Advances and Applications in Mathematical Sciences
Volume 21, Issue 11, September 2022, Pages 6179-6186

A STUDY OF MORTALITY MODELS FOR FEW DISTRICTS OF ASSAM

PALLABI SAIKIA and DIMPAL JYOTI MAHANTA

[1]Nanda Nath Saikia College
Department of Mathematics
Titabar, Assam, India, PIN 785630

[2]Assam Kaziranga University
Department of Mathematics
Assam, India, PIN 785006
E-mail: dimpaljmahanta@gmail.com

Abstract

The objective of this paper is to find a suitable mortality model for few districts of Assam. For our present analysis, we consider five high population growth rate districts: Darang, Dhubri, Goalpara, Marigaon and Nagaon and five low population growth rate districts: Golaghat, Jorhat, Sibsagar, Dibrugarh, and Tinsukia. The most commonly used parametric models namely Gompertz, Makeham, Logistic and Beard are used for describing the mortality pattern of districts of Assam. The Levenberg-Marquardt method is used to estimate the parameters. From our analysis, it can be concluded that logistic model and Makeham model are provided better results over the remaining models for male and female respectively.

1. Introduction

Mortality modeling is an old topic, which is classical and needful statistic issues. The provocation behind the mortality modeling is to detect relations and hidden discrepancies and patterns in the mortality development. Numerous attempts have been made to find mathematical formulae that will epitomize the way in which the probability of dying depends on age. Similar formulae have numerous potential applications. For illustration, they may be

2020 Mathematics Subject Classification: Primary 93A30, Secondary 92B15.
Keywords: Age-specific mortality; Gompertz; Makeham; Logistic; Beard model; Parameter estimation.
[1]Corresponding author; E-mail: pallabisaikia.math@gmail.com
Received January 12, 2022; Accepted March 5, 2022

applied in the projection of population figures and as aids in actuarial work such as the construction of life tables. Ever since Gompertz, numerous models have been suggested to mathematically describe survival and mortality curves (1), one of the foremost and still one of the most meaningful examples of such a formula is the "law of mortality" which was introduced by Gompertz (2). In this, the force of mortality would grow exponentially with age, and the alternate adds a constant, an age-independent element, to the exponential growth. The law of Makeham (3), which improves the Gompertz law by adding an redundant parameter to this model to take into account the force of accidental death, assumed to be a constant independent of age. Several models which are known under different names but which are naturally the same, because in all of them the force of mortality is a logistic function of age. These include the models introduced by Perks (4), by Beard in 1959 (5), by Vaupel et al. (6) and by Le Bras (7). We shall describe these models inclusively as the logistic model. In this study, we divide the mortality pattern by two age groups 20-60 and 60-100. These age groups are chosen, somewhat arbitrarily to separate working ages and retirement age. In addition, all the mortality models can't apply to the whole age range. In fact, Gompertz indicated (8) that there are 4 distinct periods in the life span between which different laws of mortality refer birth to 12 months, 12 months to 20 years, 20 to 60 years and 60 to 100 years.

This manuscript proposes a modification of the most frequently used six mortality models, which are Gompertz, Logistic, Makeham and Beard. In this paper, we have considered five districts of high population growth rate and five districts of low population growth rate in Assam. The high population growth rate districts of Assam are Darang, Dhubri, Marigaon, Goalpara and Nagaon and the low population growth rate districts: Golaghat, Jorhat, Dibrugarh, Sivsagar and Tinsukia. The prime focus of this manuscript is to find a appropriate mortality model for the districts of Assam for both female and male.

1.1 Mathematical Models

Following models are chosen for this study after an extensive study of a good number of literatures of this field.

Gompertz model. Gompertz [2] has first developed a parametric model to study about mortality. Gompertz modeled the elderly factor of mortality with two parameters a positive scale parameter a, which varies with level of mortality, and a positive shape parameter, which measures the rate of increase in mortality with age. The force of mortality given by Gompertz is

$$\mu_x = ae^{bx}. \tag{1}$$

Makeham model. A modification to the Gompertz model was proposed by Makeham [3] by adding a constant term. The newly added parameter c represents mortality due to causalities or sexually transmitted diseases etc. His model can be refer as

$$\mu_x = c + ae^{bx}. \tag{2}$$

Logistic model. The logistic was first established by Perks [4]. This model is known with different names. In this study, we have considered the logistic function as:

$$\mu_x = c + \frac{ae^{bx}}{1 + de^{bx}}. \tag{3}$$

It can be also observed that Makeham model is a particular case of the logistic model when $d = 0$.

Beard model. The Beard model with three parameters obtained by assuming that the parameter $c = 0$ in (3), which is

$$\mu_x = \frac{ae^{bx}}{1 + de^{bx}}. \tag{4}$$

2. Materials and Methods

2.1 Source of Data

The original data of abridged life tables of Assam are taken from sample registration system (SRS). Then we construct complete life table of the required 5 high population and 5 low population districts of Assam by using the methodology used by Saikia and Borah [9].

The force of mortality rate can be estimated by using the following approximation from empirical data at age x

$$\mu_x = \frac{1}{2}(q_{x-1} + q_x) \tag{5}$$

where q_x is referred as the probability of death at age x before reaching the age $x + 1$ [10].

The four mortality models, Logistic, Gompertz, Beard and Makeham have been considered in this paper. These models can also be written as

$$y_i = f(x_i, B) + \varepsilon_i, \tag{6}$$

$i = 1, 2, \ldots, n,$ where x is independent variable, y is the dependent variables, B is the vector of parameters to be estimated $(\beta_1, \beta_2, \ldots, \beta_p)$, ε_i is a random error term, p is defining the no of parameters, and n denotes the number of observations [10]. The estimators of the parameters can found by minimizing the sum of squares residual (SS_{Res}) function given bellow

$$(SS_{\text{Res}}) = \sum_{i=1}^{n} [y_i - f(x_i, B)]^2 \tag{7}$$

under the hypothesis that the error ε_i are normal and it is independent with mean zero and common variance σ^2. From the equations (7) we get p normal equations that must be solved for $\hat{B}$. These normal equations have the following form

$$\sum_{i=1}^{n} \{y_i - f(x_i, B)\}\left[\frac{\partial f(x_i, B)}{\partial \beta_j}\right] = 0 \tag{8}$$

for $j = 1, 2, \ldots, p.$ The Levenberg-Marquardt method is used to fit the models. The parameters are estimated with the help of MATLAB version 7.11.0.

2.2 Selection Criteria of Best Fitted Mortality Model

To select the best fit model, we are using the following selection criteria.

Sum of Squares Error. The summed square error of residuals (SSE) is given by

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

where $\hat{y}_i$ is representing the predicted values of the observe value y_i for $i = 1, 2, \ldots, n$. A value of SSE closer to 0 refers a fit more useful for prediction.

The Root Mean Square Error. The Root Mean Square Error (RMSE) of a model is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}, \tag{10}$$

A RMSE value closer to 0 refers a fit that is more useful for prediction.

Coefficient of determination. The coefficient of determination (R^2) is given by,

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \tag{11}$$

where $\bar{y}$ indicates the mean of the observed variables. If the value of R^2 is above 0.9 then it is considered as efficient, in general.

3. Results and Discussion

Four mortality models have been fitted for both female and male to 10 districts of Assam. Here we consider five high population growth rate districts and five low population growth rate districts. The high population growth rate districts considered here are Darang, Dhubri, Goalpara, Marigaon and Nagaon and low population growth rate districts are Golaghat, Jorhat, Sibsagar, Dibrugarh, and Tinsukia. We summarized the results as below based on the selection criteria given above.

The estimation of SSE, R^2, RMSE are evaluated for fitting of each model for the age groups 20-60 and 60-100 for both male and female and are presented from Table 1-Table 4. The Table 1 and Table 2 represent the values of SSE, R^2, RMSE for male and female respectively for the age group 20-60. Table 1 and Table 2 represent the same for the age group 60-100. It can be observed from Table 1 to Table 4 that for all cases the values of R-squared are significant. From Table 1 and Table 3, it is watched that estimations of the SSE and RMSE are minimum for Logistic model when contrasted with the others. It is likewise observed from Table 2 and Table 4 that Makeham show gives preferred outcomes over the others. For the age group 20-60, logistic model is the first preferred model for both high and low population growth rate districts for male. The fit is good in individual districts, the deviations small. On the other hand, for female population the three parameter Makeham model is a preferable model. Beard model also gives a better result for low population growth rate districts Dibrugarh, Golaghat, Jorhat, Tinsukia, and Sibsagar. On the other hand, for the female population, Makeham model is appropriate for high population growth rate districts as well as low population growth rate districts.

It can also be observed that for the other age group 60-100, the logistic model is more suitable model for every one of the districts for male population whereas Makeham model is the preferable model for the female population.

Table 1. Calculated SSE, R^2, RMSE for the age group 20-60 (Male).

Districts	Gompertz			Makeham			Logistic			Beard		
	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE
Darang	.00027	.9253	.00216	.00027	.9254	.00217	.00025	.9255	.00119	.00027	.9253	.00217
Dhubri	.00031	.9213	.00232	.00031	.9221	.00232	.00029	.9236	.00132	.00031	.9213	.00234
Dibrugarh	.00035	.8849	.00244	.00035	.8868	.00246	.00017	.9446	.00174	.00023	.9249	.00201
Goalpara	.00027	.9254	.00216	.00027	.9255	.00218	.00025	.9256	.00120	.00027	.9254	.00218
Marigaon	.00027	.9210	.00214	.00027	.9210	.00216	.00024	.9233	.00114	.00026	.9216	.00215
Nagaon	.00028	.9123	.00221	.00028	.9124	.00222	.00022	.9296	.00201	.00027	.9157	.00218
Golaghat	.00029	.9069	.00226	.00039	.7412	.00322	.00021	.9335	.00194	.00027	.9137	.00219
Jorhat	.00034	.8882	.00242	.00034	.8882	.00244	.00015	.9513	.00163	.00023	.9231	.00203
Sibsagar	.00031	.9005	.00232	.00038	.8514	.00366	.00017	.9455	.00174	.00026	.9153	.00215
Tinsukia	.00031	.9015	.00231	.00031	.9015	.00233	.00017	.9447	.00176	.00027	.9148	.00216

Table 2. Calculated SSE, R^2, RMSE for the age group 20-60 (Female).

Districts	Gompertz			Makeham			Logistic			Beard		
	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE
Darang	.00034	.8349	.00241	.00013	.9855	.00203	.00033	.8847	.00209	.00029	.8533	.00227
Dhubri	.00034	.8289	.00241	.00013	.9799	.00201	.00024	.8521	.00205	.00029	.8504	.00227
Dibrugarh	.00038	.7082	.00256	.00018	.9814	.00220	.00029	.7526	.00224	.00032	.7516	.00238
Goalpara	.00035	.7896	.00245	.00012	.9995	.00201	.00025	.8494	.00211	.00030	.8184	.00229
Marigaon	.00036	.7590	.00249	.00016	.9549	.00210	.00026	.8249	.00216	.00031	.7934	.00232
Nagaon	.00037	.7383	.00251	.00017	.9907	.00213	.00037	.8175	.00219	.00031	.7764	.00234
Golaghat	.00036	.7672	.00248	.00016	.9817	.00211	.00036	.8317	.00214	.00031	.8001	.00231
Jorhat	.00036	.7335	.00250	.00018	.9922	.00212	.00029	.7922	.00225	.00031	.7722	.00233
Sibsagar	.00037	.7383	.00251	.00015	.9077	.00212	.00027	.7079	.00219	.00031	.7764	.00234
Tinsukia	.00037	.7414	.00251	.00017	.9404	.00213	.00037	.8105	.00219	.00031	.7790	.00234

Table 3. Calculated SSE, R^2, RMSE for the age group 60-100 (Female).

Districts	Gompertz			Makeham			Logistic			Beard		
	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE
Darang	.00126	.9961	.00358	.00062	.9981	.00253	.00055	.9983	.00239	.00071	.9878	.00271
Dhubri	.00553	.9670	.00752	.00227	.9864	.00484	.00222	.9867	.00381	.00248	.9752	.00505
Dibrugarh	.00775	.9950	.00890	.00746	.9952	.00877	.00225	.9979	.00382	.00325	.9779	.00579
Goalpara	.00141	.9954	.00379	.00069	.9978	.00266	.00052	.9980	.00254	.00081	.9774	.00289
Marigaon	.00105	.9982	.00327	.00084	.9986	.00295	.00030	.9993	.00203	.00043	.9893	.00210
Nagaon	.00238	.9972	.00492	.00022	.9974	.00477	.00102	.9988	.00227	.00103	.9918	.00327
Golaghat	.00337	.9967	.00587	.00319	.9969	.00573	.00148	.9985	.00293	.00149	.9945	.00392
Jorhat	.00745	.9951	.00872	.00716	.9953	.00859	.00314	.9979	.00272	.00314	.9919	.00569
Sibsagar	.00467	.9961	.00690	.00446	.9993	.00678	.00204	.9983	.00361	.00205	.9913	.00459
Tinsukia	.00446	.9962	.00675	.00426	.9983	.00662	.00195	.9983	.00251	.00199	.9953	.00449

Table 4. Calculated SSE, R^2, RMSE for the age group 60-100 (Female).

Districts	Gompertz			Makeham			Logistic			Beard		
	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE	SSE	R^2	RMSE
Darang	.07156	.9291	.02702	.05076	.9498	.02103	.07086	.8298	.02717	.07133	.9293	.02712
Dhubri	.07055	.9277	.02683	.05975	.9484	.02283	.06984	.9244	.02697	.07033	.9279	.02693
Dibrugarh	.08192	.8361	.03487	.05333	.9384	.03080	.07458	.8121	.04121	.1185	.837	.03495
Goalpara	.07557	.9101	.02777	.05463	.9415	.02176	.07473	.9011	.0279	.07534	.9102	.02787
Marigaon	.08758	.8874	.02989	.05643	.9488	.02187	.08651	.8888	.03002	.08728	.8878	.03001
Nagaon	.09871	.8687	.03174	.05745	.9204	.02917	.09742	.8704	.03186	.09832	.8692	.03184
Golaghat	.08382	.8941	.02925	.05284	.9154	.02122	.08282	.8954	.02937	.08355	.8945	.02935
Jorhat	.09965	.8480	.03372	.05912	.9501	.02410	.09899	.8501	.03383	.1109	.8488	.03381
Sibsagar	.09871	.8687	.03174	.05745	.9704	.02171	.09742	.8704	.03186	.09832	.8692	.03184
Tinsukia	.09689	.8716	.03144	.05567	.9732	.02144	.09564	.8733	.03156	.09652	.8721	.03154

3. Conclusion

In this manuscript, four mortality models namely Makeham, Logistic, Gompertz, and Beard have been fitted for 10 different districts of Assam by using Levenberg-Marquardt iteration method. The investigation in this study is somewhat lengthy, but the results can be described concisely. All the models are all far closer to the observed values, and at first sight, it is not that easy to choose between them. No single model is always best. For both the age groups, the four parameter logistic model is the most successful of the four in describing the mortality in high and low population growth rate districts for male population. From a theoretical point of view, the logistic model likewise has some imperative points of interest. It is more broad than alternate models and will work in circumstances where they won't. It additionally has some informative support, there are hypotheses regarding why it works and about conditions in which it may come up short. On the other hand, in case of female mortality the Makeham model gives better results than the others. The other models also give satisfactory results. Based on our results it is noted that for both the age groups and for high and low population growth rate districts logistic and Makeham model are preferred first for male and female. As per the aftereffects of our estimation, it can be presumed that the logistic mortality model and Makeham mortality model are more reasonable over the other candidate models for describing the mortality pattern of districts of Assam for male and female respectively.

References

[1] A. R. Thatcher, V. Kannisto and J. W. Vaupel, The force of mortality at ages 80 to 120, Oden University Press, Odense, 1998.

[2] B. Gompertz, On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies, Philosophical transactions of the Royal Society of London 115 (1825), 513-583.

[3] W. M. Makeham, On the law of mortality and construction of annuity tables, The Assurance Magazine and Journal of the Institute of Actuaries 8(06) (1860), 301-310.

[4] W. Perks, On some experiments in the graduation of mortality statistics, Journal of the Institute of Actuaries 63(01) (1997), 12-57.

[5] R. E. Beard, Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes, in W. Brass (ed.), Biological Aspects of Demography, Taylor and Francis Ltd,. London, 1971.

[6] J. W. Vaupel, K. G. Manton and E. Stallard, The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography 16(3) (1979), 439-454.

[7] H. Le Bras, Lois de mortalité et age limite, Population (French edition) Institut National d'Etudes Démographiques 31(3) (1976), 655-692.

[8] S. J. Olshansky and B. A. Carnes, Ever since gompertz, Demography 34(1) (1997), 1-15.

[9] P. Saikia and M. Borah, Construction of complete life tables for 23 districts of assam for both male and female, International Journal of Science and Research 5(12) (2016), 1-11.

[10] R. C. Elandt-Johnson and N. L. Johnson, Survival models and data analysis, John Wiley and Sons 74 (1999).

[11] M. Borah and D. J. Mahanta, Rapid parameter estimation of three parameter nonlinear growth models, International Journal of Mathematical Archive 4(2) (2013), 274-282.