On the independence of integer and fractional parts

Steutel, F.W.; Thiemann, J.G.F.

Published: 01/01/1987

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Memorandum COSOR 87-29

On the independence of integer and fractional parts

by

F.W. Steutel and J.G.F. Thiemann

Eindhoven, Netherlands
October 1987
On the independence of integer and fractional parts

by F.W. Steutel and J.G.F. Thiemann
Eindhoven University of Technology,
Department of Mathematics and Computing Science,
P.O. Box 513,
5600 MB Eindhoven,
The Netherlands.

ABSTRACT

We use the independence of the integer and fractional parts of exponentially distributed random variables to obtain expressions for the order statistics from a geometric distribution. As our main result we show that a strong form of this independence characterizes the exponential distribution.

1. Introduction

It is easily verified that for an exponentially distributed random variable the integer part and the fractional part are independent. We formulate this property as a lemma.

Lemma 1. Let $Y$ have an exponential distribution, let $\lfloor Y \rfloor$ denote the largest integer not exceeding $Y$ (the integer part) and let $\{Y\} := Y - \lfloor Y \rfloor$ (the fractional part). Then

$$\lfloor Y \rfloor \text{ and } \{Y\} \text{ are independent .}$$

It is also easily verified that there is the following (well-known?) relation between exponential and geometric random variables; this relation was pointed out to me long ago by L. de Haan (1978) in the context of problems 56 and 57 in Statistica Neerlandica, where nobody had used it (see Jagers (1978)).
Lemma 2. If $X$ has a geometric distribution, i.e., if
\[ P(X = k) = p (1 - p)^k \quad (k = 0, 1, \ldots; 0 < p < 1), \]
then
\[ X \overset{d}{=} [Y], \]
where $Y$ has an exponential distribution with density
\[ f_Y(y) = \lambda e^{-\lambda y} \quad (y > 0; \lambda = \log(1 - p)). \]

In section 2 of this note we return to order statistics of geometric random variables, and in section 3 we prove that a random variable $Y$ has an exponential distribution if and only if $[cY]$ and $\{cY\}$ are independent for every $c > 0$.

2. Order statistic from the geometric distribution

It is well known and not hard to prove that the order statistics $Y_{1;n} \leq Y_{2;n} \leq \cdots \leq Y_{n;n}$ from an exponential distribution satisfy (cf. Feller (1971), p. 19)
\[ Y_{k;n} \overset{d}{=} \sum_{j=0}^{k-1} \frac{Y_{n-j}}{n-j} \quad (k = 1, 2, \ldots, n), \]
where $Y_1, \ldots, Y_n$ are iid exponential random variables. This is most easily seen as follows: waiting for $Y_{k;n}$ to finish we first wait for $Y_{1;n}$ after which the remaining $Y$'s start anew by the lack-of-memory property, so
\[ Y_{k;n} \overset{d}{=} \sum_{j=0}^{k-1} Y_{1;n-j}, \]
with the $Y_{1;n-j}$ independent and (cf. (1)) $Y_{1;n-j} = Y_{n-j} / (n - j)$.

Since the geometric distribution has a similar lack-of-memory property:
\[ P(X \geq k + m \mid X \geq m) = P(X \geq k), \]
it is tempting to conjecture that a relation similar to (2) would hold for geometric order statistics. It is easily seen, however, that this is not so. Instead of (2) we have the following result.

Theorem 1. Let $X_{k;n}$ ($k = 1, 2, \ldots, n$) denote the order statistics from a geometric distribution on $\{0,1,2,\ldots\}$. Then
where \( X_j \) and \( Y_j \) are related by \( X_j \overset{d}{=} [Y_j] \) as in Lemma 2, and all variables in the right-hand side of (3) are independent.

**Proof.** Using Lemmas 1 and 2, and the fact that \( X_{1:m} \overset{d}{=} [Y_{1/m}] \), one easily has

\[
X_{k;n} \overset{d}{=} [Y]_{k;n} = [Y_{k;n}] = \left[ \sum_{j=0}^{k-1} \frac{Y_{n-j}}{n-j} \right] = \left[ \sum_{j=0}^{k-1} \left\{ \frac{Y_{n-j}}{n-j} \right\} \right] = \sum_{j=0}^{k-1} \frac{Y_{n-j}}{n-j} + \sum_{j=0}^{k-1} \left\{ \frac{Y_{n-j}}{n-j} \right\} = \sum_{j=0}^{k-1} \frac{Y_{n-j}}{n-j} + \sum_{j=0}^{k-1} \left\{ \frac{Y_{n-j}}{n-j} \right\},
\]

where all random variables in the last expression are independent.

**Corollary.** For the probability generating function of \( X_{k;n} \) one has

\[
P_{X_{k;n}}(z) = \prod_{j=0}^{k-1} \frac{1-(1-p)^{n-j}}{1-(1-p)^{n-j}z} \cdot Q_{n,k}(z),
\]

where \( Q_{n,k} \) is a probability generating function and a polynomial of degree \( k - 1 \).

**Remark 1.** For the order statistics of exponentially distributed random variables mentioned earlier we have for the Laplace transform

\[
\phi_{Y_{k;n}}(s) = \prod_{j=0}^{k-1} \frac{\lambda(n-j)}{\lambda(n-j)+s};
\]

here the extra factor occurring in (4) is missing.

**Remark 2.** A special case of this Corollary occurs in Jagers (1979). The explicit form of \( Q_{n,k} \) seems intractable.

**Remark 3.** From Theorem 1 it follows that

\[
\sum_{j=0}^{k-1} X_{1;n-j} \leq X_{k;n} \leq \sum_{j=0}^{k-1} X_{1;n-j} + n - 1,
\]

where the inequalities should be read in distribution (i.e. in the sense of stochastic ordering); the inequality on the right in (5) holds, in fact, with probability one on the probability space defined by an infinite sequence of independent binomial trials.
3. A characterization of the exponential distribution

Let $Y$ be a random variable with distribution function $F$. For $[Y]$ and $\{Y\}$ (cf. Lemma 1) to be independent it is necessary and sufficient that $(F(n + y) - F((n)))(F(n + 1) - F(n))$ is independent of $n \in \mathbb{N}$ for $y \in (0, 1]$. There are many $F$'s with this property, but if $F$ is required to have a continuous density $f$ then it follows that $f$ must satisfy

$$f(n + v) = \alpha^n f(v),$$

for some $\alpha$, i.e. that $f$ is "globally exponential".

The object of this section is to prove the following theorem.

**Theorem 2.** If a random variable $Y$ is non-constant and such that $[eY]$ and $\{eY\}$ are independent for every $c > 0$, then $Y$ has an exponential distribution on $(0, \infty)$ or on $(-\infty, 0)$.

We first prove three lemmas. The first lemma is a direct consequence of the independence of $[cY]$ and $\{cY\}$.

**Lemma 3.** Let $Y$ satisfy the conditions of Theorem 2 and let $k \in \mathbb{Z}$, $a, b \in [0, 1]$, $c > 0$ be such that

$$P(kc \leq Y < (k + 1)c) > 0 \quad \text{and} \quad P((k + a)c < Y < (k + b)c) = 0.$$  

Then, for each $l \in \mathbb{Z}$,

$$P((l + a)c < Y < (l + b)c) = 0.$$  

**Proof:** $P([Y/c] = k) = P(kc \leq Y < (k + 1)c) > 0$, so

$$P((l + a)c < Y < (l + b)c) = P([Y/c] = l) P(a < \{Y/c\} < b) =$$

$$= P([Y/c] = l) P([Y/c] = k) P((k + a)c < Y < (k + b)c) = 0. \quad \Box$$

**Lemma 4.** Let $Y$ satisfy the conditions of Theorem 2. Then

i) \quad $P(Y < 0) = 1$ or $P(Y \geq 0) = 1$,

ii) \quad $P(|Y| < \gamma) > 0$ and $P(|Y| > \gamma) > 0$ for each $\gamma > 0$.  

Proof.

i) Suppose $P(Y < 0) < 1$ and $P(Y \geq 0) < 1$. Then $P(Y \geq 0) > 0$ and $P(Y < 0) > 0$ and therefore, for $\alpha$ large enough, we have

$$P(0 \leq Y < \alpha) > \frac{1}{2} P(0 \leq Y) \text{ and } P(-\alpha \leq Y < 0) > \frac{1}{2} P(Y < 0).$$

From this it follows that

$$P(\left\lfloor \frac{Y}{2\alpha} \right\rfloor < \frac{1}{2} \mid \left\lfloor \frac{Y}{2\alpha} \right\rfloor = -1) = \frac{P(-2\alpha \leq Y < -\alpha)}{P(-2\alpha \leq Y < 0)} = 1 - \frac{P(\alpha \leq Y < 0)}{P(Y < 0)} < \frac{1}{2} P(Y < 0) = \frac{1}{2}.$$

The inequality of these conditional probabilities contradicts the independence of $\left\lfloor \frac{Y}{2\alpha} \right\rfloor$ and $\left\lceil \frac{Y}{2\alpha} \right\rceil$, and so (i) holds.

ii) We consider the case $P(Y \geq 0) = 1$; the case $P(Y < 0) = 1$ can be handled in a similar way. Again, we argue by contradiction.

First suppose that $P(Y < \gamma) = 0$ for some $\gamma > 0$. Then

$$\delta := \sup \{\gamma > 0 \mid P(0 \leq Y < \gamma) = 0\} < \infty \quad \text{and} \quad P(0 \leq Y < \delta) = 0.$$

Also, for each $\alpha > \delta$, we have $P(0 \leq Y < \alpha) > 0$ and hence, by Lemma 3, $P(\alpha < Y < \alpha + \delta) = 0$.

So, for all $\alpha > \delta$ we have $P(\alpha < Y < \alpha + \delta) = 0$, and therefore $P(Y > \delta) = 0$. As we also have $P(0 \leq Y < \delta) = 0$, we conclude that $P(Y = \delta) = 1$, i.e. $Y$ is constant. This contradiction proves that $P(Y < \gamma) > 0$ for every $\gamma > 0$.

Next suppose that $P(Y > \gamma) = 0$ for some $\gamma > 0$. Then $\delta := \inf \{\gamma > 0 \mid P(Y > \gamma) = 0\} > 0$, since $\delta = 0$ would imply $P(Y = 0) = 1$, i.e. $Y$ is constant. Moreover $P(Y > \delta) = 0$.

Now let $\alpha \in \left(\frac{5}{6}\delta, \delta\right)$. Then, by the definition of $\delta$,

$$P(\delta < Y < 2\alpha) = 0 \quad \text{and} \quad P(\alpha \leq Y < 2\alpha) > 0.$$

So, by Lemma 3,

$$P(\delta - \alpha < Y < \alpha) = 0.$$

Now suppose that $P(0 \leq Y \leq \delta - \alpha) > 0$. Then, since $\delta - \alpha < \frac{2}{3}\delta < \alpha$, we have $P(0 \leq Y < \frac{2}{3}\delta) > 0$ and, by (7), $P(\delta - \alpha < Y < \frac{2}{3}\delta) = 0$, hence, by Lemma 3, $P(\frac{5}{3}\delta - \alpha < Y < \frac{4}{3}\delta) = 0$. The latter implies $P(\frac{5}{6}\delta < Y \leq \delta) = 0$, which contradicts the
All this implies \( P(0 \leq Y < \alpha) = 0 \) for each \( \alpha \in \left(\frac{5}{6}, \delta\right) \), so \( P(0 \leq Y < \delta) = 0 \).

Since also \( P(\delta < Y) = 0 \), we have \( P(Y = \delta) = 1 \), i.e. \( Y \) is constant. So, again we obtain a contradiction, and hence \( P(Y > \gamma) > 0 \) for every \( \gamma > 0 \).

Lemma 5. Let \( Y \) satisfy the conditions of Theorem 2. then \( [Y] \) has a geometric distribution either on \( \{0, 1, 2, \ldots \} \) or on \( \{-1, -2, -3, \ldots \} \).

Proof. We first consider the case \( P(Y \geq 0) = 1 \) (cf. Lemma 4).

Let

\[
p_i := P([Y] = i) \quad (i = 0, 1, 2, \ldots )
\]

Then, for each \( i, j \in \mathbb{N} \) with \( i < j \), we have

\[
p_i p_j = P(i \leq Y < i+1) P(j \leq Y < j+1) = P([\frac{Y}{j}] = 0) P(\frac{i}{j} \leq \frac{Y}{j} < \frac{i+1}{j}) P([\frac{Y}{j}] = 1) P(0 \leq \frac{Y}{j} < \frac{1}{j}) =
\]

\[
P(0 \leq Y < 1) P(i+j \leq Y < i+j+1) = p_0 p_{i+j}.
\]

Now \( p_0 > 0 \) by Lemma 4 ii), So, taking \( i = 1 \), we get

\[
\frac{p_1}{p_0} \cdot p_j = p_{j+1} \quad (j = 2, 3, \ldots ) ;
\]

hence

\[
p_{j+k} = \left(\frac{p_1}{p_0}\right)^k p_j \quad (j = 2, 3, \ldots , k = 0, 1, 2, \ldots ) .
\]

Also by Lemma 4 ii) we have \( P(Y \geq 3) > 0 \).

So \( 0 < \sum_{l=3}^{\infty} p_l = \sum_{k=0}^{\infty} \left(\frac{p_1}{p_0}\right)^k p_3 \), which implies \( p_3 \neq 0 \).

The foregoing now gives, for \( l = 2, 3, 4, \ldots \),

\[
p_2 p_l = p_3 \left(\frac{p_1}{p_0}\right)^{l-2} p_2 = \left(\frac{p_1}{p_0}\right)^{l-2} p_3 p_0 = \left(\frac{p_1}{p_0}\right)^l p_0
\]

so \( p_l = \left(\frac{p_1}{p_0}\right)^l p_0 \). As the last equality holds also for \( l = 0, 1 \), the random variable \( [Y] \) is geometrically distributed on \( \{0, 1, 2, \ldots \} \).
For the case \( P(Y < 0) = 1 \) we define
\[
q_i := P([Y] = -1 - i) \quad (i = 0, 1, 2, \cdots)
\]
and prove the relations \( q_i q_j = q_{i+j} q_0 \) \((0 < i < j)\) by considering the random variable \( Y/j \).
The rest of the proof is similar to the proof given above.

Proof of Theorem 2
We consider the case \( P(Y \geq 0) = 1; \) the case \( P(Y < 0) \) can be treated in the same way.

Let \( n \in \mathbb{N} \). Then the random variable \( nY \) satisfies the conditions of Theorem 2 and therefore \([nY]\) is geometrically distributed on \([0, 1, 2, \cdots]\), by Lemma 5.

So \( p \in (0, 1) \) exists such that, for each \( m \in \{0, 1, 2, \cdots\} \), we have
\[
p^m = P([nY] \geq m) = P(nY \geq m) = P(Y \geq \frac{m}{n})
\]
and, in particular, \( p^n = P(Y \geq 1) \); hence
\[
P(Y \geq \frac{m}{n}) = P(Y \geq 1)^{\frac{m}{n}}.
\]

So \( P(Y \geq r) = P(Y \geq 1)^r \) for each rational \( r \in [0, \infty) \). Since both members of the equality are non-increasing functions of \( r \) on \([0, \infty)\), the equality holds for each \( r \in [0, \infty) \). Hence \( Y \) is exponentially distributed on \([0, \infty)\).

References


de Haan, L. (1978), Personal communication.


Eindhoven University of Technology,
Department of Mathematics and Computing Science,
P.O. Box 513,
5600 MB Eindhoven,
The Netherlands.