

RECORD RANGES FOR SAMPLES FROM ASYMMETRICAL LAPLACE DISTRIBUTIONS

I.V. BELKOV

*Department of Mathematics and Mechanics, St. Petersburg State University,
St. Petersburg, Russia*

M. AHSANULLAH

*Department of Management Sciences, Rider University, Lawrenceville, New Jersey,
United States of America*

V. B. NEVZOROV

*Department of Mathematics and Mechanics, St. Petersburg State University,
St. Petersburg, Russia*

Received

Abstract

The representations of record ranges via sums of independent identically distributed exponential random variables are obtained for asymmetrical Laplace distributions. This result generalizes the corresponding relations for record values in the cases of exponential and negative exponential distributions

Keywords and phrases: Record values; Exponential distribution; Negative exponential distribution; Laplace distribution; Sample ranges

Introduction

Let X_1, X_2, \dots be a sequence of independent identically distributed random variables (r.v.'s) with an absolutely continuous distribution function (d.f.) $F(x)$. For any $n=1,2,\dots$ let us introduce r.v.'s

$$m_n = \min \{ X_1, X_2, \dots, X_n \}, \quad M_n = \max \{ X_1, X_2, \dots, X_n \} \quad (1)$$

and the corresponding sample ranges

$$W_n = M_n - m_n \quad (2)$$

Upper record times $L(n)$ and upper record values $X(n)$, $n=1,2,\dots$, are defined as follows:

$$L(1)=1, \quad L(n) = \min \{ j > L(n-1) : X_j > X_{L(n-1)} \}, \quad n=2,3,\dots,$$

and

$$X(n) = X_{L(n)}, \quad n=1,2,\dots \quad (3)$$

Analogously one can define the so-called lower record times $l(n)$ and lower record values $x(n)$, $n=1,2,\dots$:

$$l(1)=1, l(n)=\min \{j > l(n-1): X_j < X_{l(n-1)}\}, n=2,3,\dots,$$

and

$$x(n)=X_{l(n)}, n=1,2,\dots \tag{4}$$

Below the notation $X \stackrel{d}{=} Y$ means the equality in distribution of random variables or vectors X and Y .

In the record theory (see, for example, [1] - [5]) results for the initial sequences of exponentially distributed random variables are very important and popular. We give here two of the corresponding statements.

Let Z_1, Z_2, \dots be a sequence of independent $E(1)$ -distributed r.v.'s with d.f.

$$H_1(x)=\max \{0, 1-\exp(-x)\}$$

and $Z(1) = Z_1 < Z(2) < \dots$ be the corresponding upper record values. Let ξ_1, ξ_2, \dots also be independent $E(1)$ -exponential r.v.'s and $S_k = \xi_1 + \xi_2 + \dots + \xi_k, k=1,2,\dots$. The following result is valid.

- **Representation 1.** For any $n=1,2,\dots$

$$(Z(1), Z(2), \dots, Z(n)) \stackrel{d}{=} (S_1, S_2, \dots, S_n). \tag{5}$$

Now let us consider r.v.'s $V_k = -Z_k, k=1,2,\dots$, with the negative exponential d.f. $H_2(x)$, where

$$H_2(x) = 1 - H_1(-x) = \min \{e^x, 1\},$$

and the corresponding lower record values $v(1) = V_1 > v(2) > \dots$.

In this situation the relation, analogous to (5), can be written.

- **Representation 2.** For any $n=1,2,\dots$ the following equality holds:

$$(v(1), v(2), \dots, v(n)) \stackrel{d}{=} (-S_1, -S_2, \dots, -S_n). \tag{6}$$

Equalities (5) and (6) together with Smirnov's transformation allow to obtain some useful results for X 's with any continuous d.f. $F(x)$. For example, one can immediately write that

$$F(X(k)) \stackrel{d}{=} 1 - \exp(-Z(k)) \stackrel{d}{=} 1 - \exp(-S_k) = 1 - \exp(-(\xi_1 + \xi_2 + \dots + \xi_k)) \stackrel{d}{=} 1 - (1-U_1)(1-U_2)\dots(1-U_k) \stackrel{d}{=} 1 - U_1 U_2 \dots U_k, k=1,2,\dots, \tag{7}$$

where U_1, U_2, \dots are independent uniformly $U([0,1])$ -distributed random variables. Hence in this case the following relations are valid for record values:

$$(X(1), X(2), \dots, X(n)) \stackrel{d}{=} (G(1-U_1), G(1-U_1 U_2), \dots, G(1-U_1 U_2 \dots U_n)), n=1,2,\dots, \tag{8}$$

where $G(x)$ is the inverse function for d.f. $F(x)$.

Analogously,

$$F(x(k)) = \exp(v(k)) = \exp(-(\xi_1 + \xi_2 + \dots + \xi_k)) = (1-U_1)(1-U_2)\dots(1-U_k) = U_1U_2\dots U_k \quad (9)$$

and

$$(x(1), x(2), \dots, x(n)) = (G(U_1), G(U_1U_2), \dots, G(U_1U_2\dots U_n)), n=1,2,\dots$$

The aim of the given paper is to present some new relations which generalize presentations 1 and 2.

Record Ranges and Laplace Distributions

Above the exponential and negative exponential random variables with d.f.'s $H_1(x) = \max\{0, 1 - \exp(-x)\}$ and $H_2(x) = \min\{e^x, 1\}$ were under our consideration. Let us consider now the distribution which is the mixture with weights q and $p = (1-q)$, $0 \leq q \leq 1$, of these two distributions. The corresponding d.f. $H(x)$ is taken in this case as $H(x) = qH_1(x) + pH_2(x)$. Thus,

$$H(x) = pe^x, \text{ if } x \leq 0,$$

and

$$H(x) = 1 - qe^{-x}, \text{ if } x > 0 \quad (10)$$

The probability density function (pdf) $h(x)$ is given as

$$h(x) = pe^x, \text{ if } x \leq 0, \text{ and } h(x) = qe^{-x}, x > 0 \quad (11)$$

If $p = 0$, one gets $E(1)$ -exponential distribution. The value $p = 1$ corresponds to the negative exponential distribution. If $p = q = 1/2$ one deals with the classical Laplace pdf $h(x) = \exp(-|x|)/2$. Thus, we can say that relations (10) and (11) present asymmetrical Laplace distributions.

Let now independent r.v.'s X_1, X_2, \dots have the common distribution function (10) and let us add the degenerate r.v. $X_0 = 0$ in the beginning of this sequence. For X_0, X_1, X_2, \dots we consider maximal and minimal values

$$M_n = \max\{X_0, X_1, \dots, X_n\} \text{ and } m_n = \min\{X_0, X_1, \dots, X_n\}, n=1,2,\dots,$$

as well as ranges

$$W_n = M_n - m_n, n=1,2,\dots$$

Record values $W(1) = W_1 < W(2) < \dots < W(n) < \dots$ in the sequence of ranges W_1, W_2, \dots are the subjects of our interest. It appears that one can express these records via r.v.'s ξ_1, ξ_2, \dots , which are defined above. The following result is valid.

- **Representation 3.** For any $n=1,2,\dots$ and $0 \leq p = 1-q \leq 1$ the relation

$$(W(1), W(2), \dots, W(n)) \stackrel{d}{=} (S_1, S_2, \dots, S_n) \quad (12)$$

holds, where $S_k = \xi_1 + \xi_2 + \dots + \xi_k$, $k=1,2,\dots$

1. **Remark 1.** It appears that the RHS of (12) doesn't depend on p and q .
2. **Remark 2.** It is easy to see that Representation 1 is the partial (under $p = 0$) case of (12). Analogously, the result of Representation 2 immediately follows from (12) if to take $p=1$ and $q=0$.
3. **Remark 3.** We discuss here record ranges in the sequences of random variables which have different forms of Laplace distributions, which are rather close in some sense to the exponential distributions. Let us note that some analogous results for record ranges were obtained in [6] for the initial sequences of the uniformly distributed r.v.'s.

Proof of Representation 3

It is evidently that $W(1) = W_1 = \max\{0, X_1\} - \min\{0, X_1\} = |X_1|$ has the exponential $E(1)$ -distribution. Hence one can write that

$$W(1) \stackrel{d}{=} \xi_1. \quad (13)$$

Let us find now distributions of differences

$$T(n) = W(n) - W(n-1), \quad n=2,3,\dots$$

The following notation is used below:

$$M(n) = M_{L(n)} \text{ and } m(n) = m_{L(n)}, \quad n=1,2,\dots,$$

where

$$1 = L(1) < L(2) < \dots < L(n-1) < L(n) < \dots$$

represent the corresponding record moments.

We examine the behavior of conditional probabilities

$$R(x, y, v) = P\{T(n) > x \mid m(n-1) = -y, M(n-1) = v\}, \quad x > 0, y \geq 0, v \geq 0.$$

Note that the condition $\{m(n-1) = -y, M(n-1) = v\}$ corresponds to the situation, when $W(n-1) = y+v$.

Let us denote events

$$\{-y \leq \min\{X_{L(n-1)+1}, \dots, X_{L(n-1)+k}\} < \max\{X_{L(n-1)+1}, \dots, X_{L(n-1)+k}\} \leq v\}$$

as $A(y, v, k, n)$.

One can see that

$$P(A(y, v, k, n)) = (H(v) - H(-y))^k = (1 - qe^{-v} - pe^{-y})^k, \quad n=2,3, \dots, k = 0,1,2, \dots$$

Now it is possible to write that

$$R(x, y, v) = \sum_{k=0}^{\infty} P(A(y, v, k, n)) (P\{X_{L(n-1)+k+1} < -y-x\} + P\{X_{L(n-1)+k+1} > v+x\}) =$$

$$\sum_{k=0}^{\infty} (1 - qe^{-v} - pe^{-y})^k (H(-y-x) + 1 - H(v+x)) = (pe^{-(y+x)} + qe^{-(v+x)}) / (qe^{-v} + pe^{-y}) = e^{-x},$$

and

$$P\{T(n) < x \mid m(n-1) = -y, M(n-1) = v\} = 1 - e^{-x} = P\{\xi_n < x\}$$

for any $x > 0$.

It appears that conditional probabilities $R(x, y, v)$ don't depend on y and v . Thus, it follows that r.v. $T(n)$ doesn't depend on $W(n-1)$ and this difference of two neighbouring record values $W(n) - W(n-1)$ has the standard $E(1)$ -exponential distribution. Recalling relation (13) one can write now that for any record value $W(n)$, $n=1,2, \dots$, the following presentation is valid:

$$W(n) = \overset{d}{\xi_1 + \xi_2 + \dots + \xi_n}.$$

Moreover, taking into account that r.v.'s $W(1), W(2)-W(1), W(3)-W(2), \dots$ are independent we get finally that representation 3 holds for any $n=1,2, \dots$

The Number of Record Ranges

Let Z_1, Z_2, \dots be a sequence of independent $E(1)$ -distributed r.v.'s with d.f. $H_1(x) = \max\{0, 1 - \exp(-x)\}$ and $N_1(n)$ be the number of the upper records among Z_1, Z_2, \dots, Z_n .

It is known that

$$EN_1(n) = \sum_{k=1}^n \frac{1}{k} \sim \ln n, \quad n \rightarrow \infty.$$

If independent r.v.'s V_1, V_2, \dots have d.f. $H_2(x) = \min(1, e^{-x})$ and $N_2(n)$ is the number of the lower record values among V_1, V_2, \dots, V_n , then also we come to the relation

$$EN_2(n) = \sum_{k=1}^n \frac{1}{k} \sim \ln n, n \rightarrow \infty.$$

Let us consider now the number $N(n)$ of record values in the set of ranges W_1, W_2, \dots, W_n . We denote

$$Y_{k,1} = \max\{0, X_k\} \text{ and } Y_{k,2} = \min\{0, X_k\}, k=1,2,\dots$$

It is evident that

$$P\{Y_{k,1}=0\} = p, \quad P\{Y_{k,1} < x\} = 1 - q \exp(-x), x > 0,$$

and

$$P\{Y_{k,2} < x\} = p e^x, x < 0, \quad P\{Y_{k,2}=0\} = 1 - p = q.$$

The number $N(n)$ of records in the set W_1, W_2, \dots, W_n is equal to the sum $N^{(1)}(n) + N^{(2)}(n)$, where

$$N^{(1)}(n) = \eta_1 + \eta_2 + \dots + \eta_n$$

is the number of the positive upper records among r.v.'s $Y_{1,1}, Y_{2,1}, \dots, Y_{n,1}$ and

$$N^{(2)}(n) = \gamma_1 + \gamma_2 + \dots + \gamma_n$$

is the number of the negative lower records among r.v.'s $Y_{1,2}, Y_{2,2}, \dots, Y_{n,2}$. One can find that

$$P\{\eta_1=1\} = P\{Y_{1,1} > 0\} = 1 - P\{Y_{1,1}=0\} = 1 - p,$$

and

$$P\{\eta_k=1\} = 1 - P\{\eta_k=0\} = P\{Y_{k,1} > \max\{0, Y_{1,1}, Y_{2,1}, \dots, Y_{k-1,1}\}\} = (1 - p^k)/k, \quad k=2,3,\dots$$

Analogously,

$$P\{\gamma_k=1\} = 1 - P\{\gamma_k=0\} = (1 - q^k)/k, \quad k=1,2,\dots,n.$$

Thus, we get that

$$EN(n) = \sum_{k=1}^n (2 - p^k - q^k)/k, \quad n=1,2,\dots$$

The maximal value of $EN(n)$ is reached if $p=q=1/2$.

One can see also that

$$EN(n) = E(N_1(n) + N_2(n)) - \sum_{k=1}^n (p^k + q^k)/k \sim -(\ln(1-p) + \ln(1-q)) = -\ln(pq), \quad n \rightarrow \infty, \quad 0 < p=1-q < 1.$$

Acknowledgements

The authors thank the referees for valuable comments that improved the presentation of the paper. The work of the third author was partially funded by RFBR grant № 18-01-00393

References

- [1] M. Ahsanullah, *Record Statistics* (Nova Science Publishers Inc., New York, NY, 1995).
- [2] M. Ahsanullah and V.B. Nevzorov, *Ordered Random Variables* (Nova Science Publishers Inc., New York, NY, 2001).
- [3] V.B. Nevzorov, *Records: Mathematical Theory* (American Math. Soc., Providence, Rhode Island, 2001).
- [4] M. Ahsanullah and V.B. Nevzorov. Record Statistics, in *International Encyclopedia of Statistical Science*, part 18, (2011) pp. 1195-1202.

- [5] Ahsanullah M. and Nevzorov V.B. *Records via Probability Theory*. (Atlantis Press, Paris, France, 2015).
- [6] Ahsanullah M. (2005). Record Range of Uniform Distribution. *J. of the Iranian Statistical Soc.*, v.4, N1, 21-34.