Jordan Journal of Mathematics and Statistics (JJMS) 7(3), 2014, pp. 247- 256

A GENERALIZED BIVARIATE GEOMETRIC DISTRIBUTION BASED ON AN URN MODEL WITH STOCHASTIC REPLACEMENT

RAMA SHANKER[(1)] AND A. MISHRA[(2)]

ABSTRACT: A Generalized Bivariate Geometric Distribution (GBGD) for explaining data arisen from four-fold sampling has been obtained through an urn-model with stochastic replacement. The marginal distributions of this generalized distribution, as in the case of the Bivariate Geometric Distribution (BGD), are the geometric distributions, but its one of the conditional distributions is the Consul's (1974) Quasi Binomial Distribution (QBD), in place of binomial distribution in the BGD. The moments of the first and second orders of the GBGD have been obtained. As the QBD has been found to possess tremendous capability to fit to discrete data-sets of various nature, it is expected that the obtained GBGD would cover a wide range of data-sets.

1. INTRODUCTION

Suppose there is a population consisting of individuals each having one or both or none of the two characters, M and N. Suppose that the respective probabilities for an individual to possess M only, N only and both the characters M and N are θ_{10}, θ_{01} and θ_{11}. The probability that an individual possesses neither M nor N is $\theta_{00} = 1 - \theta_{10} - \theta_{01} - \theta_{11}$.

Suppose that the individuals are observed one after another from this population till the first individual possessing none of the characters M and N is observed. If X_1, X_2 and X denote the number of individuals possessing M, N, and both respectively, then the

2000 Mathematics Subject Classification. 62E05, 62E99.

Keywords and phrases. Urn –Model, Marginal and Conditional Distributions, Generalized Bivariate Geometric Distribution, Drawing with Replacement, Quasi Binomial Distribution, Additional Parameters

Received: July 14, 2013 Accepted: August 11, 2014

probability distribution of X_1 and X_2 is given by the following bivariate geometric distribution

$$P_1(X_1=x_1, X_2=x_2) = \sum_{x=0}^{\min(x_1,x_2)} \frac{\Gamma(x_1+x_2-x+1)}{\Gamma(x_1-x+1)\Gamma(x_2-x+1)\Gamma(x+1)} \theta_{10}^{\;x_1-x}\,\theta_{01}^{\;x_2-x}\,\theta_{11}^{\;x}\,\theta_{00}$$

$$x_i = 0,1,2,\ldots\;;\; i=1,2. \tag{1.1}$$

The marginal distributions of X_1 and X_2 in (1.1) are geometric distributions with parameters $\theta_{00}/(1-\theta_{01})$ and $\theta_{00}/(1-\theta_{10})$ respectively. Assuming $x_1 < x_2$, the conditional distribution of X given x_1 and $\theta_{11}/(\theta_{10}+\theta_{11})$ is binomial and that of $X_2 - X$, given $1+x_1$ and $\theta_{01}/(\theta_{01}+\theta_{11})$, is negative binomial [See Patil and Joshi (1968)].

Consul (1974), through an urn model, obtained a three-parameter 'quasi binomial distribution' (QBD) given by its probability function

$$P_2(x;n,p_1\,p_2) = \binom{n}{x} p_1\,(p_1+x\,p_2)^{x-1}\,(1-p_1-x\,p_2)^{n-x}$$

$$0<p_1<1;\; -p_1/n < p_2 < (1-p_1)/n\;;\; x=0,1,2,3,\ldots \tag{1.2}$$

The classical binomial distribution is a particular case of it and can be obtained taking $p_2 = 0$. Mishra and Singh (2000) obtained the first four moments of this distribution in terms of certain series. The first two moments of this distribution have been obtained as

$$\mu_1^{/} = np_1 \sum_{i=0}^{n-1} (n-1)^{(i)}\, p_2^{\;i} \tag{1.3}$$

$$\mu_2^{/} = n(n-1)p_1^{\,2} \sum_{i=0} \binom{-2}{i} (n-2)^{(i)} (-p_2)^i + n\,p_1 \sum_{i=0} \binom{-3}{i} (n-1)^{(i)} (-p_2)^i \tag{1.4}$$

where $n^{(i)} = n(n-1)(n-2)\ldots.(n-i+1)$

Mishra and Sinha (1981) discussed the maximum likelihood estimation of the parameters and also discussed its goodness of fits. This distribution has been found to have tremendous capability to fit discrete data-sets of various nature. A good account of this distribution can be obtained from Johnson et al (2005) and Consul and Famoye (2006).

In this paper, a generalized bivariate geometric distribution (GBGD) has been obtained with the help of a four- urn model in which the contents of some of the urns are changed in some stochastic manner by the player of the game. The marginal distributions of X_1 and X_2 are the geometric distributions, the conditional distribution of X given X_1 is the QBD and that of X_2- X , given X_1 is negative binomial distribution. The first and the second order moments of the distribution have been obtained.

2. AN URN – MODEL

Let there be four urns, namely, A, B, C and D and let there be some balls, each of which has one of the marks out of 10, 01, 11 and 00. Let A contain 'a' balls having 1 at the first place and 'b' balls having 0 at the first place. Let B contain 'c_1' balls having 0 at the first place, C contain 'c_1' balls having mark 11 and 'X_2' balls having mark 10. Let D contain 'd_1' balls having mark 01 and 'd_2' balls having mark 00.

Let n and β be two positive integers. A player is asked to choose two integers 'k_1' and 'k_2' such that $0 < k_1 < k_2 < n$. He is then asked to draw balls one by one with replacement from A till the first ball having 0 at the first place is drawn. If he gets exactly k_1 balls having 1 at the first place in this way, he is permitted to make further draws, otherwise he loses the game. Suppose out of k_1 such balls there are x balls having 1 at the second place also. The player is now asked to add $x\beta$ balls having 1 at both places to each of the urns B and C and $(n - x)\beta$ balls having mark 10 to urn C.

The final contents of the urns are shown in the following table. The marks on the balls are shown in the parentheses, a dash (-) denoting any one of 0 or 1.

(Table 2.1): Final contents of the urns

Urns	A	B	C	D
Number of balls	a (1 -) b (0 -)	c_1 (0 -) $x\beta$ (1 1)	c_1 (1 1) + $x\beta$ (1 1) c_2 (1 0) + $(n\text{-}x)\,\beta$	d_1 (0 1) d_2 (0 0)
Total	a + b	$c_1 + x\beta$	$c_1 + c_2 + n\beta$	$d_1 + d_2$

The player now draws one ball from B and if he gets this ball having 0 at the first place he is asked to draw k_1 balls with replacement from C and if he succeeds in getting exactly x balls having 1 at both the places from this urn he is allowed to make draws from the next urn D. He draws balls from this urn with replacement till he gets exactly (k_1 + 1) balls with mark 0. If for this he gets (k_2 - x) balls with mark 01 then the player is declared to be the winner of the game.

The probability that the player would be allowed to make draws from the urn D is given by

$$\left(\frac{a}{a+b}\right)^{k_1}\frac{b}{a+b}\cdot\frac{c_1}{c_1+x\beta}\binom{k_1}{x}\left(\frac{c_1+x\beta}{c_1+c_2+n\beta}\right)^{x}\left(\frac{c_2+(n-x)\beta}{c_1+c_2+n\beta}\right)^{k-x_1} \qquad (2.1)$$

and the probability of drawing balls from D to be the winner of the game is given by

$$\binom{k_1+k_2-x}{k_2-x}\left(\frac{d_1}{d_1+d_2}\right)^{k_2-x}\left(\frac{d_2}{d_1+d_2}\right)^{k_1+1} \qquad (2.2)$$

The probability of winning the game is obviously the product of (2.1) and (2.2).

Taking $\frac{a}{a+b}=p,\quad \frac{pc_1}{c_1+c_2+n\beta}=\alpha_1,\quad \frac{p\beta}{c_1+c_2+n\beta}=\alpha_2,\quad \frac{d_1}{d_1+d_2}=\theta;$ (2.3) the probability of winning the game can be put in the form

$$P_3(k_1,k_2,x)=\frac{\Gamma(k_1+k_2-x)(1-p)}{\Gamma(x+1)\,\Gamma(k_1-x+1)\,\Gamma(k_2-x+1)} \cdot \alpha_1(\alpha_1+x\alpha_2)^{x-1}\,(p-\alpha_1-x\alpha_2)^{k-x_1}\,\theta^{k_2-x}\,(1-\theta)^{k_1+1} \tag{2.4}$$

3. A GENERALIZED BIVARIATE GEOMETRIC DISTRIBUTION (GBGD)

The expression (2.4) gives the probability of winning the game for the fixed values of k_1,k_2 and x. If k_1,k_2 and x are considered as particular values of random variables, say, X_1,X_2 and X, then it can be shown that $P_3(x_1,x_2,x)$ is a probability distribution for

$0<p<1;\ 0<\alpha_1<p;\ |\alpha_2|<p;\ 0<\theta<1;\ x = 0, 1, 2, \ldots, x_1;\ x_1,x_2 = 0, 1, 2, \ldots$

For, we have

$$\sum_{x_1=0}^{\infty}\sum_{x=0}^{x_1}\sum_{x_2-x=0}^{\infty}P_3(x_1,x_2,x)=\sum_{x_1=0}^{\infty}p^{x_1}(1-p)\sum_{x=0}^{x_1}\binom{x_1}{x}\alpha_1'\left(\alpha_1'+x\alpha_2'\right)^{x-1} \left(1-\alpha_1'-x\alpha_2'\right)^{x_1-x}\sum_{x_2-x=0}^{\infty}\binom{x_1+x_2-x}{x_2-x}\theta^{x_2-x}(1-\theta)^{x_1+1} \tag{3.1}$$

where $\alpha_1'=\alpha_1/p$ and $\alpha_2'=\alpha_2/p$.

As the expression under the summation over x_2-x is the pmf of a negative binomial distribution with parameters x_1+1 and θ; the expression under the summation over x is the pmf of a QBD (1.2) with parameters x_1,α_1' and α_2' and the expression under the summation over x_1 is the pmf of a geometric distribution with parameter p, each of these summations is unity and hence $P_3(x_1,x_2,x)$ represents a true probability distribution.

It can be seen that the marginal distributions of X_1 and X_2 are geometric, the conditional distribution of X given X_1 is QBD and that of X_2- X given X_1 is negative binomial. It can also be seen that at $\alpha_2 = 0$, this distribution reduces to the bivariate geometric distribution (1.1) with

$$\theta = \theta_{01}\,;\; (p-\alpha_1)(1-\theta) = \theta_{10};\; \alpha_1(1-\theta) = \theta_{11}\,;\; (1-p)(1-\theta) = \theta_{00} \tag{3.2}$$

4. MOMENTS OF GBGD

First order moments: Obviously, the mean of X_1 is

$$\mathrm{E}\,(X_1) = p/(1-p) \tag{4.1}$$

The mean of X_2 can be obtained using the identity

$$E(X_2) = E(X_2 - X) + E(X).$$

We have

$$E(X_2 - X) = \sum_{x_1=0}^{\infty} p^{x_1}(1-p) \sum_{x=0}^{x_1} P_2\left(x;\, x_1, \alpha_1{}', \alpha_2{}'\right) \cdot \sum_{x_2-x=0}^{\infty} (x_2 - x)\binom{x_1 + x_2 - x}{x_2 - x} \theta^{x_2-x}\, (1-\theta)^{x_1+1}$$

The sum over x is unity, the function summed being a QBD, and the summation over $x_2 - x$ is the mean of a negative binomial distribution with parameters $x_1 + 1$ and θ. Writing the expression for this mean, we have

$$E(X_2 - X) = \sum_{x_1=0}^{\infty} p^{x_1}(1-p)(1+x_1)\theta/(1-\theta) = \theta/(1-\theta)(1-p) \tag{4.2}$$

Again

$$E(X) = \sum_{x_1=0}^{\infty} p^{x_1}(1-p) \sum_{x=0}^{x_1} x P_2\left(x;\, x_1, \alpha_1{}', \alpha_2{}'\right) \sum_{x_2-x=0}^{\infty} \binom{x_1 + x_2 - x}{x_2 - x} \theta^{x_2-x}\, (1-\theta)^{x_1+1}$$

As the sum over $x_2 - x$ is unity, the function summed being a negative binomial distribution with parameters $x_1 + 1$ and θ and the sum over x is the mean of a QBD with parameters x_1, α_1' and α_2'.

we have from (1.3)

$$E(X) = \sum_{x_1=0}^{\infty} p^x (1-p) x \alpha_1^{/} \sum_{i=0}^{x_1-1} (x_1 - 1)^{(-i)} \alpha_2^{/}$$

$$= \frac{\alpha_1}{(1-p)} \sum_{i=0} 2^{(-i)} \left(\frac{\alpha_2}{(1-p)} \right)^i = \frac{\alpha_1}{(1-p)} \cdot {}_2F_1\left(2;1,1,\alpha_2/(1-p)\right) \tag{4.3}$$

where ${}_2F_1\left(2;1,1,\alpha_2/(1-p)\right)$ is a Gaussian hypergeometric series.

Thus from (4.2) and (4.3), we have,

$$E(X_2) = \frac{\theta}{(1-\theta)(1-p)} + \frac{\alpha_1}{(1-p)} \cdot {}_2F_1\left(2;1,1,\alpha_2/(1-p)\right) \tag{4.4}$$

Second order moments: Obviously, the second order moment of X_1 is

$$E\left(X_1^2\right) = p(1+p)/(1-p)^2 \tag{4.5}$$

The second order moment of X_2 can be obtained using the identity

$$E\left(X_2^2\right) = E(X_2 - X)^2 + E\left(X^2\right) + 2E\left[X(X_2 - X)\right] \tag{4.6}$$

We have,

$$E(X_2 - X)^2 = \sum_{x_1=0}^{\infty} p^{x_1} (1-p) \sum_{x=0}^{x_1} P_2\left(x; x_1, \alpha_1^{/}, \alpha_2^{/}\right) \cdot \sum_{x_2-x=0}^{\infty} (x_2 - x)^2 \binom{x_1 + x_2 - x}{x_2 - x} \theta^{x_2 - x} (1-\theta)^{x_1+1} \tag{4.7}$$

The sum over x is obviously unity as the function summed is a QBD and the summation over $x_2 - x$ is the second moment about origin of a negative binomial distribution with parameters x_1+1 and θ. Writing the expression for this moment, we get

$$E(X_2 - X)^2 = \sum_{x_1=0}^{\infty} p^{x_1} (1-p) \left[\frac{(x_1+1)^2 \theta^2}{(1-\theta)^2} + \frac{(x_1+1)\theta}{(1-\theta)^2} \right]$$

$$= \frac{(1-p)\theta^2}{(1-\theta)^2} \sum_{x_1=0}^{\infty} (x_1+1)^2 \, p^{x_1} \quad + \frac{(1-p)\theta}{(1-\theta)^2} \sum_{x_1=0}^{\infty} (x_1+1)\, p^{x_1}$$

Finding the values of the two series, after a little simplification we finally get

$$E(X_2 - X)^2 = \frac{\theta\left[1-p+\theta(1+p)\right]}{(1-\theta)^2\,(1-p)^2} \tag{4.8}$$

Now

$$E(X^2) = \sum_{x_1=0}^{\infty} p^{x_1}(1-p) \sum_{x=0}^{x_1} x^2 P_2\left(x; x_1, \alpha_1^{/}, \alpha_2^{/}\right).$$

$$\sum_{x_2-x=0}^{\infty} \binom{x_1+x_2-x}{x_2-x} \theta^{x_2-x} \;(1-\theta)^{x_1+1}$$

The sum over $x_2 - x$ is obviously unity. The summation over x gives the second moment of the QBD with parameters $x_1, \alpha_1^{/}$ and $\alpha_2^{/}$. Taking its value from (1.4) we get,

$$E(X^2) = \sum_{x_1=0}^{\infty} p^{x_1}(1-p)$$

$$\left[x_1(x_1-1)\alpha_1^{/} \sum_{i=0} \binom{-2}{i}(x_1-2)^{(i)}(-\alpha_2^{/})^i \; + \; x_1\alpha_1^{/} \sum_{i=0} \binom{-3}{i}(x_1-1)^{(i)}(-\alpha_2^{/})^i\right]$$

Using the factorial moments of geometric distribution with parameter p, we get after a little simplification ,

$$E(X^2) = \frac{2\alpha_1^2}{(1-p)^2}\left[1 + 2.2!\frac{\alpha_2}{(1-p)} + 3.3!\left(\frac{\alpha_2}{(1-p)}\right)^2 + 4.4!\left(\frac{\alpha_2}{(1-p)}\right)^3 + ...\right]$$

$$+ \frac{\alpha_1}{(1-p)^2}\left[1 + 3.2!\left(\frac{\alpha_2}{(1-p)}\right) + 6.3!\left(\frac{\alpha_2}{(1-p)}\right)^2 + 10.4!\left(\frac{\alpha_2}{(1-p)}\right)^3 + ...\right] \tag{4.9}$$

Again

$$E\left[X\left(X_2-X\right)\right]=\sum_{x=0}^{\infty} p^{x_1}(1-p)\sum_{x=0}^{\infty} x\, P_2\left(x;x_1,\alpha_1^{/},\alpha_2^{/}\right)$$

$$.\sum_{x_2-x=0}^{\infty}\left(x_2-x\right)\binom{x_1+x_2-x}{x_2-x}\theta^{x_2-x}\ (1-\theta)^{x_1+1} \tag{4.10}$$

The second summation is obviously the mean of a QBD and the third summation is the mean of a negative binomial distribution. Using these means we have,

$$E\left[X\left(X_2-X\right)\right]$$

$$=\sum_{x=0}^{\infty} p^{x_1}(1-p)\left[x_1\alpha_2^{/}\left\{1+\left(x_1-1\right)\alpha_2^{/}+\left(x_1-1\right)\left(x_1-2\right)\alpha_2^{/}+.....\right\}\right]\frac{\left(x_1+1\right)\theta}{(1-\theta)}$$

$$=\frac{\theta\alpha_1^{/}}{(1-\theta)}\sum_{x_{11}=0}^{\infty}\left(x_1+1\right)\ p^{x_1}(1-p)\left[x_1\left\{1+\left(x_1-1\right)\alpha_2^{/}+\left(x_1-1\right)\left(x_1-2\right)\alpha_2^{/}+.....\right\}\right]$$

$$=\frac{\theta\,\alpha_1^{/}}{(1-p)(1-\theta)}\sum_{i=1}\mu_{(i)}^{/}\left(\alpha_2^{/}\right)^i$$

where $\mu'_{(i)}$ is the ith factorial moment of negative binomial distribution with parameters, m= 2 and p.

Thus, using $\alpha_1^{/}=\alpha_1/p$ and $\alpha_2^{/}=\alpha_2/p$, we get after a little simplification

$$E\left[X\left(X_2-X\right)\right]=\frac{\theta\alpha_1}{p(1-p)(1-\theta)}\left[2\left(\frac{\alpha_2}{1-p}\right)+2.3\left(\frac{\alpha_2}{1-p}\right)^2+........\right] \tag{4.11}$$

Now from (4.8), (4.9) and (4.11), we finally get

$$E\left(X_2^2\right)=\frac{\theta\left[1-p+\theta(1+p)\right]}{(1-\theta)^2\,(1-p)^2}$$

$$+\frac{2\alpha_1^2}{(1-p)^2}\left[1+2.2!\frac{\alpha_2}{(1-p)}+3.3!\left(\frac{\alpha_2}{(1-p)}\right)^2+4.4!\left(\frac{\alpha_2}{(1-p)}\right)^3+........\right]$$

$$+\frac{\alpha_1}{(1-p)^2}\left[1+3.2!\left(\frac{\alpha_2}{(1-p)}\right)+6.3!\left(\frac{\alpha_2}{(1-p)}\right)^2+10.4!\left(\frac{\alpha_2}{(1-p)}\right)^3+.....\right]$$

$$+\frac{2\theta\alpha_1}{p(1-p)(1-\theta)}\left[2\left(\frac{\alpha_2}{1-p}\right)+2.3\left(\frac{\alpha_2}{1-p}\right)^2+........\right] \tag{4.12}$$

REFERENCES

[1]. Consul, P.C., A simple urn model dependent on predetermined strategy, *Sankhya, Ser. B,* 36(3) (1974) 391-399.

[2]. Consul, P.C. and Famoye, F., *Lagrangian Probability Distributions,* Birkhauser, (2006).

[3]. Johnson, N.L., Kemp, A.W. and Kotz,S.*, Univariate Discrete Distributions,* Wiley-Interscience, John Wiley & Sons, (2005).

[4]. Mishra, A. and Singh, S.K., Moments of quasi binomial distribution, *Assam Statistical Review,* 13(1) (2000) 13-20.

[5]. Mishra, A. and Sinha, J.K., A generalization of binomial distribution, *Journal of Indian Statistical Association*, 19 (1981) 93– 98.

[6]. Patil, G.P. and Joshi, S.W., A Dictionary and Bibliography of Discrete Distributions, Edinburgh; Oliver and Boyd, (1968).

(1) Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea

E-mail address: (1) shankerrama2009@gmail.com

(2) Department of Statistics, Patna University, Patna -800 005, India

E-mail address: (2) mishraamar@rediffmail.com