Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Exact Finite-Sample Inference for the Exponential Proportional Hazards Model via the Empirical Characteristic Function

Nargiza Nurmukhamedova

n.nurmuhamedova@nuu.uz

National University of Uzbekistan named after Mirzo Ulugbek

Guzal Abdujalilova

National University of Uzbekistan named after Mirzo Ulugbek

Nargiza Boltaeva

National University of Uzbekistan named after Mirzo Ulugbek

Gulhayo Xalilova

National University of Uzbekistan named after Mirzo Ulugbek

Dilsuz Khamraeva

ZARMED University, Bukhara Campus

Article

Keywords: empirical characteristic function, proportional hazards model, exact bias, exact variance, exact confidence intervals, random censoring, Cram´er–Rao bound

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10570666/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Exact Finite-Sample Inference for the Exponential Proportional Hazards Model via the Empirical Characteristic Function

Nargiza Nurmukhamedova[1*], Guzal Abdujalilova[1], Nargiza Boltaeva[1], Gulhayo Xalilova[1], Dilsuz Khamraeva[2]

[1*]National University of Uzbekistan named after Mirzo Ulugbek, Universitetskaya str. 4, Tashkent, 100174, Uzbekistan.
[2]ZARMED University, Qayum Murtazoyev str. 13A, Bukhara, 200117, Uzbekistan.

*Corresponding author(s). E-mail(s): n.nurmuhamedova@nuu.uz;
Contributing authors: rasulova2406@gmail.com;
boltayevan27@gmail.com; xalilovagulhayo1@gmail.com;
mohlaroyim261291@gmail.com;

Abstract

We study estimation of the baseline rate $\boldsymbol{\theta}$ in the exponential proportional hazards model (PHM) under random right censoring, the censoring intensity $\boldsymbol{\beta}$ being an unknown nuisance parameter, using the empirical characteristic function (ECF). We obtain *exact finite-sample formulas* for the bias and variance of the ECF-based estimator $\hat{\boldsymbol{\theta}} = \bar{\boldsymbol{\delta}}/\bar{\boldsymbol{Z}}$: for every $\boldsymbol{n} \geq \mathbf{3}$ the bias equals $\boldsymbol{\theta}/(\boldsymbol{n}-\mathbf{1})$ and the variance admits a closed form, both from the exact gamma distribution of the summed censored observations. Being exact for each fixed $\boldsymbol{n}$ rather than to first asymptotic order, they yield exact confidence intervals, an exact bias correction, and an explicit sample-size rule. A single ECF/score-weighted estimating equation reproduces the maximum-likelihood equations on any support. Two extensions go beyond the closed-form regime: for the Weibull PHM with unknown shape we give a consistent, $\sqrt{\boldsymbol{n}}$-asymptotically normal moment system nearly as efficient as maximum likelihood; and for dependent censoring we obtain the exact probability limit of $\hat{\boldsymbol{\theta}}$ under arbitrary dependence. The estimator's coincidence with maximum likelihood and its asymptotic normality are known; the finite-sample formulas, their consequences, and the extensions are

new. A closed-form Fisher information computation shows $\hat{\boldsymbol{\theta}}$ attains the parametric Cramér–Rao bound for the model with $\boldsymbol{\beta}$ unknown. All formulas are validated by extensive simulation and a real-data application.

Keywords: empirical characteristic function, proportional hazards model, exact bias, exact variance, exact confidence intervals, random censoring, Cramér–Rao bound

MSC Classification: 62N02 , 62F10 , 62F12 , 62E15

1 Introduction

Estimation of a parameter from a parametric family is a classical problem, usually approached through maximum likelihood or the method of moments. Each has limitations: the likelihood equations need not admit a closed-form solution (Barnett 1966), and the method of moments is unavailable for families without finite moments (Zolotarev 1986). The method of characteristic functions provides an alternative that applies uniformly across such families, replacing the model characteristic function by its sample analogue, the empirical characteristic function (ECF)

$$\hat{\varphi}_n(t) = \frac{1}{n}\sum_{k=1}^{n} e^{itX_k}, \qquad t \in \mathbb{R}. \tag{1}$$

An estimating equation is obtained by matching $\hat{\varphi}_n$ to the model characteristic function $\varphi_\theta(t) = \mathbb{E}_\theta[e^{itX}]$ against a suitable weight; we call the resulting procedure and its estimator *ECF-based*. The theory originates with Heathcote (1977), Bryant and Paulson (1979) and Feuerverger and McDunnough (1981a), and its extension to censored data with Csörgő (1981a) and Abdushukurov and Kadirov (2001); Abdushukurov (2009). In survival analysis the ECF has been used mainly for goodness-of-fit testing (Jiménez-Gamero et al. 2009; Chen et al. 2022; Meintanis et al. 2024; Bothma et al. 2021); our focus is complementary—exact finite-sample estimation and inference in the exponential proportional hazards model (PHM).

We study the estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ of the baseline rate θ in the exponential PHM under random right censoring, with the censoring intensity β an unknown nuisance parameter. Here $\bar{\delta}$ is the fraction of uncensored observations and $\bar{Z}$ the mean observed time. This estimator is known to coincide with the maximum-likelihood estimator, and its asymptotic normality is established (Abdushukurov 2009); what has been missing are exact finite-sample formulas for its bias and variance. We supply them: for every $n \geq 3$,

$$\mathbb{E}_\theta[\hat{\theta}] - \theta = \frac{\theta}{n-1}, \qquad \mathrm{Var}_\theta[\hat{\theta}] = \frac{n^2\theta^2}{(n-1)^2(n-2)} + \frac{n\beta\theta^2}{(n-1)(n-2)}. \tag{2}$$

The word *exact* is used precisely: these are exact expressions for the first two moments at every fixed $n \geq 3$, obtained from the exact gamma distribution of $\sum_i Z_i$; they are

not statements about the full sampling distribution of $\hat{\theta}$, which is not available in closed form. They yield exact confidence intervals for the mean, an exact bias correction $\hat{\theta}_c = (n-1)\bar{\delta}/(n\bar{Z})$, and an explicit sample-size rule—of direct value in reliability and actuarial applications, where samples are small and asymptotic approximations unreliable.

The paper makes three contributions. *First*, we prove (Proposition 4.3) that the ECF/score-weighted estimating equation coincides exactly with the maximum-likelihood score equation on *any* support, giving a general frequency–time-domain equivalence that justifies ECF weights beyond the polynomial-score setting; the Cauchy location family, whose score is not polynomial, is treated as a non-trivial instance. *Second*, for the exponential PHM we derive the exact finite-sample formulas (2) and show, through a closed-form Fisher information computation (Lemma 2.12), that $\hat{\theta}$ attains the parametric Cramér–Rao bound for the model with β unknown, the factor $1/(\beta+1)$ reflecting the cost of estimating β rather than any loss of information. *Third*, we go beyond the closed-form regime in two directions: to the Weibull PHM with *unknown* shape, for which we give a consistent, $\sqrt{n}$-asymptotically normal moment system whose finite-sample efficiency essentially matches maximum likelihood; and to dependent censoring, for which we obtain the exact probability limit of $\hat{\theta}$ under an arbitrary dependence structure (Theorem 2.19), a robustness diagnosis unavailable from the likelihood alone. The efficiency claim is one of *parametric* efficiency; semiparametric efficiency in the sense of Bickel et al. (1993) is not claimed.

The main new results are, to the best of our knowledge: (i) exact bias and variance formulas for $\hat{\theta}$ (Theorem 2.5); (ii) an explicit Fisher information matrix for (θ, β) (Lemma 2.12); (iii) an extension to the Weibull PHM with known shape (Theorem 2.16) and an efficient moment system for unknown shape (Section 2.4.3); and (iv) the exact probability limit under dependent censoring (Theorem 2.19). The coincidence of $\hat{\theta}$ with the MLE and its asymptotic normality were known (Abdushukurov 2009); the finite-sample formulas, the FIM, the Weibull extension, and the dependent-censoring analysis are new. These results fill a gap in the literature, as no exact finite-sample bias and variance formulas were previously available for the exponential PHM with unknown censoring intensity.

The main results are presented first (Results), then interpreted (Discussion), with the ECF framework and regularity conditions collected in Methods; full proofs and the normal- and Cauchy-model derivations are given as Supplementary Information.

2 Results

2.1 Asymptotic Properties of $\hat{\theta} = \bar{\delta}/\bar{Z}$

2.1.1 Exact moments and consistency

We first establish the key auxiliary result.

Lemma 2.1 (Moments of the inverse sample mean) *Let* $Z_1, \ldots, Z_n \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$, $\lambda > 0$, $\bar{Z} = n^{-1}\sum Z_i$. *For* $n \geq 2$*:* $\mathbb{E}[1/\bar{Z}] = n\lambda/(n-1)$. *For* $n \geq 3$*:* $\mathbb{E}[1/\bar{Z}^2] = n^2\lambda^2/[(n-1)(n-2)]$. *For* $n = 2$*:* $\mathbb{E}[1/\bar{Z}^2] = +\infty$.

Proof See Supplementary Information (proof of Lemma 2.1). □

The central step is proving independence of $\bar{\delta}$ and $\bar{Z}$. We first establish pairwise independence of each (δ_i, Z_i) via a direct probability computation, then lift this to independence of $\bar{\delta}$ and $\bar{Z}$ via characteristic functions.

Lemma 2.2 (Pairwise independence: $\delta_i \perp Z_i$) *Under model (19), δ_i and Z_i are independent for each i.*

Proof See Supplementary Information (proof of Lemma 2.2). □

However, the following stronger result holds via a characteristic function argument, using Lemma 2.2.

Lemma 2.3 (Independence of $\bar{\delta}$ and $\bar{Z}$) *Under model (19), $\bar{\delta}$ and $\bar{Z}$ are* independent.

Proof See Supplementary Information (proof of Lemma 2.3). □

Remark 2.4 (Characteristic-function proof) The factorisation can also be seen directly at the level of characteristic functions, in a form that mirrors the ECF construction. For $s, t \in \mathbb{R}$, since the pairs (δ_i, Z_i) are i.i.d. and $\delta_j \perp Z_j$ (Lemma 2.2),

$$\mathbb{E}\big[e^{is\bar{\delta}+it\bar{Z}}\big] = \prod_{j=1}^{n} \mathbb{E}\big[e^{i(s\delta_j+tZ_j)/n}\big] = \prod_{j=1}^{n} \mathbb{E}\big[e^{is\delta_j/n}\big]\mathbb{E}\big[e^{itZ_j/n}\big] = \mathbb{E}\big[e^{is\bar{\delta}}\big]\,\mathbb{E}\big[e^{it\bar{Z}}\big],$$

so by the uniqueness theorem $\bar{\delta} \perp \bar{Z}$ (Billingsley 1995, Th. 26.3).

Theorem 2.5 (Asymptotic unbiasedness and consistency) *Let (Z_i, δ_i), $i = 1, \ldots, n$, be i.i.d. from the PHM with exponential distribution (19), $\beta, \theta > 0$ unknown, $n \geq 3$. Write $\theta_* = (\beta+1)\theta$. The estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ satisfies:*

(a) (Asymptotic unbiasedness) $\mathbb{E}[\hat{\theta}] = n\theta/(n-1)$, *and* $\mathbb{E}[\hat{\theta}] - \theta = \theta/(n-1) = O(n^{-1})$.

(b) (Exact variance)

$$\mathrm{Var}[\hat{\theta}] = \frac{n^2\theta^2}{(n-1)^2(n-2)} + \frac{n\beta\theta^2}{(n-1)(n-2)} = O(n^{-1}). \tag{3}$$

(c) (Consistency) $\hat{\theta} \xrightarrow{\mathbb{P}} \theta$.

Remark 2.6 Formula (3) shows the cost of estimating β: the second term $n\beta\theta^2/[(n-1)(n-2)]$, absent for known β, dominates when $\beta \gg 1$. For example, at $\beta = 10$, $n = 100$: $\mathrm{Var}[\hat{\theta}] \approx (100 \cdot 10/99/98)\,\theta^2 \approx 0.103\,\theta^2$. A valid upper bound is $\mathrm{Var}[\hat{\theta}] \leq n(\beta+1)\theta^2/[(n-1)(n-2)]$. As a check that no algebraic slip is hidden in the finite-sample formula, take $n = 3$. The gamma inverse moments give $\mathbb{E}[1/\bar{Z}^2] = n^2\theta_*^2/[(n-1)(n-2)] = \frac{9}{2}\theta_*^2$, and the Bernoulli count gives $\mathbb{E}[\bar{\delta}^2] = p(1-p)/n + p^2 = (\beta+3)/[3(\beta+1)^2]$ with $p = 1/(\beta+1)$; by independence $\mathbb{E}[\hat{\theta}^2] = \mathbb{E}[\bar{\delta}^2]\,\mathbb{E}[1/\bar{Z}^2]$ and, subtracting $\mathbb{E}[\hat{\theta}]^2 = \{3\theta/(3-1)\}^2 = \frac{9}{4}\theta^2$ (using $\theta_* = (\beta+1)\theta$), one obtains $\mathrm{Var}[\hat{\theta}] = \frac{9}{4}\theta^2 + \frac{3}{2}\beta\theta^2$, exactly formula (3) at $n = 3$.

Proof See Supplementary Information (proof of Theorem 2.5). □

Remark 2.7 (Bias-corrected estimator) The estimator $\hat{\theta}_c = \frac{n-1}{n}\hat{\theta} = (n-1)\bar{\delta}/(n\bar{Z})$ is exactly unbiased for all $n \geq 3$, with $\mathrm{Var}[\hat{\theta}_c] = [(n-1)^2/n^2]\mathrm{Var}[\hat{\theta}]$.

Corollary 2.8 (Known β) *For $\hat{\theta} = 1/[(\beta+1)\bar{Z}]$ with known β, the same bias $\theta/(n-1)$ holds, and*

$$\mathrm{Var}[\hat{\theta}]\big|_{\beta\ known} = \frac{n^2\theta^2}{(n-1)^2(n-2)} = O(n^{-1}),$$

which equals the first term of (3). The ratio of variances satisfies $\mathrm{Var}|_{\beta\ kn}/\mathrm{Var}|_{\beta\ unkn} = 1/(1+\beta(n-1)/n) < 1$ for $\beta > 0$.

2.1.2 Asymptotic normality

Theorem 2.9 (Asymptotic normality) *Under the conditions of Theorem 2.5,*

$$\sqrt{n}\,(\hat{\theta} - \theta) \rightsquigarrow N(0, \sigma^2), \quad \sigma^2 = (\beta+1)\theta^2 = \theta\theta_*.$$

Proof See Supplementary Information (proof of Theorem 2.9). □

Remark 2.10 (On the asymptotic variance) The limit variance $(\beta+1)\theta^2$ exceeds the variance θ^2 attainable when β is known; the excess $\beta\theta^2$ is the price of estimating the nuisance parameter β from the indicators $\delta_1, \ldots, \delta_n$. That $(\beta+1)\theta^2$ is nonetheless the Cramér–Rao bound for the model with β unknown, so that $\hat{\theta}$ is asymptotically efficient, is established in Corollary 2.13 below.

Remark 2.11 (Approximate confidence interval) From Theorem 2.9, an approximate 95% CI for θ is $\hat{\theta} \pm 1.96\,\hat{\theta}\sqrt{(\hat{\beta}+1)/n}$, where $\hat{\beta} = 1/\bar{\delta} - 1$. This is valid when $\bar{\delta} > 0$; since $\mathbb{P}(\bar{\delta} > 0) = 1 - (1-p)^n \to 1$ as $n \to \infty$, the interval is asymptotically valid. At small n this symmetric interval may produce a negative lower bound; a log-transform pivot is preferable: by the delta method applied to $\log\hat{\theta}$, the 95% CI on the log scale is $\exp\!\big(\log\hat{\theta} \pm 1.96\sqrt{(\hat{\beta}+1)/n}\big)$, which is always positive. For very small n or heavy censoring, bootstrap confidence intervals should be used.

2.1.3 Monte Carlo validation and efficiency comparison

We verify the finite-sample properties of $\hat{\theta}$ at $\theta = 1$, $\beta = 2$ ($\theta_* = 3$) using $R = 100\,000$ replications in Python 3.12 with `numpy.random.default_rng` (seed 42). For each replication, independent $X_i \sim \mathrm{Exp}(1)$ and $Y_i \sim \mathrm{Exp}(2)$ were generated; $Z_i = \min(X_i, Y_i)$, $\delta_i = \mathbf{1}[X_i \leq Y_i]$. The Monte Carlo RMSE is computed as $\widehat{\mathrm{RMSE}} = \big[R^{-1}\sum_{r=1}^{R}(\hat{\theta}_r - \theta)^2\big]^{1/2}$, where $\hat{\theta}_r$ is the estimate from the r-th replication. The Monte Carlo standard error of $\widehat{\mathrm{RMSE}}$ is approximately $\widehat{\mathrm{RMSE}}/\sqrt{2R} < 0.001$, so all reported values are accurate to the last decimal place.

Table 1 Monte Carlo ($R = 100\,000$): bias and RMSE of $\hat{\theta}$ and $\hat{\theta}_c$; theor. RMSE from (3), asymptotic SD from Theorem 2.9.

n	Bias	RMSE (MC)	RMSE (theor.)	Asymptotic SD	RMSE $\hat{\theta}_c$
50	0.0205	0.2529	0.2534	0.2449	0.2478
100	0.0097	0.1758	0.1761	0.1732	0.1741
500	0.0020	0.0774	0.0777	0.0775	0.0773

Table 2 Comparative Monte Carlo ($R = 100\,000$, $\theta = 1$, $\beta = 2$): RMSE of $\hat{\theta}$ (unknown β), oracle $\hat{\theta}_{\mathrm{or}}$ (known β), and bias-corrected $\hat{\theta}_c$. Oracle eff. $= (\theta^2/n)/\mathrm{RMSE}^2$ is relative to the infeasible known-β bound, not the full-model efficiency (which equals one; see Remark 2.14).

	$\hat{\theta}$ (unknown β)		$\hat{\theta}_{\mathrm{or}}$ (known β)		$\hat{\theta}_c$ (bias-corrected)	
n	RMSE	Oracle eff.	RMSE	Oracle eff.	RMSE	Oracle eff.
50	0.2529	0.311	0.1453	0.944	0.2478	0.324
100	0.1758	0.323	0.1013	0.974	0.1741	0.330
500	0.0774	0.334	0.0449	0.991	0.0773	0.334

Table 1 validates the theoretical formulas, and Table 2 compares $\hat{\theta}$ against the oracle estimator $\hat{\theta}_{\mathrm{or}} = 1/[(\beta+1)\bar{Z}]$ (which uses the true $\beta = 2$) and the bias-corrected version $\hat{\theta}_c = (n-1)\bar{\delta}/(n\bar{Z})$.

Table 1 exhibits the bias of order n^{-1} and the root-mean-square error of order $n^{-1/2}$ predicted by Theorem 2.5. The exact formula (3) and the simulated RMSE agree to within the Monte Carlo standard error at every n (the differences, e.g. 0.2535 versus 0.2534 at $n = 50$, are of the same order as the reported MC SE and are therefore not statistically distinguishable), which is the strongest agreement a finite simulation can demonstrate; the first-order approximation, by contrast, is visibly biased below $n = 100$. The oracle estimator of Table 2, which uses the true value of β, attains between 94 and 99 per cent of the oracle Cramér–Rao bound. That $\hat{\theta}$ itself attains the Cramér–Rao bound of the model with β unknown is the content of Corollary 2.13.

Coverage across censoring levels.

To confirm that the exact variance formula yields reliable interval inference, we repeat the experiment at three censoring intensities, $\beta = 0.5$ (light, $\approx 67\%$ events), $\beta = 2$ (moderate, $\approx 33\%$), and $\beta = 5$ (heavy, $\approx 17\%$), for $n \in \{50, 100, 500\}$, and record the empirical coverage of the nominal 95% log-transform confidence interval of Remark 2.11 ($R = 40\,000$). Table 3 shows coverage within Monte Carlo error of 0.95 in every cell, including $n = 50$ under heavy censoring, confirming that the exact variance and the log-transform pivot deliver accurate inference well outside the large-sample regime.

Table 3 Empirical coverage of the 95% log-transform confidence interval for θ (Remark 2.11) at $\theta = 1$, $R = 40\,000$; Monte Carlo standard error ≈ 0.001.

β (events)	$n = 50$	$n = 100$	$n = 500$
0.5 ($\approx 67\%$)	0.948	0.950	0.950
2 ($\approx 33\%$)	0.951	0.950	0.950
5 ($\approx 17\%$)	0.956	0.952	0.950

Table 4 Bias and RMSE of $\hat{\theta} = \bar{\delta}/\bar{Z}$ versus the censoring-ignorant $\tilde{\theta} = 1/\bar{Z}$ at $\theta = 1$, $\beta = 2$, $R = 40\,000$.

n	Bias($\hat{\theta}$)	Bias($\tilde{\theta}$)	RMSE($\hat{\theta}$)	RMSE($\tilde{\theta}$)
50	0.021	2.064	0.256	2.111
100	0.010	2.030	0.177	2.053
500	0.002	2.006	0.077	2.011

Comparison with the censoring-ignorant estimator.
The estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ coincides with the maximum likelihood estimator and so accounts for censoring correctly. To show that this correction is material, Table 4 contrasts it with the naive estimator $\tilde{\theta} = 1/\bar{Z}$ that ignores the event indicators (as would be appropriate only if every observation were an event). At $\beta = 2$ the naive estimator targets $\theta_* = (\beta + 1)\theta = 3$ rather than θ and is biased by about $+2.0$ for all n, whereas $\hat{\theta}$ carries only the exact finite-sample bias $\theta/(n - 1)$ of Theorem 2.5.

2.2 Full Fisher information matrix for (θ, β)

The observations (Z_i, δ_i) carry information about both θ and β. Since $\delta_i \perp Z_i$ (Theorem 4.7), the density of one observation factorises as

$$f(z, \delta; \theta, \beta) = \theta_* \, e^{-\theta_* z} \cdot p^{\delta}(1 - p)^{1-\delta},$$

with $\theta_* = (\beta + 1)\theta$ and $p = 1/(1 + \beta)$. The joint log-likelihood scores are:

$$s_\theta(z, \delta; \theta, \beta) = \partial \ln f/\partial\theta = 1/\theta - (\beta + 1)z,$$
$$s_\beta(z, \delta; \theta, \beta) = \partial \ln f/\partial\beta = (1 - \delta)/\beta - \theta z.$$

Lemma 2.12 (Joint Fisher information matrix; cf. Fisher 1925) *The Fisher information matrix for (θ, β) from one observation (Z, δ) under model (19) is*

$$\mathcal{I}(\theta, \beta) = \begin{pmatrix} 1/\theta^2 & 1/[\theta(\beta + 1)] \\ 1/[\theta(\beta + 1)] & 1/[\beta(\beta + 1)] \end{pmatrix}, \tag{4}$$

with determinant $\det \mathcal{I} = 1/[\theta^2 \beta(\beta + 1)^2]$.

Proof See Supplementary Information (proof of Lemma 2.12). □

Corollary 2.13 (Cramér–Rao bound for θ) *Inverting the Fisher information matrix of Lemma 2.12, the Cramér–Rao bound for θ in the full two-parameter model with (θ, β) unknown is*

$$\mathrm{CR}(\theta; \theta, \beta) = [\mathcal{I}^{-1}]_{11}/n = \frac{(\beta+1)\theta^2}{n}. \tag{5}$$

Proof By matrix inversion: $[\mathcal{I}^{-1}]_{11} = \mathcal{I}_{22}/\det\mathcal{I} = [1/(\beta(\beta+1))]\cdot\theta^2\beta(\beta+1)^2 = (\beta+1)\theta^2$. □

Remark 2.14 (Full-model efficiency and oracle efficiency) To avoid ambiguity we distinguish two notions. The *full-model efficiency* of an estimator is the ratio of the Cramér–Rao bound of the model with β unknown, namely $(\beta+1)\theta^2/n$, to its variance. The *oracle efficiency* is the ratio of the smaller bound θ^2/n of the reduced model with β *known* to its variance; the latter bound is infeasible when β must be estimated.

The estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ has full-model efficiency one: it is asymptotically efficient, a known consequence of its coincidence with ML (Proposition 4.11) under the standard regularity of the exponential PHM (thrice-differentiable log-likelihood, non-singular $\mathcal{I}$, open parameter space). Corollary 2.13 makes this quantitative: its asymptotic variance $(\beta+1)\theta^2$ (Theorem 2.9) equals $[\mathcal{I}^{-1}]_{11}$. Its oracle efficiency is $1/(\beta+1) < 1$; this is the figure reported in the oracle-ratio column of Table 2, and it measures the price of not knowing β, not a deficiency of the estimator.

2.3 Fisher information for θ at known β

For completeness, when β is known one observes only Z_i; the score $s_\theta = 1/\theta - (\beta+1)z$ gives $I_{\mathrm{or}}(\theta) = 1/\theta^2$ and CR bound θ^2/n. The oracle estimator $1/[(\beta+1)\bar{Z}]$ achieves this bound. Table 2 (oracle column) confirms 97–99% efficiency at $n \geq 100$; the small gap is due to finite-n bias $\theta/(n-1)$.

Remark 2.15 (Finite-sample bounds and the Bhattacharyya system) The Cramér–Rao bound is a first-order (linear) lower bound on the variance of an estimator. For biased or nonlinear estimators at fixed n, the sharper *Bhattacharyya bounds* (cf. Borovkov 1984, Sect. 2.5) refine it by incorporating higher-order derivatives of the log-likelihood, yielding a hierarchy $\mathrm{Var}[\hat{\theta}] \geq b_1 \geq b_2 \geq \cdots$ where b_1 is the CR bound. The exact variance formula (3) derived here gives the *attained* finite-sample variance directly, bypassing the need for the Bhattacharyya hierarchy; a worthwhile theoretical question (listed in Section 3.1) is whether $\hat{\theta}_c$ attains the second Bhattacharyya bound b_2 exactly, which would characterise its finite-sample optimality within the class of the estimator's smooth transformations. The information content of censored observations and the associated Fisher–Bhattacharyya lower bounds under right random censoring have been studied in detail by Abdushukurov et al. (2024b) and Abdushukurov et al. (2024a); in particular, informative censoring need not entail a loss of Fisher information, a phenomenon consistent with the exact accounting given here.

2.4 Weibull PHM: Extension and Efficiency

2.4.1 Model and estimator

Throughout this section the shape parameter γ is assumed *known*. This assumption is essential to the closed-form analysis: the Weibull score for γ is not polynomial in the data, so the delta-function reduction fails, and the joint estimation of (θ, γ, β) has no closed-form solution. The case of unknown γ requires numerical optimisation and is left as an open problem (Section 3.1). In many reliability applications γ is fixed by the failure physics or estimated separately from engineering knowledge, so the known-shape case is of independent interest.

Let $X_i \sim \text{Weibull}(\theta, \gamma)$ with $F(x;\theta,\gamma) = 1 - e^{-(\theta x)^\gamma}$, $f(x;\theta,\gamma) = \gamma\theta^\gamma x^{\gamma-1}e^{-(\theta x)^\gamma}$, $\theta, \gamma > 0$. Under the PHM, the density of $Z_i = \min(X_i, Y_i)$ is:

$$h(z;\theta,\gamma,\beta) = (\beta+1)\gamma\theta^\gamma z^{\gamma-1}e^{-(\beta+1)(\theta z)^\gamma} =: \gamma(\theta_W)^\gamma z^{\gamma-1}e^{-(\theta_W z)^\gamma},$$

where $\theta_W = \theta\,(\beta+1)^{1/\gamma}$ (not to be confused with $\theta_* = (\beta+1)\theta$ of the exponential model; the two coincide only at $\gamma = 1$). Thus $Z_i \sim \text{Weibull}(\theta_W, \gamma)$, and *the Weibull PHM preserves the Weibull family.*

By Theorem 4.7 (universal independence), $\delta_i \perp Z_i$ with $\Pr(\delta_i = 1) = 1/(\beta+1)$. The key observation is that

$$W_i := Z_i^\gamma \sim \text{Exp}(\theta_W^\gamma) = \text{Exp}(\theta^\gamma(\beta+1)),$$

reducing the Weibull case to an exponential model in $W_i = Z_i^\gamma$. Applying the exponential ECF estimator with W_i in place of Z_i and θ^γ in place of θ gives:

$$\widehat{\theta^\gamma} = \frac{\bar{\delta}}{\overline{Z^\gamma}}, \quad \overline{Z^\gamma} = n^{-1}\textstyle\sum Z_i^\gamma.$$

Inverting the transformation $\theta^\gamma \mapsto \theta$ yields the *Weibull PHM estimator*:

$$\hat{\theta}_W = \left(\frac{\bar{\delta}}{\overline{Z^\gamma}}\right)^{1/\gamma}. \tag{6}$$

For $\gamma = 1$ this reduces to $\hat{\theta} = \bar{\delta}/\bar{Z}$.

Theorem 2.16 (Asymptotic normality and efficiency: Weibull PHM) *Suppose $X_i \sim \text{Weibull}(\theta, \gamma)$ with known shape $\gamma \geq 1$, and let $\beta, \theta > 0$ be unknown, $n \geq 3$. The estimator (6) satisfies:*

(a) (Consistency) $\hat{\theta}_W \xrightarrow{\mathbb{P}} \theta$.

(b) (Asymptotic normality) $\sqrt{n}(\hat{\theta}_W - \theta) \rightsquigarrow N(0, \sigma_W^2)$, *where* $\sigma_W^2 = (\beta+1)\theta^2/\gamma^2$.

(c) (Asymptotic efficiency) $\hat{\theta}_W$ *achieves the Cramér–Rao bound for the full model (θ, β) unknown.*

Table 5 Monte Carlo for Weibull PHM ($\theta = 1$, $\beta = 1$, $\gamma = 2$, $R = 100\,000$): RMSE of $\hat{\theta}_W = (\bar{\delta}/\overline{Z^2})^{1/2}$; asymptotic SD $= 1/\sqrt{2n}$ equals the CR SD.

n	RMSE (MC)	Asymptotic SD	Efficiency
50	0.1019	0.1000	0.963
100	0.0713	0.0707	0.983
500	0.0317	0.0316	0.996

Proof See Supplementary Information (proof of Theorem 2.16). □

Remark 2.17 (Special cases)

- $\gamma = 1$ (exponential): $\sigma_W^2 = (\beta + 1)\theta^2$, recovering Corollary 2.13.
- $\gamma = 2$ (Rayleigh): $\sigma_W^2 = (\beta + 1)\theta^2/4$.
- As $\gamma \to \infty$: $\sigma_W^2 \to 0$, reflecting that the Weibull concentrates near its mode and is increasingly easy to estimate.

2.4.2 Monte Carlo validation for Weibull PHM

We verify Theorem 2.16 at $\theta = 1$, $\beta = 1$, $\gamma = 2$ (Rayleigh PHM, $\theta_W = \theta(1 + \beta)^{1/2} = \sqrt{2}$) with $R = 100\,000$ replications. For each replication, $X_i \sim \text{Weibull}(1, 2)$ and $Y_i \sim \text{Weibull}(1, 2)$ $^{[Y_i = (X_i^2\, U_i)^{1/2} \text{ with } U_i \sim \text{Exp}(\beta)]}$; in practice $Z_i = \min(X_i, Y_i)$ and $Z_i^\gamma = Z_i^2$ is used.

Table 5 confirms asymptotic efficiency: RMSE converges to the CR standard deviation at the $\sqrt{n}$ rate, with efficiency approaching 1 as n grows.

2.4.3 Unknown shape: estimating system and numerical scheme

When γ is unknown the closed-form reduction of Section 2.4 is no longer available, because the score for γ is not polynomial in the data and $\hat{\theta}_W$ of (6) presupposes a known exponent. We record here that the ECF approach nonetheless yields a fully specified, consistent, and $\sqrt{n}$-asymptotically normal estimator of the triple (θ, γ, β) through a system of three moment equations, each expressible in terms of derivatives of $\hat{\varphi}_n$ at the origin; only the last of the three requires a one-dimensional numerical solve. We do not obtain exact finite-sample moments in this case—hence the qualifier *partial*—but the construction shows the methodology extends beyond the closed-form regime.

Recall $Z_i \sim \text{Weibull}(\theta_W, \gamma)$ with $\theta_W = \theta(\beta + 1)^{1/\gamma}$ and $\delta_i \perp Z_i$, $p = \Pr(\delta_i = 1) = 1/(\beta + 1)$. Writing $a = 1/\theta_W$ for the Weibull scale, the raw moments are $\mathbb{E}[Z^s] = a^s\, \Gamma(1 + s/\gamma)$, and $\mathbb{E}[Z^s] = (-i)^s \varphi_Z^{(s)}(0)$ for integer s, i.e. they are ECF derivatives at the origin. Three population identities suffice to identify (θ, γ, β):

$$\mathbb{E}[\delta] = p = \frac{1}{\beta + 1}, \tag{7}$$

$$\frac{\mathbb{E}[Z^2]}{(\mathbb{E}[Z])^2} = \frac{\Gamma(1+2/\gamma)}{\Gamma(1+1/\gamma)^2} =: R(\gamma), \tag{8}$$

$$\mathbb{E}[Z^\gamma] = a^\gamma = \frac{1}{\theta^\gamma(\beta+1)}. \tag{9}$$

The scale a cancels in (8), so that equation involves γ alone. The function R is smooth and *strictly decreasing* on $(0,\infty)$, from $R(\gamma)\to\infty$ as $\gamma\downarrow 0$ to $R(\gamma)\downarrow 1$ as $\gamma\to\infty$ (indeed $R(1)=2$, $R(2)=4/\pi\approx 1.273$); hence (8) has a unique root γ for every attainable sample ratio, and γ is identified.

Estimator and algorithm.

Replacing population moments by sample averages $\bar\delta$, $m_1=\bar Z$, $m_2=n^{-1}\sum_i Z_i^2$, $m_\gamma=n^{-1}\sum_i Z_i^\gamma$ gives the plug-in system

$$\hat\beta = \frac{1}{\bar\delta} - 1, \qquad R(\hat\gamma) = \frac{m_2}{m_1^2}, \qquad \hat\theta = \Big[\frac{1}{(\hat\beta+1)\,m_{\hat\gamma}}\Big]^{1/\hat\gamma}, \tag{10}$$

solved in the order $\hat\beta\to\hat\gamma\to\hat\theta$. Only the middle step is not explicit; because R is strictly monotone, it is solved by any bracketing method (bisection) or by Newton's iteration $\gamma_{k+1}=\gamma_k-\big(R(\gamma_k)-m_2/m_1^2\big)/R'(\gamma_k)$, which in our simulations converges to a tolerance of 10^{-10} in about seven iterations from the start $\gamma_0=1$. The starting value $\gamma_0=1$ is natural because it corresponds to the exponential special case and lies in the domain of attraction for all reasonable samples; equivalently, since R is monotone, bisection on a fixed bracket such as $[0.1,10]$—which covers all plausible shapes—converges unconditionally, so the middle step is robust to the choice of initialisation. In practice the shape parameter in reliability and survival data typically lies between 0.5 and 5, so the bracket $[0.1,10]$ is deliberately conservative; when prior information constrains γ a narrower bracket may be used to accelerate the iteration. The remaining two steps are closed-form.

Proposition 2.18 (Consistency and asymptotic normality, unknown γ) *Let $(Z_i,\delta_i)_{i=1}^n$ be i.i.d. from the Weibull PHM with $\theta,\gamma,\beta>0$ all unknown and $\mathbb{E}[Z^2]<\infty$. The estimator (10) is well defined with probability tending to one, is strongly consistent for (θ,γ,β), and satisfies $\sqrt{n}\big((\hat\theta,\hat\gamma,\hat\beta)-(\theta,\gamma,\beta)\big)\rightsquigarrow N_3(0,\Sigma)$ for a covariance matrix Σ obtained by the delta method from the joint CLT of $(\bar\delta,m_1,m_2,m_\gamma)$.*

Proof sketch The vector of sample moments $(\bar\delta,m_1,m_2)$ converges a.s. to $(p,\mathbb{E}[Z],\mathbb{E}[Z^2])$ by the SLLN, and $\sqrt{n}$-CLT holds since $\mathbb{E}[Z^2]<\infty$ implies the required second moments of the moment functions (for the $\hat\theta$ step, $\mathbb{E}[Z^{2\gamma}]<\infty$ holds automatically because Z^γ is exponential). The maps $\hat\beta=1/\bar\delta-1$, $\hat\gamma=R^{-1}(m_2/m_1^2)$ (well defined and C^1 because R is a C^1 diffeomorphism onto its range), and $\hat\theta=[(\hat\beta+1)m_{\hat\gamma}]^{-1/\hat\gamma}$ are continuously differentiable in a neighbourhood of the true moment vector; consistency follows by the continuous mapping theorem and asymptotic normality by the multivariate delta method. The entries of Σ are smooth functions of (θ,γ,β) and the Weibull moments $\Gamma(1+s/\gamma)$; we omit their lengthy explicit form, as the sensitivity that matters in practice, Σ itself, is evaluated numerically. □

Table 6 Monte Carlo for the unknown-shape Weibull PHM ($\theta = 1.5$, $\gamma = 2.5$, $\beta = 1$; $R = 2000$ replications): mean estimates from system (10).

n	$\overline{\hat{\theta}}$	$\overline{\hat{\gamma}}$	$\overline{\hat{\beta}}$
200	1.503	2.531	1.013
1000	1.501	2.502	1.003
5000	1.500	2.499	1.001

Table 7 Efficiency of the moment estimator (ME) of system (10) relative to three-parameter MLE (unknown-γ Weibull PHM, $(\theta, \gamma, \beta) = (1.5, 2.5, 1.0)$, $R = 1000$–1500). Eff. $= \mathrm{MSE}_{\mathrm{MLE}}/\mathrm{MSE}_{\mathrm{ME}}$; the β-estimators coincide.

	θ		γ		β	
n	$\mathrm{RMSE}_{\mathrm{ME}}$	Eff.	$\mathrm{RMSE}_{\mathrm{ME}}$	Eff.	$\mathrm{RMSE}_{\mathrm{ME}}$	Eff.
200	0.060	1.00	0.140	0.98	0.147	1.00
1000	0.027	1.00	0.063	0.99	0.064	1.00

Table 6 reports a Monte Carlo check at $(\theta, \gamma, \beta) = (1.5, 2.5, 1.0)$: all three components are consistent, with bias decaying at the n^{-1} rate and the $\sqrt{n}$-scaled variances stabilising, in agreement with Proposition 2.18.

Numerical efficiency relative to maximum likelihood.

Although the moment system (10) uses a non-optimal (moment) weight rather than the full score, its efficiency loss is negligible for the rate parameter. We computed the three-parameter maximum likelihood estimator numerically (Nelder–Mead, started at the moment estimate) and compared it with the moment estimator over R replications at the same $(\theta, \gamma, \beta) = (1.5, 2.5, 1.0)$; Table 7 reports the resulting root-mean-square errors and the MSE-ratio efficiency. For θ and β the two estimators are statistically indistinguishable (efficiency ≈ 1 at all n); for β the moment estimator in fact coincides with the MLE $\hat{\beta} = 1/\bar{\delta} - 1$ (Proposition 4.11), so its efficiency is exactly 1 by construction and the table confirms this numerically. For the shape γ the moment estimator shows a small but genuine efficiency loss of one to two per cent (MSE-ratio efficiency 0.98 at $n = 200$ and 0.99 at $n = 1000$), which a bootstrap analysis (2000 resamples) confirms is real rather than Monte Carlo noise (at $n = 1000$ the bootstrap 95% interval for the ratio excludes 1). Importantly, this loss is confined to γ and does not propagate to θ or β, both of which remain fully efficient; the simple moment system is therefore, for practical purposes, as efficient as maximum likelihood for the parameters of primary interest (θ, β), at a fraction of the computational cost, while leaving a small, quantified efficiency gap for the shape γ. Whether an optimally weighted ECF estimator attains full efficiency for γ as well is left open (Section 3.1).

Table 8 Sensitivity of $\hat{\theta}$ to dependent censoring (Gaussian copula of correlation ρ between X and Y; $\theta = 1$, $\beta = 2$, $n = 100$, $R = 60\,000$). At $\rho = 0$ the relative bias reproduces the exact $100/(n-1) \approx 1\%$.

ρ	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3
Rel. bias (%)	22.2	15.0	7.9	1.1	-5.8	-12.5	-19.4
RMSE	0.299	0.244	0.200	0.177	0.179	0.202	0.245

2.5 Robustness: Dependent Censoring and Model Misspecification

Sections 2.5.1 and 2.5.2 examine the two ways the exponential-PHM assumptions can fail: a violation of the independent-censoring assumption, and a violation of the constant-hazard (exponential lifetime) assumption. In both cases the estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ remains consistent for an explicit functional of the data-generating law, and the exact probability limit identifies the resulting bias.

2.5.1 Sensitivity to dependent censoring

The exponential PHM, in common with parametric survival analysis generally, assumes that the censoring time Y is independent of the lifetime X. Under this assumption the independence $\delta_i \perp Z_i$ (Theorem 4.7) holds and the exact moment formulas of Theorem 2.5 apply. When independence fails, δ_i and Z_i are dependent, the factorisation $\mathbb{E}[\hat{\theta}] = \mathbb{E}[\bar{\delta}]\,\mathbb{E}[1/\bar{Z}]$ used in the proof no longer holds, and $\hat{\theta}$ is in general biased. We quantify the effect of a moderate departure by simulation.

We couple $X \sim \mathrm{Exp}(\theta)$ and $Y \sim \mathrm{Exp}(\beta\theta)$ through a Gaussian copula with correlation ρ, so that $\rho = 0$ is the independent model and $|\rho|$ measures the strength of the dependence. Concretely, we draw (U, V) from the bivariate normal law with standard margins and correlation ρ, and set $X = -\log\{1 - \Phi(U)\}/\theta$ and $Y = -\log\{1 - \Phi(V)\}/(\beta\theta)$, where Φ is the standard normal distribution function; the marginal laws of X and Y are then exactly exponential, and only their dependence is altered. Table 8 reports the relative bias and root-mean-square error of $\hat{\theta}$ at $\theta = 1$, $\beta = 2$, $n = 100$ ($R = 60\,000$ replications).

The bias grows approximately linearly in ρ near the independent model, with slope of order 0.7 per unit correlation; the relative bias remains within a few per cent for $|\rho| \lesssim 0.05$ and within about 8% for $|\rho| \le 0.1$. Positive dependence — subjects prone to early failure also prone to early censoring — biases $\hat{\theta}$ downward, and negative dependence upward.

The limiting behaviour admits an exact and completely general description.

Theorem 2.19 (Probability limit under arbitrary censoring dependence) *Let $(X_i, Y_i)_{i=1}^n$ be i.i.d. copies of (X, Y) with an arbitrary joint distribution having continuous marginals and $\mathbb{E}[\min(X, Y)] \in (0, \infty)$, and let $Z_i = \min(X_i, Y_i)$, $\delta_i = \mathbf{1}\{X_i \le Y_i\}$, $\hat{\theta} = \bar{\delta}/\bar{Z}$. Then, almost surely,*

$$\hat{\theta} = \frac{\bar{\delta}}{\bar{Z}} \xrightarrow{\text{a.s.}} \theta^* := \frac{\Pr(X \le Y)}{\mathbb{E}[\min(X, Y)]}. \tag{11}$$

In particular $\hat{\theta}$ is strongly consistent for θ^*, and for θ if and only if $\theta^* = \theta$; under independent censoring in the exponential PHM $\theta^* = \theta$, whereas under dependence $\theta^* = \theta + b(\rho)$ with $b(0) = 0$ and, to leading order near the independent model, $b(\rho) = c(\beta)\,\rho + O(\rho^2)$ for a slope $c(\beta)$ that depends on β and on the copula.

Proof Both $\bar{\delta} \xrightarrow{\text{a.s.}} \Pr(X \le Y)$ and $\bar{Z} \xrightarrow{\text{a.s.}} \mathbb{E}[\min(X, Y)]$ by the strong law of large numbers (Durrett 2010, Th. 2.4.1), the latter finite and positive by hypothesis; the continuous mapping theorem applied to the ratio gives (11). Under the exponential PHM with independent censoring, $\Pr(X \le Y) = 1/(\beta + 1)$ and $\mathbb{E}[\min(X, Y)] = 1/[(\beta + 1)\theta]$, whose ratio is θ. The expansion of b follows from differentiability of θ^* in ρ at $\rho = 0$ for the Gaussian copula, $b(0) = 0$ holding because $\rho = 0$ is the independent model. □ □

Equation (11) shows that $\hat{\theta}$ remains consistent for the functional $\theta^* = \Pr(X \le Y)/\mathbb{E}[\min(X, Y)]$ under *any* censoring mechanism, and is consistent for θ precisely when this functional equals θ — in particular under independent censoring. This is a robustness diagnosis of a kind not available from the likelihood alone: it identifies exactly what $\hat{\theta}$ estimates when the independence assumption fails, in closed form and for an arbitrary dependence structure. The simulation of Table 8 quantifies $b(\rho)$ for the Gaussian-copula family; for the Gaussian copula at $\beta = 2$ the fitted slope is $c \approx -0.7\theta$, positive ρ (early failure coupled with early censoring) shortening the mean observed time and biasing $\hat{\theta}$ downward. Here $b(\rho)$ is the asymptotic bias of the probability limit, not the finite-sample bias $\theta/(n-1)$ of Theorem 2.5. A closed-form expression for $c(\beta)$ for a general copula requires modelling the dependence explicitly and is left open.

2.5.2 Sensitivity to distributional misspecification

The exponential PHM assumes a constant hazard for the lifetime. To assess the behaviour of $\hat{\theta} = \bar{\delta}/\bar{Z}$ when this assumption is violated, we generate lifetimes X_i from alternative distributions while keeping the censoring mechanism exponential with the same β (so that the independence $\delta_i \perp Z_i$ of Theorem 4.7 is retained and only the lifetime law is misspecified). We consider three scenarios, each scaled so that $\mathbb{E}[X] = 1$ matches the exponential mean: (i) Weibull with decreasing hazard, shape $\gamma = 0.5$; (ii) Weibull with increasing hazard, shape $\gamma = 1.5$; and (iii) lognormal with $\mathbb{E}[X] = \mathrm{Var}(X) = 1$ matched to Exp(1). Censoring is $Y \sim \mathrm{Exp}(\beta\theta)$ with $\theta = 1$, $\beta = 2$.

Under misspecification $\hat{\theta}$ no longer estimates θ; by (11) it remains consistent for the functional $\Pr(X \le Y)/\mathbb{E}[\min(X, Y)]$, whose value—the "theoretical limit" below—generally differs from $\theta = 1$. Table 9 reports this limit together with the finite-sample mean, relative bias with respect to θ, and RMSE.

Two features are visible. First, in every scenario the finite-sample mean tracks the theoretical limit (11) closely, confirming that $\hat{\theta}$ is a consistent estimator of $\Pr(X \le Y)/\mathbb{E}[\min(X, Y)]$ regardless of the lifetime law—the estimator is stable, not erratic, under misspecification. Second, that functional is a faithful proxy for θ only when the hazard is (near) constant: for the strongly non-exponential Weibull with decreasing hazard ($\gamma = 0.5$) the target is more than twice θ, and the relative bias with respect to θ is correspondingly large, while for a mild increasing-hazard departure the bias is moderate. We stress that this is *not* robustness in the usual sense of retaining

Table 9 Sensitivity to lifetime misspecification ($\theta = 1$, $\beta = 2$, $R = 8000$; exponential censoring, lifetimes scaled to $\mathbb{E}[X] = 1$). Theor. limit from (11); rel. bias relative to $\theta = 1$.

Lifetime law	Theor. limit	n	Mean $\hat{\theta}$	Rel. bias (%)	RMSE
Exponential(1)	1.000	100	1.012	1.2	0.176
		500	1.002	0.2	0.076
Weibull(γ=0.5)	2.402	100	2.453	145.3	1.522
		500	2.413	141.3	1.427
Weibull(γ=1.5)	0.669	100	0.671	−32.9	0.351
		500	0.670	−33.0	0.334
Lognormal	0.789	100	0.791	−20.9	0.247
		500	0.789	−21.1	0.219

Table 10 Leukemia remission data: group summaries and ECF estimates under the exponential PHM. ECF and ML coincide (Proposition 4.11); 95% CI via log-transform; exact SD from (3).

Group	n	Events	$\hat{\beta}$	$\hat{\theta}$	Median (weeks)	95% CI	Exact SD
6-MP	21	9	1.333	0.0251	27.6	(0.013, 0.048)	0.0091
Placebo	21	21	0.000	0.1154	6.0	(0.076, 0.174)	0.0278

consistency for θ under model departures: rather, the estimator remains consistent for a *different* functional, $\Pr(X \leq Y)/\mathbb{E}[\min(X, Y)]$, which varies continuously with the data-generating law and coincides with the baseline rate θ only under (approximate) constant hazard. This behaviour is fully described by the probability limit (11), and the analysis above quantifies the resulting bias relative to θ; robustness to outliers is a separate property, not considered here. The practical recommendation is therefore unambiguous: the constant-hazard assumption should be checked—by the likelihood-ratio goodness-of-fit test used in Section 2.6—before the exact confidence intervals for θ are relied upon; when the test indicates a non-constant hazard, the Weibull extension of Section 2.4 (known or unknown shape) is the appropriate model.

2.6 Application to leukemia remission data

We apply the exponential PHM estimator to the leukemia remission-time dataset of Freireich et al. (1963), as reproduced in Kalbfleisch and Prentice (2002, Table 1.1). The trial randomised 42 patients to 6-mercaptopurine (6-MP, $n_1 = 21$) or placebo ($n_2 = 21$); time to relapse (weeks) was recorded, with censoring at study end for the 6-MP arm.

Data summary.

Table 10 summarises the two groups. For the 6-MP arm: $n = 21$, 9 relapses ($\delta_i = 1$), 12 censored ($\delta_i = 0$); $\bar{Z} = 359/21 \approx 17.10$ weeks, $\bar{\delta} = 9/21 \approx 0.4286$.

Point estimates and their interpretation.

For the 6-MP group:

$$\hat{\beta} = 1/\bar{\delta} - 1 = 12/9 \approx 1.333, \quad \hat{\theta} = \bar{\delta}/\bar{Z} = 9/359 \approx 0.0251 \text{ week}^{-1}.$$

The median remission time for 6-MP is $\log 2/\hat{\theta} \approx 27.6$ weeks, compared to 6.0 weeks for placebo, a ratio of 4.6. The estimated hazard ratio is $\hat{\theta}/\hat{\theta}_{\text{placebo}} \approx 0.0251/0.1154 \approx 0.22$, i.e. 6-MP reduces the relapse hazard by approximately 78%.

Confidence intervals and bias correction.

The log-transform 95% CI for θ (6-MP arm) is $(0.013, 0.048)$ (Remark 2.11). The bias-corrected estimate $\hat{\theta}_c = (20/21)\hat{\theta} \approx 0.0239$ lies within the interval; the finite-sample bias $\theta/(n-1)$ is small relative to the CI width. The exact SD from formula (3), evaluated by plugging in $(\hat{\theta}, \hat{\beta})$, is 0.0091, giving an exact (non-asymptotic) 95% interval of $(0.0251 \pm 1.96 \times 0.0091) = (0.007, 0.043)$ on the linear scale. For the placebo arm, where $\hat{\beta} = 0$, the same plug-in formula gives the larger SD 0.0278 (Table 10); the reason this SD is larger despite $\hat{\beta} = 0$ removing the second term of (3) is that $\hat{\theta}_{\text{placebo}}$ itself is more than four times $\hat{\theta}_{\text{6-}MP}$, and the leading term of (3) scales as θ^2.

The boundary case $\hat{\beta} = 0$.

Every placebo subject relapsed before the trial ended ($\delta_i = 1$ for all $i = 1, \ldots, 21$), so $\bar{\delta} = 1$ exactly and $\hat{\beta} = 1/\bar{\delta} - 1 = 0$. This is not an assumption imposed on the model but an empirical outcome: it reflects that, for this arm, no subject's censoring time preceded their relapse time, consistent with a trial design in which placebo patients—who relapse quickly— were followed to the event before the administrative cutoff. It does not contradict the presence of a censoring mechanism in the model; it only means the realised censoring fraction was zero for these 21 patients. At $\beta = 0$ exactly, the exponential PHM degenerates to the ordinary uncensored exponential model ($Z_i = X_i$), and the exact moment formulas (3) specialise correctly: the second (censoring) term vanishes and $\text{Var}[\hat{\theta}] = n^2\theta^2/[(n-1)^2(n-2)]$ is the classical exact variance of the reciprocal-sample-mean estimator of an exponential rate under no censoring (see, e.g., Lawless 2003, Sec. 4.1)—itself a known result, recovered here as the $\beta = 0$ special case of Theorem 2.5.

Empirical survival function.

Table 11 shows the Kaplan–Meier (KM) estimate of the survival function at selected time points for both groups, alongside the fitted exponential survival $S(t) = e^{-\hat{\theta}t}$.

The exponential fit is close for the 6-MP arm through the first half of the observation period ($t \leq 18$ weeks), where fitted and Kaplan–Meier survival agree to within about one percentage point; beyond $t \approx 24$ weeks the Kaplan–Meier curve flattens—an artefact of the heavy right-censoring in this arm, where the last observed relapse is at week 23 and all later times are censored—while the exponential continues to decay, so the two diverge in the sparsely-observed tail. The placebo arm, which is uncensored,

Table 11 Kaplan–Meier vs. fitted exponential survival $\hat{S}(t) = e^{-\hat{\theta}t}$ at selected times. KM values computed from the raw data.

	6-MP ($\hat{\theta} = 0.0251$)		Placebo ($\hat{\theta} = 0.1154$)	
t (weeks)	$\hat{S}_{\mathrm{KM}}(t)$	$e^{-\hat{\theta}t}$	$\hat{S}_{\mathrm{KM}}(t)$	$e^{-\hat{\theta}t}$
6	0.857	0.860	0.571	0.500
12	0.753	0.740	0.190	0.250
18	0.627	0.636	0.095	0.125
24	0.448	0.547	—	—
35	0.448	0.415	—	—

shows a departure at intermediate times ($t = 12$), consistent with the known limitation of the constant-hazard model for the placebo remission distribution (Kalbfleisch and Prentice 2002). These are the expected qualitative features of a first, constant-hazard fit and do not undermine the finite-sample inference for $\hat{\theta}$, which is exact under the model.

To quantify the fit we perform a likelihood-ratio test of the exponential null against the two-parameter Weibull alternative (one extra shape parameter, so the reference distribution is χ^2_1). For the 6-MP arm the maximised Weibull shape is $\hat{k} = 1.35$ and the test gives LR $= 1.03$, $p = 0.31$, indicating no significant departure from the exponential model; for the uncensored placebo arm $\hat{k} = 1.37$, LR $= 2.86$, $p = 0.09$, a mild departure consistent with the intermediate-time discrepancy noted above and with the known limitation of the constant-hazard assumption for the placebo remission distribution (Lawless 2003). Neither test rejects the exponential model at the 5% level, so the constant-hazard fit is an adequate first-order description of both arms.

Comparison with alternative estimates.

The ML estimate coincides with $\hat{\theta}$ by Proposition 4.11. A simple method-of-moments estimate (reciprocal of the mean observed time, ignoring censoring) gives $1/\bar{Z} \approx 0.0585$ week^{-1}, which is biased upward because it ignores the censoring structure. The bias-corrected ECF estimate $\hat{\theta}_c = 0.0239$ is the recommended point estimate for reporting.

We report this single, well-studied dataset in detail because the paper is methodological and the leukemia trial exhibits every feature the theory addresses—right censoring, an unknown censoring intensity, and a small sample where the exact formulas matter. The procedure applies verbatim to any right-censored sample under the exponential PHM: the point estimate is $\bar{\delta}/\bar{Z}$, the exact variance is (3), and the constant-hazard assumption is checked by the goodness-of-fit test above. We have not included a second empirical example so as not to dilute the theoretical focus; applying the same three steps to other standard right-censored datasets is immediate and yields analogous conclusions.

The continuum-GMM estimator of Carrasco and Kotchoni (2017), which uses an optimal weight over a continuum of moment conditions, yields the same ML solution

0.0251 for this dataset (since the model is correctly specified and regular). The practical advantage of the closed-form ECF estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$ over continuum-GMM is its simplicity: no numerical optimisation, no bandwidth selection, and an exact variance formula available for direct sample-size planning.

Relevance to actuarial and reliability forecasting.

With $n = 21$, the asymptotic approximation and the exact formula (3) already differ appreciably (Table 1 shows the gap at $n = 50$ is still visible). In actuarial survival-function forecasting and reliability studies, sample sizes of this order are common, and the exact variance permits confidence bounds on the fitted survival curve $\hat{S}(t) = e^{-\hat{\theta}t}$ that do not rely on large-n approximations. For instance, propagating the log-transform CI for $\hat{\theta}$ through the monotone map $t_{1/2} = \log 2/\hat{\theta}$ gives a 95% interval of (14.4, 53.1) weeks for the 6-MP median remission time — an asymmetric band that reflects the finite-sample uncertainty more honestly than a symmetric normal approximation.

3 Discussion

3.0.1 Advantages of the delta-function approach and comparison with existing methods

The delta formalism (i) eliminates numerical quadrature for the integral (12) when the score is polynomial; (ii) gives transparent, checkable derivations connecting score equations to sample moments; (iii) extends to censored models via the ECF (Csörgő 1981b); (iv) is complementary to robust methods (Ghosh and Basu 2019).

Comparison with continuum-GMM.

The continuum-GMM estimator of Carrasco and Kotchoni (2017) uses the optimal weight over a continuum of moment conditions and achieves near-ML efficiency. Because the score weight already recovers ML exactly for the exponential family, $\hat{\theta} = \bar{\delta}/\bar{Z}$ is asymptotically efficient in the full model (Corollary 2.13); the distinction is finite-sample—the delta-function method yields closed-form exact bias and variance, whereas continuum-GMM requires numerical optimisation and a regularisation choice with no comparable finite-sample formulas.

Comparison with robust PHM estimation.

The minimum density power divergence estimators of Ghosh and Basu (2019) are robust to contamination; $\hat{\theta}$ is not designed for robustness but offers exact closed-form bias and variance. Robustness favours the former; model correctness and simplicity favour $\hat{\theta}$.

3.0.2 Parametric versus semiparametric efficiency

The efficiency statement of Corollary 2.13 is *parametric*: within the exponential proportional hazards model, $\hat{\theta} = \bar{\delta}/\bar{Z}$ attains the Cramér–Rao bound for the parameter

θ when β is unknown. It is not, and is not claimed to be, semiparametrically efficient in the sense of Bickel et al. (1993), where the baseline hazard is left unspecified.

The natural semiparametric benchmark is the Cox proportional hazards estimator (Cox 1972), which estimates regression coefficients through the partial likelihood without specifying the baseline hazard and is semiparametrically efficient for those coefficients. The two procedures address different estimands. The Cox estimator targets covariate effects and treats the baseline as an infinite-dimensional nuisance, delivering no closed-form estimate of the baseline rate; the estimator $\hat{\theta}$ targets the baseline rate itself under a fully specified exponential form, and delivers closed-form finite-sample inference for it. When the exponential assumption is credible — as in many reliability and actuarial settings, where a constant hazard is a standard first model — $\hat{\theta}$ offers exact small-sample inference at negligible computational cost. When the baseline form is in doubt, the Cox estimator is more robust to misspecification, at the price of forgoing closed-form baseline inference and requiring covariates. This is the familiar trade-off between parametric and semiparametric modelling; the present paper is deliberately situated on the parametric side, and the exactness of its formulas is a direct consequence of that choice.

3.1 Conclusion and Practical Recommendations

The contribution of this paper is the pair of exact finite-sample moment formulas of Theorem 2.5 — bias $\theta/(n-1)$ and variance (3) for $\hat{\theta} = \bar{\delta}/\bar{Z}$ in the exponential PHM with both θ and β unknown — and their practical corollaries: direct sample-size calculation, exact confidence intervals for the mean, and bias correction via $\hat{\theta}_c$. We do not repeat the precise sense of "exact" or the provenance of the formulas, set out in the Introduction; we recall only that both rest on the gamma law of $\sum_i Z_i$ under independent censoring and that we have not found them in the monograph literature (Abdushukurov 2009; Kalbfleisch and Prentice 2002).

The supporting results establish context and scope. The independence $\delta_i \perp Z_i$ (Theorem 4.7) is a classical consequence of the PHM structure; we state it explicitly with a short proof because it underlies every subsequent derivation. The efficiency of $\hat{\theta}$ (Corollary 2.13) follows from its coincidence with ML (Proposition 4.11, which is known (Abdushukurov 2009)); the explicit FIM computation provides a quantitative verification and clarifies the "oracle gap" narrative in earlier literature. This efficiency is parametric, within the exponential model; semiparametric efficiency is neither achieved nor claimed (Section 3.0.2). For the Cauchy model, Lemma 4.14 fills a gap in the existing literature by verifying all regularity conditions of Feuerverger and McDunnough (1981a) for the non-smooth weight. The accompanying finite-sample study (Supplementary Table S1) shows that the ECF estimator is statistically equivalent to *global* maximum likelihood — as it must be, the two sharing the same estimating equation — and that its practical value is the unimodality of that equation, which removes the multi-start heuristic on which a naive ML implementation depends. The Weibull PHM extension (Theorem 2.16) follows the same pattern under the assumption of known shape.

Scope and limitations.

The results rest on three assumptions that delimit their scope. First, censoring is independent of the lifetime; under dependent censoring $\hat{\theta}$ is biased, though the degradation is $O(\rho)$ and mild for weak dependence (Section 2.5.1). Second, the Weibull extension requires the shape γ to be known; the unknown-shape problem has no closed-form solution. Third, the efficiency statement is parametric; the estimator is not semiparametrically efficient (Section 3.0.2). Each of these is a deliberate restriction that buys the exact finite-sample formulas, and each is listed among the open problems.

Practical recommendations.

- *Point estimation:* use the bias-corrected form $\hat{\theta}_c = (n-1)\bar{\delta}/(n\bar{Z})$.
- *Confidence intervals:* use the log-transform pivot $\exp(\log\hat{\theta} \pm 1.96\sqrt{(\hat{\beta}+1)/n})$ to guarantee positive bounds (Remark 2.11).
- *Sample size:* to achieve $\mathrm{RMSE}(\hat{\theta}) \leq \varepsilon$, take $n \geq (\beta+1)\theta^2/\varepsilon^2 + 3$.
- *Weibull PHM:* use $\hat{\theta}_W = (\bar{\delta}/\overline{Z^\gamma})^{1/\gamma}$ with known shape γ; the same exact bias formula applies to the transformed model.

3.2 Open Questions

(1) *Weibull PHM with unknown shape γ: exact finite-sample theory and efficiency.* Section 2.4.3 constructs a consistent, $\sqrt{n}$-asymptotically normal ECF estimator of (θ, γ, β) via the moment system (10), so the estimation problem itself is solved, and Table 7 shows it to be numerically near-efficient—indistinguishable from maximum likelihood for θ and β, with only a small loss for γ. What remains open is (i) the *finite-sample* theory: because the score for γ is not polynomial in the data and $\hat{\gamma}$ enters through a non-linear moment equation, exact bias and variance formulas analogous to Theorem 2.5 are not available; and (ii) a rigorous proof, rather than numerical evidence, that the moment estimator is exactly asymptotically efficient for γ (i.e. attains the Cramér–Rao bound of the three-parameter model), given that the numerical evidence of Table 7 already shows any loss to be small ($\lesssim 2\%$ at $n = 1000$); constructing an optimally weighted ECF estimator that closes this gap is the natural next step. These are the most immediate open problems.

(2) *Dependent censoring.* The exact formulas of Theorem 2.5 rest on the independence of X and Y. The sensitivity analysis of Section 2.5.1 shows a graceful $O(\rho)$ degradation under a Gaussian copula, but a systematic treatment — for example through copula or shared-frailty models, with corresponding corrections to the bias and variance — remains open.

(3) *Semiparametric refinement.* The estimator $\hat{\theta}$ is parametrically efficient within the exponential PHM (Corollary 2.13) but not semiparametrically efficient in the sense of Bickel et al. (1993). Constructing an estimator of the baseline rate that retains closed-form finite-sample inference while approaching semiparametric efficiency — bridging $\hat{\theta}$ and the Cox partial-likelihood estimator (Cox 1972) — is a natural goal.

(4) *Competing risks models.* The PHM with a single censoring source is a special case of the more general competing-risks framework, in which each subject may fail from one of several mutually exclusive causes $J \in \{1, \ldots, m\}$, and one observes (Z, J) with $Z = \min_j T_j$. The independence structure of Theorem 4.7 generalises: under proportional cause-specific hazards $\lambda_j(x) = \alpha_j \lambda_0(x)$, the cause indicator J and the time Z are independent, with $\Pr(J = j) = \alpha_j / \sum_k \alpha_k$. Extending the exact bias and variance formulas (3) to the cause-specific rate estimators $\hat{\alpha}_j$ in this setting — and identifying the joint FIM for $(\theta, \alpha_1, \ldots, \alpha_m)$ — is a natural and practically important direction for modern survival analysis (Kalbfleisch and Prentice 2002).

(5) *Robustness to PHM misspecification.* The results in this paper assume the proportional hazards structure $1 - G(x) = [1 - F(x)]^\beta$ holds exactly. A natural robustness question is the asymptotic behaviour of $\hat{\theta} = \bar{\delta}/\bar{Z}$ when the true model is a contaminated PHM, e.g. $(1 - G(x)) = (1 - F(x))^{\beta + \varepsilon h(x)}$ for a small perturbation h. The influence-function approach of Ghosh and Basu (2019) provides a framework for such analysis; applying it to the ECF estimator is open.

(6) *Gamma and log-normal PHM.* Theorem 4.7 holds for all absolutely continuous lifetimes; however, the Z_i^γ reduction specific to the Weibull family does not extend to the Gamma or log-normal families. For those the ECF estimating equation requires numerical solution, and characterising its finite-sample behaviour is open.

(7) *Optimal finite-sample weights and higher-order information.* For fixed n, the exact bias $\theta/(n-1)$ can be eliminated by $\hat{\theta}_c$; characterising the weight ω_θ that minimises the exact mean squared error $\mathbb{E}[(\hat{\theta} - \theta)^2]$ at fixed n is open. A related question is whether the finite-sample variance of $\hat{\theta}_c$ attains the second Bhattacharyya bound, the higher-order refinement of the Cramér–Rao bound, which would sharpen the finite-sample efficiency statement beyond the first-order comparison used here.

4 Methods

4.1 The ECF Estimating Equation and Its Score-Weight Reduction

4.1.1 General estimating equation

Let $X_1, \ldots, X_n$ be i.i.d. from a density $f(x; \theta)$, $\theta \in \Theta \subseteq \mathbb{R}$. In the ECF framework (Feuerverger and McDunnough 1981b; Abdushukurov and Kadirov 2001) the estimating equation is

$$\int_{-\infty}^{+\infty} \omega_\theta(t) \left[\hat{\varphi}_n(t) - \varphi_\theta(t)\right] dt = 0, \tag{12}$$

where the weight function is the Fourier transform of the score:

$$\omega_\theta(t) = \frac{1}{2\pi} \int_{\mathrm{supp}(f)} \frac{\partial \ln f(x; \theta)}{\partial \theta} \, e^{-itx} \, dx. \tag{13}$$

Remark 4.1 (Connection to ML and choice of weight) Feuerverger and McDunnough (1981a) show that the optimal weight (minimising asymptotic variance over the class of bounded measurable weights) is in general not the score Fourier transform (13). The score weight is nonetheless a natural choice: by Proposition 4.3, proved later in this section, the resulting estimating equation coincides *exactly* with the maximum-likelihood score equation for *every* regular model, so the ECF estimator with the score weight always coincides with ML, not merely for exponential families. What *is* special to exponential-family (and other polynomial-score) models is that this score equation is then also available in *closed form* as a moment equation, via Remark 4.5; for a general score, e.g. the Cauchy location family of Section 4.4, the same equivalence to ML holds but the equation must be solved numerically.

The delta function $\delta(x)$ was introduced by Dirac (Dirac 1930); its rigorous foundation is the Schwartz theory of distributions (Schwartz 1966; Gel'fand and Shilov 1964), in which $\delta \in \mathcal{D}'(\mathbb{R})$ is the functional $\delta[\varphi] := \varphi(0)$, i.e. $\int_{-\infty}^{+\infty} f(x)\,\delta(x-x_0)\,dx = f(x_0)$ for every f continuous at x_0. We list the properties used systematically in Sections 4.1–2.1.

(P1) Normalisation: $\int_{-\infty}^{+\infty} \delta(x)\,dx = 1$.
(P2) Scaling: $\delta(ax-b) = \frac{1}{|a|}\,\delta(x-b/a)$, $a \neq 0$.
(P3) Fourier representation: $\delta(x-x_0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\omega(x-x_0)}\,d\omega$ in $\mathcal{S}'(\mathbb{R})$.
(P4) Sifting with derivatives: For $k \geq 1$ and $f \in C^k$,

$$\int_{-\infty}^{+\infty} f(x)\,\delta^{(k)}(x-x_0)\,dx = (-1)^k f^{(k)}(x_0). \tag{14}$$

Remark 4.2 (Fourier calculus in $\mathcal{S}'(\mathbb{R})$) The following equalities hold *in* $\mathcal{S}'(\mathbb{R})$ (the space of tempered distributions of Schwartz (Gel'fand and Shilov 1964; Schwartz 1966)); the left-hand integrals diverge in the classical sense but are justified by differentiating **(P3)** inside the distributional pairing (Rudin 1991, Th. 7.19):

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\,dx = \delta(t), \tag{15}$$

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} x\,e^{-itx}\,dx = i\,\delta'(t), \tag{16}$$

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} x^2 e^{-itx}\,dx = -\delta''(t). \tag{17}$$

These are used only as bookkeeping: for every fixed sample $\hat{\varphi}_n$ is entire and, for the normal and exponential models, so is φ_θ, so the action $\int \delta^{(k)}(t)g(t)\,dt = (-1)^k g^{(k)}(0)$ of **(P4)** is legitimate for $g \in \{\hat{\varphi}_n, \varphi_\theta\}$; and the asymptotic arguments of Section 2.1 need only the pointwise limits $\hat{\varphi}_n^{(k)}(0) = i^k n^{-1}\sum_j X_j^k \to i^k\mathbb{E}[X^k]$ by the strong law, with no topology on $\mathcal{S}'(\mathbb{R})$ required. (A functional-central-limit treatment of $\sqrt{n}(\hat{\varphi}_n - \varphi_\theta)$ as a Gaussian process is available (van der Vaart 1998; Csörgő 1981a; Marcus 1981) and underlies the asymptotic normality via the functional delta method, but the pointwise arguments at $t = 0$ used here are more transparent.)

This pointwise calculus applies to models with $\mathbb{R}$-supported, entire characteristic functions (the normal, full-line examples); it does *not* extend, as stated, to the exponential model,

whose support is the half-line $[0, +\infty)$. For the latter the literal one-sided Fourier transform of the score contains principal-value and finite-part terms (Remark 4.4), so the formal weight (20) is a bookkeeping device rather than a literal transform; the rigorous justification of the estimating equation on any support is the separate, self-contained argument of Proposition 4.3, applied to the exponential model in Corollary 4.9.

Proposition 4.3 (The score-weighted ECF equation is the ML score equation, on any support) *Let $f(\cdot;\theta)$ be a density on a (possibly proper) support $S \subseteq \mathbb{R}$ with score $s(x;\theta) = \partial_\theta \ln f(x;\theta)$ satisfying $\mathbb{E}_\theta|s(X;\theta)| < \infty$, and let $\varphi_\theta(t) = \int_S f(x;\theta)e^{itx}\,dx$. Define the weight as the regularised transform of the score, $\omega_\theta(t) = \lim_{\varepsilon\downarrow 0} \frac{1}{2\pi}\int_S s(x;\theta)\, e^{-itx-\varepsilon x^2}\,dx$ (or any equivalent regularisation). Then*

$$\int_{-\infty}^{+\infty} \omega_\theta(t)\left[\hat{\varphi}_n(t) - \varphi_\theta(t)\right] dt = \frac{1}{n}\sum_{k=1}^{n} s(X_k;\theta), \tag{18}$$

i.e. the estimating equation (12) with the score weight is exactly *the maximum-likelihood score equation, for* every *regular model and every support S, half-line or not.*

Proof in Supplementary Information The argument is a straightforward exchange of the order of integration—performing the t-integration first, so that only the ordinary full-line identity $\frac{1}{2\pi}\int_{\mathbb{R}} e^{-it(x-y)}\,dt = \delta(x-y)$ is used and no one-sided transform of 1 or x is ever formed. The details, together with an elementary contour-integration check for the exponential model, are given in the Supplementary Information. □

Remark 4.4 (What this resolves) A literal one-sided Fourier transform of the constant and linear functions on $[0,\infty)$ is *not* a multiple of $\delta(t)$ and $\delta'(t)$; regularising gives principal-value and finite-part terms. Proposition 4.3 shows these terms never need to be tracked: performing the x- and t-integrations in the order used in the proof collapses the construction to the ordinary score equation. The formal notation $\omega_\theta(t) = \sum_k c_k(\theta)\delta^{(k)}(t)$ used from Section 4.3 onward is accordingly bookkeeping for the right-hand side of (18), valid by Proposition 4.3 on any support, rather than a literal claim about the one-sided transform of the score. The singular-term details are given in the Supplementary Information.

Remark 4.5 (Why the method works: a unifying observation) Proposition 4.3 identifies the estimating equation with the score equation $n^{-1}\sum_k s(X_k;\theta) = 0$ in general. The *closed-form* consequence used throughout is the special case in which $s(x;\theta) = \sum_{k=0}^{K} c_k(\theta)\,x^k$ is a polynomial in x: then $n^{-1}\sum_j s(X_j;\theta) = \sum_k c_k(\theta)\, n^{-1}\sum_j X_j^k$ is a polynomial equation in the sample moments, giving the mnemonic $\omega_\theta(t) = \sum_k c_k(\theta)\delta^{(k)}(t)$ and a closed-form estimator. Normal, exponential, Poisson, and binomial models admit this shortcut; the Cauchy model of Section 4.4, whose score is rational rather than polynomial, does not, and is solved by a one-dimensional root search instead.

Remark 4.6 (Normal models as a consistency check) For the normal models $N(\theta,1)$, $N(0,\theta^2)$, and $N(\theta_1,\theta_2^2)$ the score is a polynomial in x of degree at most two, so the weight ω_θ is a finite combination of $\delta, \delta', \delta''$ and, by Remark 4.5, the estimating equation (12) collapses to the sample-moment system in $\bar{X}$ and $\overline{X^2}$. The resulting estimators are the usual sample mean

and variance and coincide with both maximum likelihood and the method of moments; they introduce no new estimator and serve only as a check of the delta-function bookkeeping. One point is worth recording: for θ_2^2 in $N(\theta_1, \theta_2^2)$ the unbiased variance estimator has variance $2\theta_2^4/(n-1)$, so the classical Cramér–Rao bound $2\theta_2^4/n$ is a strict, finite-n-unattainable lower bound (a finite-sample phenomenon revisited for the PHM in Section 2.1). The elementary derivations are collected in the Supplementary Information.

4.2 PHM: General Independence and ECF Estimator

4.2.1 Model and universal independence

We follow Abdushukurov (2009); see also Kalbfleisch and Prentice (2002). Let $X \sim F(x;\theta)$ (lifetime) and $Y \sim G(x;\beta,\theta)$ (censoring time) be independent. Observations: (Z_i, δ_i), where $Z_i = \min(X_i, Y_i)$, $\delta_i = \mathbf{1}[X_i \le Y_i]$. The *proportional hazards* (PH) censoring condition: $1 - G(x;\beta,\theta) = [1 - F(x;\theta)]^\beta$, $\beta > 0$, giving $1 - H(x) = [1 - F(x)]^{\beta+1}$ and density $h(x) = (\beta+1)[1-F(x)]^\beta f(x;\theta)$.

Theorem 4.7 (Independence in PHM; Kalbfleisch and Prentice 2002, Sect. 2.2) *For any absolutely continuous lifetime distribution F satisfying the PH condition, $\delta_i \perp Z_i$ for each i, and $\Pr(\delta_i = 1) = 1/(\beta+1)$ regardless of F.*

Proof The result is a consequence of the PHM structure; we include the short proof for completeness. For $z > 0$, using $S_X(x) = 1 - F(x;\theta)$:

$$\Pr(\delta_i = 1,\, Z_i > z) = \int_z^\infty f(x;\theta)\,[S_X(x)]^\beta\,dx = \frac{[S_X(z)]^{\beta+1}}{\beta+1}.$$

The marginals are $\Pr(\delta_i = 1) = 1/(\beta+1)$ and $\Pr(Z_i > z) = [S_X(z)]^{\beta+1} = S_H(z)$, so the product factorises for all $z > 0$, proving $\delta_i \perp Z_i$. □

Remark 4.8 Theorem 4.7 is a classical consequence of the PHM structure and is implicit in standard references (Kalbfleisch and Prentice 2002; Abdushukurov 2009). We state it explicitly because the independence $\delta_i \perp Z_i$ is the key structural property used in all subsequent derivations. Notably, $\Pr(\delta_i = 1) = 1/(\beta+1)$ depends only on β, not on the form of F; this separability is what makes the two-equation ECF system (Section 4.3.3) tractable. The independence itself holds for every PH lifetime law; what is special to the *exponential* case is that the observed time $Z = \min(X, Y)$ remains exponential, with rate $\theta_* = (\beta+1)\theta$, so that $\sum_i Z_i$ has an exact Gamma(n, θ_*) law. It is this last, model-specific fact—not the independence, which is universal—that converts the general structure into the closed-form finite-sample moments of Section 2.1.

4.2.2 General ECF estimator in PHM

By Theorem 4.7 and the subsequent independence argument (analogous to Lemma 2.3), for any PHM distribution:

- The δ_i are i.i.d. Bernoulli(p), $p = 1/(1+\beta)$, giving ECF equation: $\hat{p} = \bar{\delta}$, i.e. $\hat{\beta} = 1/\bar{\delta} - 1$.
- The Z_i are i.i.d. with distribution $H(\cdot;\theta,\beta)$. One solves a separate ECF equation for θ given $\hat{\beta}$.

For distributions where H belongs to the same parametric family as F (such as Exponential or Weibull, where the PH structure preserves the family), the Z_i ECF equation is tractable in closed form, as we show in Sections 4.3–2.4.

4.3 ECF Estimation: Exponential PHM

This section serves as a benchmark: the exponential PHM is the simplest model in which the ECF/score-weighted estimating equation of Proposition 4.3 can be solved in closed form and every finite-sample quantity computed exactly. The same methodology is then extended to the Weibull model with known shape (Section 2.4), its estimating system for unknown shape and a numerical solution scheme are set out (Section 2.4.3), and its behaviour under dependent censoring is analysed (Section 2.5.1).

4.3.1 Exponential distribution in the PHM

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

$$h(x;\theta) = (\beta+1)\theta\, e^{-(\beta+1)\theta x} =: \theta_*\, e^{-\theta_* x}, \quad \theta_* = (\beta+1)\theta. \tag{19}$$

Thus $Z_i \overset{\text{iid}}{\sim} \text{Exp}(\theta_*)$, and the CF of Z is $\varphi_\theta(t) = \theta_*/(\theta_* - it)$.

4.3.2 Case: β known

Score: $\partial \ln h/\partial\theta = 1/\theta - (\beta+1)x$. Weight (using (15)–(16) over $[0,\infty)$):

$$\omega_\theta(t) = \frac{1}{\theta}\,\delta(t) - i(\beta+1)\,\delta'(t). \tag{20}$$

Corollary 4.9 (The exponential weight, rigorously) *For $h(z;\theta) = \theta_* e^{-\theta_* z}\mathbf{1}_{z\geq 0}$, Proposition 4.3 gives, directly and without ever forming $\omega_\theta(t)$ as an explicit distribution,*

$$\int_{\mathbb{R}} \omega_\theta(t)\big[\hat{\varphi}_n(t) - \varphi_\theta(t)\big]\,dt = \frac{1}{n}\sum_{k=1}^{n}\Big(\frac{1}{\theta} - (\beta+1)Z_k\Big) = \frac{1}{\theta} - (\beta+1)\bar{Z},$$

which is exactly the score equation and vanishes at $\bar{Z} = 1/\theta_*$. Both terms of Proposition 4.3 have additionally been verified for this model by exact contour integration of the ε-regularised integrand (Supplementary Information), with no principal-value or finite-part contribution surviving; the formula (20) is thus correct as a bookkeeping device for (18) (Remark 4.4).

Paired against $g(t) = \hat{\varphi}_n(t) - \varphi_\theta(t)$ in the bookkeeping sense of Remark 4.4, the weight (20) gives $\theta^{-1}\,g(0) + i(\beta+1)\,g'(0) = 0$; since $g(0) = 0$ and $g'(0) = i\big(\bar{Z} - \mathbb{E}_\theta Z\big) =$

$i\big(\bar{Z} - 1/\theta_*\big)$, this is exactly $\bar{Z} = 1/\theta_*$, matching Corollary 4.9. Hence

$$\hat{\theta} = \frac{1}{(\beta+1)\,\bar{Z}}, \tag{21}$$

which is the maximum-likelihood/method-of-moments estimator.

Proposition 4.10 *For known β, the estimator (21) coincides with ML and MM (Abdushukurov 2009). Its exact bias is $\theta/(n-1)$, as will be established in Theorem 2.5 (using Lemma 2.1).*

4.3.3 Case: β unknown

Since $\mathbb{P}(Z_i = X_i) = 1/(1+\beta)$, the LLN gives $\bar{\delta} \xrightarrow{\mathbb{P}} 1/(1+\beta)$. Substituting $\hat{\beta} = 1/\bar{\delta} - 1$ in (21):

$$\hat{\theta} = \frac{\bar{\delta}}{\bar{Z}}. \tag{22}$$

ECF derivation.

Equation (22) may equally be obtained from two independent ECF estimating equations, one for each parameter.

Equation for $\theta_* = (\beta+1)\theta$. The score for θ under (19) is $\partial \ln h/\partial\theta = 1/\theta - (\beta+1)z$. The same computation as in Section 4.3.2 (replacing θ by θ_* as the natural parameter) gives weight (20) and ECF equation:

$$\bar{Z} = 1/\hat{\theta}_*. \tag{23}$$

Equation for β (via the δ_k component). Since $\delta_i \perp Z_i$ (Lemma 2.2), the indicator sequence $\delta_1, \ldots, \delta_n$ is i.i.d. Bernoulli with success probability $p = 1/(1+\beta)$. Its empirical CF is $\hat{\phi}_\delta(t) = n^{-1}\sum_k e^{it\delta_k}$ and the theoretical CF is $\phi_p(t) = (1-p) + p\,e^{it}$. We apply the ECF estimating equation (12) to this one-dimensional family with weight $\omega_p(t) = -\delta'(t)$, the delta-function weight that extracts the first moment:

$$\int_{-\infty}^{+\infty} (-\delta'(t))\big[\hat{\phi}_\delta(t) - \phi_p(t)\big]\,dt = (-1)(-1)\big[\hat{\phi}'_\delta(0) - \phi'_p(0)\big] = i\bar{\delta} - ip = 0,$$

using property **(P4)** and $\hat{\phi}'_\delta(0) = i\bar{\delta}$, $\phi'_p(0) = ip$. This gives $\hat{p} = \bar{\delta}$, i.e.,

$$\hat{\beta} = 1/\bar{\delta} - 1. \tag{24}$$

Combining. From (23) and (24): $(\hat{\beta}+1)\,\hat{\theta} = \hat{\theta}_* = 1/\bar{Z}$, hence $(1/\bar{\delta})\,\hat{\theta} = 1/\bar{Z}$, giving $\hat{\theta} = \bar{\delta}/\bar{Z}$ as in (22). This provides the full ECF justification for the estimator, complementing the ML derivation below.

Proposition 4.11 *The estimator* (22) *coincides with ML and MM (Abdushukurov 2009).*

Proof See Supplementary Information (proof of Proposition 4.11). □

Remark 4.12 (Scope of novelty) The estimator $\hat{\theta} = \bar{\delta}/\bar{Z}$, its coincidence with ML, and asymptotic normality are known (Abdushukurov 2009; Kalbfleisch and Prentice 2002). The new contributions of Section 2.1 are: (i) exact closed-form bias $\theta/(n-1)$ and variance (3) (Theorem 2.5); (ii) a rigorous characteristic-function proof that $\bar{\delta} \perp \bar{Z}$ (Lemma 2.3); (iii) an explicit FIM computation and quantitative explanation of the efficiency gap (Section 2.2).

4.4 A Non-Trivial Example: Cauchy Distribution

To demonstrate that the ECF framework goes beyond exponential families, we apply it to the *Cauchy distribution* Cauchy$(\theta, 1)$ with density $f(x;\theta) = [\pi(1+(x-\theta)^2)]^{-1}$, $\theta \in \mathbb{R}$. For this model ML has multiple local maxima (Barnett 1966), and MM is inapplicable since $\mathbb{E}|X| = \infty$. The CF is $\varphi_\theta(t) = e^{it\theta - |t|}$ and the score is $\partial \ln f/\partial\theta = 2(x-\theta)/(1+(x-\theta)^2)$.

The score is not integrable, so its Fourier transform is taken in the L^2 (Plancherel) sense; a short residue computation (Supplementary Information) gives the weight

$$\omega_\theta(t) = e^{-it\theta}\,(-i\,\operatorname{sgn} t)\,e^{-|t|}, \tag{25}$$

which, unlike the exponential case, is *not* a combination of delta functions but a genuine function in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ (since $|\omega_\theta(t)| = e^{-|t|}$); the estimating integral (12) therefore converges absolutely and is well-posed in the classical sense. Substituting (25) into (12) and using the oddness of $\operatorname{sgn}(t)e^{-2|t|}$ to eliminate the model-CF term (the elementary steps are in the Supplementary Information), the estimating equation reduces to

$$\operatorname{Im}\left[\int_0^\infty e^{-t}\hat{\varphi}_n(t)\,e^{-it\theta}\,dt\right] = 0. \tag{26}$$

The Fisher information of the Cauchy location model is $I(\theta) = \frac{1}{2}$ (Supplementary Information), so the Cramér–Rao bound is $2/n$.

Remark 4.13 (Numerical stability of the Cauchy estimating equation) The estimating equation is well-conditioned for every finite n: since $|\hat{\varphi}_n(t)| \le 1$ and $|\omega_\theta(t)| = e^{-|t|}$, the estimating integral is bounded in absolute value by $\int_{\mathbb{R}} e^{-|t|}\,dt = 2$, uniformly in n and in the sample. The influence function $\psi(x;\theta) = (x-\theta)/[1+(x-\theta)^2]$ (Lemma 4.14(C4)) is bounded by $\frac{1}{2}$, so a single observation, however extreme, moves the estimating equation by at most $O(n^{-1})$ and the heavy Cauchy tails do not destabilise the procedure—confirmed by the simulations of Supplementary Table S1, where every replication converged.

The weight (25) contains $\operatorname{sgn}(t)$, which is non-smooth at $t = 0$; the following lemma records that the regularity conditions for asymptotic efficiency nonetheless hold, with the verification deferred to the Supplementary Information.

Lemma 4.14 (Regularity conditions for the Cauchy ECF estimator) *Let* $\omega_\theta(t) = (-i\,\operatorname{sgn} t)\,e^{-|t|}e^{-it\theta}$ *be the weight of* (25)*. The following conditions of Feuerverger and McDunnough (1981a, Sect. 2) hold for the Cauchy location family:*

(C1) (Integrability) $\omega_\theta \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$*, so the estimating integral* (12) *converges absolutely and the estimating equation is well-posed in the classical sense.*

(C2) (Identifiability) *For* $\theta \neq \theta'$*,* $\int \omega_\theta(t)[\varphi_\theta(t) - \varphi_{\theta'}(t)]\,dt \neq 0$*: the Cauchy CF* $e^{it\theta-|t|}$ *separates points, so the estimating equation has a unique zero at* θ *(van der Vaart 1998, Sect. 7.5).*

(C3) (Score identity) $\int \omega_\theta(t)\,e^{it\theta-|t|}\,dt = 0$ *and* $\partial_\theta \int \omega_\theta(t)\varphi_\theta(t)\,dt \neq 0$*: the first follows from oddness of* $\operatorname{sgn}(t)e^{-2|t|}$*; the second equals* $-I(\theta) = -1/2 \neq 0$ *by the score-weight identity of FM (1981a).*

(C4) (Non-smoothness at $t = 0$) *The weight* $(-i\,\operatorname{sgn} t)e^{-|t|}$ *is discontinuous at* $t = 0$ *but belongs to* $L^1 \cap L^2$*; the corresponding estimating function* $\psi(x;\theta) = \int_0^\infty e^{-t}\sin(t(x-\theta))\,dt = (x-\theta)/[1+(x-\theta)^2]$ *is smooth in both* x *and* θ*, with Fisher-consistent influence function. Consistency and asymptotic normality of the resulting* Z*-estimator follow by standard* M*-estimation theory (van der Vaart 1998, Thm. 5.9, Thm. 5.21), since* $\mathbb{E}|\psi(X;\theta)| < \infty$*,* ψ *is continuously differentiable in* θ*, and* $\mathbb{E}[\psi(X;\theta)] = 0$ *identifies* θ*.*

Proof in Supplementary Information Each condition reduces to an elementary computation: (C1) to $|\omega_\theta(t)| = e^{-|t|} \in L^1 \cap L^2$; (C2) to injectivity of the Cauchy characteristic function in θ; (C3) to the oddness of $\operatorname{sgn}(t)e^{-2|t|}$ and the value $I(\theta) = 1/2$; and (C4) to the Laplace-transform evaluation $\int_0^\infty e^{-t}\sin(t(x-\theta))\,dt = (x-\theta)/[1+(x-\theta)^2]$, followed by the standard M-estimation theorems of van der Vaart (1998). The full computations are given in the Supplementary Information. □

Corollary 4.15 (Asymptotic efficiency of Cauchy ECF estimator) *Under conditions* (C1)–(C4)*, the ECF estimator defined by* (26) *is consistent and satisfies*

$$\sqrt{n}\,(\hat\theta - \theta) \rightsquigarrow N(0, 2),$$

i.e. achieves the Cramér–Rao bound $1/(nI(\theta)) = 2/n$*, despite the absence of a closed-form expression for* $\hat\theta$ *and the inapplicability of MM.*

Remark 4.16 The delta-closure identified in Remark 4.5 is thus a special algebraic feature of exponential-family models, not a limitation of the ECF approach itself. The Cauchy example illustrates that ECF yields a well-defined estimator of the location parameter even without finite moments (Abdushukurov and Kadirov 2001).

4.4.1 Finite-sample comparison: ECF vs. ML for Cauchy

Since the Cauchy ECF estimator requires a one-dimensional line search, its practical standing relative to maximum likelihood is best shown numerically. A Monte Carlo comparison ($R = 50\,000$, $\theta = 0$) is reported in Supplementary Table S1, with the ECF estimator solving (26) by Brent's method and maximum likelihood computed both globally and with three random restarts. When maximum likelihood is computed globally its RMSE is indistinguishable from that of the ECF estimator at every sample

size, both attaining the Cramér–Rao limit $\sqrt{2/n}$, so the two are statistically equivalent; they differ only in the numerical robustness of a default solver, the ECF equation being solved by a single one-dimensional root search with no restart budget to tune.

Acknowledgements. The authors thank the reviewers for detailed and constructive remarks. Research was conducted within the scientific programme of the National University of Uzbekistan named after Mirzo Ulugbek.

Declarations

- **Funding.** Not applicable.
- **Competing interests.** The authors declare no competing interests.
- **Ethics approval and consent to participate.** Not applicable. The leukemia remission data analysed here are a classical, fully anonymised, publicly available dataset (Freireich et al. 1963).
- **Consent for publication.** All authors approved the manuscript for publication.
- **Data availability.** The leukemia remission dataset (Freireich et al. 1963) is in the public domain; a reformatted CSV version is included in the code repository. All simulation code (Python 3.12, `numpy`/`scipy`), pre-computed result tables, and the manuscript source are deposited at Mendeley Data, V1, doi: 10.17632/nyd7y92chg.1.
- **Materials availability.** Not applicable.
- **Code availability.** All code is deposited at Mendeley Data, V1 (doi: 10.17632/nyd7y92chg.1); the repository contains six scripts that reproduce every table and verify every closed-form formula, together with a `README.md` giving step-by-step instructions. All results are reproducible with `pip install numpy scipy pandas` under Python 3.12.
- **Author contribution.** N.N. conceived the study, developed the theory, and wrote the manuscript. G.A. carried out the Monte Carlo experiments and the data analysis. N.B. contributed to the asymptotic analysis and the verification of the closed-form formulas. G.X. contributed to the real-data application and the numerical validation. D.K. contributed to the dependent-censoring analysis and the literature review. All authors reviewed and approved the manuscript.

References

Abdushukurov, A.A., Kadirov, O. (2001). Methods of characteristic functions in statistical estimation from censored samples. In *Statistical Methods of Estimation and Hypothesis Testing*, Perm State University, Perm, pp. 4–12. [Translated: *J. Math. Sci.* **127**(1), 1737–1745 (2005). https://doi.org/10.1007/s10958-005-0162-7]

Abdushukurov, A.A. (2009). *Statistics of Incomplete Observations.* Tashkent: Universitet, 269 pp. [In Russian].

Abdushukurov, A.A., Nurmukhamedova, N.S., Abdushukurov, F.A. (2024). On Fisher–Bhattacharyya information function and lower bounds under right random censoring. *Lobachevskii Journal of Mathematics* **45**(10), 4846–4857.

Abdushukurov, A.A., Nurmukhamedova, N.S., Erisbaev, S.A. (2024). Fisher information contained in incomplete observations. *Industrial Laboratory. Materials Diagnostics* **90**(7), 84–92 [in Russian]. https://doi.org/10.26896/1028-6861-2024-90-7-84-92

Barnett, V.D. (1966). Evaluation of the maximum-likelihood estimator where the likelihood equation has multiple roots. *Biometrika* **53**, 151–165. https://doi.org/10.2307/2334214

Bickel, P.J., Klaassen, C.A.J., Ritov, Y., Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models.* Baltimore: Johns Hopkins University Press.

Billingsley, P. (1995). *Probability and Measure,* 3rd ed. New York: John Wiley & Sons.

Borovkov, A.A. (1984). *Mathematical Statistics.* Moscow: Nauka. [English translation: Gordon & Breach, Amsterdam, 1998.]

Bothma, E., Allison, J.S., Cockeran, M., Visagie, I.J.H. (2021). Characteristic function and Laplace transform-based tests for exponentiality in the presence of random right censoring. *Stat* **10**(1), e394. https://doi.org/10.1002/sta4.394

Bryant, J.L., Paulson, A.S. (1979). Some comments on characteristic function-based estimators. *Sankhyā Ser. A* **41**, 109–116.

Carrasco, M., Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory* **16**(6), 797–834. https://doi.org/10.1017/S0266466600166010

Carrasco, M., Kotchoni, R. (2017). Efficient estimation using the characteristic function. *Econometric Theory* **33**(2), 479–526. https://doi.org/10.1017/S0266466616000025

Chen, F., Jiménez-Gamero, M.D., Meintanis, S.G., Zhu, L. (2022). A general Monte Carlo method for multivariate goodness-of-fit testing applied to elliptical families. *Computational Statistics & Data Analysis* **175**, 107548. https://doi.org/10.1016/j.csda.2022.107548

Cox, D.R. (1972). Regression models and life-tables (with discussion). *J. R. Stat. Soc. Ser. B* **34**(2), 187–220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Csörgő, S. (1981a). Limit behaviour of the empirical characteristic function. *Ann. Probab.* **9**(1), 130–144. https://doi.org/10.1214/aop/1176994513

Csörgő, S. (1981b). Estimating characteristic functions under random censorship. *Carleton Math. Lecture Notes* **30**, 301–315. [Technical report.]

Dirac, P.A.M. (1930). *The Principles of Quantum Mechanics.* Oxford: Clarendon Press.

Durrett, R. (2010). *Probability: Theory and Examples,* 4th ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511779398

Feuerverger, A., McDunnough, P. (1981a). On the efficiency of empirical characteristic function procedures. *J. Roy. Statist. Soc. Ser. B* **43**(1), 20–27. https://doi.org/10.1111/j.2517-6161.1981.tb01143.x

Feuerverger, A., McDunnough, P. (1981b). On some Fourier methods for inference. *J. Amer. Statist. Assoc.* **76**(374), 379–387. https://doi.org/10.2307/2287839

Feuerverger, A., McDunnough, P. (1984). On statistical transform methods and their efficiency. *Canad. J. Statist.* **12**(4), 303–317. https://doi.org/10.2307/3315252

Fisher, R.A. (1925). Theory of statistical estimation. *Proc. Camb. Phil. Soc.* **22**, 700–725. https://doi.org/10.1017/S0305004100009580

Freireich, E.J., Gehan, E., Frei, E., et al. (1963). The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia. *Blood* **21**(6), 699–716. https://doi.org/10.1182/blood.V21.6.699.699

Gel'fand, I.M., Shilov, G.E. (1964). *Generalized Functions, Vol. 1.* New York: Academic Press.

Ghosh, A., Basu, A. (2019). Robust and efficient estimation in the parametric proportional hazards model under random censoring. *Statistics in Medicine* **38**(27), 5283–5299. https://doi.org/10.1002/sim.8377

Heathcote, C.R. (1977). The integrated squared error estimation of parameters. *Biometrika* **64**(2), 255–264. https://doi.org/10.2307/2335718

Jiménez-Gamero, M.D., Alba-Fernández, V., Muñoz-García, J., Chalco-Cano, Y. (2009). Goodness-of-fit tests based on empirical characteristic functions. *Computational Statistics & Data Analysis* **53**(12), 3957–3971. https://doi.org/10.1016/j.csda.2009.05.009

Kalbfleisch, J.D., Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data,* 2nd ed. New York: John Wiley & Sons. https://doi.org/10.1002/9781118032985

Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data,* 2nd ed. New York: John Wiley & Sons.

Marcus, M.B. (1981). Weak convergence of the empirical characteristic function. *Ann. Probab.* **9**(2), 194–201. https://doi.org/10.1214/aop/1176994460

Meintanis, S.G., Milošević, B., Jiménez-Gamero, M.D. (2024). Goodness-of-fit tests based on the min-characteristic function. *Computational Statistics & Data Analysis* **197**, 107988. https://doi.org/10.1016/j.csda.2024.107988

Reeds, J.A. (1985). Asymptotic number of roots of Cauchy location likelihood equations. *Ann. Statist.* **13**(2), 775–784. https://doi.org/10.1214/aos/1176349554

Rudin, W. (1991). *Functional Analysis,* 2nd ed. New York: McGraw-Hill.

Schwartz, L. (1966). *Théorie des Distributions.* Paris: Hermann.

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Reviews* **23**(2), 93–123. https://doi.org/10.1081/ETC-120039605

van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511802256

Zolotarev, V.M. (1986). *One-Dimensional Stable Distributions.* Providence, RI: American Mathematical Society.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.pdf