**Using Mixed Distribution for Gamma and Exponential to Estimate of Survival Function (Brain Stroke)**

**Abstract**

In this paper, we are used Bayesian on survival function estimator based on the mixed distribution of exponential distribution as primary distribution and Gama distribution as a function of probability of data, and the data was collected from Rizgari Hospital Erbil for stroke brain patients between 2015 and 2016. On the other hand, we compared with the traditional method that assumed exponential and Gamma distributions based on the Goodness of Fit tests depended on the EasyFit Program, as well as using MATLAB and SPSS statistical programs. We concluded that the mixed and proposed combination of survival function for brain stroke was expectancy, appropriate and efficient.

**Keywords**

Exponential, Gamma and mixed Distribution, Survival function, Bayesian estimation, and Maximum likelihood Estimation.

**1. Introduction**

 Brain stroke is a most important healthiness burden in Iraq specific in Kurdistan as well as worldwide. Indeed stroke happen when blood provide to the brain is not working, the hungry brain of nutrients and oxygen. Brain cells start in on to die if blood run is stopped for more than a few seconds. While longer the supply is periodic, the more possible like there is to be permanent and incapacitating brain injure. A stroke is a health check urgent situation that requires urgent treatment (Soman et al.,, 2016). On the other hand the strokes have an effect on the brain depends on which element of the brain suffers harm. The intelligence stem controls your breathing, heartbeat, and blood pressure; controls your speech, swallowing, hearing, and eye movements. The pulses sent by other parts of the brain travel through the brainstem on their way to the various parts of the body. We depend on brain stem function to survive. Stem brain stem threatens the body's vital functions, making it a life-threatening condition. (Pietrangelo, 2016). In the United States, stroke is the third leading cause of death and an effective factor for severe disability, and die more than 140,000 people each (American Heart Association, 2009) year. It is expected that the prediction of stroke and early treatment will substantially increase the risk of stroke.

 Data analysis and many medical studies have been carried out to classify effective career predictors. Framingham in Wolf et al. (1991) published a set of risk factors that may be related to stroke such as age, including age, systolic blood pressure, use of antihypertensive therapy, diabetes mellitus, smoking , Previous cardiovascular disease, an a trial test, and left ventricular hypertrophy by electrocardiogram, creatinine level, walking time 15 feet, and others (McGinn, et al., 2008).

**2. Methodology**

Let *t* denote a continuous non-negative random variable representing survival time, the probability density function for two parameter Gamma distributions has the following Gupta, R. D. and Kundu, D. (1999).

**2-1 Survival distribution**

Beginning with the simply definition for survival distribution on the possibility that the event there is no case of interest after the time *t*. If *T* specified time until death, so the survival function *S(t)* denoted probability of surviving further than time *t*. We start with probability density function which is below.

 

And the corresponding cumulative distribution function (CDF), survival and hazard functions as follows;







Where *α,β* are the shape and location parameters respectively, on the other hand the probability density, cumulative distribution, survival and hazard functions for the exponential distribution are as following Raqab and Ahsanllah (2001) and Raqab (2002):









**2-2 Bayesian Estimation ( The Prior and Posterior (Expectations) Distribution)**

A normal select for the prior of parameters and  could be assume that the three parameters are independent and that, ~G(a0,b0), ~G(a1,b1) where G(a,b) denotes the gamma distribution with mean equal to and variance , while the parameter is truncated positive exponential distribution Kotz, et.al(2003). All parameter above are chosen to reflect prior knowledge. The life time of a system with one component has the probability density function (PDF), for exponential and gamma distribution where and are known then:

We can assume by using maximum likelihood estimation as follows;

 and 

 



Where *t* is time of life patients for 2015 in equation (9), but *t* in formula (10) we can depend on the data 2016. However, to find the posterior distribution for above formula we have;



Dependent on the equation (11) we could work some mathematical on it, until get the posterior distribution;



Rearranging and some mathematical work on the above equation, we get equation (13), below:



Where ~Gamma (

Furthermore, the survival function for the posterior distribution, depend on the probability density function for exponential is as follows:



Integrating the survival function in equation (15), we get;



To find the Bayesian estimator survival function depends on the real survival function we have;

Again some mathematical work on above equation (17), then we gets equation as follow:



Thus, depend on the above equation, we can assume that cumulative density, probability density functions and hazard function for Bayesian estimator as follow;

 

After the derivatives and some mathematical work from equation (18) we can find the probability density function as follow:



So, the hazard function is;



Where  and *t* are known depend on the data 2015, and also *x* we assume on the data 2016.

**3. Techniques and Patients Data**

The data covers the 58 patients, 24 female and 34 male from January 2015 and 2016 to December 2015 and 2016 and the data was collected from Rzgary Teaching Hospital in Hawler city-Iraq. After we are organized the data and also we analyzed by using the Easyfit 5.6 standard Program.

**3.1 Statistical Analyses**

This paper considers apply on the secondary data based on laboratory investigations. These data are supplied by Rizgary Teaching Hospital, Erbil, Iraq for the year 2015 and 2016. The data collected from existing databases in terms of the survival time of death which are shows in (Appendix A & B). The data includes the factors affecting the brain stroke which are age, survival time, time inter and exit patient on hospital. The data translated into codes using a particularly designed coding sheet, after that converted to computerized database. The specific statistical advice was required and EasyFit 5.6 standard program used. First all after the estimate parameter for the gamma and exponential then mixed together, and find baysen estimation which are prior and posterior expectations, secondly compare between classically and new model.

**3.2 Results**

To determine the survival function for the gamma, exponential and mixed distributions on the brain stroke patients in the Rizgary Teaching Hospital. This is performed by finding the survival function curve for the selected time of death, beginning with first case depend on the 29 observation, to estimate scale parameter for the exponential distribution which is equal to (0.046), and the value of Kolmogrov simirnov (0.241) which is grater then 0.05, while the value of the mean, variance and standard division are equal to (21.759, 288.260 and 16.978) respectively. To estimate the parameters location and scale for the gamma distribution are equal to =1.9278, =10.607, and the value of kolmogrov simirnov is (0.0906) again grater then (0.05), which is significant value and the value of mean is (20.448), variance (216.900) and standard division (14.727) . Know we are going to the second case which is mixed exponential and gamma distributions, the observation for these state is 58. Moreover, the value of parameters andare (1.7916, 11.779) respectively, and the kolmogrov siminov is (0.1108), also it is significant value and grater then that p. value 0.05. Moreover, the value of the mean, variance and standard division are equal to (21.103, 248.590 and 15.767) respectively. The comparison of the classically result and new model are the same consequences.

As shown in Figure from 1 to 12.



**Figure 1. Probability density function curves for the exponential distribution (x)**



**Figure 2. Cumulative distribution function curves for the exponential distribution (x)**



**Figure 3. Survival function curves for the exponential distribution (x)**



**Figure 4. Hazard function curves for the exponential distribution (x)**



**Figure 5. Probability density function curves for the gamma distribution (t)**



**Figure 6. Cumulative distribution function curves for the gamma distribution (t)**



**Figure 7. Survival function curves for the gamma distribution (t)**



**Figure 8. Hazard function curves for the gamma distribution (t)**



**Figure 9. Probability density function curves for the mixed distribution (x,t)**



**Figure 10. Cumulative distribution function curves for the mixed distribution (x,t)**



**Figure 11. Survival function curves for the mixed distribution (x,t)**



**Figure 12. Hazard function curves for the mixed distribution (x,t)**

**4. Discussion and Conclusions**

They have been used in a large-scale analysis of survival in public health research, physicians

and epidemiological research in health, social and behavioral economics studies. The relationship between a particular event and "potential risk factors", Or estimating the causal or predictive model is often of interested. The survival analysis is most appropriate for longitudinal studies where the event of interest, when it happens, however, it is frequently used with display data. We construed on the Bayesian estimator with survival function, the first case to analyze survival time data. Using maximum likelihood method to estimate exponential parameter, gamma distributions and using bayesian method to estimate the parameters of both combined distributions. Although, the survival function for the gamma was very close to zero then that exponential. However, the survival function for the mixed distribution it’s also approach to zero, which mean that they model was very well. On the other hand, depended on the value of Kolmocrove simirnov for Exponential distribution is (0.241), for gamma is equal to (0.107) and for the mixed is (0.110), which were statistically significant. Again, depending on our case study, incidence of brain stroke for the year 2015 and 2016 in the Rizgary Teaching Hospital, Erbil - Kurdistan Region of Iraq. The survival for the patients from 2016 is less than in 2015, because for this year many thinks was changed for instance the life of challenging, economics, politics and etc.

**الملخص**

 تمَ في هذا البحث إستخدام مقدر بيز لدالة الحياة إعتماداً على التوزيع المختلط الحاصل من التوزيع الأسي كتوزيع أولي وتوزيع كاما كدالة ترجيح لبيانات مؤخوذة من مستشفى رزكاري لمرضى الجلطة الدماغية لعامي (2015-2016) ومقارنتها مع الأسلوب التقليدي الذي يفترض التوزيع الأسي وكاما بناءً على إختبارات جودة المطابقة (Goodness of fit) من خلال برنامج (EasyFit) فضلاً عن إستخدام برنامج لغة ماتلاب والبرنامج الإحصائي الجاهز (SPSS) وتوصلت الدراسة إلى ملائمة وكفاءة التوزيع المختلط والمقترح لتقدير دالة الحياة لمرضى الجلطة الدماغية و بالاعتماد على قيمة كولموكروف سميرموف لتوزيع الاسي يساوي (0.241) و توزي كاما (0.107) وكذلك لتوزيع المختلط (0.110) لذا فانه نتائج معنوية احصائيا.

**References**

Al-Shimmery, E.K., Amein, S. H., & Al-Tawil, N.G. (2010). Prevalence of silent stroke in Kurdistan, Iraq. Neurosciences, *15*(3), 167-171.

American Heart Association (2009). Heart Disease and Stroke Statistics 2009 Update. American Heart Association, Dallas, Texas.

Bradburn, M. J., Clark, T. G., Love, S. B., & Altman, D. G. (2003). Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *British Journal of Cancer*, *89*(3), 431–436. http://doi.org/10.1038/sj.bjc.6601119

Crowder, M. (2012). *Multivariate Survival Analysis and Competing Risks*. London: Taylor & Francis Group, LLC.

Gupta, R. D. and Kundu, D. (1999). Generalized Exponential Distribution, Austral. N. Z. Statist. 41(2), 173-188.

Health Grove (2017). Global Health Statistics. Graphiq Inc. from http://global-health.healthgrove.com/ March 31.

Ikeda, K. Kumads, H. Saitoh, S. Arase, Y. & Chayama, K. (2001). Effect of repeated transcatheter arterial embolization on the survival time in patients with hepatocellular carcinoma. *Cancer, 68*(10), 2150-2154.

Kotz, S., Lumelskii, Y. and Pensky, M. (2003). The Stress-Strength Model and its Gener- alizations, World Scienti¯c, New York.

Lumley, T. Kronmal, R.A. Cushman, M. Manolio, T.A. & Goldstein, S. (2002). A stroke prediction score in the elderly: Validation and web-based application and S. Goldstein.  *Journal of Clinical Epidemiology, 55*(2), 129-136.

Mark Stevenson. (2009). An Introduction to Survival Analysis. *EpiCentre: IVABS*, Massey University.

Pietrangelo, A. (2016). Medically Reviewed by [University of Illinois-Chicago, College of Medicine](http://www.healthline.com/health/medical-board) .

Raqab, M. Z. and Ahsanullah, M. (2001). Estimation of Location and Scale Parameters of Generalized Exponential Distribution Based on Order Statistics, Journal of Statistical Computation and Simulation 69(2), 109-124.

Soman, S., Prasad, G., Hitchner, E., Massaband, P., Moseley, M. E., Zhou, W. & Rosen, A. C. (2016), Brain structural connectivity distinguishes patients at risk for cognitive decline after carotid interventions. Hum. *Brain Mapp., 37,* pp. 2185–2194. doi:10.1002/hbm.23166

Spruance, S.L., Reid, J.E., Grace, M., & Samore, M., (2004). Hazard ratio in clinical trials. *Antimicrob Agents Chemother*. *48,* pp. 2787–92.

Wolf, P.A., D'Agostino, R.B., Belanger, J. A. and Kannel, W.B. (1991). Probability of stroke: a risk profile from the Framingham study. *Stroke, 22*, 312-318.

**Appendix A: General table for the survival function for patient’s (Censored and Death) years 2015 &2016**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Time2015&16 | Survival Time | Status | No. | Time2015&16 | Survival Time | Status | No | Time2015&16 | Survival Time | Status |
| 1 | 1 | 8 | 1 | 34 | 2 | 30 | 0 | 67 | 2 | 11 | 0 |
| 2 | 1 | 29 | 1 | 35 | 2 | 13 | 1 | 68 | 2 | 42 | 1 |
| 3 | 1 | 33 | 1 | 36 | 2 | 8 | 0 | 69 | 2 | 14 | 1 |
| 4 | 1 | 17 | 1 | 37 | 2 | 21 | 1 | 70 | 2 | 12 | 1 |
| 5 | 1 | 61 | 1 | 38 | 2 | 3 | 0 | 71 | 2 | 61 | 1 |
| 6 | 1 | 0 | 1 | 39 | 2 | 17 | 1 | 72 | 2 | 12 | 0 |
| 7 | 1 | 30 | 1 | 40 | 2 | 17 | 0 | 73 | 2 | 24 | 0 |
| 8 | 1 | 35 | 1 | 41 | 2 | 31 | 0 | 74 | 2 | 26 | 0 |
| 9 | 1 | 31 | 1 | 42 | 2 | 62 | 0 | 75 | 2 | 17 | 1 |
| 10 | 1 | 61 | 1 | 43 | 2 | 10 | 1 | 76 | 2 | 8 | 1 |
| 11 | 1 | 31 | 0 | 44 | 2 | 12 | 1 | 77 | 2 | 15 | 1 |
| 12 | 1 | 3 | 0 | 45 | 2 | 11 | 1 | 78 | 2 | 6 | 1 |
| 13 | 1 | 32 | 1 | 46 | 2 | 16 | 1 | 79 | 2 | 10 | 1 |
| 14 | 1 | 2 | 0 | 47 | 2 | 20 | 1 | 80 | 2 | 13 | 1 |
| 15 | 1 | 5 | 1 | 48 | 2 | 14 | 0 | 81 | 2 | 11 | 0 |
| 16 | 1 | 31 | 0 | 49 | 2 | 16 | 0 | 82 | 2 | 44 | 1 |
| 17 | 1 | 92 | 0 | 50 | 2 | 11 | 1 | 83 | 2 | 11 | 0 |
| 18 | 1 | 31 | 0 | 51 | 2 | 13 | 1 | 84 | 2 | 13 | 1 |
| 19 | 1 | 31 | 0 | 52 | 2 | 4 | 1 | 85 | 2 | 6 | 1 |
| 20 | 1 | 32 | 0 | 53 | 2 | 8 | 1 | 86 | 2 | 9 | 1 |
| 21 | 1 | 33 | 0 | 54 | 2 | 10 | 1 | 87 | 2 | 36 | 1 |
| 22 | 1 | 25 | 0 | 55 | 2 | 14 | 1 | 88 | 2 | 24 | 0 |
| 23 | 1 | 34 | 0 | 56 | 2 | 21 | 1 | 89 | 2 | 28 | 0 |
| 24 | 1 | 30 | 1 | 57 | 2 | 61 | 1 | 90 | 2 | 20 | 1 |
| 25 | 1 | 35 | 0 | 58 | 2 | 61 | 1 | 91 | 2 | 13 | 1 |
| 26 | 1 | 12 | 0 | 59 | 2 | 15 | 1 | 92 | 2 | 18 | 1 |
| 27 | 1 | 23 | 1 | 60 | 2 | 8 | 1 | 93 | 2 | 34 | 1 |
| 28 | 2 | 30 | 0 | 61 | 2 | 5 | 0 | 94 | 2 | 6 | 0 |
| 29 | 2 | 31 | 0 | 62 | 2 | 61 | 0 | 95 | 2 | 9 | 1 |
| 30 | 2 | 15 | 0 | 63 | 2 | 61 | 0 | 96 | 2 | 12 | 1 |
| 31 | 2 | 4 | 0 | 64 | 2 | 12 | 1 | 97 | 2 | 33 | 0 |
| 32 | 2 | 13 | 0 | 65 | 2 | 7 | 0 | 98 | 2 | 25 | 0 |
| 33 | 2 | 18 | 1 | 66 | 2 | 12 | 1 | 99 | 2 | 19 | 0 |
| 100 | 2 | 29 | 1 |

**Appendix B: Specific table for the survival function for patient’s (Censored) years 2015 &2016**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Survival Time(1) | Status | variable t | variable x | (x+(1/B)) | t+x+(1/B)) | S(t)=[(x+(1/B))/t+x+(1/B))]^@+1 |
| 1 | 8 | 1 | 0 | 6 | 6.0943 | 6.0943 | 1.0000 |
| 2 | 29 | 1 | 4 | 6 | 6.0943 | 10.0943 | 0.2282 |
| 3 | 33 | 1 | 5 | 8 | 8.0943 | 13.0943 | 0.2446 |
| 4 | 17 | 1 | 8 | 8 | 8.0943 | 16.0943 | 0.1337 |
| 5 | 61 | 1 | 8 | 9 | 9.0943 | 17.0943 | 0.1576 |
| 6 | 0 | 1 | 10 | 9 | 9.0943 | 19.0943 | 0.1140 |
| 7 | 30 | 1 | 10 | 10 | 10.0943 | 20.0943 | 0.1332 |
| 8 | 35 | 1 | 11 | 12 | 12.0943 | 23.0943 | 0.1505 |
| 9 | 31 | 1 | 11 | 12 | 12.0943 | 23.0943 | 0.1505 |
| 10 | 61 | 1 | 12 | 12 | 12.0943 | 24.0943 | 0.1329 |
| 11 | 32 | 1 | 13 | 12 | 12.0943 | 25.0943 | 0.1180 |
| 12 | 5 | 1 | 13 | 13 | 13.0943 | 26.0943 | 0.1328 |
| 13 | 30 | 1 | 14 | 13 | 13.0943 | 27.0943 | 0.1190 |
| 14 | 23 | 1 | 16 | 13 | 13.0943 | 29.0943 | 0.0966 |
| 15 | 18 | 1 | 17 | 14 | 14.0943 | 31.0943 | 0.0986 |
| 16 | 13 | 1 | 17 | 15 | 15.0943 | 32.0943 | 0.1099 |
| 17 | 21 | 1 | 18 | 15 | 15.0943 | 33.0943 | 0.1004 |
| 18 | 17 | 1 | 20 | 17 | 17.0943 | 37.0943 | 0.1035 |
| 19 | 10 | 1 | 21 | 18 | 18.0943 | 39.0943 | 0.1048 |
| 20 | 12 | 1 | 23 | 20 | 20.0943 | 43.0943 | 0.1071 |
| 21 | 11 | 1 | 29 | 21 | 21.0943 | 50.0943 | 0.0795 |
| 22 | 16 | 1 | 30 | 29 | 29.0943 | 59.0943 | 0.1256 |
| 23 | 20 | 1 | 30 | 34 | 34.0943 | 64.0943 | 0.1575 |
| 24 | 11 | 1 | 31 | 36 | 36.0943 | 67.0943 | 0.1628 |
| 25 | 13 | 1 | 32 | 42 | 42.0943 | 74.0943 | 0.1910 |
| 26 | 4 | 1 | 33 | 44 | 44.0943 | 77.0943 | 0.1948 |
| 27 | 8 | 1 | 35 | 61 | 61.0943 | 96.0943 | 0.2655 |
| 28 | 10 | 1 | 61 | 61 | 61.0943 | 122.0943 | 0.1317 |
| 29 | 14 | 1 | 61 | 61 | 61.0943 | 122.0943 | 0.1317 |