symmetry

MDPI

Article

Entropy-Regularized Likelihood Inference for Lifetime Distributions Under Progressive Type-II Censoring †

Ayse Bugatekin *, Mine Dogan and Gökhan Gökdere

Department of Statistics, Faculty of Science, Fırat University, Elazig 23119, Turkey; mine.dogan@firat.edu.tr (M.D.); ggokdere@firat.edu.tr (G.G.)

* Correspondence: aturan@firat.edu.tr

† A preliminary version of this study was presented as a conference abstract at the 8th International Conference on Engineering and Applied Natural Sciences (ICEANS 2026), held on July 25–26, 2026, in Konya, Türkiye.

Abstract

This study proposes an entropy-regularized likelihood inference framework for lifetime distributions under progressive Type-II censoring. By incorporating Shannon entropy directly into the classical likelihood function, the proposed approach aims to alleviate the information loss caused by censoring and improve the finite-sample stability of parameter estimation. Entropy-regularized maximum likelihood estimators (ERMLEs) are developed for the Exponential, Weibull, Gamma, and Lognormal lifetime distributions. Distribution-specific regularization parameters are selected by minimizing the average mean squared error across a comprehensive Monte Carlo simulation study covering different sample sizes, censoring rates, and progressive censoring schemes. Estimation performance is evaluated using bias, mean squared error, and the relative reduction in MSE achieved by ERMLE. The proposed methodology is further illustrated using two progressively Type-II censored real datasets from engineering reliability and biomedical survival analysis. Model adequacy is assessed through goodness-of-fit statistics with corresponding p-values, bootstrap confidence intervals, and graphical comparisons. The simulation results show that entropy regularization substantially improves estimation accuracy for the Exponential, Weibull, and Gamma distributions, particularly under moderate and heavy censoring, whereas only negligible improvements are observed for the Lognormal distribution. In the engineering reliability application, the Weibull distribution provides the best overall fit, while the Weibull and Gamma models exhibit the most satisfactory performance for the bladder cancer remission data. Across both applications, ERMLE yields parameter estimates and fitted models that are highly consistent with those of the classical MLE while providing stable estimation under progressive censoring. Overall, the proposed framework demonstrates that the effectiveness of entropy regularization is distribution-dependent rather than universal and provides practical guidance for selecting suitable estimation strategies in reliability and survival analysis.

Keywords: entropy regularization; progressive type-II censoring; maximum likelihood estimation; lifetime distributions; reliability analysis; survival analysis

Check for updates

Academic Editor: Calogero Vetro

Received: 10 July 2026
Revised: 3 August 2026
Accepted: 10 August 2026
Published: 13 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Lifetime and reliability analysis plays a fundamental role in many areas of engineering, industrial quality control, biomedical research, and risk assessment, where understanding the stochastic behavior of failure times is essential for system design and maintenance planning. In practice, however, complete lifetime observations are rarely available

because experiments are often terminated prematurely due to time limitations, economic considerations, or destructive testing. As a result, statistical inference must frequently be performed using censored lifetime data rather than complete observations.

Among the various censoring mechanisms, progressive Type-II censoring has received considerable attention because of its flexibility and practical relevance. Unlike conventional Type-II censoring, progressive schemes allow surviving experimental units to be removed at different stages of the life test while the experiment is still in progress. This feature enables investigators to reduce experimental costs and duration without completely sacrificing statistical information. Owing to these advantages, progressive Type-II censoring has become one of the standard censoring schemes in reliability engineering, survival analysis, and industrial life-testing experiments.

The statistical theory of progressive Type-II censoring has been extensively developed during the last three decades. Balakrishnan and Aggarwala [1] provided a comprehensive treatment of progressive censoring models and their applications, while Balakrishnan and Sandhu [2] proposed an efficient simulation algorithm for generating progressively Type-II censored samples. Subsequently, Balakrishnan et al. [3] investigated the distributional properties of progressively censored order statistics under the Exponential model, whereas Ng et al. [4] developed statistical inference procedures for exponentially distributed lifetimes under adaptive progressive Type-II censoring. These studies established the theoretical foundation for likelihood-based inference under progressive censoring and stimulated extensive research on parametric lifetime models.

Parallel to these developments, information-theoretic measures, particularly Shannon entropy, have become increasingly important for quantifying uncertainty and information loss in lifetime analysis. Entropy has been investigated for a wide range of lifetime distributions under different censoring mechanisms. For example, Gong and Yin [5] studied entropy inference for the generalized inverse exponential distribution under progressive Type-II censoring, Ahmed and Ismail [6] considered entropy estimation for the Lomax distribution under adaptive progressive censoring, Wang and Gui [7] developed Bayesian entropy estimation for the Burr Type XII model, and Cho et al. [8], Lee [9], and Hassan and Zaky [10] investigated entropy estimation for Weibull-type lifetime models. In addition, Kohansal and Rezakhah [11] proposed exponentiality tests based on Rényi entropy, whereas Alyami [12] compared Bayesian and classical estimation procedures for Tsallis entropy under censored observations. These studies demonstrate that entropy provides an effective framework for measuring uncertainty in incomplete lifetime experiments.

Considerable attention has also been devoted to parameter estimation under progressive censoring. Representative contributions include the work of Pandey [13] on the exponentiated exponential distribution, Teimouri [14] on generalized exponential models, Dutta and Kayal [15] on logistic exponential lifetimes, and Wei [16] on Bayesian inference for generalized inverse exponential distributions under hybrid censoring. Furthermore, several studies have investigated the theoretical relationship between entropy and likelihood-based inference [17–20], emphasizing the close connection between information measures and statistical estimation.

Despite these advances, an important gap remains in the existing literature. Most entropy-related studies employ entropy either as a quantity to be estimated or as an auxiliary criterion for model assessment and hypothesis testing. In contrast, relatively little attention has been devoted to incorporating entropy directly into the likelihood function as a regularization mechanism for parameter estimation under progressive Type-II censoring. More importantly, the existing literature implicitly assumes that entropy regularization improves estimation in a similar manner across different lifetime models, while

the possible dependence of regularization performance on the underlying distributional structure has received almost no theoretical or empirical investigation.

The existing literature reveals a clear progression in the analysis of censored lifetime data. Classical likelihood-based methods provide a well-established framework for parameter estimation under progressive censoring; however, their finite-sample performance may deteriorate as censoring reduces the amount of directly observed failure information. Entropy-based approaches have subsequently extended censored-data analysis from an information-theoretic perspective, but previous studies have predominantly treated entropy as a quantity to be estimated under different censoring mechanisms, as a basis for goodness-of-fit or distributional assessment, or as an information measure for comparing experimental and censoring designs. Thus, entropy has generally remained an inferential target or an evaluation criterion external to the parameter-estimation objective. Regularized likelihood methods, on the other hand, offer a mechanism for stabilizing estimation, but their regularization terms are typically introduced independently of the distributional uncertainty inherent in the assumed lifetime model. These developments reveal a methodological gap between likelihood-based inference, entropy-based characterization, and regularized estimation under progressive censoring. The present study addresses this gap by incorporating the Shannon entropy of the assumed lifetime distribution directly into the progressively censored log-likelihood as a distribution-specific regularization term. Consequently, entropy is not estimated as the primary quantity of interest; rather, it actively modifies the estimating equations and the local curvature of the likelihood surface through its gradient and Hessian. This formulation leads to an entropy-regularized maximum likelihood estimator specifically designed for parameter inference under progressive Type-II censoring. To the best of our knowledge, the direct use of distributional Shannon entropy as an additive likelihood regularizer, together with its theoretical and finite-sample investigation across multiple lifetime distributions under progressive Type-II censoring, has not been systematically studied in the existing literature.

Motivated by this gap, this study develops an entropy-regularized likelihood inference framework for progressively Type-II censored lifetime data. The proposed methodology integrates Shannon entropy into the classical likelihood function to stabilize estimation under information loss caused by censoring. Rather than restricting the investigation to a single lifetime distribution, the proposed framework is systematically examined for the Exponential, Weibull, Gamma, and Lognormal models through extensive Monte Carlo simulations and two real-data applications. Particular emphasis is placed on understanding how the effectiveness of entropy regularization varies across different parametric lifetime distributions, thereby providing practical guidance on when entropy-based regularization is expected to improve statistical inference under progressive censoring.

The main contributions of this study can be summarized as follows. First, an entropy-regularized likelihood framework is developed for progressively Type-II censored lifetime data by directly incorporating Shannon entropy into the likelihood function. Second, the proposed framework is systematically evaluated across four commonly used lifetime distributions under a unified simulation design and two real-data applications. Finally, the results demonstrate that the effectiveness of entropy regularization is distribution-dependent rather than universal, providing practical guidance for the use of entropy-based regularization in censored lifetime analysis.

2. Methodology

2.1. Progressive Type-II Censoring Scheme

Real-world lifetime data are frequently subject to censoring because complete failure times cannot always be observed due to economic costs, time limitations, or destructive

testing conditions. Consequently, statistical inference must often be performed using incomplete lifetime observations rather than complete samples. Among the various censoring mechanisms, progressive Type-II censoring has become one of the most widely used schemes in reliability engineering and survival analysis because it provides a flexible balance between experimental efficiency and statistical information.

Unlike conventional Type-I and Type-II censoring, progressive censoring allows surviving experimental units to be removed at different stages of the life test while the experiment is still in progress. This flexibility considerably reduces experimental cost and duration without completely sacrificing the information contained in the observed failures. Owing to these practical advantages, progressive censoring has received considerable attention in both theoretical and applied statistics over the past several decades.

A progressive Type-II censored experiment can be described as follows. Let $m < n$ and choose non-negative integers $R_1,\ R_2, \ldots, R_m$ satisfying

$$R_1 + R_2 + \cdots + R_m = n - m \tag{1}$$

Suppose that n identical units are placed on a life test, where the lifetimes are assumed to be independent and identically distributed random variables with common distribution function F. At the occurrence of the first observed failure, denoted by $x_{1:m:n}$, R_1 surviving units are randomly withdrawn from the experiment. Similarly, at the second observed failure $x_{2:m:n}$, R_2 surviving units are removed, and the process continues until the m-th observed failure. At the final observed failure $x_{m:m:n}$, all remaining surviving units are withdrawn from the experiment. Consequently, the observed progressively censored sample is

$$x_{1:m:n} < x_{2:m:n} < \cdots < x_{m:m:n} \tag{2}$$

Progressive Type-II censoring has been extensively investigated in the literature because of its wide applicability in reliability engineering, survival analysis, and industrial life-testing experiments [1,21–24]. The progressively censored sample defined above forms the basis for constructing the likelihood function used throughout this study.

However, progressive censoring inevitably reduces the amount of available failure information compared with complete lifetime samples. As the censoring level increases, the likelihood surface may become flatter and parameter estimation may exhibit increased variability and reduced numerical stability. This information loss provides the primary motivation for introducing an entropy-based regularization mechanism into the likelihood function. Rather than replacing the classical likelihood framework, the proposed approach aims to stabilize inference by incorporating an additional information-theoretic component while preserving the probabilistic structure of the underlying lifetime model. The construction of the entropy-regularized likelihood is presented in the next subsection.

2.2. Maximum Likelihood Structure

Let X be a continuous lifetime random variable representing the failure time of an experimental unit, and assume that X follows a lifetime distribution indexed by the parameter vector

$$\theta = (\theta_1, \theta_2, \ldots, \theta_p) \in \Theta$$

where Θ denotes the parameter space and all unknown parameters satisfy the standard regularity conditions required for likelihood-based inference.

Suppose that n identical experimental units are placed on a life test and their lifetimes are independent and identically distributed random variables, $X_1, X_2, \ldots, X_n \sim F(x;\theta)$, with probability density function $f(x;\theta)$, cumulative distribution function $F(x;\theta)$, and survival function $S(x;\theta) = 1 - F(x;\theta)$.

For the observed progressively Type-II censored sample, the likelihood function based on censored order statistics is given by

$$L(\boldsymbol{\theta}; \mathbf{x}, \boldsymbol{R}) = C_{\boldsymbol{R}} \prod_{i=1}^{m} f(x_{i:m:n}; \theta)[1 - F(x_{i:m:n}; \theta)]^{R_i} \quad (3)$$

where

$$C_{\boldsymbol{R}} = \prod_{i=1}^{m}\left(n - i + 1 - \sum_{j=1}^{i-1} R_j\right) = n(n - R_1 - 1)(n - R_1 - R_2) - (n - m + 1 - \sum_{j=1}^{m-1} R_j).$$

Here, $C_{\boldsymbol{R}}$ is the combinatorial order factor induced by the sequential progressive censoring mechanism. Thus, the likelihood is defined for the ordered progressively Type-II censored sample rather than for mmm independent and identically distributed failure observations. Although the original lifetimes $X_1, X_2, \ldots, X_n$ are i.i.d., the observed progressively censored order statistics of a random, $X_{1:m:n}, \ldots, X_{m:m:n}$ are not independent, and their joint structure is incorporated through the ordered likelihood and the censoring configuration $\boldsymbol{R}$.

Taking the natural logarithm of the likelihood function yields the classical log-likelihood function

$$l(\theta) = logL(\theta)$$

Classical likelihood inference proceeds by maximizing $l(\theta)$ with respect to the unknown parameter vector θ. The corresponding score vector is defined as

$$U(\theta) = \frac{\partial l(\theta)}{\partial \theta}$$

and the maximum likelihood estimator is obtained from the solution of the likelihood equations

$$U(\theta) = 0$$

Under standard regularity conditions, the maximum likelihood estimator is asymptotically normal and efficient. Specifically,

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(0, I^{-1}(\theta)\right)$$

where $I(\theta)$ denotes the Fisher information matrix.

However, under progressive Type-II censoring, the effective amount of observed information decreases as the censoring level increases because only a subset of the original failure times contributes directly to the likelihood function. Consequently, the likelihood surface may become flatter, the Fisher information may decrease, and the resulting score equations may become more sensitive to sampling fluctuations. These effects are particularly pronounced under moderate and heavy censoring schemes, where classical maximum likelihood estimation may exhibit increased variance, reduced accuracy, and numerical instability. This limitation motivates the development of regularized likelihood formulations capable of incorporating additional information into the estimation process while preserving the probabilistic structure of the underlying lifetime model.

2.3. Shannon Entropy and Information-Theoretic Motivation

Information-theoretic quantities have become increasingly important in modern statistical inference because they provide quantitative measures of uncertainty beyond the information contained in observed data alone. Among the various entropy measures proposed in the literature, Shannon entropy occupies a central position owing to its close relationship

with likelihood theory through the expected log-density. This connection makes Shannon entropy particularly suitable for likelihood-based inference, where both quantities are naturally expressed in terms of the probability density function. The entropy of a random variable X is a measure of the uncertainty associated with its probability distribution. If X is a continuous random variable with probability density function $f(x;\theta)$, its Shannon differential entropy is defined as follows [25–27]:

$$H(\theta) = -E_\theta[logf(X;\theta)] = -\int_{\chi} f(x;\theta)logf(x;\theta)dx \quad (4)$$

where χ denotes the support of X.

The main reason for selecting Shannon entropy in this study is its direct and parameter-free relationship with the expected log-density, as given in Equation (4). In contrast, Rényi and Tsallis entropies introduce an additional entropy-order parameter, whose specification would create a second tuning dimension alongside the regularization parameter η and complicate both interpretation and optimization. Fisher information, although fundamental for measuring parameter uncertainty and likelihood curvature, serves a conceptually different role because it is based on derivatives of the log-likelihood rather than on the distributional uncertainty represented by the density itself. Shannon entropy therefore provides a parsimonious regularization measure that is directly linked to the log-density, requires no additional entropy-order parameter, and admits closed-form expressions for the lifetime distributions considered in this study. These properties make it particularly suitable for the proposed likelihood-regularization framework.

From a statistical perspective, entropy measures the uncertainty associated with the underlying probability distribution. Larger entropy values correspond to higher uncertainty and more dispersed probabilistic structures, whereas smaller entropy values indicate more concentrated distributions. In reliability and survival analysis, entropy has been widely employed for uncertainty quantification, goodness-of-fit analysis, and information-based inference procedures.

Under progressive Type-II censoring, the observed likelihood is constructed using only a subset of the original lifetime information because a proportion of experimental units is withdrawn before failure. Consequently, the available information about the underlying lifetime distribution is reduced, and the resulting likelihood surface may become increasingly sensitive to random sampling variability. From an information-theoretic viewpoint, this loss of information motivates the incorporation of an additional quantity capable of reflecting the intrinsic uncertainty of the assumed lifetime model.

In the proposed framework, Shannon entropy is not introduced as an additional parameter to be estimated nor as an alternative goodness-of-fit measure. Instead, it is incorporated directly into the likelihood function as an information-theoretic regularization term. Consequently, parameter estimation is driven jointly by the observed data through the likelihood function and by the intrinsic information structure of the assumed probability model through Shannon entropy. This formulation aims to improve the stability of likelihood-based inference under progressive censoring while preserving the probabilistic interpretation of the underlying lifetime model.

It should be emphasized that the proposed regularization differs fundamentally from conventional penalty-based approaches such as ridge or lasso regularization. Classical penalties typically shrink the parameter vector toward a predefined target, whereas the entropy term introduced in this study is determined by the probabilistic structure of the lifetime distribution itself. Consequently, the proposed regularization is inherently distribution-dependent, and its influence is expected to vary according to the analytical properties of the underlying lifetime model.

3. Entropy-Based Likelihood Regularization

3.1. Motivation and Construction of the Regularized Likelihood

Classical likelihood-based inference constitutes the cornerstone of statistical estimation for lifetime models and has been extensively employed in reliability theory, survival analysis, and engineering applications. Under complete observations, maximum likelihood estimation possesses several desirable asymptotic properties, including consistency, asymptotic efficiency, and asymptotic normality. However, these properties may deteriorate in finite samples when the available lifetime information is substantially reduced by progressive Type-II censoring.

As discussed in the previous sections, progressively censored experiments inevitably involve information loss because only a subset of the original failure times contributes directly to the likelihood function. As the censoring proportion increases, the likelihood surface may become flatter, the observed Fisher information may decrease, and the resulting score equations may become increasingly sensitive to sampling variability. Consequently, classical maximum likelihood estimation may exhibit larger variance, reduced numerical stability, and increased finite-sample estimation error.

Motivated by these limitations, we propose an entropy-based likelihood regularization framework that incorporates Shannon entropy directly into the classical log-likelihood function. The main objective is not to modify the probabilistic model itself, but rather to stabilize the estimation process by introducing an additional information-theoretic component that reflects the intrinsic uncertainty of the underlying lifetime distribution.

Let

$$l(\theta) = logL(\theta) \tag{5}$$

denote the classical log-likelihood function corresponding to a progressively Type-II censored sample, where θ is the unknown parameter vector. Furthermore, let

$$H(\theta) = -E[logf(x;\theta)] \tag{6}$$

denote the Shannon entropy associated with the assumed lifetime distribution.

The theoretical motivation for incorporating Shannon entropy into the likelihood can be established through its direct connection with the expected log-density. The Shannon entropy for a parametric lifetime model $f(x;\theta)$ is as in Equation (6). Thus, the entropy term is constructed from the same log-density that generates the likelihood but represents a model-based expectation rather than the realized contribution of the observed sample. This relationship provides a natural information-theoretic link between likelihood inference and entropy regularization. Under progressive Type-II censoring, the likelihood is determined by the observed failures together with the survival contributions of progressively removed units, and its finite-sample curvature may be weakened as the amount of directly observed failure information decreases. Adding $\eta H(\theta)$ therefore introduces a deterministic, distribution-specific correction derived from the underlying probability model itself. Importantly, this term does not introduce external prior information or impose an arbitrary shrinkage target; instead, its gradient modifies the score equation and its Hessian modifies the local curvature of the objective function. Consequently, the proposed criterion can be interpreted as an information-regularized M-estimation objective in which the empirical information supplied by the progressively censored sample is combined with a model-based measure of distributional uncertainty. Since the entropy contribution is of order $O(1)$, meaning that it remains bounded as the sample size increases, whereas the log-likelihood accumulates with the number of observed failures, its influence is primarily finite-sample and becomes asymptotically negligible for fixed η. This interpretation is consistent with the theoretical results in Section 3.4, where the ERMLE is shown to retain the

consistency and first-order asymptotic distribution of the classical MLE while modifying the finite-sample bias–variance and curvature structure.

The proposed entropy-regularized log-likelihood function is defined as

$$l_E(\theta) = l(\theta) + \eta H(\theta) \tag{7}$$

where $\eta \geq 0$ denotes the regularization parameter controlling the contribution of the entropy component. Unlike conventional penalized likelihood methods, the proposed formulation does not shrink the parameter vector toward a predetermined target value. Instead, the entropy contribution is determined entirely by the analytical structure of the assumed lifetime distribution. Consequently, the regularization mechanism is inherently distribution-dependent, and its influence on parameter estimation is expected to vary across different lifetime models.

For fixed η, the proposed formulation is intentionally designed as a finite-sample regularization of the classical likelihood. Dividing Equation (7) by the number of observed failures m yields

$$\frac{1}{m} l_E(\theta) = \frac{1}{m} l(\theta) + \frac{\eta}{m} H(\theta)$$

which has exactly the same maximizer as Equation (7). Thus, on the normalized likelihood scale, the effective contribution of the entropy term decreases with m. This behavior is intentional, since the proposed regularization is designed to improve finite-sample inference while preserving the first-order asymptotic properties of the classical MLE.

The additive formulation in Equation (7) is adopted because it preserves the basic structure of likelihood-based estimation while allowing the entropy contribution to enter the objective function in a transparent and analytically tractable manner. In particular, the additive form yields a regularized score equation in which the classical likelihood score is modified directly by the entropy gradient, $U_E(\theta) = U(\theta) + \eta \nabla H(\theta)$, and similarly modifies the local curvature through the entropy Hessian. This structure facilitates both numerical optimization and the theoretical analysis of consistency, asymptotic normality, bias, and variance developed in Section 3.4. Moreover, when $\eta = 0$, the formulation reduces exactly to the classical MLE, providing a natural nested connection between the regularized and unregularized estimators. Alternative nonlinear or multiplicative combinations of likelihood and entropy could also be constructed; however, they would generally alter the scale and geometry of the likelihood objective and make the resulting estimating equations and asymptotic interpretation less transparent. Therefore, the additive formulation is adopted as a parsimonious and theoretically tractable way of incorporating distributional entropy while retaining the fundamental likelihood structure.

From a statistical perspective, the proposed ERMLE can be viewed as a form of penalized likelihood estimation in which the regularization component is determined by the Shannon entropy of the assumed probability model. Unlike conventional penalized likelihood methods, such as ridge- or lasso-type regularization, the entropy term does not shrink the parameter vector toward a predetermined value or impose sparsity. Instead, it modifies the likelihood according to the intrinsic distributional uncertainty of the model. The criterion also admits a formal connection with Bayesian regularization, since maximizing $l(\theta) + \eta H(\theta)$ is algebraically equivalent to maximum a posteriori estimation under a formal parameter weighting proportional to $\exp\{\eta H(\theta)\}$. However, no prior distribution or external prior information is specified in the proposed framework; the entropy term is entirely determined by the assumed lifetime distribution and is used as a frequentist regularization component. Therefore, ERMLE is most naturally interpreted as a distribution-driven penalized likelihood estimator rather than a Bayesian estimator.

Accordingly, the proposed framework should not be viewed as a replacement for conventional likelihood theory but rather as an information-theoretic extension of it. The classical likelihood remains the primary source of sample information, while Shannon entropy provides a model-derived regularization component whose finite-sample influence is controlled by η. This construction preserves the classical MLE as the special case $\eta = 0$ and, as established in Section 3.4, retains its first-order asymptotic properties for fixed η. Thus, the proposed methodology extends the likelihood framework while maintaining a direct theoretical connection with conventional maximum likelihood inference.

3.2. Properties of the Entropy-Regularized Likelihood

Differentiating the entropy-regularized log-likelihood function with respect to the parameter vector θ yields the modified score equations

$$U_E(\theta) = \frac{\partial l_E(\theta)}{\partial \theta} = U(\theta) + \eta \frac{\partial H(\theta)}{\partial \theta} \tag{8}$$

where

$$U(\theta) = \frac{\partial l(\theta)}{\partial \theta}$$

denotes the classical score vector.

Equation (8) shows that the proposed estimator preserves the original likelihood structure while introducing an additional information-theoretic correction through the entropy gradient. Consequently, parameter estimation is simultaneously influenced by the observed sample information and the intrinsic uncertainty associated with the assumed lifetime distribution.

The entropy-regularized maximum likelihood estimator (ERMLE) is obtained by solving the modified score equations

$$U_E(\theta) = 0$$

Since the entropy function depends explicitly on the underlying probability distribution, the analytical form of the modified score equations varies across lifetime models. Consequently, although the regularization framework remains identical, its statistical influence on parameter estimation is expected to differ according to the structural characteristics of the assumed distribution.

Furthermore, differentiating Equation (9) once more yields the regularized Hessian matrix

$$J_E(\theta) = \frac{\partial^2 l_E(\theta)}{\partial \theta \partial \theta^T} = J(\theta) + \eta \frac{\partial^2 H(\theta)}{\partial \theta \partial \theta^T} \tag{9}$$

where $J(\theta)$ denotes the Hessian matrix of the classical log-likelihood function.

Equation (8) indicates that entropy regularization modifies not only the score equations but also the local curvature of the likelihood surface. Therefore, the optimization process is influenced by both the observed information contained in the censored sample and the information-theoretic structure represented by the entropy term. The regularized observed information matrix is defined as

$$I_E(\theta) = -J_E(\theta).$$

Proposition 1. *Let*

$$I_E(\boldsymbol{\theta}) = -\nabla^2_{\boldsymbol{\theta}} l_E(\boldsymbol{\theta}) = I(\boldsymbol{\theta}) - \eta \nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})$$

where $I(\boldsymbol{\theta}) = -\nabla^2_{\boldsymbol{\theta}} l(\boldsymbol{\theta})$ *denotes the classical observed information matrix and* $H(\boldsymbol{\theta})$ *is the Shannon entropy of the assumed lifetime distribution.*

Assume that there exists a compact neighbourhood of the ERMLE such that $I(\boldsymbol{\theta})$ is positive definite for every $\boldsymbol{\theta} \in$. Furthermore, suppose that

$$\eta \|\nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\|_2 < \lambda_{min}\{I(\boldsymbol{\theta})\},$$

where $\|.\|_2$ denotes the spectral norm and $\lambda_{min}\{.\}$ denotes the smallest eigenvalue.

Then

$$\mathbf{v}^T I_E(\boldsymbol{\theta})\mathbf{v} > 0, \qquad \forall \mathbf{v} \neq \mathbf{0},\ \boldsymbol{\theta} \in$$

Consequently, the entropy-regularized log-likelihood $l_E(\boldsymbol{\theta})$ is strictly concave on , and every stationary point contained in is the unique local maximizer.

Proof. For any nonzero vector $\mathbf{v}$,

$$\mathbf{v}^T I_E(\boldsymbol{\theta})\mathbf{v} = \mathbf{v}^T I(\boldsymbol{\theta})\mathbf{v} - \eta \mathbf{v}^T \nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\mathbf{v}$$

Since $I(\boldsymbol{\theta})$ is positive definite,

$$\mathbf{v}^T I(\boldsymbol{\theta})\mathbf{v} \geq \lambda_{min}\{I(\boldsymbol{\theta})\}\|\mathbf{v}\|^2$$

Moreover, by the definition of the spectral norm,

$$\left|\mathbf{v}^T \nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\mathbf{v}\right| \leq \|\nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\|_2 \|\mathbf{v}\|^2$$

Therefore,

$$\mathbf{v}^T I_E(\boldsymbol{\theta})\mathbf{v} \geq \lambda_{min}\{I(\boldsymbol{\theta})\} - \eta \|\nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\|_2 \|\mathbf{v}\|^2,$$

which is strictly positive under the stated condition. Hence, $I_E(\boldsymbol{\theta})$ is positive definite throughout . Since

$$\nabla^2_{\boldsymbol{\theta}} l_E(\boldsymbol{\theta}) = -I_E(\boldsymbol{\theta})$$

the entropy-regularized log-likelihood is strictly concave on . Consequently, every stationary point in this neighbourhood is a unique local maximizer. □

Furthermore, Proposition 1 provides a sufficient theoretical upper bound for the regularization parameter through the condition

$$\eta < \frac{\lambda_{min}\{I(\boldsymbol{\theta})\}}{\|\nabla^2_{\boldsymbol{\theta}} H(\boldsymbol{\theta})\|_2},$$

which guarantees the local positive definiteness of the entropy-regularized observed information matrix. Although this bound is generally parameter-dependent, it indicates that excessively large values of η may cause the entropy contribution to dominate the likelihood information, leading to increased estimation bias and numerical instability. Consequently, the candidate values of η were restricted to a finite grid covering moderate regularization levels, and the final value was selected using the AMSE criterion.

Remark 1. *Proposition 1 provides a sufficient, but not necessary, condition for the local positive definiteness of the entropy-regularized observed information matrix. Accordingly, the proposed framework does not require the entropy-regularized log-likelihood to be globally concave over the entire parameter space. Instead, the proposition guarantees local strict concavity and local uniqueness of the stationary point within a compact neighbourhood of the ERMLE.*

Furthermore, for any fixed regularization parameter η, the entropy contribution remains of order $O(1)$, whereas the observed information increases with the number of observed failures. Consequently, the proposed entropy regularization primarily affects finite-sample estimation while preserving the first-order asymptotic behavior of the classical maximum likelihood estimator established in Section 3.4.

Remark 2. *The proposed entropy-regularized likelihood framework constitutes a direct extension of the classical maximum likelihood method. In particular, when the regularization parameter satisfies* $\eta = 0$*, the entropy-regularized maximum likelihood estimator coincides exactly with the classical maximum likelihood estimator.*

3.3. Selection of the Regularization Parameter

The performance of the proposed entropy-regularized maximum likelihood estimator depends critically on the choice of the regularization parameter η, which determines the relative contribution of the entropy term to the classical likelihood function. When $\eta = 0$, the proposed estimator reduces to the classical maximum likelihood estimator, whereas larger values of η impose stronger entropy-based regularization. Consequently, the selection of η involves a trade-off between variance reduction and the additional bias introduced by the regularization term.

Rather than selecting a different regularization parameter for each individual simulation scenario, a single distribution-specific value is determined by minimizing the average mean squared error (MSE) over all simulation settings. This strategy provides a stable, reproducible, and practically applicable selection rule while avoiding scenario-specific tuning. Specifically, the optimal regularization parameter is defined as

$$\eta^* = \underset{\eta \in \Omega}{\operatorname{argmin}}\, \overline{MSE}(\eta)$$

where

$$\overline{MSE}(\eta) = \frac{1}{S} \sum_{i=1}^{S} MSE_i(\eta),$$

S denotes the total number of simulation scenarios, and Ω represents the candidate set of regularization parameters.

In this study, the candidate set was chosen as

$$\Omega = \{0,\ 0.2,\ 0.5,\ 1,\ 1.5,\ 2,\ 3,\ 5\}$$

which spans the range from the classical maximum likelihood estimator $\eta = 0$ to increasingly stronger levels of entropy regularization. This grid was selected to investigate both mild and relatively strong regularization effects while maintaining computational efficiency.

For each candidate value of η, Monte Carlo simulations were performed under all combinations of sample size, censoring rate, and censoring scheme. The average MSE obtained across these scenarios was then computed, and the value producing the smallest average MSE was selected as the optimal regularization parameter for the corresponding lifetime distribution.

Since the primary objective of this study is parameter estimation, the regularization parameter is selected in the parameter space rather than on the survival or hazard-function space. For a given simulation scenario $s = (n, m, \boldsymbol{R},\ \boldsymbol{\theta}_0)$, let $\boldsymbol{X}_{\boldsymbol{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ denote the progressively Type-II censored sample generated under the true parameter vector $\boldsymbol{\theta}_0$ and censoring scheme $\boldsymbol{R} = (R_1, R_2, \ldots, R_m)$. The ERMLE obtained for a candidate

regularization parameter η is denoted by $\hat{\boldsymbol{\theta}}_\eta(\boldsymbol{X_R})$. The parameter-based quadratic loss is defined as

$$d(\hat{\boldsymbol{\theta}}_\eta, \boldsymbol{\theta}_0) = \|\hat{\boldsymbol{\theta}}_\eta - \boldsymbol{\theta}_0\|^2 = \sum_{j=1}^{p} (\hat{\theta}_{\eta,j} - \theta_{0,j})^2,$$

where p denotes the number of unknown parameters. Accordingly, for scenario s, the population objective functional is

$$J_s(\eta) = E_{(\boldsymbol{\theta}_0, \boldsymbol{R})} \left[d(\hat{\boldsymbol{\theta}}_\eta, \boldsymbol{\theta}_0) \right],$$

where the expectation is taken with respect to the joint sampling distribution of the progressively Type-II censored order statistics induced by $\boldsymbol{\theta}_0$ and the censoring scheme $\boldsymbol{R}$. Thus, for a p parameter model, $J_s(\eta)$ is equivalent to the sum of the parameter-specific MSEs. For example, for the Lognormal model,

$$J_s(\eta) = MSE(\hat{\mu}_\eta) + MSE(\hat{\sigma}_\eta)$$

In the Monte Carlo experiment with B replications, this quantity is estimated by

$$\hat{J}_s(\eta) = \frac{1}{B} \sum_{b=1}^{B} \sum_{j=1}^{p} (\hat{\theta}_{\eta,j}^{(b)} - \theta_{0,j})^2$$

Finally, to obtain a single distribution-specific regularization parameter, the scenario-specific MSE values are averaged over all S combinations of sample size, censoring rate, and progressive censoring scheme:

$$AMSE(\eta) = \frac{1}{S} \sum_{s=1}^{S} \hat{J}_s(\eta),$$

and the optimal regularization parameter is defined as

$$\eta^* = arg\min_{\eta \in \Omega} AMSE(\eta)$$

The resulting AMSE values for all candidate regularization parameters are reported in Table 1.

Table 1. Average Mean Squared Error for Different Candidate Values of the Regularization Parameter.

Distribution	η	Average MSE	Distribution	η	Average MSE
Exponential	0.0	0.047807	Weibull	0.0	0.170915
	0.2	0.045808		0.2	0.158287
	0.5	0.043777		0.5	0.160664
	1.0	**0.039729**		1.0	0.143665
	1.5	0.041163		1.5	0.137710
	2.0	0.040124		**2.0**	**0.130977**
	3.0	0.041944		3.0	0.522634
	5.0	0.060867		5.0	20.924189
Gamma	0.0	0.565046	Lognormal	0.0	0.054103
	0.2	0.584311		**0.2**	**0.053342**
	0.5	0.517828		0.5	0.054635
	1.0	0.517848		1.0	5.676060
	1.5	**0.470885**		1.5	11.198901
	2.0	0.493310		2.0	24.314393
	3.0	0.703514		3.0	41.093460
	5.0	76.085514		5.0	90.896935

Table 1 summarizes the average MSE values obtained for different candidate regularization parameters. For each lifetime distribution, the parameter value yielding the

smallest average MSE was selected as the optimal regularization parameter and used in all subsequent simulation studies.

Figure 1 shows the variation in the average MSE with respect to the regularization parameter for each lifetime distribution. The minimum point of each curve determines the distribution-specific optimal regularization parameter used in the subsequent simulation study.

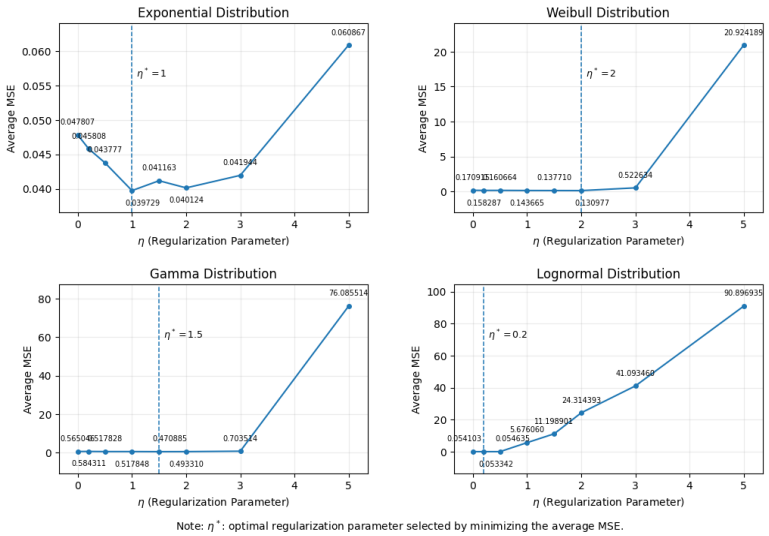


Figure 1. Selection of the optimal regularization parameter based on the average mean squared error.

As the regularization parameter increases, the entropy gradient may gradually dominate the likelihood score, leading to excessive regularization and increased estimation bias. This behavior is reflected by the sharp increase in the AMSE values observed for sufficiently large values of η in some lifetime distributions (e.g., Weibull and Gamma). Therefore, the proposed AMSE-based selection procedure naturally excludes excessively large regularization parameters and identifies a finite-sample compromise between bias reduction and variance stabilization.

The regularization parameter is selected using the AMSE criterion because the primary objective of the proposed framework is accurate parameter estimation rather than model selection or prediction. Since AMSE jointly reflects estimation bias and variance, it provides a direct measure of estimation accuracy. In contrast, information criteria such as AIC or BIC are primarily intended for comparing competing statistical models, whereas cross-validation is prediction-oriented. Because the present study considers a fixed parametric model and seeks an optimal regularization level for parameter estimation, an AMSE-based criterion provides a more natural and interpretable selection procedure.

3.4. Theoretical Properties of the ERMLE

Let $l_m(\theta)$ denote the classical log-likelihood function based on a progressively Type-II censored sample with m observed failures, and let $H(\theta)$ be the Shannon entropy of the assumed lifetime distribution. The entropy-regularized log-likelihood is defined by

$$l_E(\theta) = l_m(\theta) + \eta H(\theta),$$

where $\eta \geq 0$ 0 is the regularization parameter. The corresponding score function is

$$U_E(\theta) = \frac{\partial l_E(\theta)}{\partial \theta} = U(\theta) + \eta H'(\theta)$$

where

$$U(\theta) = \frac{\partial l_m(\theta)}{\partial \theta}$$

Thus, the entropy-regularized maximum likelihood estimator is obtained as the solution of

$$U_E\left(\hat{\theta}_{ERMLE}\right) = 0$$

This formulation shows that the proposed estimator modifies the classical likelihood equation by adding a deterministic entropy-gradient term. The following results summarize the large-sample and finite-sample behavior of the proposed estimator.

Theorem 1. *Assume that the standard regularity conditions for likelihood-based inference under progressive Type-II censoring hold, and suppose that* $H(\theta)$ *is continuously differentiable in a neighbourhood of the true parameter value* θ_0. *If* η *is fixed, then*

$$\hat{\theta}_{ERMLE} \xrightarrow{p} \theta_0$$

Proof. Consider the normalized regularized score evaluated at θ_0:

$$\frac{1}{m}U_E(\theta_0) = \frac{1}{m}U(\theta_0) + \frac{\eta}{m}H'(\theta_0)$$

Under the standard regularity conditions for progressively censored likelihoods,

$$\frac{1}{m}U(\theta_0) \xrightarrow{p} 0$$

Since $H'(\theta_0)$ does not depend on m, and η is fixed,

$$\frac{\eta}{m}H'(\theta_0) \xrightarrow{p} 0$$

Therefore,

$$\frac{1}{m}U_E(\theta_0) \xrightarrow{p} 0$$

Moreover,

$$\frac{1}{m}l_E(\theta) = \frac{1}{m}l_m(\theta) + \frac{\eta}{m}H(\theta)$$

and the second term vanishes uniformly on compact subsets of the parameter space. Hence, the normalized entropy-regularized likelihood has the same limiting objective function as the classical likelihood. By the standard consistency argument for likelihood-type estimators, the maximizer of $l_E(\theta)$ converges in probability to θ_0. This completes the proof. □

Theorem 2. *Under the assumptions of Theorem 1, and assuming that the classical Fisher information matrix*

$$I_c(\theta_0) = -\lim_{m\to\infty}\frac{1}{m}\frac{\partial^2 l_m(\theta_0)}{\partial\theta\partial\theta^T}$$

is positive definite, the ERMLE satisfies

$$\sqrt{m}\left(\hat{\theta}_{ERMLE} - \theta_0\right) \xrightarrow{d} N\left(0,\ I_c^{-1}(\theta_0)\right)$$

Proof. Using a Taylor expansion of $U_E\left(\hat{\theta}_{ERMLE}\right)$ around θ_0, we obtain

$$0 = U_E\left(\hat{\theta}_{ERMLE}\right) = U_E(\theta_0) + J_E(\theta^*)\left(\hat{\theta}_{ERMLE} - \theta_0\right),$$

where θ^* lies between $\hat{\theta}_{ERMLE}$ and θ_0, and

$$J_E(\theta) = \frac{\partial^2 l_E(\theta)}{\partial\theta\partial\theta^T}$$

Thus,

$$\sqrt{m}\left(\hat{\theta}_{ERMLE} - \theta_0\right) = -\left[\frac{1}{m}J_E(\theta^*)\right]^{-1}\frac{1}{\sqrt{m}}U_E(\theta_0)$$

Now,

$$\frac{1}{\sqrt{m}}U_E(\theta_0) = \frac{1}{\sqrt{m}}U(\theta_0) + \frac{\eta}{\sqrt{m}}H'(\theta_0)$$

The second term converges to zero because η is fixed. By the central limit theorem for progressively censored likelihood scores,

$$\frac{1}{\sqrt{m}}U(\theta_0) \xrightarrow{d} N(0,\ I_c(\theta_0))$$

Hence,

$$\frac{1}{\sqrt{m}}U_E(\theta_0) \xrightarrow{d} N(0,\ I_c(\theta_0))$$

Moreover,

$$\frac{1}{m}J_E(\theta^*) = \frac{1}{m}J(\theta^*) + \frac{\eta}{m}H''(\theta^*)$$

Since η is fixed,

$$\frac{\eta}{m}H''(\theta^*) \to 0$$

and therefore

$$\frac{1}{m}J_E(\theta^*) \xrightarrow{p} -I_c(\theta_0)$$

□

Proposition 2. *Suppose that $H(\theta)$ is twice continuously differentiable. Since*

$$l_E(\theta) = l_m(\theta) + \eta H(\theta),$$

the Hessian matrix of the entropy-regularized log-likelihood is

$$J_E(\theta) = J(\theta) + \eta H''(\theta),$$

where

$$J(\theta) = \frac{\partial^2 l_m(\theta)}{\partial\theta\partial\theta^T}$$

Consequently, the corresponding modified observed Fisher information matrix is

$$I_{c,E}(\theta) = -J_E(\theta) = I_c(\theta) - \eta H''(\theta)$$

where

$$I_c(\theta) = -J(\theta)$$

This result shows that entropy regularization changes the local curvature of the likelihood surface through the Hessian matrix of the entropy function. Therefore, the finite-sample effect of ERMLE depends on both the likelihood curvature and the analytical structure of $H(\theta)$.

Proposition 3. *For fixed* $\eta > 0$, *the ERMLE has the following first-order finite-sample bias approximation:*

$$Bias(\hat{\theta}_{ERMLE}) \approx \frac{\eta}{m} A_E^{-1}(\theta_0) H'(\theta_0)$$

where

$$A_E(\theta_0) = I_c(\theta_0) - \frac{\eta}{m} H''(\theta_0)$$

Thus,

$$Bias(\hat{\theta}_{ERMLE}) = O\left(m^{-1}\right)$$

Derivation 1. *Using the first-order Taylor expansion of the regularized score,*

$$0 \approx U_E(\theta_0) + J_E(\theta_0)(\hat{\theta}_{ERMLE} - \theta_0)$$

Hence,

$$\hat{\theta}_{ERMLE} - \theta_0 \approx -J_E^{-1}(\theta_0) U_E(\theta_0)$$

Since

$$E[U(\theta_0)] = 0$$

and

$$U_E(\theta_0) = U(\theta_0) + \eta H'(\theta_0)$$

we have

$$E[U_E(\theta_0)] = \eta H'(\theta_0)$$

Moreover, defining

$$A_E(\theta_0) = I_c(\theta_0) - \frac{\eta}{m} H''(\theta_0)$$

gives

$$J_E(\theta_0) \approx -m\left[I_c(\theta_0) - \frac{\eta}{m} H''(\theta_0)\right] = -m A_E(\theta_0)$$

and hence

$$J_E^{-1}(\theta_0) \approx \frac{1}{m} A_E^{-1}(\theta_0).$$

Therefore,

$$Bias(\hat{\theta}_{ERMLE}) = \frac{\eta}{m} A_E^{-1}(\theta_0) H'(\theta_0)$$

This shows that entropy regularization may introduce a small finite-sample bias, but this bias vanishes asymptotically.

Proposition 4. *The finite-sample mean squared error of the ERMLE can be approximated by*

$$MSE(\hat{\theta}_{ERMLE}) \approx Var(\hat{\theta}_{ERMLE}) + Bias(\hat{\theta}_{ERMLE}) Bias(\hat{\theta}_{ERMLE})^T$$

Using the first-order approximations above,

$$Var(\hat{\theta}_{ERMLE}) \approx \frac{1}{m} A_E^{-1}(\theta_0) I_c(\theta_0) A_E^{-1}(\theta_0)$$

and

$$Bias(\hat{\theta}_{ERMLE}) Bias(\hat{\theta}_{ERMLE})^T \approx \frac{\eta^2}{m^2} A_E^{-1}(\theta_0) H'(\theta_0) H'(\theta_0) A_E^{-1}(\theta_0)$$

Therefore,

$$MSE(\hat{\theta}_{ERMLE}) \approx \frac{1}{m} A_E^{-1}(\theta_0) I_c(\theta_0) A_E^{-1}(\theta_0) + \frac{\eta^2}{m^2} A_E^{-1}(\theta_0) H'(\theta_0) H'(\theta_0) A_E^{-1}(\theta_0)$$

This decomposition shows that entropy regularization affects estimation performance through two competing mechanisms. The first term represents the variance component, which depends on the modified curvature matrix $A_E(\theta_0)$ *. The second term represents the squared bias introduced by the entropy gradient. Therefore, ERMLE improves upon the classical MLE when the reduction in variance dominates the additional squared bias caused by the entropy term.*

Remark 3. *The above results clarify why entropy regularization is mainly a finite-sample improvement mechanism. For fixed* η *, the entropy term becomes asymptotically negligible after normalization by m, so the ERMLE shares the same asymptotic distribution as the classical MLE. However, in finite samples, especially under progressive Type-II censoring where the effective information may be reduced, the entropy term can alter the likelihood curvature and the bias–variance balance. This explains why the simulation results show improvements under moderate and heavy censoring, while the magnitude of improvement depends on the underlying lifetime distribution and the selected regularization level.*

4. Simulation Study

4.1. Simulation Design

A comprehensive Monte Carlo simulation study was conducted to evaluate the finite-sample performance of the proposed entropy-regularized maximum likelihood estimator under progressive Type-II censoring. The simulation design was constructed to systematically investigate the combined effects of sample size, censoring severity, censoring scheme, and entropy regularization on parameter estimation across different lifetime distributions.

Three sample sizes, n = 30, 50, and 100, were considered to represent small, moderate, and relatively large experiments commonly encountered in reliability studies. For each sample size, three numbers of observed failures were selected, namely m = (24,18,12) for n = 30, m = (40,30,20) for n = 50, and m = (80,60,40) for n = 100. These settings correspond approximately to censoring rates of 20%, 40%, and 60%, respectively.

The simulation design was applied identically to all lifetime distributions considered in this study. Distribution-specific parameter settings are introduced separately in the corresponding subsections to facilitate a fair comparison of the proposed methodology across different probabilistic models.

Estimation performance was evaluated using Bias, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). For one-parameter lifetime distributions, additional summary measures, including the Average Estimate (AE) and the percentage reduction in MSE relative to the classical maximum likelihood estimator, are also reported. For two-parameter distributions, the simulation results are presented in a compact form using Bias and MSE for each parameter in order to avoid unnecessary table expansion, since AE can be directly obtained from the reported Bias and the true parameter value, while RMSE is simply the square root of MSE.

To investigate the influence of different censoring mechanisms, three progressive Type-II censoring schemes were considered for every (n,m) combination: Early, Uniform, and Late censoring. In the Early scheme, all removable units were withdrawn immediately after the first observed failure. In the Late scheme, all removals were postponed until the last observed failure. In the Uniform scheme, the withdrawn units were distributed as evenly as possible throughout the experiment. The corresponding progressive censoring vectors used in the simulation study are summarized in Table 2.

Table 2. Progressive Type-II censoring schemes considered in the simulation study.

Scheme	Definition
Scheme I (Early)	Removals are concentrated at the beginning of the experiment.
Scheme II (Uniform)	Removals are distributed as evenly as possible throughout the experiment.
Scheme III (Late)	Removals are concentrated toward the end of the experiment.

Finally, the primary objective of the simulation study is to identify the most effective degree of entropy regularization under different progressive censoring conditions. Therefore, the main simulation tables summarize, for each scenario, the regularization level yielding the minimum MSE, thereby highlighting the best-performing entropy regularization for each combination of sample size, censoring severity, and censoring scheme.

Three representative progressive Type-II censoring schemes were employed throughout the simulation study. These schemes were selected to represent different patterns of information loss under progressive censoring while satisfying the constraint $\sum_{i=1}^{m} R_i = n - m$.

4.2. Finite-Sample Performance of the Proposed Estimator

The finite-sample performance of the proposed entropy-regularized maximum likelihood estimator was evaluated by comparing it with the classical maximum likelihood estimator under the simulation scenarios described in the previous subsection. For each lifetime distribution, parameter estimates were obtained using both estimation methods, and their statistical performance was assessed over 10,000 Monte Carlo replications.

The comparison was carried out using the Average Estimate (AE), Bias, Variance, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). These measures provide complementary information regarding estimation accuracy, variability, and overall efficiency. In particular, MSE and RMSE jointly reflect the combined effects of bias and variance and therefore serve as the primary criteria for evaluating the finite-sample performance of the competing estimators.

To quantify the practical benefit of entropy regularization, the percentage reduction in MSE relative to the classical maximum likelihood estimator was additionally calculated. Positive values indicate that the proposed estimator achieves lower estimation error than the classical MLE, whereas negative values indicate that entropy regularization is not advantageous under the corresponding experimental setting.

The simulation results are presented separately for each lifetime distribution in order to facilitate a detailed comparison of the estimation behavior under different distributional structures. Particular emphasis is placed on investigating how the effectiveness of entropy regularization changes with the sample size, censoring level, and regularization parameter.

The Exponential, Weibull, Gamma, and Lognormal distributions were selected because they represent four of the most widely used parametric lifetime models in reliability and survival analysis. These distributions cover a broad range of hazard rate behaviors and distributional characteristics. The Exponential distribution serves as the simplest constant-hazard benchmark, the Weibull distribution accommodates increasing, decreasing, and constant hazard rates, the Gamma distribution provides additional flexibility for modeling heterogeneous lifetime data, and the Lognormal distribution represents log-location-scale lifetime models commonly encountered in biomedical and engineering applications. Evaluating the proposed entropy-regularized likelihood framework across these complementary models enables a comprehensive assessment of its robustness under different probabilistic structures and censoring conditions.

4.3. Entropy-Regularized Inference for the Exponential Distribution

To illustrate the proposed entropy-regularized likelihood framework, we first consider the Exponential distribution, which represents one of the most fundamental lifetime models in reliability and survival analysis. Owing to its simple analytical structure, the Exponential distribution provides a convenient benchmark for demonstrating the effect of entropy regularization on likelihood-based inference. Let X be a lifetime random variable following an Exponential distribution with rate parameter $\lambda \geq 0$. The corresponding probability density function and survival function are given by

$$f(x;\lambda) = \lambda e^{-\lambda x},\ x > 0$$

and

$$S(x;\lambda) = e^{-\lambda x}$$

Under a progressive Type-II censoring scheme, the likelihood function is

$$L(\ \lambda) = C\ \lambda^m \exp[-\lambda \sum_{i=1}^{m}(R_i + 1)x_i]$$

where C denotes the combinatorial constant associated with the censoring scheme. For notational convenience, define

$$T_c = \sum_{i=1}^{m}(R_i + 1)x_i$$

Ignoring the terms that do not depend on λ, the corresponding log-likelihood function becomes

$$l(\lambda) = mlog\lambda - \lambda T_c$$

Differentiating with respect to λ gives the classical score equation

$$U(\lambda) = \frac{m}{\lambda} - T_c,$$

from which the classical maximum likelihood estimator is obtained as

$$\hat{\lambda}_{MLE} = \frac{m}{T_c}$$

For the Exponential distribution, the Shannon entropy is

$$H(\lambda) = 1 - log\lambda$$

Substituting this expression into Equation (8) yields

$$l_E(\lambda) = mlog\lambda - \lambda T_c + \eta(1 - log\lambda)$$

is obtained from Equation (7).

After collecting the logarithmic terms, the entropy-regularized log-likelihood function becomes

$$l_E(\lambda) = (m - \eta)log\lambda - \lambda T_c + \eta$$

Differentiating the entropy-regularized log-likelihood with respect to λ gives the modified score function

$$U_E(\lambda) = \frac{\partial l_E(\lambda)}{\partial \lambda} = \frac{m - \eta}{\lambda} - T_c$$

https://doi.org/10.3390/sym18081366

The entropy-regularized estimator is obtained by solving

$$U_E(\lambda) = 0$$

which leads to the explicit estimator

$$\hat{\lambda}_{ERMLE} = \frac{m - \eta}{T_c}$$

It is immediately seen that

$$\hat{\lambda}_{ERMLE} = \hat{\lambda}_{MLE}$$

when $\eta = 0$. Therefore, the proposed estimator constitutes a direct extension of the classical maximum likelihood estimator, with the entropy component acting as a regularization mechanism whose strength is controlled by the parameter η.

Remark 4. *For the Exponential distribution, the entropy-regularized estimator admits a closed-form expression. This analytical solution provides valuable insight into the effect of entropy regularization and serves as a benchmark for the more complex lifetime distributions considered in the following sections, where numerical optimization is generally required.*

Exponential Distribution Simulation Results

A Monte Carlo simulation study was conducted to evaluate the finite-sample performance of the proposed entropy-regularized estimator for the Exponential distribution under progressive Type-II censoring. The simulation design described in Section 4.1 was adopted without modification. The true parameter value was fixed at $\lambda = 1$. Following the distribution-specific regularization parameter selection procedure described in Section 3.3, the optimal regularization parameter was determined as $\eta^* = 1$ by minimizing the average Total MSE over all simulation scenarios. This single optimal value was subsequently employed throughout all Exponential distribution simulations. The estimation results obtained under different sample sizes, censoring rates, and progressive censoring schemes are summarized in Table 3.

Table 3. Simulation Results for the Exponential Distribution Using the Selected Regularization Parameter ($\eta^* = 1$).

				MLE			ERMLE			Improvement (%)
n	**m**	**CR**	**Scheme**	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**	
30	12	0.60	I	1.0878	0.0878	0.1168	0.9875	−0.0125	0.0862	26.22
			II	1.0741	0.0741	0.1122	1.0069	0.0069	0.0930	17.11
			III	1.0987	0.0987	0.1340	1.0006	0.0006	0.0991	26.05
	18	0.40	I	1.0747	0.0747	0.0915	0.9935	−0.0065	0.0563	38.44
			II	1.0229	0.0229	0.0611	1.0301	0.0301	0.0682	−11.68
			III	1.0478	0.0478	0.0598	0.9796	−0.0204	0.0549	8.09
	24	0.20	I	1.0412	0.0412	0.0544	1.0037	0.0037	0.0434	20.28
			II	1.0388	0.0388	0.0467	1.0150	0.0150	0.0437	6.54
			III	1.0413	0.0413	0.0440	0.9927	−0.0073	0.0405	7.85
50	20	0.60	I	1.0483	0.0483	0.0628	0.9914	−0.0086	0.0541	13.77
			II	1.0472	0.0472	0.0648	1.0050	0.0050	0.0471	27.44
			III	1.0500	0.0500	0.0683	0.9848	−0.0152	0.0430	37.03
	30	0.40	I	1.0393	0.0393	0.0380	0.9980	−0.0020	0.0377	0.70
			II	1.0383	0.0383	0.0403	1.0015	0.0015	0.0320	20.66
			III	1.0267	0.0267	0.0342	0.9926	−0.0074	0.0353	−3.24
	40	0.20	I	1.0275	0.0275	0.0293	0.9941	−0.0059	0.0253	13.85
			II	1.0278	0.0278	0.0286	0.9868	−0.0132	0.0230	19.67
			III	1.0312	0.0312	0.0258	1.0013	0.0013	0.0256	0.88

Table 3. *Cont.*

				MLE			ERMLE			Improvement (%)
n	**m**	**CR**	**Scheme**	**AE**	**Bias**	**MSE**	**AE**	**Bias**	**MSE**	
100	40	0.60	I	1.0246	0.0246	0.0273	0.9976	−0.0024	0.0267	2.38
			II	1.0354	0.0354	0.0305	1.0053	0.0053	0.0251	17.55
			III	1.0038	0.0038	0.0254	0.9938	−0.0062	0.0240	5.51
	60	0.40	I	1.0265	0.0265	0.0186	0.9992	−0.0008	0.0156	16.05
			II	1.0184	0.0184	0.0191	0.9987	−0.0013	0.0171	10.61
			III	1.0294	0.0294	0.0189	1.0084	0.0084	0.0183	3.43
	80	0.20	I	0.9991	−0.0009	0.0122	1.0040	0.0040	0.0128	−4.61
			II	1.0096	0.0096	0.0127	1.0000	0.0000	0.0154	3.69
			III	1.0108	0.0108	0.0133	1.0011	0.0011	0.0930	6.42

Table 3 summarizes the average mean squared error (MSE) obtained for different candidate values of the regularization parameter under the Exponential distribution. The results show that the average MSE decreases steadily as the regularization parameter increases from $\eta = 0$ to $\eta = 1$, indicating that moderate entropy regularization substantially improves the estimation accuracy compared with the classical maximum likelihood estimator. The minimum average MSE is attained at $\eta = 1$, after which the average MSE begins to increase, suggesting that excessive regularization introduces additional estimation error. Therefore, $\eta = 1$ was selected as the optimal distribution-specific regularization parameter and was adopted in all subsequent simulation studies for the Exponential distribution.

4.4. Entropy-Regularized Inference for the Weibull Distribution

Assume that the lifetime random variable X follows the Weibull distribution with shape parameter $\beta > 0$ and scale parameter $\theta > 0$. The corresponding density and survival functions are given by

$$f(x;\beta,\ \theta) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right],\ x > 0$$

and

$$S(x;\beta,\ \theta) = exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right]$$

Under a progressive Type-II censoring scheme, the likelihood function becomes

$$L(\beta,\ \theta) = C\left(\frac{\beta}{\theta^{\beta}}\right)^{m} \prod_{i=1}^{m} x_i^{\beta-1} exp\left[-\sum_{i=1}^{m}(R_i+1)\left(\frac{x_i}{\theta}\right)^{\beta}\right]$$

Define

$$T_W = \sum_{i=1}^{m}(R_i+1){x_i}^{\beta}$$

Using this notation

$$\sum_{i=1}^{m}(R_i+1)\left(\frac{x_i}{\theta}\right)^{\beta} = \frac{T_W}{\theta^{\beta}}$$

Therefore, ignoring the terms that do not depend on the unknown parameters, the classical log-likelihood function becomes

$$l(\beta,\ \theta) = mlog\beta - m\beta log\theta + (\beta-1)\sum_{i=1}^{m} logx_i - \frac{T_W}{\theta^{\beta}}$$

The corresponding Shannon entropy of the Weibull distribution is given by

$$H(\beta,\ \theta) = 1 + \gamma\left(1 - \frac{1}{\beta}\right) + log\left(\frac{\theta}{\beta}\right)$$

where γ denotes the Euler–Mascheroni constant.

Using the proposed entropy-based regularization framework,

$$l_E(\beta,\ \theta) = l(\beta,\ \theta) + \eta H(\beta,\ \theta)$$

$$l_E(\beta,\ \theta) = mlog\beta - m\beta log\theta + (\beta - 1)\sum_{i=1}^{m} logx_i - \frac{T_W}{\theta^\beta} + \left[1 + \gamma\left(1 - \frac{1}{\beta}\right) + log\left(\frac{\theta}{\beta}\right)\right]$$

The entropy-regularized score equations are obtained from

$$\frac{\partial l_E(\beta,\ \theta)}{\partial \beta} = 0$$

and

$$\frac{\partial l_E(\beta,\ \theta)}{\partial \theta} = 0$$

From the score equation with respect to θ, the entropy-regularized estimator of the scale parameter is obtained as

$$\hat{\theta}_{ERMLE} = \left[\frac{\beta T_W}{m\beta - \eta}\right]^{1/\beta}$$

Unlike the Exponential case, the entropy-regularized estimator of the shape parameter β does not admit a closed-form expression and therefore must be obtained numerically using iterative optimization techniques.

Weibull Distribution Simulation Results

A Monte Carlo simulation study was conducted to evaluate the finite-sample performance of the proposed entropy-regularized estimator for the Weibull distribution under progressive Type-II censoring. The simulation design described in Section 4.1 was adopted without modification. The true parameter values were fixed at $\beta = 2$ and $\theta = 1$. Following the distribution-specific regularization parameter selection procedure described in Section 3.3, the optimal regularization parameter was determined as $\eta^* = 2$ by minimizing the average Total MSE over all simulation scenarios. This single optimal value was subsequently employed throughout all Weibull distribution simulations. The estimation results obtained under different sample sizes, censoring rates, and progressive censoring schemes are summarized in Table 4.

Table 4 shows that the proposed ERMLE consistently reduces the overall estimation error relative to the classical MLE. The improvement is more pronounced under heavier censoring and smaller sample sizes, reaching approximately 28% in the most favorable scenario.

Although the bias of the Weibull shape parameter changes from positive under the MLE to slightly negative under the ERMLE in some scenarios, particularly at CR = 0.20, this sign reversal does not imply a change in the estimated failure-rate regime. The true shape parameter used in the simulation is $\beta = 2$, and even the largest negative ERMLE bias observed in Table 4 is only -0.0784, corresponding to an average estimate of approximately 1.9216, which remains well above the critical value $\beta = 1$. Therefore, the increasing-hazard interpretation of the simulated Weibull model is preserved. Nevertheless, when the true shape parameter is close to unity, regularization-induced bias could potentially affect the classification of the failure-rate regime and should therefore be interpreted with caution.

Moreover, the effect of the bias change on Weibull lifetime quantiles is not uniform across quantile levels. Since $t_p = \theta[-\log(1-p)]^{1/\beta}$, its sensitivity to the shape parameter is

$$\frac{\partial t_p}{\partial \beta} = -\frac{t_p}{\beta^2} log[-\log(1-p)]$$

Table 4. Simulation Results for the Weibull Distribution Using the Selected Regularization Parameter ($\eta^* = 2$).

				MLE				ERMLE				Imp. (%)
n	**m**	**CR**	**Scheme**	**Bias (β)**	**MSE (β)**	**Bias (θ)**	**MSE (θ)**	**Bias (β)**	**MSE (β)**	**Bias (θ)**	**MSE (θ)**	
30	12	0.60	I	0.1288	0.1518	−0.0037	0.0109	0.0699	0.1272	0.0127	0.0111	14.99
			II	0.1486	0.1703	−0.0009	0.0090	0.0729	0.1384	0.0177	0.0096	17.48
			III	0.1397	0.1645	−0.0104	0.0105	0.0435	0.1294	0.0110	0.0111	19.70
	18	0.40	I	0.1319	0.1816	−0.0154	0.0129	0.0538	0.1473	0.0081	0.0128	17.68
			II	0.1669	0.2184	−0.0122	0.0150	0.0434	0.1666	0.0221	0.0169	21.40
			III	0.2231	0.3037	−0.0161	0.0168	0.0363	0.2116	0.0313	0.0212	27.35
	24	0.20	I	0.1883	0.2548	−0.0171	0.0199	0.0712	0.1896	0.0226	0.0212	23.25
			II	0.2394	0.3741	−0.0314	0.0211	−0.0072	0.2576	0.0556	0.0321	26.70
			III	0.3342	0.6162	−0.0450	0.0296	−0.0784	0.3794	0.1199	0.0847	28.14
50	20	0.60	I	0.0775	0.0720	−0.0028	0.0068	0.0445	0.0641	0.0067	0.0068	9.95
			II	0.0567	0.0751	−0.0045	0.0064	0.0156	0.0677	0.0067	0.0065	8.98
			III	0.0769	0.0981	−0.0008	0.0061	0.0224	0.0860	0.0118	0.0063	11.36
	30	0.40	I	0.0612	0.0856	−0.0046	0.0094	0.0186	0.0764	0.0093	0.0095	9.55
			II	0.0865	0.1032	−0.0111	0.0083	0.0177	0.0885	0.0084	0.0087	12.83
			III	0.0951	0.1405	−0.0125	0.0091	−0.0084	0.1176	0.0147	0.0102	14.53
	40	0.20	I	0.0804	0.1362	−0.0162	0.0126	0.0181	0.1172	0.0065	0.0125	12.84
			II	0.1835	0.2206	−0.0242	0.0134	0.0459	0.1656	0.0202	0.0157	22.51
			III	0.1778	0.2615	−0.0233	0.0203	−0.0461	0.1943	0.0568	0.0325	19.54
100	40	0.60	I	0.0302	0.0298	−0.0004	0.0034	0.0145	0.0282	0.0044	0.0034	4.66
			II	0.0424	0.0358	−0.0017	0.0032	0.0226	0.0335	0.0037	0.0032	5.87
			III	0.0436	0.0446	−0.0042	0.0033	0.0171	0.0414	0.0018	0.0034	6.60
	60	0.40	I	0.0497	0.0478	0.0002	0.0045	0.0288	0.0445	0.0067	0.0045	6.40
			II	0.0307	0.0416	−0.0022	0.0042	−0.0015	0.0390	0.0073	0.0044	5.31
			III	0.0553	0.0589	−0.0054	0.0045	0.0052	0.0528	0.0074	0.0047	9.18
	80	0.20	I	0.0219	0.0546	−0.0100	0.0072	−0.0077	0.0513	0.0006	0.0071	5.53
			II	0.0871	0.0782	−0.0134	0.0081	0.0238	0.0670	0.0073	0.0086	12.45
			III	0.0875	0.1120	−0.0093	0.0096	−0.0177	0.0958	0.0263	0.0119	11.46

Hence, holding θ fixed, a downward shift in β tends to decrease lifetime quantiles corresponding to $p < 1 - e^{-1} \approx 0.632$, whereas it tends to increase quantiles corresponding to $p > 1 - e^{-1}$. Since $\frac{\partial t_p}{\partial \theta} = \frac{t_p}{\theta} > 0$, the actual quantile effect depends jointly on the biases of both β and θ. Therefore, the observed sign reversal in the bias of β does not by itself imply a systematic distortion of the lifetime quantiles.

4.5. Entropy-Regularized Inference for the Gamma Distribution

Assume that X follows a Gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\theta > 0$. The probability density and survival functions are respectively given by

$$f(x;\alpha,\theta) = \frac{x^{\alpha-1}e^{\frac{-x}{\theta}}}{\Gamma(\alpha)\theta^{\alpha}},\ x > 0$$

and

$$S(x;\alpha,\theta) = \frac{\Gamma\left(\alpha, \frac{x}{\theta}\right)}{\Gamma(\alpha)}$$

where $\Gamma\left(\alpha, \frac{x}{\theta}\right)$ denotes the upper incomplete gamma function.

Under progressive Type-II censoring, the likelihood contribution is constructed from the density of the observed failures and the survival function of the censored removals. Hence, the log-likelihood can be written as

$$l(\alpha,\theta) = \sum_{i=1}^{m} logf(x_i;\alpha,\theta) + \sum_{i=1}^{m} R_i logS(x_i;\alpha,\theta) + logC$$

where C is the combinatorial constant associated with the censoring scheme.

The Shannon entropy of the Gamma distribution is

$$H(\alpha,\theta) = \alpha + log\theta + log\Gamma(\alpha) + (1-\alpha)\varphi(\alpha)$$

where $\varphi(\alpha)$ denotes digamma function.

Accordingly, the entropy-regularized log-likelihood function is obtained as

$$l_E(\alpha,\theta) = l(\alpha,\theta) + \eta[\alpha + log\theta + log\Gamma(\alpha) + (1-\alpha)\varphi(\alpha)]$$

Since the resulting score equations involve the digamma function, the incomplete gamma function, and nonlinear interactions between α and θ, closed-form solutions are not available. Therefore, the entropy-regularized estimates $(\hat{\alpha}_{ERMLE}, \hat{\theta}_{ERMLE})$ are obtained numerically using constrained optimization under $\alpha > 0$ and $\theta > 0$.

Gamma Distribution Simulation Results

A Monte Carlo simulation study was conducted to evaluate the finite-sample performance of the proposed entropy-regularized estimator for the Gamma distribution under progressive Type-II censoring. The simulation design described in Section 4.1 was adopted without modification. The true parameter values were fixed at $\alpha = 2$ and $\theta = 1$, representing a moderately skewed lifetime distribution commonly used in simulation studies. Following the distribution-specific regularization parameter selection procedure described in Section 3.3, the optimal regularization parameter was determined as $\eta^* = 1.5$ by minimizing the average Total MSE over all simulation scenarios. This single optimal value was subsequently used throughout all Gamma distribution simulations. The corresponding estimation results under different sample sizes, censoring rates, and progressive censoring schemes are presented in Table 5.

Table 5 shows that the proposed ERMLE generally produces smaller MSE values than the classical MLE across most simulation scenarios. The improvement is particularly noticeable under moderate to heavy censoring and smaller sample sizes.

Table 5. Simulation Results for the Gamma Distribution Using the Selected Regularization Parameter $(\eta^* = 1.5)$.

				MLE				ERMLE				Imp. (%)
n	**m**	**CR**	**Scheme**	**Bias (α)**	**MSE (α)**	**Bias θ**	**MSE θ**	**Bias (α)**	**MSE (α)**	**Bias θ**	**MSE θ**	
30	12	0.60	I	0.2974	0.5939	−0.0581	0.0957	0.2176	0.4951	0.0051	0.1057	12.87
			II	0.2446	0.4933	−0.0336	0.0963	0.1539	0.4039	0.0396	0.1126	12.39
			III	0.2381	0.4600	−0.0364	0.0944	0.1357	0.3683	0.0437	0.1124	13.30
	18	0.40	I	0.2510	0.5327	−0.0303	0.1121	0.1522	0.4276	0.0581	0.1379	12.30
			II	0.3384	0.8584	−0.0485	0.1422	0.1917	0.6584	0.0663	0.1857	15.64
			III	0.3994	1.0421	−0.0701	0.1391	0.2108	0.7593	0.0658	0.1871	19.88
	24	0.20	I	0.3223	0.7375	−0.0487	0.1453	0.1746	0.5472	0.0879	0.2016	15.18
			II	0.5447	1.8728	−0.0948	0.1959	0.2666	1.3900	0.1216	0.3292	16.89
			III	0.5979	1.9123	−0.0479	0.3607	0.2112	1.1301	0.2888	0.9042	10.50

Table 5. *Cont.*

				MLE				ERMLE				Imp. (%)
n	**m**	**CR**	**Scheme**	**Bias (α)**	**MSE (α)**	**Bias θ**	**MSE θ**	**Bias (α)**	**MSE (α)**	**Bias θ**	**MSE θ**	
50	20	0.60	I	0.0988	0.1994	−0.0105	0.0514	0.0583	0.1797	0.0280	0.0563	5.90
			II	0.1427	0.2589	−0.0251	0.0609	0.0926	0.2293	0.0176	0.0663	7.57
			III	0.1411	0.2758	−0.0122	0.0735	0.0839	0.2422	0.0354	0.0823	7.10
	30	0.40	I	0.1135	0.2536	−0.0103	0.0663	0.0608	0.2236	0.0413	0.0755	6.50
			II	0.1714	0.2588	−0.0354	0.0712	0.0929	0.2150	0.0302	0.0819	10.03
			III	0.1700	0.3742	−0.0073	0.1012	0.0712	0.3099	0.0748	0.1267	8.16
	40	0.20	I	0.1751	0.3990	−0.0316	0.0957	0.0969	0.3353	0.0453	0.1146	9.06
			II	0.2931	0.5833	−0.0636	0.1329	0.1456	0.4425	0.0574	0.1753	13.74
			III	0.3364	0.7114	−0.0443	0.1899	0.1326	0.5038	0.1253	0.2968	11.18
100	40	0.60	I	0.0514	0.0958	0.0004	0.0287	0.0321	0.0909	0.0194	0.0303	2.65
			II	0.1014	0.0971	−0.0380	0.0261	0.0773	0.0895	−0.0175	0.0261	6.17
			III	0.0344	0.0875	−0.0002	0.0287	0.0081	0.0829	0.0232	0.0307	2.24
	60	0.40	I	0.1144	0.1553	−0.0246	0.0386	0.0881	0.1436	−0.0001	0.0400	5.31
			II	0.0697	0.1313	−0.0096	0.0436	0.0333	0.1211	0.0228	0.0473	3.72
			III	0.0814	0.1432	−0.0109	0.0467	0.0349	0.1297	0.0281	0.0515	4.58
	80	0.20	I	0.1003	0.1700	−0.0152	0.0485	0.0634	0.1541	0.0218	0.0528	5.31
			II	0.1265	0.1864	−0.0262	0.0605	0.0602	0.1609	0.0320	0.0696	6.64
			III	0.1273	0.2136	−0.0091	0.0956	0.0366	0.1797	0.0705	0.1201	3.04

4.6. Entropy-Regularized Inference for the Lognormal Distribution

Assume that X follows a Lognormal distribution with shape parameter $\mu \in R$ and scale parameter $\sigma > 0$. The probability density and survival functions are given by

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[\frac{-(logx-\mu)^2}{2\sigma^2}\right], x > 0$$

and

$$S(x;\mu,\sigma) = 1 - \Phi\left(\frac{logx-\mu}{\sigma}\right)$$

where $\Phi(.)$ denotes the standard normal distribution function.

Under progressive Type-II censoring, the log-likelihood function is expressed as

$$l(\mu,\sigma) = \sum_{i=1}^{m} logf(x_i;\mu,\sigma) + \sum_{i=1}^{m} R_i logS(x_i;\mu,\sigma) + logC$$

where C denotes the combinatorial constant associated with the censoring scheme.

The Shannon entropy of the Lognormal distribution is

$$H(\mu,\sigma) = \mu + \frac{1}{2} + \log\left(\sigma\sqrt{2\pi}\right)$$

Accordingly, the entropy-regularized log-likelihood function is defined as

$$l_E(\mu,\sigma) = l(\mu,\sigma) + \eta\ [\mu + \frac{1}{2} + \log\left(\sigma\sqrt{2\pi}\right)]$$

Since the resulting score equations are nonlinear and do not admit closed-form solutions, the entropy-regularized estimates $(\hat{\mu}_{ERMLE}, \hat{\sigma}_{ERMLE})$ are obtained numerically using constrained optimization.

Lognormal Distribution Simulation Results

A Monte Carlo simulation study was conducted to evaluate the finite-sample performance of the proposed entropy-regularized estimator for the Lognormal distribution un-

der progressive Type-II censoring. The simulation design described in Section 4.1 was adopted without modification. The true parameter values were fixed at $\mu = 0$ and $\sigma = 1$. Following the distribution-specific regularization parameter selection procedure described in Section 3.3, the optimal regularization parameter was determined as $\eta^* = 0.2$ by minimizing the average Total MSE over all simulation scenarios. This single optimal value was subsequently employed throughout all Lognormal distribution simulations. The estimation results obtained under different sample sizes, censoring rates, and progressive censoring schemes are summarized in Table 6.

Table 6. Simulation Results for the Lognormal Distribution Using the Selected Regularization Parameter ($\eta^* = 0.2$).

				MLE				ERMLE				Imp. (%)
n	**m**	**CR**	**Scheme**	**Bias (μ)**	**MSE (μ)**	**Bias (σ)**	**MSE (σ)**	**Bias (μ)**	**MSE (μ)**	**Bias (σ)**	**MSE (σ)**	
30	12	0.60	I	−0.0095	0.0376	−0.0277	0.0211	−0.0014	0.0375	−0.0236	0.0211	0.17
			II	0.0109	0.0346	−0.0282	0.0214	0.0188	0.0349	−0.0235	0.0214	−0.54
			III	−0.0064	0.0374	−0.0283	0.0230	0.0012	0.0374	−0.0231	0.0230	−0.13
	18	0.40	I	−0.0287	0.0557	−0.0374	0.0229	−0.0179	0.0555	−0.0320	0.0228	0.49
			II	−0.0324	0.0425	−0.0407	0.0283	−0.0220	0.0423	−0.0333	0.0283	0.32
			III	−0.0043	0.0364	−0.0458	0.0363	0.0062	0.0367	−0.0371	0.0364	−0.60
	24	0.20	I	0.0000	0.0760	−0.0345	0.0340	0.0169	0.0771	−0.0261	0.0342	−1.17
			II	−0.0308	0.0664	−0.0469	0.0452	−0.0131	0.0675	−0.0336	0.0458	−1.50
			III	−0.0673	0.0648	−0.0964	0.0568	−0.0490	0.0647	−0.0802	0.0563	0.44
50	20	0.60	I	−0.0062	0.0236	−0.0060	0.0124	−0.0011	0.0236	−0.0035	0.0124	−0.01
			II	0.0001	0.0235	−0.0193	0.0130	0.0048	0.0235	−0.0165	0.0130	−0.09
			III	−0.0030	0.0203	−0.0265	0.0147	0.0016	0.0203	−0.0234	0.0147	0.06
	30	0.40	I	0.0035	0.0326	−0.0265	0.0147	0.0101	0.0328	−0.0233	0.0146	−0.24
			II	−0.0112	0.0287	−0.0304	0.0182	−0.0048	0.0287	−0.0260	0.0182	0.06
			III	−0.0123	0.0289	−0.0202	0.0208	−0.0058	0.0290	−0.0148	0.0209	−0.36
	40	0.20	I	−0.0157	0.0467	−0.0303	0.0205	−0.0057	0.0467	−0.0255	0.0204	0.00
			II	−0.0236	0.0401	−0.0397	0.0261	−0.0131	0.0402	−0.0318	0.0261	−0.20
			III	−0.0322	0.0398	−0.0408	0.0305	−0.0204	0.0399	−0.0305	0.0307	−0.34
100	40	0.60	I	0.0016	0.0148	−0.0009	0.0061	0.0042	0.0149	0.0004	0.0061	−0.16
			II	−0.0078	0.0101	−0.0105	0.0058	−0.0054	0.0101	−0.0091	0.0058	0.24
			III	0.0077	0.0101	−0.0125	0.0066	0.0100	0.0102	−0.0109	0.0066	−0.21
	60	0.40	I	0.0005	0.0176	−0.0123	0.0073	0.0038	0.0176	−0.0107	0.0073	−0.01
			II	−0.0086	0.0135	−0.0092	0.0081	−0.0053	0.0135	−0.0070	0.0081	−0.03
			III	−0.0083	0.0111	−0.0147	0.0110	−0.0051	0.0111	−0.0120	0.0110	0.02
	80	0.20	I	−0.0134	0.0220	−0.0120	0.0102	−0.0083	0.0220	−0.0096	0.0102	0.19
			II	−0.0085	0.0203	−0.0127	0.0135	−0.0031	0.0203	−0.0087	0.0135	−0.42
			III	−0.0148	0.0208	−0.0255	0.0173	−0.0088	0.0209	−0.0204	0.0173	−0.21

Table 6 shows that entropy regularization provides virtually no improvement for the Lognormal distribution. This behavior can be understood from the differential structure of the Lognormal entropy. Since

$$H(\mu, \sigma) = \mu + \frac{1}{2} + \log\left(\sigma\sqrt{2\pi}\right),$$

its gradient and Hessian are

$$\nabla H(\mu, \sigma) = \begin{pmatrix} 1 \\ 1/\sigma \end{pmatrix}, \quad \nabla^2 H(\mu, \sigma) = \begin{pmatrix} 0 & 0 \\ 0 & -1/\sigma^2 \end{pmatrix}$$

Accordingly, the entropy-regularized score vector is

$$U_E(\mu, \sigma) = U(\mu, \sigma) + \eta \begin{pmatrix} 1 \\ 1/\sigma \end{pmatrix}$$

Thus, entropy regularization shifts the score equations of both parameters, including an additive perturbation of η in the μ-equation. However, the entropy Hessian contains neither curvature in the μ-direction nor a $\mu - \sigma$ cross-curvature term. Writing the classical observed information matrix as

$$I(\mu,) = \begin{pmatrix} I_{\mu\mu} & I_{\mu\sigma} \\ I_{\mu\sigma} & I_{\sigma\sigma} \end{pmatrix}$$

the entropy-regularized observed information matrix becomes

$$I_E(\mu,\sigma) = I(\mu,\sigma) - \eta\nabla^2 H(\mu,\sigma) = \begin{pmatrix} I_{\mu\mu} & I_{\mu\sigma} \\ I_{\mu\sigma} & I_{\sigma\sigma} + \frac{\eta}{\sigma^2} \end{pmatrix}$$

Hence, the regularization leaves both $I_{\mu\mu}$ and the cross-information term $I_{\mu\sigma}$ unchanged and directly modifies only the curvature associated with σ. The eigenvalues of $I_E(\mu,\sigma)$ are

$$\lambda_{\pm} = \frac{1}{2}\left[I_{\mu\mu} + I_{\sigma\sigma} + \frac{\eta}{\sigma^2} \pm \sqrt{\left(I_{\mu\mu} - I_{\sigma\sigma} - \frac{\eta}{\sigma^2} \right)^2 + 4I_{\mu\sigma}^2} \right].$$

This spectral representation shows that entropy regularization does not provide a direct curvature adjustment in the location direction; rather, it modifies the information geometry through the scale direction while leaving the original likelihood-induced cross-dependence unchanged. Therefore, the two estimation equations are not generally fully decoupled but the entropy contribution itself introduces no additional cross-coupling between μ and σ. In the present simulation setting, where $(\mu,\sigma) = (0,1)$, the entropy gradient reduces to $(1,1)^T$, consistent with the perturbation structure noted above. The resulting combination of a score shift in μ and a curvature modification only in the σ-direction explains why ERMLE yields only negligible net MSE improvement for the Lognormal model.

4.7. Overall Performance of the Proposed Estimator

To provide an overall assessment of the proposed entropy-regularized estimation approach, the simulation results obtained for all lifetime distributions and progressive Type-II censoring scenarios are summarized in Figure 2. Rather than presenting the individual censoring schemes separately, the reported values represent the average relative reduction in MSE across the Early, Uniform, and Late censoring schemes using the best-performing regularization parameter identified for each simulation scenario. This summary visualization facilitates a direct comparison of the overall effectiveness of the proposed method across different sample sizes, censoring levels, and lifetime distributions.

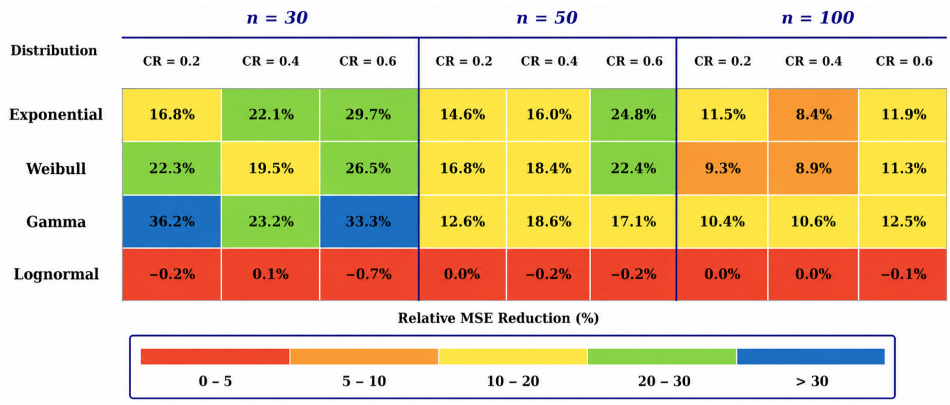


Figure 2. Average relative MSE reduction (%) achieved by ERMLE under different progressive Type-II censoring scenarios.

Figure 2 provides an overall summary of the simulation results by presenting the average relative reduction in MSE achieved by the entropy-regularized maximum likelihood estimator (ERMLE) across the Early, Uniform, and Late progressive censoring schemes. Overall, the greatest improvements are observed under heavier censoring CR = 0.6 and smaller sample sizes n = 30, where information loss is most pronounced. As the sample size increases, the relative benefit of entropy regularization gradually decreases, although ERMLE continues to provide competitive performance in most scenarios. Among the four lifetime distributions, the Gamma distribution exhibits the largest average improvements, followed by the Weibull and Exponential distributions. In contrast, the Lognormal distribution shows almost no systematic improvement, with average MSE reductions remaining close to zero across all simulation settings. This behavior is consistent with the theoretical results developed in Section 3, where the entropy function of the Lognormal distribution was shown to possess negligible curvature with respect to the location parameter, limiting the effectiveness of entropy regularization. Overall, these findings demonstrate that the benefit of entropy regularization depends strongly on the underlying lifetime distribution and is most pronounced in challenging estimation settings characterized by limited sample information and heavier censoring.

5. Real Data Analysis

To demonstrate the practical applicability of the proposed entropy-regularized likelihood estimator, two real lifetime datasets representing different application areas are analyzed in this section. For each dataset, the proposed ERMLE is compared with the classical maximum likelihood estimator under the corresponding lifetime distribution.

Unlike the simulation study, where the performance of different regularization levels was systematically evaluated, the real-data analyses employ the distribution-specific optimal regularization parameters determined in Section 3.3. Specifically, the optimal value of the regularization parameter was obtained by minimizing the average Total MSE over all simulation scenarios for each lifetime distribution. Accordingly, the Exponential, Weibull, Gamma, and Lognormal models were fitted using their corresponding optimal regularization parameters $\eta = 1,\ 2,\ 1.5,\ 0.20$, respectively. This strategy ensures methodological consistency between the simulation experiments and the real-data applications while avoiding arbitrary selection of the regularization parameter.

5.1. Real Data Analysis 1 (Engineering Reliability Dataset)

The first application considers the well-known deep-groove ball bearing failure-time dataset reported by Lieblein and Zelen [28], which has been widely used as a benchmark in engineering reliability studies. The dataset consists of 23 complete lifetime observations (measured in millions of revolutions). To emulate practical reliability experiments with incomplete observations, a progressive Type-II censoring scheme with n = 23, m = 15, and CR = 0.3478 was imposed using the censoring configuration R = (0,0,1,0,0,1,0,1,0,1,0,1,0,1,2). The proposed ERMLE and the classical MLE were then applied under the Exponential, Weibull, Gamma, and Lognormal models using the distribution-specific optimal regularization parameters described in Section 5.

Table 7 shows that the Weibull distribution provides the best overall representation of the engineering reliability data under both the MLE and ERMLE according to all three goodness-of-fit criteria. The corresponding p-values for the Weibull, Gamma, and Lognormal models exceed the 5% significance level, indicating that these distributions cannot be rejected as plausible lifetime models for the censored bearing data, whereas the Exponential model is rejected by all three tests. This result suggests that the bearing failure process cannot

be adequately described by a constant hazard-rate assumption and instead requires a more flexible lifetime model capable of accommodating age-dependent failure behaviour.

Table 7. Goodness-of-Fit Statistics for MLE and ERMLE under Progressive Type-II Censoring.

Distribution	MLE			ERMLE		
	KS	**AD**	**CvM**	**KS**	**AD**	**CvM**
Exponential	0.4239 (0.0091)	2.9922 (0.0282)	0.5819 (0.0235)	0.4488 (0.0047)	3.1940 (0.0224)	0.6363 (0.0169)
Weibull	0.2820 (0.1840)	1.7158 (0.1330)	0.3171 (0.1204)	0.2975 (0.1406)	1.7778 (0.1228)	0.3390 (0.1045)
Gamma	0.3025 (0.1284)	1.6372 (0.1473)	0.2926 (0.1414)	0.3268 (0.0812)	1.8395 (0.1135)	0.3446 (0.1008)
Lognormal	0.3204 (0.0919)	1.6745 (0.1403)	0.2952 (0.1391)	0.3242 (0.0854)	1.7096 (0.1341)	0.3041 (0.1311)

Although entropy regularization produces only small changes in the goodness-of-fit statistics, this is consistent with the objective of the proposed methodology. The proposed ERMLE is designed to improve the stability of parameter estimation under progressive Type-II censoring while preserving the underlying probabilistic structure of the fitted model. Consequently, the almost identical goodness-of-fit results demonstrate that entropy regularization does not distort the fitted lifetime distribution but instead provides a regularized estimation procedure that retains essentially the same agreement with the observed failure data.

From a reliability perspective, this finding is particularly important because engineering decisions such as reliability assessment, maintenance scheduling, warranty analysis, and remaining useful life prediction depend primarily on accurate parameter estimation rather than on marginal changes in goodness-of-fit statistics. Therefore, the proposed ERMLE offers a practical advantage by stabilizing the estimation process while maintaining the reliability characteristics of the fitted lifetime model.

Table 8 shows that the bootstrap estimates obtained from the ERMLE remain very close to those produced by the classical MLE for all four lifetime distributions. However, the principal contribution of the proposed method is not to substantially alter the point estimates, but rather to provide a more stable estimation procedure under progressive Type-II censoring. The bootstrap standard errors and the corresponding 95% confidence intervals indicate that entropy regularization preserves the overall estimation precision while avoiding noticeable deterioration in estimator variability.

Table 8. Bootstrap-Based Parameter Estimates and 95% Confidence Intervals for MLE and ERMLE under Progressive Type-II Censoring.

Distribution	Parameter	MLE			ERMLE		
		Estimate	**Boot SE**	**95% CI**	**Estimate**	**Boot SE**	**95% CI**
Exponential	λ	0.0125	0.0049	(0.0100, 0.0292)	0.0116	0.0041	(0.0088, 0.0234)
Weibull	β	3.6266	1.2432	(3.0475, 7.8059)	3.3139	0.9738	(2.4456, 6.3240)
	θ	63.9548	4.1408	(50.9310, 66.522)	64.9894	4.4892	(50.3925, 67.719)
Gamma	α	7.0984	5.8183	(5.3228, 25.9240)	6.3963	5.0741	(3.9861, 21.5735)
	θ	8.4019	2.3762	(1.9256, 10.8896)	9.5212	3.0874	(2.5207, 14.4796)
Lognormal	μ	4.0341	0.0893	(3.7790, 4.1294)	4.0370	0.0970	(3.7566, 4.1339)
	σ	0.4251	0.0703	(0.2109, 0.4769)	0.4286	0.0744	(0.2186, 0.5035)

From a practical reliability standpoint, these results imply that the proposed ERMLE can be employed without sacrificing the interpretability of the estimated lifetime model. The

resulting parameter estimates lead to essentially the same reliability characteristics as the classical MLE while benefiting from the increased numerical stability demonstrated in the simulation study. Consequently, the real-data analysis confirms that the proposed entropy regularization constitutes a robust alternative for parameter estimation in progressively censored reliability experiments, particularly in situations where numerical instability or limited failure information may adversely affect conventional likelihood-based estimation.

Figure 3 compares the empirical Kaplan–Meier survival curves with the survival functions estimated by the classical MLE and the proposed ERMLE under the four candidate lifetime distributions. Overall, both estimation methods provide very similar fitted survival curves, indicating that entropy regularization preserves the overall distributional fit while stabilizing parameter estimation.

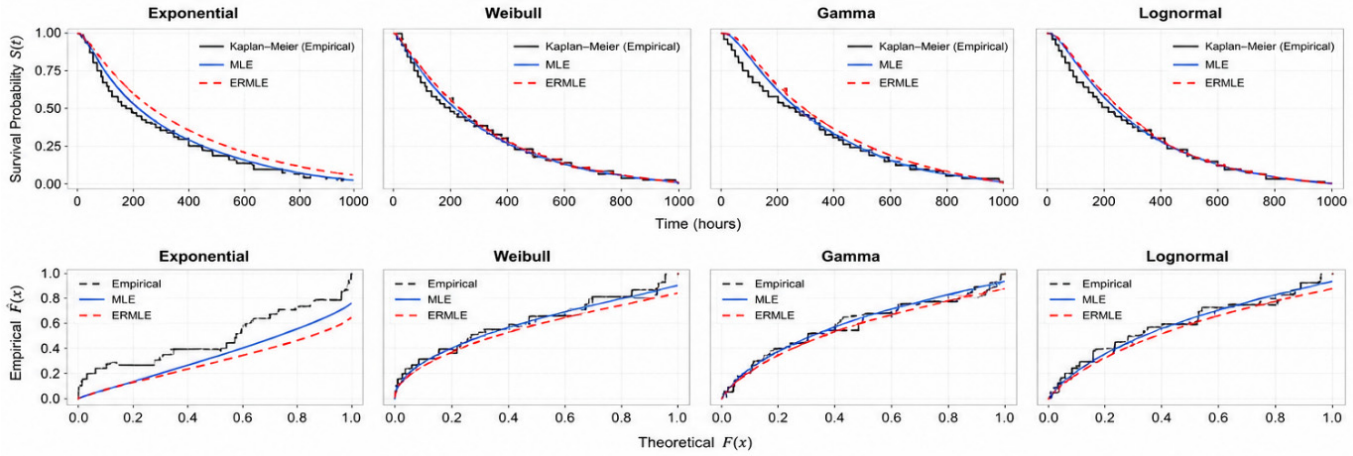


Figure 3. Survival curve comparison for the engineering reliability dataset.

Among the candidate models, the Weibull distribution exhibits the closest agreement with the empirical survival curve over the entire time range, which is consistent with the goodness-of-fit statistics reported in Table 7. The Exponential model shows the largest deviation from the empirical curve, whereas the Gamma and Lognormal distributions provide intermediate levels of agreement. The differences between the MLE and ERMLE curves are generally small, reflecting that the proposed entropy regularization modifies the parameter estimates only moderately while maintaining an adequate representation of the observed lifetime behavior.

5.2. Real Data Analysis 2 (Bladder Cancer Remission Dataset)

The second application considers the bladder cancer remission dataset, consisting of 128 remission times [29]. This dataset is a commonly used benchmark in medical survival analysis and provides a suitable setting for evaluating the proposed ERMLE in biomedical lifetime data. To mimic incomplete observations under a progressive Type-II censoring mechanism, a censoring scheme with n = 128, m = 80, and CR = 0.375 was imposed. A uniform progressive censoring scheme was used, where the remaining $n - m$ = 48 units were progressively removed across the observed failure times. The Exponential, Weibull, Gamma, and Lognormal models were fitted using the distribution-specific optimal regularization parameters determined in Section 3.3.

Table 9 shows that the Weibull and Gamma distributions provide the best overall representation of the bladder cancer remission data under both the MLE and ERMLE, as reflected by their comparatively smaller goodness-of-fit statistics. Nevertheless, the corresponding p-values indicate that all four candidate distributions are rejected at the 5% significance level. This suggests that the remission-time data exhibit a level of heterogeneity

that cannot be fully captured by any of the simple parametric lifetime models considered here. Despite this limitation, the Weibull and Gamma distributions remain the most appropriate approximating models among the candidates and therefore provide the most reliable basis for subsequent statistical inference.

Table 9. Goodness-of-Fit Statistics for MLE and ERMLE under Progressive Type-II Censoring.

Distribution	MLE			ERMLE		
	KS (*p*)	AD (*p*)	CvM (*p*)	KS (*p*)	AD (*p*)	CvM (*p*)
Exponential	0.2942 (<0.0001)	7.1327 (0.0003)	1.2807 (0.0005)	0.2987 (<0.0001)	7.3604 (0.0002)	1.3351 (0.0004)
Weibull	0.1776 (0.0129)	5.7509 (0.0013)	0.9445 (0.0032)	0.1846 (0.0086)	5.9906 (0.0010)	0.9984 (0.0024)
Gamma	0.2171 (0.0011)	5.7342 (0.0013)	0.9391 (0.0033)	0.2227 (0.0007)	6.0144 (0.0010)	1.0013 (0.0023)
Lognormal	0.2832 (<0.0001)	7.1152 (0.0003)	1.2051 (0.0008)	0.2844 (<0.0001)	7.1772 (0.0003)	1.2196 (0.0007)

Entropy regularization produces only marginal changes in the goodness-of-fit statistics and their associated *p*-values, indicating that the proposed regularization does not materially alter the fitted lifetime distributions. Instead, its principal contribution lies in stabilizing parameter estimation while preserving essentially the same level of agreement between the fitted models and the observed remission data.

From a survival-analysis perspective, these findings imply that the proposed ERMLE can be employed without compromising the clinical interpretation of the fitted lifetime model. In particular, estimates of remission-time distributions, survival probabilities, and other reliability-related characteristics remain practically unchanged, while the regularized estimation procedure offers improved numerical robustness under progressively censored observations.

Table 10 demonstrates that the bootstrap parameter estimates obtained from the ERMLE are highly consistent with those produced by the classical MLE across all four lifetime distributions. The bootstrap standard errors and the corresponding 95% confidence intervals are also very similar, confirming that entropy regularization preserves the overall estimation precision while maintaining stable parameter estimates.

Table 10. Bootstrap-Based Parameter Estimates and 95% Confidence Intervals for MLE and ERMLE under Progressive Type-II Censoring.

Distribution	Parameter	MLE			ERMLE		
		Estimate	Boot SE	95% CI	Estimate	Boot SE	95% CI
Exponential	λ	0.1543	0.0288	(0.1855, 0.2960)	0.1524	0.0270	(0.1849, 0.2874)
Weibull	β	1.8081	0.2739	(2.0169, 3.0817)	1.7845	0.2546	(1.9606, 2.9301)
	θ	5.8596	0.2823	(4.0472, 5.1297)	5.9109	0.3040	(4.0143, 5.1571)
Gamma	α	2.1134	0.5735	(2.3678, 4.4805)	2.0853	0.5866	(2.3675, 4.6567)
	θ	2.6163	0.2510	(0.8916, 1.8394)	2.6790	0.2570	(0.8524, 1.8316)
Lognormal	μ	1.5300	0.0880	(1.0646, 1.4040)	1.5322	0.0899	(1.0781, 1.4310)
	σ	0.9432	0.0640	(0.5666, 0.8223)	0.9448	0.0634	(0.5867, 0.8305)

From a practical standpoint, these results indicate that the proposed regularization framework can be incorporated into censored survival analyses without altering the interpretation of the estimated lifetime distributions. Rather than changing the underlying survival characteristics, entropy regularization provides a more stable estimation procedure while retaining essentially the same clinical conclusions regarding remission-time behaviour. Together with the simulation results, the real-data application demonstrates that

the proposed ERMLE constitutes a reliable and practically applicable alternative to the classical maximum likelihood estimator for progressively censored biomedical lifetime data.

Figure 4 compares the empirical survival and distribution functions with the fitted MLE and ERMLE curves under the four candidate lifetime distributions. The Weibull and Gamma models show the closest agreement with both the Kaplan–Meier survival estimates and the empirical distribution function, confirming the goodness-of-fit results reported in Table 9.

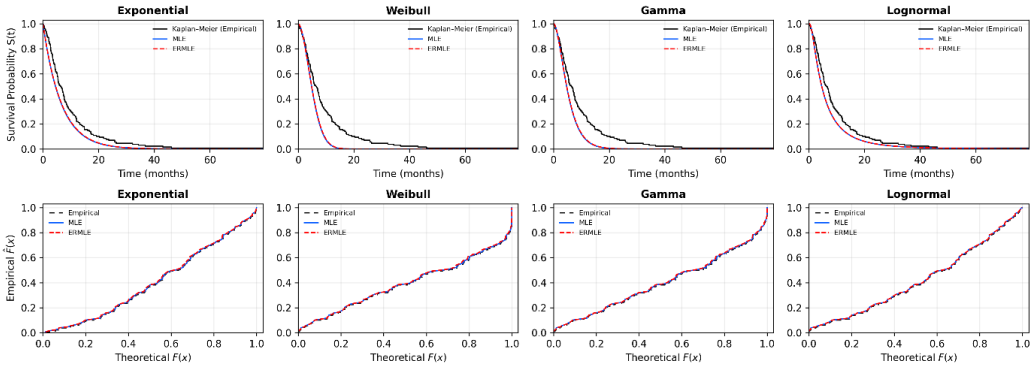


Figure 4. Survival curve comparison for the bladder cancer remission data.

The Exponential distribution exhibits the largest discrepancies, whereas the Lognormal model provides an intermediate fit. The MLE and ERMLE curves are nearly indistinguishable across all distributions, indicating that entropy regularization preserves the overall model fit while providing a more stable estimation framework.

6. Conclusions

This study developed an entropy-regularized likelihood inference framework for lifetime distributions under progressive Type-II censoring by incorporating Shannon entropy directly into the classical likelihood function. Unlike conventional penalty-based regularization methods, the proposed approach exploits the intrinsic information structure of the assumed lifetime distribution, resulting in a distribution-dependent regularization mechanism.

A comprehensive Monte Carlo simulation study demonstrated that the effectiveness of entropy regularization depends strongly on the underlying lifetime model. The largest improvements were obtained for the Gamma distribution, followed by the Weibull and Exponential distributions, particularly under moderate and heavy censoring and small sample sizes where information loss is most severe. In contrast, the Lognormal distribution exhibited almost no systematic improvement. This behavior is fully consistent with the theoretical results presented in this study, showing that the entropy function of the Lognormal model contributes only limited additional curvature to the likelihood surface. Consequently, the reduction in estimator variance is largely offset by the additional bias introduced through regularization, resulting in only negligible changes in overall estimation accuracy.

Rather than adopting a common regularization level for all models, distribution-specific regularization parameters were selected by minimizing the average mean squared error across all simulation scenarios. This strategy provides a practical and reproducible procedure for selecting the regularization parameter while ensuring methodological consistency between the simulation study and the real-data analyses.

The proposed methodology was further evaluated using two progressively Type-II censored real datasets from engineering reliability and biomedical survival analysis. The engineering reliability application showed that the Weibull distribution provided the best overall fit, whereas the Weibull and Gamma models yielded the most satisfactory per-

formance for the bladder cancer remission data. Across both applications, the entropy-regularized estimator produced parameter estimates, bootstrap confidence intervals, and fitted survival curves that were highly consistent with those obtained by the classical maximum likelihood estimator. The goodness-of-fit statistics and their corresponding p-values further demonstrated that entropy regularization preserves the probabilistic structure of the fitted lifetime models while providing a stable estimation framework under progressive censoring.

Overall, the results indicate that entropy regularization should not be regarded as a universally superior alternative to classical maximum likelihood estimation. Instead, its effectiveness depends jointly on the characteristics of the underlying lifetime distribution and the degree of information loss induced by censoring. Therefore, the proposed ERMLE should be viewed as a practical extension of the classical likelihood framework that is particularly beneficial for finite-sample problems involving moderate or heavy progressive censoring, where estimation stability is often more important than marginal improvements in model fit.

Future research may extend the proposed framework to multi-parameter lifetime models, accelerated life testing, competing risks, hybrid censoring schemes, Bayesian entropy regularization, and high-dimensional reliability problems. The proposed information-theoretic perspective also provides a promising basis for developing adaptive regularization procedures in which the regularization parameter is selected automatically according to the characteristics of the observed data.

Author Contributions: Methodology, A.B. and M.D.; Software, A.B., M.D. and G.G.; Validation, A.B., G.G.; Investigation, A.B. and G.G.; Resources, A.B. and M.D.; Writing—Original Draft, A.B.; Supervision, A.B. and M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Scientific Research Projects Coordination Unit of Firat University. Project Number: FF.26.47.

Data Availability Statement: The data used in this study are publicly available from the references cited in the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Balakrishnan, N.; Aggarwala, R. *Progressive Censoring: Theory, Methods and Applications*; Birkhäuser: Boston, MA, USA, 2000.
2. Balakrishnan, N.; Sandhu, R.A. A simple simulation algorithm for generating progressive Type-II censored samples. *Am. Stat.* **1995**, *49*, 229–230. [CrossRef]
3. Balakrishnan, N.; Cramer, E.; Kamps, U.; Schenk, N. Progressive Type-II censored order statistics from exponential distributions. *Stat. J. Theor. Appl. Stat.* **2001**, *35*, 537–556. [CrossRef]
4. Ng, H.K.T.; Kundu, D.; Chan, P.S. Statistical analysis of exponential lifetimes under an adaptive Type-II progressive censoring scheme. *Nav. Res. Logist.* **2009**, *56*, 687–698. [CrossRef]
5. Gong, Q.; Yin, B. Statistical inference of entropy functions of generalized inverse exponential model under progressive Type-II censoring test. *PLoS ONE* **2024**, *19*, e0311129. [CrossRef] [PubMed]
6. Ahmed, S.M.; Ismail, G.M. Statistical inference based on censored data of entropy for the Lomax distribution. *Eur. J. Pure Appl. Math.* **2025**, *18*, 5737. [CrossRef]
7. Wang, X.; Gui, W. Bayesian estimation of entropy for Burr Type XII distribution under progressive Type-II censored data. *Mathematics* **2021**, *9*, 313. [CrossRef]
8. Cho, Y.; Sun, H.; Lee, K. Estimating the entropy of a Weibull distribution under generalized progressive hybrid censoring. *Entropy* **2015**, *17*, 102–122. [CrossRef]
9. Lee, K. Estimation of entropy of the inverse Weibull distribution under generalized progressive hybrid censored data. *J. Korean Data Inf. Sci. Soc.* **2017**, *28*, 659–668.
10. Hassan, A.S.; Zaky, A.N. Estimation of entropy for inverse Weibull distribution under multiple censored data. *J. Taibah Univ. Sci.* **2019**, *13*, 331–337. [CrossRef]

11. Kohansal, A.; Rezakhah, S. Testing exponentiality based on Rényi entropy with progressively Type-II censored data. *arXiv* **2013**, arXiv:1303.5536. [CrossRef]
12. Alyami, S.A.; Hassan, A.S.; Elbatal, I.; Elgarhy, M.; El-Saeed, A.R. Bayesian and non-Bayesian estimation of dynamic cumulative residual Tsallis entropy for moment exponential distribution under progressive censored Type-II. *Open Phys.* **2023**, *21*, 20220264. [CrossRef]
13. Pandey, A.; Kaushik, A.; Singh, S.K.; Singh, U. On the estimation problems for exponentiated exponential distribution under generalized progressive hybrid censoring. *Aust. J. Stat.* **2021**, *50*, 24–40. [CrossRef]
14. Teimouri, M. On the parameter estimation of the generalized exponential distribution under progressive Type-I interval censoring scheme. *arXiv* **2018**, arXiv:1811.06857.
15. Dutta, S.; Kayal, S. Estimation of parameters of the logistic exponential distribution under progressive Type-I hybrid censored sample. *Qual. Technol. Quant. Manag.* **2022**, *19*, 234–258. [CrossRef]
16. Wei, S.; He, G.; Mo, D. Bayesian parameter estimation and failure time prediction under double Type-II hybrid censoring for the generalized inverse exponential distribution. *Alex. Eng. J.* **2025**, *128*, 747–766. [CrossRef]
17. Calcagnì, A. A maximum entropy procedure to solve likelihood equations. *Entropy* **2019**, *21*, 596. [CrossRef] [PubMed]
18. Machado, R.J.; van den Hout, A.; Marra, G. Penalised maximum likelihood estimation in multi-state models for interval-censored data. *Comput. Stat. Data Anal.* **2021**, *153*, 107057. [CrossRef]
19. Kim, E.; MacEachern, S.N.; Peruggia, M. Regularized exponentially tilted empirical likelihood for Bayesian inference. *arXiv* **2023**, arXiv:2312.17015.
20. Alshenawy, R. Progressive Type-II censoring schemes for extended odd Weibull exponential distribution. *Mathematics* **2020**, *8*, 1679. [CrossRef]
21. Kim, C.; Jung, J.; Chung, Y. Bayesian estimation for the exponentiated Weibull model under Type II progressive censoring. *Stat. Pap.* **2011**, *52*, 53–70. [CrossRef]
22. Wang, B.X.; Yu, K.; Sheng, Z. New inference for constant-stress accelerated life tests with Weibull distribution and progressive Type-II censoring. *IEEE Trans. Reliab.* **2014**, *63*, 807–815. [CrossRef]
23. Ma, Y.; Gui, W. Pivotal inference for the inverse Rayleigh distribution based on progressively Type-II censored samples. *J. Appl. Stat.* **2019**, *46*, 771–797. [CrossRef]
24. Sert, S.; Abusaif, I.A.; Akgenç, E.; Karakaya, K.; Kuş, C. Estimation and prediction for the half-normal distribution based on progressively Type-II censored samples. *REVSTAT* **2024**, *22*, 239–258.
25. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
26. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2006.
27. Kullback, S. *Information Theory and Statistics*; Dover Publications: Mineola, NY, USA, 1968.
28. Lieblein, J.; Zelen, M. Statistical investigation of the fatigue life of deep-groove ball bearings. *J. Res. Natl. Bur. Stand.* **1956**, *57*, 273–316. [CrossRef]
29. Lee, E.T.; Wang, J.W. *Statistical Methods for Survival Data Analysis*, 3rd ed.; Wiley-Interscience: Hoboken, NJ, USA, 2003.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.