

MAXIMUM LIKELIHOOD ESTIMATION IN GENERALIZED GAMMA TYPE MODEL

Vinod Kumar¹ and Gaurav Shukla²

1. Deptt. of Maths, Stats. & Comp. Sc., G. B. Pant Univ. of Ag. & Tech.,
 Pantnagar, India.

E-Mail: vinod_kumarbcb@yahoo.com

2. Deptt. of Statistics, G.F. (P.G.) College, Shahjahanpur, India.

E-Mail: saigaurav83@gmail.com

Abstract

In the present paper, the maximum likelihood estimates of the two parameters of a generalized gamma type model have been obtained directly by solving the likelihood equations as well as by reparametrizing the model first and then solving the likelihood equations (as done by Prentice, 1974) for fixed values of the third parameter. It is found that reparametrization does neither reduce the bulk nor the complexity of calculations. as claimed by Prentice (1974). The procedure has been illustrated with the help of an example. The distribution of MLE of θ along with its properties has also been obtained.

Key Words: Generalized Gamma Model, Maximum Likelihood Estimator (MLE), Reliability Function, Hazard Rate Function

1. Introduction

The study of life testing models begins with the estimation of the unknown parameters involved in the models. Stacy (1962) proposed a generalized gamma model and studied its characteristics. Shukla and Kumar (2006) used this model in a bit little transformed form to cover more real life situations. Stacy and Mihram (1965) and Harter (1967) have derived maximum likelihood estimators of generalized gamma model under different situations. Prantice (1974) has considered maximum likelihood estimators for generalized gamma model by using the technique of reparametrization. The probability density function of the three parameter generalized gamma type model is

$$f(t) = \frac{p}{\theta^k \Gamma(k)} t^{pk-1} e^{\left\{ \frac{-t^p}{\theta} \right\}} I_{(0, \infty)}(t) \quad ; \quad p > 0, \theta > 0, k > 0 \quad (1)$$

This distribution may be used as a generalized lifetime model, as it includes the widely used exponential ($p=k=1$), Weibull ($k=1$), gamma ($p=1$) and the model given by Stacy (1962) as special cases.

2. Maximum Likelihood Estimation

The Likelihood function (L) of (1) is given by

$$L = \left[\frac{p}{\theta^k \Gamma(k)} \right]^n e^{-\sum_{i=1}^n \frac{t_i^p}{\theta}} \prod_{i=1}^n t_i^{pk-1} \quad (2)$$

The likelihood equations are

$$\frac{\partial}{\partial \theta} \log_e L = -\frac{nk}{\theta} + \sum_{i=1}^n \frac{t_i^P}{\theta^2} = 0 \tag{3}$$

$$\frac{\partial}{\partial p} \log_e L = \frac{n}{p} + k \sum_{i=1}^n \log_e t_i - \sum_{i=1}^n \frac{t_i^P}{\theta} \log_e t_i \tag{4}$$

$$\frac{\partial}{\partial k} \log_e L = -n \log_e \theta - n \left(\log_e k - \frac{1}{2k} \right) + p \sum_{i=1}^n \log_e t_i \tag{5}$$

The MLE's of p, k and θ are solutions of these equations. It is quite difficult to obtain MLE's of p, k and θ independently, so we have obtained joint M.L.E.'s of two parameters by keeping the third parameter fixed (the joint MLE's of p and θ by keeping k constant and then joint MLE's of k and θ by keeping p constant).

If $k = k_0$ (constant), then (3) $\Rightarrow \theta = \frac{1}{nk_0} \sum_{i=1}^n t_i^P$ (6)

and (4) & (6) $\Rightarrow \frac{n}{p} + k_0 \sum_{i=1}^n \log_e t_i - \frac{nk_0 \sum_{i=1}^n t_i^P \log_e t_i}{\sum_{i=1}^n t_i^P} = 0$ (7)

This is the equation in one unknown parameter p only and may be solved for p by any numerical method, which is the MLE (p^*) of p. Then substituting the value of p^* in (6), we can find the MLE (θ^*) of θ .

Similarly, if $p = p_0$, then (3) $\Rightarrow \theta = \frac{1}{nk} \sum_{i=1}^n t_i^{P_0}$ (8)

and (5) & (8) $\Rightarrow \frac{1}{2k} = \log_e \left(\frac{\sum_{i=1}^n t_i^{P_0}}{\sum_{i=1}^n t_i^P} \right) - \log_e n - \frac{P_0}{n} \sum_{i=1}^n \log_e t_i$ (9)

Solving (9), we can find the MLE (k^*) of k and then substituting this value in (8) we can find the MLE (θ^*) of θ . The M.L.E.'s of reliability and hazard rate functions can also be obtained accordingly simply by using the property of invariance.

The MLE ($R^*(t)$) of reliability function R(t) (when $k = k_0$ is fixed) is given by

$$R^*(t) = \frac{1}{k_0} \int_{t^{P^*}}^{\infty} e^{-y} y^{k_0-1} dy \tag{10}$$

This function can be solved by using suitable numerical integration method. In particular, for k=1

$$R^*(t) = e^{-t^{P^*}} / \theta^* \tag{11}$$

The MLE ($H^*(t)$) of hazard rate function ($H(t)$) of T is given by

$$H^*(t) = \frac{\frac{p^*}{\theta^* k_0} t^{p^*} k_0^{-1} e^{-\frac{t^{p^*}}{\theta^*}} I_{(0, \infty)}(t)}{\frac{1}{k_0} \int_0^\infty e^{-y} y^{k_0 - 1} dy} \tag{12}$$

This function can also be solved by using suitable numerical integration method.

In particular, for $k=1$,
$$H^*(t) = \frac{p^*}{\theta^*} t^{p^* - 1} \tag{13}$$

MLE's of $R(t)$ and $H(t)$ (when $p = p_0$ is fixed) may be accordingly obtained.

3. Reparametrization of the Model

Prentice (1974) studied the model due to Stacy (1962) in a different but equivalent form, which makes the properties and potential difficulties with estimation in the model much more transparent. We have also reparameterized the proposed model in a way similar to that of Prentice (1974).

Let $Y = \log_e T$, Then the p.d.f. of Y is

$$g(y) = \frac{1}{b)k} \left[\exp \left\{ k \left(\frac{y-a}{b} \right) - e^{\left(\frac{y-a}{b} \right)} \right\} \right]; \quad -\infty < y < \infty \tag{14}$$

where $a = b \log_e \theta$ and $b = p^{-1}$.

Further if we consider a variate $W = (Y-a)/b$, then W has a log gamma distribution. As $k \rightarrow \infty$, the mean and variance of W become infinite, thus we make a further transformation and consider a variate $V = \sqrt{k}(W - \log_e k)$, where the p.d.f. of V is given by

$$g(v) = \frac{k^{k-1/2}}{)k} \exp \left[\sqrt{k}v - ke^{v/\sqrt{k}} \right] \quad -\infty < v < \infty, k > 0 \tag{15}$$

It can be shown that as $k \rightarrow \infty$, V approaches standard normal distribution. Now, V can also be expressed as $V = (Y - \mu)/\sigma$, where $\mu = a + b \log_e k$ and $\sigma = b/\sqrt{k}$.

From the above result, we can also find the p.d.f. $[h(y)]$ of y where

$$h(y) = \frac{k^{k-1/2}}{\sigma)k} \exp \left[\sqrt{k} \left(\frac{y - \mu}{\sigma} \right) - ke^{\left(\frac{y - \mu}{\sigma \sqrt{k}} \right)} \right]; \quad -\infty < y < \infty \tag{16}$$

Likelihood Function of (16) is given by

$$L = \frac{k^{n(k-1/2)}}{(\sigma \sqrt{k})^n} \exp \left(\sqrt{k} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) - k \sum_{i=1}^n e^{\left(\frac{y_i - \mu}{\sigma \sqrt{k}} \right)} \right) \tag{17}$$

$$\text{Log}_e L = n(k-1/2) \log_e k - n \log_e \sigma - n \log_e \sqrt{k} + \sqrt{k} \sum_{i=1}^n \left(\frac{y_i - \mu}{\sigma} \right) - k \sum_{i=1}^n e^{\left(\frac{y_i - \mu}{\sigma \sqrt{k}} \right)} \tag{18}$$

For fixed value of k, the likelihood equations for estimating μ and σ are

$$\frac{\partial}{\partial \mu} \text{Log}_e L = \sqrt{k} \sum_{i=1}^n \left[\frac{-1}{\sigma} \right] - k \sum_{i=1}^n e^{\left(\frac{y_i - \mu}{\sigma \sqrt{k}} \right)} \left[\frac{-1}{\sigma \sqrt{k}} \right] = 0 \tag{19}$$

$$\Rightarrow \mu = \sigma \sqrt{k} \log_e \left[\frac{1}{n} \sum_{i=1}^n e^{y_i / \sigma \sqrt{k}} \right] \tag{20}$$

$$\frac{\partial}{\partial \sigma} \text{Log}_e L = \frac{-n}{\sigma} - \frac{\sqrt{k}}{\sigma^2} \sum_{i=1}^n (y_i - \mu) - k \sum_{i=1}^n e^{\left(\frac{y_i - \mu}{\sigma \sqrt{k}} \right)} \left[\frac{-(y_i - \mu)}{\sigma^2 \sqrt{k}} \right] = 0 \tag{21}$$

$$\Rightarrow \frac{-n}{\sigma} - \frac{\sqrt{k}}{\sigma^2} n \bar{y} + \frac{n \mu \sqrt{k}}{\sigma^2} + \frac{n \sqrt{k}}{\sigma^2} \sum_{i=1}^n y_i \left(e^{y_i / \sigma \sqrt{k}} \right) \frac{1}{\sum_{i=1}^n e^{y_i / \sigma \sqrt{k}}} - \frac{n \mu \sqrt{k}}{\sigma^2} = 0 \tag{22}$$

$$\Rightarrow \frac{\sum_{i=1}^n y_i e^{y_i / \sigma \sqrt{k}}}{\sum_{i=1}^n e^{y_i / \sigma \sqrt{k}}} - \bar{y} - \frac{\sigma}{\sqrt{k}} = 0 \tag{23}$$

The MLE (σ^*) of σ may be obtained by solving (23) by suitable iterative procedure. Then by substituting the values of y 's and σ in (20), we get MLE (μ^*) of μ .

Now, for fixed value of p, the likelihood equations for estimating μ and k obtained from (18) by substituting $\sigma = 1/p\sqrt{k}$ give

$$\mu = \frac{1}{p} \log_e \left[\frac{1}{n} \sum_{i=1}^n e^{p y_i} \right] \tag{24}$$

$$k = \frac{1}{2 \left[\log_e \frac{1}{n} \sum_{i=1}^n e^{p y_i} - p \bar{y} \right]} \tag{25}$$

Similarly, the value of θ may be obtained by solving

$$\theta = e^{p\mu - \log_e k} \tag{26}$$

Illustration

The abovesaid procedures of obtaining MLE's of two parameters by keeping the remaining parameter fixed have been illustrated with the help of classified data of fifth bus motor failure due to Davis (1952) given in the Table-1.

Distance Interval (Thousand of Miles)	Observed Number of Failures
0-20	29
20-40	27
40-60	14
60-80	8
80-	7

Table-1 : Fifth bus failure data due to Davis (1952)

Direct Method: (i) For fixed value of k (k=1)

We have $\bar{t} = 35.1764706$, $n = 85$ and we write (7) for classified data as $g(p) = 0$, where

$$g(p) = \frac{\sum_{i=1}^n f_i t_i^p \log_e t_i}{\sum_{i=1}^n f_i t_i^p} - \frac{1}{p} - \frac{1}{N} \sum_{i=1}^n f_i \log_e t_i \quad \text{For (k=1)}$$

We note that $g(1.461) < 0$ and $g(1.462) > 0$, so we calculate $g(p)$ for intermediary values of p and repeat this process till $g(p) \approx 0$.

Final results

$$p^* \approx 1.46139812; \quad \theta^* = \frac{1}{N} \sum_{i=1}^n f_i t_i^p = 211.600546 \quad [\text{From (6)}]$$

$$g(p) = 1.6405E-10$$

(ii) For fixed value of p (p=1)

For classified data (6) $\Rightarrow \theta^* = \frac{1}{Nk} \sum_{i=1}^n f_i t_i = \bar{t} / k \quad [\text{For } p=1]$

And (7) $\Rightarrow \frac{1}{2k} = \log_e \left(\frac{\sum_{i=1}^n f_i t_i}{\sum_{i=1}^n f_i} \right) - \log_e N - \frac{1}{N} \sum_{i=1}^n f_i \log_e t_i$

or $k = \frac{1}{2 \left[\log_e \bar{t} - \frac{1}{n} \sum_{i=1}^n f_i \log_e t_i \right]}$

Final results

$$k^* = 1.78798073; \quad \theta^* = 19.6738533$$

Reparametrization Method: (i) For fixed value of k (k=1)

We have $\bar{t} = 35.17647059$ n = 85

We write (23) for classified data as $g_0(\sigma) = 0$ and calculate the values of $g_0(\sigma)$ for different values of σ , where

$$g_0(\sigma) = \frac{\sum_{i=1}^n y_i e^{y_i/\sigma}}{\sum_{i=1}^n e^{y_i/\sigma}} - \bar{y} - \sigma \quad \text{For (k=1)}$$

We note that $g_0(0.6875) < 0$ and $g_0(0.68) > 0$, so we calculate $g_0(\sigma)$ for intermediary values of p and repeat this process till $g_0(\sigma) \approx 0$.

Final results

$$\sigma^* = 0.684275; \quad \mu^* = \sigma \log_e \left[\frac{1}{n} \sum_{i=1}^n e^{y_i/\sigma} \right] = 3.664094677$$

We can also find the values of p^* and θ^* by using the relation

$$\mu = a + b \log_e k = p^{-1} \log_e \theta + p^{-1} \log_e k \quad \text{and} \quad \sigma = b / \sqrt{k} = p^{-1} / \sqrt{k}$$

For k=1, $\theta^* = 211.6027549$ and $p^* = 1.46140075$

It can be seen that the values of estimates of p and θ are almost equal in both the cases.

(ii) For fixed value of p (p=1)

Using (24), (25) and (26) for classified data of Table-1, we have

$$\mu = \log_e \left[\frac{1}{N} \sum_{i=1}^n f_i e^{y_i} \right] = 3.560377; \quad k = \frac{1}{2 \left[\log_e \frac{1}{N} \sum_{i=1}^n f_i e^{y_i} - \bar{y} \right]} = 1.787981;$$

$$\theta = e^{\mu - \log_e k} = 19.67385$$

In this case too, the values of estimates of k and θ are equal in both the cases. Moreover, the calculations need almost the same amount of labour. Thus, reparametrization of the model does not reduce the bulk and complexity of the calculations, as claimed by Prentice (1974).

4. Distribution of MLE (U) of θ given p and k

It is easy to see that for given p and k, MLE of θ is

$$U = \frac{1}{nk} \sum_{i=1}^n T_i^p \quad (27)$$

Moment generating function [$M_T(a)$] of T is

$$M_T(a) = \sum_{i=1}^{\infty} \frac{a^i}{i!} \theta \left/ \sqrt[p]{k + \frac{i}{p}} \right. \tag{28}$$

M.G.F. of U is

$$M_U(a) = E(e^{aU}) = \prod_{i=1}^n E(e^{aT_i^P / nk}) = \prod_{i=1}^n \left[1 - \frac{a\theta}{nk} \right]^{-k} = \left[1 - \frac{a\theta}{nk} \right]^{-nk} \tag{29}$$

which is the m.g.f. of Gamma distribution with parameters (nk / θ) and $(n k)$. Hence, by Uniqueness theorem, $U \sim G(nk / \theta, nk)$ and p.d.f. of U is

$$h(u) = \begin{cases} \left(\frac{nk}{\theta}\right)^{nk} \frac{1}{nk} e^{-\frac{nk}{\theta}u} u^{nk-1} ; & 0 < u < \infty \text{ and } \theta, k > 0 \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

r^{th} moment of U about origin is

$$\mu_r' = E(U^r) = \int_0^{\infty} \frac{u^r}{nk} \left(\frac{nk}{\theta}\right)^{nk} e^{-(nk/\theta)u} u^{nk-1} du = \frac{(nk+r)}{nk} \left(\frac{\theta}{nk}\right)^r \tag{31}$$

$$\text{Mean} = \mu_1' = \theta; \quad \text{Variance} = \mu_2' - (\mu_1')^2 = \left(\frac{\theta^2}{nk}\right); \quad \gamma_1 = \frac{2}{\sqrt{nk}}; \quad \gamma_2 = \frac{6}{nk} \tag{32}$$

The distribution of MLE (U) of θ for given p and k is positively skewed and leptokurtic.

5. A Test for the scale parameter θ

To find uniformly most powerful (UMP) test for testing the null hypothesis $H_0 : \theta = \theta_0$ against one sided alternative $H_1 : \theta > \theta_0$ or $\theta < \theta_0$, we use the well known Neyman-Pearson Lemma.

The problem is to test $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ where θ_1 may be $>$ or $<$ θ_0 .

Let T_1, T_2, \dots, T_n be a random sample from the given population with p.d.f.

(1), then

$$\begin{aligned} \frac{L_1}{L_0} > A &\Rightarrow \left(\frac{\theta_0}{\theta_1}\right)^{nk} e^{-\sum_{i=1}^n \frac{t_i^P}{\theta_1} + \sum_{i=1}^n \frac{t_i^P}{\theta_0}} > A \\ &\Rightarrow nk(\log_e \theta_0 - \log_e \theta_1) + \sum_{i=1}^n t_i^P \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) > \log_e A \\ &\Rightarrow \sum_{i=1}^n t_i^P \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) > A_0 \end{aligned}$$

where $A_0 = \log_e A + nk(\log_e \theta_1 - \log_e \theta_0)$ (33)

Case-I : $\theta_1 > \theta_0$, Then $U_0 > C$ is the critical region, where $C = A_0 \left(\frac{\theta_0 \theta_1}{\theta_1 - \theta_0} \right)$ is a

constant, and the value of C is determined such that

$$P[U_0 > C | H_0] = \alpha \quad (34)$$

α being a predetermined constant and $U_0 = \sum_{i=1}^n T_i^P$.

Case-II : $\theta_1 < \theta_0$, Then $U_0 < C'$ is the critical region, where the value of C' is determined such that

$$P[U_0 < C' | H_0] = \alpha \quad (35)$$

Now, since $U_0 = \sum_{i=1}^n T_i^P \sim G(1/\theta, nk)$, any decision about acceptance or rejection of H_0 may be taken on the basis of null distribution of U_0 . We have calculated right tail values of U_0 for various values of n, k and θ at different levels of significance (α). The values are given in Table 2.

Further, since $Z = 2\theta^{-1}U_0 \sim \chi^2(2nk)$, hence chi-square tables may also be used to determine critical values of U_0 . The value of Z may be read directly from the table of chi-square and then value of U_0 may be obtained by simply multiplying the value of Z by $\theta_0/2$. For the two sided alternative, Neyman-Pearson Lemma fails to provide uniformly most powerful test.

References

1. Davis, D.J. (1952). An analysis of some failure data. Journal of American Statistical Association, Vol. 47, pp 113-150.
2. Lawless, J.F. (1980). Inference in the generalized gamma and log gamma distribution. Technometrics, Vol. - 22, No. -3, pp 409-419.
3. Lawless, J.F. (1982). Statistical Models and Methods for Life Time Data. John Wiley & Sons, New York.
4. Prentice, R.L. (1974). A log-gamma model and its maximum likelihood estimation. Biometrika, Vol.-61, pp 539-544.
5. Shukla, G. and Kumar, V. (2006). Use of generalized gamma type model in life failure data. Indian Journal of Applied Statistics, Vol. 10, pp 13-20.
6. Sinha, S.K. (1986). Reliability and Life Testing, Wiley Eastern, New Delhi.
7. Stacy, E.W. (1962). A generalization of the gamma distribution, Annals of Mathematical Statistics, Vol.-33, pp 1187-1192.
8. Stacy, E.W. and Mihram, G.A. (1965). Parameter estimation for a generalized gamma distribution, Technometrics, Vol.-7, pp 349-358.

N	K	θ	α					
			0.99	0.95	0.90	0.10	0.05	0.01
2	1.1	1	0.1960	0.4384	0.6378	4.1846	5.0650	7.0070
		5	0.0391	0.0876	0.1276	0.8370	1.0130	1.4020
		10	0.0196	0.0438	0.0637	0.4184	0.5065	0.7007
		50	0.0039	0.0087	0.0127	0.0837	0.1013	0.1402
		100	0.0019	0.0043	0.0063	0.0418	0.0506	0.0700
	1.2	1	0.2488	0.5268	0.7484	4.4747	5.3798	7.3660
		5	0.0498	0.1054	0.1497	0.8950	1.0759	1.4732
		10	0.0248	0.0526	0.0748	0.4747	0.5379	0.7366
		50	0.0049	0.01053	0.0149	0.0895	0.1075	0.1473
		100	0.0024	0.0052	0.0074	0.0474	0.0539	0.0736
	1.3	1	0.3068	0.6198	0.8629	4.7608	5.6896	7.7187
		5	0.0614	0.1240	0.1726	0.9522	1.1380	1.5438
		10	0.0306	0.0619	0.0862	0.4760	0.5689	0.7718
		50	0.0061	0.0124	0.0172	0.0952	0.1138	0.1543
		100	0.0030	0.0061	0.0086	0.0476	0.0568	0.0771
	1.4	1	0.3693	0.7169	0.9810	5.0432	5.9947	8.0651
		5	0.0738	0.1434	0.1962	1.0090	1.1990	1.6131
		10	0.0369	0.0716	0.0981	0.5043	0.5994	0.8065
		50	0.0073	0.0143	0.0196	0.1009	0.1199	0.1613
		100	0.0036	0.0071	0.0098	0.0504	0.0599	0.0806
1.5	1	0.4361	0.8177	1.1021	5.3224	6.2958	8.4060	
	5	0.0872	0.1636	0.2204	1.0645	1.2591	1.6812	
	10	0.0436	0.0817	0.1102	0.5322	0.6295	0.8406	
	50	0.0087	0.0163	0.0220	0.0106	0.1259	0.0681	
	100	0.0043	0.0081	0.0110	0.0532	0.0629	0.0840	
3	1.1	1	0.5436	0.9751	1.2889	5.7359	6.7409	8.9090
		5	0.1088	0.1950	0.2578	1.1472	1.3482	1.7817
		10	0.0543	0.0975	0.1288	0.5735	0.6740	0.8909
		50	0.0108	0.0195	0.0257	0.1142	0.1348	0.1781
		100	0.0054	0.0097	0.0128	0.0573	0.0674	0.0890
	1.2	1	0.6588	1.1389	1.4812	6.1440	7.1790	9.4010
		5	0.1318	0.2278	0.2963	1.2288	1.4358	1.8802
		10	0.0658	0.1138	0.1412	0.6144	0.7179	0.9401
		50	0.0131	0.0227	0.0296	0.1228	0.1435	0.1880
		100	0.0065	0.0113	0.0141	0.0614	0.0717	0.0940
	1.3	1	0.7811	1.3087	1.6782	6.5473	7.6109	9.8854
		5	0.1562	0.2618	0.3357	1.3095	1.5222	1.9771
		10	0.0781	0.1308	0.1678	0.6547	0.7610	0.9885
		50	0.0156	0.0261	0.0335	0.1309	0.1522	0.1977
		100	0.0078	0.0130	0.0167	0.0654	0.0761	0.0988
	1.4	1	0.9097	1.4834	1.8793	6.9465	8.0376	10.362
		5	0.1819	0.2967	0.3759	1.3893	1.6078	2.0725
		10	0.0909	0.1483	0.1879	0.6946	0.8037	1.0362
		50	0.0181	0.0296	0.0375	0.1389	0.1607	0.2072
		100	0.0090	0.0148	0.0187	0.0694	0.0803	0.1036
1.5	1	1.0439	1.6628	2.0841	7.3419	8.4600	10.845	
	5	0.2088	0.3326	0.4168	1.4685	1.6919	2.1668	
	10	0.1043	0.1662	0.2084	0.7341	0.8460	1.0845	
	50	0.0208	0.0332	0.0416	0.1468	0.1691	0.2166	
	100	0.0104	0.0166	0.0208	0.0734	0.0846	0.1084	

Table-2: Right tail values of U_0 for various values of n, k and θ at different values of α