Bayesian Estimation in Some Power Series Distributions

Anwar Hassan
Department of Statistics and Operations Research
King Saud University
Riyadh, Saudi Arabia
Email: anwar.hassan2007@gmail.com

Peer Bilal Ahmad
Department of Mathematics,
Islamic University of Science & Technology,
Awantipora, Pulwama-192122 (J&K), India
Email: peerbilal@yahoo.co.in

Anwar H. Joarder
Department of Mathematics and Statistics
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia
Email: anwarj@kfupm.edu.sa

ABSTRACT

In this paper, we study the Bayesian estimation of functions of parameters of some power series distributions. These estimators are better than the classical minimum variance unbiased estimators (MVUE) as given by Patil and Joshi (1970), in the sense that these increase the range of the estimation and also have simpler forms.

1. **INTRODUCTION**

Patil (1961) defined a generalized power series distribution (GPSD) if its probability mass function (p.m.f.) is given by

$$P(X = x) = \frac{a_x \theta^x}{g(\theta)}, \qquad x \in S, \tag{1}$$

where $g(\theta)$ is a generating function i.e.,

$$g(\theta) = \sum_{x \in S} a_x \theta^x, \qquad \theta \geq 0, \, a_x \geq 0$$

so that $g(\theta)$ is positive, finite and differentiable and S is a non-empty countable sub-set of non-negative integers.

It can be easily seen that proper choice of S and $g(\theta)$, the GPSD model (1) reduces to the binomial, negative binomial, Poisson and logarithmic series distributions. When $g(\theta) = \theta$, the GPSD model (1) coincides with the class of distributions considered by Roy and Mitra (1957). Patil (1962a, 1962b) has investigated some structural properties and statistical problems associated with GPSD. Patil (1962c,1964) obtained the maximum likelihood estimation of GPSD where as Patil (1962d) obtained estimation by two moments method for GPSD. Patil (1963) and Patil and Joshi (1970) studied properties associated with minimum variance unbiased estimation (MVUE) for power series distributions. Eideh and Ahmad (1989) have investigated tests based on the Kullback-Leibler information measure, for a one parameter power series distribution.

In this paper we study the Bayesian estimation of GPSD for proper choice of S and (θ) . These estimators are better than the classical minimum variance unbiased estimators (MVUE) as given by Patil (1963), and Patil and Joshi (1970) in the sense that these increase the range of the estimation and also have simpler forms. In particular, we derive the Bayesian estimator of $\phi(\theta) = \theta^r, r \in (-\infty, \infty)$. Note that the range of estimation is increased as we have taken $r \in (-\infty, \infty)$.

For Bayesian estimation of discrete distributions we refer, among others, to Irony (1982), Howlader and Balasooriya (2003), Abdul Razak and Patil (1986), Bhattacharya (1968).

2. SOME PRELIMINARIES

Let $X_1, X_2, \cdots, X_N$ denote a random sample of size N from a given probability mass function (p.m.f.), then

$$T_N = \sum_{i=1}^{N} X_i. \tag{2}$$

We shall use the following result as given Abranowitz and Stegun (1964)

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} \, du, \tag{3}$$

$$\Gamma(x) b^{-x} = \int_0^\infty u^{x-1} e^{-bu} \, du, \tag{4}$$

$$\frac{\Gamma(b-a)\Gamma(a)M(a,b;z)}{\Gamma(b)} = \int_0^1 t^{a-1}(1-t)^{b-a-1} e^{zt} \, dt, \tag{5}$$

where $M(a, b; z)$ is the confluent hypergeometric function and has a series representation given by

$$M(a, b; z) = \sum_{j=0}^\infty \frac{(a)_j z^j}{(b)_j j!}, \text{ where } (a)_0 = 1 \text{ } and (a)_n = a(a+1)(a+2) \cdots (a+n-1). \tag{6}$$

3. BAYESIAN ESTIMATION

Let $X_1, X_2, \cdots, X_N$ be a random sample of size N from (1). The likelihood function of the random sample denoted by $L(\theta)$ is given by

$$L(\theta) = \prod_{i=1}^N \frac{a_{x_i} \theta^{x_i}}{g(\theta)}$$

$$= \theta^{t_N} (g(\theta))^{-N} \prod_{i=1}^N a_{x_i}$$

$$\text{i.e. } L(\theta) \propto \theta^{t_N} (g(\theta))^{-N}, \tag{7}$$

where t_N is an observed value of T_N defined by (2). The posterior probability density function of θ, corresponding to prior $h(\theta)$ is given by

$$\pi(\theta/t_N) = \frac{L(\theta)h(\theta)}{\int_0^\infty L(\theta)h(\theta)d\theta} \tag{8}$$

Under the squared error loss function (SELF) given by $L(\phi(\theta), d) = (\phi(\theta) - d)^2$, where $\phi(\theta)$ is a function of θ and d is a decision, the Bayes estimate $\hat{\phi}_B$ of $\phi(\theta)$ is given by

$$\hat{\phi}_B = \int_0^\infty \phi(\theta)\pi(\theta/t_N)d\theta \tag{9}$$

Similarly, under the weighted squared error loss function (WSELF) given by $L(\phi(\theta), d) = w(\theta)(\phi(\theta) - d)^2$, where $w(\theta)$ is a function of θ, the Bayes estimate $\hat{\phi}_w$ of $\phi(\theta)$ is given by

$$\hat{\phi}_w = \frac{\int_0^\infty w(\theta)\phi(\theta)\pi(\theta/t_N)d\theta}{\int_0^\infty w(\theta)\pi(\theta/t_N)d\theta} \tag{10}$$

We consider two different forms of $w(\theta)$ as given below:

(i) Let $w(\theta) = \theta^{-2}$. The Bayes estimate $\hat{\phi}_M$ of $\phi(\theta)$ known as the minimum expected loss (MEL) estimate is given by

$$\hat{\phi}_M = \frac{\int_0^\infty \theta^{-2}\phi(\theta)\pi(\theta/t_N)d\theta}{\int_0^\infty \theta^{-2}\pi(\theta/t_N)d\theta} \tag{11}$$

This loss function was used by Tummala and Sathe (1978) for estimating the reliability of certain life time distributions and by Zellner (1979) for estimating functions of parameters of some econometric models.

(ii) Let $w(\theta) = \theta^{-2} e^{-a\theta}, a > 0$. The Bayes estimate $\hat{\phi}_E$ of $\phi(\theta)$ known as the exponentially weighted minimum expected loss (EWMEL) estimate and is given by

$$\hat{\phi}_E = \frac{\int_0^\infty \theta^{-2} e^{-a\theta} \phi(\theta) \pi(\theta/t_N) d\theta}{\int_0^\infty \theta^{-2} e^{-a\theta} \pi(\theta/t_N) d\theta}. \tag{12}$$

Now we shall consider some special cases of the probability mass function given by (1) and obtain the corresponding Bayesian estimate in each case.

4. POISSON DISTRIBUTION

A random variable X is said to have Poisson distribution if its probability mass function is given by

$$P_\theta(x) = \begin{cases} \frac{\theta^x e^{-\theta}}{x!}, & x \in S, \theta > 0, \\ 0, & otherwise \end{cases} \tag{13}$$

where $= \{0,1,2,\cdots\}$. It is a special case of (1) when $a_x = \frac{1}{x!}$ and $g(\theta) = e^{-\theta}$.

In this case the likelihood function $L(\theta)$ is given by

$$L(\theta) \propto (\theta)^{t_N} e^{-\theta N}. \tag{14}$$

With the gamma prior for θ given by

$$h(\theta) = \frac{\alpha^\beta}{\Gamma\beta} e^{-\alpha\theta} \theta^{\beta-1}, \ \alpha, \beta, \theta > 0, \tag{15}$$

the posterior probability distribution function of θ is given by

$$\pi(\theta/t_N) = \frac{L(\theta)h(\theta)}{\int_0^\infty L(\theta)h(\theta)d\theta} = \frac{(N+\alpha)^{t_N+\beta}}{\Gamma(t_N+\beta)} (\theta)^{(t_N+\beta)-1} e^{-(N+\alpha)\theta} \tag{16}$$

Under the squared error loss function (SELF) given by $L(\phi(\theta), d) = (\phi(\theta) - d)^2$, the Bayes estimate $\hat{\phi}_B$ of $\phi(\theta)$ is given by

$$\hat{\phi}_B^r = \int_0^\infty \phi(\theta)\pi(\theta/t_N)d\theta = \frac{\Gamma((t_N+\beta+r)}{(N+\alpha)^r \Gamma(t_N+\beta)} \tag{17}$$

Note that if $r = 1$, then we get the Bayes estimator of θ as

$$\theta^* = \frac{\Gamma(t_N+\beta+1)}{(N+\alpha)\Gamma(t_N+\beta)} = \frac{t_N+\beta}{N+\alpha} = \frac{\sum x_i+\beta}{N+\alpha} \tag{18}$$

which is identical to maximum likelihood estimator (MLE) of θ if $\alpha = \beta = 0$.

Similarly under WSELF, when $w(\theta) = \theta^{-2}$, the MEL estimate of $\phi(\theta) = \theta^r$ is given by

$$\hat{\theta}_M^r = \frac{\int_0^\infty \theta^{r-2}(\theta)^{(t_N+\beta)-1}e^{-(N+\alpha)\theta}\,d\theta}{\int_0^\infty \theta^{-2}(\theta)^{(t_N+\beta)-1}e^{-(N+\alpha)\theta}\,d\theta}$$

$$= \frac{\Gamma(t_N+\beta+r-2)}{(N+\alpha)^r\Gamma(t_N+\beta-2)}. \tag{19}$$

Finally under WSELF, when $w(\theta) = \theta^{-2}e^{-a\theta}, a > 0$ the EWMEL estimate of $\phi(\theta) = \theta^r$ is given by

$$\hat{\theta}_E^r = \frac{\int_0^\infty \theta^{r-2}e^{-a\theta}(\theta)^{(t_N+\beta)-1}e^{-(N+\alpha)\theta}\,d\theta}{\int_0^\infty \theta^{-2}e^{-a\theta}(\theta)^{(t_N+\beta)-1}e^{-(N+\alpha)\theta}\,d\theta}$$

$$= \frac{\Gamma(t_N+\beta+r-2)}{(N+\alpha+a)^r\Gamma(t_N+\beta-2)}. \tag{20}$$

If $r = 1$, we get Bayes estimator of θ for (19) and (20). Bayes estimator of θ^r, in each case can be obtained by replacing t_N by T_N, in each equations of Bayes estimator.

Note that the MVUE of θ^r, exists as long as $z \geq r$ and is zero (0) if $z < r$, which is a serious limitation of the MVUE. The Bayes estimates, on the other hand, as given above, are free from such restrictions on t_N and r, as long as $r \geq 1$. This is another advantage of Bayesian estimation over the MVUE. However, if $r < 0$, Bayes estimates are zero (0), if $t_N + \beta < -r$ in (17) and $t_N + \beta - 2 < -r$ in (19) and (20) respectively.

5. NEGATIVE BINOMIAL DISTRIBUTION

A discrete random variable X is said to have negative binomial distribution if its probability mass function is given by

$$P_\theta(x) = \binom{n+x-1}{x}\theta^x(1-\theta)^n, \tag{21}$$

which is a special case of GPSD(1) whenever

$$a_x = \binom{n+x-1}{x}, \quad g(\theta) = (1-\theta)^{-n}, \ x \in S = \{0,1,2,\cdots,\infty\}, 0 < \theta < 1.$$

The likelihood function (θ), in this case, is given by

$$L(\theta) \propto (\theta)^{t_N}(1-\theta)^{nN}. \tag{22}$$

Since $0 < \theta < 1$, we take two different prior distributions given below

$$h_1(\theta) = \frac{\theta^{p-1}(1-\theta)^{q-1}}{B(p,q)}, \ 0 < \theta < 1, p, q > 0 \tag{23}$$

where $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ and

$$h_2(\theta) = \frac{e^{-b\theta}\theta^{p-1}(1-\theta)^{q-1}}{B(p,q)M(p,p+q;-b)}, \quad 0 < \theta < 1, p, q > 0, \tag{24}$$

with $M(a,b;z)$ is given by (6).

The posterior probability density function of θ, corresponding to $h_1(\theta)$ is given by

$$\pi_1(\theta/t_N) = \frac{L(\theta)h_1(\theta)}{\int_0^1 L(\theta)h_1(\theta)d\theta} = \frac{(\theta)^{(t_N+p)-1}(1-\theta)^{(nN+q)-1}}{B(t_N+p,nN+q)} \tag{25}$$

The posterior probability density function of θ, corresponding to $h_2(\theta)$ is given by

$$\pi_2(\theta/t_N) = \frac{L(\theta)h_2(\theta)}{\int_0^1 L(\theta)h_2(\theta)d\theta} = \frac{(\theta)^{(t_N+p)-1}(1-\theta)^{(nN+q)-1}e^{-b\theta}}{B(t_N+p,nN+q)M(t_N+p,p+q+nN+t_N;-b)} \tag{26}$$

Both $h_1(\theta)$ and $h_2(\theta)$ are natural conjugate prior density. The prior density $h_2(\theta)$ is known as the generalized beta density considered by Holla (1968) and Bhattacharya (1968).

Under SELF, the Bayes estimate of $\phi(\theta) = \theta^r$, corresponding to posterior density (25) is given by

$$\hat{\theta}_{1B}^r = \int_0^1 \theta^r \pi_1(\theta/t_N)d\theta = \frac{1}{B(t_N+p,nN+q)} \int_0^1 (\theta)^{(t_N+p+r)-1}(1-\theta)^{(nN+q)-1}d\theta$$

$$= \frac{B(t_N+p+r,nN+q)}{B(t_N+p,nN+q)}. \tag{27}$$

Similarly under WSELF, when $w(\theta) = \theta^{-2}$, the MEL estimate of $\phi(\theta) = \theta^r$, corresponding to posterior density (25) is given by

$$\hat{\theta}_{1M}^r = \frac{\int_0^1 (\theta)^{(t_N+p+r-2)-1}(1-\theta)^{(nN+q)-1}d\theta}{\int_0^1 (\theta)^{(t_N+p-2)-1}(1-\theta)^{(nN+q)-1}d\theta}$$

$$= \frac{B(t_N+p+r-2,nN+q)}{B(t_N+p-2,nN+q)} \tag{28}$$

Finally under WSELF, when $w(\theta) = \theta^{-2}e^{-a\theta}, a > 0$ the EWMEL estimate of $\phi(\theta) = \theta^r$, corresponding to posterior density (25) is given by

$$\hat{\theta}_{1E}^r = \frac{\int_0^1 (\theta)^{(t_N+p+r-2)-1}(1-\theta)^{(nN+q)-1}e^{-a\theta}d\theta}{\int_0^1 (\theta)^{(t_N+p-2)-1}(1-\theta)^{(nN+q)-1}e^{-a\theta}d\theta}$$

$$= \frac{B(t_N+p+r-2,nN+q)M(t_N+p+r-2,p+q+nN+r-2+t_N;-a)}{B(t_N+p-2,nN+q)M(t_N+p-2,p+q+nN+t_N-2;-a)} \tag{29}$$

Also under SELF, the Bayes estimate of $\phi(\theta) = \theta^r$, corresponding to posterior density (26) is given by

$$\hat{\theta}_{2B}^r = \int_0^1 \theta^r \pi_2(\theta/t_N)d\theta = \int_0^1 \frac{(\theta)^{(t_N+p+r)-1}(1-\theta)^{(nN+q)-1}e^{-b\theta}d\theta}{B(t_N+p,nN+q)M(t_N+p,p+q+nN+t_N;-b)}$$

$$= \frac{B(t_N+p+r,nN+q)M(t_N+p+r,nN+p+q+t_N+r;-b)}{B(t_N+p,nN+q)M(t_N+p,p+q+t_N+nN;-b)}. \tag{30}$$

Similarly under WSELF, when $w(\theta) = \theta^{-2}$, the MEL estimate of $\phi(\theta) = \theta^r$,corresponding to posterior density (26) is given by

$$\hat{\theta}_{2M}^r = \frac{\int_0^1 (\theta)^{(t_N+p+r-2)-1}(1-\theta)^{(nN+q)-1}e^{-b\theta}d\theta}{\int_0^1 (\theta)^{(t_N+p-2)-1}(1-\theta)^{(nN+q)-1}e^{-b\theta}d\theta}$$

$$= \frac{B(t_N+p+r-2,nN+q)M(t_N+p+r-2,nN+p+q+t_N+r-2;-b)}{B(t_N+p-2,nN+q)M(t_N+p-2,nN+p+q+t_N-2;-b)}. \tag{31}$$

Finally, under WSELF, when $w(\theta) = \theta^{-2}e^{-a\theta}, a > 0$ the EWMEL estimate of $(\theta) = \theta^r$, corresponding to posterior density (26) is given by

$$\hat{\theta}_{2E}^r = \frac{\int_0^1 (\theta)^{(t_N+p+r-2)-1}(1-\theta)^{(nN+q)-1}e^{-(a+b)\theta}d\theta}{\int_0^1 (\theta)^{(t_N+p-2)-1}(1-\theta)^{(nN+q)-1}e^{-(a+b)\theta}d\theta}$$

$$= \frac{B(t_N+p+r-2,nN+q)M(t_N+p+r-2,p+q+nN+r-2+t_N;-(a+b))}{B(t_N+p-2,nN+q)M(t_N+p+r-2,p+q+t_N+nN+r-2;-(a+b))} \tag{32}$$

Note that the MVUE of θ^r is zero (0) if $z < r$,which is a serious limitation of the MVUE. The Bayes estimates, on the other hand, are zero (0), if $r < 0$ such that $t_N + p < -r$, $t_N + p - 2 < -r$, $t_N + nN + p + q < -r$, $t_N + nN + p + q - 2 < -r$,depending upon the various loss functions. For $n = 1$, we get the estimate for geometric series distribution.

6. LOGARITHMIC SERIES DISTRIBUTION

A random variable X is said to follow logarithmic series distribution if its probability mass function is given by

$$P(X = x) = \frac{1}{x}\frac{\theta^x}{[-log(1-\theta)]}, \quad x = 1,2,3,\cdots, \tag{33}$$

which is a special case of generalized power series (1) whenever $a_x = \frac{1}{x}$ and $g(\theta) = -log(1 - \theta)$. In this case the likelihood function $L(\theta)$ is given by

$$L(\theta) \propto (\theta)^{t_N}[-log(1 - \theta)]^{-N}. \tag{34}$$

With the prior density given by

$$h(\theta) = \begin{cases} \frac{(\lambda+1)^{N+1}(1-\theta)^\lambda(-log(1-\theta))^N}{\Gamma(N+1)}, & 0 < \theta < 1, \lambda > 0 \\ 0, & otherwise \end{cases} \tag{35}$$

the posterior density would be

$$\pi(\theta/t_N) = \frac{L(\theta)h(\theta)}{\int_0^\infty L(\theta)h(\theta)d\theta} = \frac{(\theta)^{t_N}(1-\theta)^\lambda}{B(t_N+1,\lambda+1)} \tag{36}$$

Under SELF, the Bayes estimate of $\phi(\theta) = \theta^r$, is given by

$$\hat{\phi}_B^r = \int_0^1 \theta^r \pi(\theta/t_N)d\theta = \frac{B(t_N+r+1,\lambda+1)}{B(t_N+1,\lambda+1)} \tag{37}$$

Similarly under WSELF, when $w(\theta) = \theta^{-2}$, the MEL estimate of $\phi(\theta) = \theta^r$, is given by

$$\hat{\theta}_M^r = \frac{\int_0^1 (\theta)^{(t_N+r-1)-1}(1-\theta)^\lambda d\theta}{\int_0^1 (\theta)^{(t_N-1)-1}(1-\theta)^\lambda d\theta}$$

$$= \frac{B(t_N+r-1,\lambda+1)}{B(t_N-1,\lambda+1)} \tag{38}$$

Finally under WSELF, when $w(\theta) = \theta^{-2}e^{-a\theta}, a > 0$ the EWMEL estimate of $\phi(\theta) = \theta^r$, is given by

$$\hat{\theta}_E^r = \frac{\int_0^1 (\theta)^{(t_N+r-1)-1}(1-\theta)^\lambda e^{-a\theta} d\theta}{\int_0^1 (\theta)^{(t_N-1)-1}(1-\theta)^\lambda e^{-a\theta} d\theta}$$

$$= \frac{B(t_N+r-1,\lambda+1)M(t_N+r-1,t_N+r+\lambda;-a)}{B(t_N-1,\lambda+1)M(t_N-1,t_N+\lambda;-a)} \tag{39}$$

Note that the MVUE of θ^r is zero (0) if $z < r$. The Bayes estimates, on the other hand, are zero (0), if $r < 0$ such that $t_N + 1 < -r$, $t_N - 1 < -r$, $t_N + \lambda < -r$, depending upon the various loss functions.

ACKNOWLEDGEMENTS

REFERENCES

[1] Abdul Razak, R.S. and Patil, G.P. (1986). Power series distributions and their conjugates in statistical modeling and Bayesian Inference, Communications in Statistics- Theory and Methods, 15(3), 623-641.

[2] Abramowitz, M. and Stegum, I.A. (1965). Handbook of Mathematical Functions, New York: Dover.

[3] Bhattacharya, S.K. (1968). Bayes approach to compound distributions arising from truncated mixing densities. Annals of Institute of Statistical Mathematics, 20, 375-381.

[4] Eideh, A.H. and Ahmad, M.S. 1989. Some tests for the power series distributions in one parameter using the Kullback-Leibler information measure, Communications in Statistics-Theory and Methods, 18: 3649-3663.

[5] Holla, M. S. (1968). Discrete distributions with prior information, Annals of Institute of Statistical Mathematics, 20, 151-157.

[6] Howlader, H.A. and Balasooriya, U. (2003). Bayesian Estimation of the Distribution Function of the Poisson Model, Biometrical Journal, 45(7), 901-912.

[7] Irony, T.Z.(1992). Bayesian estimation of discrete distributions, Journal of Applied Statistics, 19(4), 533-549.

[8] Patil, G.P. (1961). Asymptotic bias and efficiency of ratio estimates for generalized power series distributions and certain applications, Sankhya, 23, 269-80.

[9] Patil, G.P. (1962a). On homogeneity and combined estimation for the generalized power series distribution and certain applications, Biometrics, 18, 365-374.

[10] Patil, G.P. (1962b). Certain properties of generalized power series distribution, Annals of the Institute of Statistical Mathematics,14, 179-182.

[11] Patil, G.P. (1962c). Maximum likelihood estimation for generalized power series distributions and its application to a truncated binomial distribution, Biometrika, 49, 227-238.

[12] Patil, G.P. (1962d). Estimation by two moments method for generalized power series distribution and certain applications, Sankhya, 24, 201-214.

[13] Patil, G.P. (1963). Minimum variance unbiased estimation and certain problems of additive number theory, Annals of Mathematical Statistics, 34, 1050-1056.

[14] Patil, G.P. (1964). Estimation for generalized power series distribution with two parameters and its application to binomial distribution, Contribution to Statistics, C. R. Rao (Ed.), 335-344. Calcutta Statistical Publishing Society, Oxford Pergamon Press.

[15] Patil, G.P. and Joshi, S. W. (1970). Further results on minimum variance unbiased estimation and additive number theory, Annals of Mathematical Statistics,41, 567-575.

[16] Roy,J. and Mitra, S.K. (1957). Unbiased minimum variance estimation in a class of discrete distributions, Sankhya, 18, 371-378.

[17] Tummala, V.M. and Sathe, P.T. (1978). Minimum expected loss estimators of reliability and parameters of certain life time distribution, IEEE Transactions on Reliability, 27(4), 283-285.

[18] Zellner, A. and Park, S.B. (1979). Minimum expected loss estimators of functions of parameters and structural coefficients of econometric models, Journal of American Statistical Association, 74, 185-193.