Известия НАН Армении, Математика, том 53, н. 3, 2018, стр. 84 - 96.

ON THE PARAMETERS ESTIMATORS FOR A DISCRETE ANALOG OF THE GENERALIZED EXPONENTIAL DISTRIBUTION

D. FARBOD

Quchan University of Technology, Quchan, Iran[1]
E-mails: *d.farbod@qiet.ac.ir, d.farbod@gmail.com*

Abstract. The present paper is devoted to the estimation of parameters of the so-called *Discrete Analog of the Generalized Exponential Distribution* (DGED, in short), introduced by Nekoukhou et al. (Commun. Statist. Th. Meth., 2012). We derive conditions under which a solution for the system of likelihood equations exists and coincides with the maximum likelihood (ML) estimators of the DGED. An approach for approximate computation of the ML estimations of the unknown parameters, based on Fisher's accumulation method, is presented. A simulation study is also illustrated. Some statistical properties for two special cases of the DGED are provided. We also propose a linear regression-type model for estimation of the parameters. Finally, we fit the DGED to a real data set and the results we compare with those of two other discrete distributions.

MSC2010 numbers: 62F10, 62F12, 62J05, 62P10.

Keywords: Asymptotic properties; DGED, Fisher's accumulation method; Markov Chain Monte Carlo (MCMC); ML; Parametric function; Regression-type model.

1. Introduction

In this paper we consider the problem of estimation of parameters of the so-called *Discrete Analog of the Generalized Exponential Distribution* (DGED, in short), introduced by Nekoukhou et al. [10]. The DGED, which is a two-parameter discrete probability distribution, has some interesting statistical properties and is more flexible for modeling data compared with some well-known discrete distributions. It is of interest to study statistical inferences for this model. However, it should be noted that, the lack of the closed formulas for probability mass function (pmf) and cumulative distribution function (cdf) is a drawback to the use of DGED.

Nekoukhou et al. [10] considered some distributional properties of DGED, obtained ML estimators for the parameters with the help of Newton-Raphson algorithm, and established some properties for two special cases of DGED. In addition, they applied DGED for modeling rank frequencies of graphemes in the Slavic language (Slovene).

[1]This work was supported by Quchan University of Technology under the Grant # 9463.

The main purpose of this paper is to consider statistical inferences for DGED, including estimation of the unknown parameters by using ML method, investigation of statistical properties for some special cases, as well as, an application of the DGED to fit a real data set in biology and comparison with two other discrete distributions.

Model DGED. The probability mass function of DGED is given by the following formula (see [10]):

$$f_\theta(x) = \frac{p^x(1-p^x)^{\alpha-1}}{c_\theta}, \quad x = 1, 2, ..., \tag{1.1}$$

where

$$c_\theta = \sum_{y=1}^{\infty} p^y(1-p^y)^{\alpha-1} \tag{1.2}$$

is the *normalization* factor, $\theta = (p, \alpha)$ is an unknown parameter, and

$$\theta \in \Theta := \{\theta = (p, \alpha) : \ 0 < p < 1, \ \alpha > 0\}.$$

Throughout the paper, we will use the following notation. By $X^n = (X_1, ..., X_n)$ we will denote a random sample (independent and identically distributed random variables) from the distribution of a random variable X with the distribution given by (1.1), the corresponding observed sample (a realization) will be denoted by $x^n = (x_1, ..., x_n)$. Also, we will use the following notation:

$$\begin{cases} k(x;\theta) = \frac{1}{p}x - (\alpha-1)\frac{xp^{x-1}}{1-p^x}, \quad h(x;\theta) = \ln(1-p^x), \\ \eta(x;\theta) = \left(\frac{xp^{x-1}}{1-p^x}\right)^2, \quad \delta(x;\theta) = \frac{xp^{x-1}}{1-p^x}, \\ \psi(x;\theta) = \frac{1}{p^2}x + (\alpha-1)\frac{x(x-1)p^{x-2}}{1-p^x}. \end{cases}$$

The rest of the paper is organized as follows. In Section 2, we derive conditions under which a solution for the system of likelihood equations exists and coincides with the maximum likelihood (ML) estimators of model (1.1). In Section 3, we describe an approach for approximate computation of the ML estimators of unknown parameters for model (1.1), based on Fisher's accumulation method, supported with a simulation study. Section 4 contains some statistical properties for two important special cases of DGED. In Section 5, for a special case of DGED when p is known, we establish some properties of the estimator for a parametric function and employ a linear regression-type model to obtain an estimator for the parameter α. In Appendix, some applications of DGED are provided.

2. ML ESTIMATORS

In this section, we derive conditions under which a solution for the system of likelihood equations exists and coincides with the maximum likelihood (ML) estimators of the model (1.1).

Theorem 2.1. *The ML estimator of the parameter $\theta = (p, \alpha)$ of the model (1.1) based on a sample X^n is determined from the following system of moment equations:*

$$\begin{cases} E_\theta\big[k(\xi;\theta)\big] = \overline{k^n(\theta)} \\ E_\theta\big[h(\xi;\theta)\big] = \overline{h^n(\theta)} \end{cases} \tag{2.1}$$

where $\overline{k^n(\theta)} = \frac{1}{n}\sum_{i=1}^n k(x_i;\theta)$ and $\overline{h^n(\theta)} = \frac{1}{n}\sum_{i=1}^n h(x_i;\theta)$.

Proof. By (1.1) for the logarithm of likelihood function we have

(2.2)

$$l(X^n;\theta) = \ln L(X^n;\theta) = \ln \prod_{i=1}^n f_\theta(X_i) = \Big(\sum_{i=1}^n X_i\Big)\ln p + (\alpha-1)\sum_{i=1}^n \ln\big(1-p^{X_i}\big) - n\ln c_\theta.$$

So, the ML estimator of the parameter θ is a solution of the estimation equation:

$$\frac{\partial l(X^n;\theta)}{\partial \theta_i} = 0,\ i = 1,2,\ \ \theta_1 = p,\ \theta_2 = \alpha. \tag{2.3}$$

Differentiating (2.2) with respect to parameter p, we obtain

$$\frac{\partial l(X^n;\theta)}{\partial p} = \frac{1}{p}\sum_{i=1}^n X_i - (\alpha-1)\sum_{i=1}^n \frac{X_i\, p^{X_i-1}}{(1-p^{X_i})} - n\frac{1}{c_\theta}\frac{\partial c_\theta}{\partial p}, \tag{2.4}$$

where

$$\frac{1}{c_\theta}\frac{\partial c_\theta}{\partial p} = E_\theta\big[k(\xi;\theta)\big]. \tag{2.5}$$

From (2.3) - (2.5) we obtain the first equality in (2.1).

Next, differentiating (2.2) with respect to parameter α, we get

$$\frac{\partial l(X^n;\theta)}{\partial \alpha} = \sum_{i=1}^n \ln(1-p^{X_i}) - n\frac{1}{c_\theta}\frac{\partial c_\theta}{\partial \alpha}, \tag{2.6}$$

where

$$\frac{1}{c_\theta}\frac{\partial c_\theta}{\partial \alpha} = E_\theta\big[h(\xi;\theta)\big]. \tag{2.7}$$

From (2.3), (2.6) and (2.7) we obtain the second equality in (2.1). Theorem 2.1 is proved.

Now we proceed to prove that the solution $\widehat{\theta} = \widehat{\theta}_n = (\widehat{\theta}_i^n)_{i=1,2}$ of the system (2.2) (if it exists) is the ML estimator of the unknown parameter θ. To this end, we introduce the matrix of the second derivatives:

$$\widehat{Q}_{ij}^n = Q_{ij}^n(\widehat{\theta}),\ i,j = 1,2, \quad Q_{ij}^n(\widehat{\theta}) = \frac{\partial^2 l(X^n;\theta)}{\partial\theta_i\,\partial\theta_j}\Big|_{\theta=\widehat{\theta}},$$

and show that the matrix $\widehat{Q}^n = (\widehat{Q}^n_{i,j})_{i,j=1}^2$ is negative definite. We first prove two lemmas.

Lemma 2.1. *Suppose that the solution $\widehat{\theta}$ of the system (2.2) (if it exists) satisfies the following conditions:*

$$\left\{\begin{array}{l} E_\theta\big[\eta(\xi;\theta)\big] = \overline{\eta^n(\theta)} \\ E_\theta\big[\psi(\xi;\theta)\big] = \overline{\psi^n(\theta)} \\ E_\theta\big[\delta(\xi;\theta)\big] = \overline{\delta^n(\theta)} \end{array}\right. \tag{2.8}$$

where $\overline{\eta^n(\theta)} = \frac{1}{n}\sum_{i=1}^n \eta(x_i;\theta)$, $\overline{\psi^n(\theta)} = \frac{1}{n}\sum_{i=1}^n \psi(x_i;\theta)$ *and* $\overline{\delta^n(\theta)} = \frac{1}{n}\sum_{i=1}^n \delta(x_i;\theta)$. *Then the elements of the matrix $\widehat{Q}_n$ are given by*

$$\left\{\begin{array}{l} \widehat{Q}^n_{11} = -n\ Var_\theta\Big(k(\xi;\theta)\Big) \\ \widehat{Q}^n_{12} = \widehat{Q}^n_{21} = -n\ Cov_\theta\Big(k(\xi;\theta), h(\xi;\theta)\Big) \\ \widehat{Q}^n_{22} = -n\ Var_\theta\Big(h(\xi;\theta)\Big). \end{array}\right. \tag{2.9}$$

Proof. From (2.4) and (2.6) we obtain

$$\begin{aligned} Q^n_{11} &= \tfrac{\partial^2 l(X^n;\theta)}{\partial p^2} = -n\Big(\tfrac{1}{c_\theta}\tfrac{\partial^2 c_\theta}{\partial p^2} - \Big(\tfrac{1}{c_\theta}\tfrac{\partial c_\theta}{\partial p}\Big)^2\Big) - n\ \overline{\psi^n(\theta)} - n\ (\alpha-1)\ \overline{\eta^n(\theta)}, \\ Q^n_{12} &= Q^n_{21} = \tfrac{\partial^2 l(X^n;\theta)}{\partial\alpha\partial p} = \tfrac{\partial^2 l(X^n;\theta)}{\partial p\partial\alpha} = -n\Big(\tfrac{1}{c_\theta}\tfrac{\partial^2 c_\theta}{\partial\alpha\partial p} - \big(\tfrac{1}{c_\theta}\tfrac{\partial c_\theta}{\partial\alpha}\big)\big(\tfrac{1}{c_\theta}\tfrac{\partial c_\theta}{\partial p}\big)\Big) - n\ \overline{\delta^n(\theta)}, \\ Q^n_{22} &= \tfrac{\partial^2 l(X^n;\theta)}{\partial\alpha^2} = -n\Big(\tfrac{1}{c_\theta}\tfrac{\partial^2 c_\theta}{\partial\alpha^2} - \big(\tfrac{1}{c_\theta}\tfrac{\partial c_\theta}{\partial\alpha}\big)^2\Big). \end{aligned}$$

After some algebra and simplification we get

$$\begin{aligned} Q^n_{11} &= -n\ Var_\theta\Big(k(\xi;\theta)\Big) + n(\alpha-1)\Big(E_\theta\Big[\eta(\xi;\theta)\Big] - \overline{\eta^n(\theta)}\Big) + n\Big(E_\theta\big[\psi(\xi;\theta)\big] - \overline{\psi^n(\theta)}\Big), \\ Q^n_{12} &= Q^n_{21} = -n\ Cov_\theta\Big(k(\xi;\theta), h(\xi;\theta)\Big) + n\Big(E_\theta\Big[\delta(\xi;\theta)\Big] - \overline{\delta^n(\theta)}\Big), \\ Q^n_{22} &= -n\ Var_\theta\Big(h(\xi;\theta)\Big). \end{aligned}$$

Since by assumption the solution $(\widehat{p},\widehat{\alpha})$ of the system (2.1) satisfies the conditions (2.8), the result follows. Lemma 2.1 is proved.

Lemma 2.2. *Assume that the conditions in* (2.8) *are satisfied. Then, the matrix $\widehat{Q}^n$ with elements given by* (2.9) *is negative definite.*

Proof. It is enough to show that $\widehat{Q}^n_{11} < 0$ and $\det(\widehat{Q}^n) > 0$. In view of the first equality in (2.9), it is obvious that $\widehat{Q}^n_{11} < 0$. To establish that $\det(\widehat{Q}^n) > 0$, we write $\det(\widehat{Q}^n) = \widehat{Q}_{11}\widehat{Q}_{22} - (\widehat{Q}_{12})^2$. Now the inequality $\det(\widehat{Q}^n) > 0$ follows from (2.9) and Cauchy-Schwarz's inequality. Lemma 2.2 is proved.

As an immediate consequence of Lemmas 2.1 and 2.2 we have the following result.

Theorem 2.2. *If the solution of the system* (2.1) *satisfies the condition* (2.8), *then it coincides with the ML estimator of the parameter θ.*

3. Approximate computation of ML estimators

In Section 2, we have shown that the ML estimators of the unknown parameters of the model (1.1) coincide with the solution of the system (2.1). However, it is not easy to obtain a closed form for the solution (2.1). In this section, we propose an approach for approximate computation of the ML estimators by using Fisher's accumulation method. We refer the readers to [7] (p. 88) for details concerning Fisher's accumulation method.

Let $\theta(0) = \big(p(0), \alpha(0)\big)$ be an initial value of the parameter $\theta = (p, \alpha)$. Following [3], for $z = 0, 1, 2, \ldots$ we can use a recurrent formula to obtain $(z+1)^{th}$ approximation as follows:

$$\theta_j(z+1) = \theta_j(z) + \frac{\Upsilon_j(\theta(z))}{n \det I(\theta(z))}, \quad j = 1, 2; \quad \theta_1(0) = p(0), \ \theta_2(0) = \alpha(0), \tag{3.1}$$

where

$$I(\theta) = \begin{pmatrix} I_{11}(\theta) & I_{12}(\theta) \\ I_{21}(\theta) & I_{22}(\theta) \end{pmatrix}$$

is the Fisher's information matrix for one observation X_1 (put $X_1 = x$ and $n = 1$), and also

$$\Upsilon_1(\theta) = \begin{vmatrix} U_1(\theta) & I_{12}(\theta) \\ U_2(\theta) & I_{22}(\theta) \end{vmatrix}, \qquad \Upsilon_2(\theta) = \begin{vmatrix} I_{11}(\theta) & U_1(\theta) \\ I_{21}(\theta) & U_2(\theta) \end{vmatrix},$$

where $U_1(\theta) = \frac{\partial l(X^n;\theta)}{\partial p}$ and $U_2(\theta) = \frac{\partial l(X^n;\theta)}{\partial \alpha}$ are *contribution* functions, given by

$$U_1(\theta) = -n\Big[E_\theta\big[k(\xi;\theta)\big] - \overline{k^n(\theta)}\Big], \quad U_2(\theta) = -n\Big[E_\theta\big[h(\xi;\theta)\big] - \overline{h^n(\theta)}\Big].$$

Using formula (3.1), we introduce the following iterative algorithm (cf. [3]).

Algorithm.

1. Generate data based on Markov Chain Monte Carlo (MCMC) method.
2. Use (3.1) to calculate $\theta_j(z)$ for $j = 1, 2$; $z = 0, 1, 2, \ldots$.
3. If $|\theta_j(z+1) - \theta_j(z)| < \varepsilon$ (where ε is a small positive number), then $\theta_j(z+1) = \widehat{\theta}$ is the desired ML estimator, otherwise go to the step 2.

Simulation. In order to support the above stated theoretical results, we propose a simulation study. We apply the MCMC method to generate random samples from the model (1) (for details about MCMC, see [6]). To simplify the numerical calculations, we consider a truncated version of the random variable X, by restricting the possible values to 100 (cf. [3]).

Let $\theta_0 = (p = 0.6, \alpha = 3)$ be the true value of the parameter θ. We do simulation for 1000 times to illustrate the behavior of the ML estimators. Namely, for simulation study, we consider $M = 1000$ (M is the number of iteration), $N = 50, 100, 200$ (N is the sample size), and $\varepsilon = 0.0005$.

By the above described Algorithm and with the help of the *statistical software R*, the point estimates (means) and the mean square errors (MSE) are calculated and tabulated in Tables 1-3.

Table 1. For $N = 50$

	Mean	MSE
$p = 0.6$	0.9668756	0.1345977
$\alpha = 3$	3.24448654	17.6000100
Iteration	131	—

Table 2. For $N = 100$

	Mean	MSE
$p = 0.6$	0.8080627	0.04329008
$\alpha = 3$	1.1838108	3.29854332
Iteration	54	

Table 3. For $N = 200$

	Mean	MSE
$p = 0.6$	0.7667197	0.02779546
$\alpha = 3$	1.0895628	2.77626936
Iteration	18	—

From Tables 1-3, it is easily seen that with increasing sample size the MSE decreases.

4. Special case I

Nekoukhou et al. [10] showed that in the special cases where $\alpha = 2$ or $\alpha = 3$, the model (1.1) possesses some important statistical properties. Specifically, they proved that under some regularity conditions the ML estimators of the parameters of the model (1.1) are consistent and asymptotically normal.

In this section, we obtain more interesting properties of the estimators for these special cases. For simplicity, we consider the case $\alpha = 3$ (the case $\alpha = 2$ can be treated similarly), and denote $\theta = p := (p, 3)$. To state the main result of this section (Theorem 4.1), we first list the regularity conditions (cf. [2]).

C1. There exists a compact subset $\overline{H}$ of the parametric set Θ, which contains an open neighborhood of the true value p_0 of the parameter p;

C2. The distributions $\mathbb{P}_p$ are identifiable,, that is, $f_{p_1}(x) \neq f_{p_2}(x)$ for all $p_1 \neq p_2$ $(p_1, p_2 \in \overline{H})$ and for all $x \in Supp\ \mathbb{P}_p = \{x : f_p(x) > 0\}$;

C3. The function $f_p(x)$ is continuous in $p \in \Theta$, and has continuous first and second order derivatives in $p \in \overline{H}$;

C4. The distributions $\mathbb{P}_p$ have common support, namely the set $Supp\ \mathbb{P}_p$ does not depend on p;

C5. Put $V(x;p) := \frac{\partial^2 \ln f_p(x)}{\partial p^2}$. Then for $p \in \overline{H}$ and $x \in Supp\ \mathbb{P}_p$, there exists a function $G(x)$ (independent of p) such that $\left|V(x;p)\right| \leq G(x)$, and $E_p[G(X_1)] < \infty$;

C6. The Fisher information $I(p) = E_p\Big[\frac{\partial \ln f_p(X_1)}{\partial p}\Big]^2$ is continuous in p and satisfies the condition $0 < I(p) < \infty$.

Proposition 4.1. *The regularity conditions C1-C6 are satisfied for the model (1.1) with $\alpha = 3$.*

Proof. The conditions C1 – C4 and C6 can easily be verified. So, we have to verify only C5.

Since $\overline{H}$ is a compact set and $V(x;p)$ is continuous in $p \in \overline{H}$, for a fixed x and $p \in \overline{H}$ it can be concluded that $|V(x;p)|$ is bounded by a function $G(x)$, which itself is bounded for any fixed point x . We examine the behavior of $G(x)$ for sufficiently large x. We have

$$V(x;p) = \frac{f_p''(x)}{f_p(x)} - \Big(\frac{f_p'(x)}{f_p(x)}\Big)^2, \tag{4.1}$$

where

$$f_p'(x) = \frac{\partial f_p(x)}{\partial p} = \frac{xp^{x-1}(1-p^x)^2}{c_p} - \frac{2xp^x p^{x-1}(1-p^x)}{c_p} - \frac{c_p'}{(c_p)^2} p^x(1-p^x)^2, \quad c_p' = \frac{\partial c_p}{\partial p}, \tag{4.2}$$

and

$$\begin{aligned} f_p''(x) = \frac{\partial^2 f_p(x)}{\partial p^2} &= \frac{x(x-1)p^{x-2}(1-p^x)^2}{c_p} - \frac{2x^2 p^{2x-2}(1-p^x)}{c_p} - \frac{c_p'}{(c_p)^2} \cdot xp^{x-1}(1-p^x)^2 \\ &- \frac{2x(2x-1)\ p^{2x-2}(1-p^x)}{c_p} + \frac{2x^2 p^{2x-1}p^{x-1}}{c_p} + \frac{c_p'}{(c_p)^2} \cdot 2xp^{2x-1}(1-p^x) - \frac{c_p'}{(c_p)^2} \cdot xp^{x-1}(1-p^x)^2 \\ &+ \frac{c_p'}{(c_p)^2} \cdot \big(2xp^x p^{x-1}(1-p^x)\big) + \frac{(c_p')^2}{(c_p)^3} p^x(1-p^x)^2 - \frac{1}{c_p}\Big(\frac{c_p'}{c_p}\Big)' p^x(1-p^x)^2. \end{aligned} \tag{4.3}$$

Since $0 < p < 1$ and $x \geq 1$, we have $0 < p^x < 1$ and $\frac{1}{1-p^x} > 1$. Now, substituting (4.2) and (4.3) into (4.1) and using some calculations, a polynomial (based on x) is received for the $V(x;p)$. This polynomial may be considered of degree at most $\frac{x^2}{(1-p^x)^2}$. Therefore, for sufficiently large x, we obtain $G(x) = O\Big(\frac{x^2}{(1-p^x)^2}\Big)$. Also, it is easy to see that $E_p\big[\frac{X_1^2}{(1-p^{X_1})^2}\big] < \infty$. Thus, the condition C5 is satisfied. Proposition 4.1 is proved.

Theorem 4.1. *Under the regularity conditions C1-C6, the likelihood equation*

$$\frac{\partial l(X^n;p)}{\partial p} = 0$$

has a unique solution $\widehat{p}_n = \widehat{p}_n(X^n)$ in H (H is a subset of Θ whose closure $\overline{H}$ is also contained in Θ). Moreover, $\widehat{p}_n$ is a ML estimator for p and possesses the following properties:

(1). $\widehat{p}_n$ is consistent and asymptotically normal;

(II). $\widehat{p}_n$ is asymptotically efficient, that is, we have

$\widehat{u}_n = \sqrt{n}(\widehat{p}_n - p) \xrightarrow{d} \xi \in N\big(0, I^{-1}(p)\big)$, ($\xrightarrow{d}$ *means convergence in distribution*).

(III). For the moments of $\widehat{p}_n$ we have

$$E_p\big[\widehat{u}_n^k\big] \longrightarrow E_p\big[\xi^k\big], \quad \text{for any} \quad k \geq 1 \tag{4.4}$$

According to (4.4), *for $k = 1$ the property of asymptotic unbiasedness is also satisfied, namely we have*

$$E_p[\widehat{p}_n] = p + o\left(\frac{1}{\sqrt{n}}\right).$$

(IV). If $\phi(p)$ is a differentiable function on $\mathbb{R}$ such that $\phi'(p) \neq 0$, then

$$\sqrt{n}\big(\phi(\widehat{p}_n) - \phi(p)\big) \xrightarrow{d} \xi \in N\Big(0, \frac{[\phi'(p)]^2}{I(p)}\Big), \tag{4.5}$$

where $\phi'(p) = \frac{\partial \phi(p)}{\partial p}$.

Proof. The assertions of the theorem follow from Proposition 4.1 and a general theorem of mathematical statistics (see [2], [5]).

Remark 4.1. For $k = 2$ from (4.4) we obtain

$$E_p\big(\widehat{p}_n - p\big)^2 = \frac{1 + o(1)}{nI(p)}.$$

Also, the relation (4.5) can be stated in the following form:

$$E_p\big(\phi(\widehat{p}_n) - \phi(p)\big)^2 = \frac{[\phi'(p)]^2}{nI(p)}\big(1 + o(1)\big).$$

In the examples that follow we consider two special parametric functions of p, and use Theorem 4.1 to establish some statistical properties for the corresponding estimators.

Example 4.1. Let $\phi(p) = f_p(x) = \frac{p^x(1-p^x)^2}{c_p}$, where $x \in \mathbb{N}$ is fixed. In view of Theorem 4.1, we have

$$f_{\widehat{p}_n}(x) \xrightarrow{\mathbb{P}} f_p(x), \quad (\xrightarrow{\mathbb{P}} \textit{ means convergence in probability}).$$

From property (IV) of Theorem 4.1, we have

$$\sqrt{n}\Big(f_{\widehat{p}_n}(x) - f_p(x)\Big) \xrightarrow{d} N\Big(0, \frac{[f_p'(x)]^2}{I(p)}\Big).$$

Now we evaluate $f_p'(x)$. To this end, observe that

$$\frac{c_p'}{c_p} = E_p\Big[U(p^{X_1}(1-p^{X_1})^2)\Big], \tag{4.6}$$

where

$$U\big(p^x(1-p^x)^2\big) = \frac{1}{p^x(1-p^x)^2} \cdot \frac{\partial(p^x(1-p^x)^2)}{\partial p}$$

is the so-called *contribution* function of x. Hence, from (4.6) we get

$$f'_p(x) = f_p(x)\Big[U\big[p^{X_1}(1-p^{X_1})^2\big] - E_p\big[U(p^{X_1}(1-p^{X_1})^2)\big]\Big].$$

Next, in view of property (III), for $k=1$ we have $E_p\big[f_{\widehat{p}_n}(x)\big] = f_p(x) + o\Big(\frac{1}{\sqrt{n}}\Big)$, and for $k=2$, we get

$$E_p\big[f_{\widehat{p}_n}(x) - f_p(x)\big]^2 = \frac{[f'_p(x)]^2}{nI(p)} + o\Big(\frac{1}{n}\Big). \tag{4.7}$$

Finally, using the following formula

$$E_p\big(f_{\widehat{p}_n}(x) - f_p(x)\big)^2 = Var_p\big(f_{\widehat{p}_n}(x)\big) + \Big(E_p[f_{\widehat{p}_n}(x)] - f_p(x)\Big)^2,$$

from (4.7) we obtain

$$Var_p\big[f_{\widehat{p}_n}(x)\big] \sim \frac{[f'_p(x)]^2}{nI(p)}.$$

Example 4.2. Let $\phi_t(p) = \overline{F}_p(t) \equiv 1 - F_p(t) = \sum_{x=t}^{\infty} f_p(x)$ for all $t \in (0,\infty)$. Applying Theorem 4.1, we get

$$\overline{F}_{p_n}(t) \xrightarrow{\mathbf{P}} \overline{F}_p(t), \text{ for all } t \in (0,\infty).$$

From property (IV) of Theorem 4.1, we have

$$\sqrt{n}\big(\overline{F}_{\widehat{p}_n}(t) - \overline{F}_p(t)\big) \xrightarrow{d} N\Big(0, \frac{[\overline{F}'_p(t)]^2}{I(p)}\Big),$$

where

$$\begin{aligned}\overline{F}'_p(t) = \sum_{x=t}^{\infty} f'_p(x) &= E_\alpha\Big[U\big(p^{X_1}(1-p^{X_1})^2\big)\Big]\cdot 1_{(X_1\geq t)} - E_p\Big[U\big(p^{X_1}(1-p^{X_1})^2\big)\Big]\cdot \overline{F}_p(t)\\ &= \Big(1_{(X_1\geq t)} - \overline{F}_p(t)\Big)\cdot E_p\Big[U\big(p^{X_1}(1-p^{X_1})^2\big)\Big],\end{aligned}$$

and 1_A is the indicator function of a set A.

Next, using property (III) of Theorem 4.1, for the cases $k=1$ and $k=2$, we obtain

$$E_p\big[\overline{F}_{\widehat{p}_n}(t)\big] = \overline{F}_p(t) + o\Big(\frac{1}{\sqrt{n}}\Big),$$

$$E_p\big[\overline{F}_{\widehat{p}_n}(t) - \overline{F}_p(t)\big]^2 = \frac{[\overline{F}'_p(t)]^2}{n\,I(p)} + o\Big(\frac{1}{n}\Big).$$

Therefore, we have

$$Var_p\big(\overline{F}_{\widehat{p}_n}(t)\big) \sim \frac{[\overline{F}'_p(t)]^2}{n\,I(p)}.$$

Remark 4.2. The results obtained in Proposition 4.1, Theorem 4.1 and Examples 4.1 and 4.2 can also be stated for $\alpha = 2$ (generally for any α).

5. Special case II

In this section, for a special case of DGED when p is known, we establish some properties of the estimators of a parametric function of unknown α, and employ a linear regression-type model to obtain an estimator for the parameter α. Specifically, we find ML estimator and uniformly minimum variance unbiased estimator (UMVUE) for a parametric function of α. Also, we fit a linear regression-type model and then propose the least square (LS) estimator for the parameter α.

Estimation of a parametric function. Consider DGED model (1.1) with known p and unknown α. Here we are interested in the estimation of the following parametric function:

$$\tau(\alpha) = \frac{c_\alpha'}{c_\alpha}, \tag{5.1}$$

where $c_\alpha' = \frac{\partial c_\alpha}{\partial \alpha}$, and c_α is as in (1.2). As an estimator of function $\tau(\alpha)$, based on a sample $X^n = (X_1, ..., X_n)$ from (1.1), we consider the statistic:

$$M(X^n) = \frac{1}{n}\sum_{i=1}^{n} \ln(1 - p^{X_i}). \tag{5.2}$$

Theorem 5.1. *Under the regularity conditions C3, C4 and C6, the statistic $M(X^n)$ defined by* (5.2) *is UMVUE and also an efficient estimator for the parametric function $\tau(\alpha)$ defined by* (5.1).

Proof. The density $f_\alpha(x)$ we write in exponential form (see (1.1)):

$$f_\alpha(x) = \exp\Big\{(\ln p)x + (\alpha - 1)\ln(1 - p^x) - \ln c_\alpha\Big\}. \tag{5.3}$$

It follows from (5.3) that the statistic $\sum_{i=1}^{n} \ln(1 - p^{x_i})$ is a complete sufficient statistic for the model (1.1) when p is known. Hence applying the well-known theorems of statistics (see [7] and [2], Ch. II, Sec. 26), we conclude that the statistic $M(X^n)$ defined by (5.2), is the UMVUE and also an efficient estimator for the parametric function $\tau(\alpha)$ defined by (5.1). Theorem 5.1 is proved.

Remark 5.1. It follows from Theorem 5.1 and Theorem 26.2 of [2] that the statistic $M(X^n)$ is ML estimator for parametric function $\tau(\alpha)$. Therefore we can conclude that $M(X^n)$ is consistent and asymptotically normal estimator for $\tau(\alpha)$.

A regression-type model. In this subsection, we explore a linear regression-type method for the model (1.1), and then provide a LS estimator for the unknown parameter α. To this end, we first take logarithm from both sides of (1.1) to obtain

$$\ln f_\alpha(x) = \ln p^x + (\alpha - 1)\ln(1 - p^x) - \ln c_\alpha. \tag{5.4}$$

Obviously, we have (with $x = x_i$)

$$f_\alpha(x_i) = F_\alpha(x_i) - F_\alpha(x_{i-1}), \quad i = 1, 2, ..., n. \tag{5.5}$$

Substituting (5.5) into (5.4), we get

$$\ln\left(F_\alpha(x_i) - F_\alpha(x_{i-1})\right) = \ln p^{x_i} + (\alpha - 1)\ln(1 - p^{x_i}) - \ln c_\alpha. \tag{5.6}$$

Observe that the relation (5.6) cannot be used to fit a regression-type model, because its left-hand side depends on the unknown parameter α. To solve the problem, we use the empirical distribution function (edf):

$$F_n(x_i) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le x)},$$

and consider the variables (cf. [4]):

$$\zeta_i = \ln\left(F_n(x_i) - F_n(x_{i-1})\right) + \ln\left(\frac{1 - p^{x_i}}{p^{x_i}}\right) = \left(\ln(1 - p^{x_i})\right)\alpha + \beta,$$

where $\beta = -\ln c_\alpha$.

Now, we can suggest the estimation of α by regressing $\zeta_i = \left(\ln(1 - p^{x_i})\right)\alpha + \beta$ on $\ln(1 - p^{x_i})$ as follows:

$$\zeta_i = \left(\ln(1 - p^{x_i})\right)\alpha + \beta + e_i, \tag{5.7}$$

where $e_i \sim N(0, \sigma^2)$, $i = 1, 2, ..., n$ and $x = (x_1, ..., x_n)$ is an observed (non-random) sample. Thus, we can use (5.7), to estimate the parameter α by regressing ζ_i on $\ln(1 - p^{x_i})$.

Note that in [8] was used a different method, based on empirical characteristic function. It is of interest to consider the LS estimator for the model (5.7). As an unbiased LS estimator $\tilde{\alpha}$ of the parameter α we consider the statistic:

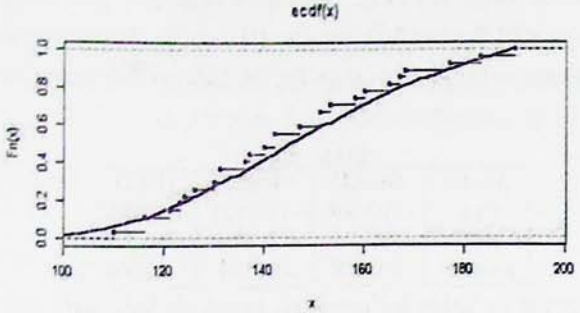
$$\tilde{\alpha} = \frac{\sum_{i=1}^{n}\left(\ln(1 - p^{x_i}) - \overline{\ln(1 - p^{x})}\right)\cdot\left(\zeta_i - \overline{\zeta}\right)}{\sum_{i=1}^{n}\left(\ln(1 - p^{x_i}) - \overline{\ln(1 - p^{x})}\right)^2}. \tag{5.8}$$

As an example, consider the data set $1, 9, 23, 17, 13, 12, 10, 9, 9, 3, 6$ (see [7]). Taking $p = 0.7$, and using (5.8), for this data set we obtain the point estimate $\tilde{\alpha} = 4.995138$.

For the considered model, we have the following result.

Theorem 5.2. *The LS estimator $\tilde{\alpha}$ of the unknown parameter α is consistent, asymptotically normal and is the best unbiased linear estimator.*

Proof. The proof is similar to that of Theorem 1 of [4], and so is omitted.



PIC. 1. Fitting of the truncated DGED to the data of Table 4. The dashed line is the ecdf of data and the solid line is the fitted cdf.

6. APPENDIX

As it was pointed out in Introduction, Nekoukhou et al. [10], fitted the DGED for modeling rank frequencies of graphemes in a Slavic language (Slovene). Here, we fit the DGED with a real data set in biology. In addition, we compare the DGED with two other discrete distributions. To this end, we first consider the following example.

Numerical Example. The data in the following table represent the systolic blood pressures (mm HG) for 27 women at age group 45-74 years old (see [11]).

For this data set, the ML estimates for the parameters p and α, the maximized log-likelihood ($\ln L$), the Akaike information criterion (AIC) and the p-value can be calculated to obtain the following numbers:

$$\widehat{p} = 0.9666823, \quad \widehat{\alpha} = 121.4182461, \quad \ln L = -119.550,$$

$$AIC = 243.100, \quad p - value = 0.9543.$$

Moreover, we can run an informal goodness of fit test (see the plots of the edf and fitted cdf of the systolic blood pressure data in Figure 1).

Now, we compare the DGED with *Discrete Distribution Generated by Levy's Density (DLD)* (see [4], Eq. (2)) and *Power-Law Distribution (PLD)* (see [12], Eq. (1)). Notice that the DLD and PLD are unimodal discrete distributions (supported on the set of natural numbers $\mathbb{N}$), which can be used for modeling phenomena arising, for example, in biology.

Table 4

110	116	121	126	131	136	142	151	165
116	124	131	137	142	151	158	168	183
123	130	140	147	153	160	167	177	190

For data given in Table 4, can be calculated lnL, AIC and the p-values when fitting the data using DLD and PLD (see also [1], Sec. 6). The corresponding results, together with the above obtained results for DGED, are tabulated in the following table.

Table 5

Model	DGED	DLD	PLD
$\ln L$	-119.550	-124.921	-127.391
AIC	243.100	251.842	256.783
p-value	0.9543	0.1964	0.1006

The results presented in Table 5 show that, based on lnL, AIC and the p-values, the DGED provides a better fit than DLD and PLD. Figure 1 shows a good fit for the DGED as well.

Acknowledgment. The author is thankful to the anonymous referee for valuable comments and suggestions, which led to considerable improvement in the original manuscript.

Список литературы

[1] H. Bidram, M. H. Alamatsaz, V. Nekoukhou, "On an extension of the exponentiated Weibull distribution", Communications in Statistics - Simulation and Computation, 44, 1389 – 1404 (2015).

[2] A. A. Borovkov, Mathematical Statistics, Gordon and Science Breach Publishing, translated from original Russian edition, (1998).

[3] D. Farbod, K. Gasparian, "On the maximum likelihood estimators for some generalized Pareto-like frequency distribution", Journal of the Iranian Statistical Society (JIRSS), 12, no. 2, 211 – 233 (2013).

[4] D. Farbod, "Some statistical inferences for two frequency distributions arising in bioinformatics", Applied Mathematics E-Notes, 14, 151 – 160 (2014).

[5] D. Farbod, "Asymptotic properties of maximum likelihood estimators for a generalized Pareto-type distribution", Journal of Contemporary Mathematical Analysis, 50, no. 1, 44 – 51 (2015).

[6] G. H. Givens, J. A. Hoeting, Computational Statistics, Wiley and Sons (2005).

[7] G. I. Ivchenko, Yu. Medvedev, Mathematical Statistics, Mir Press, Moscow (1990).

[8] I. A. Koutrouvelis, "Regression type estimation of the parameters of stable laws", Journal of American Statistical Association (JASA), 75, 918 928 (1980).

[9] K. Mohammad, H. Malekafzali, V. Nahapetian, Statistical Methods and Health Indices, (in Persian) (1994).

[10] V. Nekoukhou, M. H. Alamatsaz, and H. Bidram, "A discrete analog of the generalized exponential distribution", Communications in Statistics - Theory and Methods, 41, 2000 – 2013 (2012).

[11] S. Port, L. Demer, R. Jennrich, D. Walter, A. Garfinkel, "Systolic blood pressure and mortality", The Lancet, 355, no. 9199, 175 – 180 (2000).

[12] A. Rzhetsky, Sh. M. Gomez, "Birth of scale-free molecular networks and the number of distinct DNA and protein domains per genome", Bioinformatics, 17, no. 10, 988 – 996 (2001).

Поступила 10 января 2017