ON TWO PARAMETER MITTAG-LEFFLER COUNT MODEL

SEETHA LEKSHMI, V. CATHERINE THOMAS

Abstract: In this paper a new count model with Two parameter Mittag-Leffler Inter arrival time distribution is introduced and its properties are studied. Various characteristics like mean function, variance function, probability generating function, hazard rate function, etc are derived.The model is applied to a real data set on inter arrival times of customers in a bank counter at Muvattupuzha, Kerala, India.

Key words : Count model, Exponential distribution, Mittag-Leffler distribution, Mittag-Leffler count model.

Introduction : Poisson model is the basic regression model for count data. Poisson process is a stochastic process in continuous time and discrete state space, and it can be taken as a sequence of independently and identically exponentially distributed waiting times. (see Cox(1972). Poisson model is valid only in the case of equi dispersed data (Variance=Mean). Because of such limitations, statisticians developed many other count models for the over dispersed (negative duration dependence) as well as under dispersed (positive duration dependence) data. Negative duration dependence means decreasing hazard function and positive duration dependence means increasing hazard function. Weibull count model by Mc shane et al(2008), q-Weibull count model by Naik and Jose(2008), q-exponential count model by Seethalekshmi and Catherine(2012) (communicated), Gamma count model by Winkelmann (1995), Mittag-Leffler count model by Jose and Abraham(2011) etc are some count models which have been recently developed. Seethalekshmi and Catherine (2012) also developed some generalizations of Poisson Process. In this paper we develop a count model called two parameter Mittag Leffler count model which is a generalization of the one parameter Mittag-Leffler count model by Jose and Abraham(2011).

Mittag-Leffler Distribution: Mittag-Leffler introduced the function $E_\alpha(z)=\sum_{k=0}^{\infty} z^k/\Gamma(1+\alpha k)$ in 1903. Pillai(1990) introduced Mittag-Leffler distribution in terms of Mittag-Leffler functions as follows. He proved that the function

$$F_\alpha(z)=1-E_\alpha(-z^\alpha)=1-\sum_{k=0}^{\infty}(-z)^{\alpha k}/\Gamma(1+\alpha k),$$

$0<\alpha\leq 1$ is a distribution function having the Laplace Transform $\psi(t)=(1+t^\alpha)^{-1}$, $t>0$ which is completely monotone for $0<\alpha\leq 1$. He called $F_\alpha(x)$, $0<\alpha\leq 1$, as the distribution function of Mittag -Leffler.

Now we consider a two parameter Mittag -Leffler distribution.

A random variable X is said to follow two parameter Mittag -Leffler distribution denoted by $ML(\alpha,\nu)$, if its cumulative distribution function (c.d.f) has the form

$$F_\alpha(x;\alpha;\nu)=\sum_{k=1}^{\infty}(-1)^{k+1}(\nu x)^{k\alpha}/\Gamma(1+k\alpha), x>0, 0<\alpha\leq 1, \nu>0. \quad (1)$$

The probability density function is given by

$$f_\alpha(x;\alpha;\nu)=\sum_{k=1}^{\infty}(-1)^{k+1}k\alpha\nu(\nu x)^{k\alpha-1}/\Gamma(1+k\alpha),$$

$$x>0, 0<\alpha\leq 1, \nu>0 \quad (2)$$

The Laplace Transform of (1) is given by

$$\psi(t)=\nu^\alpha/(\nu^\alpha+t^\alpha)$$

When $\nu=1$, it reduces to one parameter Mittag-Leffler. When $\alpha=1$ it reduces to exponential.

The rest of the paper is organized as follows. In section 2, a two parameter count model is introduced and various properties are discussed. In section 3 the above model is applied to a real life data set.

A Two Parameter Mittag Leffler Count Model

A general frame work of the count model can be derived based on the relationship between inter arrival times and the number of events. Let W_n be the time from the measurement origin at which the n^{th} event occurs. Let $X(t)$ denotes the number of events that have occurred until time t. The relation between inter arrival times and the number of events is given by $W_n\leq t \Leftrightarrow X(t)\geq n$

$$Hence \quad P_n(t)=P[X(t)=n]=F_n(t)-F_{n+1}(t)$$

where $F_n(t)$ denote the c.d.f of W_n. It can be obtained as the n-fold convolution of the common inter arrival time distribution. Assume that the inter arrival times are independent and identically distributed as $ML(\alpha,\nu)$. To obtain (3) we have the recursive relationship

$$P_n(t) = \int_0^t F_{n-1}(t-s)f(s)ds - \int_0^t F_n(t-s)f(s)ds = \int_0^t P_{n-1}(t-s)f(s)ds \qquad (4)$$

Now $F_0(t) = 1$ for every t and $F_1(t) = F(t)$

$$Also \quad P_0(t) = F_o(t) - F_1(t) = \sum_{k=0}^{\infty}(-1)^k (\nu t)^{k\alpha}/\Gamma(1+\alpha k) \qquad P_1(t) = \int_0^t P_0(t-s)f(s)ds$$

$$= \int_0^t \sum_{j=0}^{\infty} \frac{(-1)^j [\nu(t-s)]^{j\alpha}}{\Gamma(1+j\alpha)} \sum_{k=1}^{\infty} \frac{(-1)^{k-1} k\alpha (\nu s)^{k\alpha-1}\nu}{\Gamma(1+k\alpha)} \qquad = \sum_{j=0}^{\infty}\sum_{k=1}^{\infty} \frac{(-1)^j (-1)^{k-1} (\nu t)^{k\alpha+j\alpha}}{\Gamma(1+j\alpha+k\alpha)}$$

Using change of variables, m=j and l=m+k, we get,

$$P_1(t) = \sum_{l=1}^{\infty} lC_1 \frac{(-1)^{l-1}(\nu t)^{l\alpha}}{\Gamma(1+l\alpha)}$$

$$P_2(t) = \sum_{l=2}^{\infty} \frac{lC_2 (-1)^{l-2}(\nu t)^{l\alpha}}{\Gamma(1+l\alpha)}$$

Hence a general form can be obtained as

$$P_{n+1}(t) = \int_0^t P_n(t-s)f(s)ds$$

$$= \sum_{l=n+1}^{\infty} \frac{lC_{n+1}(-1)^{l-(n+1)}(\nu t)^{l\alpha}}{\Gamma(1+l\alpha)}$$

Thus the p.d.f of the model can be given as

$$P[N(t)=n] = \sum_{j=n}^{\infty} \frac{jC_n(-1)^{(j-n)}(\nu t)^{j\alpha}}{\Gamma(1+j\alpha)}, n = 0,1,2,... \qquad (5)$$

The following tables gives the probabilities of the two parameter Mittag-Leffler count model for different values of α and ν.

					$\nu = 1$				
		$t=1$			$t=2$			$t=3$	
α	$P_1(t)$	$P_2(t)$	$P_3(t)$	$P_1(t)$	$P_2(t)$	$P_3(t)$	$P_1(t)$	$P_2(t)$	$P_3(t)$
0.1	0.2508	0.1290	0.0661	0.2500	0.1330	0.0705	0.2492	0.1351	0.0730
0.2	0.2533	0.1339	0.0697	0.2501	0.1412	0.0787	0.2471	0.1447	0.0837
0.3	0.2577	0.1398	0.0733	0.2502	0.1500	0.0874	0.2433	0.1537	0.0947
0.4	0.2642	0.1466	0.0766	0.2505	0.1596	0.0968	0.2377	0.1624	0.1063
0.5	0.2732	0.1544	0.0792	0.2510	0.1705	0.1073	0.2304	0.1709	0.1189
0.6	0.2852	0.1628	0.0807	0.2518	0.1832	0.1191	0.2207	0.1794	0.1329
0.7	0.3006	0.1713	0.0801	0.2534	0.1986	0.1326	0.2084	0.1881	0.1491
0.8	0.3197	0.1786	0.0768	0.2563	0.2176	0.1478	0.1929	0.1977	0.1686
0.9	0.3424	0.1834	0.0704	0.2615	0.2413	0.1642	0.1734	0.2090	0.1929
1.0	0.3679	0.1839	0.0613	0.2707	0.2707	0.1804	0.1494	0.2240	0.2240

.

					$\nu = 3$				
		$t = 1$			$t = 2$			$t = 3$	
α	$P_1(t)$	$P_2(t)$	$P_3(t)$	$P_1(t)$	$P_2(t)$	$P_3(t)$	$P_1(t)$	$P_2(t)$	$P_3(t)$
0.1	0.2493	0.1351	0.0730	0.2475	0.1385	0.0772	0.2462	0.1402	0.0796
0.2	0.2471	0.1447	0.0837	0.2401	0.1490	0.0914	0.2351	0.1505	0.0954
0.3	0.2433	0.1537	0.0947	0.2277	0.1560	0.1046	0.2169	0.1549	0.1086
0.4	0.2378	0.1624	0.1063	0.2105	0.1590	0.1161	0.1926	0.1525	0.1175
0.5	0.2304	0.1709	0.1189	0.1884	0.1571	0.1252	0.1631	0.1430	0.1209
0.6	0.2207	0.1794	0.1329	0.1617	0.1499	0.1310	0.1299	0.1261	0.1174
0.7	0.2084	0.1881	0.1491	0.1305	0.1364	0.1324	0.0947	0.1023	0.1056
0.8	0.1929	0.1977	0.1686	0.0952	0.1154	0.1277	0.0597	0.0726	0.0847
0.9	0.1734	0.2090	0.1929	0.0564	0.0855	0.1145	0.0275	0.0390	0.0542
1.0	0.1494	0.2240	0.2240	0.0149	0.0446	0.0892	0.0011	0.0050	0.0150

Table.1:Probabilities of Mittag-Leffler count model for ν =1,3

Mean function

Expected number of counts for the two parameter Mittag-Leffler Count model is

$$M(t) = E(N(t)) = \Sigma_{n=1}^{\infty} F_n(t) \tag{6}$$

Using Laplace transform technique,we get

$$M^*(s) = \frac{\nu^\alpha}{s^{\alpha+1}}$$

Hence $M(t) = \dfrac{(\nu t)^\alpha}{\Gamma(1+\alpha)}$, (7)

Variance function

The variance function can be obtained from the relation,

$$V(t) = E[N(t)]^2 - [E[N(t)]]^2$$

Second moment of $P_n(t)$ is

$$M_1(t) = E[N(t)]^2$$

$$= \sum_{n=0}^{\infty} n^2 [F_n(t) - F_{n+1}(t)]$$

$$M_1^*(s) = \frac{(s^\alpha + 2\nu^\alpha)(\nu^\alpha)}{s^{2\alpha+1}}$$

By Laplace Transform technique, $E[N(t)]^2 = t^\alpha \nu^\alpha [\dfrac{1}{\Gamma(1+\alpha)} + \dfrac{2t^\alpha \nu^\alpha}{\Gamma(1+2\alpha)}]$.

Hence $V(t) = t^\alpha \nu^\alpha [\dfrac{1}{\Gamma(1+\alpha)} + \dfrac{2t^\alpha \nu^\alpha}{\Gamma(1+2\alpha)}] - (\dfrac{(\nu t)^\alpha}{\Gamma(1+\alpha)})^2$ (8)

The mean function (7) and Variance function (8) are tabulated in the following table for different values of ν and α .

.

			$\nu=1$			
	$t=1$		$t=2$		$t=3$	
α	*Mean*	*Variance*	*Mean*	*Variance*	*Mean*	*Variance*
0.1	1.0511	2.1245	1.1266	2.3596	1.1732	2.5103
0.2	1.0891	2.1571	1.2511	2.6602	1.3568	3.0140
0.3	1.1142	2.1111	1.3718	2.8827	1.5492	3.4763
0.4	1.1271	2.0041	1.4872	3.0143	1.7490	3.8612
0.5	1.1284	1.8551	1.5958	3.0493	1.9544	4.1347
0.6	1.1192	1.6818	1.6963	2.9890	2.1636	4.2663
0.7	1.1005	1.4994	1.7878	2.8405	2.3746	4.2316
0.8	1.0737	1.3199	1.8694	2.6157	2.5856	4.0135
0.9	1.0398	1.1516	1.9403	1.3298	2.7947	3.6030
1	1	1	2	2	3	3

.

			$\nu=3$			
	$t=1$		$t=2$		$t=3$	
α	*Mean*	*Variance*	*Mean*	*Variance*	*Mean*	*Variance*
0.1	1.1732	2.5103	2.2574	2.7933	1.3094	2.9751
0.2	1.3568	3.0140	1.5585	3.7453	1.6902	4.2620
0.3	1.5492	3.4763	1.9073	4.8282	2.1540	5.8793
0.4	1.7490	3.8612	2.3079	5.9854	2.7142	7.8009
0.5	1.9544	4.1347	2.7640	7.1245	3.3851	9.9260
0.6	2.1636	4.2663	3.2794	8.1102	4.1826	12.0410
0.7	2.3746	4.2316	3.8576	8.7583	5.1236	13.7690
0.8	2.5856	4.0135	4.5019	8.8303	6.2268	14.5076
0.9	2.7947	3.6030	5.2152	8.0298	7.5119	13.3515
1	3	3	6	6	9	9

. Table.2:Mean and Variance of Mittag-Leffler Count model for ν =1,3

Probability generating function

The probability generating function of the two parameter Mittag-Leffler count model is given by

$$P(t,s,\nu)=\Sigma_{n=0}^{\infty}s^n P[N(t)=n] \qquad =\Sigma_{n=0}^{\infty}s^n\sum_{j=n}^{\infty}\frac{jC_n(-1)^{(j-n)}(\nu t)^{j\alpha}}{\Gamma(1+j\alpha)}$$

$$=\Sigma_{j=0}^{\infty}(-1)^j[\nu t(1-s)^{\frac{1}{\alpha}}]^{j\alpha} \qquad =1-F[\nu t(1-s)^{\frac{1}{\alpha}}]$$

Hazard rate function

The hazard rate function is given by, $h(t)=\frac{f(t)}{\overline{F}(t)}$

$$=\frac{\sum_{k=1}^{\infty}(-1)^{k+1}k\nu\alpha(\nu t)^{k\alpha-1}/\Gamma(1+k\alpha)}{\sum_{k=0}^{\infty}(-1)^k(\nu t)^{k\alpha}/\Gamma(1+k\alpha)}$$

The following figure displays the hazard rate function of two parameter Mittag Leffler count model for $\alpha=0.2$ and 0.9 ,for $\nu=1,3,0.4$. Clearly it is a decreasing function of time t.

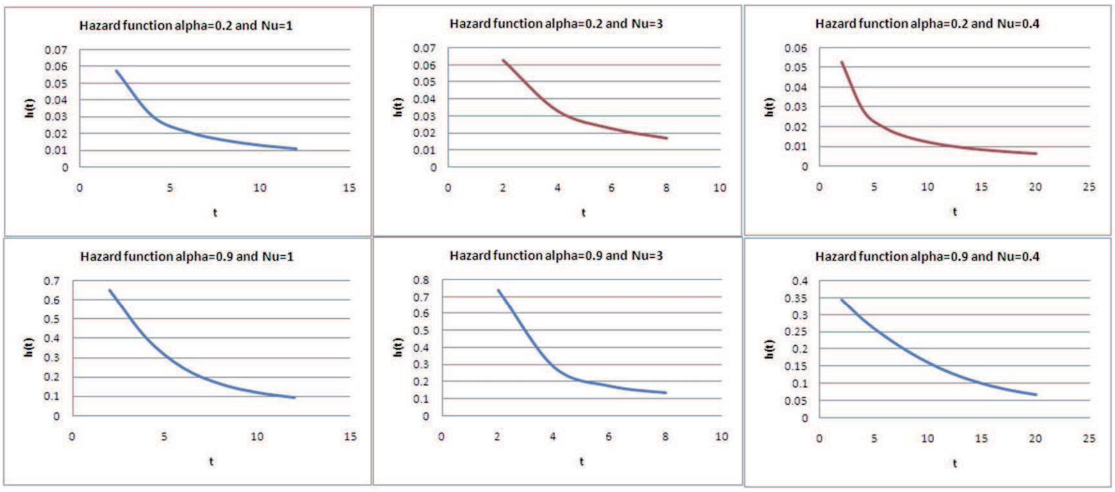

Figure.1 Hazard function of two parameter Mittag-Leffler count model.

If the hazard function is a decreasing function of time, the distribution displays negative duration dependence. If the hazard function is an increasing function of time, the distribution displays positive duration dependence. In both cases the conditional probability of a current occurrence depends on the time since the last occurrence rather than on the number of previous events.

There is a link between duration dependence and dispersion. Without making the assumptions on the exact distribution of τ , a limiting result can be obtained. Denote the mean and the variance of the waiting time distribution by $E(\tau)=\mu$ and $Var(\tau)=\sigma^2$ and the coefficient of variation by $\nu=\frac{\sigma}{\mu}$. Let $\lambda(t)$ be the hazard function. The distribution displays negative duration dependence for $\frac{d\lambda(t)}{dt}<0$ and positive duration dependence for, $\frac{d\lambda(t)}{dt}>0$. Assume that the hazard function is monotonic. Then by Barlow and Proschan (1965), $\frac{d\lambda(t)}{dt}<0=0>0 \Rightarrow \nu(>1=1<1)$

Application to a real data set

In this section we apply the model to a data on the inter arrival time of customers arrived on a given day in 2012 at a bank counter in Muvattupuzha, Kerala, India.All inter arrival times are expressed in minutes and total number of customers arrived in the bank on that day is 300. Here the conditional mean is smaller than the conditional standard deviation (mean=4.16 and S.D=4.37). Thus this data set is over dispersed and hence we can apply the $ML(\alpha,\nu)$ count model.Using fractional method of estimation of Mittag-Leffler parameters (Kozubowski 2001), the values of the parameters are given by $\alpha=.95$ and $\nu=.27$. To test whether there is significant difference between the observed inter arrival time distribution and two parameter Mittag-Leffler distribution, we use the Kolmogrov Smirnov test. Let H_o : the data follows Two parameter Mittag-Leffler distribution . The calculated value of the Kolmogrov Smirnov test statistic is 0.05723 and critical value at 5 percent level of significance is 0.077 showing that the two parameter Mittag Leffler assumption for inter arrival times is valid

.

P-P plot of the Two parameter Mittag Leffler distribution is given in the following figure.

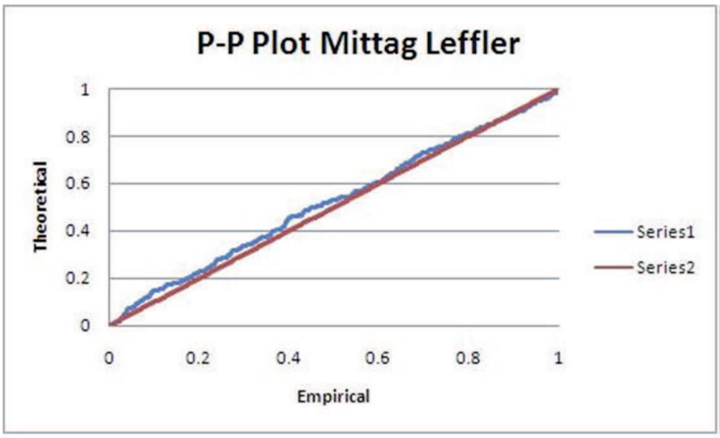

Figure 2: P-P plot of two parameterMittag-Leffler distribution.

The P-P plot is very close to the straight line joining (0,0) and (1,1). This shows that the Mittag Leffler distribution is a good fit to the data. Thus based on Kolmogrov Smirnov statistic and P-P plot we conclude that the null hypothesis of Mittag Leffler distribution as a good fit to the data is acceptable.

References:

1. Barlow, R.E. Proschan F. (1965), Mathematical Theory of Reliability, *John Wiley New York*.
2. Cox, D.R.(1972), Renewal Theory, *John Wiley New York*.
3. Jose, K.K. Abraham, B. (2011), A Count model based on Mittag Leffler inter arrival times, *Statistica*.,4,501-514.
4. Jose, K.K. Naik, S.R.(2008),Pathway Distributions,Auto regressive Processes and their Applications. (unpublished Ph.D. Thesis Dec 2008).
5. Kozubowski,T.J.(2001),Fractional moment estimation of Linnik and Mittag-Leffler Parameters,*Mathematical and Computer Modelling*, 34,1023-1035.
6. Mcshane,B. Adrian,M. Barlow,E.T. Fader,P.S.(2008), Count models based on Weibull interarrival times,*Journal of Business and Economic Statistics,* 26, No.3, 369-378.
7. Pillai, R.N.(1990), On Mittag-Leffler and related distributions, *Ann. Inst.Statist.Math.* ,42.No. 1,157-161.
8. Seetha Lekshmi, V. and Catherine, T. (2012),Some generalizations of Poisson processes.*Journal of Statistical Theory and Applications*,Vol 11,No.3,225-235)
9. Seetha Lekshmi, V. and Catherine, T. (2012),On q-Exponential Count Models, *Journal of Applied Statistical Science* (Communicated)
10. Winklemann, R. (1995), Duration Dependence and Dispersion in count-Data Models,*Journal of Business and Economic Statistics*,13,467-474.

***

Dept. of Statistics, Nirmala College, Muvattupuzha, Kerala, India
[2] Dept. of Statistics, Govt. College, Tripunithura , Kerala, India