AIMS Mathematics, 11(5): 13567–13588.
DOI: 10.3934/math.2026558
Received: 04 February 2026
Revised: 19 April 2026
Accepted: 27 April 2026
Published: 15 May 2026

https://www.aimspress.com/journal/Math

Research article

A two-parameter family of Hollander–Proschan–type tests against NBUE alternatives

Juan Ding[1,2], Chi Zhou[1], Anqi Xia[1], Wenxin Zhou[1], and Wenjun Xiong[3,*]

[1] School of Mathematics, Hohai University, Nanjing 211100, China
[2] Laboratory of Mathematical Modeling and Intelligent Computing for Water Systems, Hohai University, Nanjing 211100, China
[3] School of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China

* **Correspondence:** Email: wjxiong@gxnu.edu.cn.

Abstract: Testing for aging behavior is a fundamental problem in reliability analysis. We study tests of exponentiality against new better than used in expectation (NBUE) alternatives. Many Hollander–Proschan type procedures can be interpreted as weighted functionals of a common mean residual life gap, but are often presented in isolated forms, making it difficult to understand how weighting affects sensitivity to different alternatives. We develop a unified weighted-functional framework in which several classical procedures arise as special or limiting cases. Within this framework, we construct a centered and studentized statistic with an asymptotic normal null distribution and an explicit centering term. Pitman efficiency analysis and Monte Carlo experiments illustrate how weighting influences sensitivity and finite-sample performance, while real-data applications demonstrate practical interpretability.

Keywords: NBUE; Hollander–Proschan test; Pitman asymptotic efficiency; reliability; mean residual life
Mathematics Subject Classification: 62N05, 62G10

1. Introduction

In lifetime distributions, identifying aging behavior is a fundamental problem in reliability analysis, and testing exponentiality is a natural starting point in this line of research. Under the exponential model, the elapsed operating time of a device does not alter the distribution of its remaining lifetime, so exponentiality is commonly regarded as the benchmark of no aging [1]. Let

$$X_t \overset{d}{=} (X - t \mid X > t)$$

denote the residual lifetime at age t. By comparing X_t with the initial lifetime X, several standard aging classes can be derived [2]. Among them, the new better than used (NBU) class is defined by $P(X_t > x) \leq P(X > x)$ for all $t, x \geq 0$, while the weaker new better than used in expectation (NBUE) is defined by $\mathbb{E}(X_t) \leq \mathbb{E}(X)$ for all $t \geq 0$. Since NBUE characterizes aging through the mean residual lifetime [3], it has been extensively studied in reliability [4] and survival analysis [5]. Testing exponentiality against NBUE alternatives can therefore be interpreted as assessing whether a lifetime distribution departs from the no-aging benchmark in the direction of deterioration in expected lifetime. In addition, when maintenance, repair, and replacement decisions depend on residual lifetime information, modeling based on the mean residual life (MRL) plays an important role in maintenance planning [6], condition-based maintenance [7], and optimization of multi-component systems [8].

Since the seminal work of Hollander and Proschan [9], testing exponentiality against NBUE alternatives has remained an active area of research [10]. Subsequent studies have proposed various modifications and extensions of Hollander–Proschan-type (HP-type) methods, such as the alternative test of Koul [11], the simplified test introduced by de Souza Borges et al. [12], and the generalized family of HP-type statistics developed by Anis et al. [13]. More recent work has extended NBUE-related methodologies beyond the classical HP framework, including testing procedures under censoring [14–16], tests for other aging properties [17, 18], and nonparametric inference based on the total time on test transform [19]. These developments have substantially broadened the scope of NBUE testing. Recently, empirical studies indicate that HP-type and related tests may exhibit markedly different performances across alternative distributions and sample sizes [20, 21]. This heterogeneity is particularly relevant for HP-type procedures, since different constructions can often be traced back to different ways of organizing the same underlying aging information.

Many HP-type and related procedures can be interpreted through a common aging gap, $\Delta(t) = m(0) - m(t)$, while differing mainly in how this quantity is weighted or aggregated. Although the literature has produced a rich collection of procedures [22], these methods are often introduced in isolated forms, so the effect of the weighting choice on sensitivity to different NBUE departures is not fully transparent. As a result, comparison across procedures is often tied to specific constructions, and the practical choice of a test may remain largely empirical [20]. This lack of a unified weighting perspective motivates the present work.

This paper studies HP-type tests for exponentiality against NBUE alternatives through weighted functionals of the mean residual life gap. This formulation provides a more systematic treatment of HP-type NBUE tests and clarifies the connections among several classical procedures. The contributions of the paper are twofold. First, we develop a unified weighted-functional framework of HP-type tests for NBUE alternatives, under which several classical procedures arise as special cases. Second, we construct a centered and studentized test within this framework, derive a closed-form centering term and obtain the corresponding asymptotic null calibration. We further study its practical performance through local-efficiency analysis, numerical experiments, and a real-data illustration, highlighting how weighting influences sensitivity to different alternatives. The remainder of this article is organized as follows. Section 2 develops the proposed weighted-functional framework for HP-type tests under NBUE alternatives and introduces the corresponding centered and studentized statistics, together with explicit centering terms and asymptotic normal calibration under the null. Section 3 examines finite-sample performance through Monte Carlo experiments and comparisons with existing methods, and also presents real-data applications. Section 4 concludes.

2. Method

2.1. A unified weighted-functional framework for HP-type tests

Let X be a nonnegative random variable with distribution function F and finite mean $\mu = \mathbb{E}(X) \in (0, \infty)$. Denote the survival function by $\bar{F}(x) = 1 - F(x)$. For $t \geq 0$, the MRL function associated with F is

$$m(t) = E[X - t \mid X > t] = \begin{cases} \dfrac{\int_t^\infty \bar{F}(u)\, du}{\bar{F}(t)}, & \text{if } \bar{F}(t) > 0, \\ 0, & \text{if } \bar{F}(t) = 0. \end{cases} \tag{2.1}$$

The MRL function $m(t)$ represents the conditional expected residual life of a unit that has survived up to time t. Then the NBUE property $\mathbb{E}(X_t) \leq \mathbb{E}(X)$ for all $t \geq 0$ is equivalent to the condition

$$m(t) \leq m(0) \qquad \text{for all } t \geq 0. \tag{2.2}$$

As in Hollander and Proschan [9], we consider testing

$$H_0 : F \text{ is exponential on } [0, \infty),\ \bar{F}(x) = \exp(-x/\mu),\ x \geq 0,\ \mu > 0, \tag{2.3}$$

$$H_1 : F \text{ is NBUE but not exponential.} \tag{2.4}$$

Under the NBUE property (2.2), a natural quantity of interest is the aging contrast

$$\Delta(t) = m(0) - m(t), \tag{2.5}$$

which is nonnegative under NBUE alternatives and vanishes identically under exponentiality.

Many tests for exponentiality against NBUE alternatives can be viewed as different ways of aggregating the same aging signal $\Delta(t)$. Their main distinction lies not in the underlying aging contrast itself, but in how that contrast is weighted over the distribution. Motivated by this observation, we consider the weighted functional

$$T_w(F) = \int_0^\infty w\left(F(t)\right) \Delta(t)\, dF(t), \tag{2.6}$$

where $w : [0, 1] \to \mathbb{R}$ is a measurable weight function.

Equation (2.6) provides a unified structural description for a broad group of HP-type procedures. It separates the common aging signal from the weighting mechanism: $\Delta(t)$ represents the underlying NBUE contrast, whereas $w\{F(t)\}$ determines how that contrast is aggregated over the probability scale. This decomposition clarifies the connections among existing procedures and makes the effect of weighting directly analyzable.

Within the unified representation (2.6), we consider the following two-parameter working class of weights:

$$w(x; a, b) = (1 - x)(a + 1 - x)^b, \qquad 0 \leq x \leq 1. \tag{2.7}$$

This formal working class includes the boundary subfamily $a = 0$ [13], for which

$$w(x; 0, b) = (1 - x)^{b+1},$$

thereby connecting the proposed framework to power-weighted HP-type procedures. In particular, $(a, b) = (0, 0)$ recovers the classical Hollander–Proschan test [9], while $(a, b) = (0, -1)$ corresponds to the Aly-type configuration [23]. More generally, $(a, b) = (0, j - 1)$ for integer j corresponds to the generalized HP family of Anis and Mitra [13]. Indeed, these connections follow by direct substitution into (2.7). For $(a, b) = (0, 0)$, $w(x; 0, 0) = (1 - x)$, which is the classical Hollander–Proschan weight. For $(a, b) = (0, -1)$, $w(x; 0, -1) = 1$, which yields the Aly-type configuration. More generally, for integer j, $w(x; 0, j - 1) = (1 - x)^j$, thereby recovering the generalized HP family.

Throughout the paper, we restrict (a, b) to the admissible region

$$\mathcal{A} = \{(a, b) : a > 0,\ b \in \mathbb{R}\} \ \cup\ \{(0, b) : b > -\tfrac{3}{2}\}.$$

This restriction is adopted to ensure that the subsequent closed-form representations and variance-based inference are well defined.

The two parameters also have distinct and interpretable roles. The exponent b governs how strongly the aging contrast is emphasized over the probability scale, whereas a acts as an offset that regularizes the tail region. This structure also retains sufficient analytical tractability to permit explicit centering, studentization, and asymptotic normal calibration under the null hypothesis.

2.2. Asymptotic inference for the proposed procedure

By restricting the general weighted functional to the working class (2.7), we define the HP-type functional studied in this paper. For $(a, b) \in \mathcal{A}$, let

$$T_w(F; a, b) = \int_0^\infty w(F(t); a, b)\Delta(t)dF(t). \tag{2.8}$$

Since $1 - F(t) = \bar{F}(t)$, the weight can also be written as

$$w(F(t); a, b) = \bar{F}(t)\,[a + \bar{F}(t)]^b. \tag{2.9}$$

For both computation and theoretical analysis, it is useful to rewrite (2.8) in equivalent kernel forms. Under the conditions detailed in Appendix A, the following equivalent representations hold:

$$T_w(F; a, b) = \int_0^\infty W(F(t); a, b)\,\bar{F}(t)\,dt = \int_0^\infty t\,J_w(F(t); a, b)\,dF(t), \tag{2.10}$$

where $W(x; a, b) = \int_0^1 w(v; a, b)\,dv - \int_0^x \frac{w(v;a,b)}{1-v}\,dv$, and $J_w(x; a, b) = w(x; a, b) + W(x; a, b)$. For $(a, b) \in \mathcal{A}$ and $0 \le x < 1$, $W(\cdot; a, b)$ has a closed-form expression,

$$W(x; a, b) = \begin{cases} 1 + \ln(1 - x), & a = 0, b = -1, \\ 1 + a\ln\left(\frac{a}{a+1}\right) - \ln\left(\frac{a+1}{a+1-x}\right), & a > 0, b = -1, \\ C(a + 1 - x, b + 1) + \dfrac{C(a, b + 2) - C(a + 1, b + 2)}{b + 1}, & b \neq -1, \end{cases} \tag{2.11}$$

where $C(x, y) = x^y/y$ for $y \neq 0$, and $C(x, 0) = \ln(x)$.

The first representation in (2.10) is particularly convenient for computation. For statistical inference, we work with the empirical counterpart of $T_w(F; a, b)$. Let F_n be the empirical distribution function based on i.i.d. observations $x_1, \ldots, x_n$. Let $x_{(1)} \le \cdots \le x_{(n)}$ be the order statistics and set $x_{(0)} = 0$. The sample analogue of $T_w(F; a, b)$ is

$$T_w(F_n; a, b) = \int_0^\infty W(F_n(t); a, b)\, \bar{F}_n(t)\, dt = \sum_{i=1}^{n} x_{(i)} \int_{(i-1)/n}^{i/n} J_w(x; a, b)\, dx$$
$$= \sum_{i=1}^{n} W\left(\frac{i-1}{n}; a, b\right) \frac{n-i+1}{n} (x_{(i)} - x_{(i-1)}). \tag{2.12}$$

The empirical functional (2.12) is the basic object for inference in the proposed framework. Its large-sample behavior provides the foundation for subsequent standardization and testing. To ensure that both the population functional and its empirical counterpart are well defined, and to support the subsequent asymptotic analysis, we impose the following conditions throughout this subsection: $(a, b) \in \mathcal{A}$, the equivalent representation (2.10) holds, F is absolutely continuous on $[0, \infty)$ with finite mean, and the asymptotic variance $\sigma^2(J_w, F; a, b)$ is finite and positive. These conditions are also directly related to finiteness, integrability, and numerical stability of the proposed statistic. The following theorem establishes its asymptotic normality under these regularity conditions.

Theorem 1. *Let F be an absolutely continuous distribution on $[0, \infty)$ with finite mean, and let $(a, b) \in \mathcal{A}$. Assume that the equivalent representation* (2.10) *holds and that the asymptotic variance $\sigma^2(J_w, F; a, b)$ is finite and positive, where*

$$\sigma^2(J_w, F; a, b) = \int_0^\infty \int_0^\infty J_w(F(x); a, b)\, J_w(F(y); a, b)\, [F(\min\{x, y\}) - F(x)F(y)]\, dx\, dy. \tag{2.13}$$

Then

$$\frac{\sqrt{n}\, [T_w(F_n; a, b) - T_w(F; a, b)]}{\sigma(J_w, F; a, b)} \xrightarrow{d} N(0, 1). \tag{2.14}$$

Under the exponential null H_0 in (2.3), it suffices to work with the unit-mean exponential baseline

$$F_0(x) = 1 - e^{-x}, \qquad x \ge 0. \tag{2.15}$$

To justify this reduction, note that for an exponential distribution with mean $\mu > 0$, the mean residual life is constant, $m(t) \equiv \mu$, and hence $T_w(F; a, b) = 0$ for all (a, b). Moreover, $T_w(F; a, b)$ and its asymptotic standard deviation $\sigma(J_w, F; a, b)$ are scale equivariant within the exponential family. Specially, if X has exponential distribution with mean $\mu > 0$ and $Y = X/\mu$, then Y has distribution F_0 and

$$T_w(F_Y; a, b) = \mu^{-1} T_w(F_X; a, b), \qquad \sigma(J_w, F_Y; a, b) = \mu^{-1} \sigma(J_w, F_X; a, b).$$

Consequently, the standardized statistic in (2.14) is invariant to the nuisance scale μ, so it suffices to work under (2.15).

Specializing (2.13) to F_0 yields a closed-form expression for the asymptotic variance. With $C(x, y)$

defined as above and $(a, b) \in \mathcal{A}$,

$$\sigma^2(J_w, F_0; a, b) = \begin{cases} 1, a = 0,\ b = -1, \\ 1 - a(a+1)[\ln a - \ln(a+1)]^2, a > 0,\ b = -1, \\ \dfrac{C(a+1, 3+2b) - C(a, 3+2b)}{(b+1)^2} - \left[\dfrac{C(a, b+2) - C(a+1, b+2)}{b+1}\right]^2, b \neq -1. \end{cases} \tag{2.16}$$

2.3. Studentized statistic and the algorithm

For practical implementation, we use a centered and studentized version of the empirical functional. Centering is introduced to reduce the finite-sample null drift, whereas studentization places different members of the working class on a common scale. We begin with the scale-free statistic $T_w(F_n; a, b)/\overline{X}_n$, where $\overline{X}_n = n^{-1}\sum_{i=1}^n x_i$. Since $\mathbb{E}_0\{T_w(F_n; a, b)/\mu\}$ is generally nonzero under H_0 at finite n, we define

$$\Delta_T(a, b) = \frac{\sqrt{n}\left\{T_w(F_n; a, b)/\overline{X}_n - \zeta_n(a, b)\right\}}{\sigma(J_w, F_0; a, b)}, \tag{2.17}$$

where the centering term is

$$\zeta_n(a, b) = \frac{1}{n}\sum_{i=1}^n W\left(\frac{i-1}{n}; a, b\right). \tag{2.18}$$

Under H_0, one has $\mathbb{E}_0[T_w(F_n; a, b)/\mu] = \zeta_n(a, b)$, where $\mathbb{E}_0$ denotes expectation under H_0. A detailed derivation is presented in Appendix B. Hence, by Theorem 1, Slutsky's theorem, and the consistency of $\overline{X}_n$ for μ,

$$\Delta_T(a, b) \xrightarrow{d} N(0, 1) \qquad \text{under } H_0.$$

This asymptotic normality provides a common large-sample reference distribution across different members of the class. Detailed derivations are given in Appendix C.

Accordingly, for large samples we reject H_0 at level α if $\Delta_T(a, b) > z_{1-\alpha}$ or equivalently if $p_{\text{asy}}(a, b) = 1 - \Phi(\Delta_T(a, b)) < \alpha$, where Φ is the standard normal distribution function. For finite samples, calibration may instead be based on the empirical critical value $c_{n,\alpha}(a, b)$, defined as the $(1-\alpha)$-quantile of $\Delta_T(a, b)$ under the exponential null distribution. Then the rejection rule becomes $\Delta_T(a, b) > c_{n,\alpha}(a, b)$.

For practical implementation, Algorithm 1 summarizes the resulting testing procedure.

Algorithm 1 Testing procedure based on $\Delta_T(a, b)$

Input: Independent lifetime observations $x_1, \ldots, x_n$, parameters (a, b), and significance level α
Output: The statistic $\Delta_T(a, b)$ and the testing decision

1: Compute $T_w(F_n; a, b)$ from (2.12)
2: Compute the centering term $\zeta_n(a, b)$ from (2.18)
3: Compute the null standard deviation $\sigma(J_w, F_0; a, b)$ from (2.16)
4: Compute $\Delta_T(a, b)$ from (2.17)
5: Reject H_0 if $\Delta_T(a, b) > c_{n,\alpha}(a, b)$ for finite samples, or if $\Delta_T(a, b) > z_{1-\alpha}$ for large samples

2.4. Pitman local efficacy under contiguous alternatives

To compare different choices of (a, b) within the proposed working class, we study Pitman local efficacy under contiguous alternatives [24]. The aim is not to claim universal optimality over all possible weighting schemes, but to provide a transparent, criterion-guided basis for tuning within the present class. The analysis is carried out for the centered and studentized statistic $\Delta_T(a, b)$ in (2.17), which is asymptotically standard normal under H_0.

Let $\{F_\theta : \theta \in \Theta\}$ be a one-parameter model with $F_{\theta_0} = F_0$, where θ_0 corresponds to the unit-mean exponential baseline. We focus on contiguous sequences of alternatives of the form

$$\theta_n = \theta_0 + \frac{h}{\sqrt{n}}, \qquad h \in \mathbb{R}. \tag{2.19}$$

Because $\Delta_T(a, b)$ is normalized by $\overline{X}_n$ and is therefore scale-free under H_0, we study local drift through the scale-normalized population functional $T_w(F_\theta; a, b)/\mu(\theta)$, where $\mu(\theta) = \mathbb{E}_\theta(X) \in (0, \infty)$. The resulting local slope is defined by

$$\lambda(a, b; F_\theta) = \frac{\partial}{\partial\theta}\left\{\frac{T_w(F_\theta; a, b)}{\mu(\theta)}\right\}\Bigg|_{\theta=\theta_0} \Big/ \sigma(J_w, F_0; a, b), \tag{2.20}$$

with $T_w(F_\theta; a, b)$ given by the population representation in (2.10). Under standard differentiability, domination, and moment conditions ensuring that differentiation under the integral sign is valid and that the asymptotic variance is finite, the statistic $\Delta_T(a, b)$ has an asymptotically shifted normal limit

$$\Delta_T(a, b) \xrightarrow{d} N(h\,\lambda(a, b; F_\theta),\ 1). \tag{2.21}$$

Accordingly, the Pitman local efficacy of $\Delta_T(a, b)$ against the family $\{F_\theta\}$ is

$$\mathcal{E}(a, b; F_\theta) = \lambda(a, b; F_\theta)^2. \tag{2.22}$$

In the numerical study, we report $\lambda(a, b; F_\theta)$, since it is the noncentrality coefficient in (2.21) and therefore provides a direct measure of local sensitivity under contiguous alternatives.

We consider three standard reliability families representing distinct directions of departure from exponentiality. Each family reduces to F_0 at θ_0:

(1) Weibull: $\bar{F}_1(x; \theta) = \exp\{-x^\theta\}$, $\quad x > 0,\ \theta > 0,\ \theta_0 = 1$;

(2) Linear failure rate (LFR): $\bar{F}_2(x; \theta) = \exp\left\{-x - \frac{\theta}{2}x^2\right\}$, $\quad x > 0,\ \theta \geq 0,\ \theta_0 = 0$;

(3) Makeham: $\bar{F}_3(x; \theta) = \exp\left\{-x - \theta\,(x + e^{-x} - 1)\right\}$, $\quad x > 0,\ \theta \geq 0,\ \theta_0 = 0$.

For each family, $\lambda(a, b; F_\theta)$ in (2.20) is obtained by differentiating the scale-normalized functional $T_w(F_\theta; a, b)/\mu(\theta)$ at θ_0 and then normalizing by the null standard deviation $\sigma(J_w, F_0; a, b)$ in (2.16). A detailed derivation for the Weibull family, illustrating the evaluation of $\lambda(a, b; F_\theta)$, is provided in Appendix D.

Figures 1–3 display $\lambda(a, b; F_\theta)$ over a grid of (a, b). As discussed in Section 2.1, $(a, b) = (0, 0)$ and $(0, -1)$ recover the original HP test of Hollander and Proschan [9] and the test proposed by Aly [23], while $(a, b) = (0, j - 1)$ for integer j corresponds to the generalized HP family of Anis and Mitra [13].

The figures show that the preferable parameter region depends on the alternative family. For the Weibull and Makeham families, the local slope tends to be larger at nontrivial choices of a for fixed b, indicating that the offset parameter a can materially affect local sensitivity through the weight $w(x; a, b) = (1 - x)(a + 1 - x)^b$. By contrast, for the LFR family, the largest values on the grid occur near $(a, b) = (0, -1)$, which is consistent with the Aly-type choice.

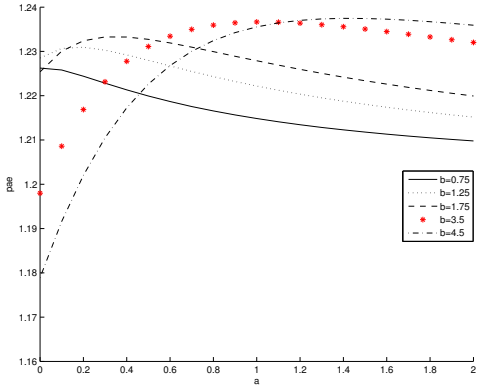

Figure 1. Pitman local slope $\lambda(a, b; F_\theta)$ for the Weibull family under contiguous alternatives.

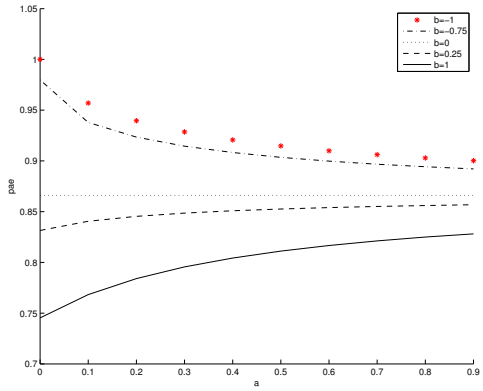

Figure 2. Pitman local slope $\lambda(a, b; F_\theta)$ for the linear failure rate (LFR) family.

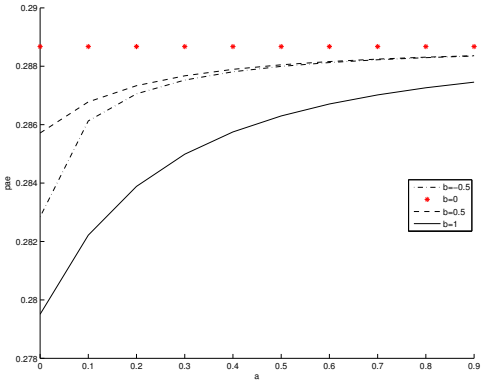

Figure 3. Pitman local slope $\lambda(a, b; F_\theta)$ for the Makeham family.

3. Numerical study

3.1. Simulations

To investigate the finite-sample performance of the proposed test statistic $\Delta_T(a, b)$, we conducted Monte Carlo simulation experiments. Two commonly used classes of increasing failure rate alternatives were considered. The first class is the Weibull distribution, with survival function

$$\bar{F}_1(x;\theta) = \exp\{-x^{\theta}\}, \qquad \theta \in \{1.2, 1.3, 1.4, 1.5\}.$$

The second class is the Gamma distribution, with density function

$$f(x;\theta) = \Gamma(\theta)^{-1} x^{\theta-1} e^{-x}, \qquad \theta \in \{1.4, 1.6, 1.8, 2.0\}.$$

The sample sizes were taken as $n \in \{5, 15, 30, 50, 75, 100\}$. For comparison, we also included the statistic $\gamma^*_{j=0.25}$ recommended by Anis et al. [13] and the statistic T_1 proposed by Hollander and Proschan [9]. Four parameter settings are considered $(a, b) \in \{(0.4, 3.25), (0, 0), (0, -1), (0.5, -0.5)\}$. The choice $(0, 0)$ corresponds to the classical HP weight, whereas the other three parameter combinations are used to reflect finite-sample differences under different tail-weighting schemes. The empirical size under the null and the empirical power under alternatives are defined as the rejection frequencies over $M = 10^5$ independent replications [25]. Random number generation was controlled by a fixed seed to ensure reproducibility of the results [26]. All numerical experiments were implemented in MATLAB 2021a.

Before comparing power, we first examine the empirical size under the null hypothesis. For a given (a, b) and significance level α, finite-sample rejection for $\Delta_T(a, b)$ is calibrated by empirical critical values simulated under H_0; representative critical values for selected parameter settings are reported in Appendix E. Table 1 reports the empirical size at the nominal level $\alpha = 0.05$, that is, the Monte Carlo estimate of the type I error rate under this finite-sample calibration. Overall, the empirical sizes for all four parameter choices remain close to the nominal level across the sample sizes considered. As the sample size increases, the empirical critical values gradually stabilize and become closer to the corresponding standard normal quantiles, which is consistent with the asymptotic normal calibration established in Section 2.

Table 1. Empirical size at the nominal level $\alpha = 0.05$ under the exponential null hypothesis.

n	$\Delta_T(a, b)$			
	$(0.4, 3.25)$	$(0, 0)$	$(0, -1)$	$(0.5, -0.5)$
5	0.0500	0.0495	0.0498	0.0498
15	0.0504	0.0498	0.0503	0.0497
30	0.0484	0.0487	0.0500	0.0489
50	0.0493	0.0496	0.0499	0.0495
75	0.0498	0.0501	0.0498	0.0496
100	0.0495	0.0498	0.0496	0.0500

The empirical power under the Weibull alternatives is reported in Table 2. When the departure from the exponential distribution is weak, for example at $\theta = 1.2$, the differences in power among the methods are small for the extremely small sample size $n = 5$, whereas these differences gradually become more apparent as the sample size increases. For instance, when $\theta = 1.2$ and $n = 100$, the powers of $\Delta_T(0.4, 3.25)$ and $\Delta_T(0, 0)$ are both close to 69%, exceeding that of $\Delta_T(0, -1)$ and being essentially comparable to that of T_1. As the degree of departure increases to $\theta = 1.3$ and above, the power of all methods rises with the sample size, but $\Delta_T(0.4, 3.25)$ and $\Delta_T(0, 0)$ remain at similar levels in most settings, whereas $\Delta_T(0, -1)$ tends to have relatively lower power. For example, at $\theta = 1.4$ and $n = 30$, the powers of the former two are both around 67%, higher than those of $\Delta_T(0, -1)$ and $\gamma^*_{j=0.25}$.

Table 2. Estimated power (%) under the Weibull alternatives.

θ	n	$\Delta_T(a, b)$				Benchmarks	
		$(0.4, 3.25)$	$(0, 0)$	$(0, -1)$	$(0.5, -0.5)$	$\gamma^*_{j=0.25}$	T_1
1.2	5	9.03	9.10	8.90	9.04	9.17	9.22
	15	17.92	17.65	16.24	17.39	16.30	17.36
	30	29.54	28.81	25.02	28.32	26.43	29.04
	50	43.63	42.72	36.50	42.03	38.74	42.66
	75	58.71	57.77	49.02	56.81	52.32	57.11
	100	69.88	69.20	59.13	68.15	63.61	69.32
1.3	5	11.86	11.81	11.37	11.73	11.80	11.98
	15	27.59	27.11	24.39	26.71	24.77	26.77
	30	48.69	48.05	41.78	47.21	44.19	48.27
	50	69.91	69.36	60.84	68.50	64.35	69.31
	75	85.46	85.30	77.03	84.50	80.47	84.88
	100	93.58	93.48	87.15	92.98	90.40	93.53
1.4	5	14.71	14.80	14.23	14.63	14.70	15.02
	15	39.34	39.04	34.98	38.44	35.89	38.65
	30	67.43	67.17	59.51	66.31	62.54	67.42
	50	87.98	87.85	80.89	87.32	84.10	87.82
	75	97.05	97.21	93.42	96.94	95.29	97.08
	100	99.32	99.36	97.89	99.27	98.77	99.37
1.5	5	18.36	18.35	17.53	18.15	18.13	18.58
	15	51.37	51.27	46.01	50.49	47.21	50.86
	30	82.08	82.42	75.10	81.64	78.24	82.59
	50	96.40	96.64	93.12	96.36	94.89	96.62
	75	99.62	99.68	98.81	99.64	99.33	99.66
	100	99.96	99.97	99.83	99.96	99.92	99.97

The empirical power under the Gamma alternatives is reported in Table 3. Compared with the Weibull case, the differences induced by parameter choice are more pronounced under the Gamma alternatives. For example, when $\theta = 1.4$ and $n = 50$, the power of $\Delta_T(0.4, 3.25)$ is about 51%, higher than the approximately 45% achieved by $\Delta_T(0, 0)$ and also higher than that of the other competing

methods. When $\theta = 1.6$, this difference becomes even more pronounced. At $n = 75$, the power of $\Delta_T(0.4, 3.25)$ exceeds 91%, whereas $\Delta_T(0, -1)$ and $\gamma^*_{j=0.25}$ are clearly lower. Even under stronger departures, such as $\theta = 1.8$ and 2.0, $\Delta_T(0.4, 3.25)$ still attains the highest or jointly highest power for most sample sizes.

Table 3. Estimated power (%) under the Gamma alternatives.

θ	n	$\Delta_T(a,b)$				Benchmarks	
		$(0.4, 3.25)$	$(0, 0)$	$(0, -1)$	$(0.5, -0.5)$	$\gamma^*_{j=0.25}$	T_1
1.4	5	9.75	9.69	9.45	9.62	9.76	9.84
	15	20.54	18.81	16.67	18.43	16.94	18.55
	30	34.20	30.61	24.65	29.64	26.51	30.81
	50	51.12	45.07	34.98	43.42	38.21	45.03
	75	68.00	60.75	46.09	58.60	50.82	60.09
	100	79.08	71.66	55.13	69.43	61.19	71.77
1.6	5	12.53	12.36	11.88	12.24	12.32	12.53
	15	31.26	28.41	24.12	27.62	24.96	28.06
	30	55.07	49.38	39.07	47.66	42.49	49.62
	50	77.06	70.05	55.38	68.01	60.47	70.00
	75	91.11	85.58	70.34	83.83	76.03	85.17
	100	96.89	93.65	80.68	92.42	86.48	93.71
1.8	5	15.38	15.20	14.48	15.04	15.02	15.39
	15	42.95	38.87	32.41	37.62	33.80	38.47
	30	72.90	66.39	53.30	64.42	57.76	66.62
	50	91.79	86.71	72.87	85.03	78.09	86.68
	75	98.45	96.44	86.76	95.55	91.09	96.28
	100	99.77	99.18	93.70	98.83	96.72	99.19
2.0	5	18.28	17.98	16.98	17.82	17.74	18.23
	15	54.32	49.36	41.10	47.90	43.12	48.95
	30	85.37	79.51	65.86	77.53	70.66	79.68
	50	97.62	95.08	84.86	94.07	89.30	95.06
	75	99.83	99.39	94.86	99.12	97.26	99.35
	100	99.99	99.93	98.36	99.89	99.37	99.93

Tables 2 and 3 show that the power of $\Delta_T(0, 0)$ is broadly comparable to, and often slightly higher than, that of T_1. This comparison indicates that the unified centering and studentization preserves the competitive performance of the classical HP statistic, while allowing different weight choices to be compared on a common standardized scale.

In addition to the comparison among different parameter combinations, the tabulated results also reflect the role of centering and studentization. Since $T_w(F_n; 0, 0)$ reduces to the classical HP statistic T_1, the difference between $\Delta_T(0, 0)$ and T_1 mainly arises from standardization and finite-sample drift correction.

3.2. Real data analysis

To assess the performance of the proposed method on real data, we selected six classical datasets arising from different application backgrounds, denoted by Elec_Y [27], Fiber [28], Precip [29], Carbon [30], Relief [31], and Wind [32]. These datasets come from various application areas, with sample sizes equal to 30, 30, 30, 63, 20, and 23, respectively. For ease of reference, Table 4 summarizes the background information of these six datasets, while Table 5 reports the corresponding test statistics, p-values, and final decisions for all methods.

Table 4. Summary of the real datasets analyzed in the study.

Dataset	Description	Application area	n
Elec_Y	Failure times of Type Y electrical insulation	Engineering	30
Fiber	Tensile strength measurements of polyester fibers	Materials science	30
Precip	March precipitation observations in the Minneapolis–St. Paul area	Meteorology	30
Carbon	Tensile strength measurements of carbon fibers at 10 mm gauge length	Materials science	63
Relief	Relief times following analgesic treatment	Medical data	20
Wind	Daily wind speed observations	Meteorology	23

Table 5. Real-data analysis for the proposed $\Delta_T(a, b)$ procedures and benchmark tests.

Dataset	n	Proposed $\Delta_T(a,b)$ family								Benchmark methods				Decision
		(0.4, 3.25)		(0, 0)		(0, −1)		(0.5, −0.5)		$\gamma^*_{j=0.25}$		T_1		
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value	
Elec_Y	30	0.358	0.3603	0.563	0.2867	0.578	0.2815	0.575	0.2827	0.601	0.2740	0.879	0.2847	Fail to reject H_0
Fiber	30	1.149	0.1253	1.435	0.0756	1.331	0.0916	1.466	0.0713	1.449	0.0737	1.751	0.0722	Fail to reject H_0
Precip	30	4.152	< 0.0001	3.139	0.0008	1.888	0.0295	2.957	0.0016	2.325	0.0100	3.455	0.0006	Reject H_0
Carbon	63	15.429	< 0.0001	10.770	< 0.0001	6.383	< 0.0001	10.033	< 0.0001	7.820	< 0.0001	10.989	< 0.0001	Reject H_0
Relief	20	7.074	< 0.0001	4.566	< 0.0001	2.369	0.0089	4.178	< 0.0001	3.047	0.0012	4.953	< 0.0001	Reject H_0
Wind	23	5.393	< 0.0001	3.365	0.0004	1.921	0.0274	3.082	0.0010	2.364	0.0090	3.727	0.0002	Reject H_0

Notes: For $\Delta_T(a, b)$, decisions are based on empirical critical values, and the reported p-values are asymptotic right-tail references.

For each dataset, we computed $\Delta_T(0.4, 3.25)$, $\Delta_T(0, 0)$, $\Delta_T(0, -1)$, and $\Delta_T(0.5, -0.5)$, and compared them with $\gamma^*_{j=0.25}$ and T_1. The significance level was fixed at $\alpha = 0.05$. For $\Delta_T(a, b)$, the final decision is made by comparing the test statistic with the corresponding empirical critical value, while the reported p-values are based on the standard normal approximation and are included for reference. For the benchmark methods, we report the corresponding test statistics, p-values, and decisions from the original procedures.

From Table 5, for the Elec_Y and Fiber datasets, the reported decisions for all methods are to fail to reject the null hypothesis of exponentiality at the 5% significance level. For the tests considered in this paper, these decisions are based on the empirical critical-value calibration described above, and the corresponding asymptotic right-tail p-values are also greater than 0.05. This indicates that no sufficiently strong evidence of NBUE-type departure is observed in these two datasets.

In contrast, for the Precip, Carbon, Relief, and Wind datasets, the reported decisions for all methods are to reject the null hypothesis at the 5% level, indicating significant NBUE-type departures from the

exponential model. A further comparison of the magnitudes of the statistics reveals that $\Delta_T(0.4, 3.25)$ always attains the largest value within each of these four datasets, followed in most cases by $\Delta_T(0, 0)$, whereas the other two weight choices tend to produce smaller values overall. For example, for the Carbon dataset, the statistics of $\Delta_T(0.4, 3.25)$ and $\Delta_T(0, 0)$ are 15.429 and 10.770, respectively. This pattern suggests that, when the departure is pronounced, the choice of weights affects the relative magnitude of the test statistic.

It should also be noted that cross-method comparison needs to be interpreted cautiously. For the four datasets Precip, Carbon, Relief, and Wind, $\Delta_T(0.4, 3.25)$ gives the largest statistic among the four parameter settings considered in the proposed framework. However, the benchmark procedures $\gamma^*_{j=0.25}$ and T_1 are based on different constructions and calibrations, so their raw statistic values are not directly comparable with those of $\Delta_T(a, b)$. Overall, these results mainly suggest that $(a, b) = (0.4, 3.25)$ is a favorable choice among the parameter settings considered within the proposed framework, while remaining consistent with the benchmark methods in terms of the final decision.

4. Conclusions

This paper studies tests of exponentiality against NBUE alternatives through a unified weighted-functional formulation. By expressing Hollander–Proschan type procedures as weighted aggregations of a common mean residual life contrast, we obtain a structural framework that clarifies the relationships among a range of existing methods. Within this framework, we construct a centered and studentized statistic $\Delta_T(a, b)$, which enables comparison across different weighting schemes on a common inferential scale.

We derive equivalent representations of the functional, obtain an explicit centering term under the exponential null, and establish an asymptotic normal approximation for the standardized statistic. These results facilitate direct implementation and place different weight choices within a unified framework. The Pitman analysis highlights how the choice of weight affects sensitivity to different alternatives, offering guidance for parameter selection within the working class. Simulation results indicate that the empirical critical values gradually approach their normal-reference counterparts as the sample size increases, and that the proposed procedures achieve competitive performance across several alternatives.

The current development is restricted to complete lifetime data. Extending the framework to other aging classes [33, 34] and to right-censored data remains of interest [35]. In the censored setting, a natural approach is to consider Kaplan–Meier plug-in versions of the functional. However, the resulting statistic becomes a nonlinear functional of the survival curve and is no longer reducible to a finite-order statistic or a simple closed-form expression. This introduces additional challenges for asymptotic analysis and variance estimation, particularly in the tail region. We leave a detailed investigation of this extension to future work.

Author contributions

Juan Ding: Conceptualization, Methodology, Writing-original draft, Formal analysis, Funding acquisition; Chi Zhou: Methodology, Writing-original draft; Wenxin Zhou: Methodology, Formal analysis; Anqi Xia: Formal analysis; Wenjun Xiong: Conceptualization, Methodology, Funding

acquisition, Writing-review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was supported by the fundamental research funds for the central universities (B240201095).

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. C. D. Lai, M. Xie, Concepts and applications of stochastic aging in reliability, In: *Handbook of Reliability Engineering*, London: Springer, 2003, 165–180. http://doi.org/10.1007/1-85233-841-5_9
2. R. E. Barlow, F. Proschan, *Statistical Theory of Reliability and Life Testing: Probability Models*, New York: Holt, Rinehart and Winston, 1975.
3. F. Guess, F. Proschan, Mean residual life: Theory and applications, In: *Handbook of Statistics, Vol. 7: Quality Control and Reliability*, Amsterdam: Elsevier, **7** (1988), 215–224. http://doi.org/10.1016/S0169-7161(88)07014-2
4. C. D. Lai, M. Xie, *Stochastic Ageing and Dependence for Reliability*, New York: Springer, 2006. http://doi.org/10.1007/0-387-34232-X
5. V. Poynor, A. Kottas, Nonparametric Bayesian inference for mean residual life functions in survival analysis, *Biostatistics*, **20** (2019), 240–255. http://doi.org/10.1093/biostatistics/kxx075
6. R. Ahmadi, A bivariate process-based mean residual lifetime model for maintenance and inspection planning, *Comput. Indust. Eng.*, **163** (2022), 107792. http://doi.org/10.1016/j.cie.2021.107792
7. R. Ahmadi, I. T. Castro, L. Bautista, Reliability modeling and maintenance planning for a parallel system with respect to the state-dependent mean residual time, *J. Oper. Res. Soc.*, **75** (2024), 297–313. http://doi.org/10.1080/01605682.2023.2194316
8. P. Do, C. Bérenguer, Residual life-based importance measures for predictive maintenance decision-making, *Proc. Instit. Mech. Eng., Part O: J. Risk Reliab.*, **236** (2022), 98–113. http://doi.org/10.1177/1748006X211028112
9. M. Hollander, F. Proschan, Tests for the mean residual life, *Biometrika*, **62** (1975), 585–593. http://doi.org/10.1093/biomet/62.3.585

10. B. Ebner, The test of exponentiality based on the mean residual life function revisited, *J. Nonparam. Statist.*, **35** (2023), 601–621. http://doi.org/10.1080/10485252.2023.2178831

11. H. L. Koul, Testing for new is better than used in expectation, *Commun. Statist.–Theory Meth.*, **7** (1978), 685–701. http://doi.org/10.1080/03610927808827658

12. W. de Souza Borges, F. Proschan, J. Rodrigues, A simple test for new better than used in expectation, *Commun. Statist.–Theory Meth.*, **13** (1984), 3217–3223. http://doi.org/10.1080/03610928408828888

13. M. Z. Anis, M. Mitra, A generalized Hollander–Proschan type test for NBUE alternatives, *Statist. Probab. Lett.*, **81** (2011), 126–132. http://doi.org/10.1016/j.spl.2010.10.002

14. S. K. Kattumannil, D. C. Mathew, A Gini-based exact test for exponentiality against NBUE alternatives with censored observations, *J. Nonparam. Statist.*, **27** (2015), 503–515. http://doi.org/10.1080/10485252.2015.1077242

15. Y. Q. Chen, N. P. Jewell, X. Lei, S. C. Cheng, Semiparametric estimation of proportional mean residual life model in presence of censoring, *Biometrics*, **61** (2005), 170–178. http://doi.org/10.1111/j.0006-341X.2005.030224.x

16. L. Sun, Z. Zhang, A class of transformed mean residual life models with censored survival data, *J. Amer. Statist. Assoc.*, **104** (2009), 803–815. http://doi.org/10.1198/jasa.2009.0130

17. S. K. Kattumannil, P. Anisha, A simple non-parametric test for decreasing mean time to failure, *Statist. Papers*, **60** (2019), 73–87. http://doi.org/10.1007/s00362-016-0827-y

18. S. Ghosh, M. Mitra, A new test for exponentiality against HNBUE alternatives, *Commun. Statist.–Theory Meth.*, **49** (2020), 27–43. http://doi.org/10.1080/03610926.2018.1528370

19. T. Lando, Nonparametric inference for NBUE distributions based on the TTT transform, *Statist. Probab. Lett.*, **212** (2024), 110157. http://doi.org/10.1016/j.spl.2024.110157

20. M. Z. Anis, K. Basu, Tests for exponentiality against NBUE alternatives: A Monte Carlo comparison, *J. Statist. Comput. Simul.*, **84** (2014), 231–247. http://doi.org/10.1080/00949655.2012.704517

21. J. S. Allison, L. Santana, N. Smit, I. J. H. Visagie, An "apples to apples" comparison of various tests for exponentiality, *Comput. Statist.*, **32** (2017), 1241–1283. http://doi.org/10.1007/s00180-017-0733-3

22. E. O. Ossai, M. S. Madukaife, A. V. Oladugba, A review of tests for exponentiality with Monte Carlo comparisons, *J. Appl. Statist.*, **49** (2022), 1277–1304. http://doi.org/10.1080/02664763.2020.1854202

23. E. E. A. A. Aly, Tests for monotonicity properties of the mean residual life function, *Scand. J. Statist.*, **17** (1990), 189–200.

24. F. Belzunce, J. F. Pinar, J. M. Ruiz, On testing the dilation order and HNBUE alternatives, *Ann. Instit. Statist. Math.*, **57** (2005), 803–815. http://doi.org/10.1007/BF02915440

25. T. P. Morris, I. R. White, M. J. Crowther, Using simulation studies to evaluate statistical methods, *Statist. Med.*, **38** (2019), 2074–2102. http://doi.org/10.1002/sim.8086

26. C. Williams, Y. Yang, M. Lagisz, K. Morrison, L. Ricolfi, D. I. Warton, et al., Transparent reporting items for simulation studies evaluating statistical methods: Foundations for reproducibility and reliability, *Meth. Ecology Evol.*, **15** (2024), 1926–1939. http://doi.org/10.1111/2041-210X.14415

27. N. Alsadat, E. M. Almetwally, M. Elgarhy, H. Ahmad, G. A. Marei, Bayesian and non-Bayesian analysis with MCMC algorithm of stress-strength for a new two parameters lifetime model with applications, *AIP Adv.*, **13** (2023), 095203. http://doi.org/10.1063/5.0167295

28. A. I. Al-Omari, S. A. Dobbah, On the mixture of Shanker and gamma distributions with applications to engineering data, *J. Rad. Res. Appl. Sci.*, **16** (2023), 100533. http://doi.org/10.1016/j.jrras.2023.100533

29. R. S. Gomaa, E. A. Hebeshy, M. M. El Genidy, B. S. El-Desouky, Alpha-power of the power Ailamujia distribution: properties and applications, *J. Statist. Appl. Probab.*, **12** (2023), 701–723. http://doi.org/10.18576/jsap/120231

30. D. Kundu, R. D. Gupta, Estimation of $P[Y < X]$ for Weibull distributions, *IEEE Trans. Reliab.*, **55** (2006), 270–280. http://doi.org/10.1109/TR.2006.874918

31. R. A. Ganaie, V. Rajagopalan, S. R. Nair, M. Kanan, A new generalization of power garima distribution with applications in blood cancer and relief times, *J. Statist. Appl. Probab.*, **12** (2023), 371–394. http://doi.org/10.18576/jsap/120205

32. M. G. M. Ghazal, M. M. Hasaballah, R. M. EL-Sagheer, O. S. Balogun, M. E. Bakr, Bayesian analysis using joint progressive Type-II censoring scheme, *Symmetry*, **15** (2023), 1884. http://doi.org/10.3390/sym15101884

33. S. Ghosh, M. Mitra, A weighted integral approach to testing against HNBUE alternatives, *Statist. Probab. Lett.*, **129** (2017), 58–64. http://doi.org/10.1016/j.spl.2017.05.003

34. E. Lorenzo, G. Malla, H. Mukerjee, A new test for decreasing mean residual lifetimes, *Commun. Statist.–Theory Meth.*, **47** (2018), 2805–2812. http://doi.org/10.1080/03610926.2014.985841

35. Y. Q. Chen, S. Cheng, Semiparametric regression analysis of mean residual life with censored survival data, *Biometrika*, **92** (2005), 19–29. http://doi.org/10.1093/biomet/92.1.19

A. Appendix

Appendix A. Equivalent representations of $T_w(F; a, b)$

For $(a, b) \in \mathcal{A}$, recall

$$T_w(F; a, b) = \int_0^\infty w(F(t); a, b)\,[m(0) - m(t)]\,dF(t), \qquad w(x; a, b) = (1 - x)(a + 1 - x)^b. \tag{A.1}$$

Assume that F is absolutely continuous with density f and finite mean $\mu = \int_0^\infty \bar{F}(u)\,du < \infty$. In addition, assume that the integrals appearing below are finite so that Tonelli's theorem and the change of variable $v = F(t)$ are applicable. For the second representation, we further assume that

$$t\,W(F(t); a, b)\,\bar{F}(t) \to 0, \qquad t \to \infty.$$

Under these conditions, the following equivalent representations are valid.

Since

$$m(t) = \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(u)\, du,$$

we have

$$m(0) - m(t) = \mu - \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(u)\, du.$$

Substituting this into (A.1) gives

$$T_w(F; a, b) = \mu \int_0^\infty w(F(t); a, b)\, dF(t) - \int_0^\infty \frac{w(F(t); a, b)}{\bar{F}(t)} \left(\int_t^\infty \bar{F}(u)\, du \right) dF(t).$$

Using $\mu = \int_0^\infty \bar{F}(u)\, du$, the change of variable $v = F(t)$, and Tonelli's theorem, we obtain

$$\begin{aligned} T_w(F; a, b) &= \left(\int_0^\infty \bar{F}(u)\, du \right) \left(\int_0^1 w(v; a, b)\, dv \right) - \int_0^\infty \bar{F}(u) \left(\int_0^u \frac{w(F(t); a, b)}{\bar{F}(t)}\, dF(t) \right) du \\ &= \int_0^\infty \bar{F}(u) \left[\int_0^1 w(v; a, b)\, dv - \int_0^{F(u)} \frac{w(v; a, b)}{1 - v}\, dv \right] du. \end{aligned}$$

Therefore,

$$T_w(F; a, b) = \int_0^\infty W(F(t); a, b)\, \bar{F}(t)\, dt,$$

where

$$W(x; a, b) = \int_0^1 w(v; a, b)\, dv - \int_0^x \frac{w(v; a, b)}{1 - v}\, dv.$$

To obtain the second representation, define

$$A(t) = W(F(t); a, b)\, \bar{F}(t).$$

Then

$$A'(t) = W'(F(t); a, b) f(t) \bar{F}(t) - W(F(t); a, b) f(t).$$

Since

$$W'(x; a, b) = -\frac{w(x; a, b)}{1 - x}, \qquad 1 - F(t) = \bar{F}(t),$$

it follows that

$$A'(t) = -\{w(F(t); a, b) + W(F(t); a, b)\} f(t) = -J_w(F(t); a, b) f(t),$$

where

$$J_w(x; a, b) = w(x; a, b) + W(x; a, b).$$

If, in addition, $tA(t) \to 0$ as $t \to \infty$, then integration by parts yields

$$\begin{aligned} T_w(F; a, b) &= \int_0^\infty A(t)\, dt = \Big[tA(t) \Big]_0^\infty - \int_0^\infty tA'(t)\, dt \\ &= \int_0^\infty t\, J_w(F(t); a, b)\, f(t)\, dt = \int_0^\infty t\, J_w(F(t); a, b)\, dF(t). \end{aligned}$$

This establishes the two equivalent representations stated in (2.10).

Appendix B. Closed form of $\zeta_n(a,b)$ under H_0

Write the empirical functional in the spacing form

$$T_w(F_n;a,b)=\sum_{i=1}^{n} W\!\left(\frac{i-1}{n};a,b\right)\frac{n-i+1}{n}(x_{(i)}-x_{(i-1)}),\qquad X_{(0)}:=0, \tag{A.2}$$

where $x_{(1)}\le\cdots\le x_{(n)}$ are the order statistics. Under $H_0:\bar{F}(x)=\exp(-x/\mu)$, the exponential spacings satisfy

$$\mathbb{E}_0[x_{(i)}-x_{(i-1)}]=\frac{\mu}{n-i+1},\qquad i=1,\dots,n.$$

Substituting this into (A.2) gives

$$\mathbb{E}_0[T_w(F_n;a,b)]=\sum_{i=1}^{n} W\!\left(\frac{i-1}{n};a,b\right)\frac{n-i+1}{n}\cdot\frac{\mu}{n-i+1}=\frac{\mu}{n}\sum_{i=1}^{n} W\!\left(\frac{i-1}{n};a,b\right).$$

Therefore,

$$\zeta_n(a,b)=\mathbb{E}_0\Big[\frac{T_w(F_n;a,b)}{\mu}\Big]=\frac{1}{n}\sum_{i=1}^{n} W\!\left(\frac{i-1}{n};a,b\right). \tag{A.3}$$

This is the centering term used in $\Delta_T(a,b)$.

Appendix C. Centering and asymptotic normality of $\Delta_T(a,b)$

Let $T_w(F_n;a,b)$ be the empirical version of (2.10). Under the exponential null hypothesis with mean $\mu>0$, one has $T_w(F;a,b)=0$. By Theorem 1 and the scale equivariance discussed in the main text, it is sufficient to work with the unit-mean exponential baseline F_0. Hence

$$\frac{\sqrt{n}\,T_w(F_n;a,b)}{\mu\,\sigma(J_w,F_0;a,b)}\ \xrightarrow{d}\ N(0,1).$$

Next, by Appendix B,

$$\mathbb{E}_0\left[\frac{T_w(F_n;a,b)}{\mu}\right]=\frac{1}{n}\sum_{i=1}^{n} W\!\left(\frac{i-1}{n};a,b\right)=\zeta_n(a,b),$$

where $\mathbb{E}_0$ denotes expectation under H_0. Since $\int_0^1 W(x;a,b)\,dx=0$ and $(a,b)\in\mathcal{A}$, standard Riemann-sum arguments (with Stirling's formula in the logarithmic case $a=0$, $b=-1$) yield $\sqrt{n}\,\zeta_n(a,b)\to 0$. Hence the centering term is asymptotically negligible, and therefore

$$\frac{\sqrt{n}\,\{T_w(F_n;a,b)/\mu-\zeta_n(a,b)\}}{\sigma(J_w,F_0;a,b)}\ \xrightarrow{d}\ N(0,1).$$

Since $\overline{X}_n/\mu\to_p 1$, Slutsky's theorem yields

$$\Delta_T(a,b)=\frac{\sqrt{n}\,\{T_w(F_n;a,b)/\overline{X}_n-\zeta_n(a,b)\}}{\sigma(J_w,F_0;a,b)}\ \xrightarrow{d}\ N(0,1).$$

This establishes the asymptotic normal calibration stated in the main text.

Appendix D: Derivation of the local slope for the Weibull family

We illustrate the evaluation of the local slope $\lambda(a, b; F_\theta)$ in (2.20) for the Weibull family

$$\bar{F}_\theta(x) = \exp(-x^\theta), \qquad x > 0,\ \theta > 0,$$

with baseline $\theta_0 = 1$, so that $F_0(x) = 1 - e^{-x}, \qquad x \geq 0$.

Using the representation $T_w(F_\theta; a, b) = \int_0^\infty W\{F_\theta(t); a, b\}\, \bar{F}_\theta(t)\, dt$, differentiate under the integral sign. Since

$$\frac{\partial}{\partial\theta}\bar{F}_\theta(t) = -\, t^\theta \log(t)\, \bar{F}_\theta(t), \qquad \frac{\partial}{\partial\theta}F_\theta(t) = t^\theta \log(t)\, \bar{F}_\theta(t),$$

and

$$W'(x; a, b) = -\frac{w(x; a, b)}{1 - x},$$

it follows that

$$\frac{\partial}{\partial\theta}\Big[W\{F_\theta(t); a, b\}\, \bar{F}_\theta(t)\Big] = -\, J_w\{F_\theta(t); a, b\}\, \frac{\partial}{\partial\theta}F_\theta(t),$$

where $J_w(x; a, b) = w(x; a, b) + W(x; a, b)$. Hence

$$\frac{\partial}{\partial\theta}T_w(F_\theta; a, b) = -\int_0^\infty J_w\{F_\theta(t); a, b\}\, t^\theta \log(t)\, \bar{F}_\theta(t)\, dt. \tag{A.4}$$

Evaluating at $\theta = 1$ gives

$$\frac{\partial}{\partial\theta}T_w(F_\theta; a, b)\Big|_{\theta=1} = -\int_0^\infty J_w\{F_0(t); a, b\}\, t \log(t)\, e^{-t}\, dt. \tag{A.5}$$

Now consider the scale-normalized functional $T_w(F_\theta; a, b)/\mu(\theta)$, where $\mu(\theta) = \mathbb{E}_\theta(X)$. Since $T_w(F_0; a, b) = 0$ under the exponential baseline and $\mu(1) = 1$, the quotient rule yields

$$\frac{\partial}{\partial\theta}\left\{\frac{T_w(F_\theta; a, b)}{\mu(\theta)}\right\}\Big|_{\theta=1} = \frac{\partial}{\partial\theta}T_w(F_\theta; a, b)\Big|_{\theta=1}. \tag{A.6}$$

Combining (A.5) and (A.6), we obtain

$$\frac{\partial}{\partial\theta}\left\{\frac{T_w(F_\theta; a, b)}{\mu(\theta)}\right\}\Big|_{\theta=1} = -\int_0^\infty J_w\{F_0(t); a, b\}\, t \log(t)\, e^{-t}\, dt.$$

Therefore, the local slope for the Weibull family is

$$\lambda(a, b; F_\theta) = -\frac{\displaystyle\int_0^\infty J_w\{F_0(t); a, b\}\, t \log(t)\, e^{-t}\, dt}{\sigma(J_w, F_0; a, b)}. \tag{A.7}$$

For numerical evaluation, one may equivalently use the transformation $u = F_0(t) = 1 - e^{-t}$, which gives

$$\lambda(a, b; F_\theta) = -\frac{\displaystyle\int_0^1 J_w(u; a, b)\, [-\log(1 - u)]\, \log[-\log(1 - u)]\, du}{\sigma(J_w, F_0; a, b)}. \tag{A.8}$$

Formulae (A.7) and (A.8) are used to evaluate the Weibull local slope in Figure 1.

Appendix E. Finite-sample critical values for selected $\Delta_T(a, b)$ statistics

This appendix reports representative simulated finite-sample critical values for selected $\Delta_T(a, b)$ statistics under H_0.

Table 6. Critical values for $\Delta_T(0.4, 3.25)$.

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.50$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.99$
2	-2.1228	-1.9490	-1.7309	0.0015	1.7347	1.9497	2.1235
3	-2.1749	-1.8010	-1.5198	-0.1639	1.7970	2.2620	2.8845
4	-2.1101	-1.7194	-1.4582	-0.1457	1.7036	2.2402	3.0874
5	-2.0846	-1.6820	-1.4128	-0.1369	1.6249	2.1682	3.1247
6	-2.0705	-1.6546	-1.3826	-0.1262	1.5752	2.1087	3.1002
7	-2.0663	-1.6389	-1.3653	-0.1200	1.5353	2.0584	3.0514
8	-2.0628	-1.6280	-1.3516	-0.1136	1.5097	2.0260	3.0198
9	-2.0665	-1.6183	-1.3383	-0.1075	1.4896	1.9935	2.9671
10	-2.0647	-1.6124	-1.3292	-0.1042	1.4713	1.9697	2.9426
11	-2.0681	-1.6084	-1.3223	-0.1031	1.4555	1.9479	2.9150
12	-2.0717	-1.6019	-1.3164	-0.0983	1.4441	1.9277	2.8904
13	-2.0762	-1.5992	-1.3110	-0.0919	1.4350	1.9171	2.8699
14	-2.0784	-1.5974	-1.3083	-0.0886	1.4280	1.9046	2.8486
15	-2.0790	-1.5940	-1.3015	-0.0853	1.4211	1.8925	2.8250
20	-2.0971	-1.5909	-1.2943	-0.0779	1.3902	1.8503	2.7490
25	-2.1066	-1.5880	-1.2834	-0.0660	1.3748	1.8230	2.7010
30	-2.1217	-1.5894	-1.2796	-0.0596	1.3645	1.8071	2.6650
35	-2.1353	-1.5886	-1.2795	-0.0590	1.3513	1.7842	2.6313
40	-2.1336	-1.5885	-1.2765	-0.0536	1.3464	1.7757	2.6125
45	-2.1490	-1.5923	-1.2750	-0.0498	1.3399	1.7640	2.5883
50	-2.1570	-1.5939	-1.2731	-0.0484	1.3388	1.7564	2.5762
55	-2.1594	-1.5893	-1.2696	-0.0447	1.3328	1.7492	2.5645
60	-2.1649	-1.5950	-1.2733	-0.0431	1.3311	1.7484	2.5606
65	-2.1655	-1.5922	-1.2708	-0.0412	1.3276	1.7428	2.5524
70	-2.1774	-1.5952	-1.2727	-0.0412	1.3256	1.7371	2.5342
75	-2.1759	-1.5975	-1.2716	-0.0386	1.3236	1.7317	2.5268
80	-2.1806	-1.5970	-1.2705	-0.0372	1.3203	1.7298	2.5162
85	-2.1931	-1.5969	-1.2701	-0.0372	1.3198	1.7282	2.5120
90	-2.1903	-1.5998	-1.2710	-0.0362	1.3173	1.7231	2.5099
95	-2.1886	-1.5988	-1.2687	-0.0345	1.3148	1.7192	2.4960
100	-2.1931	-1.6012	-1.2698	-0.0327	1.3181	1.7219	2.5012

Notes. Entries report the simulated α-quantiles of $\Delta_T(0.4, 3.25)$.

Table 7. Critical values for $\Delta_T(0.5, -0.5)$.

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.50$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.99$
2	-1.0549	-0.9685	-0.8601	0.0007	0.8620	0.9688	1.0552
3	-1.6061	-1.2764	-1.0279	0.0196	0.9966	1.2274	1.5365
4	-1.8477	-1.3715	-1.0756	0.0166	1.0517	1.3155	1.7383
5	-1.9725	-1.4282	-1.1167	0.0162	1.0947	1.3712	1.8455
6	-2.0480	-1.4670	-1.1432	0.0198	1.1236	1.4133	1.9149
7	-2.1007	-1.4958	-1.1648	0.0174	1.1427	1.4399	1.9595
8	-2.1306	-1.5198	-1.1815	0.0173	1.1593	1.4626	2.0029
9	-2.1596	-1.5330	-1.1915	0.0156	1.1707	1.4779	2.0250
10	-2.1875	-1.5459	-1.2008	0.0178	1.1827	1.4954	2.0527
11	-2.2007	-1.5600	-1.2113	0.0134	1.1879	1.5046	2.0714
12	-2.2160	-1.5677	-1.2174	0.0132	1.1946	1.5149	2.0937
13	-2.2246	-1.5749	-1.2193	0.0139	1.2019	1.5256	2.1072
14	-2.2391	-1.5826	-1.2285	0.0139	1.2083	1.5329	2.1241
15	-2.2391	-1.5848	-1.2276	0.0133	1.2115	1.5373	2.1301
20	-2.2768	-1.6072	-1.2456	0.0105	1.2268	1.5620	2.1715
25	-2.2906	-1.6129	-1.2505	0.0123	1.2378	1.5754	2.1974
30	-2.3061	-1.6201	-1.2567	0.0131	1.2456	1.5871	2.2125
35	-2.3163	-1.6288	-1.2644	0.0089	1.2491	1.5945	2.2220
40	-2.3079	-1.6276	-1.2649	0.0090	1.2534	1.5964	2.2334
45	-2.3190	-1.6334	-1.2669	0.0093	1.2549	1.5988	2.2408
50	-2.3212	-1.6323	-1.2680	0.0063	1.2573	1.6044	2.2461
55	-2.3204	-1.6341	-1.2697	0.0078	1.2601	1.6077	2.2546
60	-2.3298	-1.6388	-1.2717	0.0085	1.2629	1.6128	2.2601
65	-2.3197	-1.6342	-1.2700	0.0086	1.2639	1.6175	2.2659
70	-2.3318	-1.6418	-1.2752	0.0072	1.2651	1.6165	2.2641
75	-2.3229	-1.6413	-1.2762	0.0076	1.2625	1.6138	2.2691
80	-2.3292	-1.6387	-1.2723	0.0074	1.2652	1.6152	2.2699
85	-2.3296	-1.6376	-1.2726	0.0077	1.2651	1.6198	2.2774
90	-2.3320	-1.6411	-1.2763	0.0081	1.2670	1.6204	2.2735
95	-2.3302	-1.6364	-1.2747	0.0065	1.2651	1.6177	2.2710
100	-2.3301	-1.6395	-1.2744	0.0058	1.2702	1.6233	2.2808

Notes. Entries report the simulated α-quantiles of $\Delta_T(0.5, -0.5)$.

Table 8. Critical values for $\Delta_T(0,0)$.

n	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.50$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.99$
2	-1.2003	-1.1021	-0.9793	0.0022	0.9808	1.1025	1.2003
3	-1.7202	-1.3702	-1.1072	-0.0003	1.1077	1.3697	1.7199
4	-1.9208	-1.4412	-1.1391	-0.0015	1.1332	1.4357	1.9206
5	-2.0110	-1.4764	-1.1666	-0.0003	1.1674	1.4748	2.0153
6	-2.0633	-1.5045	-1.1856	0.0026	1.1870	1.5060	2.0665
7	-2.1137	-1.5251	-1.1994	-0.0032	1.2005	1.5257	2.1045
8	-2.1326	-1.5422	-1.2107	0.0010	1.2104	1.5436	2.1390
9	-2.1549	-1.5515	-1.2182	0.0007	1.2202	1.5538	2.1596
10	-2.1758	-1.5646	-1.2258	0.0007	1.2238	1.5621	2.1787
11	-2.1909	-1.5711	-1.2307	0.0000	1.2322	1.5715	2.1910
12	-2.2070	-1.5808	-1.2387	0.0000	1.2349	1.5759	2.2056
13	-2.2093	-1.5809	-1.2390	-0.0010	1.2400	1.5844	2.2134
14	-2.2234	-1.5881	-1.2426	0.0005	1.2444	1.5882	2.2227
15	-2.2257	-1.5874	-1.2442	-0.0003	1.2419	1.5878	2.2249
20	-2.2515	-1.6040	-1.2518	-0.0021	1.2529	1.6029	2.2538
25	-2.2708	-1.6117	-1.2586	-0.0012	1.2583	1.6117	2.2621
30	-2.2775	-1.6208	-1.2656	0.0015	1.2617	1.6178	2.2796
35	-2.2905	-1.6202	-1.2628	-0.0014	1.2645	1.6200	2.2780
40	-2.2927	-1.6235	-1.2673	0.0000	1.2676	1.6245	2.2909
45	-2.2886	-1.6255	-1.2671	0.0008	1.2677	1.6257	2.2942
50	-2.2933	-1.6279	-1.2695	-0.0005	1.2710	1.6334	2.2942
55	-2.2987	-1.6292	-1.2713	0.0023	1.2732	1.6322	2.3051
60	-2.2930	-1.6282	-1.2728	0.0029	1.2751	1.6349	2.3068
65	-2.3035	-1.6326	-1.2728	0.0005	1.2732	1.6357	2.3005
70	-2.3093	-1.6358	-1.2761	0.0023	1.2728	1.6323	2.3049
75	-2.3060	-1.6333	-1.2743	-0.0005	1.2739	1.6343	2.3083
80	-2.3026	-1.6337	-1.2749	-0.0003	1.2739	1.6354	2.3070
85	-2.3038	-1.6364	-1.2749	0.0005	1.2760	1.6372	2.3078
90	-2.3173	-1.6358	-1.2743	-0.0030	1.2759	1.6358	2.3137
95	-2.3137	-1.6385	-1.2776	-0.0015	1.2753	1.6365	2.3113
100	-2.3165	-1.6392	-1.2765	0.0018	1.2771	1.6399	2.3152

Notes. Entries report the simulated α-quantiles of $\Delta_T(0,0)$.

AIMS Press