SURVIVAL ANALYSIS OF PROSTATE CANCER PATIENTS: PARAMETRIC INSIGHTS

Gitanjali Pradhani[1], Anirban Dutta[2] and Jonali Gogoi[3,*]

[1]Research Scholar, Department of Statistics, Assam University, Silchar

E-mail id: pradhanig18@gmail.com

[2]Research Scholar, Department of Statistics, Assam University, Silchar

E-mail id: anirban7991@gmail.com

[3]Assistant Professor, Department of Statistics, Assam University, Silchar

*Corresponding Author E-mail id: zonalie04@gmail.com

Abstract

This research focuses on employing parametric survival analysis to model the mortality rates of prostate cancer patients. By utilizing parametric distributions, such as the Weibull, Exponential, Gompertz, Generalized-Gamma and Log-logistic distributions, we aim to characterize the patterns of death events in this patient population. The Exponential model is identified as the best fit, supported by the lowest AIC value (4758.391) and competitive Log Likelihood (-2378.196), signifying a constant hazard over time and alignment with observed survival patterns. The survival plots which have been generated using estimated parameters from the Exponential distribution and alternative models, visually depict their ability to capture prostate cancer survival patterns. These plots offer an accessible means of comparing and understanding the temporal aspects of survival predictions for prostate cancer patients.

Key words: Parametric Survival Analysis, Prostate cancer, Parametric distributions.

1. INTRODUCTION

Prostate cancer typically grows slowly and does not show noticeable symptoms. This often goes unnoticed by patients during their daily activities. As a result of which the condition becomes worse and the mortality rate goes up. Thus, the current study feels a need to understand the time until a patient survives if diagnosed with prostate cancer. This is one of the most frequently occurring cancers among men worldwide [5]. Known for its typically slow progression and asymptomatic nature, prostate cancer poses a unique challenge as it often evades early detection. This leads to a more severe condition and an elevated mortality rate. The dangerous nature of prostate cancer underscores the critical need for a comprehensive understanding of its temporal dynamics and the factors influencing patient survival.

Survival analysis encompasses various statistical methods for analyzing time-to-event data. Two widely used approaches are the Kaplan-Meier (KM) analysis and the Cox proportional hazards model. The Kaplan-Meier method is a non-parametric technique that estimates the survival function, providing the probability of an individual surviving beyond a given time point. On the other hand, Cox regression considers the covariates associated with the particular target variable. A notable trend in recent literature involves the adoption of parametric survival analysis to examine the survival experiences of prostate cancer patients. This statistical approach goes beyond conventional methods, such as Kaplan-Meier curves, by fitting parametric models to survival data.

Parametric survival analysis of prostate cancer involves analyzing the survival times of patients using parametric models. Several studies have explored this topic. Leuchter et al. compared relative survival rates of prostate cancer patients from a regional cancer registry and emphasized the importance of considering stage differentiation for accurate survival analysis [4]. Han et al. constructed a prognostic nomogram model based on Gleason grade, total prostate-specific antigen, alkaline phosphate (ALP), and TNM stage to predict progression-free survival in prostate cancer patients [7]. Chan performed parametric survival analysis to compare the survivorship of prostate cancer patients by race, specifically comparing White and African American men [6]. Balogun analyzed survival times of prostate cancer patients using the Kaplan-Meier Product Limit method and compared the survival function, median survival time, and hazard rates [3].

Taking a note of this, the present study seeks to shed light on the temporal aspects of survival in prostate cancer patients. Given the inherent complexities of the disease progression, this research adopts a parametric survival analysis approach, leveraging statistical models to explore mortality rates. By employing parametric distributions, including the Weibull, Exponential, Gompertz, Generalized-Gamma and Log-logistic distributions and our aim is to characterize the patterns associated with death events in this specific patient population.

2. METHODOLOGY

This section provides the methods and materials utilized for the study. The dataset for this study was sourced from The Cancer Genome Atlas (TCGA) website, a comprehensive repository of genomic and clinical data for various cancer types. The focus of data collection cantered on prostate cancer patients, aiming to extract information pertinent to their survival times. The dataset includes details on patients' outcomes and is enriched with censoring information, a crucial component in survival analysis.

To model the survival times of prostate cancer patients, five parametric distributions—Weibull, Exponential, Generalized- Gamma, Gompertz and Log-Normal were selected. These distributions offer diverse shapes and characteristics, allowing for a nuanced exploration of the survival patterns in prostate cancer survival.

The R software was employed for the implementation of parametric survival analysis.

For each selected distribution, parametric survival models were fitted to the survival data using maximum likelihood estimation. This involved optimizing the parameters of the chosen distribution to maximize the likelihood of observing the given survival times and censoring information in the dataset.

The Akaike Information Criterion (AIC) was employed as a key metric for model evaluation. AIC balances the goodness of fit with the complexity of the model, aiding in the selection of the most appropriate model. Lower AIC values indicate a better trade-off between model fit and complexity.

3. RESULTS AND DISCUSSION

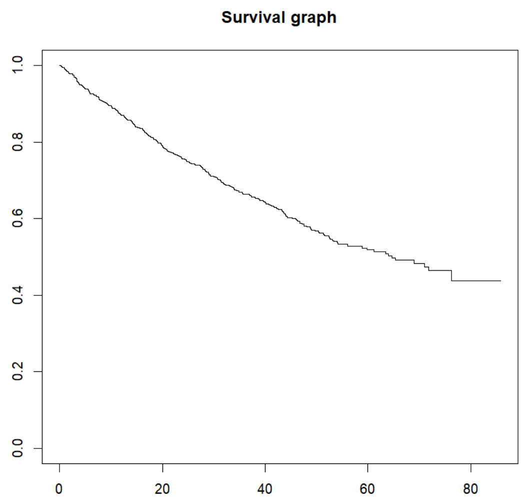
The fitted parametric survival models, encompassing Weibull, Lognormal, Exponential, Generalized-Gamma, and Gompertz distributions, were rigorously evaluated using Maximum Likelihood Estimates (MLEs), Akaike Information Criterion (AIC) and Log Likelihood metrics. **Table 1** provides a comprehensive overview of the model estimates, AIC values, and Log Likelihoods for each distribution in the context of prostate cancer survival.

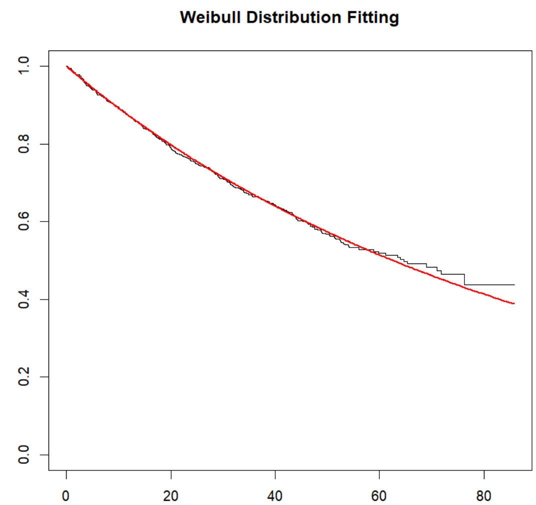
Table 1: MLE's, AIC and $-2lnL$ of the fitted distributions of prostate cancer data

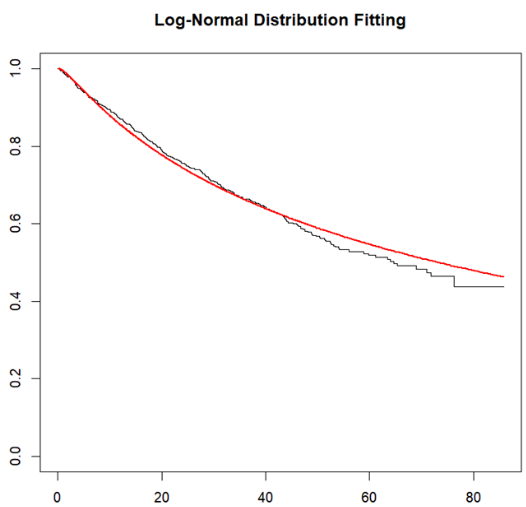
Model	Estimates	Log Likelihood	AIC
Weibull	shape=0.982 scale=90.801	-2378.095	4760.191
Lognormal	meanlog=4.2936 sdlog=1.7020	-2383.127	4770.254
Exponential	rate=0.011194	-2378.196	4758.391
Generalized-Gamma	μ=4.4554 σ=1.2331 Q=0.6686	-2376.961	4759.921
Gompertz	shape=0.003274 rate=0.011993	-2377.572	4759.145

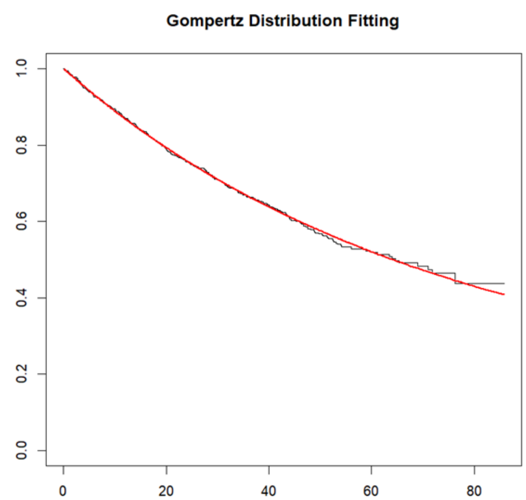
While each distribution provides valuable insights into prostate cancer survival, the Exponential model emerges as the best fit, as indicated by the lowest AIC value (4758.391) and a competitive log likelihood (-2378.196). The Exponential distribution is characterized by a constant hazard over time, suggesting a consistent rate of events, aligning well with the observed survival patterns in the dataset.

After fitting various distributions to the prostate cancer dataset, we generated survival plots using the estimated parameters from the best-fitting Exponential distribution and other models like Weibull, Lognormal, Generalized-Gamma, and Gompertz. These plots visually represent how well each distribution captures the survival patterns in prostate cancer. The survival plots provide an accessible way to compare and understand the different models' predictions regarding the temporal aspects of survival in prostate cancer patients.


Survival graph
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80


Weibull Distribution Fitting
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80


Log-Normal Distribution Fitting
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80


Gompertz Distribution Fitting
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80

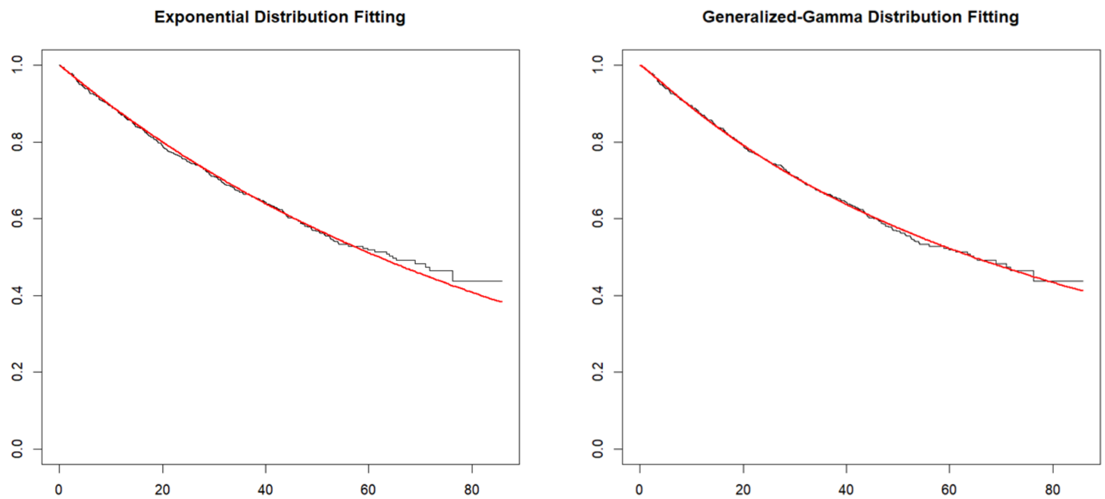


Figure 1: Comparative Visualization of Prostate Cancer Survival Patterns: Parametric Models vs. Kaplan-Meier Plot

In addition to the evaluation of parametric survival models, the study compares the generated plots from these distributions with the Kaplan-Meier plot. Remarkably, the visual comparison indicates that all mentioned distributions, including the Exponential, Weibull, Lognormal, Generalized-Gamma, and Gompertz, provide a good fit in terms of visualizing prostate cancer survival patterns.

4. CONCLUSION

Our evaluation of parametric survival models identified the Exponential distribution as the best fit for prostate cancer survival, supported by low AIC and competitive log likelihood values. Survival plots visually compared the Exponential model with alternatives. However, limitations include assumptions of distributional and proportional hazards, and the exclusion of external factors. Future research should explore more flexible modelling techniques, consider time-dependent covariates and competing risks, and leverage advancements in data collection for more accurate predictions. Despite valuable insights, ongoing refinement is crucial for a comprehensive understanding of prostate cancer survival dynamics.

REFERENCES

[1] Aladwani, M., Lophatananon, A., Ollier, W., & Muir, K. (2020). Prediction models for prostate cancer to be used in the primary care setting: a systematic review. BMJ open, 10(7), e034661.

[2] Baghestani, A. R., Moghaddam, S. S., Majd, H. A., Akbari, M. E., Nafissi, N., & Gohari, K. (2016). Survival analysis of patients with breast cancer using weibull parametric model. Asian Pacific Journal of Cancer Prevention, 16(18), 8567-8571.

[3] Chan, Y. M. (2013). Parametric survival analysis: A comparison of prostate cancer survivorship by race. International Journal of Mathematical Sciences in Medicine (IJMSM), 1(1).

[4] Dai, X., Park, J. H., Yoo, S., D'Imperio, N., McMahon, B. H., Rentsch, C. T., ... & Justice, A. C. (2022). Survival analysis of localized prostate cancer with deep learning. Scientific Reports, 12(1), 17821.

[5] Hamdy, F. C., Donovan, J. L., Lane, J. A., Metcalfe, C., Davis, M., Turner, E. L., ... & Neal, D. E. (2023). Fifteen-year outcomes after monitoring, surgery, or radiotherapy for prostate cancer. New England Journal of Medicine, 388(17), 1547-1558.

[6] Han, Y., Wen, X., Chen, D., Li, X., Leng, Q., Wen, Y., ... & Zhu, W. (2022). Survival analysis and a novel nomogram model for progression-free survival in patients with prostate cancer. Journal of Oncology, 2022.

[7] Leuchter, M., Kalata, P., Hildebrandt, G., Zettl, H., & Hakenberg, O. W. (2016). The parameter "relative survival" Analysis of regional cancer registry data for prostate cancer. Der Urologe, 55, 156-166.

[8] Sabouri, S., Esmaily, H., Shahid Sales, S., & Emadi, M. (2018). Determining related factors to survival of colorectal cancer patients using cox regression. Medical journal of mashhad university of medical sciences, 61(4), 1083-1092.

[9] Sainani, K. L. (2016). Introduction to survival analysis. PM&R, 8(6), 580-585.

[10] Zare, A., Mahmoodi, M., Mohammad, K., Zeraati, H., Hosseini, M., & Naieni, K. H. (2014). A comparison between Kaplan-Meier and weighted Kaplan-Meier methods of five-year survival estimation of patients with gastric cancer. Acta Medica Iranica, 764-767.