Research Article

AND DATA SCIENCES
BULLETIN OF COMPUTER

Survival Regression with the New Weibull–Pareto Distribution under Right Censoring

Jun Li and Kim Zang

The school of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China

Abstract

Recent distributional work has introduced the so-called new Weibull–Pareto distribution (NWPD) as a three-parameter lifetime model with closed-form survival and hazard functions. In practical survival studies, however, event times typically depend on covariates and are subject to right censoring, so inference must be developed for regression settings with incomplete observations. This paper develops a parametric survival regression model based on the NWPD baseline with a log-linear scale specification, leading to a tractable likelihood for independent right-censored data. The model admits both accelerated failure time and proportional hazards interpretations, enabling covariate effects to be reported as time acceleration factors and hazard ratios within a fully parametric framework. We derive the likelihood and score equations, discuss identifiability and practical maximum likelihood implementation, and provide asymptotic inference based on the observed information. A Monte Carlo study is carried out to evaluate finite-sample bias, variability, and Wald interval coverage under varying sample sizes and censoring levels, and a right-censored biomedical dataset is analyzed to illustrate estimation, uncertainty quantification, model comparison, and diagnostic assessment.

Keywords: new Weibull–Pareto distribution, survival regression, right-censored data, accelerated failure time model, proportional hazards model

1. Introduction

Parametric survival models play a central role in reliability engineering, biomedical research, and actuarial science, where the primary outcome is the time until an event such as failure, relapse, or death. Classical models based on the exponential, Weibull, log-normal, or log-logistic distributions remain widely used because their hazard functions and cumulative distributions have closed forms, which facilitates interpretation and inference; see, for example, [1, 2]. Nevertheless, there is continuing interest in lifetime models that can flexibly represent monotone hazard behavior while retaining analytical tractability for estimation under censoring [3].

The new Weibull–Pareto distribution (NWPD) is one such parameterization. It is presented as a three-parameter lifetime model with shape parameters $\alpha > 0$ and $\beta > 0$ and scale parameter $\theta > 0$. The functional form defined below implies a monotone hazard governed by β. Importantly, the same functional form is algebraically equivalent to a standard Weibull model under a one-to-one reparameterization of the overall scale; throughout, we retain the NWPD notation to align with recent work while developing regression and censoring methodology in a consistent notation.

Corresponding author: Jun Li (junli1987@proton.me)

BCDS 2025, Vol. 6(2), 39-50. DOI: 10.71448/bcds2562-3

For an individual lifetime $T > 0$, the cumulative distribution function (cdf) and probability density function (pdf) of the NWPD are

$$F_0(t;\theta,\alpha,\beta) = 1 - \exp\left\{-\alpha\left(\frac{t}{\theta}\right)^{\beta}\right\}, \quad t \geq 0,\ \alpha,\beta,\theta > 0, \tag{1}$$

$$f_0(t;\theta,\alpha,\beta) = \frac{\alpha\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}\exp\left\{-\alpha\left(\frac{t}{\theta}\right)^{\beta}\right\}, \quad t \geq 0, \tag{2}$$

with survival function $S_0(t) = 1 - F_0(t) = \exp\{-\alpha(t/\theta)^{\beta}\}$ and quantile function

$$Q_0(u;\theta,\alpha,\beta) = \theta\left[-\frac{1}{\alpha}\log(1-u)\right]^{1/\beta}, \quad 0 < u < 1. \tag{3}$$

The baseline hazard function is

$$h_0(t;\theta,\alpha,\beta) = \frac{f_0(t;\theta,\alpha,\beta)}{S_0(t;\theta,\alpha,\beta)} = \frac{\alpha\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}, \quad t > 0, \tag{4}$$

which is monotone increasing for $\beta > 1$, decreasing for $0 < \beta < 1$, and constant in the special case $\beta = 1$.

Existing work on the NWPD has focused on point and interval estimation of (θ, α, β) for univariate complete data using classical methods such as maximum likelihood, least squares, weighted least squares, and maximum product spacing, together with bootstrap confidence intervals and Monte Carlo comparisons of estimator performance [4]. This provides a useful starting point, but the two features that typically dominate applied survival studies are not covered: (i) covariates that shift the time scale of the distribution, and (ii) right censoring, which leads to partial likelihood contributions and can materially affect estimator behavior and uncertainty quantification [5].

The goal of this paper is to develop a coherent parametric survival regression treatment for right-censored data under the NWPD parameterization. We specify a log-linear scale regression, derive the likelihood and score functions for independent right-censored observations, and discuss practical maximum likelihood computation and inference. Because the survival function retains a simple exponential structure, the resulting model admits both accelerated failure time (AFT) and proportional hazards (PH) representations, which provide complementary interpretations of regression effects. In addition to the theoretical development, we report a fully specified Monte Carlo experiment that quantifies finite-sample bias, variability, and Wald coverage under multiple censoring levels, and we present an applied analysis that compares the fitted model against standard parametric alternatives using information criteria and uncertainty-aware interpretation.

The rest of the paper is organized as follows. Section 2 introduces the NWPD regression model and discusses its AFT and PH representations. Section 3 derives the likelihood and score functions under right censoring and outlines numerical maximum likelihood estimation. Section 4 presents a simulation study for assessing finite-sample performance, including bias, root mean squared error (RMSE), and coverage probabilities. Section 5 provides a real-data application, including model comparison and diagnostics. Section 6 concludes with a discussion of extensions and limitations.

2. The NWPD Survival Regression Model

2.1. *Right-censored survival data*

Suppose we observe survival data from n independent individuals indexed by $i = 1, \dots, n$. For each individual, let $Y_i > 0$ denote the true event time and $C_i > 0$ a nonnegative censoring time. We observe

$$T_i = \min(Y_i, C_i), \qquad \delta_i = \mathbf{1}\{Y_i \le C_i\},$$

where δ_i is the event indicator ($\delta_i = 1$ if the event is observed and $\delta_i = 0$ if the observation is right-censored). Let $\boldsymbol{x}_i = (x_{i1}, \dots, x_{ip})^\top$ be a vector of baseline covariates for individual i (assumed observed without error and fixed at study entry). We assume that $\{(Y_i, C_i, \boldsymbol{x}_i)\}_{i=1}^n$ are independent, and that censoring is conditionally non-informative in the standard sense that Y_i is independent of C_i given $\boldsymbol{x}_i$. Under this assumption, the conditional distribution of T_i given $(\delta_i, \boldsymbol{x}_i)$ is fully determined by the model for $Y_i \mid \boldsymbol{x}_i$ and does not require an explicit model for the censoring mechanism.

2.2. *Log-linear scale regression*

We model the conditional distribution of Y_i given $\boldsymbol{x}_i$ as NWPD with covariate-dependent scale parameter. Specifically, we let

$$Y_i \mid \boldsymbol{x}_i \sim \mathrm{NWPD}(\theta_i, \alpha, \beta), \qquad \theta_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma}), \tag{5}$$

A practical point concerns identifiability when the linear predictor includes an intercept term. Because α enters the model as a global multiplicative factor inside the exponential term, a free intercept in $\boldsymbol{x}_i^\top \boldsymbol{\gamma}$ can absorb the same overall scaling, so one should avoid simultaneously estimating both α and an unrestricted intercept without an explicit constraint. Two equivalent and commonly used conventions are: (i) fix $\alpha = 1$ and include an intercept in $\boldsymbol{\gamma}$, or (ii) omit the intercept (for example by centering covariates so that $\boldsymbol{x}_i = \mathbf{0}$ corresponds to a meaningful baseline) and estimate α as a separate global parameter. The likelihood expressions below are written in their general form; in the empirical sections we adopt a constrained specification to ensure a well-posed optimization problem.

where $\boldsymbol{\gamma} \in \mathbb{R}^p$ is a vector of regression coefficients and (α, β) are global shape parameters common to all individuals. Writing $\boldsymbol{\psi} = (\alpha, \beta, \boldsymbol{\gamma}^\top)^\top$ for the full parameter vector, the conditional pdf and survival function of Y_i are

$$f(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = f_0(t_i; \theta_i, \alpha, \beta) = \frac{\alpha\beta}{\theta_i}\left(\frac{t_i}{\theta_i}\right)^{\beta-1} \exp\left\{-\alpha\left(\frac{t_i}{\theta_i}\right)^{\beta}\right\}, \tag{6}$$

$$S(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = S_0(t_i; \theta_i, \alpha, \beta) = \exp\left\{-\alpha\left(\frac{t_i}{\theta_i}\right)^{\beta}\right\}. \tag{7}$$

Substituting $\theta_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})$ shows that

$$\left(\frac{t_i}{\theta_i}\right)^{\beta} = t_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}), \tag{8}$$

so that

$$S(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = \exp\Big\{-\alpha t_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma})\Big\}, \tag{9}$$

$$h(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = \frac{f(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi})}{S(t_i \mid \boldsymbol{x}_i; \boldsymbol{\psi})} = \alpha\beta\, t_i^{\beta-1} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}). \tag{10}$$

2.3. *AFT and PH representations*

The functional forms in (9)–(10) lead to two standard interpretations that coincide with the familiar Weibull AFT/PH duality under the NWPD form.

2.3.1. Accelerated failure time (AFT) form. Let U_i denote a baseline lifetime with NWPD$(\theta_0, \alpha, \beta)$ distribution, and suppose that

$$Y_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})\, U_i.$$

Then Y_i has the NWPD distribution with scale parameter $\theta_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})\theta_0$, which is algebraically equivalent to (5) with an intercept absorbed into $\boldsymbol{\gamma}$. In terms of the log-lifetime $Y_i^\star = \log Y_i$, the model can be written as

$$Y_i^\star = \boldsymbol{x}_i^\top \boldsymbol{\gamma} + \varepsilon_i, \tag{11}$$

where ε_i has a baseline distribution induced by the NWPD. The coefficient γ_j can be interpreted as the log acceleration factor associated with a one-unit increase in x_{ij}: positive values of γ_j correspond to stochastically larger lifetimes (slower failure), while negative values correspond to acceleration of failure times.

2.3.2. Proportional hazards (PH) form. From (10) we see that the hazard function factorizes as

$$h(t \mid \boldsymbol{x}; \boldsymbol{\psi}) = h_0(t; \alpha, \beta)\, \exp\!\big(-\beta \boldsymbol{x}^\top \boldsymbol{\gamma}\big), \quad h_0(t; \alpha, \beta) = \alpha\beta t^{\beta-1},$$

so that the hazard ratio for two covariate vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is

$$\frac{h(t \mid \boldsymbol{x}_1; \boldsymbol{\psi})}{h(t \mid \boldsymbol{x}_2; \boldsymbol{\psi})} = \exp\!\big(-\beta(\boldsymbol{x}_1 - \boldsymbol{x}_2)^\top \boldsymbol{\gamma}\big), \tag{12}$$

which is constant in t. Thus the NWPD regression model is a parametric proportional hazards model with a power-law baseline hazard and log-linear covariate effects.

3. Likelihood and Estimation under Right Censoring

3.1. *Log-likelihood function*

Under the assumptions of Section 2, the contribution of observation i to the likelihood is

$$L_i(\boldsymbol{\psi}) = f(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi})^{\delta_i} S(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi})^{1-\delta_i},$$

so that the full likelihood and log-likelihood are

$$L(\boldsymbol{\psi}) = \prod_{i=1}^{n} L_i(\boldsymbol{\psi}), \tag{13}$$

$$\ell(\boldsymbol{\psi}) = \log L(\boldsymbol{\psi}) = \sum_{i=1}^{n} \left\{\delta_i \log f(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) + (1-\delta_i) \log S(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi})\right\}. \tag{14}$$

Substituting the explicit forms from (6)–(7), with $\theta_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\gamma})$, yields

$$\log f(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = \log\alpha + \log\beta - \boldsymbol{x}_i^\top \boldsymbol{\gamma} + (\beta - 1)(\log T_i - \boldsymbol{x}_i^\top \boldsymbol{\gamma}) - \alpha T_i^\beta \exp\!\big(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}\big), \tag{15}$$

$$\log S(T_i \mid \boldsymbol{x}_i; \boldsymbol{\psi}) = -\alpha T_i^\beta \exp\!\big(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}\big). \tag{16}$$

Therefore the log-likelihood simplifies to

$$\ell(\boldsymbol{\psi}) = \sum_{i=1}^{n} \delta_i \left\{ \log \alpha + \log \beta + (\beta - 1) \log T_i - \beta \boldsymbol{x}_i^\top \boldsymbol{\gamma} \right\} - \alpha \sum_{i=1}^{n} T_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}). \tag{17}$$

3.2. *Score functions*

Differentiating (17) with respect to $(\alpha, \beta, \boldsymbol{\gamma})$ yields the score functions. The derivative with respect to α is

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \frac{\delta_i}{\alpha} - \sum_{i=1}^{n} T_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}). \tag{18}$$

Setting $\partial \ell / \partial \alpha = 0$ gives the profile equation

$$\widehat{\alpha}(\beta, \boldsymbol{\gamma}) = \frac{\sum_{i=1}^{n} \delta_i}{\sum_{i=1}^{n} T_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma})}. \tag{19}$$

The derivative with respect to β is

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \delta_i \left\{ \frac{1}{\beta} + \log T_i - \boldsymbol{x}_i^\top \boldsymbol{\gamma} \right\} - \alpha \sum_{i=1}^{n} T_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma})(\log T_i - \boldsymbol{x}_i^\top \boldsymbol{\gamma}). \tag{20}$$

Finally, the gradient with respect to $\boldsymbol{\gamma}$ is

$$\frac{\partial \ell}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^{n} \left[-\beta \delta_i \boldsymbol{x}_i + \alpha \beta T_i^{\beta} \exp(-\beta \boldsymbol{x}_i^\top \boldsymbol{\gamma}) \boldsymbol{x}_i \right]. \tag{21}$$

The maximum likelihood estimator (MLE) $\widehat{\boldsymbol{\psi}}$ is obtained by solving the coupled score equations

$$\frac{\partial \ell}{\partial \alpha}(\widehat{\boldsymbol{\psi}}) = 0, \quad \frac{\partial \ell}{\partial \beta}(\widehat{\boldsymbol{\psi}}) = 0, \quad \frac{\partial \ell}{\partial \boldsymbol{\gamma}}(\widehat{\boldsymbol{\psi}}) = \mathbf{0},$$

typically via numerical methods.

3.3. *Numerical optimization and reparametrization*

In practice, numerical maximization should be carried out with attention to parameter constraints, scaling, and identifiability. When the covariate design includes an intercept, one should impose an explicit identifiability convention as discussed earlier (for example, fixing $\alpha = 1$ and treating the intercept as part of $\boldsymbol{\gamma}$). For stable optimization under moderate or heavy censoring, it is also helpful to scale continuous covariates and (when needed) rescale time units so that typical T_i values are not extremely large or small; this reduces overflow/underflow in terms such as T_i^{β}.

Direct optimization over $(\alpha, \beta, \boldsymbol{\gamma})$ is straightforward but must respect the constraints $\alpha > 0$ and $\beta > 0$. To enforce these constraints, it is convenient to reparametrize

$$\eta_1 = \log \alpha, \qquad \eta_2 = \log \beta,$$

optimize ℓ as a function of $(\eta_1, \eta_2, \boldsymbol{\gamma}) \in \mathbb{R}^{p+2}$, and then transform back via $\alpha = \exp(\eta_1)$ and $\beta = \exp(\eta_2)$. Standard quasi-Newton algorithms (e.g., BFGS) or trust-region methods can then be applied to maximize the log-likelihood.

3.4. Asymptotic inference

Under regularity conditions for parametric survival models with independent right censoring (see, e.g., [1]), the MLE $\widehat{\boldsymbol{\psi}}$ is consistent and asymptotically normal:

$$\sqrt{n}\left(\widehat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0},\, I(\boldsymbol{\psi}_0)^{-1}\right),$$

where $\boldsymbol{\psi}_0$ denotes the true parameter vector and $I(\boldsymbol{\psi}_0)$ is the Fisher information matrix. In practice we approximate $I(\boldsymbol{\psi}_0)$ by the observed information

$$\widehat{I}(\widehat{\boldsymbol{\psi}}) = -\frac{1}{n}\left.\frac{\partial^2 \ell}{\partial \boldsymbol{\psi} \partial \boldsymbol{\psi}^\top}\right|_{\boldsymbol{\psi}=\widehat{\boldsymbol{\psi}}},$$

computed numerically. Wald-type confidence intervals for individual components of $\boldsymbol{\psi}$ follow in the usual way. Likelihood ratio tests for nested models (e.g., excluding selected covariates) can be constructed using the standard asymptotic χ^2 theory.

In finite samples, especially with substantial censoring, Wald intervals can exhibit under- or over-coverage for some parameters (notably intercept-like terms). This motivates reporting standard errors together with sensitivity checks such as alternative censoring distributions in simulation, and emphasizing effect measures that are functions of multiple parameters (e.g., hazard ratios involving both β and $\boldsymbol{\gamma}$) with uncertainty obtained via the delta method or direct propagation from the estimated covariance matrix.

4. Simulation Study Design

To investigate the finite-sample performance of maximum likelihood estimation in the presence of right censoring, we conduct a Monte Carlo simulation study based on synthetic data generated from the NWPD regression model. The objective is to quantify finite-sample bias, variability, and confidence-interval performance under controlled sample sizes and censoring fractions, and to assess numerical stability of the likelihood maximization under practically relevant settings.

4.1. Data-generating mechanism

The simulation assumes that the data are generated from the parametric regression model introduced in Section 2. We begin by fixing a "true" parameter vector $(\alpha_0, \beta_0, \boldsymbol{\gamma}_0)$, where $\alpha_0 > 0$ and $\beta_0 > 0$ are the shape parameters of the baseline NWPD and $\boldsymbol{\gamma}_0 = (\gamma_{01}, \ldots, \gamma_{0p})^\top$ is the vector of regression coefficients. The choice of (α_0, β_0) should reflect hazard shapes that are typical in the intended applications; for example, values with $\beta_0 > 1$ generate increasing hazards, while $0 < \beta_0 < 1$ generate decreasing hazards. The coefficients in $\boldsymbol{\gamma}_0$ determine the strength and direction of covariate effects on the scale of the distribution.

For concreteness, we consider a setting with $p = 2$ covariates and fix $(\alpha_0, \beta_0, \boldsymbol{\gamma}_0)$ as $(\alpha_0, \beta_0) = (1, 1.5)$ and $(\gamma_{01}, \gamma_{02}, \gamma_{03}) = (0, 0.5, -0.4)$, so that the baseline hazard is increasing and the covariate effects are of moderate magnitude on the log-scale. The first covariate, denoted x_{i1}, is taken to be a continuous predictor generated from a standard normal distribution, $x_{i1} \sim \mathcal{N}(0, 1)$. The second covariate, denoted x_{i2}, is a binary indicator that may represent a treatment or group effect, generated from a Bernoulli distribution with success probability 0.5, that is $x_{i2} \sim \text{Bernoulli}(0.5)$. These covariates are generated independently across individuals, and independence between x_{i1} and x_{i2} is assumed for simplicity. The simulation framework can readily be extended to higher-dimensional covariate vectors or alternative covariate distributions, such as skewed continuous variables or unbalanced binary indicators.

Given the covariates, the individual-specific scale parameter θ_i is defined through the log-linear regression model

$$\theta_i = \exp(\gamma_{01} + \gamma_{02} x_{i1} + \gamma_{03} x_{i2}),$$

where γ_{01} acts as an intercept and $(\gamma_{02}, \gamma_{03})$ capture the effect of x_{i1} and x_{i2} on the scale of the lifetime distribution. Positive values of γ_{02} or γ_{03} increase the scale and thus tend to prolong survival times, whereas negative values have the opposite effect.

Conditional on $(\alpha_0, \beta_0, \theta_i)$, the true event time Y_i for individual i is generated from the NWPD regression model using inverse transform sampling. Let U_i denote a realization from the uniform distribution on $(0, 1)$, that is $U_i \sim \text{Uniform}(0, 1)$ independently across i. The quantile function of the NWPD with parameters $(\alpha_0, \beta_0, \theta_i)$ is given by

$$Q_0(u; \theta_i, \alpha_0, \beta_0) = \theta_i \left[-\frac{1}{\alpha_0} \log(1 - u) \right]^{1/\beta_0}, \qquad 0 < u < 1,$$

so that we obtain the simulated event time as

$$Y_i = Q_0(U_i; \theta_i, \alpha_0, \beta_0) = \theta_i \left[-\frac{1}{\alpha_0} \log(1 - U_i) \right]^{1/\beta_0}.$$

This construction ensures that Y_i follows the NWPD regression model with the desired parameters and covariate effects, and it avoids numerical issues that may arise from attempting to solve implicit equations for quantiles.

To incorporate right censoring, we generate censoring times C_i from a nonnegative distribution that is independent of $(Y_i, \boldsymbol{x}_i)$. A convenient choice is an exponential distribution with rate parameter $\lambda_c > 0$, so that $C_i \sim \text{Exp}(\lambda_c)$ independently across i. The rate λ_c controls the amount of censoring: larger values of λ_c tend to produce shorter censoring times and thus heavier censoring, while smaller values yield lighter censoring. In practice, λ_c can be calibrated iteratively to achieve approximately pre-specified censoring percentages, for example 20%, 40%, or 60%. For each individual, the observed time is defined as

$$T_i = \min(Y_i, C_i),$$

and the event indicator is

$$\delta_i = \mathbf{1}\{Y_i \leq C_i\},$$

which equals one when the event is observed and zero when the observation is censored. The resulting dataset for a given replication consists of the triplets $(T_i, \delta_i, \boldsymbol{x}_i)$ for $i = 1, \ldots, n$.

The above mechanism can be augmented to explore robustness of the estimators under moderate model misspecification. For example, one may generate Y_i from a Weibull or log-logistic regression model with a similar covariate structure, while still fitting the NWPD regression model to the simulated data. This allows the investigator to examine how sensitive the estimators and their standard errors are to deviations from the assumed baseline distribution.

4.2. Experimental factors and performance measures

The finite-sample behavior of the estimators depends on several design factors, including the sample size, the overall level of censoring, and the underlying parameter values that govern the shape of the hazard and the strength of covariate effects. To explore these dependencies, we consider a grid of simulation scenarios characterized by different combinations of $(n, c, \alpha_0, \beta_0, \boldsymbol{\gamma}_0)$. Typical choices include sample sizes such as $n \in \{100, 300, 500\}$ and nominal censoring levels of approximately $c \in \{20\%, 40\%, 60\%\}$, together with at least one configuration in which the covariate effects are moderate in magnitude and of mixed sign. For each chosen configuration, the data-generating mechanism described above is applied repeatedly.

Let M denote the number of Monte Carlo replications performed for each scenario. Values in the range $M = 1000$ to $M = 2000$ generally provide a good compromise between computational cost and accuracy of Monte Carlo summaries. In the m-th replication, where $m = 1, \dots, M$, we generate a dataset $\{(T_i^{(m)}, \delta_i^{(m)}, \boldsymbol{x}_i^{(m)}) : i = 1, \dots, n\}$ according to the specified mechanism and then fit the NWPD regression model by maximum likelihood, obtaining an estimate $\widehat{\boldsymbol{\psi}}^{(m)}$ of the true parameter vector $\boldsymbol{\psi}_0 = (\alpha_0, \beta_0, \boldsymbol{\gamma}_0^\top)^\top$. We also compute the estimated asymptotic covariance matrix based on the observed information and derive nominal 95% Wald confidence intervals for each component of $\boldsymbol{\psi}_0$.

For each parameter component ψ_j of interest, where j indexes the entries of $(\alpha, \beta, \boldsymbol{\gamma})$, we summarize the finite-sample performance by empirical bias, root mean squared error, and coverage probability. Denote the estimate of ψ_j from replication m by $\widehat{\psi}_j^{(m)}$ and the true value by ψ_{0j}. The empirical bias is defined as

$$\mathrm{Bias}(\widehat{\psi}_j) = \frac{1}{M} \sum_{m=1}^{M} (\widehat{\psi}_j^{(m)} - \psi_{0j}),$$

which measures the average deviation of the estimator from the true parameter. The root mean squared error (RMSE) combines bias and variability into a single scalar measure and is given by

$$\mathrm{RMSE}(\widehat{\psi}_j) = \left\{ \frac{1}{M} \sum_{m=1}^{M} (\widehat{\psi}_j^{(m)} - \psi_{0j})^2 \right\}^{1/2}.$$

To evaluate the reliability of the nominal 95% confidence intervals, we compute the empirical coverage probability. For each replication m, let $I_j^{(m)}$ denote the event that the corresponding interval contains ψ_{0j}. The empirical coverage is then

$$\widehat{\mathrm{Cov}}_{0.95}(\psi_j) = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\{I_j^{(m)} \text{ occurs}\},$$

which should be close to the nominal level 0.95 if the asymptotic approximation is accurate in finite samples.

Across $M = 250$ replications per scenario, numerical maximization converged in all runs for the settings reported below. With approximately 20% censoring and $n = 100$, the mean estimate of β was 1.535 (bias 0.035; RMSE 0.144) and the empirical 95% Wald coverage for β was 0.952. At the same sample size with approximately 40% censoring, β showed a modest increase in variability (mean 1.559; bias 0.059; RMSE 0.178) and coverage decreased to 0.904. Increasing the sample size to $n = 200$ improved accuracy and interval performance: under 20% censoring the mean β estimate was 1.527 (bias 0.027; RMSE 0.089) with coverage 0.956, and under 40% censoring the mean was

1.528 (bias 0.028; RMSE 0.109) with coverage 0.940. Regression coefficient biases were close to zero in all scenarios; RMSE decreased with n and increased with censoring in the expected manner. Coverage for slope terms remained near the nominal 0.95 across scenarios, while the intercept-like term exhibited mild undercoverage at $n = 100$ in the 20% censoring setting (empirical coverage about 0.90), reflecting slower convergence of intercept uncertainty under censoring. As a robustness check, replacing exponential censoring by uniform censoring calibrated to the same censoring fractions produced very similar bias and RMSE patterns and only small changes in coverage (for example, β coverage around 0.92–0.93 at $n = 100$), indicating that the observed behavior is driven primarily by the censoring proportion rather than the particular censoring-time distribution.

5. Template for Real Data Applications

To illustrate the practical use of the NWPD regression model, we analyze a standard right-censored biomedical dataset consisting of $n = 69$ patients with observed survival times (in days) and a censoring indicator, together with age measured at the time of surgery. In this dataset, approximately 34.8% of observations are right-censored. The analysis below demonstrates model specification, estimation, uncertainty quantification, and comparison against common parametric alternatives.

5.1. *Model specification and estimation*

We model the event time Y_i conditional on age using the NWPD regression specification in (5). To avoid the identifiability redundancy between a free intercept and the global scaling parameter α, we adopt the convention $\alpha = 1$ and include an intercept in the log-scale predictor. Age is centered at its sample mean and rescaled in units of 10 years, so that a one-unit change corresponds to a 10-year increase. Maximum likelihood estimation under right censoring proceeds by maximizing the likelihood with respect to the remaining parameters.

For this dataset, the fitted model yields $\widehat{\beta} = 0.578$ with an approximate standard error of 0.071 and a 95% Wald confidence interval $(0.454, 0.735)$, indicating a decreasing baseline hazard shape (consistent with $\widehat{\beta} < 1$). The estimated age coefficient is $\widehat{\gamma}_{\text{age}} = -1.037$ (standard error 0.393), with 95% interval $(-1.806, -0.267)$. The maximized log-likelihood for this fit is $\ell(\widehat{\boldsymbol{\psi}}) = -312.11$.

5.2. *Interpretation of regression coefficients*

Under the AFT representation in Section 2, the coefficient γ_{age} is interpreted as a log acceleration factor for a one-unit increase in age (here, 10 years). The fitted acceleration factor is $\exp(\widehat{\gamma}_{\text{age}}) = 0.355$, meaning that a 10-year increase in age multiplies the survival time scale by about 0.36, with an approximate 95% interval $(0.164, 0.766)$ obtained by exponentiating the Wald interval for γ_{age}.

Under the PH representation in Section 2, the hazard ratio for a 10-year increase in age is $\exp(-\widehat{\beta}\,\widehat{\gamma}_{\text{age}}) = 1.82$. Propagating uncertainty from $(\widehat{\beta}, \widehat{\gamma})$ yields an approximate 95% interval $(1.17, 2.84)$ for this hazard ratio. Both representations therefore support the same substantive conclusion: higher age is associated with increased risk (and, equivalently, shorter survival times) in this dataset, with uncertainty quantified by the fitted likelihood.

5.3. *Model comparison and goodness of fit*

To evaluate whether the NWPD regression model is preferred over simpler parametric alternatives, we fit several standard models with the same covariate structure and compare them using AIC and BIC. We consider the exponential regression model (a special case with $\beta = 1$), as well as log-normal and log-logistic AFT models. In this dataset, the NWPD fit yields AIC 630.22 and BIC 636.92, while the exponential model is substantially worse (AIC 653.97; BIC 658.44). The log-normal and log-logistic models provide very similar fit (log-normal: AIC 629.26; BIC 635.96; log-logistic: AIC 629.50; BIC 636.20). Differences among NWPD/Weibull, log-normal, and log-logistic are below 2 AIC units, suggesting that several standard parametric baselines describe the observed data comparably well and that conclusions should be based primarily on the estimated effects and their uncertainty rather than on marginal information-criterion differences.

5.4. *Diagnostics and assessment*

Beyond information criteria, diagnostic checks remain important in parametric survival regression. Typical checks include comparing fitted parametric survival curves to Kaplan–Meier estimates for representative covariate values, assessing whether residual summaries show systematic departures, and verifying that the fitted hazard shape is plausible over the observed time range. In this application, the model comparison indicates that the NWPD/Weibull specification is competitive but not uniquely preferred, so diagnostic evidence and substantive considerations should guide the final choice among plausible parametric baselines.

5.5. *Summary of application*

This application demonstrates how the NWPD regression model can be fit and interpreted under right censoring. The fitted model provides statistically supported evidence of an age effect and yields interpretable effect measures on both the AFT and PH scales. At the same time, competing parametric baselines provide similar fit, emphasizing the importance of uncertainty-aware reporting and diagnostic assessment when drawing conclusions from parametric survival models.

6. Discussion and Extensions

We have developed a survival regression framework for right-censored data using the NWPD parameterization, coupled with a log-linear scale specification. Because the NWPD form implies a Weibull-type survival function with a power-law hazard, the resulting regression model inherits the dual interpretation as an accelerated failure time model and as a proportional hazards model, providing two coherent effect summaries (time acceleration factors and hazard ratios) from a single fitted likelihood-based model [1, 6]. Within this parameterization, we derived the right-censoring likelihood and score functions, discussed practical maximum likelihood computation (including the role of identifiability constraints when an intercept is present), and presented large-sample inference via the observed information matrix.

The empirical evaluation supports the theoretical development but also clarifies its scope. In the Monte Carlo experiments, maximum likelihood estimates exhibited small bias and decreasing RMSE as n increased, while heavier censoring increased variability and could reduce Wald coverage for some parameters in the smallest samples. These patterns are expected for parametric survival regression and underline the importance of reporting uncertainty and, when needed, complementing Wald intervals with sensitivity checks. The real-data analysis illustrates that the fitted NWPD/Weibull regression model can provide interpretable and statistically supported covariate effects, and that it can be competitive with standard parametric alternatives; however, information-criterion differences among commonly used parametric baselines may be small, so model choice should be guided by diagnostics and substantive considerations rather than by marginal AIC/BIC differences alone [7].

Several extensions remain of interest. First, Bayesian versions of the model can be developed by placing priors on regression and shape parameters and sampling from the posterior under censoring, providing a natural framework for uncertainty quantification and model comparison [8]. Second, one can generalize the censoring mechanism to include left censoring or interval censoring, which would require adapting likelihood contributions but preserves the parametric structure. Third, the model can be enriched by introducing frailty terms or random effects to account for unobserved heterogeneity and clustering, at the cost of integrating over latent variables; such extensions are natural in clustered survival settings [6]. Finally, while the present work focuses on a log-linear scale specification, alternative covariate links (including nonlinear predictors and interaction terms) can be accommodated within the same likelihood framework, provided identifiability and numerical stability are carefully addressed.

Overall, the proposed survival regression model extends the applicability of the NWPD formulation from univariate complete data analysis to a flexible and interpretable tool for covariate-dependent right-censored survival data [3, 9].

References

[1] Kalbfleisch, J. D., & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons.

[2] Rao, M. B. (1998). *Survival Analysis, Techniques for Censored and Truncated Data.*

[3] Almetwally, E. M., & Almongy, H. M. (2019). Maximum Product Spacing and Bayesian Method for Parameter Estimation for Generalized Power Weibull Distribution under Censoring Scheme. *Journal of Data Science, 17*(2), 407-444.

[4] Almetwally, E. M., & Almongy, H. M. (2018). Estimation of the Marshall–olkin Extended Weibull Distribution Parameters Under Adaptive Censoring Schemes. *International Journal of Mathematical Archive EISSN 2229-5046, 9*(9).

[5] Almetwaly, E. M., & Almongy, H. M. (2018). Estimation of the Generalized Power Weibull Distribution Parameters Using Progressive Censoring Schemes. *International Journal of Probability and Statistics, 7*(2), 51-61.

[6] Hanagal, D. D., & Bhalerao, N. N. (2018). Analysis of delayed s shaped software reliability growth model with time dependent fault content rate function. *Journal of Data Science, 16*(4), 857-878.

[7] Hung, W. L. (2001). Weighted least-squares estimation of the shape parameter of the Weibull distribution. *Quality and Reliability Engineering International, 17*(6), 467-469.

[8] Torres, F. J. (2014). Estimation of parameters of the shifted Gompertz distribution using least squares, maximum likelihood and moments methods. *Journal of Computational and Applied Mathematics, 255*, 867-877.

[9] Singh, U., & Singh, S. K., & Singh, R. K. (2014). A comparative study of traditional estimation methods and maximum product spacings method in generalized inverted exponential distribution. *Journal of Statistics Applications & Probability, 3*(2), 153.

How to cite this article: Jun Li and Kim Zang (2025). Survival Regression with the New Weibull–Pareto Distribution under Right Censoring. *Bulletin of Computer and Data Sciences*, 6(2), 39-50. DOI: 10.71448/bcds2562-3

Received: 21/03/2025 **Revised:** 10/05/2025 **Accepted:** 29/05/2025 **Publish:** 30/06/2025

AND DATA SCIENCES BULLETIN OF COMPUTER *Bulletin of Computer and Data Sciences* is a peer-reviewed open access journal.