,* *N. V. Philips, Eindhoven*

Abstract

It is proved that for a Poisson process $\{N(t);\ t \geqq 0\}$ there exists a one-to-one relation between the distribution of the random variable $N(Y)$ and the distribution of the non-negative random variable Y. This relation is used to characterize the gamma distribution by the negative binomial distribution. Furthermore it is applied to obtain some characterizations of the exponential distribution.

GAMMA DISTRIBUTION; NEGATIVE BINOMIAL DISTRIBUTION; CHARACTERIZATION; POISSON PROCESS; GEOMETRIC DISTRIBUTION; EXPONENTIAL DISTRIBUTION; MOMENT PROBLEM

1. Introduction

During the last decades several characterization results of the gamma distribution have been obtained, in particular in the special case of the exponential distribution. Surveys may be found in Johnson and Kotz (1969), Kotz (1974) and the recent monograph by Galambos and Kotz (1978), in which a unified treatment is given of characterizations dealing with the exponential distribution and its monotonic transformations.

Investigating the following practical problem we recently met a characterization, not mentioned in these surveys. Suppose a factory supplies quantities (e.g. one truckload) of some product to a central warehouse. The lead times of such production quantities constitute a renewal process, while the orders for the product arrive at the warehouse according to a Poisson process. It turned out that the distribution of the number of orders during a lead time is negative binomial if and only if the lead time has a gamma distribution.

We now give a more formal description of the situation we want to study.

Consider a homogeneous Poisson process with parameter $\lambda = 1$ and let $N(Y)$

Received 5 April 1979; revision received 4 September 1979.
* N. V. Philips, ISA-Quantitative Methods, Building HSM-301, 5600 MD Eindhoven, The Netherlands.

be the number of points in the time-interval $[0, Y)$ with Y a non-negative random variable.

The range of $N(Y)$ is $\{0, 1, 2, \cdots\}$. One may say that by means of the Poisson process each probability distribution on $[0, \infty)$ is mapped on a discrete probability distribution defined on $\{0, 1, 2, \cdots\}$. It will be shown that the mapping is one-to-one. We denote the set of all probability distributions on $[0, \infty)$ by C and its image by D.

2. A characterization of the gamma distribution

Proposition 1. Let Y and Z be independent non-negative random variables with distribution functions F and G respectively. Then F is equal to G if $N(Y)$ and $N(Z)$ have the same probability distribution.

Proof. $N(Y)$ and $N(Z)$ equally distributed means

$$\Pr(N(Y) = k) = \Pr(N(Z) = k), \qquad k = 0, 1, 2, \cdots.$$

This is equivalent to

$$\int_0^\infty \frac{y^k e^{-y}}{k!} dF(y) = \int_0^\infty \frac{z^k e^{-z}}{k!} dG(z), \qquad k = 0, 1, 2, \cdots. \tag{2.1}$$

Let f and g be the Laplace–Stieltjes transforms of the probability measures defined by F and G respectively; then we have for f and its kth derivative $f^{(k)}$ (and similarly for g and $g^{(k)}$):

$$\begin{aligned} f(s) &= \int_0^\infty e^{-ys} dF(y), && \operatorname{Re}(s) > 0 \\ f^{(k)}(s) &= \int_0^\infty (-y)^k e^{-ys} dF(y), && \operatorname{Re}(s) > 0. \end{aligned} \tag{2.2}$$

Using (2.2) we observe that (2.1) is equivalent to

$$f^{(k)}(1) \frac{(-1)^k}{k!} = g^{(k)}(1) \frac{(-1)^k}{k!}, \qquad k = 0, 1, 2, \cdots. \tag{2.3}$$

As the functions $f(s)$ and $g(s)$ both are analytical for $\operatorname{Re}(s) > 0$ we have

$$\begin{aligned} f(1+z) &= \sum_{k=0}^\infty \frac{f^{(k)}(1)}{k!} z^k, && |z| < 1 \\ g(1+z) &= \sum_{k=0}^\infty \frac{g^{(k)}(1)}{k!} z^k, && |z| < 1. \end{aligned} \tag{2.4}$$

From (2.3) and (2.4) it follows that $f(1+z) = g(1+z)$, $|z| < 1$. As both functions can be continued analytically, we have $f(s)=g(s)$ for all s. But this implies $F=G$.

Remark. The proposition may be considered as a variant of the well-known moment problem, cf. Feller (1971), Chapter VII, 3 and 6. From the proof it will be clear that the proposition simply states that the distribution of a non-negative random variable Y is uniquely determined by the sequence $E\{Y^k e^{-Y}\}$, $k = 0, 1, 2, \cdots$.

Proposition 2. $N(Y)$ has a negative binomial distribution with parameters r and $\alpha\ (1+\alpha)^{-1}$ if and only if Y is gamma distributed with parameters r and α; $r>0$, $\alpha>0$.

Proof. Let Y have a gamma distribution with density

$$h_{\alpha,r}(y)=\frac{1}{\Gamma(r)}\alpha^r y^{r-1}e^{-\alpha y}, \qquad r>0,\ \alpha>0,\ y\geqq 0.$$

Then we have

$$\begin{aligned}\Pr(N(Y)=k)&=\int_0^\infty \frac{y^k}{k!}e^{-y}h_{\alpha,r}(y)dy\\ &=\frac{\Gamma(k+r)}{\Gamma(r)k!}\left(\frac{\alpha}{1+\alpha}\right)^r\left(\frac{1}{1+\alpha}\right)^k\int_0^\infty\frac{1}{\Gamma(k+r)}(1+\alpha)^{k+r}y^{k+r-1}e^{-(1+\alpha)y}dy\\ &=\binom{k+r-1}{k}\left(\frac{\alpha}{1+\alpha}\right)^r\left(1-\frac{\alpha}{1+\alpha}\right)^k, \qquad k=0,1,2,\cdots.\end{aligned}$$

So the distribution of $N(Y)$ is negative binomial with parameters r and $\alpha(1+\alpha)^{-1}$, cf. Feller (1967), Chapter VI, 8. The only if part of the proposition now follows from Proposition 1.

In the special case $r=1$, i.e. Y is exponentially distributed with parameter α, $N(Y)$ has a geometric distribution with parameter $\alpha(1+\alpha)^{-1}$. If the distribution of Y is Erlang (i.e. $r\geqq 1$ and integer) it is known that $Y=Y_1+\cdots+Y_r$ with $Y_1,\cdots,Y_r$ independent and identically exponentially distributed random variables with parameter α.

Let $N(Y_i)$ be the number of Poisson points in the interval

$$[Y_0+Y_1+\cdots+Y_{i-1},\ Y_0+Y_1+\cdots+Y_i),\ 1\leqq i\leqq r,\ Y_0\equiv 0.$$

From $N(Y)=N(Y_1)+\cdots+N(Y_r)$ it follows just as a consequence of Proposition 2 that the negative binomial distribution with parameters r and $\alpha(1+\alpha)^{-1}$ has to be the convolution of r geometric distributions with parameter $\alpha(1+\alpha)^{-1}$, a well-known property of the negative binomial distribution, cf. Feller (1967), Chapter XI.2.

3. Some additional characterizations of the exponential distribution

We shall discuss a few examples of the way in which Propositions 1 and 2 can be applied to obtain certain characterizations of the exponential distribution.

Example 1. If $\{0,1,2,\cdots\}$ is the range of the non-degenerate random variable U then U is geometrically distributed if and only if, see Johnson and Kotz (1969),

$$\Pr(U \geqq k+m) = \Pr(U \geqq k)\Pr(U \geqq m) \quad \text{for all positive integers } k \text{ and } m. \tag{3.1}$$

It is easily seen that this 'lack of memory' property is equivalent to:
U is geometrically distributed if and only if

$$\Pr(U \geqq k) = [\Pr(U \geqq 1)]^k \qquad \text{for } k = 1,2,\cdots. \tag{3.2}$$

It will be clear that this statement remains valid if we restrict U to a random variable with a distribution in the subset D, i.e. U can be considered as the image $N(Y)$ of a uniquely determined non-negative random variable Y. Replacing (3.2) by the equivalent expression in terms of Y and applying Proposition 1 we obtain a characterization of the distribution in C, which is corresponding to the geometric distribution in D.

Characterization 1. The non-negative and non-degenerate random variable Y is exponentially distributed if and only if

$$\Pr(Y \geqq X_1 + \cdots + X_k) = [\Pr(Y \geqq X_1)]^k \tag{3.3}$$

for $k = 1,2,\cdots$ and for Y, X_1, $X_2,\cdots$ independent with the X_i exponentially distributed with parameter $\lambda = 1$.

Characterization 1 is rather similar to the following variant of the well-known characterization by the lack of memory property: the non-negative and non-degenerate random variable Y is exponentially distributed if and only if

$$\Pr(Y > kx) = [\Pr(Y > x)]^k \qquad \text{for } k = 1,2,\cdots \quad \text{and} \quad x \geqq 0.$$

This characterization may be found in Galambos and Kotz (1978), Section 1.5. The same result, formulated however in a different way, has been obtained by Bosch (1977).

More interesting is a relationship between Characterization 1 and a characterization given by Krishnaji (1971), which may be found also in Galambos and Kotz (1978), p. 26.

Krishnaji states that under very restrictive assumptions concerning the distribution function $F(y)$ of Y this function is exponential if and only if

$$\Pr(Y \geqq W + Z) = \Pr(Y \geqq W)\Pr(Y \geqq Z)$$

for all W from D_1 and all Z from D_2, in which D_1 and D_2 are two families of random variables and Y is independent of each member of both D_1 and D_2.

Replacing (3.3) by the equivalent expression

$$\Pr(Y \geqq X_1 + \cdots + X_k) = \Pr(Y \geqq X_1)\Pr(Y \geqq X_2 + \cdots + X_k) \quad \text{for } k = 2, 3, \cdots \tag{3.4}$$

we see that in Characterization 1 a special choice of D_1 and D_2 is made, namely D_1 is containing the exponentially distributed random variable with parameter 1 and D_2 is containing all Erlang-distributed random variables with parameters $k = 1, 2, \cdots$ and $\lambda = 1$. We observe that due to the special choice in Characterization 1 all very restrictive assumptions about F given in Krishnaji's characterization can be dropped. However, in a sense, a special choice of D_1 and D_2 means more restrictions with respect to D_1 and D_2.

In the following characterization that is closely related to Characterization 1 we observe the same phenomenon for another choice of D_1 and D_2 in the special case $k = 2$.

Characterization 2. The non-negative and non-degenerate random variable Y is exponentially distributed if and only if for Y, X_1, $X_2, \cdots, X_k$ independent and the X_i identically exponentially distributed with parameter λ,

$$\Pr(Y \geqq X_1 + \cdots + X_k) = [\Pr(Y \geqq X_1)]^k \tag{3.5}$$

for at least one $k \geqq 2$ and all λ with $a < \lambda < b$; a and b are non-negative reals.

We prove the 'if' part. Let F be the distribution function of Y, then conditioning on $X_1 + \cdots + X_k$ and X_1 in (3.5) leads to

$$\lambda^k f^{(k-1)}(\lambda)(-1)^{k-1} = (k-1)!\,[\lambda f(\lambda)]^k, \qquad a < \lambda < b, \tag{3.6}$$

in which $f(\lambda)$ is the Laplace transform of $1 - F$. The unique solution of this differential equation is $f(\lambda) = (\alpha + \lambda)^{-1}$, $a < \lambda < b$ and $\alpha > 0$. So for λ with $a < \lambda < b$ the function f is equal to the function $g(\lambda) = (\alpha + \lambda)^{-1}$, which is analytical for $\mathrm{Re}(\lambda) > -\alpha$. Continuing f analytically we get $f(\lambda) = g(\lambda)$ for $\mathrm{Re}(\lambda) > -\alpha$. However, $g(\lambda) = (\alpha + \lambda)^{-1}$ is the Laplace transform of $1 - F(y) = \exp(-\alpha y)$, i.e. Y is exponentially distributed. The 'only if' part of the proof is trivial.

Example 2. Puri (1973) gives the following characterization of the geometric distribution. If U and V are mutually independent random variables with range $\{0, 1, 2, \cdots\}$ then U is geometrically distributed with parameter $\alpha(1 + \alpha)^{-1}$ if and only if

$$\Pr(V+U \geqq k) - \Pr(V \geqq k+1) = \left(\frac{1+\alpha}{\alpha}\right) \Pr(V+U=k) \tag{3.7}$$

holds for $k = 0, 1, 2, \cdots$.

It is obvious that Puri's statement remains valid if we restrict U and V to random variables with distributions in the subset D, i.e. U and V can be considered as the image $N(Y)$ and $N(Z)$ respectively of uniquely determined random variables Y and Z. Replacing (3.7) by the equivalent expression in terms of Y and Z, and applying Proposition 1 we obtain a characterization of the exponential distribution.

Characterization 3. The non-negative and non-degenerate random variable Y is exponentially distributed with parameter α if and only if for $k = 0, 1, 2, \cdots$

$$\begin{aligned} &\Pr(Z+Y \geqq X_1 + \cdots + X_k) - \Pr(Z \geqq X_1 + \cdots + X_{k+1}) \\ &\qquad = \left(\frac{1+\alpha}{\alpha}\right) \Pr(X_1 + \cdots + X_k \leqq Z + Y < X_1 + \cdots + X_{k+1}) \end{aligned} \tag{3.8}$$

for the non-negative and independent random variables Z, Y, X_1, $X_2, \cdots$, in which the X_i are identically exponentially distributed with parameter $\lambda = 1$.

Obviously (3.7) holds if V is degenerate at 0. In that case also Z is degenerate at 0 and (3.8) can be written as

$$\frac{\Pr(X_1 + \cdots + X_k \leqq Y < X_1 + \cdots + X_{k+1})}{\Pr(Y \geqq X_1 + \cdots + X_k)} = \frac{\alpha}{1+\alpha}. \tag{3.9}$$

The characterization by (3.9) resembles the characterization of the exponential distribution by a constant hazard rate, cf. Galambos and Kotz (1978), p. 15. Notice that a characterization by means of the hazard rate requires the existence of a probability density function. This is not needed in Characterization 3.

4. Acknowledgements

We wish to thank our colleague Dr Paul van Beek for drawing our attention to the practical problem mentioned in the introduction.

We are also indebted to the referee for some valuable comments and especially for drawing attention to the recent monograph by Galambos and Kotz.

References

Bosch, K. (1977) Eine Charakterisierung der Exponentialverteilungen. *Z. Angew. Math. Mech.* **57**, 609–610.

FELLER, W. (1967) *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd edn. Wiley, New York.

FELLER, W. (1971) *An Introduction to Probability Theory and its Applications*, Vol. 2, 2nd edn. Wiley, New York.

GALAMBOS, J. AND KOTZ, S. (1978) *Characterizations of Probability Distributions.* Lecture Notes in Mathematics **675**, Springer–Verlag, Berlin.

JOHNSON, N. L. AND KOTZ, S. (1969) *Distributions in Statistics*, Vols, 1–4. Wiley, New York.

KOTZ, S. (1974) Characterizations of statistical distributions: a supplement to recent surveys. *Internat. Statist. Rev.* **42**, 39–65.

KRISHNAJI, N. (1971) Note on a characterizing property of the exponential distribution. *Ann. Math. Statist.* **42**, 361–362.

PURI, P. A. (1973) On a property of exponential and geometric distributions and its relevance to multivariate failure rate. *Sankhyā* **35**, 61–68.