单位 Shiha 分布及其在工程和医学数据中的应用

F. A. Shiha

数学系，曼苏拉大学理学院，曼苏拉 35516，埃及 联系方式：电子邮箱：fshiha@mans.edu.eg

arXiv:2602.04400v1 [stat.ME] 4 Feb 2026

摘要

人们日益需要能够准确建模定义在单位区间上的数据的灵活统计分布。本文介绍了一种新的单位分布，称为单位 Shiha（USh）分布，该分布通过反指数变换由原始 Shiha（Sh）分布推导而来。USh 分布的概率密度函数具有足够的灵活性，能够对左偏和右偏数据进行建模；其风险率函数则能够捕捉多种失效率模式，包括递增型、浴盆形和 J 形。本文研究了所提出分布的若干统计性质，包括矩及相关测度、分位数函数、熵以及应力-强度可靠性。参数估计采用极大似然法进行，并通过模拟研究评估其性能。通过使用四个实际数据集，展示了 USh 分布的实际应用价值，并将其性能与若干著名的竞争性单位分布进行了比较。比较结果表明，与本研究中应用的竞争模型相比，所提出的模型对数据具有更好的拟合效果。

关键词：单位Shiha分布；可靠性；极大似然估计；数据分析。数学学科分类：60E05，62H12，62N05，62H10。

1 引言

近年来，人们对开发用于实际数据分析的新概率分布的兴趣日益浓厚。现有的大多数分布定义在区间 $(0,\infty)$ 上，因此不适合对取值范围介于 0 和 1 之间的有界数据进行建模。这类数据在实践中十分常见，包括标准化测量、比例和百分比。因此，显然需要定义在该单位区间上的简单而灵活的概率分布。构造单位分布的一种常见且有效的方法，是对取正值的随机变量应用适当的变换。诸如 $X = Y/(1+Y)$ 和 $X = e^{-Y}$ 这样的变换（其中 Y 定义在 $(0,\infty)$ 上）被广泛用于将变量的取值范围映射到单位区间。基于这一思路，文献中提出了许多单位分布，用于对各类应用领域中取值于 (0,1) 的数据进行建模。值得注意的例子包括单位X伽马分布 [10], 对数-XLindley分布 [6], 单位指数分布 [4], 单位Weibull分布 [15], 单位指数化 Weibull分布 [18], 单位Teissier分布 [13], 单位指数化半逻辑分布[11], 单位 Haq 分布 [3], 对数-余弦-幂单位分布 [17], 单位指数化 Lomax 分布 [8], 以及单位修正Burr-III分布 [2]。关于 (0,1) 上连续概率分布的全面综述见 [7]。

Shiha [19] 将 Shiha 分布（ShD）引入为 Exp(ω)、Exp(2ω) 和 Gamma(2,2ω) 分布的双参数混合分布。ShD 具有灵活的风险率模式，因此适用于对 $(0,\infty)$ 上各种寿命数据进行建模。受这些性质启发，本研究的主要目标是引入 ShD 的单位版本，称为单位Shiha分布（UShD），该分布通过变换 $\mathrm{X} = \mathrm{e}^{-Y}$ 得到，其中 Y 服从 Sh 分布。所提出的 UShD 将原始 ShD 的适用范围扩展至单位区间，同时保留了其灵活性和良好的解析性质。它能够对左偏和右偏数据进行建模，并捕捉多种失效率行为，包括递增型、浴盆形和 J形模式。这些特性使 UShD 成为一种实用且灵活的模型，可用于分析可靠性、医疗和工程应用中产生的有界数据。

本文其余部分安排如下。第 2 节介绍所提出的分布及其风险率函数和应力–强度可靠性函数。第 3 节介绍主要统计性质，包括矩及相关度量、分位数函数和熵。第 4 节和第 5 节分别讨论基于最大似然的参数估计和模拟研究。第 6 节通过四个实际数据集以及与若干竞争分布的比较，展示所提出模型的性能。最后，第 7 节给出结论。

2 单位Shiha分布

Shiha分布的概率密度函数（概率密度函数） [19] 定义为

$$f(y) = \frac{\omega}{\omega+3\eta}\left[\omega + (2\eta + 8\,\omega\,\eta\, y)e^{-\omega y}\right]e^{-\omega y}, \quad y \geq 0, \omega > 0, \eta > 0, \tag{2.1}$$

相应的累积分布函数（累积分布函数）为

$$F(y) = 1 - \frac{1}{\omega+3\eta}\left[\omega + (3\eta + 4\,\omega\,\eta\, y)e^{-\omega y}\right]e^{-\omega y}, \quad y \geq 0, \omega > 0, \eta > 0. \tag{2.2}$$

令 $\mathrm{X} = \mathrm{e}^{-Y}$，由于 $\mathrm{Y} \geq 0$，我们有 $\mathrm{X} \in (0,1]$ 且 $\mathrm{Y} = -\ln \mathrm{X}$，则变换后的（单位）分布的累积分布函数可推导为

$$\begin{aligned} F(x) &= P(X \leq x) = P(e^{-Y} \leq x) \\ &= P(Y \geq -\ln x) = 1 - P(Y < -\ln x) \\ &= 1 - F_Y(-\ln x), \end{aligned} \tag{2.3}$$

其对应的概率密度函数为

$$f(x) = f_Y(-\ln x)\left|\frac{dy}{dx}\right| = \frac{1}{x} f_Y(-\ln x). \tag{2.4}$$

根据 (2.4) 和 (2.3)，单位Shiha分布（USh(ω,η）的概率密度函数和累积分布函数分别为

$$f(x;\omega,\eta) = \frac{\omega}{\omega+3\eta}\left[\omega x^{\omega-1} + (2\eta - 8\omega\eta\ln x)x^{2\omega-1}\right], \qquad 0 < x \leq 1,, \ \omega > 0, \ \eta > 0. \tag{2.5}$$

$$F(x;\omega,\eta) = \frac{1}{\omega+3\eta}\left[\omega x^{\omega} + (3\eta - 4\omega\eta\ln x)x^{2\omega}\right], \qquad 0 < x < 1, \tag{2.6}$$

其中，当 $\mathrm{x} \leq 0$ 时，$\mathrm{F}_X(\mathrm{x}) = 0$；当 $\mathrm{x} \geq 1$ 时，$\mathrm{F}_X(\mathrm{x}) = 1$。注意，当 $\eta = 0$ 时，USh(ω,η) 分布可化为 Beta(ω, 1) 分布。

UShD 的风险率函数为

$$h(x) = \frac{\omega x^{\omega-1}\left[\omega + 2\eta x^{\omega}(1 - 4\omega\ln x)\right]}{\omega + 3\eta - x^{\omega}\left[\omega + \eta x^{\omega}(3 - 4\omega\ln x)\right]}, \quad 0 < x < 1. \tag{2.7}$$

图1展示了参数 ω 和 η 取不同值时 UShD 的概率密度函数和风险率函数。可以看出，USh 分布的概率密度函数足够灵活，能够对左偏数据和右偏数据进行建模。此外，风险率函数呈现出多种形状，包括浴盆形、递增形和 J 形。这种灵活性表明，所提出的分布能够对具有不同失效行为的多种寿命数据进行建模，因此适用于工程和医学领域的应用。

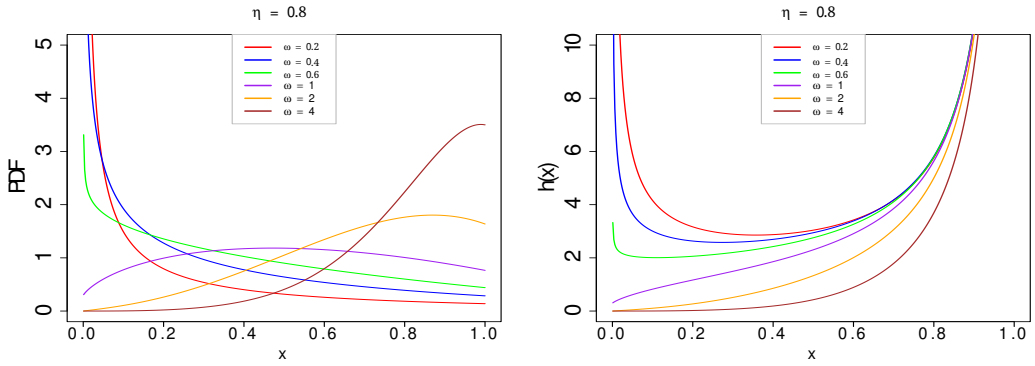


图1：不同参数值下 UShD 的概率密度函数和风险函数 s.

2.1 应力-强度可靠性

应力-强度可靠性模型是可靠性工程和统计分析中的一个基本概念，用于衡量系统强度超过所施加应力的概率；当应力超过强度时，系统便会发生故障。形式上，对于两个相互独立的随机变量 X_1和 X_2, ，其中 X_1 表示组件或系统的“强度”，X_2 表示施加于其上的“应力”，应力-强度可靠度定义为 $R = P(X_2 < X_1)$。

设 X_1 和 X_2 为相互独立的 USh随机变量，其参数分别为 (ω_1,η_1) 和 (ω_2,η_2)。则应力-强度可靠性参数 R 为

$$R = \int_0^1 f(x;\omega_1,\eta_1)F(x;\omega_2,\eta_2)\,dx.$$

将 (2.5) 和 (2.6) 代入，可得

$$R = \frac{\omega_1}{(\omega_1+3\eta_1)(\omega_2+3\eta_2)}\int_0^1 \left[\omega_1 x^{\omega_1-1} + (2\eta_1 - 8\omega_1\eta_1\ln x)\,x^{2\omega_1-1}\right] \times \left[\omega_2 x^{\omega_2} + (3\eta_2 - 4\omega_2\eta_2\ln x)\,x^{2\omega_2}\right] dx.$$

展开被积函数并逐项积分，可得如下形式的积分

$$\int_0^1 x^a\,dx = \frac{1}{a+1}, \qquad \int_0^1 x^a \ln x\,dx = -\frac{1}{(a+1)^2}, \qquad \int_0^1 x^a(\ln x)^2\,dx = \frac{2}{(a+1)^3}.$$

化简后，得到应力-强度可靠度

$$R = \frac{\omega_1}{(\omega_1+3\eta_1)(\omega_2+3\eta_2)}\left[\frac{\omega_1\omega_2}{\omega_1+\omega_2} + \omega_2\eta_1\left(\frac{2}{2\omega_1+\omega_2} + \frac{8\omega_1}{(2\omega_1+\omega_2)^2}\right) + \omega_1\eta_2\left(\frac{3}{\omega_1+2\omega_2} + \frac{4\omega_2}{(\omega_1+2\omega_2)^2}\right) + \eta_1\eta_2\left(\frac{6}{2\omega_1+2\omega_2} + \frac{24\omega_1+8\omega_2}{(2\omega_1+2\omega_2)^2} + \frac{64\omega_1\omega_2}{(2\omega_1+2\omega_2)^3}\right)\right]. \tag{2.8}$$

值得注意的是，随机变量 $X_i \sim \mathrm{USh}(\omega_i,\eta_i)$ 是由变换 $X_i = e^{-Y_i}$ 得到的，其中 $Y_i \sim \mathrm{Sh}(\omega_i,\eta_i)$。由于变换 $x \mapsto e^{-x}$ 严格单调递减，因此

$$P(X_2 < X_1) = P(Y_2 > Y_1) = 1 - P(Y_2 < Y_1).$$

因此，UShD 的应力-强度可靠度即为 ShD 的失效概率，也就是 $R_{USh} = 1 - R_{Sh}$。现在我们来研究 (2.8) 的一些特殊情形。

1. 令 $\eta_1 = \eta_2 = 0$，可得两个相互独立、服从 Beta(ω_1, 1) 和 Beta(ω_2, 1) 分布的随机变量的应力-强度可靠度：$R = \frac{\omega_1}{\omega_1+\omega_2}$；这也等同于两个相互独立的 Exp($\omega_1$) 和 Exp($\omega_2$) 随机变量的失效概率。 2. 令 $\eta_1 = \eta_2 = \eta$ 和 $\omega_1 = \omega_2 = \omega$，即 $X_1 \sim \mathrm{U\,Sh}(\omega,\eta)$ 和 $X_2 \sim \mathrm{U\,Sh}(\omega,\eta)$，可得 $R = \frac{1}{2}$。 3. 令 $\omega_1 = \omega_2 = \omega$ 和 $\eta_2 = 0$，即 $X_1 \sim \mathrm{U\,Sh}(\omega,\eta)$、$X_2 \sim \mathrm{Beta}(\omega, 1)$，可得 $R = \frac{9\omega+28\eta}{18(\omega+3\eta)}$。

---

3 统计性质

本节讨论所提出的 UShD 的主要统计性质。

3.1 矩及相关测度

定理1。单位Shiha分布随机变量的矩母函数为

$$M_X(t) = \frac{\omega}{\omega+3\eta}\sum_{\kappa=0}^{\infty}\frac{t^\kappa}{\kappa!}\left[\frac{\omega}{\kappa+\omega} + \frac{2\eta}{\kappa+2\omega} + \frac{8\omega\eta}{(\kappa+2\omega)^2}\right]. \tag{3.1}$$

证明。根据定义，$M_X(t) = E[e^{tX}]$。利用泰勒展开 $e^{tX} = \sum_{\kappa=0}^{\infty} t^\kappa X^\kappa/\kappa!$，并交换求和与期望的顺序，

$$M_X(t) = \sum_{\kappa=0}^{\infty}\frac{t^\kappa}{\kappa!}E[X^\kappa], \tag{3.2}$$

且 κ 阶矩为

$$\begin{aligned}\mu'_\kappa = E[X^\kappa] &= \int_0^1 x^\kappa f(x;\omega,\eta)\,dx \\ &= \frac{\omega}{\omega+3\eta}\left[\omega\int_0^1 x^{\kappa+\omega-1}dx + 2\eta\int_0^1 x^{\kappa+2\omega-1}dx - 8\omega\eta\int_0^1 (\ln x)x^{\kappa+2\omega-1}dx\right] \\ &= \frac{\omega}{\omega+3\eta}\left[\frac{\omega}{\kappa+\omega} + \frac{2\eta}{\kappa+2\omega} + \frac{8\omega\eta}{(\kappa+2\omega)^2}\right].\end{aligned} \tag{3.3}$$

□

在式（3.3）中分别令 $\kappa = 1$、2、3 和 4，即可得到 UShD 的前四阶矩，

$$\mu'_1 = E[X] = \frac{\omega}{\omega+3\eta}\left[\frac{\omega}{1+\omega} + \frac{2\eta}{1+2\omega} + \frac{8\omega\eta}{(1+2\omega)^2}\right], \tag{3.4}$$

$$\mu'_2 = E[X^2] = \frac{\omega}{\omega+3\eta}\left[\frac{\omega}{2+\omega} + \frac{\eta}{1+\omega} + \frac{2\omega\eta}{(1+\omega)^2}\right]. \tag{3.5}$$

$$\mu'_3 = E[X^3] = \frac{\omega}{\omega+3\eta}\left[\frac{\omega}{3+\omega} + \frac{2\eta}{3+2\omega} + \frac{8\omega\eta}{(3+2\omega)^2}\right] \tag{3.6}$$

$$\mu'_4 = E[X^4] = \frac{\omega}{\omega+3\eta}\left[\frac{\omega}{4+\omega} + \frac{\eta}{2+\omega} + \frac{2\omega\eta}{(2+\omega)^2}\right]. \tag{3.7}$$

表1列出了 UShD 在若干组 η和 ω 取值下的前四阶矩数值。可以看出，前四阶矩随 ω增大而增大，随 η增大而减小。利用前四阶矩，可求得偏度系数 (SK) 和峰度系数 (KU)

表1：不同模型参数取值下UShD的前四阶矩。

ω	$\eta = 0.1$				$\eta = 0.8$				$\eta = 1.5$			
	μ_1'	μ_2'	μ_3'	μ_4'	μ_1'	μ_2'	μ_3'	μ_4'	μ_1'	μ_2'	μ_3'	μ_4'
0.1	0.078	0.039	0.026	0.019	0.075	0.036	0.024	0.018	0.074	0.036	0.024	0.018
0.2	0.156	0.081	0.054	0.041	0.151	0.075	0.05	0.037	0.15	0.075	0.049	0.036
0.3	0.225	0.121	0.082	0.062	0.22	0.114	0.076	0.056	0.22	0.114	0.075	0.056
0.4	0.283	0.159	0.11	0.084	0.281	0.152	0.102	0.076	0.28	0.151	0.101	0.075
0.5	0.333	0.194	0.136	0.104	0.333	0.188	0.128	0.096	0.333	0.187	0.127	0.095
0.6	0.377	0.227	0.161	0.124	0.379	0.221	0.153	0.116	0.38	0.22	0.152	0.115
0.7	0.415	0.257	0.184	0.144	0.419	0.252	0.177	0.135	0.42	0.251	0.176	0.134
0.8	0.448	0.284	0.207	0.162	0.454	0.281	0.2	0.154	0.456	0.281	0.199	0.152
0.9	0.478	0.31	0.228	0.18	0.485	0.308	0.222	0.172	0.487	0.308	0.221	0.17
1	0.504	0.333	0.248	0.197	0.513	0.333	0.243	0.19	0.515	0.333	0.242	0.188
1.1	0.528	0.355	0.267	0.213	0.538	0.357	0.263	0.206	0.54	0.357	0.262	0.205
1.2	0.55	0.376	0.285	0.228	0.56	0.378	0.282	0.223	0.563	0.379	0.281	0.221
1.3	0.57	0.395	0.302	0.243	0.58	0.399	0.3	0.238	0.583	0.4	0.299	0.237
1.4	0.588	0.413	0.318	0.258	0.599	0.418	0.317	0.254	0.602	0.419	0.317	0.252
1.5	0.604	0.43	0.333	0.271	0.615	0.436	0.333	0.268	0.619	0.437	0.333	0.267
1.6	0.619	0.447	0.348	0.285	0.631	0.452	0.349	0.282	0.634	0.454	0.349	0.281
1.7	0.634	0.462	0.362	0.298	0.645	0.468	0.364	0.296	0.649	0.47	0.364	0.295
1.8	0.647	0.476	0.376	0.31	0.658	0.483	0.378	0.309	0.662	0.485	0.379	0.308
1.9	0.659	0.49	0.389	0.322	0.67	0.497	0.391	0.321	0.674	0.499	0.392	0.321
2	0.67	0.502	0.401	0.333	0.681	0.51	0.404	0.333	0.685	0.513	0.406	0.333
2.1	0.681	0.515	0.413	0.344	0.692	0.523	0.417	0.345	0.696	0.526	0.418	0.345
2.2	0.691	0.526	0.424	0.355	0.701	0.534	0.429	0.356	0.705	0.538	0.43	0.357
2.3	0.7	0.537	0.435	0.366	0.711	0.546	0.44	0.367	0.715	0.549	0.442	0.368
2.4	0.709	0.548	0.446	0.376	0.719	0.556	0.451	0.378	0.723	0.56	0.453	0.378
2.5	0.717	0.558	0.456	0.385	0.727	0.567	0.462	0.388	0.731	0.57	0.464	0.389
2.6	0.725	0.568	0.466	0.395	0.735	0.576	0.472	0.398	0.739	0.58	0.474	0.399
2.7	0.732	0.577	0.475	0.404	0.742	0.586	0.481	0.407	0.746	0.589	0.484	0.408
2.8	0.739	0.586	0.484	0.413	0.749	0.595	0.491	0.416	0.753	0.598	0.493	0.418
2.9	0.746	0.594	0.493	0.421	0.755	0.603	0.5	0.425	0.759	0.607	0.502	0.427
3	0.752	0.602	0.502	0.43	0.761	0.611	0.508	0.434	0.765	0.615	0.511	0.435

计算如下：

$$\mathrm{SK} = \frac{\mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3}{(\mathrm{Var}(Y))^{3/2}}, \qquad \mathrm{KU} = \frac{\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3{\mu_1'}^4}{(\mathrm{Var}(Y))^2}.$$

表2报告了UShD在ω与 η的各种组合下的方差、偏度和峰度值。从表2可以观察到，

1. 对于固 定的 ω，方差随 η 增大而减小，而当 $\omega \in (0, 1)$时偏度增大 .

2. 对于固定的η，

   • 方差在 ω取较小值时先增大，随后减小。

   • 偏度随着 ω 增大而减小，在 $\omega \in (0,1)$ 时为正，在$\omega > 1$时为负，

   • 当 $0 < \omega < 1$ 时，峰度会降低；而对于 $\omega \geq 1$，峰度则会增加。

图2进一步说明了参数的影响。这些结果表明，参数 ω主要决定分布偏度的方向：当 $\omega < 1$时产生右偏分布，当 $\omega > 1$时产生左偏分布，而在 $\omega = 1$时则接近对称。另一方面，参数 η控制离散程度和尾部行为；η的值越大，分布越集中，且尾部越厚。这种结构上的灵活性使UShD非常适合对具有不同程度非对称性和尾部厚度的各类实际数据进行建模。

表2： 不同 ω 和 η 取值下 UShD 的方差、偏度和峰度。

ω	$\eta = 0.1$			$\eta = 0.8$			$\eta = 1.5$			$\eta = 3$		
	Var(Y)	SK	KU	Var(Y)	SK	KU	Var(Y)	SK	KU	Var(Y)	SK	KU
0.1	0.033	2.968	11.66	0.031	3.066	12.389	0.031	3.075	12.456	0.03	3.08	12.495
0.2	0.056	1.78	5.285	0.053	1.847	5.63	0.052	1.854	5.668	0.052	1.859	5.691
0.3	0.071	1.235	3.436	0.066	1.281	3.638	0.065	1.287	3.663	0.065	1.29	3.68
0.4	0.079	0.897	2.636	0.073	0.929	2.775	0.072	0.934	2.796	0.072	0.937	2.809
0.5	0.083	0.656	2.23	0.077	0.679	2.342	0.076	0.683	2.36	0.075	0.686	2.372
0.6	0.085	0.471	2.014	0.077	0.488	2.112	0.076	0.492	2.129	0.075	0.495	2.141
0.7	0.085	0.322	1.9	0.077	0.334	1.993	0.075	0.338	2.011	0.074	0.341	2.023
0.8	0.083	0.196	1.848	0.075	0.206	1.94	0.073	0.21	1.958	0.072	0.213	1.971
0.9	0.081	0.089	1.836	0.073	0.096	1.928	0.071	0.1	1.947	0.069	0.105	1.96
1	0.079	-0.005	1.849	0.07	0	1.944	0.068	0.005	1.964	0.066	0.01	1.976
1.1	0.076	-0.088	1.881	0.067	-0.085	1.98	0.065	-0.079	1.999	0.063	-0.073	2.011
1.2	0.074	-0.163	1.925	0.065	-0.162	2.028	0.062	-0.155	2.048	0.06	-0.147	2.059
1.3	0.071	-0.231	1.979	0.062	-0.231	2.086	0.06	-0.223	2.105	0.058	-0.214	2.116
1.4	0.068	-0.292	2.038	0.059	-0.293	2.15	0.057	-0.285	2.17	0.055	-0.275	2.179
1.5	0.065	-0.349	2.103	0.057	-0.351	2.219	0.054	-0.343	2.239	0.052	-0.331	2.247
1.6	0.063	-0.401	2.17	0.055	-0.405	2.292	0.052	-0.396	2.312	0.05	-0.383	2.317
1.7	0.06	-0.449	2.24	0.052	-0.454	2.366	0.05	-0.445	2.387	0.047	-0.431	2.391
1.8	0.058	-0.494	2.311	0.05	-0.501	2.443	0.047	-0.491	2.464	0.045	-0.475	2.465
1.9	0.056	-0.537	2.383	0.048	-0.544	2.52	0.045	-0.534	2.542	0.043	-0.517	2.541
2	0.053	-0.576	2.456	0.046	-0.584	2.597	0.043	-0.574	2.62	0.041	-0.556	2.617
2.1	0.051	-0.613	2.528	0.044	-0.623	2.675	0.042	-0.612	2.698	0.039	-0.594	2.694
2.2	0.049	-0.648	2.6	0.042	-0.659	2.752	0.04	-0.648	2.776	0.038	-0.629	2.77
2.3	0.047	-0.681	2.672	0.041	-0.693	2.829	0.038	-0.682	2.854	0.036	-0.662	2.846
2.4	0.045	-0.712	2.744	0.039	-0.725	2.905	0.037	-0.714	2.932	0.035	-0.693	2.922
2.5	0.044	-0.742	2.814	0.038	-0.756	2.98	0.035	-0.745	3.008	0.033	-0.723	2.998
2.6	0.042	-0.77	2.884	0.036	-0.785	3.054	0.034	-0.774	3.084	0.032	-0.752	3.072
2.7	0.041	-0.797	2.953	0.035	-0.812	3.127	0.033	-0.802	3.159	0.031	-0.779	3.146
2.8	0.039	-0.823	3.021	0.034	-0.839	3.199	0.032	-0.829	3.233	0.029	-0.805	3.219
2.9	0.038	-0.847	3.088	0.033	-0.864	3.27	0.03	-0.854	3.306	0.028	-0.83	3.292
3	0.036	-0.871	3.154	0.031	-0.888	3.34	0.029	-0.879	3.378	0.027	-0.854	3.363

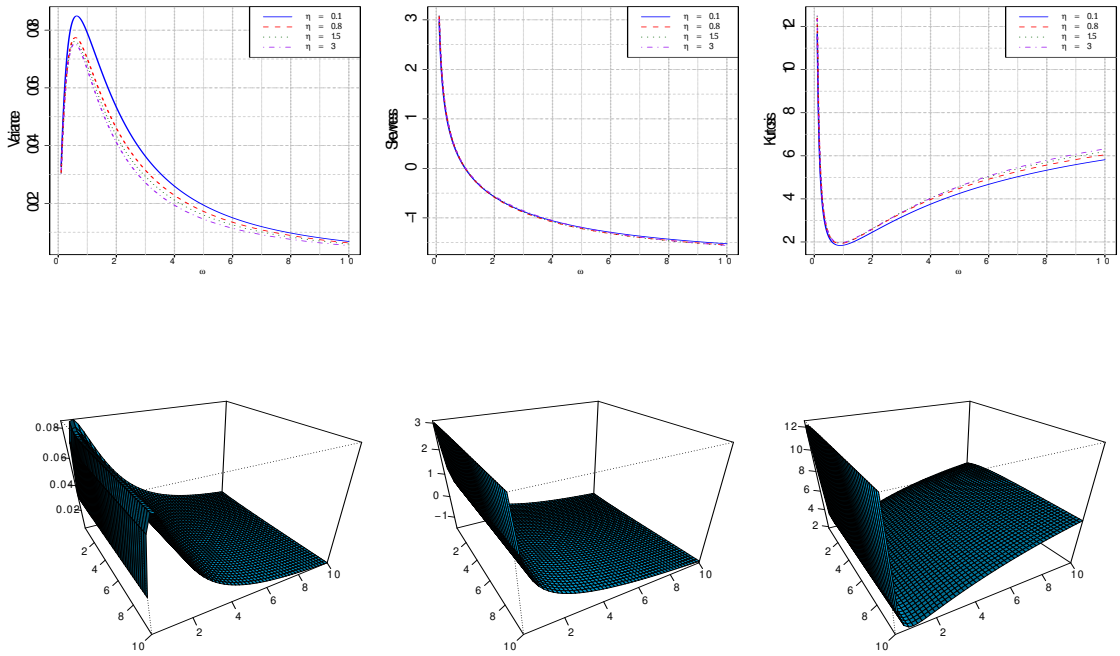


图2： $0 < \omega, \eta \leq 10$时 UShD 的方差、偏度和峰度。

3.2 分位数函数

UShD 的分位数函数 $x_p = F^{-1}(p)$，其中 $p \in (0,1)$，是下式的解

$$\omega x_p^{\omega} + (3\eta - 4\omega\eta \ln x_p) x_p^{2\omega} = p(\omega + 3\eta), \quad p \in (0,1) \tag{3.8}$$

方程 (3.8) 无法得到 x_p 的闭式解。因此，使用 R 统计软件中的 uniroot 函数以数值方式计算分位数。表 3 列出了 ω 和 η 的不同组合对应的部分分位数值。可以观察到，对于固定的 η，分位数随 ω 增大而增大。对于固定的 ω，增大 η 会使概率 $p \in (0, 0.57]$ 对应的分位数值升高；而当 $p \in (0.57, 1$ 时，分位数会随 η 增大而略微降低。

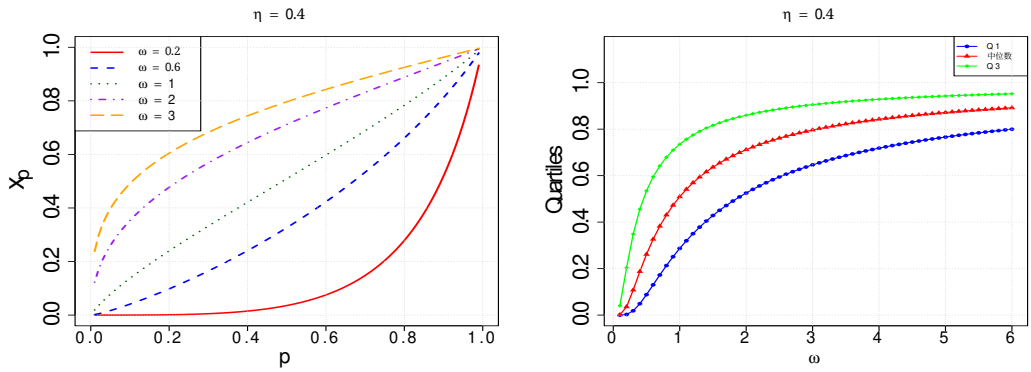


图 3：左： $\eta = 0.4$ 和不同 ω 值对应的 UShD 分位数函数。右：UShD 的四分位数作为 ω 的函数，对应 $\eta = 0.4$。

图 3 展示了在 $\eta = 0.4$ 取固定值时，形状参数 ω 对 UShD 分位数的影响。左图显示了不同概率水平下的完整分位数函数，右图则展示了第一四分位数、中位数和第三四分位数随 ω 变化的情况；随着 ω 增大，所有四分位数均单调递增。

表 3：不同参数组合下 UShD 的分位数值。

p	$\omega=0.5$ $\eta=0.4$	$\omega=0.5$ $\eta=1$	$\omega=1$ $\eta=0.4$	$\omega=1$ $\eta=1$	$\omega=1.5$ $\eta=0.4$	$\omega=1.5$ $\eta=1$
0.01	0.0006	0.0011	0.0192	0.0272	0.0650	0.0818
0.05	0.0080	0.0102	0.0779	0.0926	0.1723	0.1954
0.10	0.0224	0.0262	0.1374	0.1533	0.2564	0.2782
0.25	0.0880	0.0933	0.2869	0.2992	0.4284	0.4424
0.40	0.1821	0.1861	0.4214	0.4281	0.5589	0.5657
0.50	0.2615	0.2635	0.5091	0.5119	0.6363	0.6390
0.55	0.3067	0.3075	0.5531	0.5540	0.6734	0.6743
0.57	0.3260	0.3262	0.5708	0.5710	0.6880	0.6882
0.58	0.3359	0.3358	0.5797	0.5795	0.6952	0.6952
0.60	0.3562	0.3555	0.5975	0.5967	0.7097	0.7090
0.75	0.5347	0.5292	0.7353	0.7301	0.8166	0.8124
0.90	0.7770	0.7693	0.8860	0.8802	0.9244	0.9201
0.95	0.8800	0.8741	0.9411	0.9372	0.9616	0.9588
0.99	0.9743	0.9726	0.9879	0.9868	0.9922	0.9915

3.3 熵

熵是衡量随机变量相关不确定性或信息的基本指标。对于连续分布，微分熵量化了结果的不可预测性，广泛应用于统计学、信息论和工程领域，用于模型比较、风险评估和系统分析。对于随机变量 $X \sim \mathrm{USh}(\omega,\eta)$，微分熵定义为

$$H(x) = -\int_0^1 f(x;\omega,\eta)\,\ln\big(f(x;\omega,\eta)\big)\,dx = -\mathrm{E}\big[\ln f(X;\omega,\eta)\big].$$

代入式（2.5）中的概率密度函数，可得

$$\begin{aligned} H(x) &= -\mathrm{E}\left[\ln\left(\frac{\omega}{\omega+3\eta}\left(\omega X^{\omega-1}+(2\eta-8\omega\eta\ln X)X^{2\omega-1}\right)\right)\right] \\ &= -\ln\left(\frac{\omega}{\omega+3\eta}\right) - \mathrm{E}\left[\ln\left(\omega X^{\omega-1}+(2\eta-8\omega\eta\ln X)X^{2\omega-1}\right)\right]. \end{aligned} \tag{3.9}$$

式（3.9）中的期望值无法通过解析方法推导，但可以针对特定参数值进行数值计算，如表4和图4所示。可以观察到，H(X) 为负值，并且对于 $0<\omega<1$，随着 ω 增大而增大，随后对于 $\omega \geq 1$ 缓慢减小。

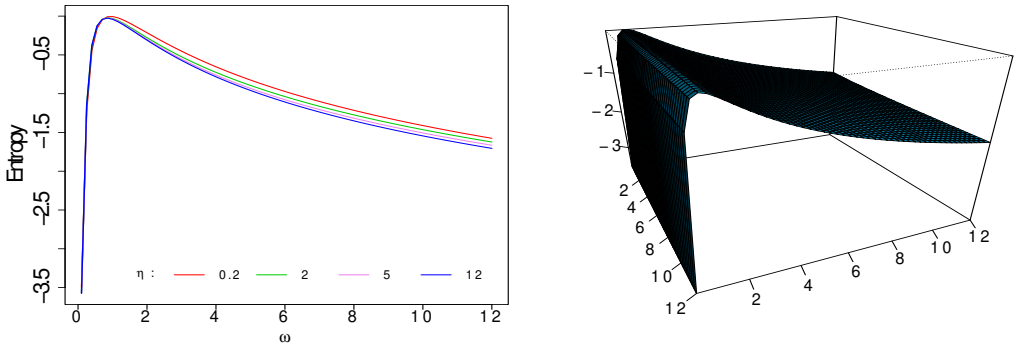
4 参数估计

在本节中，我们使用极大似然估计法估计 UShD 的参数 ω和η 。设 $X1,\ldots,X_n$为来自 UShD 的随机样本，$x1,\ldots,x_n$表示该样本的取值。似然函数和对数似然函数分别为

$$L(\omega,\eta) = \prod_{i=1}^{n} \frac{\omega}{\omega+3\eta}\left[\omega x_i^{\omega-1}+(2\eta-8\omega\eta\ln x_i)x_i^{2\omega-1}\right].$$

表4：所选参数值下 UShD 的熵。

ω\η	0.2	0.5	0.7	0.9	1	1.1	1.5	2	3	4
0.1	-3.4946	-3.5412	-3.5511	-3.5568	-3.5588	-3.5605	-3.5649	-3.5680	-3.5712	-3.5728
0.2	-1.7715	-1.7418	-1.7350	-1.7311	-1.7297	-1.7286	-1.7254	-1.7232	-1.7210	-1.7198
0.3	-0.8691	-0.8240	-0.8132	-0.8069	-0.8046	-0.8027	-0.7975	-0.7938	-0.7901	-0.7882
0.5	-0.2268	-0.2040	-0.1984	-0.1951	-0.1939	-0.1930	-0.1904	-0.1886	-0.1868	-0.1859
0.7	-0.0469	-0.0414	-0.0407	-0.0405	-0.0405	-0.0405	-0.0408	-0.0411	-0.0417	-0.0420
0.9	-0.0043	-0.0092	-0.0120	-0.0142	-0.0151	-0.0159	-0.0185	-0.0207	-0.0233	-0.0248
1.0	-0.0052	-0.0138	-0.0179	-0.0210	-0.0223	-0.0235	-0.0271	-0.0301	-0.0336	-0.0356
1.1	-0.0143	-0.0259	-0.0311	-0.0350	-0.0367	-0.0381	-0.0426	-0.0463	-0.0507	-0.0533
1.5	-0.0910	-0.1093	-0.1176	-0.1239	-0.1265	-0.1288	-0.1361	-0.1423	-0.1498	-0.1542
2.0	-0.2143	-0.2357	-0.2457	-0.2536	-0.2570	-0.2600	-0.2696	-0.2780	-0.2884	-0.2946
2.5	-0.3370	-0.3590	-0.3697	-0.3784	-0.3821	-0.3855	-0.3966	-0.4065	-0.4191	-0.4268
3.0	-0.4515	-0.4730	-0.4840	-0.4930	-0.4969	-0.5005	-0.5124	-0.5233	-0.5376	-0.5466



Fig图4： $0<\omega$、$\eta\leq 12$ 下 UShD 的熵，熵范围为 [-3.576、-0.003]

$$\ell(\omega,\eta)=n\ln\omega-n\ln(\omega+3\eta)+\sum_{i=1}^{n}\ln\left[\omega x_i^{\omega-1}+(2\eta-8\omega\eta\ln x_i)x_i^{2\omega-1}\right].$$

$\ell(\omega,\eta)$关于 ω 和 η 的一阶偏导数为

$$\frac{\partial\ell}{\partial\omega}=\frac{n}{\omega}-\frac{n}{\omega+3\eta}+\sum_{i=1}^{n}\frac{x_i^{\omega-1}\left(1+\omega\ln x_i\right)+x_i^{2\omega-1}\left[(2\eta-8\omega\eta\ln x_i)(2\ln x_i)-8\eta\ln x_i\right]}{\omega x_i^{\omega-1}+(2\eta-8\omega\eta\ln x_i)x_i^{2\omega-1}}.$$

$$\frac{\partial\ell}{\partial\eta}=-\frac{3n}{\omega+3\eta}+\sum_{i=1}^{n}\frac{(2-8\omega\ln x_i)x_i^{2\omega-1}}{\omega x_i^{\omega-1}+(2\eta-8\omega\eta\ln x_i)x_i^{2\omega-1}}.$$

通过求解似然方程组，可以得到参数 ω和 η 的最大似然估计

$$\frac{\partial\ell}{\partial\omega}=0\quad\text{and}\quad\frac{\partial\ell}{\partial\eta}=0.$$

然而，由于这些方程较为复杂，因此不存在闭式解。因此，采用数值优化方法（例如带盒约束的有限内存 Broyden–Fletcher–Goldfarb–Shanno 算法（L-BFGS-B））来获得最大似然估计。

5 模拟研究

在本节中，我们评估所提出模型参数的最大似然估计和渐近置信区间的性能。模拟的主要目的是评估以下内容：

- $\hat{\beta}$ 的偏差：$\text{Bias}(\hat{\beta}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}_i - \beta)$
- $\hat{\beta}$ 的均方误差（MSE）：$\text{MSE}(\hat{\beta}) = \frac{1}{M}\sum_{i=1}^{M}\left(\hat{\beta}_i - \beta\right)^2$
- $\hat{\beta}$ 的平均相对误差（MRE）：$\text{MRE}(\hat{\beta}) = \frac{1}{M}\sum_{i=1}^{M}\left|\frac{\hat{\beta}_i - \beta}{\beta}\right|$
- β 的覆盖概率（CP）：$\text{CP}_\beta = \frac{1}{M}\sum_{i=1}^{M} I(\beta \in \text{CI}_i(\beta))$
- 收敛率（CR），定义为优化算法成功收敛的模拟运行次数所占的比例，

$$\text{CR} = \frac{\text{Number of successful estimation runs}}{\text{Total number of simulation runs}}$$

$\hat{}$其中，β_i是第 i 次模拟运行中 β（ω 或 η）的最大似然估计（MLE），$i = 1, \ldots, M$，$\text{CI}_i(\beta)$ 是根据第 i 个样本得到的 β 的置信区间，$I(\cdot)$ 是指示函数，即，如果 α 为真，则 $I(\alpha) = 1$，否则为 0。

模拟过程如下（使用 R 软件）：

1. 使用基于推导出的概率密度函数的拒绝抽样算法，从 UShD 中生成样本量为 $n = 30$、60、100、150 和 200 的独立随机样本。2. 对于每个生成的样本，使用具有适当上下界的 L-BFGS-B 优化方法，通过最大化似然函数估计参数（ω、η）。3. 对于每个估计参数，使用 $B = 100$ 个重抽样构造基于自助法的置信区间，以近似抽样分布并计算覆盖概率。4. 对上述步骤重复进行 $M = 1000$ 次模拟运行，以获得 Bias、MSE、MRE、CP 和 CR 的可靠估计。

模拟结果呈现在表 5 中，展示了不同样本量下的性能指标。该模拟研究揭示了最大似然估计的有限样本行为，说明随着样本量增加，估计精度和置信区间覆盖率均有所提高。

6 应用

本节旨在通过将 USh 模型应用于实际数据集，并将其与若干替代模型进行比较，来说明该模型的有效性与灵活性。所提出模型的拟合优度将对照表 6 中给出的七个竞争模型进行评估。通过对 USh 模型及其竞争模型进行拟合，考察了四个实际数据集，并使用多种信息准则（包括赤池信息准则、修正赤池信息准则、贝叶斯信息准则和 Hannan–Quinn信息准则）评估其表现。此外，采用 Kolmogorov–Smirnov检验来衡量经验分布与拟合理论分布之间的一致性，并使用相应的 p 值进行模型比较。该检验被广泛用于评估理论分布对观测数据的代表程度。一般而言，信息准则取值最小且 p值最大的模型被视为拟合效果最佳。

本研究考虑的数据集来自不同来源，并已在文献中被广泛使用。为确保一致性并适合对单位概率分布进行建模，我们选取了要么自然定义在单位区间 (0, 1) 上、要么可以适当变换到该范围的数据集。其中一些数据集对应于容量因子，这些

表 5：模拟结果。

ω	η	n	偏差($\hat{\omega}$)	偏差($\hat{\eta}$)	均方误差($\hat{\omega}$)	均方误差($\hat{\eta}$)	均对数误差($\hat{\omega}$)	均对数误差($\hat{\eta}$)	CP_{ω}	CP_{η}	CR
0.6	0.2	30	0.0305	0.0889	0.0139	0.0590	0.3019	0.7618	0.9456	0.9379	0.972
		60	0.0230	0.0840	0.0081	0.0535	0.2285	0.7066	0.9474	0.9395	0.969
		100	0.0185	0.0389	0.0045	0.0285	0.1707	0.5170	0.9419	0.9498	0.973
		150	0.0145	0.0303	0.0033	0.0262	0.1507	0.4784	0.9529	0.9476	0.969
		200	0.0124	0.0254	0.0024	0.0224	0.1288	0.4487	0.9498	0.9480	0.978
0.6	1.8	30	0.0401	0.1502	0.0144	0.7765	0.0993	0.4454	0.9465	0.9404	0.976
		60	0.0290	0.1157	0.0076	0.6970	0.0749	0.4080	0.9454	0.9428	0.973
		100	0.0170	0.0239	0.0041	0.1814	0.0548	0.1409	0.9415	0.9457	0.983
		150	0.0127	0.0190	0.0026	0.1068	0.0435	0.1013	0.9549	0.9485	0.975
		200	0.0114	0.0081	0.0019	0.0454	0.0377	0.0871	0.9545	0.9465	0.980
1	0.7	30	0.0582	0.2135	0.0424	0.5668	0.1882	0.7689	0.9418	0.9438	0.975
		60	0.0349	0.1939	0.0201	0.5018	0.1347	0.6930	0.9491	0.9436	0.977
		100	0.0260	0.0689	0.0125	0.2747	0.1060	0.5931	0.9531	0.9379	0.962
		150	0.0182	0.0505	0.0078	0.1354	0.0846	0.4375	0.9574	0.9418	0.959
		200	0.0144	0.0017	0.0067	0.0973	0.0768	0.3305	0.9474	0.9455	0.967
1.2	0.8	30	0.0686	0.0679	0.0572	0.3763	0.1948	0.6159	0.9309	0.9350	0.974
		60	0.0439	0.0397	0.0280	0.3584	0.1366	0.5947	0.9452	0.9426	0.987
		100	0.0264	0.0110	0.0155	0.2556	0.1000	0.4914	0.9483	0.9439	0.973
		150	0.0206	0.0103	0.0108	0.1594	0.0846	0.3589	0.9498	0.9427	0.996
		200	0.0165	0.0098	0.0095	0.0763	0.0804	0.2952	0.9469	0.9498	0.985
1.5	0.2	30	0.0305	0.0802	0.0805	0.0598	0.6346	0.6568	0.9471	0.9430	0.988
		60	0.0264	0.0486	0.0435	0.0519	0.4689	0.6072	0.9478	0.9407	0.983
		100	0.0200	0.0072	0.0265	0.0305	0.3663	0.4623	0.9504	0.9574	0.992
		150	0.0141	0.0049	0.0179	0.0275	0.2933	0.4337	0.9473	0.9564	0.991
		200	0.0091	0.0017	0.0116	0.0153	0.1599	0.2855	0.9476	0.9507	0.995
2	0.6	30	0.0851	0.0803	0.1461	0.2972	0.3322	0.5627	0.9396	0.9366	0.996
		60	0.0446	0.0788	0.0767	0.2329	0.2319	0.5338	0.9589	0.9548	0.998
		100	0.0390	0.0351	0.0474	0.1952	0.1872	0.4601	0.9476	0.9406	0.995
		150	0.0180	0.0527	0.0321	0.1205	0.1589	0.4183	0.9469	0.9368	0.998
		200	0.0072	0.0065	0.0142	0.0261	0.1244	0.3298	0.9579	0.9469	0.998
2	1.4	30	0.0666	0.0669	0.1237	0.6743	0.1700	0.5361	0.9493	0.9438	0.991
		60	0.0562	0.0610	0.0711	0.5084	0.1264	0.5098	0.9457	0.9486	0.996
		100	0.0472	0.0539	0.0480	0.3296	0.1043	0.4539	0.9600	0.9530	1.000
		150	0.0209	0.0170	0.0301	0.0692	0.0819	0.3410	0.9648	0.9557	0.997
		200	0.0080	0.0062	0.0247	0.0495	0.0720	0.2020	0.9509	0.9539	0.999

表 6：单位Shiha分布的若干竞争模型

分布	概率密度函数	参考资料
Kumaraswamy (Kw)	$f_{Kw}(x) = \omega\eta x^{\eta-1}(1-x^{\eta})^{\omega-1}, \quad \omega>0, \eta>0$	[14]
双参数单位 Bilal 分布 (UB)	$f_{UB}(x) = \frac{6\omega}{\eta}\left(1-x^{\omega/\eta}\right)x^{2\omega/\eta-1}, \quad \omega,\eta>0$	[20]
单位指数分布 (UE)	$f_{UE}(x) = \frac{2\omega\eta}{1-}x^{2}(\frac{1+x}{1-}x)^{\eta}$指数分布 $\left[\omega\left(1-(\frac{1+x}{1-}x)^{\eta}\right)\right], \quad \omega,\eta\geq 0$	[4]
幂化 UEHL 分布（EUEHL）	$f_{EUEHL}(x) = 2\omega\eta\alpha\frac{x^{\omega-1}(1+x^{\omega})^{2}}{(}\left(\frac{1-x^{\omega}}{1+}x^{\omega}\right)^{\eta-1}\left[1-\left(\frac{1-x^{\omega}}{1+}x^{\omega}\right)^{\eta}\right]^{\alpha-1}, \quad \omega>0,\eta>0,\alpha>0$	[9]
单位指数化 Lomax 分布 (UEL)	$f_{UEL}(x) = \frac{\eta\omega\alpha}{x}(1-\eta\ln(x))^{-\omega-1}\{1-(1-\eta\ln(x))^{-\omega}\}^{\alpha-1}, \quad \omega>0,\ \eta>0,\ \alpha>0$	[8]
Beta	$f_{Beta}(x) = \frac{\Gamma(\omega+\eta)}{\Gamma(\omega)\Gamma(\eta)}x^{\omega-1}(1-x)^{\eta-1}, \quad \omega,\eta>0$	[12]
Topp-Leone分布	$f_{TL}(x) = 2\omega x^{\omega-1}(1-x)(2-x)^{\omega-1}, \quad \omega>0$	[21]

被定义为发电机组的实际电力输出与最大可能输出之比，因此位于区间 (0, 1) 内。这些数据集通常用作评估能源发电系统效率的可靠性度量，并已使用不同的

估计算法。此外，本研究还包括寿命和可靠性数据，例如医学生存时间和 I 型删失失效时间数据，这些数据经过转换后落在单位区间内，以便进行建模。下文将详细介绍各数据集。数据I：以下数据集包含沙特阿拉伯卫生部一家医院中 43 名血癌患者的寿命（以天为单位） [1]:

115, 181, 255, 418, 441, 461, 516, 739, 743, 789, 807, 865, 924, 983, 1025, 1062, 1063, 1165, 1191, 1222, 1222, 1251, 1277, 1290, 1357, 1369, 1408, 1455, 1478, 1519, 1578, 1578, 1599, 1603, 1605, 1696, 1735, 1799, 1815 ,1852, 1899, 1925, 1965。将每个观测值除以 1970，所得数据值介于 0 和 1 之间。

以下数据集II和III由 [5] 报告，近期已被许多作者研究，例如 [6, 3]。这些数据提供了两种不同算法（即SC16和P3）估计单位容量因子的比较。

数据II：SC16算法产生的观测值如下：
0.853，0.759，0.866，0.809，0.717，0.544，0.492，0.403，0.344，0.213，0.116，0.116，0.092，0.070，0.059，0.048，0.036，0.029，0.021，0.014，0.011，0.008，0.006。

Data III: The values, ranked from smallest to largest, are: 0.853, 0.759, 0.874, 0.800, 0.716, 0.557, 0.503, 0.399, 0.334, 0.207, 0.118, 0.118, 0.097, 0.078, 0.067, 0.056, 0.044, 0.036, 0.026, 0.019, 0.014, 0.010。

数据 IV： [16], 讨论的 I 型删失数据集由 50 个经过测试的项目组成，测试在 12 小时后终止。测试终止前观察到的失效时间如下：

0.80、1.26、1.29、1.85、2.41、2.47、2.76、3.35、3.68、4.46、4.65、4.83、5.21、5.26、5.36、5.39、5.53、5.64、5.80、6.08、6.38、7.02、7.18、7.60、8.13、8.46、8.69、10.52、11.25、11.90。将每个观测值除以12，得到介于0和1之间的数据值。表 7给出了该数据的一些描述性统计量。图 5–8展示了四个数据集的TTT图、箱线图和小提琴图。

表 7：四个数据集的一些描述性统计量

Data	Min	Q1	中位数	Mean	Q3	Max	方差	SK	KU
I	0.058	0.424	0.635	0.605	0.806	0.998	0.066	-0.442	5.279
II	0.006	0.032	0.116	0.288	0.518	0.866	0.101	0.768	4.974
III	0.01	0.047	0.118	0.304	0.544	0.874	0.101	0.711	4.884
IV	0.067	0.286	0.448	0.459	0.595	0.992	0.058	0.388	5.645

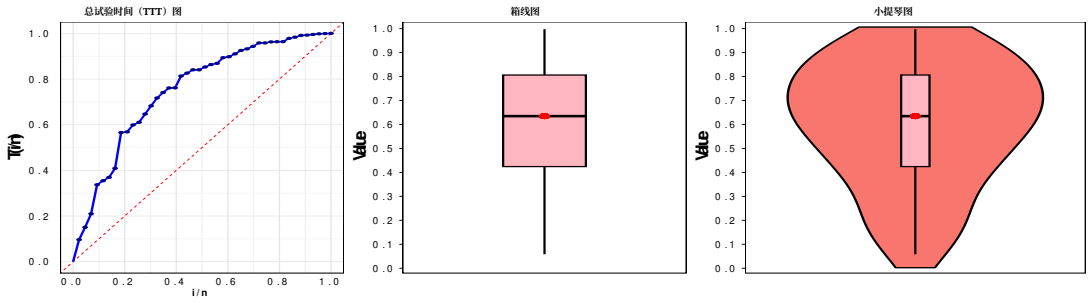


图 5：数据I 的 TTT、箱线图和小提琴图。

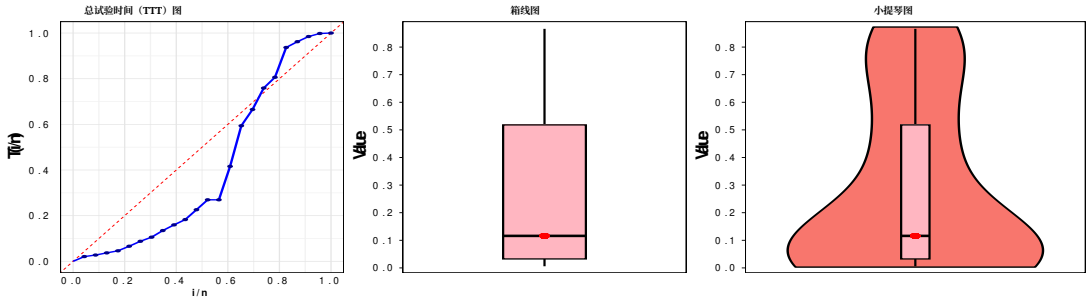


图6：数据II的TTT图、箱线图和小提琴图。

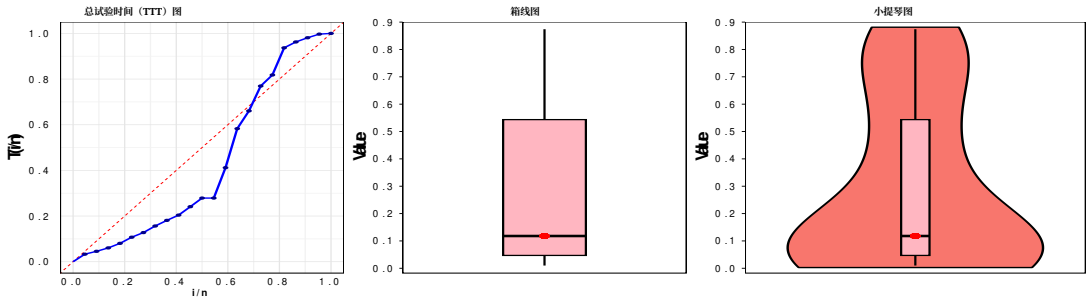


图7：数据 III 的 TTT 图、箱线图和小提琴图。

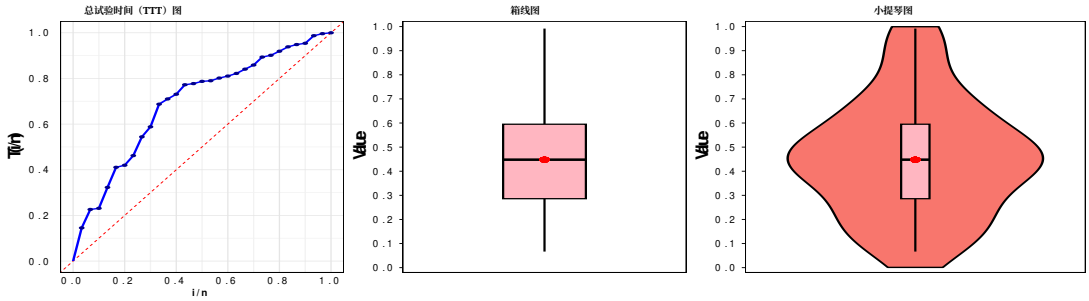


图8：数据 IV 的 TTT 图、箱线图和小提琴图。

根据表 7中报告的描述性统计量以及相应的箱线图和小提琴图可以看出，数据集I呈左偏分布，而数据集II、III和IV为右偏。此外，数据集I和IV的TTT图呈凹形，表明失效率（风险率）递增。相比之下，数据集II和III的TTT图起初为凸形、随后为凹形，表明其失效率呈浴盆形，即先下降后上升。因此，USh 分布似乎是拟合这些数据集的合适且灵活的模型。

表 8–11 展示了四个数据集在所有竞争模型下的最大似然估计（MLEs）和拟合优度信息准则。报告的各项指标一致表明，所提出的 USh 分布优于其他模型，在大多数情况下拟合效果最佳。图形诊断进一步支持了这些数值结果。具体而言，通过将每个模型拟合到每个数据集，我们评估了这些模型捕捉不同类型数据中潜在分布模式和特征的能力。叠加了拟合概率密度函数的数据集直方图，以及经验和估计的累积分布函数，均表明 USh 模型拟合充分。此外，相应的 PP 图和 Q-Q图也提供了进一步验证：基于 USh 的点与参考线高度吻合，表明与竞争模型相比，该模型与观测数据的契合度更高。

总体而言，这些结果凸显了 USh 分布的有效性和灵活性，证实了它适用于对各种单位值数据集进行建模。参见图 9–20。

表8：数据I的最大似然估计和拟合优度度量。

分布	$\hat{\omega}$	$\hat{\eta}$	$\hat{\alpha}$	AIC	AICC	BIC	HQIC	KS	p值
USh	1.4957	0.1221	–	**-3.1059**	**-2.8059**	**0.4165**	**-1.8069**	**0.0726**	**0.9772**
Kw	1.0177	1.559	–	-3.0158	-2.7158	0.5066	-1.7169	0.0813	0.9390
UB	1.2819	0.9874	–	11.4709	11.7709	14.9933	12.7699	0.1457	0.3205
UE	1.7603	0.2293	–	2.5036	2.8036	6.0260	3.8026	0.1317	0.4447
EUEHL	0.075	1.0012	40.3039	-1.0279	-0.4125	4.2557	0.9205	0.0799	0.9465
UEL	54.3925	0.0294	1.0319	-0.9877	-0.3724	4.2959	0.9607	0.0819	0.9350
Beta	1.5613	1.0173	–	-3.0159	-2.7159	0.5065	-1.7170	0.0812	0.9393
TL	3.0122	–	–	15.9571	16.0546	17.7183	16.6065	0.1877	0.0965

表9：数据II的最大似然估计和拟合优度度量。

分布	$\hat{\omega}$	$\hat{\eta}$	$\hat{\alpha}$	AIC	AICC	BIC	HQIC	KS	p值
USh	0.374	193.1469	–	**-15.7077**	**-15.1077**	**-13.4367**	**-15.1366**	**0.1512**	**0.6694**
Kw	1.1862	0.5044	–	-15.3416	-14.7416	-13.0706	-14.7704	0.1790	0.4529
UB	2.0102	5.3038	–	-11.2158	-10.6158	-8.9448	-10.6446	0.1786	0.4553
UE	107.6512	0.0126	–	-9.0515	-8.4515	-6.7805	-8.4804	0.2941	0.0374
EUEHL	0.3132	1.0549	2.4651	-14.6813	-13.4181	-11.2748	-13.8246	0.1584	0.6112
UEL	52.1005	0.0099	1.2034	-13.1600	-11.8968	-9.7535	-12.3033	0.1808	0.4400
Beta	0.4869	1.1679	–	-15.2149	-14.6149	-12.9439	-14.6438	0.1836	0.4202
TL	0.5943	–	–	-14.2302	-14.0398	-13.0948	-13.9447	0.1690	0.5272

表10：数据III的最大似然估计和拟合优度度量。

分布	$\hat{\omega}$	$\hat{\eta}$	$\hat{\alpha}$	AIC	AICC	BIC	HQIC	KS	p值
USh	0.4159	1000	–	**-10.1225**	**-9.4910**	**-7.9405**	**-9.6085**	**0.1760**	**0.5029**
Kw	1.2305	0.5718	–	-9.6872	-9.0557	-7.5051	-9.1732	0.1963	0.3650
UB	1.7316	4.1212	–	-6.3672	-5.7357	-4.1852	-5.8532	0.1827	0.4546
UE	93.4921	0.0137	–	-6.7942	-6.1626	-4.6121	-6.2802	0.2824	0.0598
EUEHL	0.34	1.0911	2.6393	-9.0181	-7.6847	-5.7449	-8.2470	0.1788	0.4828
UEL	54.3238	0.0107	1.2387	-7.5177	-6.1844	-4.2446	-6.7467	0.1968	0.3616
Beta	0.554	1.2198	–	-9.5638	-8.9323	-7.3818	-9.0498	0.2002	0.3413
TL	0.6778	–	–	-8.9965	-8.7965	-7.9055	-8.7395	0.1848	0.4401

表11：数据 IV 的最大似然估计和拟合优度度量。

分布	$\hat{\omega}$	$\hat{\eta}$	$\hat{\alpha}$	AIC	AICC	BIC	HQIC	KS	p值
USh	0.8854	545.452	–	**-0.2354**	**0.2090**	2.5670	**0.6611**	0.1296	0.6478
Kw	1.4896	1.3281	–	1.5721	2.0166	4.3745	2.4686	0.1304	0.6401
UB	0.9848	1.1287	–	2.9225	3.3670	5.7249	3.8190	**0.0881**	**0.9581**
UE	7.8993	0.0945	–	4.6855	5.1300	7.4879	5.5820	0.1540	0.4318
EUEHL	1.2418	1.2371	1.4679	1.4424	2.3655	5.6460	2.7872	0.1020	0.8830
UEL	91.3279	0.0148	1.501	3.7091	4.6322	7.9127	5.0539	0.1316	0.6287
Beta	1.3856	1.5036		1.4478	1.8922	4.2502	2.3443	0.1288	0.6548
TL	1.8705	–	–	1.0529	1.1957	**2.4541**	1.5011	0.0909	0.9461

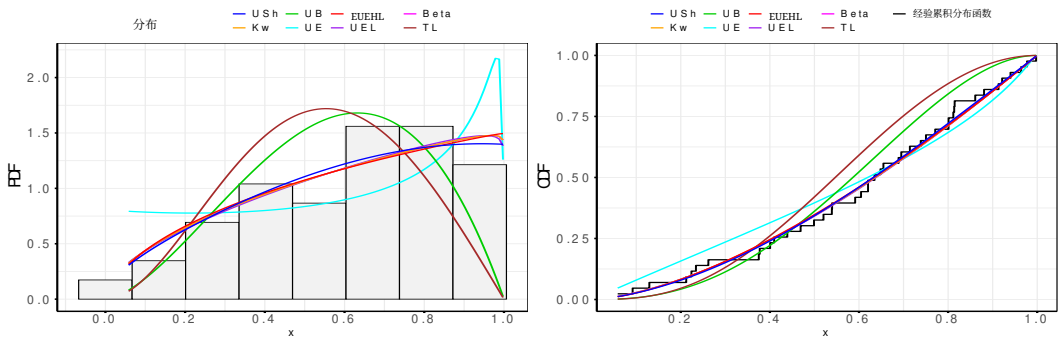


图9：数据I的竞争模型的拟合概率密度函数和累积分布函数。

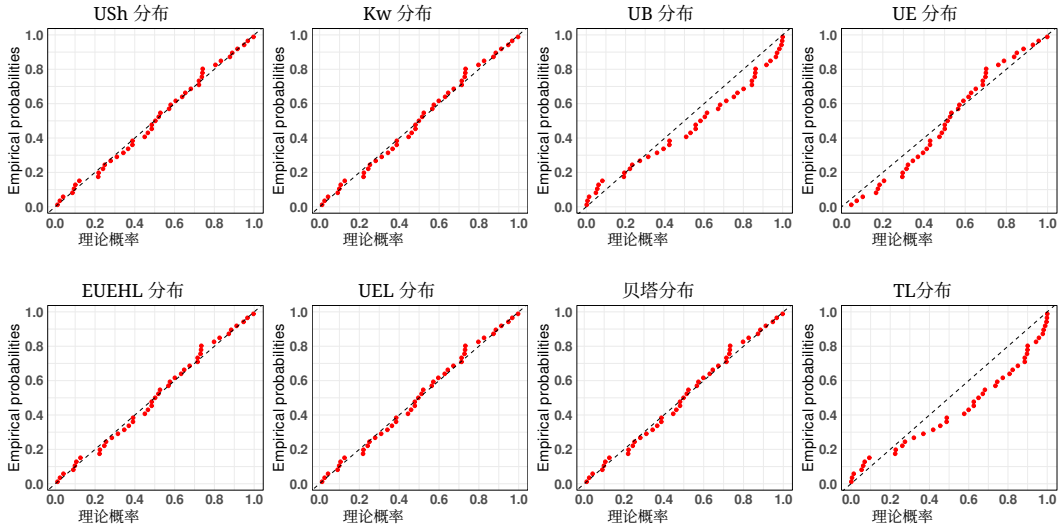


图10：数据I的竞争模型概率-概率（PP）图 .

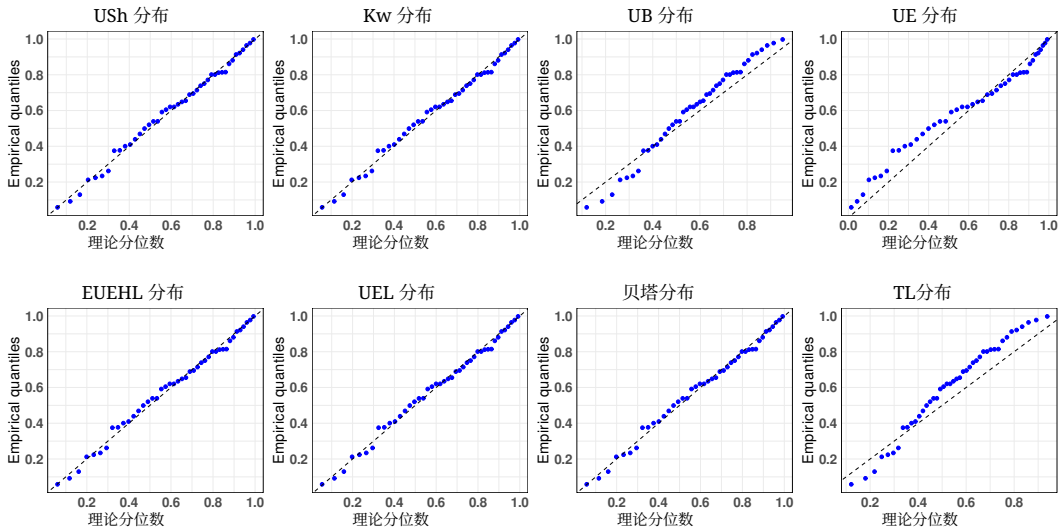


图11：数据I的竞争模型分位数-分位数（QQ）图。

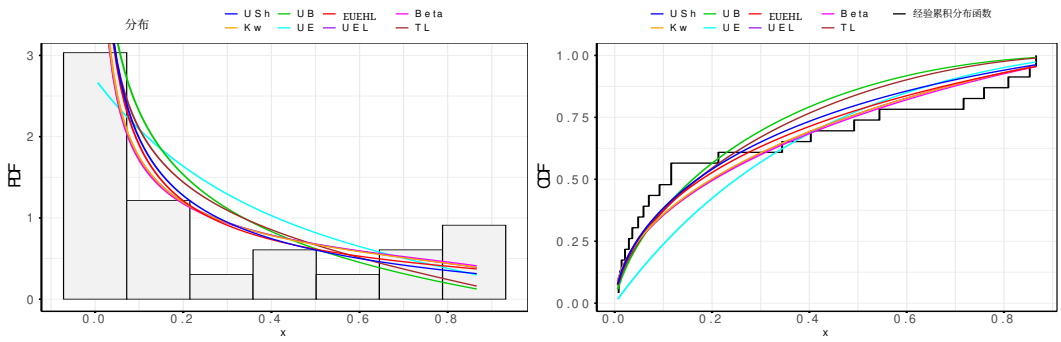


图12：数据II的竞争模型的拟合概率密度函数和累积分布函数。

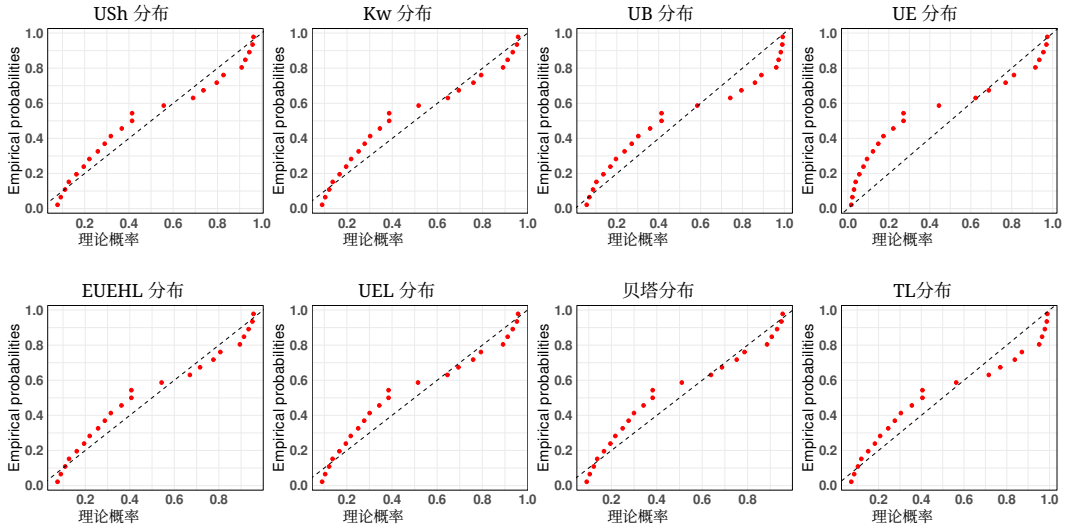


图13：数据II的竞争模型概率-概率（PP）图 .

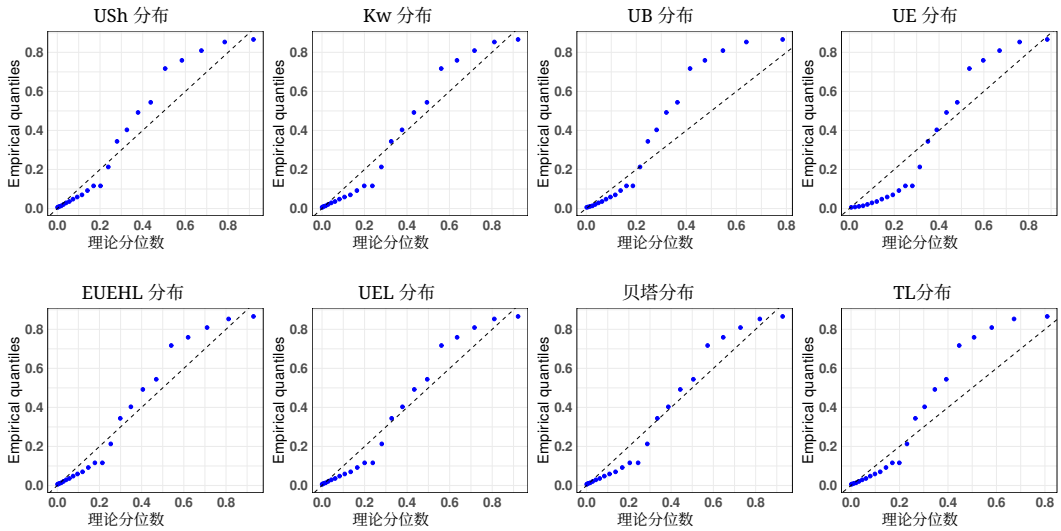


图14：数据II的竞争模型分位数-分位数（QQ）图。

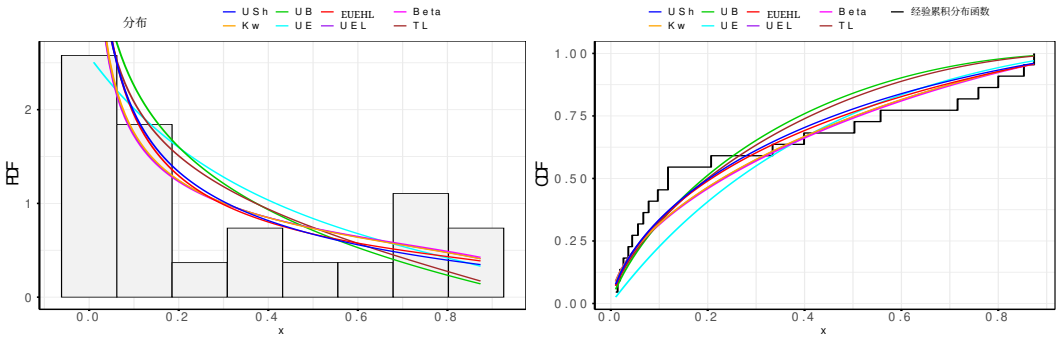


图15：数据III的竞争模型拟合概率密度函数和累积分布函数。

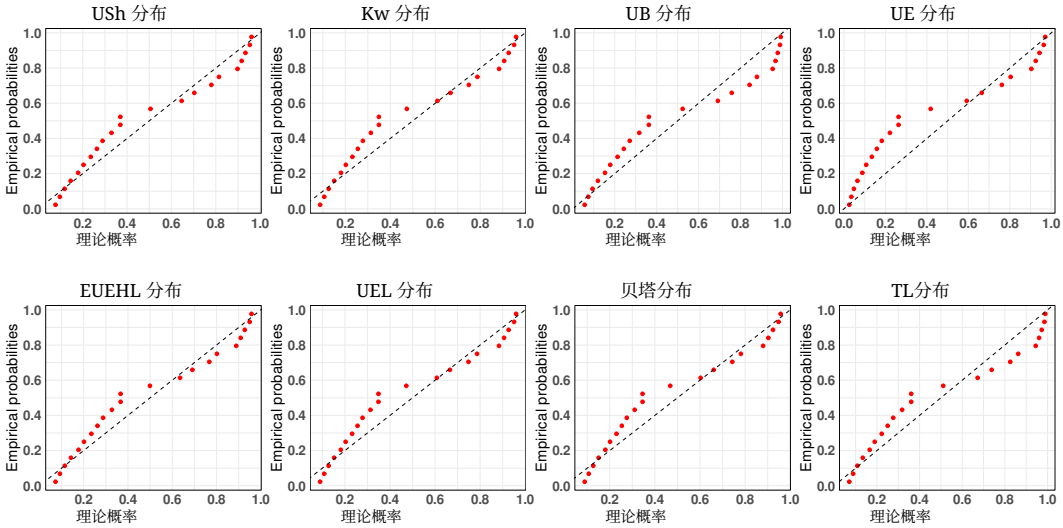


图16：数据II的竞争模型概率-概率（PP）图 I.

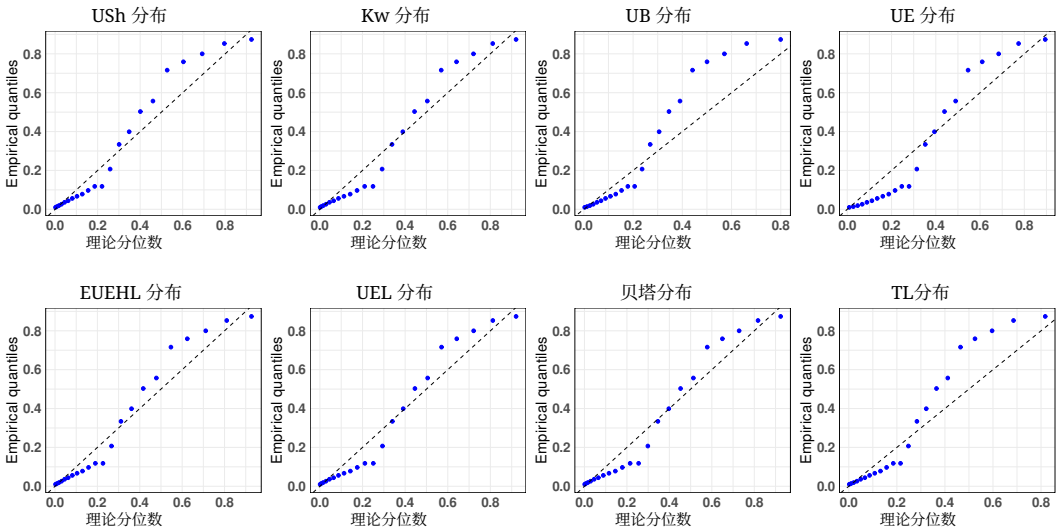


图17：数据III的竞争模型的分位数-分位数（QQ）图。

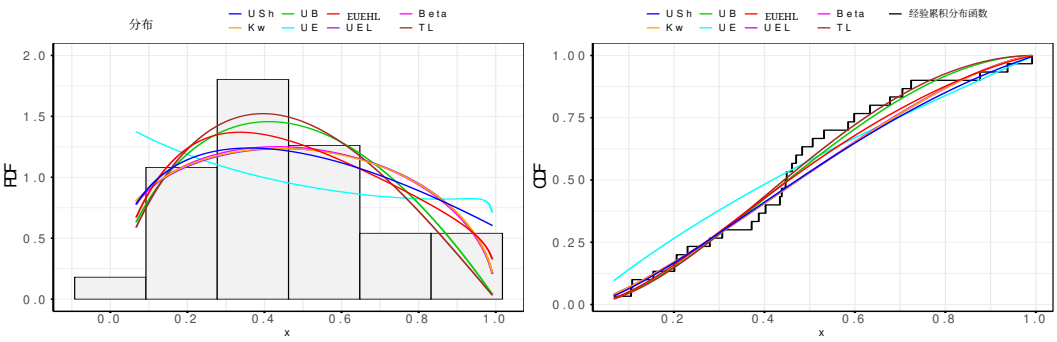


图18：数据 IV 的竞争模型拟合概率密度函数和累积分布函数。

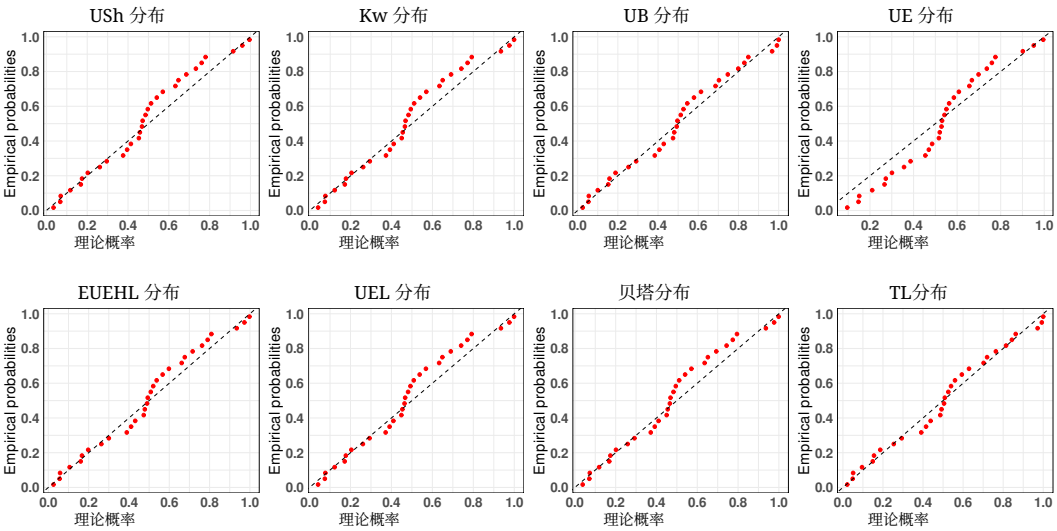


图19：数据 IV 的竞争模型概率-概率（PP）图 .

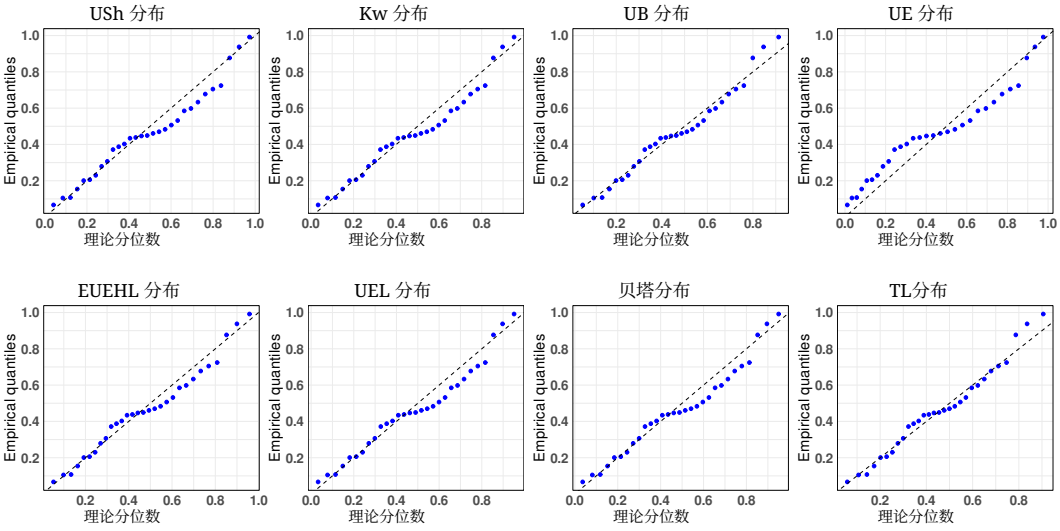


图 20：数据集 IV 的竞争模型分位数-分位数（QQ）图。

7 结论

本文提出了一种新的单位分布，称为单位 Shiha（USh）分布，该分布由 Sh 分布经反指数变换得到。本文推导了所提模型的若干统计性质，包

括矩、分位数函数、熵和应力-强度可靠度。得益于其概率密度函数和风险率函数的灵活性，USh 分布能够对左偏和右偏数据进行建模，并刻画不同的失效率行为，例如递增型和浴盆形模式。本文采用最大似然估计进行参数推断，并开展模拟研究以评估估计量的有限样本性能。模拟结果表明，估计精度令人满意，收敛表现良好，相关置信区间的覆盖概率也较为充分，且随着样本量增加，各项表现均有明显改善。本文使用四个实际数据集展示了所提模型的实际适用性，并将其与若干广为人知的竞争性单位分布进行了比较。拟合优度准则和图形诊断均表明，在大多数情况下，USh 分布对所考察数据集的拟合效果更优。总体而言，所提出的 USh 分布为单位区间数据提供了灵活而有效的建模框架，也为单位概率分布系列增添了有价值的新成员。

利益冲突

作者声明，就本文的发表而言不存在利益冲突。

参考文献

[1] Abouammoh, A. M., Abdulghani, S. A., Qamber, I. S. (1994). 关于新优于更新使用类的偏序关系与检验。Reliab. Eng. Syst. Safe. 43, 37–41.[2] Ahsan-ul Haq, M., Hashmi, S., Aidi, K., Ramos, P. L., Louzada, F. (2023). 单位修正 Burr-III 分布：估计、特征刻画与验证检验，Ann. Data. Sci., 10, 415–440. https://doi.org/10.1007/s40745-020-00298-6[3] Alzahrani, M. R., Almohaimeed, M. (2025). 单位 Haq 分布在工程数据中的分析、推断与应用，Alexandria Engineering Journal 117, 193–204. https://doi.org/10.1016/j.aej.2025.01.050[4] Bakouch, H. S., Hussain, T., Toši´c, M., Stojanovi´c, V. S., Qarmalah, N. (2023). 单位指数概率分布：在环境与工程数据建模中的特征刻画与应用。Mathematics , 11, 4207. https://doi.org/10.3390/math11194207[5] Caramanis, M., Stremel, J., Fleck, W., Daniel, S. (1983). 概率生产成本核算：替代算法的研究，Int. J. Electr. Power Energy Syst. 5(2), 75–86.[6] Eliwa, M. S., Ahsan-ul-Haq, M., Al-Bossly, A., El-Morshedy, M. (2022). 一种用于比例与非对称数据的单位概率模型：性质、估计方法及其在 SC16 数据和 P3算法建模中的应用，Math. Probl. Eng. 2022, 1–13. https://doi.org/10.1155/2022/9289721[7]Emmanuel Afuecheta, Idika E. Okorie, Haady Jallow, Saralees Nadarajah. (2025). 单位连续概率分布综述[J]. AIMS Mathematics, 10(11): 25939-26057. https://doi:10.3934/math.20251146[8]Fayomi, A., Hassan, A. S., Almetwally, E. M. (2023). 单位指数化 Lomax 分布的推断与分位数回归。PLOS ONE 18(7): e0288635. https://doi.org/10.1371/journal.pone.0288635

[9] Genç, M., Özbilen, Ö. (2023). 幂指数化UEHL分布：性质与应用。雷杰普·塔伊普·埃尔多安大学科学与工程学报，4(2), 232-241. https://doi.org/10.53501/rteufemud.1388416

[10]Hashmi, S., Ahsan-ul Haq, M., Zafar, J., Khaleel, M. (2022). 单位X伽马分布：其性质、估计与应用。《巴基斯坦科学院学报：A辑，物理与计算科学》，59(1)，15–28。https://doi.org/10.53560/PPASA(59-1)636

[11] Hassan, A. S., Fayomi, A., Algarni, A., Almetwally, E. M. (2022). 单位幂指数化半逻辑斯蒂分布的贝叶斯与非贝叶斯推断及数据分析。《应用科学》，12(21)，11253。https://doi.org/10.3390/app122111253[12] Johnson, N. L., Kotz, S., Balakrishnan, N. (1955). 连续一元分布。John Wiley & Sons, INC.[13] Krishna, A., Maya, R., Chesneau, C., Irshad, M. R. (2022). 单位Teissier分布及其应用。《数学与计算应用》2022，27(1)，12。https://doi.org/10.3390/mca27010012[14]Kumaraswamy, P. (1980). 双界随机过程的广义概率密度函数。《水文学杂志》，46(1-2)，79-88。https://doi.org/10.1016/0022-1694(80)90036-0[15] Mazucheli, J., Menezes, A., Ghitany, M. (2018). 单位威布尔分布及相关推断。《应用概率与统计学杂志》，1–22。[16] Murthy, D. N. P. , Xie, M., (2004). Jiang, R. 威布尔模型，John Wiley & Sons INC., Hoboken, NJ.[17]Nasiru, S., Chesneau, C., Ocloo, S. K. (2024). 对数余弦幂单位分布：一种用于比例数据分析的新单位分布，《决策分析杂志》，10，100397。https://doi.org/10.1016/j.dajour.2024.100397[18] Sarhan, A. M., Sobh M. E., (2025). 带应用的单位幂指数化威布尔模型，《非洲科学》27，e02606, https://doi.org/10.1016/j.sciaf.2025.e02606[19] Shiha, F. A. (2026) Shiha分布：统计性质及其在可靠性工程和环境数据中的应用。https://doi.org/10.48550/arXiv.2602.02875[20] Sindhu, T. N., Shafiq, A., Riaz, M. B., Abushal, T. A. (2024). 一种新型双参数单位 Bilal 分布的统计框架及其在非对称数据建模中的应用。《Heliyon》，第11卷(16)，e38340[21] Topp, C. W., Leone, F. C. (1955). 一族J形频率函数。《美国统计学会杂志》，50，209–219。