

ON THE RESOLUTION OF A MIXTURE OF  
OBSERVATIONS FROM TWO MODIFIED  
POWER SERIES DISTRIBUTIONS

by

M.S.ABU-SALIH

**ABSTRACT.** The modified power series distributions (MPSD) introduced by Gupta (1974) includes a number of the well known discrete distributions. In this note we assume that a sample of  $N$  observations is available and that it consists of  $N_1$  observations from a MPSD and  $N_2$  observations from another MPSD. The maximum likelihood method is used to identify the population of origin of each observation, and to estimate the mean and parameter of that population. Special cases are dealt with in detail. The variance of the estimate of  $N_i$ ,  $i = 1, 2$ , is derived.

§1. Introduction. The estimation of the parameters of mixtures of specified distributions has been

discussed by many authors. In the majority of published work in this field it is assumed that a sample of size  $N$  is chosen from a population having the probability density function

$$f(x) = \delta f_1(x) + (1-\delta)f_2(x)$$

where  $f_1$  and  $f_2$  belong to the same family with different parameters. Rider (1961 a) used the method of moments to estimate the parameters of mixed Poisson, binomial and Weibull distributions. The mixture of exponential distributions was discussed by Rider (1961 b), and Tallis and Light (1968). Blischke (1964) estimated the parameters of mixtures of binomial distribution. The mixture of normal distributions was discussed by Hasselbland (1966) and Cohen (1967). John (1970 a,b) considered a different model where the sample available is assumed to consist of  $N_1$  observations originating from one population and  $N_2$  observations from another population. He discussed the identification of the population of origin of each observation in the case of two normal and two gamma populations. Dickinson (1974) gave an extension of John's work on the gamma mixture.

The method of maximum likelihood is used in this note to identify the population of origin of each observation in a mixture of observations from two modified power series distributions. Estimates of the parameters are provided.

§2. The Model. Assume a random sample of size  $N$  is available and that it is a mixture of two random samples, where  $N_i$  ( $i = 1, 2$ ) observations belong to the  $i^{\text{th}}$  population, with a modified power series distribution (MPSD) with probability distribution function

$$P_{\theta_i}(X=x) = \frac{a(x)(g(\theta_i))^x}{f(\theta_i)}, \quad i = 1, 2; \quad (1)$$

$x \in T$  where  $T$  is a subset of the set of non-negative integers,  $a(x) > 0$ ,  $g(\theta_i)$  and  $f(\theta_i)$  are positive, finite, and differentiable. The mean of a MPSD is  $\mu(\theta) = \frac{g(\theta)}{g'(\theta)} \frac{f'(\theta)}{f(\theta)}$ , and its variance is  $\mu_2(\theta) = \frac{g(\theta)}{g'(\theta)} \frac{d\theta}{d\theta}$  (see Gupta 1974).

The numbers  $N_1$  and  $N_2$  of observations originating from the first and second populations, respectively, are considered fixed but unknown, and the likelihood functions are conditioned on them. Let the random sample be  $X_1, X_2, \dots, X_N$  and let  $\alpha_{ir} = 1$  if  $X_r = x_r \in i^{\text{th}}$  population,  $\alpha_{ir} = 0$  otherwise,  $i = 1, 2$ . Then the likelihood function is

$$L = \sum_{r=1}^N \sum_{i=1}^2 \left[ \frac{a(x_r)(g(\theta_i))^{x_r}}{f(\theta_i)} \right]^{\alpha_{ir}} \quad (2)$$

Taking logarithms, differentiating with respect to  $\theta_i$  and equating to zero, we get the maximum likelihood equations

$$\frac{f'(\theta_i)}{f(\theta_i)} \sum_{r=1}^N \alpha_{ir} = \frac{g'(\theta_i)}{g(\theta_i)} \sum_{r=1}^N \alpha_{ir} x_r, \quad i=1,2. \quad (3)$$

Equation (3) can be written as

$$\sum_{r=1}^N \alpha_{ir} x_r / \sum_{r=1}^N \alpha_{ir} = \frac{g(\theta_i)}{g'(\theta_i)} \frac{f'(\theta_i)}{f(\theta_i)} = \mu(\theta_i), \quad (4)$$

$$i = 1,2,$$

where  $\mu(\theta_i) = E_{\theta_i}[X]$ .

Before we can evaluate the MLE's  $\hat{\mu}(\theta_i)$  and  $\hat{\theta}_i$  of  $\mu(\theta_i)$  and  $\theta_i$ , respectively, the  $\alpha_{ir}$  ( $i = 1,2$ , and  $r = 1,2,\dots,N$ ) must be determined. Let A be the set of all N-tuples of ones and zeroes. Any sequence  $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1N})$ , where the  $\alpha_{1r}$ 's were defined earlier, belongs to A and determines an identification of the observations with their respective populations of origin. A has  $2^N$  elements. Every element in A determines  $\mu(\theta_i)$ ,  $i = 1,2$ , from (4). If  $\mu(\theta)$  is invertible then  $\theta_i$  and the likelihood function can be evaluated. The sequence  $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1N})$  leading to  $\theta_1$  and  $\theta_2$  which actually maximize the likelihood function will determine the MLE of  $\mu(\theta_i)$  and  $\theta_i$  ( $i = 1,2$ ), and identifies the observations with their respective population of origin. In section 3, a general case is discussed where A may be significantly decreased.

### §3. General case, $g(\theta)$ is monotone increasing.

First, we remark that in this case  $\mu(\theta)$  is invertible because the variance of the MPSD is  $\frac{g(\theta)}{g'(\theta)} \frac{d\mu}{d\theta} > 0$ ,  $g(\theta) > 0$  and  $g'(\theta) > 0$ , hence  $\mu(\theta)$  is a monotone increasing function of  $\theta$ .

Writing (2) in the form

$$L = \prod_{r=1}^N \left[ \frac{a(x_r)(g(\theta_1))^{x_r}}{f(\theta_1)} \right]^{\alpha_{1r}} \left[ \frac{a(x_r)(g(\theta_2))^{x_r}}{f(\theta_2)} \right]^{\alpha_{2r}} \quad (5)$$

we conclude that if

$$\frac{a(x_r)(g(\theta_1))^{x_r}}{f(\theta_1)} > \frac{a(x_r)(g(\theta_2))^{x_r}}{f(\theta_2)} \quad (6)$$

then  $\alpha_{1r} = 1, \alpha_{2r} = 0$  and  $x_r$  originated from the first population. If

$$\frac{a(x_r)(g(\theta_1))^{x_r}}{f(\theta_1)} = \frac{a(x_r)(g(\theta_2))^{x_r}}{f(\theta_2)} \quad (7)$$

then we randomize by taking  $\alpha_{ir} = 1$  with probability  $\frac{1}{2}$ ,  $i = 1, 2$ . Inequality (6) is satisfied iff

$$x_r \left[ \log \frac{g(\theta_1)}{g(\theta_2)} \right] > \log \frac{f(\theta_1)}{f(\theta_2)} \quad (8)$$

and (7) holds iff the inequality in (8) is replaced by equality.

Without loss of generality assume  $\theta_1 > \theta_2$ . Since  $g(\theta)$  is monotone increasing the left hand side of (8) is a monotone increasing function of  $x$ .

Hence, after arranging the observations in an ascending order, the sequence  $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1N})$  will be of the form  $(0, 0, \dots, 0, 1, \dots, 1)$ . In this case A contains  $N+1$  sequences. Each sequence determines the values of  $\mu(\theta_i)$  and  $\theta_i$ ,  $i = 1, 2$ , from (4). Each pair  $\theta_1, \theta_2$  gives a value of the likelihood function. The pair  $\hat{\theta}_1$  and  $\hat{\theta}_2$  which actually maximize (2) is the MLE of  $\theta_1$  and  $\theta_2$  respectively. The sequence  $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1N})$  leading to this solution determines the population of origin of each observation. The case where  $g(\theta)$  is monotone decreasing is treated similarly.

§4. Special cases. In each of the following special cases it is easily verified that  $g(\theta)$  is monotone increasing, henceforth  $\mu(\theta)$  is invertible and can be obtained explicitly or by iterative methods.

a. *Generalized Poisson Distribution (GPD)*. Let (1) be the generalized Poisson distribution given by

$$P(X=x) = \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{x!} \frac{(\theta e^{-\lambda_2 \theta})^x}{e^{\lambda_1 \theta}}, \quad x = 0, 1, 2, \dots \quad (9)$$

$\lambda_1 \theta > 0$ ,  $|\lambda_2 \theta| < 1$  (see Consul and Jain 1973). Let  $\lambda_1, \lambda_2$  be known and  $\theta$  be the unknown parameter. Here  $g(\theta) = \theta e^{-\lambda_2 \theta}$  and  $f(\theta) = e^{\lambda_1 \theta}$ . Equation (4) becomes

$$\frac{\sum \alpha_{ir} x_r}{\sum \alpha_{ir}} = \frac{\lambda_1 \theta_i}{1 - \lambda_2 \theta_i}, \quad i = 1, 2, \quad (10)$$

(all summation signs run from  $r = 1$  to  $r = N$  and therefore the limits will not be shown), and equation (10) has the solution

$$\theta_i = \frac{\sum \alpha_{ir} x_r / \sum \alpha_{ir}}{\lambda_1 + \lambda_2 \sum \alpha_{ir} x_r / \sum \alpha_{ir}}, \quad i=1,2, \quad (11)$$

b. *Decapitated Generalized Poisson Distribution.*

Let (1) be the decapitated generalized Poisson distribution given by

$$P(X=x) = \frac{\lambda_1 (\lambda_1 + \lambda_2 x)^{x-1}}{x!} \frac{(\theta e^{-\lambda_2 \theta})^x}{e^{\lambda_1 \theta} - 1}, \quad x=1,2,\dots; \\ \lambda_1 \theta > 0, \quad |\lambda_2 \theta| < 1.$$

Here  $g(\theta)$  is the same as of the GPD, but  $f(\theta) = e^{\lambda_1 \theta} - 1$  and

$$\mu(\theta) = \frac{\lambda_1 \theta}{(1 - \lambda_2 \theta_i)(1 - e^{-\lambda \theta})},$$

and equations (4) becomes

$$\sum \alpha_{ir} x_r / \sum \alpha_{ir} = \frac{\lambda_1 \theta_i}{(1 - \lambda_2 \theta_i)(1 - e^{-\lambda_1 \theta_i})}, \quad i=1,2. \quad (12)$$

The solution of (12) has been given by Barton, David and Merrington (1960), in the case  $\lambda_1 = 1, \lambda_2 = 0$  (decapitated Poisson distribution). Hence, using the procedure of section 3, we can find  $\hat{\theta}_1, \hat{\theta}_2, \{\alpha_{ir}\}_{r=1}^N$  ( $i = 1,2$ ) which maximize the likelihood function (2).

We remark that (a) and (b) reduce for  $\lambda_2 = 1$  to the Borel-Tanner and decapitated Borel-Tanner distributions, respectively (Haight and Breuer (1960)).

c. *Generalized (decapitated) Negative Binomial Distribution* (GNBD). Let (1) be the GNBD given by

$$P(X=x) = \frac{n\Gamma(n+\beta x)}{x!\Gamma(n+\beta x-x+1)} \frac{(\theta(1-\theta)^{\beta-1})^x}{(1-\theta)^{-n}}, \quad x=0,1,2,\dots, \quad (13)$$

$0 < \theta < 1$  and  $|\theta\beta| < 1$ . (See Jain and Consul 1971). Here  $g(\theta) = \theta(1-\theta)^{\beta-1}$ ,  $f(\theta) = (1-\theta)^{-n}$ , and equation (4) reduces to

$$\Sigma \alpha_{ir} x_r / \Sigma \alpha_{ir} = \mu(\theta_i) = \frac{n\theta_i}{1-\beta\theta_i}, \quad i = 1,2$$

which gives

$$\theta_i = \frac{\Sigma \alpha_{ir} x_r / \Sigma \alpha_{ir}}{n + \beta \Sigma \alpha_{ir} x_r / \Sigma \alpha_{ir}}; \quad i = 1,2.$$

Letting  $f(\theta)$  in (13) to be  $(1-\theta)^{-n-1}$  and  $x = 1,2,\dots$  we get the *decapitated* GNBD with

$$\Sigma \alpha_{ir} x_r / \Sigma \alpha_{ir} = \frac{n\theta_i}{(1-\beta\theta_i)(1-(1-\theta_i)^{n+1})} \quad i=1,2. \quad (14)$$

We also remark that (c) reduces to the binomial and negative binomial distributions for  $\beta = 0$  and  $\beta = 1$ , respectively.

§5. Case where  $N_1$  and  $N_2$  are known. If in addition to the condition of the general case we suppose  $N_1$  and  $N_2$  to be known then the problem is solved in one step; namely, since the sequence  $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1N}) = (0, 0, \dots, 0, 1, \dots, 1)$  and  $N_1$  is



known then there will be  $N_1$  ones in the sequence. Hence, the largest  $N_1$  observations belong to the first population (the one with the bigger  $\theta$ ). The MLE of  $\theta_1$  in the particular cases of section 4 remain valid with  $\sum_{r=1}^N \alpha_{1r} x_r / \sum_{r=1}^N \alpha_{1r}$  replaced by  $\sum_{r=N_2+1}^N x_r / N_1$  and  $\sum_{r=1}^N \alpha_{2r} x_r / \sum_{r=1}^N \alpha_{2r}$  replaced by  $\sum_{r=1}^{N_2} x_r / N_2$ .

§6. Asymptotic variances of estimates. Since the variances of the MLE  $\hat{\theta}_1$  and  $\hat{\theta}_2$  are extremely complex, and the standard maximum likelihood theory may not be used (see Dickinson 1974) we confine ourselves to deriving the variance of  $\Sigma \alpha_{1r}$ . The variance of  $\Sigma \alpha_{2r}$  is derived similarly. Let

$$b = \frac{\log f(\theta_1) - \log f(\theta_2)}{\log g(\theta_1) - \log g(\theta_2)},$$

from (8) we get  $\alpha_{1r} = 1$  if  $x_r > b$ , and  $\alpha_{1r} = 0$  with probability  $\frac{1}{2}$  if  $x_r = b$ . Moreover,  $\alpha_{1r} = 0$  if  $x_r < b$ . For  $i = 1, 2$ , let

$$B_i = P_{\theta_i}(X > b) = \sum_{\substack{x \in T \\ x > b}} P_{\theta_i}(X = x),$$

$E_i = P_{\theta_i}(X = b)$ , and  $A_i$  be the  $i^{\text{th}}$  population. Then

$$\begin{aligned} E[\Sigma \alpha_{1r}] &= \sum P(\alpha_{1r} = 1) \\ &= \sum_{i=1}^2 \sum_{r=1}^2 |P(X_r > b | X_r \in A_i) P(X_r \in A_i) + \end{aligned}$$

$$\begin{aligned}
& + P(X_r \in A_i)P(X_r=b | X_r \in A_i)P(\alpha_{1r}=1 | X_r \in A_i, X_r=b) | \\
& = \left[ B_1 \frac{N_1}{N} + B_2 \frac{N_2}{N} + \frac{1}{2} \frac{N_1}{N} E_1 + \frac{1}{2} \frac{N_2}{N} E_2 \right] \\
& = N_1 \left( B_1 + \frac{1}{2} E_1 \right) + N_2 \left( B_2 + \frac{1}{2} E_2 \right).
\end{aligned} \tag{15}$$

$$E[\sum \alpha_{1r}]^2 = E[\sum \alpha_{1r}^2 + \sum_{r \neq s} \alpha_{1r} \alpha_{1s}]. \tag{16}$$

From the definition of  $\alpha_{1r}$  and (15) we get

$$E[\sum \alpha_{1r}^2] = N_1 \left( B_1 + \frac{1}{2} E_1 \right) + N_2 \left( B_2 + \frac{1}{2} E_2 \right) \tag{17}$$

For  $r \neq s$ ,  $E(\alpha_{1r} \alpha_{1s})$  is

$$\begin{aligned}
& \sum_{i,j=1,2} [P(X_r > b, X_s > b, X_r \in A_i, X_s \in A_j) \\
& + P(X_r > b, X_r \in A_i, X_s = b, \alpha_{1s} = 1, X_s \in A_j) \\
& + P(X_r = b, \alpha_{1r} = 1, X_r \in A_i, X_s > b, X_s \in A_j) \\
& + P(X_r = b, \alpha_{1r} = 1, X_r \in A_i, X_s = b, \alpha_{1s} = 1, X_s \in A_j)].
\end{aligned}$$

Using the multiplication rule and summing over  $r$  and  $s$  we find,

$$\begin{aligned}
E[\sum_{r \neq s} \alpha_{1r} \alpha_{1s}] & = B_1^2 N_1 (N_1 - 1) + 2 B_1 B_2 B_1 N_2 \\
& + B_2^2 N_2 (N_2 - 1) + B_1 E_1 N_1 (N_1 - 1) \\
& + (B_1 E_2 + B_2 E_1) N_1 N_2 + B_2 E_2 N_2 (N_2 - 1) \\
& + E_1^2 N_1 (N_1 - 1) + 2 E_1 E_2 N_1 N_2 + E_2^2 N_2 (N_2 - 1).
\end{aligned} \tag{18}$$

From (15) to (18),

$$\begin{aligned} \text{Var}(\sum \alpha_{1r}) &= N_1 B_1 (1-B_1) + N_2 B_2 (1-B_2) \\ &+ N_1 E_1 \left(\frac{1}{2}-B_1\right) + N_2 E_2 \left(\frac{1}{2}-B_2\right) \\ &- \frac{1}{4}(N_1 E_1^2 + N_2 E_2^2) . \end{aligned} \quad (19)$$

From (15) we conclude that the MLE,  $\sum \alpha_{1r}$ , of  $N_1$  is heavily biased except if  $(B_1 + \frac{1}{2}E_1)$  is very close to 1 and  $(B_2 + \frac{1}{2}E_2)$  is very close to zero. If  $b$  is not an integer then  $E_1 = E_2 = 0$  and the terms involving  $E_1, E_2$  in (15) and (19) vanish.

ACKNOWLEDGMENTS. The author is very thankful to the referee for his comments and to Prof. M.M. Siddique for his useful discussions. This work was partially supported by Yarmouk University.

#### REFERENCES

- Barton, D.E. , David, F.N., and Merrington, M. (1960). *Tables for the solution of the exponential equation,  $\exp(-a)+ka=1$* , *Biometrika* 47, 439-445.
- Blischke, W.R. (1964). *Estimating the parameters of Mixtures of Binomial Distributions*, *J. Amer. Stat. Assoc.* 59, 510-528.
- Cohen, A.C. (1967). *Estimation in Mixture of two Normal Distributions*, *Technometrics* 9, 15-28.
- Consult, P.C. and Jain, G.C. (1973), *A generalization of the Poisson distribution*, *Technometrics* 15, 791-799.
- Dickinson, J.P. (1974). *On the Resolution of a Mix-*

- ture of observations from two Gamma Distributions by the Method of Maximum Likelihood. *Metrika* 21, 133-141.
- Gupta, R.C. (1974). *Modified Power Series Distribution and Some of its Applications*, *Sankhya* 36 B, 288-298.
- Haight, F.A. and Breuer, M.A. (1960). *The Borel-Tanner distribution*, *Biometrika* 47, 143-146.
- Hasselblad, V. (1966). *Estimation of Parameters for a Mixture of Normal Distributions*, *Technometrics* 8, 431-444.
- Jain, G.C. and Consul, P.C. (1971). *A Generalized Negative Binomial Distribution*, *Siam J. Appl. Math.* 21, 501-513.
- John, S. (1970 a). *On Identifying the Population of Origin of Each Observation in a Mixture of Observations from two Normal Populations*, *Technometrics* 12, 555-563.
- . (1970 b). *On Identifying the Population of Each Observation in a Mixture of Observations from two Gamma Populations*, *Technometrics* 12, 565-568.
- Rider, P.R. (1961 a). *Estimating the Parameters of Mixed Poisson, Binomial and Weibull Distributions by the Method of Moments*, *Bulletin of the International Statistical Institute* 39, 225-232.
- . (1961 b). *The Method of Moments Applied to a Mixture of two Exponential Distributions*, *Ann. Math. Stat.* 32, 143-147.
- Tallis, G.M.M and Light, R. (1968). *The Use of Fractional Moments for Estimating the Parameters of Mixed Exponential Distribution*, *Technometrics* 10, 161-175.

\* \*

Department of Mathematics  
Yarmouk University  
Irbid, JORDANIA.

(Recibido en Septiembre de 1979).