

Bias in Classical Life Tables Under Censoring: A Comparative Study With Kaplan-Meier Estimation and Actuarial Estimation Using Real and Simulated Data

Lubomír Seif[†], Ondřej Vít[‡]

[†]ORCID: 0009-0003-7444-9425

[‡]ORCID: 0009-0003-7317-5856

Department of Statistics and Probability

Faculty of Informatics and Statistics

Prague University of Economics and Business

W. Churchill's square 4, 13067 Prague, Czech Republic

[†]Email: lubomir.seif@vse.cz

[‡]Email: Ondrej.vit@vse.cz

Lubomír Štěpánek^{1, 2, 3}

ORCID: 0000-0002-8308-4304

¹Department of Statistics and Probability

²Department of Mathematics

Faculty of Informatics and Statistics

Prague University of Economics and Business

W. Churchill's square 4, 13067 Prague, Czech Republic

Email: lubomir.stepanek@vse.cz

&

³Institute of Biophysics and Informatics

First Faculty of Medicine

Charles University

Salmovská 1, 12000 Prague, Czech Republic

Email: lubomir.stepanek@lf1.cuni.cz

Abstract—Classical life tables assume fully observed lifespans and ignore censoring, which can bias survival estimates. In this study, we compare the classical life table with two censoring-aware approaches: the Kaplan-Meier estimator applied to simulated censored lifetimes, and the actuarial estimator assuming uniform censoring within intervals.

Using mortality data for the Czech Republic (from year 2021), we show that both the mentioned alternative methods yield systematically lower survival and life expectancy estimates, compared to life tables-based approach. The actuarial estimator is the most conservative, and we provide a formal proof that it underestimates survival relative to Kaplan-Meier.

These differences have practical implications for actuarial pricing and longevity risk. We advocate incorporating survival analysis techniques into actuarial workflows when dealing with incomplete or censored data.

Index Terms—life tables, survival analysis, censoring, Kaplan-Meier estimator, actuarial estimator, life expectancy overestimation, life expectancy underestimation, actuarial science

I. INTRODUCTION

LIFE tables are a cornerstone of actuarial science and demographic research, offering standardized tools to assess mortality and survival in a given population. Based on aggregated counts of deaths and population exposures at each age, these tables enable the computation of key quantities such as age-specific death probabilities and life expectancy. In standard construction, however, life tables operate under the

assumption of fully observed lifespans and do not incorporate censoring the fact that, in many contexts, we do not observe the complete lifespan for every individual.

Although this assumption holds in retrospective national mortality data, it becomes increasingly problematic in settings involving survey-based samples, incomplete registration systems, or real-time monitoring, where censoring is inherent in the data structure. Censoring occurs naturally when the full path of survival is not observed, either due to loss to follow-up, study end, or delayed entry. Right-censoring means that we do not know when an individual would experience the event of interest because the monitoring period ended before the event could occur.

The distinction could be nontrivial. As noted by Slud [1], the classical construction of life tables, based on central death rates and complete cohorts, fails to account for right-censoring, which can bias survival estimates if present but ignored. Missov and Vaupel [2] further explore the impact of censoring on mortality measures, showing that even small amounts of right-censoring can influence life expectancy and related metrics. Modern survival analysis, in contrast, is equipped to handle censored observations through estimators such as Kaplan-Meier, Nelson-Aalen, and parametric survival models.

In this study, we revisit classical life table methods using mortality data from the Human Mortality Database (HMD) [3] for the Czech Republic in 2021. We begin by constructing a traditional life table and estimating survival probabilities and life expectancy under the assumption of complete data.

This research was supported by the grant no. F4/36/2025 which has been provided by the Internal Grant Agency of the Prague University of Economics and Business.

We then introduce synthetic censoring at the aggregate level and evaluate its effect on survival metrics. Next, we simulate individual-level lifetimes with right-censoring and estimate the survival function [4] using the Kaplan-Meier estimator, a non-parametric method that accommodates incomplete observation [5], [6].

In addition, we implement the actuarial estimator of the survival function, which assumes uniformly distributed censoring within fixed intervals and is widely used when exact death times are not available [7]. We analytically demonstrate that this estimator yields lower survival estimates than Kaplan-Meier under the same data, and we confirm this property empirically.

By comparing all three approaches, i.e. the classical life table, Kaplan-Meier, and actuarial estimator, we quantify the extent of deviation in survival curves and life expectancy measures due to censoring and method choice. Finally, we discuss the implications of these differences and possible underestimation of the survivals for actuarial pricing, reserving, and longevity risk assessment.

II. PRELIMINARIES

Life tables are traditionally constructed using a deterministic sequence of transformations, starting from observed age-specific death rates or death probabilities. The standard components of a complete life table include the probability of dying between ages x and $x+1$ (denoted as q_x), the number of survivors at age x (marked as l_x), the number of deaths at age x (d_x), the person-years lived between ages x and $x+1$ (here denoted as L_x), the total person-years remaining beyond a given age x (T_x), and the remaining life expectancy at age x (marked as e_x). These definitions are extensively documented in actuarial and demographic literature, for instance [8], [9], [10].

The death probability q_x is typically calculated from empirical data as

$$q_x = \frac{D_x}{N_x}, \quad (1)$$

where D_x is the number of deaths observed at age x , and N_x is the number of individuals exposed to risk at that age.

Starting from a radix $l_0 = 100,000$, the number of survivors at each age is recursively defined as

$$l_{x+1} = l_x(1 - q_x). \quad (2)$$

From this, the number of deaths is given by

$$d_x = l_x q_x, \quad (3)$$

and the number of person-years lived between ages x and $x+1$ is typically approximated by

$$L_x = l_x - \frac{1}{2}d_x, \quad (4)$$

assuming that deaths are uniformly distributed within each age interval. The cumulative future person-years T_x and the corresponding life expectancy e_x are then computed as

$$T_x = \sum_{k=x}^{\omega} L_k, \quad e_x = \frac{T_x}{l_x}, \quad (5)$$

respectively, where ω denotes the maximum age in the life table.

III. METHODS

In this section, we implement the above methodology using mortality and exposure data from the Human Mortality Database (HMD) for the Czech Republic in 2021. The resulting classical life table will serve as the benchmark for later comparison with methods that incorporate synthetic censoring.

All computations and visualizations were performed using the R language [11] with the packages `ggplot2` [12], `dplyr` [13], and `survival` [14], [15].

A. Classical Life Table Construction

Using mortality and exposure data from the Human Mortality Database for the Czech Republic in 2021, we constructed a complete life table following the standard actuarial approach. The key steps involved computing age-specific probabilities of death q_x , the number of survivors l_x , the number of deaths d_x , person-years lived L_x , cumulative person-years T_x , and remaining life expectancy e_x , always considering age x . The radix was set at $l_0 = 100,000$, and the last observed age $\omega = 109$ was treated as an absorbing state with $q_{109} = 1$, meaning that all remaining individuals are assumed to die within the interval $(109, 110)$, as is standard in demographic practice, [8], [9], and is feasible even for research purposes, including this proceeding.

Part of the resulting table for ages 0–9 is shown in Tables I and II. It illustrates the decreasing pattern of life expectancy with age and the consistent structure of death probabilities and exposures.

Fig. 1 presents the curve of remaining life expectancy e_x as a function of age. The smooth and decreasing trajectory reflects the typical aging pattern in low-mortality populations.

B. Incorporating Synthetic Censoring

To assess how censoring could influence life table calculations even when working with aggregated population-level data, we construct a synthetic scenario in which a portion of the population at each age is assumed to be right-censored. This mimics conditions observed in longitudinal surveys, clinical studies, or incomplete registration systems, where not all individuals' times of death are observed. Since the dataset Human Mortality Database (HMD) does not include (or even

TABLE I
FIRST 10 OBSERVATIONS OF HMD DATA (CZECH REPUBLIC, 2021)

Age	Deaths	Exposure
0	246	110751.7
1	15	109581.9
2	12	112244.1
3	13	114184.9
4	11	113540.5
5	6	112191.5
6	18	111140.5
7	14	109107.7
8	8	108884.5
9	3	109051.5

TABLE II
FIRST 10 OBSERVATIONS OF THE CONSTRUCTED CLASSICAL LIFE TABLE
(CZECH REPUBLIC, 2021)

Age	q_x	l_x	d_x	L_x	T_x	e_x
0	0.002221	100000.00	222.12	99888.94	7699599	76.996
1	0.000137	99777.88	13.66	99771.05	7599710	76.166
2	0.000107	99764.22	10.67	99758.89	7499939	75.177
3	0.000114	99753.56	11.36	99747.88	7400180	74.185
4	0.000097	99742.20	9.66	99737.37	7300433	73.193
5	0.000054	99732.54	5.33	99729.87	7200695	72.200
6	0.000162	99727.20	16.15	99719.13	7100965	71.204
7	0.000128	99711.05	12.79	99704.66	7001246	70.215
8	0.000073	99698.26	7.33	99694.60	6901542	69.224
9	0.000028	99690.93	2.74	99689.56	6801847	68.229

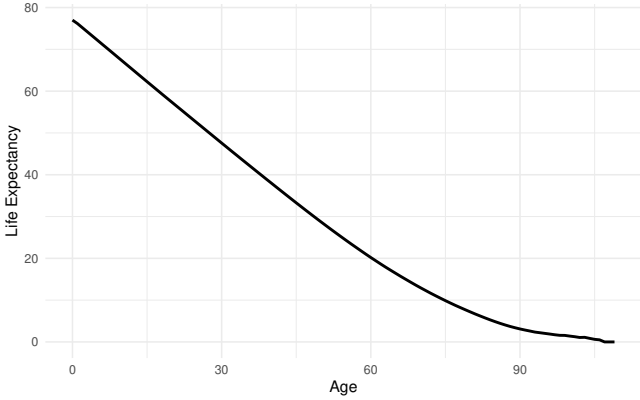


Fig. 1. Remaining life expectancy (in years) at age x based on classical life table construction.

consider) censoring, the synthetic, i.e. simulated censoring is the only way how to investigate its impact.

In our synthetic setup, we assume a fixed censoring probability c_x at each age. Specifically, we define $c_x = 0.10$, i.e., 10 % of the exposure at each age is assumed to be censored. This means that instead of observing N_x person-years fully at risk, we effectively observe only $(1 - c_x)N_x$ as uncensored exposure, while $c_x N_x$ is assumed to be censored.

Following the framework suggested by Slud [1] and the analytical concerns raised in Missov and Vaupel [2], we adjust the death probability coming from formula (1) calculation accordingly,

$$\tilde{q}_x = \frac{D_x}{(1 - c_x)N_x}, \quad (6)$$

where $\tilde{q}_x$ represents the adjusted probability of death under censoring. All subsequent life table quantities $\tilde{l}_x, \tilde{d}_x, \tilde{L}_x, \tilde{T}_x, \tilde{e}_x$ from formulae (2, 3, 4, 5) are computed using this modified death probability $\tilde{q}_x$, replicating the structure of the classical method.

We emphasize that this adjustment does not imply individual-level censoring but rather a proxy for its effect at the population level. The goal is to examine whether such synthetic censoring, even when introduced into aggregate mortality data, leads to meaningful differences in key survival metrics such as life expectancy.

In the next subsection, we present the life table results under the synthetic censoring assumption and compare them to the classical baseline.

C. Comparison of Results under Censoring

Fig. 2 shows the comparison between the classical and synthetically censored life tables in terms of remaining life expectancy. As expected, the life expectancy under synthetic censoring is systematically lower in all ages. This is a direct consequence of inflating the estimated mortality probabilities when only partial exposures are assumed to be observed. In other words, if classical life tables-based approach assumes non-censored framework when probability of death at given age x is, using formula (1), equal to $q_x = \frac{D_x}{N_x}$, while the death probability under censoring is (6), $\tilde{q}_x = \frac{D_x}{(1 - c_x)N_x}$, obviously is $\tilde{q}_x > q_x$ for any $c_x > 0$, i.e., whenever there is nonzero censoring we are not aware of, the probability of death is in fact a bit higher and, consequently, the true remaining life expectancy is a bit lower.

When we assume that data are not censored we might get falsely optimistic classical estimate used in life tables. When incorporating censoring we can always see the downward shift in estimated life expectancy curve which might be closer to the reality.

In further sections we discuss methods for estimating survival function which is then used to model life expectancy. We present the actuarial estimator and the Kaplan-Meier estimator.

IV. ACTUARIAL (LIFETABLE) ESTIMATOR

One of the approaches for estimating the survival function, i.e., a probability that an object will survive longer than a specified time point [4], is the actuarial or life table estimator [7]. This method is suitable for interval-censored or grouped data where exact times of death or censoring are not known. It is commonly used in demography and actuarial science when data are available in age or time intervals.

Let the time axis be partitioned into intervals (t_{j-1}, t_j) . For each interval, we define

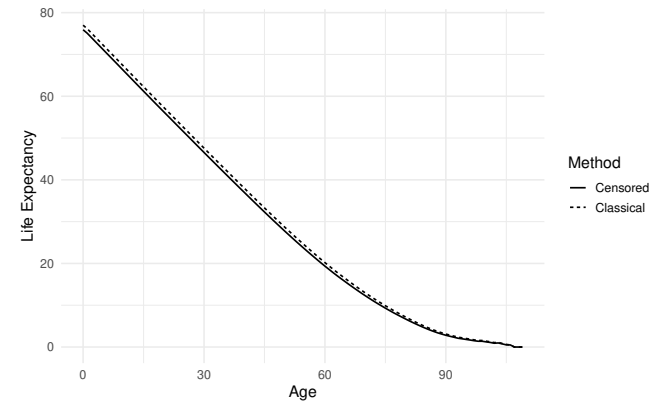


Fig. 2. Comparison of remaining life expectancy between classical life table and synthetic censoring scenario ($c_x = 10\%$).

- r_j as the number of individuals entering interval j ,
- d_j as the number of deaths observed in interval j ,
- c_j as the number of right-censored observations in interval j .

Assuming uniform censoring within the interval, the effective number at risk is approximated by

$$r'_j = r_j - \frac{c_j}{2}, \quad (7)$$

clearly saying that at a midpoint of the given interval (j), there is just half the censoring ($\frac{c_j}{2}$) of the entire interval (c_j). The conditional survival probability in interval j is then

$$p_j = \frac{r'_j - d_j}{r'_j} = 1 - \frac{d_j}{r'_j}. \quad (8)$$

The actuarial estimate of the survival function $\hat{S}(t)$ is obtained as the cumulative product,

$$\hat{S}_A(t) = \prod_{j:t_j \leq t} p_j, \quad (9)$$

since, if an individual survived up to interval j , they had to survive also up to intervals $j-1, j-2, \dots, 1$. Thus, we have to mutually multiply the interval-related probabilities p_j .

This estimator is generally more conservative in estimating survival than Kaplan-Meier, especially when censoring is heavy or uneven. The assumption of uniformly distributed censoring within intervals can lead to a downward bias in estimated survival compared to methods that use exact event times.

V. KAPLAN-MEIER ESTIMATOR

The Kaplan-Meier estimator is a non-parametric method for estimating the survival function $S(x)$, which represents the probability that a subject survives beyond age x [5], [6]. It is particularly well-suited for handling right-censored data, where the exact time of death is not observed for some individuals.

Let $t_1 < t_2 < \dots < t_k$ denote the distinct observed failure times in the sample, d_j the number of events (deaths) between time points t_{j-1} and t_j , and r_j the number of individuals at risk just before t_j . Then the Kaplan-Meier estimator is given by

$$\hat{S}_{KM}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{r_j}\right). \quad (10)$$

This step function decreases only at observed event times and remains constant in intervals where no events occur. Censored observations reduce the risk set r_j at subsequent time points but do not directly contribute to d_j .

Once the survival function is estimated, the life expectancy at birth e_0 can be computed as the area under the curve (which is not differentiable but is continuous, though),

$$e_0^{KM} = \int_0^\infty \hat{S}(x) dx, \quad (11)$$

which in practice is approximated as the sum over the discrete age grid,

$$e_0^{KM} \approx \sum_x \hat{S}(x), \quad (12)$$

where $\hat{S}(x)$ is an estimate of survival function, modeled using actuarial estimate $\hat{S}_A$ or Kaplan-Meier estimate $\hat{S}_{KM}$, as introduced before. This provides a consistent estimator of the mean remaining lifetime even in the presence of censoring, unlike classical life table methods that assume fully observed (non-censored) data.

We compare the actuarial estimator with the Kaplan-Meier estimator in the next section and quantify their divergence in terms of survival function and life expectancy.

VI. THE ACTUARIAL ESTIMATOR UNDERESTIMATES SURVIVAL COMPARED TO KAPLAN-MEIER ESTIMATOR

We now provide an analytical justification for why the actuarial estimator $\hat{S}_A(t)$ is systematically lower than the Kaplan-Meier estimator $\hat{S}_{KM}(t)$ under the same data.

Recall that in Kaplan-Meier from formula (10), the survival function decreases only at exact failure times,

$$\hat{S}_{KM}(t_j) = \hat{S}_{KM}(t_{j-1}) \cdot \left(1 - \frac{d_j}{r_j}\right), \quad (13)$$

where r_j is the number at risk just before time point t_j .

By contrast, the actuarial estimator assumes that censoring is uniformly distributed within interval $[t_{j-1}, t_j]$, so the effective number at risk is adjusted to

$$r'_j = r_j - \frac{c_j}{2}, \quad (14)$$

and the survival estimate is

$$\hat{S}_A(t_j) = \hat{S}_A(t_{j-1}) \cdot \left(1 - \frac{d_j}{r'_j}\right). \quad (15)$$

As comes from formulae (13, 15), terms $1 - \frac{d_j}{r_j}$ and $1 - \frac{d_j}{r'_j}$ are the key. Because $r'_j = r_j - \frac{c_j}{2} < r_j$ in most practical cases (i.e., if any censoring $c_j > 0$ occurs before the end of the interval), we have

$$\frac{d_j}{r'_j} > \frac{d_j}{r_j}, \quad (16)$$

which implies

$$1 - \frac{d_j}{r'_j} < 1 - \frac{d_j}{r_j}, \quad (17)$$

for every time point $t_j \leq t$, and therefore, multiplying the terms from previous formula (17), we get

$$\begin{aligned} 1 - \frac{d_j}{r'_j} &< 1 - \frac{d_j}{r_j} \quad \forall t_j \in \{t_1, t_2, \dots, t\} \\ \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{r'_j}\right) &< \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{r_j}\right) \\ \hat{S}_A(t) &\stackrel{(8,9,10)}{<} \hat{S}_{KM}(t). \end{aligned} \quad (18)$$

This inequality propagates multiplicatively, so the entire survival curve under the actuarial method lies below the Kaplan-Meier curve, if any censoring occurs. Hence, the actuarial estimator systematically underestimates survival compared to Kaplan-Meier, especially when censoring is substantial.

This property also directly impacts life expectancy estimates derived as area under the survival curve.

VII. ESTIMATION WITH SIMULATED CENSORING

To analyze the effect of censoring using a statistically rigorous survival model, we simulated individual-level lifetimes from the classical life table probabilities q_x and applied a 10 % right-censoring mechanism. We then estimated the survival function using the Kaplan-Meier estimator.

In parallel, we computed the actuarial survival function based on interval-level data and uniform censoring assumptions. The corresponding life expectancy was obtained as the discrete area under the actuarial survival curve,

$$e_0^A = \sum_x \hat{S}_A(x), \quad (19)$$

where $\hat{S}_A(x)$ denotes the actuarial survival estimate at the beginning of age x . This summation provides a consistent approximation of the expected remaining lifetime under the actuarial model.

Fig. 3 shows the comparison between the Kaplan-Meier estimate, actuarial estimate and the classical life table survival function $S(x) = l_x/l_0$, as the first two come from formulae (10, 9). The Kaplan-Meier curve lies systematically below the classical curve, indicating shorter estimated survival due to the presence of censoring. The curve estimated by the actuarial estimate lies below both curves. Fig. 4 shows estimated survival for ages above 75 years. The differences between the estimates, and, consequently, between the treatment of the censoring, are more clear in this figure.

In terms of overall life expectancy at birth¹, the classical method based on l_x gave $e_0^{LT} = 77.50$ (95% CI [77.43, 77.57]), while the Kaplan-Meier estimator yielded $e_0^{KM} = 76.54$ (95% CI [76.46, 76.63]) and the actuarial estimate gave $e_0^A = 75.97$ (95% CI [75.90, 76.04]), respectively. The differences here are quite obvious. Between classical method and Kaplan-Meier estimate, the difference is approximately 0.96 years. Between classical estimate and actuarial estimate, the difference is 1.53 years and when comparing Kaplan-Meier and actuarial estimate the difference shows 0.57 years.

VIII. DISCUSSION

The results of our study provide a comprehensive assessment of how censoring affects survival estimation using three different approaches: the classical life table, the Kaplan-Meier

¹To report variability of the simulated estimates, 95% confidence intervals are constructed using the percentile bootstrap, i.e., using the 0.025-th and 0.975-th quantiles of the bootstrapped distribution of each estimator. Confidence intervals are reported in square brackets, preceded by the prefix “95% CI”.

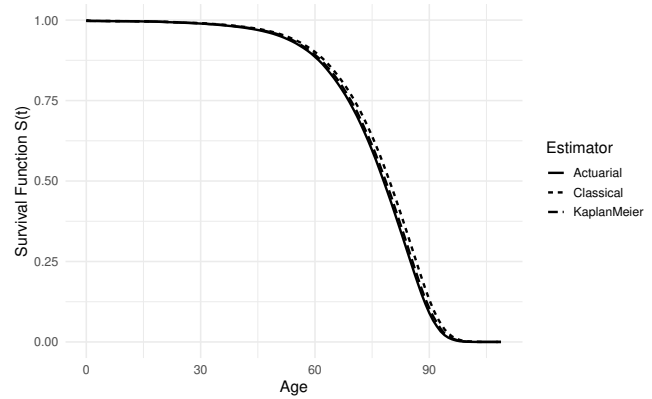


Fig. 3. Comparison of survival functions: Kaplan-Meier estimate, actuarial estimate and life table estimate.

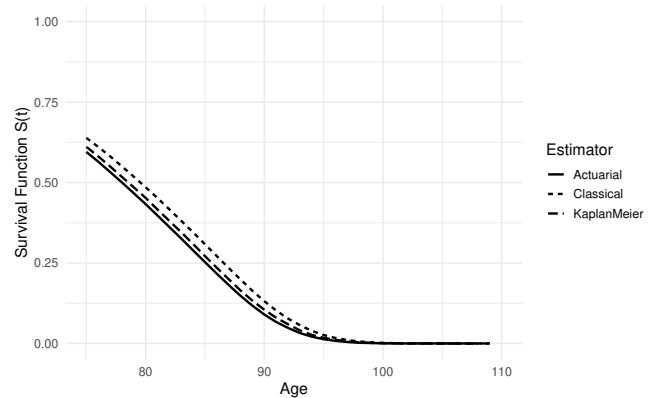


Fig. 4. Comparison of survival functions for ages 75+ years.

estimator, and the actuarial survival estimator. All methods were applied consistently to the same underlying mortality data for the Czech Republic in 2021, using synthetic (simulated) censoring.

Our analyses revealed that the assumptions of no censoring in life table estimates might lead to falsely optimistic estimates. Incorporating censoring could bring the estimate closer to the reality. The classical life table estimate of life expectancy at birth was 77.50 years (95% CI [77.43, 77.57]), while the Kaplan-Meier estimator applied to simulated censored lifetimes yielded 76.54 years (95% CI [76.46, 76.63]), a nearly one-year difference. When the actuarial estimator was applied under uniform censoring assumptions (75.97 with 95% CI [75.90, 76.04]), life expectancy was lower still, confirming analytically and empirically that this method is the most conservative among the three.

From a demographic perspective, this highlights the critical importance of censoring-aware methodologies, especially in datasets where full lifespan observation is implausible. Ignoring censoring can lead to biased survival estimates and misleading conclusions about population longevity.

In an actuarial context, these differences carry tangible

consequences. Life expectancy estimates directly influence pricing, reserving, and solvency calculations in life insurance and pension products. A lower estimated life expectancy typically leads to higher premiums for annuity products and lower reserves for life insurance, potentially introducing pricing inaccuracies if censoring is not handled correctly. Conversely, underestimating survival in products with longevity risk exposure may lead to systematic underreserving, e.g. two people with the same product but different life expectancy should lead to higher reserves for the person with higher probability of death. But when censoring is not accounted for these reserves could be calculated to lower volume leading to underreserving. In other words when calculating death insurance under falsely optimistic estimates, lower probability of death (higher life expectancy) could lead to cheaper insurance policies and vice versa.

These findings reinforce the need to integrate non-parametric survival estimators such as Kaplan-Meier into actuarial workflows when individual-level or censored data are available. Even when working with grouped data, awareness of the limitations of life table and actuarial approximations is essential.

IX. FUTURE OUTLOOKS

This study opens several directions for further research. First, extending the analysis to various censoring levels (e.g., censoring of 5 %, 20 %, 30 % of available data) would reveal how estimator performance changes with increasing data incompleteness. Second, stratifying results by subgroups such as sex or region could uncover heterogeneity in censoring effects, which is relevant in pricing and reserving. Third, comparing classical and non-parametric estimators with parametric survival models, such as Weibull or Gompertz, may enhance accuracy where model assumptions are plausible. Finally, quantifying how survival differences (e.g., in life expectancy) translate into financial impacts, such as annuity prices or reserves, would strengthen the actuarial relevance of this work.

X. CONCLUSION

This study demonstrates that classical life tables, while effective under complete observation, can yield upwardly biased survival estimates in the presence of censoring. We showed that both synthetic censoring-like adjustments to exposure and non-parametric survival methods yield systematically lower survival curves and life expectancy values.

Among the three methods considered, the actuarial estimator produced the lowest survival estimates due to its structural assumption of uniform censoring. The Kaplan-Meier estimator, by leveraging precise event and censoring times, offered a more accurate and interpretable estimate under realistic data conditions.

Although the absolute difference in life expectancy across methods may seem modest, around one year at birth, such

discrepancies can significantly impact actuarial outcomes. For annuity products, underestimating survival leads to financial strain due to underestimated liabilities, while for life insurance, overestimation may affect competitive pricing and risk pooling.

We conclude that integrating survival analysis techniques, particularly those accounting for censoring, is essential for reliable longevity modeling. These tools not only improve the statistical fidelity of survival estimates but also support more robust decision-making in insurance, demography, and public health.

REFERENCES

- [1] Slud, Eric V. *Actuarial Mathematics and Life-Table Statistics*. University of Maryland, 2006. Available at <https://www.math.umd.edu/~slud/s470/BookChaps/01Book.pdf>.
- [2] Missov, Tiago V., and James W. Vaupel. "How much can we trust life tables? Sensitivity of mortality measures to right-censoring." *Palgrave Communications* 1 (2015): 1–6. <https://doi.org/10.1057/palcomms.2015.49>.
- [3] HMD. Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at www.mortality.org (data downloaded on 22.5.2025).
- [4] Štěpánek, L., Habarta, F., Malá, I., Marek, L. (2024). Non-parametric comparison of survival functions with censored data: A computational analysis of greedy and Monte Carlo approaches. *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*, 39, 725–730. <http://dx.doi.org/10.15439/2024F223>
- [5] Kaplan, E. L., & Meier, P. "Nonparametric estimation from incomplete observations." (1958). *Journal of the American Statistical Association*, 53(282), 457–481. <https://doi.org/10.1057/palcomms.2015.49>.
- [6] Štěpánek, L., Habarta, F., Malá, I., Marek, L. (2020). Analysis of asymptotic time complexity of an assumption-free alternative to the log-rank test. *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems*, 21, 453–460. <http://dx.doi.org/10.15439/2020F198>
- [7] Clark, T. G., Bradburn, M. J., Love, S. B., and Altman, D. G. (2003). Survival analysis part I: Basic concepts and first analyses. *British Journal of Cancer*, 89(2), 232–238. <https://doi.org/10.1038/sj.bjc.6601118>
- [8] Social Security Administration. *Definitions of Life Table Functions*. SSA, (2016). <https://www.ssa.gov/oact/Downloadables/LifeTableDefinitions.pdf>
- [9] Office for National Statistics. *Guide to calculating national life tables*. ONS, (2019). <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/methodologies/guidetocalculatingnationallifetables>
- [10] Furman, Edward. *Actuarial Mathematics*. York University, (2020). <https://edfurman.info.yorku.ca/files/2020/09/MATH3280-2020-lec2.pdf>
- [11] R Core Team (2024). *_R: A Language and Environment for Statistical Computing_*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- [12] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- [13] Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *_dplyr: A Grammar of Data Manipulation_*. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.
- [14] Therneau T (2024). *_A Package for Survival Analysis in R_*. R package version 3.6-4, <https://CRAN.R-project.org/package=survival>.
- [15] Terry M. Therneau, Patricia M. Grambsch (2000). *_Modeling Survival Data: Extending the Cox Model_*. Springer, New York. ISBN 0-387-98784-3.