International Journal of Statistics and Applied Mathematics 2023; 8(5): 130-134

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; 8(5): 130-134
© 2023 Stats & Maths
https://www.mathsjournal.com
Received: 05-06-2023
Accepted: 06-07-2023

Onchere Walter
Department of Mathematics and Actuarial Science, Kisii University, Kenya

Maina Calvin
Department of Mathematics and Actuarial Science, Kisii University, Kenya

Monari Fred
Department of Mathematics and Actuarial Science, Kisii University, Kenya

Corresponding Author:
Onchere Walter
Department of Mathematics and Actuarial Science, Kisii University, Kenya

Mortality graduation using a rating up approach

Onchere Walter, Maina Calvin and Monari Fred

DOI: https://dx.doi.org/10.22271/maths.2023.v8.i5b.1302

Abstract
Modeling the risk profile of a heterogeneous life from unreported risks is complex. Insurance firms routinely disregard unreported risk factors perhaps because of difficulties in modeling. The scientific interest of the study is to account for unreported heterogeneity factors using a compound process. For comparison reasons the gamma (mostly applied in literature) and non-central gamma compound process are suggested with the generalized exponential and generalized Weibull baselines to account for unreported heterogeneity. In this study, maximum likelihood estimation is used to calibrate the base force of mortality distributions using a large Kenyan insurer term insurance data. Subsequently, the performance of the candidate models described above is compared following the criteria information values. The findings show that the non-central gamma generalized Weibull is significant to model the insurers' claims events.

Keywords: Rating up, unreported heterogeneity, term insurance, non-central gamma, generalized Weibull, generalized exponential

1. Introduction

Rating up theory was first proposed by Beard (1959) [2] to account for age-pattern of mortality, where the term "longevity factor" was used instead of rating up. Homogeneity is presumed with respect of reported risk factors in standard models. The implication is that study subjects are pooled in the same risk profile at a given age. However, statistical evidence suggest a different model as indicated by several researchers, such as Su & Sherris (2012) [12]; Gatzert *et al.*, (2012) [7]; Onchere (2013) [10]; Fong (2015) [6]; Olivieri & Pitacco (2016) [9] and Pitacco (2018) [11] with references therein.

To adequately price and allocate reserves that represent the insurance contracts all relevant factors affecting mortality needs to be considered. Neglecting unreported heterogeneity could result in biased insurance products pricing and reserves allocation. This research outlines a heterogeneous model that will improve the underwriting process to ensure fair pricing and reserving of life products consistent with the insured risk.

The mortality model selected in valuation determines how term insurance and annuity products are priced (Batty *et al.*, (2010) [1]; Gildas *et al.*, (2018) [8]. Life insurance-based rating up models measure population-level heterogeneity caused by unreported risks. On the other hand, heterogeneity caused by reported risk factors is determined during underwriting before issuing a policy to guarantee an optimal assignment of premium equivalent to insured risk for each contract. Excluding relevant factors or relying solely on age and sex may contribute to incorrectly priced assurance products. Term insurance contract is a policy in which a certain payment, say, KES *A* per annum, is provided if the insured passes on within a stated period, say *m* years. This benefit is determined from the EPV as

$$A \int_0^m v^t S_x(t) h(x+t) dt \tag{1}$$

Where: x is the insured's age, v^t represents the present value interest factor, $S_x(t)$represents the survivor function and $h(x+t)$ the intensity rate. The type of mortality model applied to $h(x+t)$ influences how the policy is priced.

2. Methodology

For this study, we considered 732 term insurance contracts between calendar years 2010-2015 from a large Kenyan insurance company. Demographic information of policyholders includes; age, date of signing the contract and mortality date. This data-set will be used to compute the times-to-death and crude intensity rates experienced by the insured in the 22-64 age group. Our aims of study are firstly to show that when the gamma is applied as a rating up distribution the intensity rates are overestimated at all ages compared to the non-central gamma (NCG). Secondly, is to show the relevance of the NCG rating up mixture to graduate the insurance firm's crude intensity rates. Here, force mortality is presumed to be a pair-wise constant, assuming a fixed value across all ages consistent with Brouhns *et al.*, (2002) [3] and Dodd *et al.*, (2018) [5] assumption. We further assumed our rating up model has no observed covariates as only survival data was obtained for this analysis also in life insurance due to the underwriting procedures groups should be homogeneous with respect to observed covariates. Finally, and as given in the actual dataset we presumed that policyholders purchase term assurance policy at the age of 22-64 years.

2.1 The proposed model

The generalized Weibull (GW) density function is represented as

$$f(t) = b(1 - exp(-\lambda t^{\rho}))^{b-1} \cdot \lambda \rho t^{\rho-1} \cdot exp(-\lambda t^{\rho});\ t > 0, \rho, b, \lambda > 0.$$

We can express the survivor, baseline hazard and cumulative hazard rates as

$$S(t) = \int_0^{\infty} b(1 - exp(-\lambda x^{\rho}))^{b-1} \cdot \lambda \rho x^{\rho-1} \cdot exp(-\lambda x^{\rho}) dx.$$

$$Let\ y = exp(-\lambda x^{\rho})\ and\ \frac{dy}{dx} = -\lambda \rho x^{\rho-1} exp(-\lambda x^{\rho}).$$

$$\therefore S(t) = \int_{exp(-\lambda x^{\rho})}^{0} -b(1 - y)^{b-1}\, dy = 1 - [1 - exp(-\lambda t)]^{b}.$$

$$h_0(t) = \frac{b(1 - exp(-\lambda t^{\rho}))^{b-1} \cdot \lambda \rho t^{\rho-1} \cdot exp(-\lambda t^{\rho})}{1 - [1 - \exp(-\lambda t^{\rho})]^{b}}. \quad (2)$$

$$H_0(t) = -log\, S(t) = -log(1 - [1 - exp(-\lambda t^{\rho})]^{b}). \quad (3)$$

The generalized exponential (GE) function is derived from the GW distribution when $\rho = 1$. Putting $\rho = 1$ in Equations (2, 3) leads to

$$h_0(t) = \frac{b(1 - exp(-\lambda t))^{b-1} \cdot \lambda \cdot exp(-\lambda t)}{1 - [1 - exp(-\lambda t)]^{b}}$$

$$H_0(t) = -\log(1 - [1 - \exp(-\lambda t)]^{b}$$

Firstly, using GE distribution as the baseline intensity for gamma rating up, we can generate an intensity rate for gamma-GE rating up mixture as follows

$$h(t) = \frac{b(1 - \exp(-\lambda t))^{b-1} \cdot \lambda \cdot \exp(-\lambda t)}{1 - [1 - \exp(-\lambda t)]^{b}}.(1 - \sigma^2 \log(1 - [1 - \exp(-\lambda t)]^{b})^{-1} \quad (4)$$

The NCG is further adapted as a rating up distribution resulting in the NCG-GE intensity rate given by

$$h(t) = \frac{b(1 - \exp(-\lambda t))^{b-1} \cdot \lambda \cdot \exp(-\lambda t)}{1 - [1 - \exp(-\lambda t)]^{b}}.(1 - 0.5\sigma^2 \log(1 - [1 - \exp(-\lambda t)]^{b})^{-2} \quad (5)$$

Where, $b < 1, b > 1, and\ b = 1$ represent decreasing, increasing, and constant GE intensity rate, respectively.

Secondly, the GW distribution is proposed as the base force of mortality with the gamma rating up giving the gamma-GW (G-GW) intensity rate expressed as

$$h(t) = \frac{b(1 - \exp(-\lambda t^{\rho}))^{b-1} \cdot \lambda \rho t^{\rho-1} \cdot \exp(-\lambda t^{\rho})}{1 - [1 - \exp(-\lambda t^{\rho})]^{b}}(1 - \sigma^2 \log(1 - [1 - \exp(-\lambda t^{\rho})]^{b}))^{-1} \quad (6)$$

Similarly, the NCG-GW intensity rate is described as

$$h(t) = \frac{b(1 - exp(-\lambda t^{\rho}))^{b-1} \cdot \lambda \rho t^{\rho-1} \cdot exp(-\lambda t^{\rho})}{1 - [1 - exp(-\lambda t^{\rho})]^{b}}(1 - 0.5\sigma^2\, log(1 - [1 - exp(-\lambda t^{\rho})]^{b}))^{-2} \quad (7)$$

where $(\rho \leq 1)$ and $(b\rho \leq 1)$ represent monotonically decreasing function and $(\rho \geq 1)$ and $(b\rho \geq 1)$ represents monotonically increasing function; unimodal if $(\rho < 1)$ and $(b\rho > 1)$ and bath-tub shaped if $(\rho > 1)$ and $(b\rho < 1)$.

3. Data analysis, results and discussions

3.1 Parameter Estimation

Butt & Haberman (2004) [4] apply the rating up-based survival model to insurance. The authors first consider different choices of models and then apply them to two large life insurance mortality datasets. The results indicate a potential range $\sigma^2 \approx (2.916, 14.444)$ in an insured population with $\sigma^2 = 14\%$ for the heterogeneous case. From the findings of the investigation by Butt & Haberman (2004) [4], in this study we consider $\sigma^2 = 14\%$ for the heterogeneous case.

The MLE approach is concerned with obtaining parameter values say, (b, λ, ρ) that maximizes the probability of observing the data D given those parameters, $p(D|\, b, \lambda, \rho)$. The likelihood function gives the probability of the observed sample generated by the model. Generally, maximization of the likelihood function to find the ML estimates is done algebraically, but can be computational intensive. In this study, the MLE algorithm is implemented using stats 4 package in R.

For the Generalized Weibull with PDF

$$f\,(t) = b(1 - exp\,(-\lambda t^{\rho}))^{b-1} \cdot \lambda\rho t^{\rho-1} \cdot exp\,(-\lambda t^{\rho});\ t > 0, \rho, b, \lambda > 0.$$

The log-likelihood function is derived as

$$L = \prod_{i=1}^{n} b(1 - \exp(-\lambda t^{\rho}))^{b-1} \cdot \lambda\rho t^{\rho-1} \cdot \exp(-\lambda t^{\rho})$$

$$logL = \sum_{i=1}^{n} \log(b(1 - \exp(-\lambda t^{\rho}))^{b-1}) + \log(\lambda\rho t^{\rho-1}) - \lambda t^{\rho}$$

For the Generalized Exponential with PDF

$$f\,(t) = b(1 - exp\,(-\lambda t))^{b-1} \cdot \lambda \cdot exp\,(-\lambda t);\ t > 0, b, \lambda > 0.$$

The log-likelihood function is derived as

$$L = \prod_{i=1}^{n} b(1 - exp\,(-\lambda t_i))^{b-1} \cdot \lambda \cdot exp\,(-\lambda t_i)$$

$$logL = \sum_{i=1}^{n} \log(b(1 - \exp(-\lambda t_i))^{b-1}) + \log(\lambda) - \lambda t_i$$

The parameter estimates and criteria values is shown in the table below;

Table 1: Base Force of Mortality Parameter Estimates

Baseline model	Parameter estimates	AIC	BIC
1. Generalized Exponential	λ = 0.14537, b = 286.57	4995.128	5004.32
2. Generalized Weibull	λ = 0.031367, b = 80.3076, ρ = 1.3477	4960.588	4974.376

Discussion

Stats 4 provides the AIC and BIC measures that penalizes both excessive use of parameters and poor data fitting. From the above results, the least AIC and BIC suggests that the GW gives a better fit. The G-GW and NCG-GW rating up models given in Equations (6) and (7) respectively are as shown in Figure 1 where t is the time-to-death, $\sigma^2 = 0.14$, $\lambda = 0.031367$, $b = 80.3076$, $\rho = 1.3477$; $h_0(t) \sim GW\,(0.031367, 80.3076, 1.3477)$.

In Figure 1 graduation is done using the G-GW and NCG-GW rating up model both calibrated on the real term assurance times-to-death data. This is compared with the real term assurance intensity rates. As shown the G- GW overestimates the intensity rate at all ages compared to the NCG-GW model. The NCG-GW is observed to fit well to the actual claims experience hazards. The chi-square test Table 2 and Kolmogorov-Smirnov (KS) hypothesis test Table 3 for overall goodness of fit is significant for NCG-GW. The chi-squared goodness-of-fit test has p-value greater than 0.01, indicating that the model fits well. Similarly, for KS goodness of fit test p-value > 0.01, indicating goodness of fit for the distribution.

Table 2: Chi-squared Goodness-of-fit of NCG-GW to the Crude Intensity Rates

Name	Value
Chi-squared statistic	1722
Degree of freedom	1681
Chi-squared p-value	0.2379

CONTRACTS INFORCE BETWEEN 2010-2020							
Age x	dx	Ex	sx	fx	CRUDE_hx	NCG_GW_hx	G_GW_hx
24	3	729	0.995893	0.004098	0.004115	0.001630	0.001654
25	1	726	0.998624	0.001376	0.001377	0.002210	0.002530
26	5	725	0.993127	0.006849	0.006897	0.003420	0.003748
27	6	720	0.991701	0.008264	0.008333	0.005340	0.005390
28	8	714	0.988858	0.011080	0.011204	0.007220	0.007533
29	8	706	0.988733	0.011204	0.011331	0.010250	0.010252
30	7	698	0.990021	0.009929	0.010029	0.013600	0.013601
31	9	691	0.987060	0.012856	0.013025	0.017625	0.017627
32	15	682	0.978246	0.021516	0.021994	0.022342	0.022343
33	11	667	0.983643	0.016222	0.016492	0.027742	0.027744
34	17	656	0.974418	0.025252	0.025915	0.033792	0.033795
35	15	639	0.976799	0.022930	0.023474	0.040434	0.040439
36	10	624	0.984102	0.015771	0.016026	0.047588	0.047599
37	33	614	0.947673	0.050934	0.053746	0.055161	0.055179
38	34	581	0.943160	0.055194	0.058520	0.063045	0.063077
39	27	547	0.951838	0.046983	0.049360	0.071129	0.071182
40	34	520	0.936707	0.061246	0.065385	0.079300	0.079385
41	24	486	0.951817	0.047003	0.049383	0.087448	0.087578
42	14	462	0.970152	0.029399	0.030303	0.095471	0.095663
43	170	448	0.684228	0.259640	0.379464	0.103276	0.103554
44	99	278	0.700392	0.249420	0.356115	0.110784	0.111174
45	45	179	0.777714	0.195515	0.251397	0.117927	0.118462
46	18	134	0.874303	0.117444	0.134328	0.124652	0.125372
47	8	116	0.933359	0.064370	0.068966	0.130921	0.131867
48	6	108	0.945959	0.052553	0.055556	0.136708	0.137928
49	10	102	0.906613	0.088884	0.098039	0.141997	0.143544
50	14	92	0.858839	0.130693	0.152174	0.146784	0.148715
51	4	78	0.950011	0.048718	0.051282	0.151076	0.153449
52	10	74	0.873598	0.118054	0.135135	0.154884	0.157762
53	11	64	0.842084	0.144733	0.171875	0.158225	0.161673
54	13	53	0.782483	0.191930	0.245283	0.161122	0.165205
55	9	40	0.798516	0.179666	0.225000	0.163600	0.168385
56	9	31	0.748022	0.217168	0.290323	0.165684	0.171239
57	3	22	0.872525	0.118981	0.136364	0.167401	0.173792
58	6	19	0.729213	0.230278	0.315789	0.168778	0.176072
59	1	13	0.925961	0.071228	0.076923	0.169840	0.178102
60	1	12	0.920044	0.076670	0.083333	0.170611	0.179906
61	3	11	0.761300	0.207627	0.272727	0.171113	0.181504
62	1	8	0.882497	0.110312	0.125000	0.171368	0.182915
63	2	7	0.751477	0.214708	0.285714	0.171394	0.184157
64	1	5	0.818731	0.163746	0.200000	0.171209	0.185243
65	4	4	0.367879	0.367879	1.000000	0.170828	0.186188

Fig 1: Construction of Crude Intensity Rates from Real Term Insurance Dataset

Table 3: Goodness of fit using Kolmogorov-Smirnov test

Name	p-value	test statistic
Kolmogorov-Smirnov test	0.06448	0.28571

4. Discussion

Our goal in this paper is to model unreported heterogeneity by applying the gamma and NCG rating up distributions. We applied our model to real term insurance times-to-death data. Based on maximum likelihood, the GW baseline fits better as the AIC & BIC is lower than the GE baseline. As shown in Figure 1 the G-GW model overestimates the intensity rates at all ages compared to the NCG-GW model. According to chi-squared goodness of fit test Table 2 and KS hypothesis test Table 3, the NCG-GW shows better fitness to insurer's claims experience.

5. Conclusion

Life Insurance-Based rating up models measure population-level heterogeneity caused by unreported risks. On the other hand, heterogeneity caused by reported risk factors is determined during underwriting before issuing a policy to guarantee an optimal assignment of premium equivalent to insured risk for each contract. Excluding relevant factors or relying solely on age and sex may contribute to incorrectly priced assurance products. The scientific interest of the study is to account for unreported heterogeneity factors using a compound process. For comparison reasons the gamma (mostly applied in literature) and non-central gamma compound process are suggested with the generalized exponential and generalized Weibull baselines to account for unreported heterogeneity. The conclusion arrived at is that using the gamma as the rating up distribution may lead to inappropriate term assurance valuations resulting in high prices that negatively impacts marketability of term contracts. The gamma rating up index is time invariant and unreported heterogeneity effects remains constant throughout life. The NCG compound process represents time-varying unreported heterogeneity effects and is recommended for better term assurance valuations.

6. Acknowledgment

This paper has benefited from discussions with many people, particularly Prof. Patrick Weke from the University of Nairobi. The research was partially supported by ISP-IPMS-APREA project.

ORCID ID

Walter Onchere https://orcid.org/0000-0002-4492-0023

7. References

1. Batty M, Tripathi A, Kroll A, Wu C, Moore D, Stehno C, *et al*. Predictive Modeling for Life Insurance, Deloitte Consulting LLP; c2010.
2. Beard RE. Note on some mathematical mortality models. The Lifespan of Animals. In: Wolstenholme, G.E.W. and O Conner, M., Eds., Ciba Foundation Colloquium on Ageing, Little, Brown and Company, Boston. 1959;5:302-311. DOI:10.1002/9780470715253.app1
3. Brouhns N, Denuit M, Vermunt KJ. A Poisson log-bilinear regression approach to the construction of projected life tables. Insurance: Mathematics and Economics. 2002;31:373-393.
4. Butt Z, Haberman S. Application of frailty-based mortality models using generalized linear models: Astin Bulletin. 2004;34(1):175-197.
5. Dodd E, Forster J, Bijak J, Smith P. Smoothing mortality data: the English Life Tables; c2010-12. J. R. Statist. Soc. A 0964-1998/18/181000; c2018.
6. Fong HJ. Beyond Age and Sex: Enhancing Annuity Pricing: The Geneva Risk and Insurance Review. 2015;40:133-170.
7. Gatzert N, Gudrun SH, Hato S. Optimal Risk Classification with an Application to Substandard Annuities, North American Actuarial Journal. 2012;16(4):462-486.
8. Gildas R, Francois D, Enkelejd H, Youssouf T. On age difference in joint lifetime modeling with life insurance annuity applications: Annals of Actuarial Science. 2018;12(2):350-371.
9. Olivieri A, Pitacco E. Frailty for Life Annuity Portfolios: Risks. 2016;4:39.
10. Onchere WO. Frailty models' applications in pension schemes. Masters Thesis, University of Nairobi, 2013.
11. Pitacco E. Heterogeneity in mortality. A survey with an actuarial focus: Working research 2018/19. Australian Research Council Centre of Excellence in Population Ageing Research, University of New South Wales, Sydney; c2018.
12. Su S, Sherris M. Heterogeneity of Australian population mortality and implications for a viable life annuity market. Insurance: Mathematics and Economics. 2012;51(2):322-332.