ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN—POLONIA
VOL. XXXIII, 9 SECTIO A 1979

Instytut Matematyki, Uniwersytet Marii Curie-Skłodowskiej, Lublin

Zofia GRUDZIEŃ

On Distributions and Moments of i-th Record Statistic with Random Index

O rozkładach i momentach i-tej statystyki rekordowej z losowym indeksem

О распределениях и моментах и-той рекордной статистики со случайным индексом

INTRODUCTION

Let $\{X_n, n \geq 1\}$ be a sequence of independent random variables with a common absolutely continuous distribution function $F(x)$ and the density function $f(x)$, and let $X_1^{(n)} \leq X_2^{(n)} \leq \ldots \leq X_n^{(n)}$ denote order statistics of the sample $(X_1, X_2, \ldots, X_n)$.

By

$$Y_0^{(i)} = X_1^{(i)}, \quad Y_n^{(i)} = X_{L_i(n)}^{(L_i(n)+i-1)}, \quad n = 0,1,2,..; \quad i \geq 1$$

where

$$L_i(0) = 1$$

$$L_i(n+1) = \min\{j : X_{L_i(n)}^{(L_i(n)+i-1)} < X_j^{(j+i-1)}\}, \quad n = 0,1,2,..$$

we define a sequence of i-th record statistics.

Properties of the first record statistic (the case $i = 1$) has been studied in [2], [3], [4] and the case $i \geq 1$ has been considered e.g. in [1].

In this note we give the distribution and moments of i-th record statistic $Y_N^{(i)}$, where N is a random variable.

2. DISTRIBUTION OF RECORD STATISTIC

(i) N has a power series distribution. A random variable N is said to have the power series distribution (PSD), if the probability function of N is of the form

$$(1)\qquad p(k;\Theta) = P[N = k] = \frac{a(k)\,\Theta^k}{f(\Theta)} \quad \text{for } k \in T,$$

where $T \subset \mathbb{N} \cup \{0\}$, $a(k) \geq 0$, $f(\Theta) = \sum_{k \in T} a(k)\,\Theta^k$ for $\Theta \in \Omega = \{\Theta : 0 < \Theta < \rho\}$ - the parameter space, and ρ is the radius of convergence of the power series of $f(\Theta)$, and $\mathbb{N}$ denotes the set of all integers.

In what follows we write f_i for $f(y_i)$, F_j for $F(y_j)$ etc. and put

$$A_i(\Theta, F) = \sum_{k \in T} \frac{a(k)[i\Theta a_i(F)]^k}{k!}, \quad a_i(F) = \log(1 - F_i)^{-1}$$

$$M_k = \sum_{l=0}^{n-k} \binom{n-k}{l}(-1)^l E P^{k+l},$$

$$L_n^i(F) = \sum_{k=0}^{n} \binom{n}{k}[i a_i(F)]^k \frac{M_k}{k!},$$

$$M_k^* = \sum_{r=0}^{n-k} \binom{n-k}{r}(-1)^r \frac{1}{k + r + 1}, \quad S_k = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} E\Lambda^{k+r}$$

Under these denotations we prove the following lemma

LEMMA 1. Let $\{X_n, n \geq 1\}$ be a sequence of independent identically distributed random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ having a common absolutely continuous distribution function $F(x)$ and the density function $f(x)$. Suppose that N is a positive integer-valued random variable defined on the same probability space with the probability function given by (1). Then the density function of $Y_N^{(i)}$ is

$$g(y_i) = \frac{i(1 - F_i)^{i-1} f_i}{f(\Theta)} A_i(\Theta, F) \tag{2}$$

P r o o f. Let G denote the distribution function of $Y_N^{(i)}$ e.g.

$$G(y_i) = P\left[Y_N^{(i)} < y_i\right]$$

Put $H(y_i|k) = P[Y_N^{(i)} < y_i | N = k]$ and $h(y_i|k) = H'(y_i)$.
We have $G(y_i) = \sum_{k \in T} P\left[Y_N^{(i)} < y_i | N = k\right] P[N = k] =$
$= \sum_{k \in T} H(y_i|k) \, P[N = k]$.
Hence, we get

$$g(y_i) = \sum_{k \in T} h(y_i|k) P[N = k]$$

From [1], we have

$$h(y_i|k) = \frac{i}{k!} [- i \log(1 - F_i)]^k (1 - F_i)^{i-1} f_i$$

By (1), we obtain

$$g(y_i) = \sum_{k \in T} \frac{i}{k!} [- i \log(1 - F_i)]^k (1 - F_i)^{i-1} f_i \frac{a(k) \Theta^k}{f(\Theta)} =$$

$$= \frac{i(1 - F_i)^{i-1} f_i}{f(\theta)} A_i(\theta, F)$$

(ii) Particular cases.

It is known that (1) with $T = \{0,1,\dots,n\}$, $a(k) = \binom{n}{k}$, $f(\theta) = (1+\theta)^n$ $\theta = \frac{p}{q}$ where $0<p<1$, $p + q = 1$ reduces to the binomial distribution with parameters p and n.

If we put $T = \mathbb{N} \cup \{0\}$, $a(k) = (-1)^k \binom{-n}{k}$, $f(\theta) = (1-\theta)^{-n}$ $\theta = q$, $0<q<1$, then (1) is the negative binomial distribution with parameters q and n.

Putting in (1) $T = \mathbb{N} \cup \{0\}$, $a(k) = \frac{1}{k!}$, $f(\theta) = e^{\theta}$, $\theta = \lambda > 0$, we obtain the Poisson distribution with parameter λ.

In the case when $T = \mathbb{N} \cup \{0\}$, $a(k) = 1$, $f(\theta) = \frac{1}{1-\theta}$, $\theta = p$, $0<p<1$, (1) reduces to the geometric distribution with parameter p.

We have then

COROLLARY 1. If the random variable N has the binomial distribution with parameters p and n, then

$$g(y_i) = iq^n (1 - F_i)^{i-1} f_i \sum_{k=0}^{n} \frac{1}{k!} \binom{n}{k} \left[i \frac{p}{q} a_i(F)\right]^k$$

COROLLARY 2. If the random variable N has the negative binomal distribution with parameters q and n, then

$$g(y_i) = i(1 - q)^n (1 - F_i)^{i-1} f_i \sum_{k=0}^{\infty} \frac{[iqa_i(F)]^k}{k!} \binom{n+k+1}{k}$$

COROLLARY 3. If the random variable N has the Poisson distribution with parameter λ, then

$$g(y_i) = \frac{i(1-F_i)^{i-1}f_i}{e^{\lambda}} \sum_{k=0}^{\infty} \frac{[i\lambda a_i(F)]^k}{(k!)^2}$$

COROLLARY 4. If the random variable N has the probability function

$$(3) \qquad P[N = k] = -\frac{1}{\log p} \frac{(1-p)^k}{k}, \quad k = 1,2,\ldots, \quad 0<p<1$$

then

$$g(y_i) = \frac{i(1-F_i)^{i-1}f_i}{\log p^{-1}} \sum_{k=1}^{\infty} \frac{[i(1-p)a_i(F)]^k}{k \cdot k!}$$

COROLLARY 5. If the random variable N has the geometric distribution with parameter p, then

$$g(y_i) = i(1-p)(1-F_i)^{i(1-p)-1} f_i$$

(iii) N has the compound binomial and Poisson distribution

A random variable N is said to have the compound binomial distribution if the probability function of N is of the form

$$(4) \qquad p(k;P) = P[N = k] = \binom{n}{k} \int_0^1 p^k q^{n-k} f(p)dp,$$

$$k = 0,1,\ldots,n; \quad 0<p<1, \quad p+q = 1$$

where f(p) denotes the density function of the random variable P.

Using this definition we can prove the following lemma

LEMMA 2. If N is a random variable having the distribution (4), then the probability density function of $Y_N^{(i)}$ is given by

$$g(y_i) = i(1 - F_i)^{i-1} f_i L_n^i(F) \tag{5}$$

P r o o f. By (4) we have

$$g(y_i) = \sum_{k=0}^{n} h(y_i|k) P[N = k] =$$

$$= \sum_{k=0}^{n} \frac{1}{k!} [- i \log(1 - F_i)]^k (1 - F_i)^{i-1} f_i \binom{n}{k} \int_0^1 p^k q^{n-k} f(p)dp =$$

$$= i(1 - F_i)^{i-1} f_i \sum_{k=0}^{n} \binom{n}{k} [ia_i(F)]^k \frac{1}{k!} \sum_{r=0}^{n-k} \binom{n-k}{r} (-1)^r EP^{k+r}$$

Using the above denotations we obtain (5).

COROLLARY 1. If the random variable P has uniform distribution on (0,1) then

$$g(y_i) = i(1 - F_i)^{i-1} f_i \sum_{k=0}^{n} \frac{[ia_i(F)]^k}{k!} \binom{n}{k} M_k^*$$

COROLLARY 2. If the random variable P has the beta distribution e.g.

$$f(p) = \frac{p^a q^{b-a}}{B(a + 1, b - a + 1)}, \quad 0 < p < 1, \quad -1 < a < b + 1, \quad p + q = 1 \tag{6}$$

then

$$g(y_i) = i(1 - F_i)^{i-1} f_i \sum_{k=0}^{n} \frac{[ia_i(F)]^k}{k!} \binom{n}{k} \tilde{M}_k$$

where

$$\tilde{M}_k = \sum_{r=0}^{n-k} \binom{n-k}{r} (-1)^r \frac{\Gamma(k + r + a + 1)\Gamma(b + 2)}{\Gamma(a + 1)\Gamma(k + r + b + 2)}$$

Further we consider the case, where N has a compound Poisson distribution.

A random variable N is said to have the compound Poisson dostribution if the probability function of N is of the form

$$(7) \qquad p(k) = P[N = k] = \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} dG(\lambda), \qquad k = 0,1,2,\ldots$$

where G denotes distribution function of the parameter Λ.

We now give the formula for distribution of $Y_N^{(i)}$ in the case when N has the distribution (7). It can be easily seen, that in this case the following lemma is true

LEMMA 3. If N is a random variable having the distribution (7), then the density function of $Y_N^{(i)}$ is given by

$$(8) \qquad g(y_i) = i(1 - F_i)^{i-1} f_i K_i(F)$$

where

$$K_i(F) = \sum_{k=0}^{\infty} \frac{[ia_i(F)]^k}{(k!)^2} S_k$$

P r o o f. Taking into account that

$$\int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} dG(\lambda) = \frac{1}{k!} \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} E\Lambda^{k+r}$$

we get (8).

COROLLARY. If the random variable Λ is distributed according to

$$(9) \qquad f(\lambda) = \begin{cases} \dfrac{a^v}{\Gamma(v)} \lambda^{v-1} e^{-av} & \text{for } \lambda > 0 \\ 0 & \text{for } \lambda \leq 0 \end{cases}$$

where $a > 0$, $v > 0$, then

$$g(y_i) = i(1 - F_i)^{i-1} f_i \sum_{k=0}^{\infty} \left[\frac{ia_i(F)}{a}\right]^k \frac{1}{k!} S_k^*$$

and

$$S_k^* = \sum_{r=0}^{\infty} \binom{k+r}{k} \binom{v+k+r-1}{k+r} \frac{(-1)^r}{a^r}$$

2. MOMENTS OF RECORD STATISTICS

We now consider the case where the distribution of random variables $\{X_n, n \geq 1\}$ is the uniform distribution in (0,1), e.g.

$$F(x) = x, \quad x \in (0,1).$$

One can prove

LEMMA 4. <u>If</u> $i \geq 1$, $k \geq 1$ <u>and</u> $m \geq 1$ <u>are integers then</u>

$$\int_0^1 (1 - x)^{i-1} x^m [\log(1 - x)]^{k-1} dx = \Gamma(k) \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{k+r+1}}{(r + i)^k}$$

Using the above lemma we get the following

THEOREM 1. If $Y_N^{(i)}$ is the i-th record statistic of the sequence of independent random variables $\{X_n, n \geqslant 1\}$ with distribution function $F(x) = x, \; x \in (0,1)$ and N is a random variable distributed according to (1), then

$$g(y_i) = \frac{i(1-y_i)^{i-1}}{f(\theta)} \; \bar{D}_i(\theta, F) \tag{10}$$

where

$$\bar{D}_i(\theta, F) = \sum_{k \in T} \frac{a(k)\left[i\theta \log(1-y_i)^{-1}\right]^k}{k!}$$

Moreover, for $m \geqslant 1$, we have

$$E\left[Y_N^{(i)}\right]^m = \frac{i}{f(\theta)} \sum_{k \in T} a(k)\left[i\theta\right]^k \bar{E}_m(i,k) \tag{11}$$

where

$$\bar{E}_m(i,k) = \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^r}{(r+i)^{k+1}}$$

P r o o f. (10) follows directly from (2).
(11) can be obtained after using Lemma 4 by simple evaluations.

COROLLARY 1. If a random variable N has the binomial distribution with parameters p and n, then we have in the considered case

$$g(y_i) = iq^n(1-y_i)^{i-1} \sum_{k=0}^{n} \binom{n}{k} \frac{\left[i\,\frac{p}{q} \log(1-y_i)^{-1}\right]^k}{k!}$$

and for $m \geqslant 1$

$$E\left[Y_N^{(i)}\right]^m = iq^n \sum_{k=0}^{n} \binom{n}{k} \left[i\,\frac{p}{q}\right]^k \bar{E}_m(i,k)$$

COROLLARY 2. If in the considered case a random variable N has the negative binomial distribution with parameter p and n, then

$$g(y_i) = i(1-p)^n(1-y_i)^{i-1} \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{1}{k!}\left[ip \log(1-y_i)^{-1}\right]^k$$

and for $m \geq 1$

$$E\left[Y_N^{(i)}\right]^m = i(1-p)^n \sum_{k=0}^{\infty} \binom{n+k-1}{k}\left[ip\right]^k E_m(i,k)$$

COROLLARY 3. If in the considered case N has the Poisson distribution, then

$$g(y_i) = \frac{i(1-y_i)^{i-1}}{e^{\lambda}} \sum_{k=0}^{\infty} \frac{\left[i\lambda \log(1-y_i)^{-1}\right]^k}{(k!)^2}$$

and for $m \geq 1$

$$E\left[Y_N^{(i)}\right]^m = ie^{-\lambda} \sum_{k=0}^{\infty} \frac{[i\lambda]^k}{k!} E_m(i,k)$$

COROLLARY 4. If in the considered case N has the distribution (3), then

$$g(y_i) = \frac{i(1-y_i)^{i-1}}{\log p^{-1}} \sum_{k=1}^{\infty} \frac{\left[i(1-p)\log(1-y_i)^{-1}\right]^k}{k \cdot k!}$$

and for $m \geq 1$

$$E\left[Y_N^{(i)}\right]^m = \frac{1}{\log p^{-1}} \sum_{k=1}^{\infty} \frac{[i(1-p)]^k}{k} E_m(i,k)$$

COROLLARY 5. If in the considered case N has the geometric distribution with parameter p, then

$$g(y_i) = i(1-p)(1-y_i)^{i(1-p)-1}$$

and for $m \geq 1$

$$E[Y_N^{(i)}]^m = i(1-p) \sum_{k=0}^{\infty} \binom{i(1-p)-1}{k} \frac{(-1)^k}{k+m+1}$$

We are going to present the analogous results in the case when N has the compound binomial and Poisson distribution.

After using Lemma 2 we have

THEOREM 2. If $Y_N^{(i)}$ is the i-th record statistic of the sequence $\{X_n, n \geq 1\}$ of independent random variables with distribution function $F(x) = x$, $x \in (0,1)$ and N is a random variable distributed according to (4), then

$$g(y_i) = i(1-y_i)^{i-1} \sum_{k=0}^{n} \binom{n}{k} [-i \log(1-y_i)]^k \frac{1}{k!} M_k$$

and for $m \geq 1$

$$E[Y_N^{(i)}]^m = \sum_{k=0}^{n} i^{k+1} \binom{n}{k} \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{r+1}}{(r+1)^{k+1}} M_k$$

COROLLARY 1. If in the considered case P is uniformly distributed in $(0,1)$, then

$$g(y_i) = i(1-y_i)^{i-1} \sum_{k=0}^{n} \binom{n}{k} [-i \log(1-y_i)]^k \frac{M_k^*}{k!}$$

and for $m \geq 1$

$$E[Y_N^{(i)}]^m = \sum_{k=0}^{n} i^{k+1} \binom{n}{k} \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{r+1}}{(r+i)^{k+1}} M_k^*$$

COROLLARY 2. If in the considored case P has the beta distribution (6), then

$$g(y_i) = i(1 - y_i)^{i-1} \sum_{k=0}^{n} \binom{n}{k}[- i \log(1 - y_i)]^k \frac{\tilde{M}_k}{k!}$$

and for $m \geq 1$

$$E[Y_N^{(i)}]^m = \sum_{k=0}^{m} i^{k+1}\binom{n}{k} \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{r+1}\tilde{M}_k}{(r+i)^{k+1}}$$

REMARK. Using the relation

$$EP^k(1 - P)^{n-k} = \sum_{r=0}^{n-k} (-1)^r \binom{n-k}{r} EP^{k+r}$$

we can get the following identities

$$B(k + 1, n - k + 1) = \sum_{r=0}^{n-k} (-1)^r \binom{n-k}{r} \frac{1}{k + r + 1}$$

$$B(k+a+1, n-k+b-a+1) = \sum_{r=0}^{n-k} (-1)^r \binom{n-k}{r} B(k+r+a+1, b-a+1)$$

We are going to discuss the case when N has the compound Poisson distribution.

One can see that the following theorem is true.

THEOREM 3. If $Y_N^{(i)}$ is the i-th record statistic of a sequence of independent random variables $\{X_n, n \geq 1\}$ with the distribution function $F(x) = x$, $x \in (0,1)$ and N is a random variable distributed according to (7), then

$$g(y_i) = i(1 - y_i)^{i-1} \sum_{k=0}^{\infty} \frac{[-i \log(1 - y_i)]^k}{(k!)^2} S_k \tag{12}$$

and for $n \geq 1$, we have

$$E[Y_N^{(i)}]^m = \sum_{k=0}^{\infty} \frac{i^{k+1}}{k!} S_k \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{r+1}}{(r+i)^{k+1}} .$$

COROLLARY. If in the considered case, the random variable Λ is distributed according to (9), then

$$g(y_i) = i(1-y_i)^{i-1} \sum_{k=0}^{\infty} \left[\frac{-i\log(1-y_i)}{a}\right]^k \frac{1}{k!} S_k^*$$

and for $m \geqslant 1$, we have

$$E[Y_N^{(i)}]^m = \sum_{k=0}^{\infty} \frac{i^{k+1}}{a^k} \sum_{r=0}^{m} \binom{m}{r} \frac{(-1)^{r+1}}{(r+i)^{k+1}} S_k^*$$

REFERENCES

[1] Dziubdziela, W., Kopociński, B., Limiting properties of the k-th record values, Zastos. Mat., 15(1976), 187-190.

[2] Nagaraja, H.N., On expected values of record values, Austral. J. Statist., 20(2)(1978), 176-182.

[3] ,, , On characterization based on record values, Austral. J. Statist., 19(1)(1977), 70-73.

[4] Resnick, S.I., Extremal processes and record value times, J. Appl. Probab., 10(4)(1973), 864-868.

STRESZCZENIE

Niech $\{X_n, n \geqslant 1\}$ będzie ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie a $Y_N^{(i)}$, $i \geqslant 1$, $n = 0,1,2,..$ ciągiem i-tych statystyk rekordowych. W pracy badano rozkłady

$Y_N^{(i)}$ i ich momenty, gdzie N jest zmienną losową o dodatnich wartościach całkowitych. Rozważono między innymi, przypadki w których N ma rozkład dwumianowy, ujemny dwumianowy, Poissona, logarytmiczny i geometryczny a X_1 ma rozkład jednostajny.

Резюме

Пусть $\{X_n, n \geq 1\}$ - последовательность независимых одинаково распределенных случайных величин, а $Y_N^{(i)}$, $i \geq 1$, $n=0,1,2$ последовательность u-тых рекордных статистик.

В работе исследуются распределения $Y_N^{(i)}$ и их моменты,когда N случайная величина принимающая неотрицательные целые значения. Рассматриваются, среди других, случаи, в которых N имеет биноминальное, отрицательно биноминальное, Пуассона, логарифмическое и геометрическое распределение и X_1- равномерное распределение.