Z. Wahrscheinlichkeitstheorie verw. Gebiete
66, 173-183 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Characterization by Truncated Moments and Its Application to Pearson-Type Distributions

W. Glänzel, A. Telcs, and A. Schubert

Library of the Hungarian Academy of Sciences, Department for Informatics and Science Analysis, Budapest, POB 7, H-1361, Hungary

Summary. A characterization theorem based on the proportional relation between two truncated moments is proved for both continuous and discrete distributions. The results are applied for characterizing distributions of Pearson's system and its discrete analogon.

1. Introduction

Recently, growing interest has been focussed on characterization of both continuous and discrete distributions by truncated moments (Galambos and Kotz 1978). The most general results (Kotz and Shanbhag 1980) will be reformulated in our Propositions 2.1 and 3.1. However general, these characterizations reduce to a form simple enough for any kind of application only in case of a rather limited class of distributions. A possible way of overcoming this limitation is to use the relation between two different moments for characterizing the distribution.

In our paper, a characterization theorem based on a simple proportionality between two different moments truncated from the left at the same point will be proved. Applications to a wide class of continuous and discrete distributions of great practical significance will be presented.

2. The Continuous Case

Let $(\Omega, \mathscr{A}, P)$ be a given probability space and let X be a continuous random variable such that $X\colon \Omega \to H$, where $H = [0, a)$ for some $a \in \mathbb{R}_0^+$ or $H = \mathbb{R}_0^+$. For a given real function h defined on H we consider the function

$$E(h(X) \mid X \geqq x) = e_h(x); \qquad x \in H \tag{2.1}$$

provided it is defined. Assume that the distribution function $F(x)$ of the random variable X is absolutely continuous and let $f(x)$ be its density. Finally, put $G(x)=1-F(x)$. We give the following proposition without proof as it is a reformulation of the results of Kotz and Shanbhag (1980).

Proposition 2.1. *Let H and X be the same as above. Further let h be a continuous and monotonic function defined on H such that*

$$h|_{[y,\infty)\cap H} \not\equiv \text{const}$$

for any finite $y \in H$. Assume that

$$A(x)=\int_{[0,x]} \frac{e_h'}{e_h-h}\,d\lambda \tag{2.2}$$

is infinite for $x=a$ if H is bounded; else, if $H=\mathbb{R}_0^+$, $\lim_{x\to\infty} A(x)$ is infinite. Then the distribution function $F(x)$ of the random variable X is absolutely continuous and uniquely determined by e_h; particularly

$$F(x)=\begin{cases} 0 & \textit{for } x<0 \\ 1-\exp(-A(x)) & \textit{for } x\geqq 0 \ (\textit{if } H \textit{ is bounded then for } x\in[0,a)) \\ 1 & \textit{for } x\geqq a \textit{ if } H \textit{ is bounded.} \end{cases} \tag{2.3}$$

In the following we try to answer the question: under which conditions the distribution of a given random variable can be characterized by a simple proportional relation

$$e_g=\lambda_h^g e_h$$

between e_g and e_h, the truncated moments belonging to the real functions g and h defined on H in the sense of Eq. (2.1); λ_h^g is a real function defined on H.

In the following we use the notation $C_{sm}^2(H)$ for the class of two times continously differentiable and strictly monotonic functions defined on the set H.

Theorem 2.1. *Let $X:\Omega\to H$ be a continuous random variable with the distribution function F and let g and h be two real functions defined on H such that*

$$e_g=e_h\lambda_h^g \tag{2.4}$$

is defined. Assume that $g,h\in C^1(H)$, $\lambda_h^g\in C^2(H)$ and $F\in C_{sm}^2(H)$. Finally, assume that the equation

$$h\lambda_h^g=g \tag{2.5}$$

has no solution on int H. *Then F is uniquely determined by the functions g, h and λ_h^g.*

Proof. Both sides of Eq. (2.4) are differentiable by assumption. After differentiation and putting $\lambda_h^g=\lambda$ we have

$$\lambda'(x)\int_x^a h(t)\,f(t)\,dt=f(x)(\lambda(x)\,h(x)-g(x)), \tag{2.6}$$

where $a=\max H$ if H is bounded or $a=\infty$ if $H=\mathbb{R}_0^+$.

Because by assumption neither f nor $(\lambda h - g)$ can be zero on $\operatorname{int} H$, λ' cannot change its sign. Thus λ is a strictly monotonic function. On the other hand

$$e_h(x) \neq 0; \qquad x \in \operatorname{int} H$$

follows. After dividing Eq. (2.6) with λ' and after differentiation of the resulting equation

$$\frac{f'}{f} = \frac{\lambda''(\lambda h - g) - \lambda'(\lambda' h + \lambda h' - g') - \lambda'^2 h}{\lambda'(\lambda h - g)} \tag{2.7}$$

is obtained. The solution of Eq. (2.7) is

$$f(x) = C \left| \frac{\lambda'}{\lambda h - g} \right| e^{-\int_0^x \frac{\lambda' h}{\lambda h - g} dt}, \tag{2.8}$$

where the constant C has to be chosen so that $\int_0^a f(x)\, dx = 1$.

Remark 2.1. If $h \equiv \text{const}$, or more specially if $h \equiv 1$, then a version of Proposition 2.1 restricted to distribution functions of the class $C^2_{sm}(H)$ is obtained; the milder condition $e_g \neq g$ on $\operatorname{int} H$ substitutes the monotonicity condition of g in Proposition 2.1.

Remark 2.2. Let g, h and λ^g_h be real functions defined on H and assume that the conditions of Theorem 2.1 are satisfied. Let g^* and h^* be real functions defined on H such that

$$h^* = \alpha h + \beta; \qquad \alpha \in \mathbb{R} \setminus \{0\},\ \beta \in \mathbb{R}$$

and

$$g^* = \gamma g + \delta; \qquad \gamma \in \mathbb{R} \setminus \{0\},\ \delta \in \mathbb{R}.$$

Then Eq. (2.4) is equivalent with

$$e_{g^*} = \psi e_{h^*} + \varphi, \tag{2.9}$$

where

$$\psi = \frac{\gamma}{\alpha} \lambda^g_h \quad \text{and} \quad \varphi = -\beta\psi + \delta. \tag{2.10}$$

Applications

Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $X: \Omega \to H \subseteq \mathbb{R}^+_0$ be a continuous random variable with differentiable distribution function, where $H = [0, a]$ for $a \in \mathbb{R}^+_0$ or $H = \mathbb{R}^+_0$.

Definition 2.1. The distribution of a continuous random variable belongs to Pearson's system, if its density function f is differentiable and satisfies the following equation:

$$\frac{1}{f}\frac{df}{dx} = \frac{d \log f}{dx} = -\frac{x + A}{Bx^2 + Cx + D}; \qquad A, B, C, D \in \mathbb{R}. \tag{2.11}$$

Theorem 2.2. *Let X be a nonnegative random variable. Then X has a distribution belonging to Pearson's system if and only if the functions in Eq.* (2.10) *have the forms*

$$\varphi(t)=at+b; \qquad a>0,\ b>0$$

and

$$\psi(t)=ct; \qquad c+a_0b>0\ (a_0=\min(a,1))$$

for $r=-1$ and $s=1$, provided e^{-1} and e^1 exist and are differentiable. In the case $a>1$ the sum $c+b$ may be zero.

Proof. Assume that e^{-1} and e^1 exist and consider Eq. (2.10)

$$e^1(x)=cxe^{-1}(x)+ax+b. \tag{2.12}$$

It is clear that both sides of Eq. (2.12) are differentiable. Since φ and ψ are differentiable, the density function f is differentiable too. From Eq. (2.12) we get by repeated differentiation

$$\frac{d\log f(x)}{dx}=-\frac{(2a-1)\,x+c}{(a-1)\,x^2+(b+c)\,x} \tag{2.13}$$

i.e., under the given conditions the distribution of X is a member of Pearson's system indeed.

Assume that the distribution function of X belongs to Pearson's system. Then the density is differentiable and satisfies Eq. (2.11). Since $X\geqq 0$ we have $D=0$, i.e. $Z(x)=Bx^2+Cx+D=0$ (cf. Eq. (2.11)) has always the solution $x=0$. We can derive four types of distribution functions:

(i) $C>0$, $B\neq 0$ and one solution of $Z(x)$ is negative: infinite Beta distribution

(ii) $C>0$, $B\neq 0$ and one solution of $Z(x)$ is positive: finite Beta distribution

(iii) $C>0$, $B=0$, $A<C$: Gamma distribution

(iv) $C=0$, $B>1$ and $A<0$: a distribution with the density function $f(x)=K\cdot x^{-B}e^{-\left|\frac{A}{B}\right|\frac{1}{x}}$, where K is a positive value depending on A and B such that $\int_0^\infty f(x)\,dx=1$.

Assume that X has a distribution of type (i). Then the density function can be written in the form

$$f(x)=\frac{1}{B(\alpha,\beta)}\,\frac{\gamma^\alpha\cdot x^{\beta-1}}{(\gamma+x)^{\alpha+\beta}}; \qquad \alpha,\beta,\gamma>0. \tag{2.14}$$

Since the expectation exists, $\alpha>1$ has to be assumed. By definition,

$$e^1(x)=\frac{K\int_x^\infty \frac{t^\beta}{(\gamma+t)^{\alpha+\beta}}\,dt}{K\int_x^\infty \frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta}}\,dt}=\frac{\int_x^\infty \frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta-1}}\,dt}{\int_x^\infty \frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta}}\,dt}-\gamma$$

or, equivalently

$$-\gamma-\frac{\frac{1}{\beta}\frac{x^{\beta}}{(\gamma+x)^{\alpha+\beta-1}}}{\int\limits_x^{\infty}\frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta}}dt}+\frac{\alpha+\beta-1}{\beta}e^{-1}(x)=e^{1}(x);$$

on the other hand

$$1=\frac{\int\limits_x^{\infty}\frac{t^{\beta-2}(t+\gamma-\gamma)}{(\gamma+t)^{\alpha+\beta}}dt}{\int\limits_x^{\infty}\frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta}}dt}=-\gamma e^{-1}(x)-\frac{\frac{1}{\beta-1}\frac{x^{\beta-1}}{(\gamma+x)^{\alpha+\beta-1}}}{\int\limits_x^{\infty}\frac{t^{\beta-1}}{(\gamma+t)^{\alpha+\beta}}dt}+\frac{\alpha+\beta-1}{\beta-1}.$$

From the last two equations Eq. (2.12) can directly be derived with $a=\frac{\alpha}{\alpha-1}$, $b=\frac{\beta\gamma}{\alpha-1}$ and $c=-\frac{\beta-1}{\alpha-1}\gamma$. The validity of Eq. (2.12) for the types (ii)-(iv) can be shown similarly.

Discussion. Consider the following cases resulting from special choices of the parameters in Eq. (2.12); note that b has to be positive.

Case 1. $a>1$.

1.1. $c=-b$: Type (iv) of Theorem 2.2. The density function is

$$f(x)=\frac{\left(\frac{b}{a-1}\right)^{\frac{a}{a-1}}}{\Gamma\left(\frac{a}{a-1}\right)\cdot x^{\frac{a}{a-1}+1}}e^{-\frac{b}{a-1}\frac{1}{x}}.$$

1.2. $-b-c<0$: Type (i) of Theorem 2.2. The parameters of this infinite Beta distribution are $\alpha=\frac{a}{a-1}$, $\beta=\frac{b}{b+c}>1$ and $\gamma=\frac{b+c}{a-1}$ (cf. Eq. (2.14)).

1.3. $c=0$: Infinite Pareto distribution (a special case of the Beta distribution). The density function results from the preceding one with $\beta=\frac{b}{b+c}=1$:

$$f(x)=\frac{a}{a-1}\left(\frac{b}{a-1}+x\right)^{-\frac{a}{a-1}-1}.$$

1.4. $c>0$: Same as case 1.2 with $0<\frac{b}{b+c}<1$.

Cases 1.2.-1.4 coincide with type (i) of Theorem 2.2.

Case 2. $a=1$.

2.1. $-b<c<0$: Gamma distribution; its density function is

$$f(x)=\frac{x^{\left(\frac{b}{b+c}\right)-1}e^{-\frac{x}{b+c}}}{\Gamma\left(\frac{b}{b+c}\right)(b+c)^{\frac{b}{b+c}}}\quad\text{with}\quad\frac{b}{b+c}>1.$$

2.2. $c=0$: Exponential distribution (a special case of the Gamma distribution). The density is $f(x)=\frac{1}{b}\cdot e^{-\frac{x}{b}}$.

2.3. $c>0$: Same as case 2.1 with $0<\frac{b}{b+c}<1$.

Case 2 coincides with type (iii) of Theorem 2.

Case 3. $0<a<1$. Type (ii) of Theorem 2.2. Since $a<1$, we obtain a finite Beta distribution if $c>-a\cdot b$. The density takes the form

$$f(x)=\frac{x^{\left(\frac{b}{b+c}\right)-1}\left(\frac{b+c}{1-a}-x\right)^{\left[\frac{c+ab}{(1-a)(b+c)}\right]-1}}{\left(\frac{b+c}{1-a}\right)^{\frac{a}{1-a}}}.$$

Note that for $c=0$ the right-hand side of Eq. (2.12) does not depend on $e^{-1}(x)$.

3. The Discrete Case

Let $(\Omega,\mathscr{A},P)$ be a probability space and suppose in the following that X is a nonnegative random variable such that $X\colon\Omega\to H$ with $H=\{0,1,2,\dots,n\}$ for some $n\in\mathbb{N}$ or $H=\mathbb{N}_0$. Parallel to the continuous case, consider

$$E(h(X)|X\geqq k)=e_h(k);\qquad k\in H,\tag{3.1}$$

where h is a given function defined on H such that Eq. (3.1) is defined. Put $p_k=P(X=k)$, $F_k=P(X<k)$ and $G_k=1-F_k$ for each $k\in H$. The following proposition is a consequence of the results of Kotz and Shanbhag (1980).

Proposition 3.1. Let X and H be the same as above. Further let h be a monotonic function defined on H such that

$$h|_{\{m,m+1,\dots\}\cap H}\not\equiv\text{const}$$

for any finite $m\in H$. Assume that if H is finite and $n=\max H$ the equality $e_h(n)=h(n)$ holds; else, if $H=\mathbb{N}_0$, assume that

$$A(k)=\sum_{i=0}^{k-1}\log\frac{e_h(i)-h(i)}{e_h(i+1)-h(i)}\tag{3.2}$$

is infinite for $k\to\infty$. Then the distribution of the random variable X is uniquely determined by e_h; particularly,

$$p_k=\begin{cases}\left\{\prod\limits_{i=0}^{k-1}\dfrac{e_h(i)-h(i)}{e_h(i+1)-h(i)}\right\}\dfrac{e_h(k+1)-e_h(k)}{e_h(k+1)-h(k)} & \text{if } H=\mathbb{N}_0 \text{ or } H \text{ is finite and } k<n=\max H\\ \prod\limits_{i=0}^{n-1}\dfrac{e_h(i)-h(i)}{e_h(i+1)-h(i)} & \text{if } H \text{ is finite and } n=\max H.\end{cases}\tag{3.3}$$

To find another way for characterizing a discrete distribution we proceed in analogy to the continuous case from a simple proportional relation between two truncated moments, now in the form of the ratio:

$$\lambda_h^g = \frac{e_g}{e_h}.$$

Note that in the discrete case λ_h^g can be defined only if e_h does not vanish on H.

Theorem 3.1. *Let X: $\Omega \to H$ be a discrete random variable and let g and h be two functions defined on H such that*

$$\lambda_h^g = \frac{e_g}{e_h} \tag{3.4}$$

is defined. Assume that λ_h^g is strictly monotonic and $p_k \neq 0$ for each finite $k \in H$. Then the distribution of X is uniquely determined by the functions g, h and λ_h^g.

Proof. Put $\lambda = \lambda_h^g$ and $B_k = \sum_{i=k}^{n} h(i) p_i$, where $n = \max H$, if H is bounded, or $n = \infty$, if $H = \mathbb{N}_0$.

Consider the case $H = \{0, 1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. For $k < n$ the following two equations can be derived from Eq. (3.4) by elementary calculations:

$$h(k)\lambda(k+1) - g(k) = \{\lambda(k+1) - \lambda(k)\} \frac{B_k}{p_k} \tag{3.5}$$

and

$$h(k)\lambda(k) - g(k) = \{\lambda(k+1) - \lambda(k)\} \frac{B_{k+1}}{p_k}. \tag{3.6}$$

According to the assumptions the right hand sides of Eqs. (3.5) and (3.6) cannot be zero; thus the corresponding left hand sides cannot be zero either. Using Eq. (3.5) for $k = i+1 < n-1$ and Eq. (3.6) for $k = i < n-1$ we obtain the following result:

$$\frac{p_{k+1}}{p_k} = \frac{\lambda(k)h(k) - g(k)}{\lambda(k+2)h(k+1) - g(k+1)} \, \frac{\lambda(k+2) - \lambda(k+1)}{\lambda(k+1) - \lambda(k)}. \tag{3.7}$$

For $k = n-1$

$$\frac{p_n}{p_{n-1}} = \frac{\lambda(n-1)h(n-1) - g(n-1)}{\{\lambda(n) - \lambda(n-1)\}\, h(n)} \tag{3.8}$$

results directly from Eq. (3.6). p_0 has to be chosen so that $\sum_{i \in H} p_i = 1$. Consider now the case $H = \mathbb{N}_0$. Equation (3.7) is obviously valid for each finite $k \in H$. Consequently we have

$$A_k = \frac{p_k}{p_0} = \frac{\lambda(k+1) - \lambda(k)}{\lambda(1) - \lambda(0)} \prod_{i=1}^{k} \frac{\lambda(i-1)h(i-1) - g(i-1)}{\lambda(i+1)h(i) - g(i)}; \qquad k \in H$$

and $p_0 = \left(1 + \sum_{k=1}^{\infty} A_k\right)^{-1}$. This completes the proof.

Remark 3.1. If $h \equiv \text{const}$, a version of Proposition 3.1 analogous to that described in Remark 2.1 is obtained.

Applications

Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $X: \Omega \to H \subseteq \mathbb{N}_0$ be a discrete random variable, where $H = \{0, 1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$ or $H = \mathbb{N}_0$. For given $j \in \mathbb{Z}$ the j-th descending factorial is defined as

$$k^{(j)} = \begin{cases} 0 & \text{if } k < j \\ \dfrac{k!}{(k-j)!} & \text{if } k \geq j \end{cases}; \quad k \in \mathbb{N}_0.$$

The j-th descending factorial moment of the random variable X is defined as

$$E(X^{(j)}) = \sum_{i \in H} \frac{i!}{(i-j)!} \cdot P(X = i),$$

provided it exists. Consequently, we define the truncated j-th descending factorial moment of X as

$$E(X^{(j)} | X \geqq k) = \sum_{i \in H} \frac{i!}{(i-j)!} \cdot P(X = i | X \geqq k); \qquad k \in H,$$

provided the series at the right-hand side is convergent for all $k \in H$. Put $e^{(j)}(k) = e_h(k)$ for $h(k) = k^{(j)}$ and $j \neq 0$. Taking into account that the statement of Remark 2.2 is valid in the discrete case too, it is reasonable to consider the following equation

$$e^{(j)}(k) = e^{(i)}(k)\, \psi(k) + \varphi(k); \qquad k \in H, \tag{3.9}$$

where φ and ψ are functions according to Eqs. (2.10) such that Eq. (3.4) has sense.

Definition 3.1. We say that the distribution of a given discrete random variable X belongs to Irwin's system (Irwin 1975) if

$$p_{k+1} = \frac{(k+\beta)(k+\gamma)}{(k+\alpha+\beta+\gamma)(k+1)} p_k; \qquad k \in H \tag{3.10}$$

where α, β and γ are parameters such that the distribution has a sense.

Remark 3.1. Condition (3.10) is obviously equivalent with the following one:

$$\frac{p_{k+1} - p_k}{p_k} = -\frac{Ak + B}{k^2 + (C+1)k + C}; \tag{3.11}$$

$$A = \alpha + 1, \qquad B = \alpha + 1 - (1-\beta)(1-\gamma), \qquad C = \alpha + \beta + \gamma.$$

Thus we can consider Irwin's system as a discrete analogon to the nonnegative class of Pearson's system (cf. proof of Theorem 2.2). Note that the denominator of the right-hand side of Eq. (3.11) has always a root in $k=-1$. The following lemma (cf. Chen 1980) is given without proof.

Lemma 3.1. *Consider a distribution of Irwin's system as defined by Eq.* (3.10) *with such parameters that the distribution is infinite. Then* p_k *is characterized by the following property:*

$$k^{1+\alpha}p_k \to d \quad as \quad k\to\infty \qquad \alpha\in\mathbb{R}^+, \tag{3.12}$$

where d is a positive constant depending on the parameters but independent of k.

Theorem 3.2. *Let X be a nonnegative discrete random variable. Then X has a distribution belonging to Irwin's system if and only if the functions in Eq.* (3.9) *have the forms*

$$\varphi(k)=ak+b; \qquad a>0,\ b>0$$

and

$$\psi(k)=ck; \qquad c+a_0b+2a-1>0 \ \left(a_0=\min\left(2-\frac{1}{a},1\right)\right)$$

for $i=-1$ *and* $j=1$, *provided* $e^{(1)}$ *exists and the distribution has a sense.*

Proof. Note at first that $e^{(-1)}$ is defined for any nonnegative discrete distribution. Assume that $e^{(1)}$ exists too, and consider Eq. (3.9) in its actual form:

$$e^{(1)}(k)=cke^{(-1)}(k)+ak+b. \tag{3.13}$$

By definition, we obtain

$$\sum_{i\in H_k} ip_i=ck\sum_{i\in H_k}\frac{p_i}{i+1}+akG_k+bG_k,$$

where $H_k=H-\{0,1,2,\ldots,k-1\}$. By subtracting from this equation the analogous one for $k+1$, the following result is obtained:

$$k\cdot p_k=c\frac{k}{k+1}p_k-c\sum_{i\in H_{k+1}}\frac{p_i}{i+1}+akp_k-aG_k+bp_k.$$

Repeating the same procedure on the result

$$p_k\{(a-1)\cdot k^2+(b+c+a-1)\cdot k+b\}=p_{k+1}(k+1)\{(a-1)\cdot k+(b+c+2a-1)\}$$

follows, leading directly to

$$\frac{p_{k+1}-p_k}{p_k}=-\frac{(2a-1)\cdot k+(c+2a-1)}{(k+1)[(a-1)k+(c+2a-1)+b]}. \tag{3.14}$$

According to Remark 3.1, the distribution $\{p_k\}_{k\in H}$ belongs to Irwin's system, provided it has a sense.

Assume that the distribution of the random variable X is a member of Irwin's system. Let $k\in H$ be fixed and put $q_i=P(X=i|X\geqq k)$. Then Eq. (3.10) is obviously valid also for the pair (q_i, q_{i+1}), provided $i\geqq k$:

$$(i+\alpha+\beta+\gamma)(i+1)\,q_{i+1}-(\beta+i)(\gamma+i)\,q_i=0. \tag{3.15}$$

Suppose that the distribution is infinite. Since the expectation exists we have to assume that $\alpha>1$. Consider now Eq. (3.15) for the pair (q_k, q_{k+1}) and add to it the same equation for the pairs (q_{k+1}, q_{k+2}), (q_{k+2}, q_{k+3}), ... (q_{j-1}, q_j), where j is an arbitrary natural number with $j>k+1$. The resulting equation is:

$$(\alpha+\beta+\gamma-1)\sum_{i=k+1}^{j} iq_i+j^2q_j-k^2q_k-\beta\gamma\sum_{i=k}^{j-1} q_i-(\beta+\gamma)\sum_{i=k}^{j-1} iq_i=0. \tag{3.16}$$

If $j\to\infty$ each term of Eq. (3.16) is convergent because the expectation exists by assumption and according to Lemma 3.1,

$$\lim_{j\to\infty} j^2 q_j=\lim_{j\to\infty} j^2\frac{d}{j^{1+\alpha}}=\lim_{j\to\infty}\frac{d}{j^{\alpha-1}}=0.$$

Thus we have

$$(\alpha-1)\,e^{(1)}(k)=\beta\gamma+(\alpha+\beta+\gamma-1)\,kq_k+k^2q_k. \tag{3.17}$$

On the other hand, after dividing Eq. (3.15) by $i+1$ the following equation is obtained in analogous manner:

$$(\beta-1)(\gamma-1)\,e^{(-1)}(k)=\alpha-(\alpha+\beta+\gamma-1)\,q_k-kq_k. \tag{3.18}$$

Equations (3.17) and (3.18) lead directly to our statement:

$$e^{(1)}(k)=-\frac{(\beta-1)(\gamma-1)}{\alpha-1}\,ke^{(-1)}(k)+\frac{\alpha}{\alpha-1}k+\frac{\beta\gamma}{\alpha-1}. \tag{3.19}$$

Equation (3.19) is obviously valid for all $k\in H$.

Consider now the case $H=\{0,1,2,\ldots,n\}$. Equation (3.16) holds for $j<n$ in this case too. For $i=n$ Eq. (3.15) takes the following form:

$$-\beta\gamma q_n-(\beta+\gamma)\,nq_n-n^2q_n=0.$$

From this equation and Eq. (3.16) with $j=n-1$ Eq. (3.17) can be obtained. Equation (3.18) and consequently Eq. (3.19) can be derived similarly. Thus the proof is completed.

Remark 3.2. According to Theorem 3.2 the case $c+a_0b+2a=0$ can be considered as defined for $a=0.5$, if $2b\in\mathbb{N}$. Then $c=0$ and a discrete uniform distribution with $H=\{0,1,2,\ldots,2b\}$ and the expectation b is obtained.

Remark 3.3. Substituting the parameters β and γ by $\gamma'=\beta+\gamma-1$ and $\beta'=\frac{\beta\gamma}{\beta+\gamma-1}$ we obtain Eq. (3.19) in the form

$$e^{(1)}(k)=-\frac{\beta'-1}{\alpha-1}\,\gamma' ke^{(-1)}(k)+\frac{\alpha}{\alpha-1}k+\frac{\beta'\gamma'}{\alpha-1} \tag{3.20}$$

that corresponds directly to the continuous case (cf. proof of Theorem 2.2).

Discussion

Consider the following cases resulting from special choices of the parameters in Eq. (3.13); note that b has to be positive.

Case 1. $a>1$.

1.1. $-2a-b+1<c\leqq -a-b+1-2\sqrt{b(a-1)}$: This is possible only if $a>\sqrt{b(a-1)}$, i.e. $\alpha>2\sqrt{\beta\gamma}$. Then the quadratic expression $(a-1)k^2+(b+c+a-1)k+b$ has two positive roots. The condition for obtaining a distribution is that the smaller root is a natural number (extensions for real roots were given by Kemp and Kemp 1956). Under these conditions the distribution is a hygergeometric one.

1.2. $\max\{-2a-b+1, -a-b+1-2\sqrt{b(a-1)}\}<c<-a-b+1+2\sqrt{b(a-1)}$: This corresponds to the case when the parameters β and γ are complex conjugate (Irwin 1975).

1.3. $-a-b+1+2\sqrt{b(a-1)}\leqq c$: Inverse Pólya-Eggenberger distribution. The special case $c=0$ corresponds to a Waring distribution.

Case 2. $a=1$.

2.1. $-b-1<c<-b$: Binomial distribution if $\frac{b}{b+c}\in\mathbb{Z}$. The parameters are

$$n=-\frac{b}{b+c} \quad \text{and} \quad p=-(b+c).$$

2.2. $c=-b$: Poisson distribution with the parameter b.

2.3. $c>-b$: Negative binomial distribution, particularly

$$p_k=\binom{\frac{b}{b+c}+k-1}{k}\left(\frac{1}{b+c+1}\right)^{\frac{b}{b+c}}\left(\frac{b+c}{b+c+1}\right)^k.$$

Case 3. $a<1$.

3.1. $c>-2a-\left(2-\frac{1}{a}\right)b+1$: Pólya-Eggenberger (negative hypergeometric) distribution. Note that for $a<0.5c$ has to be positive.

3.2. $c=0$ and $a=0.5$: Discrete uniform distribution, provided $2b\in\mathbb{N}$ (cf. Remark 3.2).

References

Chen, W.-C.: On the Weak Form of Zipf's Law. J. Appl. Probability **17**, 611–622 (1980)

Galambos, J., Kotz, S.: Characterizations of Probability Distributions. Lecture Notes in Mathematics **675**. Berlin-Heidelberg-New York: Springer 1978

Irwin, J.O.: The Generalized Waring Distribution; Parts I, II, III. J.R. Statist. Soc. A **138**, 18–31, 204–227, 374–384 (1975)

Kemp, C.D., Kemp, A.W.: Generalized Hypergeometric Distributions. J.R. Statist. Soc. B **18**, 202–211 (1956)

Kotz, S., Shanbhag, D.N.: Some New Approaches to Probability Distributions. Adv. Appl. Probability **12**, 903–921 (1980)

Received April 28, 1983; in revised form August 10, 1983