Ann. Inst. Statist. Math.
Vol. 50, No. 3, 543–565 (1998)

DISTRIBUTIONAL PROPERTIES OF EXCEEDANCE STATISTICS

JACEK WESOŁOWSKI[1] AND MOHAMMAD AHSANULLAH[2]

[1] *Mathematical Institute, Warsaw University of Technology, Plac Politechniki 1, 00-661 Warsaw, Poland*
[2] *Department of Management Sciences, Rider University, Lawrenceville, NJ 08648-3099, U.S.A.*

(Received October 21, 1996; revised May 26, 1997)

Abstract. Behaviour of a sequence of independent identically distributed random variables with respect to a random threshold is investigated. Three statistics connected with exceeding the threshold are introduced, their exact and asymptotic distributions are derived. Also distribution-free properties, leading to some common and some new discrete distributions, are considered. Identification of equidistribution of observations and the threshold are discussed. In this context relations between the exponential and gamma distributions are studied and a new derivation of the celebrated Laplace expansion for the standard normal distribution function is given.

Key words and phrases: Exceedance statistics, random threshold, approximation, discrete distribution, distribution-free properties, characterization, standard normal distribution function.

1. Introduction

For a random level we are interested in properties of a sequence of observations connected with exceedance of that level. A special kind of such problems for the level being an order statistic from the original sample and for the sequence being a new sample from the same distribution, closely related to the theory of tolerance limits, were widely investigated in fifties, see for instance: Gumbel and Schelling (1950), Epstein (1954), Sarkadi (1957) or Wilks (1959), and even earlier in Wilks (1941, 1942). The chapter "The distribution of exceedance" in the Gumbel (1958) monograph is the basic source in this area. Another aspect of exceedance in the same model, connected with order statistics from extended samples was considered in Siddiqui (1970). For more recent surveys on that approach see David (1981) and Johnson *et al.* (1992).

Here we are concerned with a general model assuming only independence of the random threshold and the sequence of observations, which are independent and have the same distribution. Three basic exceedance statistics are investigated: number of observations in a finite sample not exceeding the level, number of observations not exceeding the level counted to moments of consecutive exceedances and number of records of the sequence of observations not exceeding the threshold. Exact distributions are derived in Section 2. They have forms of generalized binomial, negative binomial and Poisson mixtures. Asymptotic results given in Section 3, due to their generality, allow to obtain approximations for a wide family of discrete distributions. Section 4 is devoted to the distribution-free properties. Except of some common discrete distributions appearing here, also three new discrete laws, named record exceedance distribution of the first and second kind and record-order exceedance distribution, are derived in a very natural way. Also asymptotic properties follow easily from the results of Section 3. On the basis of considerations of Section 4 it is shown in Section 5 that some special forms of the distribution of some exceedance statistics imply equidistribution, while others do not. This can give hints for possible applications of the exceedance statistics in hypothesis testing. Also some remarks on general identifiabiliy properties are given in the final part of that section. The exponential model investigated in Section 6 comes as an example of potential developments of our approach. Here new characterizations of the gamma and exponential laws are given and also a new derivation of the celebrated Laplace expansion of the standard normal distribution function is presented.

Many special cases of exceedance properties are discussed at numerous places scattered throughout the literature. However no such general approach is known to the authors. Hence a simplicity of the model and a considerable level of universality of the results (especially Sections 2 and 3) seem to promise wide applicability in many different fields as mixture models, approximation of distributions, simulations, discrete distributions, characterizations as well as estimation or hypothesis testing. Also, it is obvious that many natural complements and extensions, are to be considered, including also new exceedance statistics, for example along the lines of Aki and Hirano (1994) or Mohanty (1994) (waiting time for consequtive k exceedances, Markov dependent observations).

2. Definitions and exact distributions

Consider a random variable (rv) X with a distribution function (df) F and a sequence $\boldsymbol{Y} = (Y_i)_{i\geq 1}$ of independent identical distributed (iid) rv's with a common df G, independent of X, and denote $\bar{G} = 1 - G$. We are interested in different characteristics describing behaviour of the sequence $\boldsymbol{Y}$ with respect to crossing the random threshold X.

First consider number of observations S_n in the sample $(Y_1, \ldots, Y_n)$ which do not exceed the level X. More precisely

Definition 1. For any integer $n \geq 1$

$$S_n = \#\{j \leq n : Y_j \leq X\}$$

denotes the number of Y's falling below the threshold X among the first n Y's in the sequence.

Observe that S_n can be defined using order statistics by

$$S_n = \max\{j \leq n : Y_{j:n} \leq X\}.$$

The exact distribution of S_n, its mean and variance are given in the following

THEOREM 1. *For any integer $n \geq 1$*

$$P(S_n = j) = \binom{n}{j} E(G^j(X)\bar{G}^{n-j}(X)), \qquad j = 0, 1, \ldots, n,$$

and $E(S_n) = nE(G(X))$,

$$\operatorname{Var}(S_n) = nE(G(X)\bar{G}(X)) + n^2 \operatorname{Var}(G(X)).$$

PROOF. Denote

$$A_{j,\epsilon_j} = \begin{cases} \{Y_j \leq X\} & \text{if} \quad \epsilon_j = 1, \\ \{Y_j > X\} & \text{if} \quad \epsilon_j = 0, \end{cases} \qquad j = 1, \ldots, n.$$

Consequently for any $n \geq 1$ and $j = 0, 1, \ldots, n$

$$P(S_n = j) = P\left(\bigcup_{\sum_{k=1}^n \epsilon_k = j} \bigcap_{k=1}^n A_{k,\epsilon_k} \right) = \sum_{\sum_{k=1}^n \epsilon_k = j} P\left(\bigcap_{k=1}^n A_{k,\epsilon_k} \right).$$

Since each member of the sum given above has the same value then for any vector $(\epsilon_1, \ldots, \epsilon_n)$ such that $\sum_{k=1}^n \epsilon_k = j$ we have

$$P\left(\bigcap_{k=1}^n A_{k,\epsilon_k} \right) = P\left(\bigcap_{k=1}^j A_{k,1} \cap \bigcap_{k=j+1}^n A_{k,0} \right).$$

Then the total probability rule together with independence of $\boldsymbol{Y}$ and X implies

$$P(S_n = j) = \binom{n}{j} \int_{\boldsymbol{R}} P(Y_1 \leq x, \ldots, Y_j \leq x, Y_{j+1} > x, \ldots, Y_n > x) dF(x).$$

Now, since Y's are iid it follows finally that

$$P(S_n = j) = \binom{n}{j} \int_{\boldsymbol{R}} G^j(x)\bar{G}^{n-j}(x) dF(x)$$

which proves the first part.

Since

$$E(S_n) = \sum_{j=0}^n j \binom{n}{j} E\left(G^j(X)\bar{G}^{n-j}(X)\right) = E\left(\sum_{j=0}^n j \binom{n}{j} G^j(X)(1 - G(X))^{n-j} \right)$$

then by the well known formula for the binomial distribution we have $E(S_n) = nE(G(X))$. Similarly for the second moment one has

$$E(S_n^2) = E\left(\sum_{j=0}^{n} j^2 \binom{n}{j} G^j(X)(1-G(X))^{n-j}\right) = n(n-1)E(G^2(X)) + nE(G(X)).$$

Consequently the formula for the variance follows. □

Second, we are interested in the number R_n of Y's not exceeding the level X and counted to the moment of the n-th exceedance. More precisely

DEFINITION 2. For any integer $n \geq 1$

$$R_n = \min\{j \geq 0 : S_{n+j-1} = j, Y_{n+j} > X\}$$

denotes the number of Y's below the level X in a sample of the size $R_n + n$ with the last observation exceeding the level X.

Observe that the equivalent definition in terms of order statistics has the form

$$R_n = \min\{j \geq 0 : Y_{j+1:n+j} > X\}.$$

The exact distribution of R_n and its two first conditional moments are given in

THEOREM 2. *Assume that $P(G(X) < 1) > 0$. For any integer $n \geq 1$*

$$P(R_n = j) = \binom{n+j-1}{n-1} E(G^j(X)\bar{G}^n(X)), \qquad j = 0, 1, \ldots,$$

and $E(R_n) = nE(\frac{G(X)}{\bar{G}(X)} I_{(0,1)}(G(X)))$,

$$\mathrm{Var}(R_n) = nE\left(\frac{G(X)}{\bar{G}^2(X)} I_{(0,1)}(G(X))\right) + n^2 \,\mathrm{Var}\left(\frac{G(X)}{\bar{G}(X)} I_{(0,1)}(G(X))\right).$$

Observe that if $G(X) = 1$ a.s. then the right end of the support of Y's is less than the left end of the support of X and then with probability 1 no exceedance is possible. The indicator $I_{(0,1)}(G(X))$ appearing in the expressions for moments is necessary due to the fact that, in general, it may happen that $\bar{G}(X) = 0$ with a positive probability. Instead also conditional distribution of R_n given the event $\{G(X) < 1\}$, and its moments can be considered—see Section 3.

PROOF OF THEOREM 2. Similarly as in the proof of Theorem 1 we have

$$
\begin{aligned}
P(R_n = j) \\
&= P\left(\bigcup_{\sum_{k=1}^{n+j-1}\epsilon_k = j}\ \bigcap_{i=1}^{n+j-1} A_{i,\epsilon_i} \cap A_{n+j,0}\right) \\
&= \binom{n+j-1}{n-1}\int_{\boldsymbol{R}} P(Y_1 \le x, \ldots, Y_j \le x, Y_{j+1} > x, \ldots, Y_{j+n} > x)dF(x) \\
&= \binom{n+j-1}{n-1}\int_{\boldsymbol{R}} G^j(x)(1-G(x))^n dF(x).
\end{aligned}
$$

To compute the moments it is convenient first to establish that one can change the order of summation and taking expectations. To this end observe that $\exists c \in (0,1)$ such that $\forall j \ge 1$ $E(G^j(X)) < c^j$. To show this assume the converse. Then $\forall c \in (0,1)$ $\exists j \ge 1$ such that $E(G^j(X)) \ge c^j$. Denote such j by j_0. Hence $E(G^{j_0}(X)) = 1$ yielding $G(X) = 1$ a.s., which is impossible. Consequently it is allowed to write the expected value of R_n in the form

$$
E(R_n) = E\left(I_{(0,1)}(G(X))\sum_{j=0}^{\infty} j\binom{n+j-1}{n-1}G^j(X)(1-G(X))^n\right)
$$

and the final formula follows from the respective expression for the negative binomial distribution. Similarly for the second moment one has

$$
\begin{aligned}
E(R_n^2) &= E\left(I_{(0,1)}(G(X))\sum_{j=0}^{\infty} j^2\binom{n+j-1}{n-1}G^j(X)(1-G(X))^n\right) \\
&= n(n+1)E\left(\left(\frac{G(X)}{1-G(X)}\right)^2 I_{(0,1)}(G(X))\right) \\
&\quad + nE\left(\frac{G(X)}{1-G(X)} I_{(0,1)}(G(X))\right)
\end{aligned}
$$

and hence the formula for the variance follows immediately. □

Third, we are concerned with number of records K of the sequence $\boldsymbol{Y}$ falling below the threshold X. More precisely

DEFINITION 3. Let $U(j)$ denotes the j-th record time for the sequence $\boldsymbol{Y}$, $j = 1, 2, \ldots$, i.e. $U(1) = 1$ and $U(j) = \min\{i > U(j-1) : Y_i > Y_{U(j-1)}\}$, $j = 2, 3, \ldots$. Then

$$
K = \min\{j \ge 0 : Y_{U(j+1)} > X\}
$$

denotes the number of records of the sequence $\boldsymbol{Y}$ not exceeding the level X.

Again, as in the latter case, it is reasonable to exclude the case $G(X) = 1$ a.s., hence then all Y's remain below the threshold with probability one.

THEOREM 3. *Assume that* $P(G(X) < 1) > 0$. *Then*

$$P(K = j) = \frac{1}{j!} E\left(\bar{G}(X)(-\log(\bar{G}(X)))^j I_{(0,1)}(G(X))\right), \qquad j = 0, 1, \ldots$$

and $E(K) = -E(\log(\bar{G}(X)) I_{(0,1)}(G(X)))$,

$$\mathrm{Var}(K) = -E(\log(\bar{G}(X)) I_{(0,1)}(G(X))) + \mathrm{Var}(\log(\bar{G}(X)) I_{(0,1)}(G(X))).$$

Similarly as in Theorem 2 instead of indicators one can consider the conditional distribution of K given the event $\{G(X) < 1\}$, i.e.

$$P(K = j \mid G(X) < 1) = \frac{1}{j!} E\left(\bar{G}(X)(-\log(\bar{G}(X)))^j \mid G(X) < 1\right), \qquad j = 0, 1, \ldots,$$

$E(K \mid G(X) < 1) = -E(\log(\bar{G}(X)) \mid G(X) < 1)$ and

$$\begin{aligned}\mathrm{Var}(K \mid G(X) < 1) = &-E(\log(\bar{G}(X)) \mid G(X) < 1) \\ &+ \mathrm{Var}(\log(\bar{G}(X)) \mid G(X) < 1).\end{aligned}$$

Observe also that it follows from Theorem 3, in the case $G(X) = 0$ a.s., i.e. all the Y's exceed X with probability one, that $P(K = 0) = 1$.

PROOF OF THEOREM 3. Let $Y_{U(0)} = -\infty$. Then for any $j = 0, 1, \ldots$

$$P(K = j) = P(Y_{U(j)} \leq X < Y_{U(j+1)}) = \int_{\boldsymbol{R}} P(Y_{U(j)} \leq x < Y_{U(j+1)}) dF(x).$$

Since the joint df of $(Y_{U(j)}, Y_{U(j+1)})$ has the form

$$P(Y_{U(j)} \leq u, Y_{U(j+1)} \leq v) = \frac{1}{j!}(-\log(\bar{G}(u)))^j G(v) I_{(0,\infty)}(v - u)$$

(see for instance Ahsanullah (1995)), then

$$\begin{aligned}P(K = j) &= \int_{\boldsymbol{R}} \int_x^{\infty} \int_{-\infty}^{x} \frac{1}{j!} d((-\log(\bar{G}(u)))^j) dG(v) dF(x) \\ &= \int_{\boldsymbol{R}} \frac{1}{j!}(-\log(\bar{G}(x)))^j (1 - G(x)) I_{(0,1)}(G(x)) dF(x)\end{aligned}$$

for any $j = 0, 1, \ldots$. Now, similarly as in the proof of Theorem 2 it follows that $\exists c \in (0, \infty)$ such that $\forall j \geq 1$ $E(-\log(\bar{G}(X))^j) < c^j$.

Observe that the converse of that statement leads to $G(X) = 1$ a.s. which is contradictory. Consequently the order of summation and taking expectations in the formulas for the moments of K can be reversed. Hence, via respective expressions for the Poisson distribution, one easily gets

$$\begin{aligned} E(K) &= E\left(I_{(0,1)}(G(X))\bar{G}(X) \sum_{j=0}^{\infty} j \frac{(-\log(\bar{G}(X)))^j}{j!} \right) \\ &= -E(\log(\bar{G}(X)) I_{(0,1)}(G(X))), \\ E(K^2) &= E\left(I_{(0,1)}(G(X))\bar{G}(X) \sum_{j=0}^{\infty} j^2 \frac{(-\log(\bar{G}(X)))^j}{j!} \right) \\ &= -E[\log(\bar{G}(X)) I_{(0,1)}(G(X))] + E[(\log(\bar{G}(X)))^2 I_{(0,1)}(G(X))] \end{aligned}$$

and the final variance formula follows immediately. □

Remark 2.1. Observe that the exceedance model, we discuss, gives a natural mechanism for producing various types of generalized binomial, negative binomial and Poisson mixtures with respect to the continuous parameters, since it follows from the proofs of theorems of this section that:

$$\begin{aligned} \mu_{S_n|X} &= b(n, G(X)), \\ \mu_{R_n|X} &= nb(n, \bar{G}(X)), \\ \mu_{K|X} &= \mathcal{P}(-\log(\bar{G}(X))), \end{aligned}$$

where b, nb and $\mathcal{P}$ denote binomial, negative binomial and Poisson distribution, respectively, and μ's with subscripts denote conditional distributions. Consequently, for the degenerate X, say $P(X = c) = 1$ it follows that S_n is binomial with parameters n, $p = G(c)$, R_n is negative binomial with parameters n, $p = \bar{G}(c)$ and K is Poisson with the paramater $\lambda = -\log(\bar{G}(c))$.

3. General limit results

In this section we derive asymptotic properties of distributions of properly normalized exceedance statistics connected with tending to infinity with their parameter n. It is included in the first two statistics and it will be introduced in a natural way to the third one later on in this section. It should be emphasized that the asymptotic results obtained here give immediately numerous approximations for a variety of discrete distributions of respective statistics, obtained for special choices of F and G.

We start with the limiting behaviour of S_n/n as n tends to infinity.

THEOREM 4. *For $n \to \infty$*

$$\frac{1}{n} S_n \xrightarrow{d} G(X).$$

PROOF. Consider the chf of S_n/n:

$$E(e^{itS_n/n}) = E\left(\sum_{j=0}^{n}\binom{n}{j}(e^{it/n}G(X))^j(1-G(X))^{n-j}\right)$$
$$= E[(1-G(X)+G(X)\exp(it/n))^n].$$

Observe that for any real a and complex b

$$\lim_{x\to\infty}(1-a+a\exp(b/x))^x = \exp(ab). \tag{3.1}$$

It follows easily from

$$x\log(1-a+a\exp(b/x)) = x\log(1+ab/x+o(1/x))$$
$$= x(ab/x+o(1/x)) \overset{x\to\infty}{\longrightarrow} ab.$$

Applying (3.1) with $a = G(X)$ and $b = it$ to the formula for the chf of S_n/n it follows that

$$E(e^{itS_n/n}) \to E(e^{itG(X)})$$

as $n \to \infty$. (Taking the limit under the integral sign is permitted since $|1-G(X)+G(X)\exp(it/n)| \le 1$ a.s. for any $n \ge 1$.) Therefore the theorem follows. □

Remark 3.1. Due to the fact that the variance of S_n is of the order n^2 for a nondegenerate X and of the order n for a degenerate X it follows that for $n \to \infty$

$$\frac{S_n - E(S_n)}{\sqrt{\mathrm{Var}(S_n)}} \overset{d}{\to} \begin{cases} \mathcal{N}(0,1) & \text{if } X \text{ is degenerate,} \\ \dfrac{G(X)-E(G(X))}{\sqrt{\mathrm{Var}(G(X))}} & \text{if } X \text{ is nondegenerate.} \end{cases}$$

The first assertation is just the de Moivre-Laplace theorem while the second follows from Theorem 4 and the following observations (see Theorem 1):

$$\frac{E(S_n)}{\sqrt{\mathrm{Var}(S_n)}} \overset{n\to\infty}{\longrightarrow} \frac{E(G(X))}{\sqrt{\mathrm{Var}(G(X))}},$$
$$\frac{n}{\sqrt{\mathrm{Var}(S_n)}} \overset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{\mathrm{Var}(G(X))}}.$$

Now we turn to the asymptotic behaviour of R_n/n.

THEOREM 5. *Assume that $G(X) < 1$ a.s. Then for $n \to \infty$*

$$\frac{1}{n}R_n \overset{d}{\to} \frac{G(X)}{\bar{G}(X)}.$$

Remark 3.2. If the condition $G(X) < 1$ a.s. does not hold then the right hand side of the above formula is not a proper rv. In that case, it is preferable to use a more general limit property for the conditional distribution as $n \to \infty$:

$$\mu_{(R_n/n)|G(X)<1} \xrightarrow{w} \mu_{(G(X)/\bar{G}(X))|G(X)<1},$$

where $\xrightarrow{w}$ denotes the weak convergence of probability measures. This asymptoticity follows along the lines of the proof of Theorem 5, given beneath, since as in Theorem 2 one can easily get that

$$P(R_n = j \mid G(X) < 1) = \binom{n+j-1}{n-1} E(G^j(X)\bar{G}^n(X) \mid G(X) < 1),$$

$$j = 0, 1, \ldots.$$

PROOF OF THEOREM 5. For the chf of the rv R_n/n we have for any real t

$$\begin{aligned} E(e^{itR_n/n}) &= E\left(\sum_{j=0}^{\infty}\binom{n+j-1}{n-1}(e^{it/n}G(X))^j\bar{G}^n(X)\right) \\ &= E\left[\left(\frac{\bar{G}(X)}{1-G(X)\exp(it/n)}\right)^n\right] \\ &= E\left[\left(\frac{1}{1-G(X)} - \frac{G(X)}{1-G(X)}e^{it/n}\right)^{-n}\right]. \end{aligned}$$

Now (3.1) with $a = -G(X)/G(X)$ and $b = it$ (see the proof of Theorem 4) yields

$$E(e^{itR_n/n}) \xrightarrow{n\to\infty} E[(e^{-itG(X)/\bar{G}(X)})^{-1}], \qquad t \in \boldsymbol{R}.$$

Therefore the theorem follows. □

Remark 3.3. Similarly as for S_n (see Remark 3.1), under an additional assumption $G(X) < 1$ a.s. (or using conditioning), we have

$$\frac{R_n - E(R_n)}{\sqrt{\mathrm{Var}(R_n)}} \xrightarrow{d} \begin{cases} \mathcal{N}(0,1) & \text{if } X \text{ is degenerate,} \\ \dfrac{G(X)/\bar{G}(X) - E(G(X)/\bar{G}(X))}{\sqrt{\mathrm{Var}(G(X)/\bar{G}(X))}} & \text{if } X \text{ is nondegenerate.} \end{cases}$$

Now the first limit property follows from ($q = 1 - p = G(c)$, where $P(X = c) = 1$)

$$\begin{aligned} E\left[\exp\left(it\frac{R_n - E(R_n)}{\sqrt{\mathrm{Var}(R_n)}}\right)\right] &= \exp(-it\sqrt{nq})E\left[\exp\left(\frac{itpR_n}{\sqrt{qn}}\right)\right] \\ &= \left(\frac{p\exp(-it\sqrt{q/n})}{1-q\exp(itp/\sqrt{qn})}\right)^n \\ &= (p^{-1}[1 + it\sqrt{q/n} - qt^2/(2n) + o(1/n) \\ &\qquad - q(1 + it/\sqrt{qn} - t^2/(2qn) + o(1/n))])^{-n} \\ &= \frac{1}{(1 + t^2/(2n) + o(n))^n} \xrightarrow{n\to\infty} -\frac{t^2}{2}, \qquad t \in \boldsymbol{R}, \end{aligned}$$

while the second is an immediate consequence of Theorems 2 and 5.

To study asymptotic properties of the number of records below the level X, first we introduce, in a very natural way, the asymtoticity parameter n. Consider the following extension of our scheme.

Let $(Y_{1,l},\ldots,Y_{n,l})$, $l=1,2,\ldots$, be independent random vectors with iid components with a common df G, where $n \geq 1$ is an integer. Consider the sequence of their minima: $\boldsymbol{Y}_{1:n}=(Y_{1:n,l})_{l=1,\ldots}$. Observe that $\boldsymbol{Y}_{1:n}$ is again a sequence of iid rv's with a common df $1-(1-G)^n$, $n \geq 1$. Now define K_n as the number of records of the sequence $\boldsymbol{Y}_{1:n}$ falling below the threshold X:

DEFINITION 4. For any integer $n \geq 1$

$$K_n = \min\{j \geq 0 : Y_{1:n,U(j+1)} > X\},$$

where $U(j)$, $j=1,2,\ldots$, are record times for the sequence $\boldsymbol{Y}_{1:n}$, is the number of records of minima not exceeding the random level X.

It follows immediately from Theorem 3 that in the case $G(X)<1$ a.s.

$$\begin{aligned}
&P(K_n = j) = \frac{n^j}{j!}E(\bar{G}^n(X)(-\log(\bar{G}(X)))^j), \qquad j=0,1,\ldots,\\
&\mu_{K_n|X} = \mathcal{P}(-n\log(\bar{G}(X))),\\
&E(K_n) = -nE(\log(\bar{G}(X))),\\
&\operatorname{Var}(K_n) = -nE(\log(\bar{G}(X))) + n^2\operatorname{Var}(\log(\bar{G}(X))).
\end{aligned}$$

If the assumption $G(X)<1$ a.s. is not satisfied again one has to use in the above formulas the indicator $I_{(0,1)}(G(X))$ or conditioning with respect to the event $\{G(X)<1\}$. Of course $K_1 = K$ introduced in Section 2. Observe that the distribution of the sequence $\boldsymbol{Y}_{1:n}$ coincides with the distribution of the sequence of, so called, n-th records introduced in Dziubdziela and Kopocinski (1976) (see also Ch. 9 in Arnold *et al.* (1992) for a more convenient reference). Recall that n-th records are defined as the n-th largest value in the sequence. Since we are interested here only in distributional properties of K_n alternatively anyone of these two specific models: records of minima of n observations or n-th records can be kept in mind. An extension of this scheme will be given in Remark 3.6 below.

Now we study asymptotic properties of K_n:

THEOREM 6. *Assume that $G(X)<1$ a.s. Then for $n\to\infty$*

$$\frac{1}{n}K_n \xrightarrow{d} -\log(\bar{G}(X)).$$

PROOF. Again, as in the previous two proofs, we consider the chf of the rv K_n:

$$E(e^{itK_n/n}) = E\left[\bar{G}^n(X)\sum_{j=0}^{\infty}\frac{1}{j!}(-ne^{it/n}\log(\bar{G}(X)))^j\right]$$
$$= E(\exp[-n\log(\bar{G}(X))(e^{it/n}-1)]).$$

Since

$$n(e^{it/n}-1) = n(it/n + o(1/n)) \overset{n\to\infty}{\longrightarrow} it,$$

then $\forall t \in \boldsymbol{R}$

$$E(e^{itK_n/n}) \overset{n\to\infty}{\longrightarrow} E(e^{-it\log(\bar{G}(X))}). \qquad \square$$

Remark 3.4. Similarly as in the case of R_n the condition $G(X) < 1$ a.s. can not be simply removed since $\log(\bar{G}(X))$ is not a proper rv if it does not hold. Again then one can turn to the conditional distribution:

$$P(K_n = j \mid G(X) < 1) = \frac{n^j}{j!}E[\bar{G}^n(X)(-\log(\bar{G}(X)))^j \mid G(X) < 1],$$

$j = 0, 1, \ldots$. Hence repeating the argument from the above proof one gets for the conditional distributions

$$\mu_{(K_n/n)G(X)<1} \overset{w}{\to} \mu_{-\log(\bar{G}(X))|G(X)<1}.$$

Remark 3.5. Analogously as for S_n and R_n, under an additional condition $G(X) < 1$ a.s. (or upon conditioning), it follows easily that

$$\frac{K_n - E(K_n)}{\sqrt{\mathrm{Var}(K_n)}} \overset{d}{\to} \begin{cases} \mathcal{N}(0,1) & \text{if } X \text{ is degenerate,} \\ \dfrac{-\log(\bar{G}(X)) + E(\log(\bar{G}(X)))}{\sqrt{\mathrm{Var}(\log(\bar{G}(X)))}} & \text{if } X \text{ is nondegenerate.} \end{cases}$$

Remark 3.6. For a sequence of random vectors $(Y_{1,l}, \ldots, Y_{n,l})$, $l = 1, 2, \ldots$, consider now the sequence of k-th order statistics $\boldsymbol{Y}_{k:n} = (Y_{k:n,l})_{l=1,\ldots}$ for some fixed $k \in \{1, \ldots, n\}$. It follows that $\boldsymbol{Y}_{k:n}$ is a sequence of iid rv's with a common df

$$G_{k:n} = \sum_{r=k}^{n}\binom{n}{r}G^r(1-G)^{n-r}.$$

Analogously to the definition of K_n we consider the number of records of the sequence of k-th order statistics not exceeding the level X:

$$K_{(k:n)} = \min\{j \geq 0 : Y_{k:n,U(j+1)} > X\},$$

where $U(j)$, $j = 1, 2, \ldots$ are the record times of the sequence $\boldsymbol{Y}_{k:n}$.

Obviously Theorem 3 implies

$$P(K_{(k:n)} = j) = \frac{1}{j!}E\left[\left(\sum_{r=0}^{k-1}\binom{n}{r}G^r(X)\bar{G}^{n-r}(X)\right)\right.$$
$$\left.\cdot\left(-\log\left(\sum_{r=0}^{k-1}\binom{n}{r}G^r(X)\bar{G}^{n-r}(X)\right)\right)^j\right],$$
$$j = 0, 1, \ldots.$$

Similarly to the asymptotic behaviour of $K_n = K_{(1:n)}$ given in Theorem 6 it can be proved that under the assumption $G(X) < 1$ a.s. for $n \to \infty$ and fixed but arbitrary k

$$\frac{1}{n}K_{(k:n)} \xrightarrow{d} -\log(\bar{G}(X)).$$

To establish this convergence result we follow the argument from the proof of Theorem 6. Hence

$$E(e^{itK_{(k:n)}}) = E[\exp(-\log(\bar{G}_{k:n}(X))(e^{it/n} - 1))].$$

Consequently, the form of $G_{k:n}$ implies that it suffices to prove that

$$\lim_{n\to\infty} \log\left(\sum_{r=0}^{k-1}\binom{n}{r}p^r(1-p)^{n-r}\right)(e^{b/n} - 1) = b\log(1-p)$$

for any $G(X)(\omega) = p \in (0,1)$ and any $b = it$. And the last convergence follows from the double inequality

$$(1-p)^n \le \sum_{r=0}^{k-1}\binom{n}{r}p^r(1-p)^{n-r} \le (1-p)^{n-k+1}n^k k.$$

Hence upon taking logarithms

$$n\log(1-p) \le \log\left(\sum_{r=0}^{k-1}\binom{n}{r}p^r(1-p)^{n-r}\right)$$
$$\le (n-k+1)\log(1-p) + k\log(n) + \log(k).$$

Then multiplying the above inequalities by $\exp(b/n) - 1$ and passing to the limit as n tends to infinity we easily get: $n\log(1-p)(\exp(b/n) - 1) \to b\log(1-p)$, $k\log(n)(\exp(b/n) - 1) \to 0$ and $(\log(k) - (k-1)\log(1-p))(\exp(b/n) - 1) \to 0$.

4. Distribution-free properties

Here we are concerned with distributions of exceedance statistics in a special, but very important case when the distribution of the rv X and that of the sequence of observations coincide. Such a model leads to an interesting family of discrete distributions containing some widely known laws and some completely new ones appearing in a quite natural way. The considerations contained in this section can also be used as a starting point for deriving tests for equidistribution.

Since we are going to consider different random thresholds, instead of symbols S_n, R_n and K_n we will use $S_n(Z)$, $R_n(Z)$ and $K_n(Z)$, where Z denotes the random threshold. The following versions of Z will be used: X, $X_{k:m}$ $(1 \le k \le m)$, $X_{U(k)}$ $(k \ge 1)$. Obviously, to have $X_{k:m}$ or $X_{U(k)}$ we have to assume that a sequence $X_1, X_2, \dots$ of iid X's with a df F is given such that the sequences of X's and Y's are independent.

Observe that in the model we consider throughout this section we have

$$G(X) \sim U(0,1),$$

where $U(0,1)$ denotes the uniform distribution in the interval $(0,1)$;

$$G(X_{k:m}) \sim beta_I(k, m-k+1),$$

where $beta_I(k, m-k+1)$ is the beta distribution of the first kind defined by the density $f(x) = \frac{m!}{(k-1)!(m-k)!} x^{k-1}(1-x)^{m-k} I_{(0,1)}(x)$;

$$G(X_{U(k)}) \sim U_{U(k)}(0,1)$$

where $U_{U(k)}(0,1)$ denotes the distribution of the k-th record from $U(0,1)$ distribution defined by the density $f(x) = [(-\log(1-x))^k/k!] I_{(0,1)}(x)$. This density defines a special case of log-gamma (or unit-gamma) distribution studied by Grassia (1977), and considered as an alternative to the beta distribution by, for example, Ratnaparkhi (1981) and more recently in the damage model set-up by Fosam and Sapatinas (1995).

From Theorem 1, upon integrating, one easily gets the following special cases:

COROLLARY 1. *If $F \equiv G$ then*
(a) *$S_n(X)$ has the uniform distribution over the set $\{0, 1, \dots, n\}$, U_n, i.e.*

$$P(S_n(X) = j) = \frac{1}{n+1}, \qquad j = 0, 1, \dots, n;$$

(b) *$S_n(X_{k:m})$ has the negative hypergeometric distribution of the first kind, $nh_I(k, m, n)$, i.e.*

$$P(S_n(X_{k:m}) = j) = \frac{\binom{k+j-1}{k-1}\binom{n+m-j-k}{m-k}}{\binom{n+m}{m}}, \qquad j = 0, 1, \dots, n;$$

(c) $S_n(X_{U(k)})$ *has the record exceedance distribution of the first kind,* $re_I(k,n)$, *i.e.*

$$P(S_n(X_{U(k)})=j)=\binom{n}{j}\sum_{l=0}^{j}\binom{j}{l}\frac{(-1)^{j-l}}{(n-l+1)^k},\qquad j=0,1,\ldots,n.$$

Remark 4.1. Properties of $S_n(X_{k:m})$ were widely investigated in fifties as was pointed out in Section 1—see Gumbel ((1958), pp. 58–60).

Remark 4.2. While the discrete distributions appearing in (a) and (b) are well known, the one appearing in (c) seems to be quite new. It has been derived in the integral form

$$P(S_n(X_{U(k)})=j)=\frac{1}{(k-1)!}\binom{n}{j}\int_0^\infty(1-e^{-t})^j e^{-(n-j+1)t}t^{k-1}dt,\qquad j=0,1,\ldots,n$$

in a recent paper by Bairamov (1997).

Remark 4.3. Observe that the case (a) is included in (b) or (c) since $U_n = nh_I(1,1,n)=re_I(1,n)$.

Remark 4.4. By the asymptotic properties of S_n given in Theorem 4 it follows easily that for $n\to\infty$

$$\frac{V_n}{n}\xrightarrow{d}V$$

for the following choices of distributions of V_n and V:

(1^o) U_n and $U(0,1)$;

(2^o) $nh_I(k,m,n)$ and $beta_I(k,m+k-1)$;

(3^o) $re_I(k,n)$ and $U_{U(k)}$.

Observe that 1^o follows from 2^o or 3^o since $U(0,1)=beta_I(1,1)=U_{U(1)}$.

Similarly Theorem 2 implies the following results for R_n.

COROLLARY 2. *If* $F\equiv G$ *then*

(a) $R_n(X)$ *has the Waring distribution,* $W(n,n+1)$, *i.e.*

$$P(R_n(X)=j)=\frac{n}{(n+j)(n+j+1)},\qquad j=0,1,\ldots;$$

(b) $R_n(X_{k:m})$ *has the negative hypergeometric distribution of the second kind,* $nh_{II}(k,m,n)$, *i.e.*

$$P(R_n(X_{k:m})=j)=\frac{k}{k+j}\frac{\binom{n+j-1}{n-1}\binom{m}{k}}{\binom{m+n+j}{m+n-k}},\qquad j=0,1,\ldots;$$

(c) $R_n(X_{U(k)})$ *has the record exceedance distribution of the second kind,* $re_{II}(k,n)$, *i.e.*

$$P(R_n(X_{U(k)}) = j) = \binom{n+j-1}{n-1} \sum_{l=0}^{j} \binom{l}{j} \frac{(-1)^l}{(n+l+1)^k}, \qquad j = 0, 1, \ldots.$$

Remark 4.5. The result of part (b) was derived in Wilks (1959). The $nh_{II}(k,m,n)$ is also called the beta-Pascal distribution—see Johnson *et al.* ((1992), pp. 242–243). Here we prefer to keep the name of negative hypergeometric of the second kind because of some analogy between both kinds of negative hypergeometric and beta distributions—see Remark 4.8 below.

Remark 4.6. The integral representation of $re_{II}(k,n)$ is the following

$$P(R_n(X_{U(k)}) = j) = \frac{1}{(k-1)!} \binom{n+j-1}{n-1} \int_0^\infty (1-e^{-t})^j e^{-(n+1)t} t^{k-1} dt,$$
$$j = 0, 1, \ldots.$$

Remark 4.7. Observe that $W(n, n+1) = nh_{II}(1,1,n) = re_{II}(1,n)$ and consequently (b) or (c) imply (a).

Remark 4.8. Via Theorem 5 it is easily obtained that for $n \to \infty$

$$\frac{V_n}{n} \xrightarrow{d} V$$

for the following choices of distributions for V_n and V, respectively:

(1^o) $W(n, n+1)$ and $Pareto(2)$, where $Pareto(\nu)$, $\nu > 1$, is defined by the density function $f(x) = (\nu - 1)(1+x)^{-\nu} I_{(0,\infty)}(x)$;

(2^o) $nh_{II}(k,m,n)$ and $beta_{II}(k, m-k+1)$, where the second kind beta distribution $beta_{II}(a,b)$, $a > 0$, $b > 0$, is defined by the density $f(x) = B^{-1}(a,b) x^{a-1} (1+x)^{-(a+b)} I_{(0,\infty)}(x)$;

(3^o) $re_{II}(k,n)$ and the distribution of $\log(Z_k + 1)$, where Z_k has the $gamma(1,k)$ distribution, where the gamma distribution $gamma(a,p)$, $a > 0$, $p > 0$, is defined by the density $f(x) = a^p \Gamma^{-1}(p) x^{p-1} \exp(-ax) I_{(0,\infty)}(x)$, i.e. the density of the limit distribution has the form $f(x) = [\log(1+x)]^{k-1}(1+x)^{-2}(k!)^{-1} I_{(0,\infty)}(x)$.

Observe that $Pareto(2) = beta_{II}(1,1) = \mathcal{L}(\log(Z_1+1))$ ($\mathcal{L}(\xi)$ denotes the distribution of the rv ξ), hence 1^o follows immediately from 2^o or 3^o.

Finally let us consider the distribution-free properties of K_n's. This time the following results follow immediately from Theorem 3.

COROLLARY 3. *If* $F \equiv G$ *then:*

(a) $K_n(X)$ *has the geometric distribution*, $ge(1/(n+1))$, *i.e.*

$$P(K_n(X)=j)=\frac{n^j}{(n+1)^{j+1}},\qquad j=0,1,\ldots;$$

(b) $K_n(X_{k:m})$ *has the record-order exceedance distribution*, $roe(k,m,n)$, *i.e.*

$$P(K_n(X_{k:m})=j)=kn^j\binom{m}{k}\sum_{l=0}^{k}\frac{(-1)^{k-l-1}}{(n+m-l)^{j+1}},\qquad j=0,1,\ldots;$$

(c) $K_n(X_{U(k)})$ *has the negative binomial distribution*, $nb(k,1/(n+1))$, *i.e.*

$$P(K_n(X_{U(k)})=j)=\binom{k+j-1}{k-1}\frac{n^j}{(n+1)^{j+k}},\qquad j=0,1,\ldots.$$

Remark 4.9. Observe that $roe(1,m,n)=ge(1/(n+m))$ hence it follows from (a) and (b) of Corollary 3 that the number of $(n+m-1)$-th records not exceeding the level X has the same distribution as the number of n-th records not exceeding the minimum of the sample of size m from the distribution of X, i.e.

$$K_{n+m-1}(X)\stackrel{d}{=}K_n(X_{1:m})=ge(1/(n+m)).$$

Remark 4.10. The integral representation of the pmf of the $roe(k,m,n)$ distribution is of the following form:

$$P(K_n(X_{k:m})=j)=\frac{kn^j}{j!}\binom{m}{k}\int_0^\infty(1-e^{-t})^{k-1}e^{-(n+m-k+1)t}t^{j-1}dt.$$

Remark 4.11. Observe that again (a) is covered by (b) or (c) due to the identities $ge(1/(n+1))=roe(1,1,n)=nb(1,1/(n+1))$.

Remark 4.12. From the asymptotic properties of K_n statistics, see Theorem 6, it follows easily that for $n\to\infty$

$$\frac{V_n}{n}\stackrel{d}{\to}V$$

for the following choices of distributions of V_n and V, respectively:

(1^o) $ge(1/(n+1))$ and exponential $Exp(1)$, where $Exp(\lambda)=gamma(\lambda,1)$;

(2^o) $roe(k,m,n)$ and $\mathcal{L}(-\log(Z_{k,m}))$, where $Z_{k,m}$ has the $beta_I(m-k+1,k)$ distribution, i.e. the pdf of the limit distribution has the form $f(x)=B^{-1}(m-k+1,k-1)\exp(-(m-k+1)x)(1-\exp(-x))^{k-1}I_{(0,\infty)}(x)$ and as a special case we have $roe(1,m,n)=ge(1/(m+n))$ and $Exp(m)$ for the distributions of V_n and V, respectively;

(3^o) $nb(k,1/(n+1))$ and $gamma(1,k)$.

Observe that 1^o is covered by anyone of 2^o or 3^o since $Exp(1)=\mathcal{L}(-\log(Z_{1,1}))=gamma(1,1)$.

5. Characterization of equidistribution

In this section we return to the general model, i.e. we do not assume $F \equiv G$. Conversely to what was done in Section 4, we would like to use exact distributions of exceedance statistics derived there, to conclude that $F \equiv G$. It appears that for some of them such characterizations of equidistribution really holds, and they can be preferred for developing tests of identity of distributions, while for others counterexamples can be constructed.

We start with a negative result involving the distributions of S_n.

Assume that $S_n(X)$ has the discrete uniform distribution over the set $\{0,1,\dots,n\}$, U_n, as it is in the case of equidistribution—see Corollary 1. Then it follows that

$$\binom{n}{j} E(G^j(X)(1-G(X))^{n-j}) = \frac{1}{n+1}, \qquad j = 0,1,\dots,n.$$

Take $X \sim U(0,1)$. Then the above identity yields

$$\binom{n}{j} \int_0^1 G^j(x)(1-G(x))^{n-j} dx = \frac{1}{n+1}, \qquad j = 0,1,\dots,n.$$

Obviously $G(x) = x$, $x \in (0,1)$, satifies the above requirement, see Corollary 1 (a). The question is wheteher other df's G can fit this scheme. Observe that the above system of equations can be rewritten as

$$E(Z_n^j) = \binom{n}{j}^{-1}, \qquad j = 0,1,\dots,n,$$

where $Z_n = T_n/(1-T_n)$, and T_n is a rv with the df H defined by

$$dH(x) = (n+1)(1-x)^n dG^{-1}(x), \qquad x \in (0,1),$$

where G^{-1} is the unique inverse defined by $G^{-1}(x) = \sup\{y : G(y) \le x\}$. Now we have restrictions on the values of the first n moments of Z_n imposed, which obviously is not sufficient to determine uniquely its distribution. Hence H and, consequently, G are not unique.

Similarly in the remaining two cases of the S_n statistics the assumption

$$\binom{n}{j} E(G^j(X)(1-G(X))^{n-j}) = \alpha_j, \qquad j = 0,1,\dots,n,$$

where α_j's are given in Corollary 1 (b) or (c), can be reduced to the restrictions on the first n moments of the rv Z_n and the argument repeats.

This observation is very close (but not the same) to nonidentifiability of binomial mixtures announced in Teicher (1961).

On the other hand, for both the remaining statistics R_n and K_n, their exact distributions as for $F \equiv G$, imply equidistribution. Since, as was mentioned at the end of Section 2, these distributions are mixtures of the modified power series form, the result given beneath is closely related to the main result of Sapatinas (1995) (see Remark 5.2). That result gives a Carleman like condition for identifiabiliy of a wide class of modified power series mixtures, while not covering exactly the situation we consider here. Nevertheless the method, referring explicitly to the Carleman criterion, used in the present proof, is quite close to that from Sapatinas (1995) and also similar to one used in Wesołowski (1995).

THEOREM 7. *Let n be an arbitrary (fixed) positive ineteger. If*
(a) *$R_n(X)$ has the Waring distribution, $W(n, n+1)$, or*
(b) *$K_n(X)$ has the geometric distribution, $ge(1/(n+1))$,*
then $F \equiv G$.

PROOF. (a) By Theorem 2 for any $j = 0, 1, \ldots$,

$$\frac{n}{(n+j)(n+j+1)} = \binom{n+j-1}{n-1} E(G^j(X)(1-G(X))^n) = \binom{n+j-1}{n-1} \int_0^1 x^j (1-x)^n dF(G^{-1}(x)).$$

Define now a new df H by the formula

$$dH(x) = (n+1)(1-x)^n dF(G^{-1}(x)), \qquad x \in (0,1).$$

Let Z_n be a rv with a df H. Then the above system of equations yields

$$E(Z_n^j) = \frac{(n+1)!j!}{(n+j+1)!}, \qquad j = 0, 1, \ldots.$$

Hence all the moments of Z_n are uniquely determined and since $supp(Z_n) = [0,1]$ we conclude that its distribution is uniquely determined by the moment sequence. By Corollary 2 (a) it follows that $dH(x) = (n+1)(1-x)^n dx$. Consequently $F(G^{-1}(x)) = x$, implying $F \equiv G$.

(b) Similarly as for (a) we have

$$\frac{1}{j!} \int_0^1 (1-x)^n (-\log(1-x))^j dF(G^{-1}(x)) = \frac{1}{(n+1)^{j+1}}, \qquad j = 0, 1, \ldots.$$

Now defining H and Z_n as in (a) consider a new rv $T_n = -\log(1 - Z_n)$. The above system of equation leads to

$$E(T_n^j) = \frac{j!}{(n+1)^{j+1}}, \qquad j = 0, 1, \ldots.$$

Since $E(T_n^j) < j^j$, $j = 1, 2, \ldots$, then

$$\sum_{j=1}^{\infty} E^{-1/j}(T_n^j) > \sum_{j=1}^{\infty} \frac{1}{j} = \infty.$$

Hence the Carleman criterion implies that the distribution of T_n is uniquely determined by the moment sequence $(E(T_n^j))_{j=1,2,\ldots}$. It means that the distribution of $Z_n = 1 - \exp(-T_n)$ is also unique and the final result follows as for (a), but this time via Corollary 3 (a). □

The remaining cases of the exact distributions for $R_n(X_{k:m})$, $R_n(X_{U(k)})$, $K_n(X_{k:m})$ and $K_n(X_{U(k)})$ follow along the same lines. All these results point out to the respective statistics as convenient tools for testing the equidistribution hypothesis for data connected with exceeding a random level.

Remark 5.1. Assume now that for the general model, i.e. without the assumption $F = G$, the distribution of, say, $K_n(X)$ is known. It means we are given the numbers $p_j(n) = P(K_n(X) = j)$, $j = 0, 1, \ldots$. Then similarly as in the proof of Theorem 7 we define

$$dH(x) = p_0^{-1}(1-x)^n dF(G^{-1}(x)), \qquad x \in (0,1),$$

which appears to be uniquely determined by the sequence $(p_j(n))_{n=1,2,\ldots}$ since

$$E(T_n^j) = \frac{p_j(n)j!}{p_0(n)n^j} < \frac{j^j}{p_0(n)}, \qquad j = 1, 2, \ldots.$$

Now for

$$\hat{H}(x) = p_0 \int_0^x (1-u)^{-n} dH(u),$$

$x \in (0,1)$, we have

$$\hat{H} \circ G \equiv F, \tag{5.1}$$

where $\hat{H}$ is a df on $(0,1)$, again uniquely determined by $(p_j(n))_{n=1,2,\ldots}$. The above relation (as given for $K_n(X)$ or its analogues for other exceedance statistics) can be a source of many new characterization results, since assuming one of F or G to be known, the second can be derived. To some extent this observation will be exploited in the next section for characterizing the gamma and exponential distribution.

Remark 5.2. Observe that assuming G to be the df of the $U(0,1)$ distribution, the question of seeking for F while knowing one of the distributions of $S_n(X)$, $R_n(X)$ or $K_n(X)$ is nothing else but the identifiability problem for the binomial, negative binomial or Poisson mixtures—see Teicher (1961), Barndorff-Nielsen (1965) or the last chapter in a recent monograph Prakasa Rao (1992). Hence the possible approach announced in Remark 5.1 can be viewed as identifiability of generalized b, nb and $\mathcal{P}$ mixtures. Consequently if G is known and it is a 1-1 function then by Theorem 2.1 (together with Remark 1(iii)) of Sapatinas (1995) it follows that F is unique. This will be the case considered in Theorem 8 of Section 6.

6. Records exceedance for exponential sequences

Let $Y_1, Y_2, \ldots$ be iid exponential with the density $f(x) = \lambda \exp(-\lambda x) I_{(0,\infty)}(x)$. Then it follows immediately from general properties of $K_n = K_n(X)$ given in Section 3 that

$$P(K_n = j) = \frac{(\lambda n)^j}{j!} E(X^j e^{-n\lambda X}), \qquad j = 0, 1, \ldots.$$

From now on we will consider, without much loss of generality, the case $n = 1$, and denote as in Section 2, $K_1 = K$. Consequently we have

$$P(K = j) = \frac{\lambda^j}{j!} E(X^j e^{-\lambda X}), \qquad j = 0, 1, \ldots.$$

In the sequel we are going to consider two different distributions for the random threshold X: gamma and normal. In the first case we derive the exact distribution for K and then respective characterization results, while in the second case after computing the exact distribution of K we use it to obtain an unexpected derivation of the Laplace expansion formula for the df of the standard normal distribution.

6.1 *The gamma threshold*

Assume that X is distributed according to the gamma law, $gamma(r, a)$. Then

$$P(K = j) = \frac{a^r \lambda^j}{\Gamma(r) j!} \int_0^\infty x^{j+r-1} e^{-(\lambda+a)x} dx = \binom{r+j-1}{r-1} \left(\frac{\lambda}{\lambda + a}\right)^j \left(\frac{a}{\lambda + a}\right)^r,$$

$j = 0, 1, \ldots$, i.e. K is a negative binomial, $nb(r, a/(a+\lambda))$, rv. This observation will be used now for characterizations of the gamma via exponential and exponential via gamma laws. The result can be looked at as a generalization of the Engel, Zijlstra (1980) characterization of the gamma distribution—see also Cacoullos and Papageorgiou (1982).

THEOREM 8. *Let K have the negative binomial distribution $nb(r, p)$, where $r > 0$ and $p \in (0, 1)$ are given numbers. Then*

$$F \sim gamma(r, a) \Leftrightarrow G \sim Exp(\lambda),$$

where $a > 0$ and $\lambda > 0$ satisfy the relation $\lambda p = a(1-p)$.

PROOF. Assume that G is exponential. Then Theorem 2.1 and Remark 1(iii) of Sapatinas (1995) imply that F is unique, hence it has to be a df of the $gamma(r, a)$ distribution.

Assume now that F is a df of the gamma distribution, $gamma(r, a)$. Since $\hat{H}$, as defined in Remark 5.1 is uniquely determined by $p_n(j)$'s, then it follows that in the case, we consider, it is a df of $1 - \exp(-V)$, where V is a gamma$(r, p/(1-p))$ rv. Consequently $\hat{H}$ restricted to $(0, 1)$ is invertible, i.e. $G \equiv \hat{H}^{-1}(F)$ is unique. □

6.2 *The normal threshold*

Now keeping the sequence of Y's exponential with the parameter λ we consider the threshold X having the normal $\mathcal{N}(\lambda\sigma^2, \sigma^2)$ distribution. Observe that this time supports of F and G do not coincide. According to the formula given in the beginning of this section it follows that

$$P(K = j) = \begin{cases} \dfrac{\lambda^j}{j!} E(X^j e^{-\lambda X} I_{(0,\infty)}(X)), & j \geq 1 \\ E(I_{(-\infty,0]}(X)) + E(e^{-\lambda X} I_{(0,\infty)}), & j = 0. \end{cases}$$

Observe that

$$E(I_{(-\infty,0]}(X)) = P(X < 0) = 1 - \Phi(\lambda\sigma).$$

On the other hand for any $j = 0, 1, \dots$ we have

$$E(X^j e^{-\lambda X} I_{(0,\infty)}(X)) = \sigma^j \exp\left(-\frac{(\lambda\sigma)^2}{2}\right) J_j,$$

where

$$J_j = \frac{1}{\sqrt{2\pi}} \int_0^\infty x^j \exp\left(-\frac{x^2}{2}\right) dx.$$

Consequently $J_j = (j-1)J_{j-1}$ for any $j = 2, 3, \dots,$ $J_0 = 1/2$ and $J_1 = 1/\sqrt{2\pi}$. Hence

$$J_j = (j-1)!! \begin{cases} 1/2 & j = 2, 4, \dots \\ 1/\sqrt{2\pi} & j = 1, 3, \dots \end{cases}.$$

Finally we obtain

$$P(K = j) = \begin{cases} 1 - \Phi(\lambda\sigma) + 0.5\exp(-0.5(\lambda\sigma)^2) & j = 0 \\ \dfrac{(\lambda\sigma)^j}{j!!} e^{-0.5(\lambda\sigma)^2} \cdot \begin{cases} 1/2 \\ 1/\sqrt{2\pi} \end{cases} & \begin{matrix} j = 2, 4, \dots \\ j = 1, 3, \dots \end{matrix} \end{cases}.$$

This pmf can be used to obtain characterizations of the exponential and normal distributions, similarly as it was done in the last subsection for the exponential and gamma ones. Instead we will use it to give a quite surprising derivation of the Laplace (1812) expansion formula for the df Φ of the standard normal law. Observe that the given above form of the pmf implies

$$\begin{aligned} 1 &= \sum_{j=0}^{\infty} P(K = j) \\ &= 1 - \Phi(\lambda\sigma) + 0.5\exp(-0.5(\lambda\sigma)^2) + 0.5\sum_{i=1}^{\infty} \frac{(\lambda\sigma)^{2i}}{(2i)!!} \exp(-0.5(\lambda\sigma)^2) \\ &\quad + \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\infty} \frac{(\lambda\sigma)^{2i-1}}{(2i-1)!!} \exp(-0.5(\lambda\sigma)^2). \end{aligned}$$

Observe that combining the third and the fourth terms in the above sum we get

$$0.5\exp(-0.5(\lambda\sigma)^2)\sum_{i=0}^{\infty}\frac{(\lambda\sigma)^{2i}}{(2i)!!} - 0.5\exp(-0.5(\lambda\sigma)^2)\sum_{i=0}^{\infty}\frac{(0.5(\lambda\sigma)^2)^i}{i!} = 0.5.$$

Consequently for any $z > 0$ (hence also for any real z)

$$\Phi(z) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\exp(-0.5z^2)\sum_{i=0}^{\infty}\frac{z^{2i-1}}{(2i-1)!!},$$

which is exactly the celebrated Laplace formula, rediscovered in Polya (1949) (see also the comments in Kerridge and Cook (1976)).

Acknowledgements

The authors are greatful to the referees for their thorough and useful remarks which contributed to substantial improvements in the presentation of the paper.

References

Ahsanullah, M. (1995). *Record Statistics*, Nova Science Publishers, Commack.

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consequtive successes, *Ann. Inst. Statist. Math.*, **46**, 193–202.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, Wiley, New York.

Bairamov, I. G. (1997). Some distribution free properties of statistics based on record values and characterization of distributions through a record, *Journal of Applied Statistical Science*, **5**, 17–25.

Barndorff-Nielsen, O. (1965). Identifiability of mixtures of exponential families, *Journal of Mathematical Analysis and Applications*, **12**, 115–121.

Cacoullos, T. and Papageorgiou, H. (1982). Characterizing the negative binomial distribution, *J. Appl. Probab.*, **19**, 742–743.

David, H. A. (1981). *Order Statistics*, Wiley, New York.

Dziubdziela, W. and Kopocinski, B. (1976). Limiting properties of the k th record values, *Applicationes Mathematicae*, **15**, 187–190.

Engel, J. and Zijlstra, M. (1980). A characterization of the gamma distribution by the negative binomial distribution, *J. Appl. Probab.*, **17**, 1138–1144.

Epstein, B. (1954). Tables for the distribution of the number of exceedances, *Annals of Mathematical Statistics*, **25**, 762–768.

Fosam, E. B. and Sapatinas, T. (1995). Characterisations of some income distributions based on multiplicative damage models, *Austral. J. Statist.*, **37**, 89–93.

Grassia, A. (1977). On a family of distributions with arguments between 0 and 1 obtained by transformations of the gamma and derived distributions, *Austral. J. Statist.*, **19**, 108–114.

Gumbel, E. J. (1958). *Statistics of Extremes*, Columbia University Press, New York.

Gumbel, E. J. and Schelling, H. von (1950). The distribution of number of exceedances, *Annals of Mathematical Statistics*, **21**, 247–262.

Johnson, N. L., Kotz, S. and Kemp, A. (1992). *Univariate Discrete Distributions*, Wiley, New York.

Kerridge, D. F. and Cook, G. W. (1976). Yet another series for the normal integral, *Biometrika*, **63**, 401–403.

Laplace, P. S. (1812). *Théorie Analitique de Probabilités*, Vol. 1, Courcier, Paris.

Mohanty, S. G. (1994). Success runs of length k in Markov dependent trials, *Ann. Inst. Statist. Math.*, **46**, 777–796.

Pólya, G. (1949). Remarks on computing the probability integral in one and two dimensions, *Proc. Berkeley Symp. on Math. Statist. Probab.* (ed. J. Neyman), 63–78, University of California Press, Berkeley.

Prakasa Rao, B. L. S. (1992). *Identifiability in Stochastic Models*, Academic Press, Boston.

Ratnaparkhi, M. V. (1981). Some bivariate distributions of (X, Y) where the conditional distribution of Y, given X, is either beta or unit-gamma, *Statistical Distributions in Scientific Work*, Vol. 4 (eds. C. Taillie, G. P. Patil and B. A. Baldessari), 389–400, Reidel, Dordrecht.

Sapatinas, T. (1995). Identifiability of mixtures of power series distributions and related characterizations, *Ann. Inst. Statist. Math.*, **47**, 447–459.

Sarkadi, K. (1957). On the distribution of the number of exceedances, *Annals of Mathematical Statistics*, **28**, 1021–1022.

Siddiqui, M. M. (1970). Order statistics for a sample and extended sample, *Nonparametric Techniques in Statistical Inference* (ed. M. L. Puri), 417–423, Cambridge University Press, Cambridge.

Teicher, H. (1961). Identifiability of mixtures, *Annals of Mathematical Statistics*, **32**, 244–248.

Wesołowski, J. (1995). Bivariate discrete measures via a power series conditional distribution and a regression function, *J. Multivariate Anal.*, **55**, 219–229.

Wilks, S. S. (1941). On the determination of sample sizes for setting tolerance limits, *Annals of Mathematical Statistics*, **12**, 91–96.

Wilks, S. S. (1942). Statistical prediction with special references to the problem of tolerance limits, *Annals of Mathematical Statistics*, **13**, 400–409.

Wilks, S. S. (1959). Recurrence of extreme observations, *J. Austral. Math. Soc.*, **1**, 106–112.