Article

Bayesian Analysis of Exponentiated-Weibull Weibull Distribution for Modelling Student Dropout Times

Mohammed R. Alzahrani [*]

Article

Bayesian Analysis of Exponentiated-Weibull Weibull Distribution for Modelling Student Dropout Times

Mohammed R. Alzahrani (iD)

Department of Psychology, Faculty of Education, Umm Al-Qura University, Al-Abidiyah, Makkah, Saudi Arabia; mrzahrani@uqu.edu.sa

Abstract: This paper presents the five-parameter Exponentiated Weibull Weibull (EWW) distribution with Weibull base failure rate parameterization, developed to model the dropout times of students in an online class module. Parameters of the model were estimated using both Maximum Likelihood and Bayesian estimation procedures to determine the most effective estimation method. Simulation results indicate that, across various hazard types and censoring levels, Bayesian estimates are more accurate and precise than Maximum Likelihood estimates. Real-life data analysis of online students module presentation supports these findings, showing significantly lower standard deviations for Bayesian estimates compared to Maximum Likelihood estimates, highlighting the efficiency of Bayesian estimation for the EWW distribution. Additionally, a comparative goodness-of-fit analysis demonstrated that the EWW distribution better fits the student dropout times dataset than Weibull Weibull (WW), Weibull Exponential (WE), and Exponential Weibull (EW) distributions. The median dropout time estimated using the EWW distribution is 24 days, suggesting that instructors should closely monitor student attendance around this period.

1. Introduction

The issue of student dropout in online education has become a critical concern for educators and institutions alike. Online education offers unparalleled flexibility and accessibility, allowing students from diverse backgrounds to pursue their academic goals. However, the flexibility that makes online learning appealing also presents unique challenges, particularly regarding student engagement and retention [1,2]. High dropout rates in online courses not only affect individual students' academic progress but also undermine the efficacy of educational programs and the reputation of institutions offering these courses. With the rapid expansion of online learning platforms, understanding the factors that influence student retention and dropout rates has become crucial for designing effective interventions aimed at mitigating dropout rates and enhancing student success [1–4].

Traditional models of student retention, such as those used in face-to-face learning environments [5–8], often fail to capture the complex and dynamic patterns observed in online dropout data. These models typically assume linear relationships and do not account for the myriad factors that can influence a student's decision to drop out [9]. Online learning environments introduce variables such as technological issues, varying levels of digital literacy, and the need for self-regulation and motivation, all of which can significantly impact student retention [9]. Consequently, there is a pressing need for the development of more sophisticated statistical tools that can accurately model these complexities and provide coverage for dropout time. While most of the existing studies on modelling dropout focuses of the use of machine learning classification techniques [8,10–13], only few [2] has utilized survival analysis and probability modelling to analyze students dropout time. Thus, this study introduces the Exponentiated Weibull Weibull (EWW) distribution, a five-parameter model specifically designed to better represent dropout times in online learning environments. The EWW distribution extends traditional survival analysis models by incorporating additional parameters that

allow for greater flexibility in modeling the hazard function. The hazard function, which represents the instantaneous risk of dropout at any given time, can take on various shapes, including increasing, decreasing, and bathtub-shaped hazard rates. This flexibility is particularly important for capturing the uncommon patterns of student engagement and disengagement in online courses.

The motivation for modelling student dropout times by the EWW distribution comes from the requirement to include the unusual patterns of online learning engagement and disengagement. Earlier studies have applied different statistical techniques in the analysis of dropout rates, including logistic regression and Cox proportional hazards models [2,14–17]. Moreover, these models are often inadequate to properly predict the non-monotonic hazard rates of dropout that appear in educational data, where the risk of student leaving his/her course is lower at the beginning and then increases thereafter. Using the EWW distribution in this setting allows us to investigate the dropout dynamics in a more fine-grained manner. EWW distribution provides a better fit for the data by introducing parameters which is able to capture both the initial decrease, as well as long-term increase in dropout hazard rates. A closer fit is important in generating appropriate interventions that directly target dropout and retention rate for online learning courses [2].

The EWW distribution is built on the foundational principles of the Weibull distribution, a widely used model in reliability engineering and survival analysis due to its flexibility in modelling various types of data [18]. The Exponentiated Weibull (EW) distribution are known for characterization of both increasing and decreasing hazard rates, making it suitable for diverse applications [18]. The EWW distribution extends this versatility by incorporating additional parameters, allowing it to accommodate more complex hazard functions often observed in real-world data, such as the bathtub-shaped hazard function which maybe observed in dropout studies [2,19,20]. This enhancement is important for capturing the unusual patterns of student engagement and disengagement, providing a more accurate and detailed understanding of dropout dynamics in online learning environments.

Bayesian estimation has gained popularity in recent years due to its ability to incorporate prior information and provide more accurate parameter estimates, especially in small sample sizes or complex models [18,21–26]. Unlike the traditional Maximum Likelihood Estimation (MLE), which relies solely on the observed data, Bayesian methods combine prior distributions with the likelihood of the observed data to generate posterior distributions. This approach can yield more robust estimates, particularly in scenarios with high variability or limited data [27,28].

In this paper, I used both MLE and Bayesian estimation methods to estimate the parameters of the EWW distribution. The comparison of these methods is essential to determine the most effective approach for parameter estimation in this context. Simulation studies and real-life data analysis are conducted to evaluate the accuracy and precision of these estimates.

2. Exponentiated-Weibull Distribution

The exponentiated Weibull (EW) distribution was introduced by [20]. This distribution serves as an extension of the popular Weibull distribution, offering a more flexible model for analyzing lifetime data. Unlike the traditional Weibull distribution, the EW distribution can model both monotone and nonmonotone failure rates, making it particularly useful for practical applications where data often exhibit bathtub-shaped or upside-down bathtub-shaped failure rates. In many real-world scenarios, such as reliability engineering and survival analysis, lifetime data do not always follow a simple increasing or decreasing hazard function. Instead, they might show a complex pattern where the failure rate decreases initially, stabilizes, and then increases, resembling a bathtub shape. Similarly, an upside-down bathtub shape might be observed, where the failure rate initially increases, stabilizes, and then decreases. The EW distribution captures these patterns more accurately than distributions with purely monotone failure rates. The flexibility of the EW distribution is derived from its two shape parameters η and λ and one rate parameter δ. The probability density function (pdf) of the three-parameter exponentiated Weibull distribution is given by:

$$f(t) = \eta\delta\lambda(\delta t)^{\eta-1}\left(1 - \exp[-(\delta t)^{\eta}]\right)^{\lambda-1}\exp[-(\delta t)^{\eta}], \tag{1}$$

and its cumulative distribution function (CDF) is:

$$F(t) = \left(1 - \exp[-(\delta t)^{\eta}]\right)^{\lambda}, \tag{2}$$

where $t > 0$ is within the support of the distribution, and $\eta > 0$, $\delta > 0$ and $\lambda > 0$ are the parameters. If $\lambda = 1$, the exponentiated Weibull distribution simplifies to the standard Weibull distribution, which has the pdf:

$$f(t) = \eta\delta(\delta t)^{\eta-1}\exp[-(\delta t)^{\eta}], \tag{3}$$

and the cumulative distribution function (CDF) of the Weibull distribution is given by

$$F(t) = 1 - \exp\left[-(\delta t)^{\eta}\right]. \tag{4}$$

2.1. Moments and Median Survival Time

According to [29,30] , the *r*th moment of EW for positive integer λ can be defined as:

$$E(X^r) = \lambda\delta^r\Gamma\left(\frac{r}{\eta} + 1\right)A_r(\eta)$$

where

$$A_r(\eta) = 1 + \sum_{j=1}^{\lambda-1}\frac{(-1)^j\binom{\lambda-1}{j}}{(j+1)^{\left(\frac{r}{\eta}+1\right)}}$$

For the mean, with $r = 1$,

$$E(X) = \lambda\delta\Gamma\left(\frac{1}{\eta} + 1\right)A_1(\eta)$$

$$E(X) = \lambda\delta\Gamma\left(\frac{1}{\eta} + 1\right)\left(1 + \sum_{j=1}^{\lambda-1}\frac{(-1)^j\binom{\lambda-1}{j}}{(j+1)^{\left(\frac{1}{\eta}+1\right)}}\right)$$

For the variance, with $r = 2$,

$$E(X^2) = \lambda\delta^2\Gamma\left(\frac{2}{\eta} + 1\right)A_2(\eta)$$

$$E(X^2) = \lambda\delta^2\Gamma\left(\frac{2}{\eta} + 1\right)\left(1 + \sum_{j=1}^{\lambda-1}\frac{(-1)^j\binom{\lambda-1}{j}}{(j+1)^{\left(\frac{2}{\eta}+1\right)}}\right)$$

$$\mathrm{Var}(X) = \lambda\delta^2\Gamma\left(\frac{2}{\eta} + 1\right)\left(1 + \sum_{j=1}^{\lambda-1}\frac{(-1)^j\binom{\lambda-1}{j}}{(j+1)^{\left(\frac{2}{\eta}+1\right)}}\right) - \left[\lambda\delta\Gamma\left(\frac{1}{\eta} + 1\right)\left(1 + \sum_{j=1}^{\lambda-1}\frac{(-1)^j\binom{\lambda-1}{j}}{(j+1)^{\left(\frac{1}{\eta}+1\right)}}\right)\right]^2$$

Consequently, according to [31] the median survival time is derived as:

$$M(t) = \frac{1}{\delta}\left[-\log\left(1 - 0.5^{\frac{1}{\lambda}}\right)\right]^{\frac{1}{\eta}}. \tag{5}$$

2.2. Survival, Hazard, and Cumulative Hazard Functions

The survival function, hazard function, and cumulative hazard function of the exponentiated Weibull distribution are as follows:

1. Survival Function

$$S(t) = 1 - (1 - \exp[-(\delta t)^\eta])^\lambda,$$

2. Hazard Function

$$h(t) = \frac{\eta \delta \lambda (\delta t)^{\eta-1} (1 - \exp[-(\delta t)^\eta])^{\lambda-1} \exp[-(\delta t)^\eta]}{1 - (1 - \exp[-(\delta t)^\eta])^\lambda},$$

3. Cumulative Hazard Function

$$H(t) = -\log\left\{1 - (1 - \exp[-(\delta t)^\eta])^\lambda\right\}.$$

4. Hazard Function Behavior

The behaviour of the hazard function can be characterized as follows:

- It is monotone increasing if $\eta \geq 1$ and $\eta\lambda \geq 1$.
- It is monotone decreasing if $\eta \leq 1$ and $\eta\lambda \leq 1$.
- It is unimodal if $\eta < 1$ and $\eta\lambda > 1$.
- It is bathtub-shaped if $\eta > 1$ and $\eta\lambda < 1$.

3. Five-Parameter Exponentiated Weibull Weibull Distribution

The cumulative distribution function (CDF) of the exponentiated Weibull-generated family [32] for a random variable $t > 0$ is given by

$$F(t) = \left\{1 - \exp\left[-\delta\left(\frac{G(t)}{1 - G(t)}\right)^\eta\right]\right\}^\lambda ; \ \delta, \ \eta, \ \lambda > 0, \tag{6}$$

where η and λ are the two shape parameters, and δ is the rate parameter. This CDF provides a broader family of continuous distributions. The corresponding probability density function (PDF) is given by

$$f(t) = \frac{\lambda \delta \eta (G(t))^{\eta-1} g(t)}{[1 - (G(t))]^{\eta+1}} \exp\left[-\delta\left(\frac{G(t)}{1 - G(t)}\right)^\eta\right]\left\{1 - \exp\left[-\delta\left(\frac{G(t)}{1 - G(t)}\right)^\eta\right]\right\}^{\lambda-1} ; \ \delta, \ \eta, \ \lambda > 0. \tag{7}$$

In this study, we propose a novel five-parameter model that serves as a competitive extension to the Weibull distribution by utilizing the EW-G distribution. By replacing $G(t)$ in equation (6) with the cumulative distribution function of the Weibull distribution from equation (4), we obtain the following result.

$$F(t) = \left\{1 - \exp\left[-\delta\left(\frac{1 - \exp[-(\nu t)^\gamma]}{\exp[-(\nu t)^\gamma]}\right)^\eta\right]\right\}^\lambda$$

Simplifying the expression inside the exponential:

$$\frac{1 - \exp[-(\nu t)^\gamma]}{\exp[-(\nu t)^\gamma]} = \exp[(\nu t)^\gamma] - 1$$

So the function becomes:

$$F(t) = \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^\lambda$$

Therefore, the cumulative distribution function (CDF) of the five-parameter Exponentiated-Weibull Weibull (EWW) distribution for a random variable $t > 0$ is given by:

$$F(t) = \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^\lambda; \ \delta, \ \eta, \ \lambda, \ \nu, \ \gamma > 0, \tag{8}$$

where η, δ and γ are the three shape parameters, and δ and ν are the rate parameters.

Correspondingly, substituting equations (3) and (4) in (7), we have the pdf of the five-parameter exponentiated-Weibull Weibull distribution given by

$$
\begin{aligned}
f(t) = {} & \lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\left\{1 - \exp\left[-(\nu t)^\gamma\right]\right\}^{\eta-1}\exp[\eta(\nu t)^\gamma] \\
& \times \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right] \\
& \times \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^{\lambda-1}; \ \delta, \ \eta, \ \lambda, \ \nu, \ \gamma > 0.
\end{aligned}
\tag{9}
$$

The survival, hazard and cumulative hazard functions follow from the cumulative density function as follows respectively:

$$
\begin{aligned}
S(t) &= 1 - F(t) \\
&= 1 - \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^\lambda,
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
h(t) &= \frac{f(t)}{1 - F(t)} \\
&= \frac{f(t)}{1 - \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^\lambda},
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
H(t) &= -\log(1 - F(t)) \\
&= -\log\left(1 - \left\{1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta\right]\right\}^\lambda\right).
\end{aligned}
\tag{12}
$$

In addition, the median survival time is given by:

$$M(t) = \frac{1}{\nu}\left(\ln\left\{1 + \left[-\frac{\ln\left(1 - 0.5^{\frac{1}{\lambda}}\right)}{\delta}\right]^{\frac{1}{\eta}}\right\}\right)^{\frac{1}{\gamma}}. \tag{13}$$

Note that the derivations in equations (8) to (13) are different from the one presented by [32] in terms of parameterization of the Weibull base distribution. While the approach of [32] is applicable in medicine, our own parameterization is applicable for modelling failure times [33,34] which is relevant for the specific situation addressed in this paper.
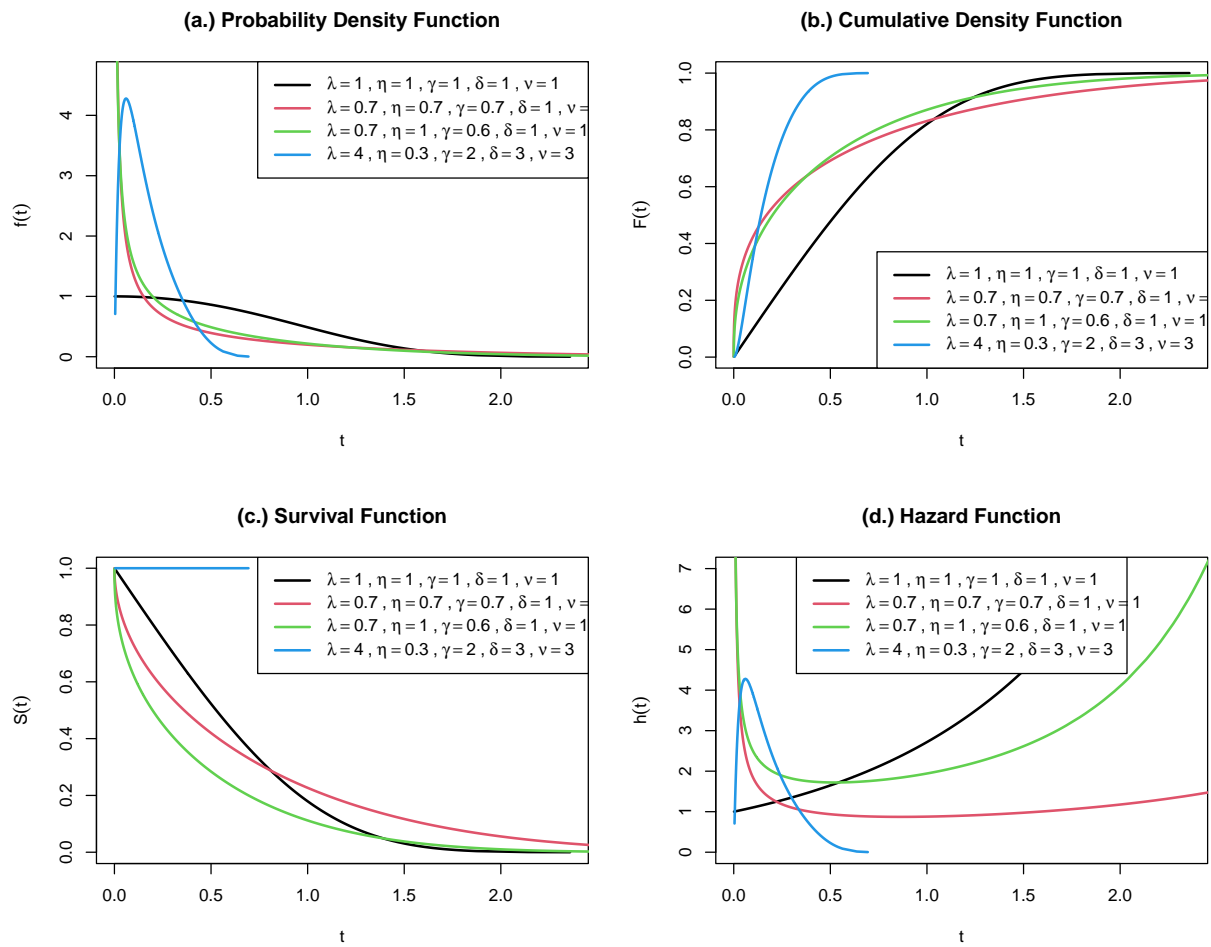
(a.) Probability Density Function

(b.) Cumulative Density Function

(c.) Survival Function

(d.) Hazard Function

Figure 1. Plot of pdf, cdf, survival and hazard functions for the five-parameter exponentiated-Weibull Weibull Distribution for increasing (black), decreasing (red), bathtub (green) and unimodal (blue) hazard types.

Figure 1 shows that when $\lambda, \eta, \gamma, \delta, \nu$ are all equal to 1, we have an increasing hazard and when λ, η, γ are less than 1, we have a decreasing hazard. Similarly, when λ, γ are less than 1 but γ at least 1, we have a bathtub hazard and η is less than 1 but λ and γ are greater than 1, we have a unimodal hazard.

4. Modelling Student Dropout Time

Suppose we have a dataset $D = \{t_i, Y_i\}$ that corresponds to dropout times t_i and outcomes (withdrawn: W and completed: C) Y_i for a randomly selected student i from an online class module presentation. Thus, the joint density function for simultaneously modelling dropout and completion times is given by:

$$f(t_W, t_C) = [f(W)]^{\alpha} \times [S(C)]^{1-\alpha} \tag{14}$$

where $\alpha = \{0, 1\}$ with 1 denoting withdrawn and 0 denoting complete, $f(W)$ and $S(C)$ are the EWW probability density function and survival functions for the times of withdrawn and completed students, respectively. Substituting (9) and (10) in (12) we have:

$$
\begin{aligned}
f(t_W, t_C) = &\Big(\lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\{1-\exp[-(\nu t)^{\gamma}]\}^{\eta-1}\exp[\eta(\nu t)^{\gamma}] \\
&\times \exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big] \\
&\times \{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda-1}\Big)^{\alpha} \\
&\times \Big(1-\{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda}\Big)^{1-\alpha}.
\end{aligned}
\tag{15}
$$

5. Parameter Estimation

5.1. Maximum Likelihood Estimation

The maximum likelihood estimates of the unknown parameter set $\Omega = \{\lambda, \eta, \gamma, \delta, \nu\}$ can be obtained from the random sample of students withdraw and completion times $t_i, i \in \{1, \ldots, n\}$. The likelihood and log-likelihood functions for the joint density in (15) are given as:

$$
\begin{aligned}
L(t_1, \ldots, t_n; \Omega) &= \prod_{i=1}^{n} f(t_W, t_C; \Omega) \\
&= \prod_{i=1}^{n}\Big(\lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\{1-\exp[-(\nu t)^{\gamma}]\}^{\eta-1}\exp[\eta(\nu t)^{\gamma}] \\
&\quad\times \exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big] \\
&\quad\times\{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda-1}\Big)^{\alpha} \\
&\quad\times\Big(1-\{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda}\Big)^{1-\alpha}
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
\log(L(t_1, \ldots, t_n; \Omega)) &= \sum_{i=1}^{n} \log(f(t_W, t_C; \Omega)) \\
&= \sum_{i=1}^{n} \log\Big[\big(\lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\{1-\exp[-(\nu t)^{\gamma}]\}^{\eta-1}\exp[\eta(\nu t)^{\gamma}] \\
&\quad\times \exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big] \\
&\quad\times\big\{1-\exp[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}]\}^{\lambda-1}\big)^{\alpha} \\
&\quad\times\Big(1-\{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda}\Big)^{1-\alpha}\Big]
\end{aligned}
\tag{17}
$$

The MLE estimates of the parameters correspond to the solutions of the partial derivatives of the log-likelihood function concerning λ, η, γ, δ, and ν. Given the $\log(L(t_1, \ldots, t_n; \Omega))$, we can define $f(t; \Omega)$ as:

$$
\begin{aligned}
f(t; \Omega) = &\Big[\big(\lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\{1-\exp[-(\nu t)^{\gamma}]\}^{\eta-1}\exp[\eta(\nu t)^{\gamma}] \\
&\times \exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big] \\
&\times\big\{1-\exp[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}]\}^{\lambda-1}\big)^{\alpha} \\
&\times\Big(1-\{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda}\Big)^{1-\alpha}\Big].
\end{aligned}
$$

Suppose we define $k(t)$ and $m(t)$ for simplification:

$$
\begin{aligned}
k(t) = &\lambda\delta\eta\gamma\nu(\nu t)^{\gamma-1}\{1-\exp[-(\nu t)^{\gamma}]\}^{\eta-1}\exp[\eta(\nu t)^{\gamma}] \\
&\times \exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big] \\
&\times \{1-\exp\big[-\delta(\exp[(\nu t)^{\gamma}]-1)^{\eta}\big]\}^{\lambda-1}
\end{aligned}
$$

$$m(t) = 1 - \left\{ 1 - \exp\left[-\delta(\exp[(\nu t)^\gamma] - 1)^\eta \right] \right\}^\lambda$$

Then, the log-likelihood function becomes:

$$\log(L(t_1, \ldots, t_n; \Omega)) = \sum_{i=1}^{n} \log\left(k(t_i)^\alpha \cdot m(t_i)^{1-\alpha} \right)$$

$$= \sum_{i=1}^{n} (\alpha \log k(t_i) + (1 - \alpha) \log m(t_i))$$

The corresponding partial derivatives with respect to the parameters are as follows:

1. With respect to λ:

$$\frac{\partial \log(L)}{\partial \lambda} = \sum_{i=1}^{n} \left(\alpha \frac{1}{k(t_i)} \frac{\partial k(t_i)}{\partial \lambda} + (1 - \alpha) \frac{1}{m(t_i)} \frac{\partial m(t_i)}{\partial \lambda} \right)$$

2. With respect to η:

$$\frac{\partial \log(L)}{\partial \eta} = \sum_{i=1}^{n} \left(\alpha \frac{1}{k(t_i)} \frac{\partial k(t_i)}{\partial \eta} + (1 - \alpha) \frac{1}{m(t_i)} \frac{\partial m(t_i)}{\partial \eta} \right)$$

3. With respect to γ:

$$\frac{\partial \log(L)}{\partial \gamma} = \sum_{i=1}^{n} \left(\alpha \frac{1}{k(t_i)} \frac{\partial k(t_i)}{\partial \gamma} + (1 - \alpha) \frac{1}{m(t_i)} \frac{\partial m(t_i)}{\partial \gamma} \right)$$

4. With respect to δ:

$$\frac{\partial \log(L)}{\partial \delta} = \sum_{i=1}^{n} \left(\alpha \frac{1}{k(t_i)} \frac{\partial k(t_i)}{\partial \delta} + (1 - \alpha) \frac{1}{m(t_i)} \frac{\partial m(t_i)}{\partial \delta} \right)$$

5. With respect to ν:

$$\frac{\partial \log(L)}{\partial \nu} = \sum_{i=1}^{n} \left(\alpha \frac{1}{k(t_i)} \frac{\partial k(t_i)}{\partial \nu} + (1 - \alpha) \frac{1}{m(t_i)} \frac{\partial m(t_i)}{\partial \nu} \right)$$

Combining these partial derivatives, we get the final expressions for the partial derivatives of the log-likelihood function with respect to λ, η, γ, δ, and ν. Solving these partial derivatives manually is very hard and hence in practice the Newton-Raphson procedure [34] is often used.

5.2. Bayesian Estimation

In this section, we discuss the Bayesian estimation of the parameters of the five-parameter exponentiated Weibull Weibull (EWW) distribution. We use the uninformative Uniform (a, b) prior since there is no prior information or elicitation may be difficult [18]. Let the parameters of the EWW distribution be λ, η, γ, δ, and ν. We suggest five independent Uniform (a, b) distributions for these parameters. The joint density function for the prior of these five parameters can be defined as:

$$f(\lambda, \eta, \gamma, \delta, \nu \mid a_1, b_1; a_2, b_2; a_3, b_3; a_4, b_4; a_5, b_5) = \frac{1}{b_1 - a_1} \times \frac{1}{b_2 - a_2} \times \frac{1}{b_3 - a_3} \\ \times \frac{1}{b_4 - a_4} \times \frac{1}{b_5 - a_5} \tag{18}$$

where $a_1, b_1; a_2, b_2; a_3, b_3; a_4, b_4; a_5, b_5$ are the prior hyperparameters for the parameters $\lambda, \eta, \gamma, \delta, \nu$ respectively. The posterior distribution of the parameters $\lambda, \eta, \gamma, \delta, \nu$ for the EWW model can be defined as the product of the likelihood $L(t_1, \ldots, t_n; \Omega)$ and the prior density which is:

$$f(\lambda, \eta, \gamma, \delta, \nu \mid t, a_1, b_1; a_2, b_2; a_3, b_3; a_4, b_4; a_5, b_5) = L(t_1, \ldots, t_n; \mid \lambda, \eta, \gamma, \delta, \nu) \times f(\lambda, \eta, \gamma, \delta, \nu) \quad (19)$$

where the likelihood function $L(t_1, \ldots, t_n; \Omega)$ is (16).

The posterior distribution in (19) does not have a closed form as it is an approximate distribution since the marginal distribution that ensures it scales to one has been dropped. One of the ways of sampling from this distribution is by using the Metropolis-Hastings algorithm [24].

5.2.1. Metropolis-Hastings Algorithm (MHA) for EWW

The Metropolis-Hastings algorithm [24] is a hybrid random walk that employs the accept/reject algorithm to arrive at the desired distribution. The desired Bayesian inference can be summarized as follows:

Algorithm 1: Metropolis-Hastings Algorithm

---

1. Define a starting solution $\Omega_i^0 = \{\lambda, \eta, \gamma, \delta, \nu\}^0$, for which $P(\Omega_i \mid a_p, b_p; p \in \{1:5\}) > 0$.

2. For iteration $q = \{1, 2, 3, \ldots, Q\}$ do:

    (a) Sample $\hat{\Omega}_q$ from the Uniform distribution; $B(\hat{\Omega}_q \mid a, b) \sim U(a_q, b_q)$.

    (b) Calculate the acceptance probability:

    $$\phi(\Omega_i^q, \hat{\Omega}_q) = \min\left(\frac{P(\hat{\Omega}_q \mid a_1, b_1; a_2, b_2; a_3, b_3; a_4, b_4; a_5, b_5)}{P(\Omega_i^q \mid a_1, b_1; a_2, b_2; a_3, b_3; a_4, b_4; a_5, b_5)}, 1\right)$$

    (c) Generate a uniform random variable $U \sim U(0, 1)$.

    (d) Then:
    $$\Omega_i^{q+1} = \begin{cases} \hat{\Omega}_q & \text{if } U \leq \phi(\Omega_i^q, \hat{\Omega}_q) \\ \Omega_i^q & \text{if } U > \phi(\Omega_i^q, \hat{\Omega}_q) \end{cases}$$

3. End for q

---

6. Simulation study and results

To illustrate the EWW dropout model using both the maximum likelihood estimator and Bayesian estimator, we simulated random samples that corresponds to the four hazard types presented in figure 1. Specifically, the following simulation parameters were used for the various hazard types.

1. Monotone increasing: $\lambda = 1, \eta = 1, \gamma = 1, \delta = 1, \nu = 1$.
2. Monotone decreasing: $\lambda = .7, \eta = .7, \gamma = .7, \delta = 1, \nu = 1$.
3. Unimodal: $\lambda = 4, \eta = .3, \gamma = 2, \delta = 3, \nu = 3$.
4. Bathtub-shaped: $\lambda = .7, \eta = 1, \gamma = .6, \delta = 1, \nu = 1$.

The dropout (t_i) time was generated using the following equation:

$$t_i = \frac{1}{\nu}\left(\ln\left\{1 + \left[-\frac{\ln\left(1 - U^{\frac{1}{\lambda}}\right)}{\delta}\right]^{\frac{1}{\eta}}\right\}\right)^{\frac{1}{\gamma}}, \quad (20)$$

where U is the standardized uniform random variable with parameters $0, 1$. To achieve censoring, a random censoring scheme was adopted. In this scheme, a random variable Z that follows a Bernoulli

distribution with parameter π, denoted as Bernoulli(π), is defined. If $Z_i = 1$, then t_i corresponds to the dropout time of student i. Otherwise, t_i corresponds to the completion time of student i. To study the effect of censoring on the estimators, we set $\pi = 0, 0.1, 0.2, 0.3, 0.4, 0.5$ which corresponds to $10\% - 50\%$ censoring percentages. In addition, the estimators were evaluated using bias and root mean square error. Bias and root mean square error for each parameter set are given below:

$$Bias(\theta) = \hat{\theta} - \theta, \tag{21}$$

$$Bias(\Omega) = \frac{\sum_{i=1}^{5}(\hat{\Omega} - \Omega)}{5}, \tag{22}$$

$$RMSE(\theta) = \sqrt{\left(\hat{\theta} - \theta\right)^2}, \tag{23}$$

$$RMSE(\Omega) = \frac{\sqrt{\sum_{i=1}^{5}\left(\hat{\Omega} - \Omega\right)^2}}{5}. \tag{24}$$

The sample size n was fixed at 200 throughout the experiment. To achieve stability, each simulation run was replicated 1000 times [35]. For the Bayesian estimation, 20,000 posterior samples were generated with a thinning value set to 10 (i.e., every 10th estimate was saved). The initial burn-in period, during which samples were discarded, was set at 2000.

In Table 1, we present the parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods. The table summarizes these statistics for five different parameters: λ, η, γ, δ, and ν, evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for scenario 1: increasing hazard type. The results show that Bayesian estimates consistently outperform Maximum Likelihood Estimates (MLE) across all parameters. For λ, Bayesian estimates are close to 1 the true value with low standard deviations and root mean squared errors (RMSE), whereas MLEs exhibit higher variability and larger biases, particularly at the 10% quantile. For η, Bayesian estimates also display consistency and low biases, while MLEs vary significantly with large biases and higher RMSE values.

Similarly Table 1 shows that the Bayesian estimates for γ are accurate with small positive biases and low RMSE values, whereas MLEs show greater variability and higher biases, notably at the 20% quantile. For δ, Bayesian estimates are close to the true value with low biases and RMSE, while MLEs are more variable with larger biases and higher RMSE values. Lastly, Bayesian estimates for ν are accurate with small biases and low RMSE, in contrast to MLEs which exhibit higher variability, larger biases, and higher RMSE values. Overall, Bayesian estimation demonstrates higher accuracy and precision compared to MLE for all parameters, indicating its robustness in this context.

Table 2 presents the parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for a decreasing hazard type. The results show that Bayesian estimates consistently exhibit lower biases and RMSE values compared to MLEs across all parameters and censoring levels. Bayesian estimates for λ, η, γ, δ, and ν are more accurate and precise, with smaller SDs and biases. In contrast, MLEs show higher variability, larger biases, and greater RMSE values, indicating less reliable estimates. For instance, the bias for λ in Bayesian estimates remains small and close to zero, while MLEs show a significant bias, especially at the 10% censoring level. Similarly, Bayesian estimates for η and ν demonstrate consistent accuracy, whereas MLEs vary widely with substantial biases and higher RMSE values.

Table 1. Parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods were evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for an increasing hazard type.

Parameter	π	Bayesian				MLE			
		Estimate	SD	Bias	RMSE	Estimate	SE	Bias	RMSE
λ		0.991	0.014	-0.009	0.017	1.910	4.459	0.910	4.551
η		0.946	0.025	-0.055	0.060	2.001	1.871	1.001	2.122
γ	10%	1.033	0.030	0.033	0.045	1.741	0.570	0.741	0.935
δ		1.013	0.010	0.013	0.017	0.450	0.920	-0.550	1.072
ν		1.030	0.022	0.030	0.037	0.398	0.347	-0.602	0.695
Ω				0.003	0.035			0.300	1.875
λ		0.977	0.016	-0.023	0.028	1.187	0.645	0.187	0.671
η		0.959	0.025	-0.041	0.048	0.434	0.348	-0.566	0.664
γ	20%	1.043	0.033	0.043	0.055	2.284	3.407	1.284	3.641
δ		0.996	0.011	-0.004	0.011	1.385	1.658	0.385	1.702
ν		1.047	0.026	0.047	0.054	1.444	1.067	0.444	1.155
Ω				0.004	0.039			0.347	1.567
λ		0.961	0.021	-0.039	0.044	0.457	0.861	-0.543	1.018
η		0.947	0.026	-0.053	0.059	1.335	5.281	0.335	5.291
γ	30%	1.049	0.031	0.049	0.058	1.079	3.170	0.079	3.171
δ		0.980	0.016	-0.021	0.026	1.204	3.943	0.204	3.948
ν		1.044	0.026	0.044	0.051	0.753	1.150	-0.247	1.176
Ω				-0.004	0.048			-0.034	2.921
λ		0.952	0.024	-0.048	0.054	0.831	1.252	-0.169	1.263
η		0.934	0.028	-0.066	0.072	1.210	1.368	0.210	1.384
γ	40%	1.044	0.030	0.044	0.053	0.754	1.096	-0.246	1.123
δ		0.958	0.020	-0.042	0.046	0.733	1.172	-0.267	1.202
ν		1.040	0.023	0.040	0.046	1.190	1.216	0.190	1.231
Ω				-0.014	0.054			-0.056	1.241
λ		0.921	0.030	-0.079	0.085	0.457	1.036	0.636	1.216
η		0.949	0.024	-0.051	0.056	1.335	0.348	-0.385	0.519
γ	50%	1.068	0.034	0.068	0.077	1.079	0.131	1.076	1.084
δ		0.926	0.029	-0.074	0.079	1.204	0.489	-0.297	0.572
ν		1.052	0.028	0.052	0.059	0.753	0.736	0.466	0.871
Ω				-0.017	0.071			0.299	0.853

Table 3 presents parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods, evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for a unimodal hazard type. For parameter λ, Bayesian estimates are consistently close to 4 with low SD and RMSE, whereas MLEs show significant variability and larger biases, especially at the 10% censoring level. Bayesian estimates for η remain close to the true value of 0.3, displaying small biases and low RMSEs, while MLEs vary considerably with large biases and higher RMSEs. For γ, Bayesian estimates are accurate with low biases and RMSEs, but MLEs exhibit higher variability and significant biases. In the case of δ, Bayesian estimates are near the true value of 3 with low biases and RMSEs, while MLEs show greater variability and higher biases. Lastly, Bayesian estimates for ν demonstrate accuracy with small biases and low RMSEs, in contrast to MLEs which display high variability, larger biases, and elevated RMSEs.

Table 2. Parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods were evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for decreasing hazard type.

Parameter	π	Bayesian				MLE			
		Estimate	SD	Bias	RMSE	Estimate	SE	Bias	RMSE
λ		0.703	0.014	0.003	0.015	0.442	0.436	-0.258	0.506
η		0.591	0.036	-0.109	0.115	0.991	1.085	0.291	1.124
γ	10%	0.704	0.024	0.004	0.024	0.812	0.826	0.112	0.833
δ		1.014	0.011	0.014	0.018	1.623	3.481	0.623	3.537
ν		0.982	0.013	-0.019	0.023	0.472	0.257	-0.528	0.588
Ω				-0.021	0.039			0.048	1.318
λ		0.722	0.017	0.022	0.027	0.945	0.548	0.245	0.601
η		0.607	0.031	-0.093	0.098	0.373	0.246	-0.328	0.410
γ	20%	0.685	0.020	-0.015	0.025	1.926	1.649	1.226	2.055
δ		1.027	0.013	0.027	0.030	3.613	3.452	2.613	4.330
ν		0.970	0.012	-0.030	0.032	0.627	0.193	-0.374	0.420
Ω				-0.018	0.042			0.677	1.563
λ		0.660	0.020	-0.040	0.044	0.442	1.913	0.528	1.985
η		0.606	0.038	-0.094	0.101	0.991	0.813	-0.319	0.873
γ	30%	0.752	0.034	0.052	0.062	0.812	7.484	1.799	7.697
δ		0.983	0.013	-0.017	0.021	1.623	1.883	0.192	1.893
ν		0.999	0.017	-0.001	0.017	0.472	0.394	-0.286	0.487
Ω				-0.020	0.049			0.383	2.587
λ		0.740	0.019	0.040	0.044	0.304	0.645	-0.396	0.757
η		0.595	0.029	-0.105	0.109	2.514	3.627	1.814	4.055
γ	40%	0.664	0.019	-0.036	0.041	1.167	0.752	0.467	0.885
δ		1.021	0.013	0.021	0.024	2.512	7.925	1.512	8.068
ν		0.954	0.013	-0.046	0.047	0.126	0.141	-0.874	0.885
Ω				-0.025	0.053			0.505	2.930
λ		0.687	0.016	-0.013	0.021	0.444	1.294	-0.256	1.320
η		0.579	0.041	-0.121	0.128	0.729	3.649	0.029	3.650
γ	50%	0.716	0.027	0.016	0.031	0.827	4.394	0.127	4.395
δ		1.001	0.009	0.001	0.010	1.609	10.079	0.609	10.098
ν		0.980	0.014	-0.021	0.025	0.608	0.297	-0.392	0.492
Ω				-0.027	0.043			0.023	3.991

Table 3. Parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods were evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for unimodal hazard type.

Parameter	π	Bayesian				MLE			
		Estimate	SD	Bias	RMSE	Estimate	SE	Bias	RMSE
λ		3.956	0.019	-0.044	0.048	16.078	0.354	12.078	12.083
η		0.315	0.015	0.015	0.021	1.066	1.844	0.766	1.997
γ	10%	2.038	0.028	0.038	0.048	5.577	3.687	3.577	5.137
δ		2.930	0.035	-0.070	0.078	0.226	0.350	-2.774	2.796
ν		3.018	0.018	0.018	0.026	0.573	1.167	-2.427	2.693
Ω				-0.008	0.044			2.244	4.941
λ		3.954	0.021	-0.047	0.051	2.479	2.450	-1.521	2.883
η		0.329	0.015	0.029	0.032	0.412	0.275	0.112	0.297
γ	20%	2.036	0.028	0.036	0.045	2.901	5.622	0.901	5.694
δ		2.928	0.035	-0.072	0.080	2.869	4.027	-0.131	4.029
ν		3.028	0.022	0.028	0.036	1.809	1.962	-1.191	2.296
Ω				-0.005	0.049			-0.366	3.040
λ		3.952	0.022	-0.048	0.053	3.264	1.574	-0.736	1.738
η		0.332	0.015	0.032	0.036	0.304	0.339	0.004	0.339
γ	30%	2.038	0.029	0.038	0.048	1.640	0.972	-0.360	1.037
δ		2.920	0.038	-0.080	0.088	1.944	0.787	-1.056	1.317
ν		3.023	0.022	0.023	0.032	3.748	5.723	0.748	5.772
Ω				-0.007	0.051			-0.280	2.041
λ		3.951	0.022	-0.049	0.054	2.462	0.854	-1.538	1.759
η		0.329	0.015	0.029	0.033	0.266	0.144	-0.035	0.148
γ	40%	2.038	0.030	0.038	0.049	1.737	1.341	-0.263	1.366
δ		2.916	0.039	-0.085	0.093	2.204	0.752	-0.796	1.096
ν		3.024	0.022	0.024	0.033	3.832	2.858	0.832	2.976
Ω				-0.008	0.052			-0.360	1.469
λ		3.938	0.023	-0.062	0.066	3.264	0.583	-0.722	0.928
η		0.352	0.016	0.052	0.054	0.304	0.047	-0.210	0.215
γ	50%	2.042	0.032	0.042	0.053	1.640	4.560	4.567	6.454
δ		2.910	0.044	-0.090	0.100	1.944	0.289	-0.731	0.786
ν		3.027	0.024	0.027	0.036	3.748	3.781	3.643	5.250
Ω				-0.006	0.062			1.309	2.727

The results presented in Table 4 indicate that Bayesian estimates consistently outperform Maximum Likelihood Estimates (MLE) across various parameters and censoring levels for the bathtub hazard type. Bayesian estimates for all parameters (λ, η, γ, δ, ν) show low standard deviations (SD) and root mean squared errors (RMSE), with biases generally close to zero. In contrast, MLEs exhibit higher variability, larger biases, and higher RMSE values across the board. For example, Bayesian estimates of λ at the 10% censoring level have a low bias of 0.024 and an RMSE of 0.029, whereas MLEs show a high bias of 0.152 and an RMSE of 0.729. Similarly, Bayesian estimates of δ are close to the true value with low biases and RMSEs across all censoring levels, while MLEs are more variable with larger biases and higher RMSE values.

Overall across the four hazard types and censoring levels, the Bayesian estimates demonstrate higher accuracy and precision compared to the MLEs for all parameters, as indicated by lower biases and RMSE values. The standard deviations for Bayesian estimates are also consistently smaller than the standard errors for MLEs, further highlighting the robustness of the Bayesian approach in this context. This simulation underscores the advantages of Bayesian estimation in providing more reliable parameter estimates for the Exponentiated Weibull Weibull (EWW) distribution.

Table 4. Parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE), biases, and root mean squared errors (RMSE) for both Bayesian and MLE methods were evaluated at different censoring levels (π) of 10%, 20%, 30%, 40%, and 50% for bathtub hazard type.

Parameter	π	Bayesian				MLE			
		Estimate	SD	Bias	RMSE	Estimate	SE	Bias	RMSE
λ		0.724	0.016	0.024	0.029	0.852	0.713	0.152	0.729
η		0.892	0.036	-0.108	0.114	2.094	1.288	1.094	1.690
γ	10%	0.556	0.018	-0.044	0.048	1.370	0.522	0.770	0.930
δ		1.010	0.011	0.010	0.015	0.760	0.941	-0.241	0.972
ν		0.941	0.015	-0.060	0.061	0.184	0.103	-0.816	0.823
Ω				-0.035	0.053			0.192	1.029
λ		0.710	0.016	0.010	0.019	0.831	0.574	0.131	0.589
η		0.897	0.037	-0.103	0.110	0.873	0.647	-0.127	0.659
γ	20%	0.576	0.020	-0.024	0.031	0.818	0.598	0.218	0.637
δ		1.011	0.011	0.011	0.016	1.166	1.339	0.166	1.349
ν		0.969	0.012	-0.032	0.034	0.696	0.330	-0.304	0.448
Ω				-0.027	0.042			0.017	0.736
λ		0.685	0.015	-0.015	0.021	0.344	0.684	-0.356	0.771
η		0.889	0.043	-0.111	0.119	1.381	7.424	0.381	7.434
γ	30%	0.597	0.022	-0.003	0.022	0.821	3.444	0.221	3.451
δ		0.997	0.012	-0.003	0.012	1.878	16.820	0.878	16.843
ν		0.968	0.012	-0.032	0.034	0.403	0.427	-0.598	0.734
Ω				-0.033	0.042			0.105	5.847
λ		0.680	0.017	-0.020	0.026	0.457	0.515	-0.243	0.570
η		0.885	0.044	-0.115	0.123	1.148	2.762	0.148	2.766
γ	40%	0.606	0.023	0.006	0.024	0.662	1.520	0.062	1.521
δ		0.986	0.014	-0.014	0.020	1.085	3.978	0.085	3.979
ν		0.974	0.014	-0.026	0.030	0.598	0.423	-0.402	0.584
Ω				-0.034	0.045			-0.070	1.884
λ		0.653	0.021	-0.047	0.052	0.344	1.623	1.187	2.011
η		0.901	0.038	-0.099	0.106	1.381	0.110	-0.384	0.400
γ	50%	0.651	0.031	0.051	0.060	0.821	0.978	0.219	1.002
δ		0.969	0.017	-0.031	0.035	1.878	0.355	-0.631	0.724
ν		0.997	0.018	-0.003	0.018	0.403	1.272	0.413	1.337
Ω				-0.026	0.054			0.161	1.095

7. Application to Modelling Time to Students' Dropout from an Online Class Presentation

The dataset used in this study was extracted from the student behaviour and performance dataset for 23,344 students in the United Kingdom, accessible at [36]. This dataset includes various data points such as assessment results and extensive logs of student interactions with the Virtual Learning Environment (VLE). The final result of each student is categorized as Distinction, Pass, Withdrawn, or Fail, along with the submission date measured in days since the start of the module presentation. For the analysis in this study, a strategic sample of 500 students was selected to reflect the actual distribution of the final results. The goodness of fit for the proposed EWW distribution, alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) distributions [32], was assessed using the log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). These three competing distributions were selected because they are special cases of the EWW distribution. The WW can be derived from the EWW if we set $\lambda = 1$. The EWW reduces to the WE if $\lambda = 1$ and $\gamma = 1$. Similarly, the EWW reduces to the EW if $\lambda = 1$ and $\eta = 1$.
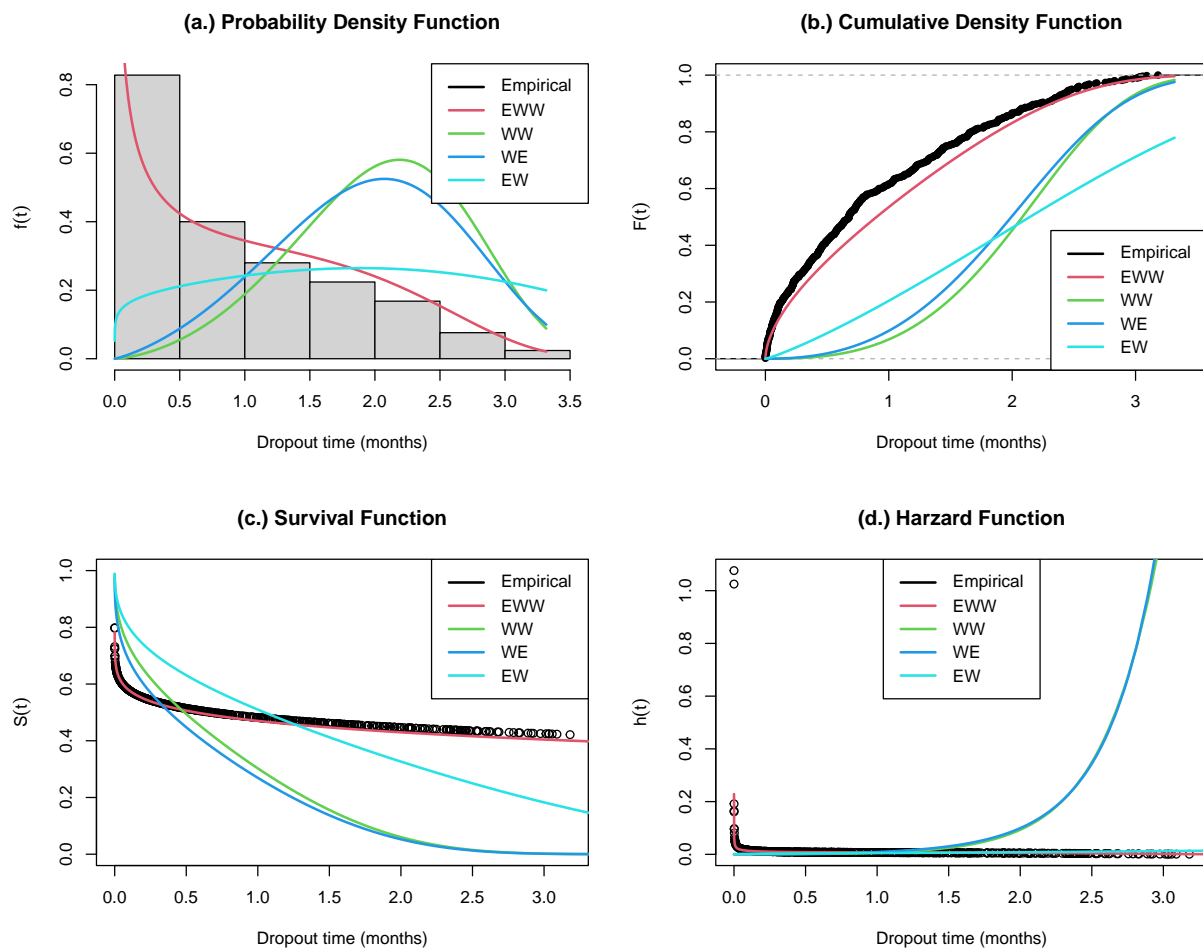
Figure 2. Plot of cdf, pdf, for the five-parameter exponentiated-Weibull Weibull (EWW) alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) distributions for the student dropout dataset based on the Bayesian parameter estimates.

Figure 2 shows the plots of cdf, pdf, for the five-parameter exponentiated-Weibull Weibull (EWW) alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) distributions for the student dropout dataset based on the Bayesian parameter estimates. The plot shows that the most appropriate distribution for the student dropout dataset is EWW.

The results presented in Table 5 highlight the superior performance of Bayesian estimates over Maximum Likelihood Estimates (MLE) for various parameters for the student dropout dataset. The Bayesian estimates for λ, η, γ, δ, and ν exhibit lower standard deviations (SD), indicating higher precision. For instance, the Bayesian estimate of λ is 0.233 with an SD of 0.008, compared to the MLE estimate of 0.536 with a standard error (SE) of 0.137. Similarly, Bayesian estimates for η and γ are close to their true values with minimal SDs (0.002), while MLE estimates show greater variability and higher SEs (0.149 and 0.154, respectively). For the parameter δ, the Bayesian method yields an estimate of 0.635 with an SD of 0.005, significantly more precise than the MLE estimate of 0.759 with an SE of 0.129. Lastly, the Bayesian estimate for ν is 0.358 with an SD of 0.003, whereas the MLE estimate is 1.190 with a considerably higher SE of 0.196. Overall, these findings underscore the robustness and precision of Bayesian estimation in parameter estimation compared to the MLE method for the student dropout dataset.

Table 5. Parameter estimates, standard deviations (SD) for Bayesian estimates, standard errors (SE) for Maximum Likelihood Estimates (MLE) for the student dropout dataset.

Parameter	Bayesian		MLE Estimate	
	Estimate	SD	Estimate	SE
λ	0.233	0.008	0.536	0.137
η	2.149	0.002	0.704	0.149
γ	1.149	0.002	0.720	0.154
δ	0.635	0.005	0.759	0.129
ν	0.358	0.003	1.190	0.196

Table 6. Goodness of fit for for the five-parameter exponentiated-Weibull Weibull (EWW) alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) distributions for the student dropout dataset.

Model	LogLikelihood	AIC	BIC
EWW	-750.75	1511.5	1532.57
WW	-1799.22	3606.44	3623.3
WE	-1859.54	3725.09	3737.73
EW	-1029.02	2064.04	2076.68

The goodness-of-fit results presented in Table 6 for the five-parameter Exponentiated Weibull Weibull (EWW) model, alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) models for the student dropout dataset, indicate that the EWW model provides the best fit. This is evidenced by the EWW model having the highest LogLikelihood (-750.75) and the lowest values for the Akaike Information Criterion (AIC = 1511.5) and the Bayesian Information Criterion (BIC = 1532.57). In contrast, the WW, WE, and EW models exhibit significantly lower LogLikelihood values (-1799.22, -1859.54, and -1029.02 respectively) and higher AIC (3606.44, 3725.09, and 2064.04 respectively) and BIC (3623.3, 3737.73, and 2076.68 respectively) values. These findings underscore the superior performance of the EWW model in capturing the underlying characteristics of the student dropout data, making it the most suitable choice among the compared models.

Table 7. Median, 25-percentile, 75-percentile and Interquartile range (IQR) of student dropout times (in months) for the five-parameter exponentiated-Weibull Weibull (EWW) alongside the Exponential Weibull (EW), Weibull Weibull (WW), and Weibull Exponential (WE) distributions.

	EWW	WW	WE	EW
25%	0.3	1.6	1.5	1.2
Median ($M(t)$)	0.8	2.0	1.9	2.1
75%	1.6	2.5	2.5	3.1
IQR	1.3	0.9	1.0	1.9

Table 7 shows the estimate of the dropout times for the student from the online class module using the four distributions based on the Bayesian parameter estimate. The result showed that median dropout time using the EWW distribution from the online class presentation is 0.8 months (24 days) from the day the class commenced. On the other hand, if any of the WW, WE, EW distributions is assumed, the median dropout time from the online class presentation is approximately 2 months (60 days) from the day the class commenced. The median estimate using any of WW, WE, EW distribution is significantly larger than EWW and also unrealistic given that most students dropout from online class modules in the first few days after the commencement of the online presentation [11,15].

8. Conclusion

In this paper, the five-parameter Exponentiated Weibull Weibull (EWW) distribution with Weibull base failure rate parameterization was presented. Specifically, the EWW distribution was developed to model the dropout times of students in an online class module. The parameters of the model were

estimated using both Maximum Likelihood and Bayesian estimation procedures to recommend the best estimation method. The simulation results revealed that, at various hazard types and censoring levels, the Bayesian estimates are more accurate and precise than the Maximum Likelihood estimates. Furthermore, the real-life analysis corroborated the findings from the simulation study, with the standard deviation of the Bayesian estimates found to be significantly lower than that of the Maximum Likelihood estimates, indicating that the Bayesian estimates are more efficient for the EWW distribution. Additionally, the comparative goodness-of-fit of the EWW distribution alongside Weibull Weibull (WW), Weibull Exponential (WE), and Exponential Weibull (EW) distributions was examined. The goodness-of-fit results showed that the EWW distribution fits the student dropout times dataset better than the three competing nested distributions. Finally, the estimate of the median dropout time of the students using the EWW distribution revealed that the median dropout time from the online presentation is 24 days. This indicates that the instructor should pay more attention to student attendance around this time.

Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the article.

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

References

1. Wang, W.; Zhao, Y.; Wu, Y.J.; Goh, M. Factors of dropout from MOOCs: a bibliometric review. *Library Hi Tech* **2023**, *41*, 432–453.
2. Lee, Y.; Choi, J. A review of online course dropout research: Implications for practice and future research. *Educational Technology Research and Development* **2011**, *59*, 593–618.
3. Rahmani, A.M.; Groot, W.; Rahmani, H. Dropout in online higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education* **2024**, *21*, 19.
4. Park, J.H. Factors Related to Learner Dropout in Online Learning. *Online Submission* **2007**.
5. Kim, S.; Choi, E.; Jun, Y.K.; Lee, S. Student dropout prediction for university with high precision and recall. *Applied Sciences* **2023**, *13*, 6275.
6. Choi, J.; Lee, J.; Park, M.Y.; Kim, H.K. Heterogeneity in Korean school dropouts and its associations with emerging adulthood adjustment. *Journal of Applied Developmental Psychology* **2023**, *85*, 101509.
7. Lee, J.; Song, H.D.; Kim, Y. Quality factors that influence the continuance intention to use MOOCs: An expectation-confirmation perspective. *European Journal of Psychology Open* **2023**.
8. Christou, V.; Tsoulos, I.; Loupas, V.; Tzallas, A.T.; Gogos, C.; Karvelis, P.S.; Antoniadis, N.; Glavas, E.; Giannakeas, N. Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications* **2023**, *225*, 120079.
9. Zhou, Y.; Zhao, J.; Zhang, J. Prediction of learners' dropout in E-learning based on the unusual behaviors. *Interactive Learning Environments* **2023**, *31*, 1796–1820.
10. Rodríguez, P.; Villanueva, A.; Dombrovskaia, L.; Valenzuela, J.P. A methodology to design, develop, and evaluate machine learning models for predicting dropout in school systems: the case of Chile. *Education and Information Technologies* **2023**, *28*, 10103–10149.
11. Bañeres, D.; Rodríguez-González, M.E.; Guerrero-Roldán, A.E.; Cortadas, P. An early warning system to identify and intervene online dropout learners. *International Journal of Educational Technology in Higher Education* **2023**, *20*, 3.
12. Huo, H.; Cui, J.; Hein, S.; Padgett, Z.; Ossolinski, M.; Raim, R.; Zhang, J. Predicting dropout for nontraditional undergraduate students: a machine learning approach. *Journal of College Student Retention: Research, Theory & Practice* **2023**, *24*, 1054–1077.

13.   Cardona, T.; Cudney, E.A.; Hoerl, R.; Snyder, J. Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice* **2023**, *25*, 51–75.

14.   Ali, D.A.; Hussein, A.M. Analysis of cox proportional hazard model for dropout students in university: case study from SIMAD university. *Journal of Applied Research in Higher Education* **2024**, *16*, 820–830.

15.   Ameri, S.; Fard, M.J.; Chinnam, R.B.; Reddy, C.K. Survival analysis based framework for early prediction of student dropouts. Proceedings of the 25th ACM international on conference on information and knowledge management, 2016, pp. 903–912.

16.   Masci, C.; Giovio, M.; Mussida, P.; others. Survival models for predicting student dropout at university across time. *Education and new developments* **2022**, p. 203.

17.   Gury, N. Dropping out of higher education in France: a micro-economic approach using survival analysis. *Education Economics* **2011**, *19*, 51–64.

18.   Abubakar, J.; Abdullah, M.A.A.; Olaniran, O.R. Variational bayesian inference for exponentiated weibull right censored survival data. *Statistics, Optimization & Information Computing* **2023**, *11*, 1027–1040.

19.   Al-Essa, L.A.; Muhammad, M.; Tahir, M.H.; Abba, B.; Xiao, J.; Jamal, F. A new flexible four parameter bathtub curve failure rate model, and its application to right-censored data. *IEEE Access* **2023**.

20.   Mudholkar, G.S.; Srivastava, D.K. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability* **1993**, *42*, 299–302.

21.   Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. *Bayesian data analysis*; Chapman and Hall/CRC, 1995.

22.   Olaniran, O.; Abdullah, M. Subset selection in high-dimensional genomic data using hybrid variational Bayes and bootstrap priors. Journal of Physics: Conference Series. IOP Publishing, 2020, Vol. 1489, p. 012030.

23.   Olaniran, O.R.; Alzahrani, A.R.R. On the Oracle Properties of Bayesian Random Forest for Sparse High-Dimensional Gaussian Regression. *Mathematics* **2023**, *11*, 4957.

24.   Olaniran, O.R.; Abdullah, M.A.A. Bayesian weighted random forest for classification of high-dimensional genomics data. *Kuwait Journal of Science* **2023**, *50*, 477–484.

25.   Olaniran, O.R.; Abdullah, M.A.A. Bayesian analysis of extended cox model with time-varying covariates using bootstrap prior. *Journal of Modern Applied Statistical Methods* **2020**, *18*, 7.

26.   Olaniran, O.R.; Yahya, W.B. Bayesian hypothesis testing of two normal samples using bootstrap prior technique. *Journal of Modern Applied Statistical Methods* **2017**, *16*, 34.

27.   Robert, C.P.; others. *The Bayesian choice: from decision-theoretic foundations to computational implementation*; Vol. 2, Springer, 2007.

28.   Olaniran, O.R.; Abdullah, M.A.A.B. Bayesian Random Forest for the Classification of High-Dimensional mRNA Cancer Samples. Proceedings of the Third International Conference on Computing, Mathematics and Statistics (iCMS2017) Transcending Boundaries, Embracing Multidisciplinary Diversities. Springer, 2019, pp. 253–259.

29.   Nassar, M.M.; Eissa, F.H. On the exponentiated Weibull distribution. *Communications in Statistics-Theory and Methods* **2003**, *32*, 1317–1336.

30.   Nadarajah, S.; Gupta, A.K. On the moments of the exponentiated Weibull distribution. *Communications in Statistics-Theory and Methods* **2005**, *34*, 253–256.

31.   Khan, S.A. Exponentiated Weibull regression for time-to-event data. *Lifetime data analysis* **2018**, *24*, 328–354.

32.   ELGARHY, M.; HASSAN, A. Exponentiated weibull weibull distribution: Statistical properties and applications. *Gazi University Journal of Science* **2019**, *32*, 616–635.

33.   Kalbfleisch, J.D.; Prentice, R.L. *The statistical analysis of failure time data*; John Wiley & Sons, 2011.

34.   Therneau, T.; others. A package for survival analysis in S. *R package version* **2015**, *2*, 2014.

35.   Popoola, J.; Yahya, W.B.; Popoola, O.; Olaniran, O.R. Generalized self-similar first order autoregressive generator (gsfo-arg) for internet traffic. *Statistics, Optimization & Information Computing* **2020**, *8*, 810–821.

36.   Kuzilek, J.; Hlosta, M.; Zdrahal, Z. Open university learning analytics dataset. *Scientific data* **2017**, *4*, 1–8.