Chapter 3: A Comparison between Gompertz and Makeham Law of Mortality for Projecting Survivors of Assam Ppoluation

3.1 Introduction

The relation between mortality and age is the most established subject in demography. The spearheading work of [[17], [25], [27]] set up the life table as vital and explanatory tool. The quest for a mathematical model of age variety in mortality dangers (mortality law) likewise has a long history. Mortality modelling is one of the conventional and major demographic issues. The first informative model, and the most persuasive parametric mortality modelling, was that suggested by Benjamin Gompertz [23] in 1825. He noticed that an exponential pattern in age captured the behaviour of human mortality for large portions of the life table [30]. Gompertz's model was really intended to speak to just "fundamental" mortality, i.e. mortality cleansed of accidental or irresistible causes. Keeping in mind the end goal to incorporate these two arrangements of mortality causes which are accepted to act freely of age; Makeham [46] improved on the Gompertz law by adding a further term which does not depend on age. Gompertz and Makeham models are still regularly used to smooth data, particularly at older ages [32]. Since the time that Gompertz, several models were suggested mathematically to describe survival and mortality rates. The Gompertz model and the Weibull model are the most generally used at present [[21], [22]] for this purpose. But in case of our Assam mortality, Weibull model does not give the fit. In this chapter Gompertz and Makeham models have been analysed.

3.1.1 The Models Used for Extrapolating Mortality Curves

3.1.1.1 Gompertz Model:

Gompertz first observed that a law of geometric progression pervades, after a certain age, in many population and modelled the mortality risk as:

$$\mu_x = ae^{bx}. \tag{3.1}$$

Since, $l_x = \exp(-\int \mu_x dx)$.

Therefore $$l_x = kg^{h^x}, \tag{3.2}$$

where k, g and h are parameters; $k = e^{c_1}$,. c_1 is an integrating constant, $g = e^{-\left(\frac{a}{b}\right)}$ and $h = \exp b$.

The equation (**3.2**) has been used to project the l_x values in a life table, where l_x denotes the number of persons living at any specified age x.

3.1.1.2 Makeham Model:

Makeham [46] suggested adding a constant term to the Gompertz model to include accidental or infectious cause of death as;

$$\mu_x = ae^{bx} + c. \tag{3.3}$$

The parameter c denotes the mortality resulting from causes, such as accidents or sexually transmitted diseases, unrelated to either maturation or senescence.

Since, $l_x = \exp(-\int \mu_x dx)$.

Therefore $l_x = ks^x g^{h^x}$, (3.4)

where k, g and h are parameters; $k = e^{c_1}$,. c_1 is an integrating constant, $s = \exp(-c)$, $g = e^{-\left(\frac{a}{b}\right)}$ and $h = \exp b$.

The equation **(3.4**) is used to graduate the l_x values in a life table. Where l_x denotes the number of persons living at any specified age x.

In this chapter, a comparison has been made between the Gompertz and Makeham model for projecting survivors up to the last age in a life table for Assam. In this investigation, the total, rural and urban areas are considered for both male and female. The parameters of the model have been estimated using different methods of estimation. These models can simplify the task of preparing demographic projections.

3.2 Objectives

In this Chapter, the main focus is to select the best fit mortality model to extrapolate survivors for Assam for total, rural and urban population for both the genders. The Gompertz and Makeham models have been examined for extrapolating survivors in a life table past beyond the last age. Using the abridged life tables of Assam for the period 2009-13 as input, the parameters of the mortality models are estimated. The parameters of these two models have been estimated using two methods of estimation.

3.3 Methods and Materials

3.3.1 Estimation of Parameters of Gompertz model

3.3.1.1 Method of Three Equidistant Points:

The general equation of Gompertz model for (l_x) with three parameters is given by,

$$l_x = k.g^{h^x}, \quad (3.5)$$

where k, g and h are parameters to be estimated.

The model takes the linear form after taking log on both sides of the equation **(3.5**). Then,

$$Y = A + Bh^x, \tag{3.6}$$

where $Y = \log l_x$, $A = \log k$, and $B = \log g$.

Now, let us consider x_1, x_2, x_3 be the three equidistant points from the survival function of the life table. Then from the equation **(3.6)**,

$$Y_1 = A + Bh^{x_1}, \tag{3.7}$$

$$Y_2 = A + Bh^{x_2}, \tag{3.8}$$

$$Y_3 = A + Bh^{x_3}, \tag{3.9}$$

Using the equations from (**3.7**) to **(3.9)** and after simplification the parameters $k, g,$ and h of the Gompertz model can be estimated as

$$\hat{h} = \left(\frac{Y_3 - Y_2}{Y_2 - Y_1}\right)^{1/m}, \tag{3.10}$$

$$\hat{g} = exp\left(\frac{(Y_2 - Y_1)^2}{Y_3 - 2Y_2 + Y_1}\left(\frac{Y_2 - Y_1}{Y_3 - Y_2}\right)^{\frac{x_1}{m}}\right), \tag{3.11}$$

$$\hat{k} = Y_1 - \left(\frac{(Y_2 - Y_1)^2}{Y_3 - 2Y_2 + Y_1}\left(\frac{Y_2 - Y_1}{Y_3 - Y_2}\right)^{\frac{x_1}{m}}\right), \tag{3.12}$$

where $y_i = \ln x_{x_i}$ for $i = 1, 2$ and 3 and $m = x_2 - x_1 = x_3 - x_2$.

The parameters $k, g,$ and h of the Gompertz model have been estimated using the equations (**3.10**), (**3.11**) and (**3.12**).

3.3.1.2 Method of Three Partial Sums:

In this method, the range of observations is divided into three equal parts. That is if the number of observations is n then take m such that $m = \frac{n}{3}$.

Now let $s_1$be the sum of first m observations, s_2 be the sum of second observations and s_3 be the last observations. The parameters $k, g,$ and h of the Gompertz model estimated using the method of three partial sums can be given as

$$\hat{k} = \exp\left[\frac{s_1 s_3 - s^2{}_2}{m(s_3 - 2s_2 + s_1)}\right], \tag{3.13}$$

$$\hat{g} = \exp\left[\frac{(s_2 - s_1)^2}{s_3 - 2s_2 + s_1}\left(\frac{s_2 - s_1}{s_3 - s_2}\right)^{\frac{m+1}{2m}}\right], \tag{3.14}$$

$$\hat{h} = \left(\frac{s_3 - s_2}{s_2 - s_1}\right)^{\frac{1}{m}}. \tag{3.15}$$

3.3.2 Estimation of Parameters of Makeham Model

3.3.2.1 **Method of Four Equidistant Points*:***

The general equation of Makeham model for (l_x) is given by,

$$l_x = ks^x g^{h^x}, \tag{3.16}$$

where $k,\ g,\ s$ and h are parameters.

In terms of the logarithms of l_x, the equation **(3.16)** can be written as

$$\log l_x = \log k + x \log s + h^x \log g\,.$$

In this method, consider four equidistant points x_1, x_2, x_3 and x_4 such that the distance between two consecutive points is m. Then after simplifying the parameters $\hat{k},\ \hat{g},\ \hat{s}$ and $\hat{h}$ can be estimated as

$$\hat{h} = \left(\frac{d_3 - d_2}{d_2 - d_1}\right)^{\frac{1}{m}}, \tag{3.17}$$

$$\hat{g} = \exp\left(\frac{d_1 - d_2}{(1 - h^{-m})^2 h^{t_4}}\right), \tag{3.18}$$

$$\hat{s} = \exp\left\{\frac{1}{m}\left(d_1 - h^{t_4}(1 - h^{-m})\frac{d_1 - d_2}{(1 - h^{-m})^2 h^{t_4}}\right)\right\}, \tag{3.19}$$

$$\hat{k} = l_4 exp\left\{\frac{t_4}{m}(u_2 - d_1) - h^{t_4} u_1\right\}, \tag{3.20}$$

where $u_1 = \frac{d_1 - d_2}{(1 - h^{-m})^2 h^{t_4}}$ and $u_2 = h^{t_4}(1 - h^{-m})u_1$.

3.3.2.2 Method of Four Partial Sums:

In this method, the total number of observations is divided into four equal parts.

Now let s_0 be the sum of first m observations, s_1 be the sum of second m observations and s_3 be sum of third m observations and s_3 be the sum of fourth m observations. Let us take

$$d_1 = s_1 - s_0,$$

$$d_2 = s_2 - s_1,$$

$$d_3 = s_3 - s_2.$$

Then after calculating the nonlinear parameter estimations for Makeham model are:

$$\hat{c} = \left(\frac{d_3 - d_2}{d_2 - d_1}\right)^{\frac{1}{n}}, \tag{3.21}$$

$$\hat{g} = \exp\left(\frac{(d_2 - d_1)(\hat{c} - 1)}{(\hat{c}^n - 1)^3}\right), \tag{3.22}$$

$$\hat{s} = \exp\left\{\frac{1}{n^2}(d_1 - u_4)\right\}, \tag{3.23}$$

$$\hat{k} = \exp\left\{\frac{1}{n}\left(s_0 - \frac{n(n-1)}{2}u_3 - \left(\frac{\hat{c}^n - 1}{\hat{c} - 1}\right)\log\hat{g}\right)\right\}, \tag{3.24}$$

where $u_4 = u_3 \times \frac{(\hat{c}^n - 1)^2}{\hat{c} - 1}$.

The above equations can be used for estimating the parameters of the Makeham model by the method of four partial sums.

3.3.3 Methodology for Calculating Life Expectancy at Age x ($e_x{}^0$)

After estimating the survivors l_x the expectation of life $e_x{}^0$ at age x is obtained from the relation

$$e_x{}^0 = \frac{T_x}{l_x}, \tag{3.25}$$

Where

T_x: Total number of person-years lived after the age $x = L_x + L_{x+1} + L_{x+2} + \cdots$.

n^{L_x}: No of person-years lived by the l_x persons during the age interval $(x, x+1) = n * l_{x+n} + \frac{n}{2} n^{d_x}$.

n^{d_x}: No of persons who attain age and die before reaching the age $x + 1 = l_x - l_{x+n}$.

The life expectancies $e_x{}^0$ at age x can be estimated using the formula (**3.25**).

3.4 Results and Discussion

The Gompertz and Makeham model have been fitted for Assam population for projection of mortality at oldest old ages. The estimated survivors and observed values along with the estimated parameters are presented in **Table 3.1** to **Table 3.6** for total, rural and urban area for both male and female. After fitting the models the RMSE and R^2 have been evaluated for each cases and presented are in **Table 3.7** and **Table 3.8**.

Table 3.1: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for female in rural area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	94126	91494	91408	93193	92968
5	91355	91355	91249	92258	92101
10	90818	91155	91022	91315	91229
15	90370	90866	90699	90356	90345
20	89322	90452	90241	89363	89433
25	88314	89857	89591	88314	88472
30	87506	89006	88673	87169	87425
35	86410	87791	87381	85869	86229
40	84885	86068	85571	84318	84784
45	83641	83641	83055	82372	82928
50	80150	80259	79595	79810	80403
55	76309	75618	74911	76309	76815
60	70866	69390	68704	71425	71594
65	64376	61296	60734	64602	64018
70	52315	51251	50943	55302	53397
75	40639	39585	39649	43332	39632
80	25217	27269	27736	29452	24187
85	15926	15926	16664	15927	10630
90		7330	8060	5965	2693
95		2392	2861	1238	271
100		475	654	99	6
105		46	80	2	0
110		2	4	0	0
115		0	0	0	0
Parameters	h	1.443	1.426	1.607	1.677
	g	0.998	0.997	0.999	0.999

	s	------	------	0.990	0.991
	k	91808.9	91784.5	93239.34	92996.77

Table 3.2: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for male in rural area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	94383	92395	92483	93682	93520
5	92217	92217	92295	92960	92828
10	91863	91959	92025	92202	92102
15	91323	91586	91637	91388	91322
20	90617	91047	91079	90487	90457
25	89451	90271	90279	89451	89458
30	88080	89158	89138	88209	88253
35	86755	87567	87515	86654	86732
40	85002	85309	85224	84632	84731
45	82136	82136	82023	81920	82012
50	77762	77742	77612	78210	78245
55	73103	71783	71659	73103	72994
60	64274	63940	63862	66143	65760
65	55954	54060	54077	56926	56111
70	44759	42377	42535	45369	43998
75	30510	29767	30076	32130	30249
80	17132	17832	18236	18977	16949
85	8479	8479	8856	8479	6907
90		2884	3121	2468	1715
95		603	693	372	197
100		62	79	20	7
105		2	3	0	0
110		0	0	0	0

115		0	0	0	0
Parameters	h	1.451	1.444	1.537	1.556
	g	0.997	0.997	0.998	0.999
	s			0.993	0.993
	k	92792	92907	93826	93647

Table 3.3: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for female in total area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	94403	92014	91897	93538	93375
5	91887	91887	91754	92672	92563
10	91383	91704	91550	91799	91746
15	90950	91439	91258	90911	90919
20	89981	91056	90841	89993	90067
25	89024	90503	90246	89024	89171
30	88262	89707	89400	87967	88196
35	87248	88565	88201	86766	87085
40	85828	86934	86509	85330	85742
45	84622	84622	84138	83520	84011
50	81432	81378	80852	81118	81642
55	77802	76892	76362	77802	78243
60	72528	70821	70353	73115	73237
65	66186	62857	62553	66472	65864
70	54365	52870	52850	57269	55354
75	42438	41136	41502	45214	41474
80	27074	28582	29344	30978	25580
85	16860	16860	17849	16859	11345
90		78393	8749	6313	2878

95		2582	3147	1288	283
100		516	726	98	6
105		50	89	2	0
110		2	4	0	0
115		0	0	0	0
Parameters	h	1.450	1.434	1.624	1.695
	g	0.998	0.998	0.999	0.999
	s			0.991	0.991
	k	92295.2	92225.7	93577	93398

Table 3.4: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for male in total area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	94615	92788	92849	93953	93827
5	92614	92614	92667	93270	93171
10	92262	92362	92404	92552	92481
15	91738	91998	92028	91779	91737
20	91043	91474	91488	90920	90908
25	89930	90720	90716	89930	89947
30	88675	89639	89614	88739	88784
35	87307	88096	88050	87245	87312
40	85564	85909	85843	85296	85374
45	82837	82837	82759	82677	82742
50	78709	78582	78509	79089	79098
55	74142	72806	72764	74142	74025
60	65846	65184	65215	67380	67037
65	57652	55539	55692	58389	57697
70	46524	44043	44360	47033	45897
75	32487	31479	31958	33871	32320

80	18834	19355	19922	20541	18847
85	9570	9570	10082	9570	8207
90		3451	3777	2975	2274
95		788	918	497	313
100		93	119	32	15
105		4	6	0	0
110		0	0	0	0
115		0	0	0	0
Parameters	h	1.448	1.441	1.534	1.548
	g	0.997	0.997	0.998	0.999
	s			0.9935	0.9938
	k	93179	93263	94095	93958

Table 3.5: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for female in urban area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	96749	96471	96220	96475	96840
5	96409	96409	96160	96194	96488
10	96173	96315	96069	95901	96126
15	95871	96173	95933	95589	95750
20	95488	95960	95728	95246	95346
25	94850	95637	95420	94850	94896
30	94306	95152	94957	94369	94371
35	93690	94423	94265	93749	93719
40	92753	93332	93232	92904	92857
45	91706	91706	91697	91697	91650
50	89925	89302	89431	89910	89881
55	87212	85786	86124	87212	87209
60	83210	80732	81371	83114	83121

65	77771	73652	74702	76954	76904
70	67723	64109	65673	67962	67711
75	54214	51977	54088	55554	54886
80	40704	37851	40377	40016	38761
85	23435	23435	25994	23434	21750
90		11352	13389	9782	8318
95		3795	4928	2346	1678
100		724	1093	227	117
105		59	113	5	1
110		1	4	0	0
115		0	0	0	0
Parameters	h	1.512	1.506	1.636	1.668
	g	0.999	0.999	0.999	0.999
	s			0.9973	0.9965
	k	96593	96339	96505	96863

Table 3.6: Estimated survivors using Gompertz and Makeham model along with the estimated parameters for male in urban area.

Age	Observed	Gompertz		Makeham	
		Method I	Method II	Method I	Method II
1	96771	96362	96240	96469	96682
5	96208	96208	96088	96138	96328
10	95900	95988	95871	95763	95929
15	95522	95673	95562	95323	95462
20	94879	95224	95122	94789	94896
25	94114	94586	94497	94114	94186
30	93514	93679	93612	93234	93266
35	91885	92397	92361	92055	92043
40	90169	90591	90602	90444	90384
45	88069	88069	88147	88213	88108

50	85050	84580	84754	85109	84969
55	80806	79826	80131	80806	80655
60	74988	73482	73961	74916	74799
65	67485	65269	65959	67049	67034
70	56945	55085	56005	56954	57127
75	44940	43210	44331	44775	45213
80	31503	30525	31744	31385	32095
85	18562	18562	19697	18562	19415
90		9108	9960	8534	9284
95		3287	3760	2702	3142
100		765	935	492	639
105		95	128	40	62
110		5	7	1	2
115		0	0	0	0
Parameters	h	1.431	1.429	1.482	1.470
	g	0.997	0.997	0.998	0.998
	s			0.9975	0.9974
	k	96721	96596	96663	96897

From our results it is observed that the estimation of the parameter k for urban area population is much bigger than total and rural area population for both male and female. The values of the parameter g are almost identical for total, rural and urban area for both male and female. It is likewise watched that for total and urban area female population the estimation of the parameter h is somewhat more prominent than male. In case of rural area, the value of the parameter h is greater for male than female. It is also observed that, the estimation of the parameter g is almost identical for total, rural and urban area for both male and female. The values of the parameter s for total, rural and urban area are larger for male than female. But the estimations of the parameter h are smaller for male than female for total, rural and urban zones. The evaluated values for k for urban area population is larger than total and rural area population for both male and female.

Table 3.7: Estimated values of $\boldsymbol{R^2}$ for Gompertz and Makeham model with Method I and Method II.

Sex	Area	Gompertz		Makeham	
		Three equidistant points method	Three partial sums method	Four equidistant points method	Four partial sums method
Male	Total	0.9984	0.9985	0.9993	0.9996
	Rural	0.9985	0.9986	0.9990	0.9995
	Urban	0.9984	0.9993	0.9999	0.9998
Female	Total	0.9962	0.9956	0.9963	0.9960
	Rural	0.9964	0.9956	0.9960	0.9964
	Urban	0.9942	0.9971	0.9996	0.9992

The estimation of R^2 is evaluated for all the technique for estimation for each model and is presented in **Table 3.7**. It is clear from the table that for all cases R^2 value is significant. R-squared is a statistical measure of how close the data are to the fitted regression line. In general, the higher the R-squared, the better the model fits your data.

From **Table 3.8** it is observed that value of the RMSE is least for Makeham model when contrasted with Gompertz model. The Makeham model has been fitted by two methods of estimation namely the method of four equidistant points and the method of four partial sums. It is also seen that the method of four partial sums gives better result for total and rural area for male population. In case of urban area male population, the method of four equidistant points performed well than the other method. For total and urban area female population, the method of four equidistant points seems better RMSE than the method of four partial sums. The method of four partial sums gives better RMSE for rural area female population. In case of rural area female population Gompertz model also give satisfactory result.

Table 3.8: Estimated values of $\boldsymbol{RMSE}$ for Gompertz and Makeham model with Method I and Method II.

Sex	Area	Gompertz		Makeham	
		Three equidistant points method	Three partial sums Method	Four equidistant points method	Four partial sums Method
Male	Total	1055	1004	715	523
	Rural	1026	1019	831	621
	Urban	965	644	183	300
Female	Total	1404	151	1399	1443
	Rural	1397	1547	1469	1393
	Urban	1738	1221	433	663

The projected survivors from the age 90 to the last age are presented in **Table 3.9** and **Table 3.10**. The projected life expectancies at oldest old ages using best fit model (Makeham model) is given in **Table 3.11**.

Table 3.9: Projection of l_x values using Makeham model for male in Assam.

	Four equidistant points			Four partial sums		
Age	Total	Rural	Urban	Total	Rural	urban
90	2975	2468	8534	2274	1715	9284
95	497	372	2702	313	197	3142
100	32	20	492	15	7	639
105	0	0	40	0	0	62
110	0	0	0	0	0	2
115	0	0	0	0	0	0

It is seen that the number of survivors for urban area is greater than rural area for both the method of estimations. It is also remarkable that only a man from urban area can expect to live at age 105 while other area can expect to live at age 100.

Table 3.10: Projection of l_x values using Makeham model for female in Assam.

Age	Four equidistant points			Four partial sums		
	Total	Rural	Urban	Total	Rural	urban
90	6313	5965	9782	2878	2693	8318
95	1288	1238	2346	283	271	1678
100	98	99	227	6	6	117
105	2	2	5	0	0	1
110	0	0	0	0	0	0
115	0	0	0	0	0	0

It is seen from **Table 3.10** that the same fact is happened for female also. That is, the number of survivors for urban area is greater than rural area. It is also remarkable that the projected values of l_x at age 105 are nonzero for urban male population while for total and rural are zero.

Table 3.11: Projected Life Expectancy at Older Ages using Makeham model for Assam.

Age	Male			Female		
	Total	Rural	Urban	Total	Rural	Urban
90	3.09	3.29	4.39	3.60	3.62	3.82
95	2.68	2.77	3.48	2.89	2.91	2.99
100	2.50	2.50	2.91	2.60	2.60	2.61
105	0	0	0	0	2.50	0
110	0	0	0	0	0	0
115	0	0	0	0	0	0

3.5 Conclusion

In this study, two mortality models, in particular, Gompertz and Makeham models are analyzed for extrapolating survivors in a life table past the last age for Assam for total, rural and urban populace for both the sexual orientations. The best fit model is selected based on the premise of RMSE and R^2 value. As a result of the comparison, it can be concluded that Makeham model lead to the best results for projecting the survivors for Assam for total, rural and urban population for both male and female. From the obtained results it is seen that the number of survivors for urban area is greater than rural area. A woman in Assam has higher life expectancy at ages 90, 95, 100 than her male counterpart within the State in rural and total areas but a woman in Assam from urban area has lower life expectancy than her male at the above age group.

*****