Journal of Data Science 0 (0), 1–19
September 2024

DOI: 0000
Research Article

Predictive Modeling of Short-Term Recidivism: A Mixture Cure Rate Approach with Diverse Link Functions

Chaeyeon Yoo[1,*] and Dipak K. Dey[1,†]

[1]*215 Glenbrook Rd, Storrs, University of Connecticut, USA*

Abstract

Addressing short-term recidivism presents a significant challenge within the criminal justice system, necessitating robust predictive models that can accurately identify risk factors and predict outcomes. This paper introduces application of Bayesian cure rate models to predict short-term recidivism using data from individuals released from Iowa prisons in 2018. The cure rate models effectively capture the recidivism risk by distinguishing between cured individuals—those unlikely to re-offend-and those susceptible to relapse. By implementing Accelerated Failure Time (AFT) models combined with logistic regression and various link functions, including logit, skewed logit, reversed power logit, and flexible generalized logit, we enhance the model's flexibility and interpretability. Through a comparative analysis of these models, key predictors of recidivism are identified and discussed in detail.

Keywords *Accelerated failure time; Bayesian inference; generalized logistic distribution; time-to-event*

1 Introduction

Addressing short-term recidivism presents a significant challenge within the criminal justice system, necessitating robust predictive models that can accurately identify risk factors and predict outcomes. Criminal recidivism, defined as the return to criminal behavior following intervention by the criminal justice system, has serious implications for both rehabilitation and public safety. Statistical modeling has become a crucial tool in analyzing recidivism data, aiding in the identification of factors that contribute to re-offending and predicting future behaviors of individuals involved with the justice system. In particular, survival analysis has been widely applied in criminology since Partanen (1969), providing valuable insights into the complexities of data by analyzing time-to-event data. Within the framework of survival analysis, individuals undergo a period of treatment through correctional programs aimed at rehabilitation, and under satisfactory completion of this program, they are released back into society. The main concern is the probability of recidivism, defined as the likelihood that a randomly selected participant will re-offend and be re-incarcerated after release.

One approach that has earned attention is the use of cure rate models (Maltz, 1984; Maller and Zhou, 1996), which acknowledge the presence of distinct sub-populations within the justice-involved community. These models recognize two distinct groups: individuals who are never subject to the event of interest (in this case, recidivism), and those who remain susceptible to relapse over time despite interventions. This partitioning of the population into the subsets, also

*Corresponding author. Email: chaeyeon.yoo@uconn.edu.
†Simple footnote.

referred to as a split population (Schmidt and Witte, 1989, 2012), offers a unique perspective on recidivism by implying rapidly decreasing hazard rates due to the proportion of individuals who are cured. Recent work by De la Cruz et al. (2022) introduced a Bayesian mixture cure rate model tailored to recidivism analysis, allowing for the simultaneous estimation of short-term and long-term reoffending risks. This approach enhances the modeling of heterogeneous reoffending behavior by distinguishing between individuals who are temporarily versus permanently desisted from crime. The ability of cure rate models to parameterize both the probability of an eventual return to criminal behavior and the timing of such events allows for a comprehensive analysis.

Previous work on the cure rate model dates back to the conceptualization of the cure fraction after cancer therapy by Boag (1949). The latency portion, which relates to the survival modeling of the susceptible group, various framework has been applied. Parametric mixture cure rate models have been studied since Berkson and Gage (1952) expanded upon this cured notion to divide the population into susceptible and insusceptible individuals, refining the mixture model framework. Farewell (1977, 1982) introduced a novel application of mixture models by integrating logistic regression on the incidence portion, and Weibull distribution on the latency distribution. Yamaguchi (1992) introduced the accelerated failure time (AFT) models to the cure rate model, offering interpretable insights into the effects of baseline covariates on event timing. More recently, semiparametric mixture cure models have garnered attention. While keeping the logistic regression form for predicting the cure rate, the survival function was altered into a semiparametric form, where the baseline survival function, say $S_0(t)$, unspecified. This approach allows for greater flexibility in modeling the survival distribution without assuming a specific parametric form. Kuk and Chen (1992) and Sy and Taylor (2000) adopted a Cox proportional hazards model for the conditional survival function, while Li and Taylor (2002) and Zhang and Peng (2007) focused on a semiparametric AFT model.

While most studies focus on latency, the mixture model also incorporates an additional parameter, known as the cure factor, which estimates the proportion of individuals in the risk set who will never experience a failure. Typically, a logistic model with a logit link function is used. However, the standard logit link function assumes symmetry, which may not always be appropriate for all datasets. Previous research has highlighted the potential issues that arise from using symmetric link functions like logit and probit inappropriately. For instance, Nagler (1994) introduced the scobit model as an alternative to the logit and probit link functions, specifically addressing asymmetric cases by allowing for different rates of change in the response variable. Similarly, Chen et al. (1999) and Kim et al. (2008) utilized generalized t-link models as skewed links for dichotomous response data, offering alternatives for analyzing asymmetric behavior. Building on well-known link functions, Bazán et al. (2017) proposed power and reversal power links, while Prasetyo et al. (2020) developed the flexible generalized logit link function, expanding the repertoire of available link functions. These studies underscore the need for more flexible and asymmetric link functions to accurately capture the complexities in binary response models. By moving beyond the standard logit and probit link functions, more precise models that better reflect the data structures can be developed.

In this paper, we present an application of the cure rate model in the context of short-term recidivism prediction with a Bayesian approach. We extend the AFT cure rate model by integrating a variety of link functions for logistic regression, specifically logit, skewed logit, reversed power logit, and flexible generalized logit. These link functions, which have not been previously applied in recidivism modeling, allow us to better capture the complex relationship between predictors and recidivism likelihood. This approach provides more flexible and interpretable modeling of real data, facilitating a deeper understanding of the factors influencing recidivism.

The data used in this paper consist of information on a cohort of releases from the Iowa prison system in 2018. This dataset, available from the Iowa government correctional system database (https://data.iowa.gov/Correctional-System/Iowa-Prison-Recidivism-and-Change-by-Cohort/dnzw-paxg/about_data), includes variables such as age, sex, race, original offense, convicted offense causing recidivism, and level of risk. Our analysis employs Bayesian methods, utilizing Stan for model fitting. Model performance is assessed based on Deviance Information Criterion (DIC), Leave-One-Out Information Criterion (LOOIC) and Watanabe-Akaike Information Criterion (WAIC), ensuring robust evaluation of the predictive accuracy and model fit.

2 Data Description

The Iowa 3 year recidivism dataset identifies case-level records of 5010 individuals leaving prison and reentering the community in the 2018 cohort. It provides individual's status over three years to assess whether their reentry into the community was successful or whether the individual was subject to recidivism. If individuals who returned to a correctional facility, the dataset includes the time until that event occurred. The dataset encompasses information on age, sex, race, original offense committed, convicted offense causing recidivism (if applicable), and the individual's level of risk. Table 1 provides a detailed overview of categorical variables in our dataset, summarizing both the distribution of each category and the proportion of individuals within each category who experienced recidivism. Specifically, the proportion column presents the percentage of the total 5010 subjects represented by each category. For example, females constitute 14.77% of the total subjects.

Kaplan–Meier estimates were used to evaluate the proportion of individuals free from recurrence across different covariate groups. For individuals with no observed re-incarceration event, the time to recurrence was set to the maximum follow-up duration of 1,095 days. As illustrated in Figure 1, the survival curves indicate that supervision type, offense type, race, and sex are all associated with time to re-incarceration among the susceptible population. Comparing the supervision types, those under work release shows a higher recurrence of events compared to those in prison. Regarding offense type, subjects with public order and violence related offenses tend to have a higher rate of recidivism. When examining the race graph, the gaps between white and black against Asian or Pacific Islander are noticeable, indicating there is a higher rate of certain race being re-incarcerated. Similarly, gender differences are evident, which female than male group showing a higher trend of recurrence.

3 Mixture Cure Rate Model

In survival analysis, the survival function $S(t)$ quantifies the probability of observing a survival time T greater than a given value t, expressed as $S(t) = P(T > t)$. Conventionally, it is assumed that $S(t)$ approaches 0 as t approaches ∞. However in real-world scenarios, not all subjects experience the event of interest, leading to situations where $S(t)$ does not converge to 0. This highlights the significance of cure rate models, which are particularly effective in accounting for a subset of subjects who are cured and thereby permanently immune to the event.

The observed survival time for subject i is defined as $y_i = \min(T_i, C_i)$, where T_i represents the true survival time and C_i denotes the censoring time. The censoring status is indicated by the variable δ_i, where

Table 1: Overview of categorical input variables.

	Proportion (%)
Age	
Months.Supervised	
Race.Asian.or.Pacific.Islander	0.68
RaceBlack	21.40
RaceHispanic	5.75
RaceWhite	70.12
RaceAmerican.Indian.or.Alaska.Native	2.06
SexMale	85.23
SexFemale	14.77
Supervising.Unit_Anamosa.State.Penitentiary	3.59
Supervising.Unit_Clarinda.Correctional.Facility	10.00
Supervising.Unit_Districts.Compact.Regions	0.66
Supervising.Unit_Fort.Dodge.Correctional.Facility	38.63
Supervising.Unit_Iowa.Correctional.Institution.for.Women	14.71
Supervising.Unit_Iowa.Medical.Classification.Center	13.91
Supervising.Unit_Iowa.State.Penitentiary	1.46
Supervising.Unit_Mount.Pleasant.Correctional.Facility	15.71
Supervising.Unit_Newton.Correctional.Facility	17.13
Supervising.Unit_North.Central.Correctional.Facility	9.40
Supervision.TypeWork.Release	29.08
Supervision.TypeWork.Prison	70.92
Supervision.End.ReasonDischarged-Expiration.of.Sentence	23.79
Supervision.End.ReasonParole.Granted	64.05
Supervision.End.ReasonParoled.to.Detainer.INS	0.90
Supervision.End.ReasonParoled.to.Detainer.Iowa	0.68
Supervision.End.ReasonParoled.to.Detainer.Out.of.State	1.18
Supervision.End.ReasonParoled.to.Detainer.U.S..Marshall	0.98
Supervision.End.ReasonParoled.w.Immediate.Discharge	2.59
Supervision.End.ReasonReleased.to.Special.Sentence	5.82
Supervision.Offense.TypeDrug	30.32
Supervision.Offense.TypeOther	8.12
Supervision.Offense.TypeProperty	26.05
Supervision.Offense.TypePublic.Order	12.12
Supervision.Offense.TypeViolent	23.40
Risk.Ranking_Low	51.54
Risk.Ranking_Moderate	15.91
Risk.Ranking_High	51.54

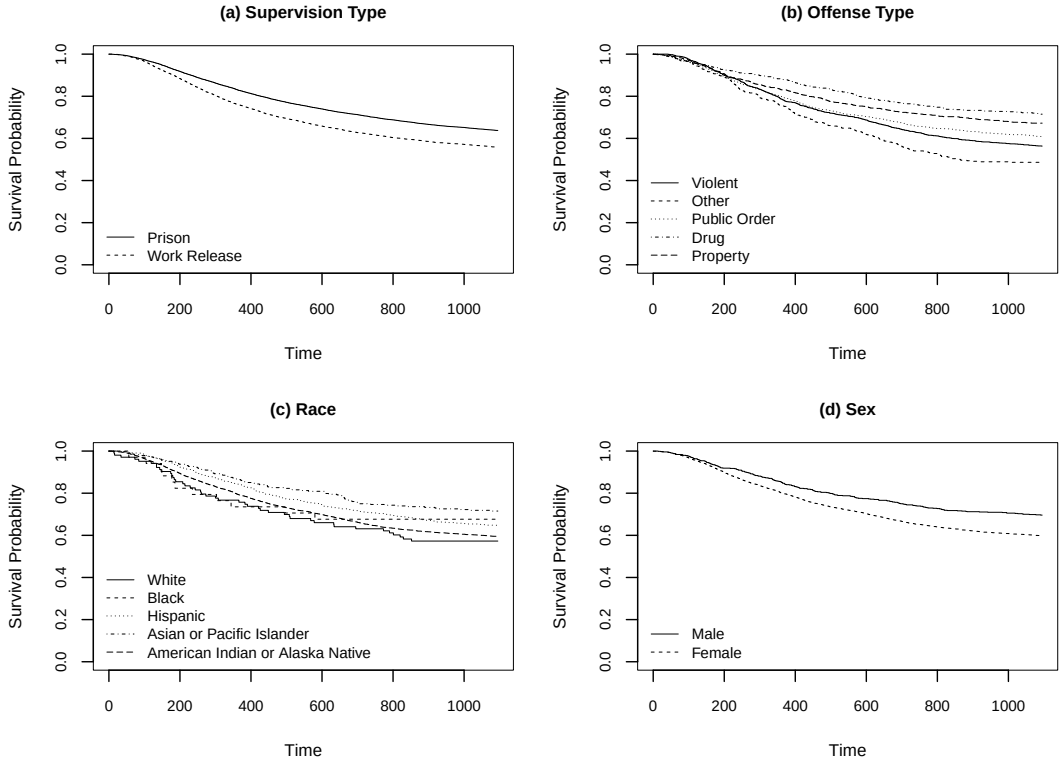

Figure 1: Kaplan-Meier survival curves by (a) Release type, (b) Offense type, (c) Race, (d) Sex.

$$\delta_i = \begin{cases} 1 & \text{if } T_i \leqslant C_i \text{ (Uncensored)} \\ 0 & \text{if } T_i > C_i \text{ (Censored).} \end{cases}$$

The model includes two sets of covariates: $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^T \in \mathbb{R}^{n \times p}$, which represents the matrix of covariates used in the cure rate component, and $\boldsymbol{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)^T \in \mathbb{R}^{n \times q}$, which corresponds to the covariates in the survival component. The covariate matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$ consist of subsets of available covariates, and they may share common elements.

The cure rate model is comprised of two distinct components: a binary component that models the probability of being cured or permanently non-vulnerable to the event of interest, and a survival component that models the time until event occurs for the uncured group. The cured group represents individuals who will never encounter the event.

Defining the cure indicator for individual i as

$$w_i = \begin{cases} 1 & \text{if subject } i \text{ is uncured} \\ 0 & \text{if subject } i \text{ is cured,} \end{cases} \tag{1}$$

the decomposition for event time T consisting of the susceptible and insusceptible individuals is

$$T = wT_s + (1 - w)\infty, \tag{2}$$

where $T = \infty$ is considered that the event never happens for the cured group. Let $F(t)$ and $S(t)$ denote the cumulative distribution function (cdf) and survival function, respectively, for the overall population. In contrast, $F_s(t)$ and $S_s(t)$ represent the corresponding functions for the susceptible group. The functions $F_s(t)$ and $S_s(t)$ are proper distribution functions of the event time T, meaning they are well-defined only for the susceptible group and satisfy the properties of a complete probability distribution (e.g., $S_s(0) = 1$, $\lim_{t\to\infty} S_s(t) = 0$). Thus, we have

$$\begin{aligned} &\Pr(T \leqslant t | w = 1) = F_s(t) \\ &\Pr(T \leqslant t | w = 0) = 0, \end{aligned} \tag{3}$$

and the functions of the overall population can be re-written with $\Pr(w = 1) = p_w$ as

$$\begin{aligned} F(t) = \Pr(T \leqslant t) &= \Pr(T \leqslant t \mid w = 1) \Pr(w = 1) + \Pr(T \leqslant t \mid w = 0) \Pr(w = 0) \\ &= p_w F_s(t) + 0 = p_w F_s(t) \end{aligned} \tag{4}$$

or equivalently,

$$S(t) = (1 - p_w) + p_w S_s(t). \tag{5}$$

4 Link Functions

In the context of cure rate models, the binary component often relies on logistic regression for modeling the probability of being cured or permanently non-vulnerable to the event of interest. Consider a vector $\boldsymbol{w} = (w_1, w_2, ..., w_n)'$ consisting of n independent dichotomous random variables with probability $p_{w_i} = \Pr(w_i = 1 | \boldsymbol{x}_i)$. $\boldsymbol{x}_i = (x_{i1}, ..., x_{ik})'$ denote a $k \times 1$ vector of covariates for the ith subject, and $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_k)'$ denotes a $k \times 1$ vector of regression coefficients. Then the probability p_{w_i}, which is the probability of the ith individual to be susceptible to the event, is predicted by applying a link function to a linear combination of predictor variables.

The logit link function is a standard choice in logistic regression, serving to connect the linear predictor to to probability of the binary response. However, it is not entirely accurate to claim that the logit link function is suitable for all types of data in the context of binary response models. While the logit link is indeed one of the most commonly used link in logistic regression, it is important to recognize that it is a symmetric link function. The symmetry implies that it assumes an equal rate of change in the response variable with respect to the predictor variable, which may not always be the case in real world scenarios. The relationships between predictors and binary outcomes can be asymmetric and even exhibit varying degrees of nonlinearity. In these situations, link functions based on the inverse of non-standard cdfs can better accommodate the asymmetry and nonlinearity present in the data.

4.1 Logit Link

The inverse of the logistic function, or the logit function transforms probabilities into their corresponding log-odds. Then the probability p_{w_i} of the binary response variable $w_i = 1$ is expressed as

$$p_{w_i} = \frac{\exp(\boldsymbol{x_i'\gamma})}{1 + \exp(\boldsymbol{x_i'\gamma})}, \quad i = 1, \ldots, n,$$

and the logit link is represented as

$$\text{logit}(p_{w_i}) = \log\left(\frac{p_{w_i}}{1 - p_{w_i}}\right) = \boldsymbol{x_i'\gamma}.$$

4.2 Skewed and Reversed Power Logit Link

While the standard logit link function is symmetric there are instances where asymmetry in the relationship between the predictors and the response variable exists in real life. This is where the use of an asymmetric link distribution is required. The asymmetric link distributions offer greater flexibility in modeling complex relationships between predictors and binary outcomes, while effectively capturing skewness. Couple of the well-known distributions to derive the asymmetric link functions are derived from the Type I and II generalized logistic distribution (Bazán et al., 2017; Nagler, 1994).

The Type I generalized logistic distribution is also called exponentiated logistic distribution or the skewed logistic distribution. It is an exponential distribution form of the standard logistic distribution, by adding a parameter α_s to the definition of the distribution. Simplifying the formula with $u_i = \boldsymbol{x_i'\gamma}$, we have

$$G(u_i) = \frac{1}{\{1 + \exp(-u_i)\}^{\alpha_s}}, \qquad -\infty < u_i < \infty, \alpha > 0, \tag{6}$$

and the corresponding probability density function (pdf) is given by

$$g(u_i) = \frac{\alpha_s \exp(-u_i)}{\{1 + \exp(-u_i)\}^{\alpha_s + 1}}, \qquad -\infty < u_i < \infty, \alpha > 0. \tag{7}$$

This distribution is used to generate an alternative estimator to the logit function, and is called the skewed-logit, because it allows for a skewed response curve, with α serving as a parameter to measure skewness. Compared to the standard logit function, but rather the shape of the curve changes as the value of α_s varies. The link function, denoted by slogit, is represented as

$$\text{slogit}(p_{w_i}) = \log\left\{(1 - p_{w_i}^{1/\alpha_s})^{-1} - 1\right\} = \boldsymbol{x}_i'\boldsymbol{\gamma}. \tag{8}$$

Another alternative to the logit function is derived in a similar direction. Under Type II generalized logistic distribution, with an additional parameter λ_{rp}, giving the reversed exponentiated cdf form as follows:

$$H(u_i) = 1 - \left[1 - \frac{1}{\{1+\exp(-u_i)\}}\right]^{\lambda_{rp}}, \qquad -\infty < u_i < \infty, \lambda > 0, \tag{9}$$

with the corresponding pdf

$$h(u_i) = \frac{\lambda_{rp}\exp(u_i)}{\{1+\exp(u_i)\}^2}\left\{\frac{1}{1+\exp(u_i)}\right\}^{\lambda_{rp}-1}, \qquad -\infty < u_i < \infty, \lambda_{rp} > 0. \tag{10}$$

Finally the link function denoted by rplogit is represented by

$$\text{rplogit}(p_{w_i}) = \log\left\{(1-p_{w_i})^{-1/\lambda_{rp}} - 1\right\} = \boldsymbol{x}_i'\boldsymbol{\gamma}. \tag{11}$$

Again, as the λ_{rp} value varies, the shape of the curve changes with some skewness.

4.3 Flexible Generalized Logit Link

The flexible generalized logit (fglogit) link function is derived from the exponentiated-exponential logistic (EEL) distribution (Ghosh and Alzaatreh, 2018), which itself is based on the exponentiated-exponential distribution introduced by Gupta and Kundu (2001). This distribution combines features of both the exponentiated-exponential and logistic distributions, resulting in a flexible family capable of modeling skewness and heavy tails. With $\mu \in \mathrm{R}$ and $s > 0$ as the location and scale parameters of logistic distribution and $\alpha_{fg} > 0$ and $\lambda_{fg} > 0$ as the shape and scale parameters, the pdf and cdf are given as

$$e(u_i) = \frac{\alpha_{fg}\lambda_{fg}\exp\left(u_i-\mu\right)/s}{s\{1+\exp\left(u_i-\mu\right)/s\}^{\lambda_{fg}+1}}\left[1-\{1+\exp\left(u_i-\mu\right)/s\}^{-\lambda_{fg}}\right]^{\alpha_{fg}-1}, \tag{12}$$

$$E(u_i) = [1-\{1+\exp(u_i-\mu)/s\}^{-\lambda_{fg}}]^{\alpha_{fg}}. \tag{13}$$

When $\mu = 0$ and $s = 1$, specific values of λ_{fg} and α_{fg} lead to well-known distributions: when $\lambda_{fg} = \alpha_{fg} = 1$, the distribution reduces to that of the logistic distribution, while $\lambda_{fg} = 1$ results in the Type I Generalized logistic distribution, and $\alpha_{fg} = 1$ yields the Type II Generalized logistic distribution. Additionally, depending on the inequality of the parameters, the direction of skewness differs. The distribution exhibits negative skewness for $\alpha_{fg} < \lambda_{fg}$, positive skewness for $\lambda_{fg} < \alpha_{fg}$, and symmetry for $\alpha_{fg} = \lambda_{fg}$.

Utilizing the inverse of the standard EEL cdf, where $\mu = 0$, and $s = 1$, a flexible generalized logit function, denoted by fglogit, is formed (Prasetyo et al., 2020):

$$\text{fglogit}(p_{w_i}) = \log\left\{(1-p_{w_i}^{1/\alpha_{fg}})^{-1/\lambda_{fg}} - 1\right\} = \boldsymbol{x}_i'\boldsymbol{\gamma} \tag{14}$$

For the following sections on simulation and real data application, we employ a mixture cure rate model, which combines three accelerated failure time models (AFT)—Weibull (WB), lognormal (LN), and loglogistic (LLG)—for the latency component with logistic regression models using different link functions for the incidence component: Logit (L), skewed logit (S), reversed power logit (RP), and flexible generalized logit (FG). This result in 12 different models, where each model is denoted by the corresponding combination of acronyms, for example, Weibull with reversed power logit link is represented as WBRP.

5 Bayesian inference

The proposed model combines the elements of the accelerated failure time model and logistic regression to account for both the latency and incidence components of survival data. Let $\boldsymbol{D}_{\text{obs}} = (\boldsymbol{D}_1, \ldots, \boldsymbol{D}_n)^T = (\boldsymbol{y}, \boldsymbol{\delta}, \boldsymbol{X}, \boldsymbol{Z})$ be the observed data. The model consists of two components: incidence and latency. The incidence component is governed by the parameter vector $\boldsymbol{\nu} = (\boldsymbol{\gamma}^T, \boldsymbol{\eta}^T)^T$, where $\boldsymbol{\gamma}$ represents the regression coefficients for the incidence model, and $\boldsymbol{\eta} = (\alpha_s, \lambda_{rp}, \alpha_{fg}, \lambda_{fg})^T$ includes additional parameters associated with the chosen link function. The latency component is characterized by the parameter vector $\boldsymbol{\upsilon} = (\boldsymbol{\beta}^T, \boldsymbol{\sigma}^T)^T$, where $\boldsymbol{\beta}$ denotes the regression coefficients for the latency model, and $\boldsymbol{\sigma} = (\sigma_{wb}, \sigma_{ln}, \sigma_{llg})^T$ contains parameters for the latency distribution. The survival function for the uncured individuals can be express as $S_s(y_i|\boldsymbol{z}_i; \boldsymbol{\upsilon}) = S_s(y_i; \boldsymbol{z}_i^T\boldsymbol{\beta}, \boldsymbol{\sigma})$, where $S_s(\cdot)$ denotes the survival function corresponding to the chosen latency distribution. Let $p_{w_i}(\boldsymbol{x}_i; \boldsymbol{\nu}) = \Pr(w_i = 1 \mid \boldsymbol{x}_i; \boldsymbol{\nu})$. The corresponding likelihood function with the observed data is as follows

$$L_{\text{obs}} = \Pi_{i=1}^n \{p_{w_i}(\boldsymbol{x}_i; \boldsymbol{\nu}) f(y_i|\boldsymbol{z}_i; \boldsymbol{\upsilon})\}^{\delta_i} \{1 - p_{w_i}(\boldsymbol{x}_i; \boldsymbol{\nu}) + p_{w_i}(\boldsymbol{x}_i; \boldsymbol{\nu}) S(y_i|\boldsymbol{z}_i; \boldsymbol{\upsilon})\}^{(1-\delta_i)}. \tag{15}$$

5.1 Bayesian assessment

Model performance is evaluated using three Bayesian model comparison criteria: Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002), Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010), and Leave-One-Out Information Criterion (LOOIC) (Vehtari et al., 2017). These criteria balance model fit and complexity, providing insight into the predictive accuracy and generalizability of the models.

The DIC assesses model fit by penalizing model complexity, making it suitable for comparing models with different numbers of parameters. It is computed as

$$\text{DIC} = \text{Dev}(\overline{\boldsymbol{\Theta}}) + 2p_D,$$

where $\boldsymbol{\Theta}$ represents the collection of all model parameters, $\text{Dev}(\boldsymbol{\Theta}) = -2\log L(\boldsymbol{\Theta}|\boldsymbol{D}_{\text{obs}})$ is the deviance function, and $p_D = \overline{\text{Dev}}(\boldsymbol{\Theta}) - \text{Dev}(\overline{\boldsymbol{\Theta}})$ represents the effective number of parameters.

The WAIC is a fully Bayesian alternative to DIC, which asymptotically approximates Bayesian cross-validation. It is computed as

$$\text{WAIC} = -2\sum_{i=1}^{n} \log \mathbb{E}_{\boldsymbol{\Theta}}[p(y_i|\boldsymbol{\Theta})] + 2p_{\text{WAIC}},$$

where the first term represents the log pointwise predictive density (lpd) averaged over the posterior, and $p_{\text{WAIC}} = \sum_{i=1}^n \text{Var}_{\boldsymbol{\Theta}}(\log p(y_i|\boldsymbol{\Theta}))$ accounts for model complexity by estimating the effective number of parameters.

The LOOIC evaluates out-of-sample predictive performance using leave-one-out cross-validation (LOO-CV). It estimates the expected log predictive density for new observations by approximating

$$\text{LOOIC} = -2\sum_{i=1}^{n} \log p(y_i|y_{-i}),$$

where $p(y_i|y_{-i})$ represents the leave-one-out predictive density for observation i. A practical approximation is given by

$$\text{LOOIC} = -2\sum_{i=1}^{n} \log \mathbb{E}_{\boldsymbol{\Theta}}[p(y_i|\boldsymbol{\Theta})] + 2p_{\text{LOOIC}},$$

where p_{LOOIC} is an estimate of the effective number of parameters. Since LOOIC is computed directly from posterior samples, it provides a more robust assessment of model generalizability than DIC.

These model selection criteria are used to compare different models and determine the best fit for the given data, where smaller values indicate better model performance.

6 Simulation Study

6.1 Prior setting

In fitting our cure rate models, we adopted a Bayesian approach with independent prior distributions for all parameters to ensure proper regularization while allowing flexibility.

For the logistic regression coefficient vector $\boldsymbol{\gamma}$, we used weakly informative normal priors to provide reasonable regularization without imposing strong assumptions. Specifically, each coefficient follows $\gamma_j \sim N(0,1),\ j = 1,\ldots,m$. For the link function parameters of slogit, rplogit, and fglogit, we assigned lognormal priors, such that, $\alpha_s, \lambda_{rp}, \alpha_{fg}, \lambda_{fg} \sim \text{Lognormal}(0,1)$. These priors reflect the assumption that the shape parameters are likely near 1 while allowing flexibility in their estimation. For the latency component of the cure rate model, the scale parameter σ follows a Cauchy prior to accommodate heavy-tailed behavior, with $\sigma_{wb}, \sigma_{ln}, \sigma_{llg} \sim \text{Cauchy}(0, 2.5)$. The regression coefficient $\boldsymbol{\beta}$, which model the survival component, follow the same weakly informative normal prior as γ, which is $\beta_k \sim N(0,1),\ k = 1,\ldots,p$.

6.2 Simulation result

In this simulation study, we evaluated the performance of various mixture cure rate models under different link function settings for the cure fraction. To assess the model performance on true value recovery, we generate a total of 12 simulated datasets based on the twelve models with sample size of $n = 1000$. The covariates for the latency include an intercept and two numeric data generated from the standard normal distribution, while the incidence component include two scaled covariates generated from a uniform distribution without an intercept. For true parameter values, we have $\boldsymbol{\beta} = (1,-1,-2)^T, \boldsymbol{\gamma} = (0.5, 0.6)^T, \boldsymbol{\sigma} = (1.5, 1.5, 1.5)^T, \boldsymbol{\eta} = (1.5, 1.5, 2.0, 1.5)^T$. Event times for uncured individuals were simulated based on the specified latency distribution, while cure status was drawn from a Bernoulli distribution according to the calculated cure probabilities. Random uniform censoring times were applied to simulate right-censoring, and individuals identified as cured were automatically assigned censored status.

Across all scenarios, the models demonstrated strong parameter recovery, with most true values falling within the 95% highest posterior density (HPD) intervals in Tables 2, 3, and 4. The models consistently produced accurate and stable estimates for both the survival and cure components. Flexible link functions such as slogit and rplogit also showed high recovery rate, with the estimated parameters providing additional modeling flexibility. Although the uncertainty for these parameters was somewhat larger with the fglogit, the overall inference remained robust, supporting the use of flexible link functions when modeling potential asymmetry in the cure fraction.

In addition to the main simulation results, we conducted a comprehensive assessment where each link function (logit, slogit, rplogit, and fglogit) was used to fit data generated from all other link function settings under Weibull, lognormal, and loglogistic latency models—resulting in twelve distinct true model scenarios. These analyses further confirmed the strong performance

Table 2: Weibull AFT mixture cure rate model: Comparison of True vs. Estimated Parameter Values.

Parameter	True Value	WBL	HPD 95% Interval	WBS	HPD 95% Interval
betaU[1]	1.000	1.000	[0.861, 1.148]	1.005	[0.874, 1.139]
betaU[2]	−1.000	−0.997	[−1.124, −0.872]	−0.980	[−1.082, −0.875]
betaU[3]	−2.000	−1.977	[−2.115, −1.841]	−0.464	[−2.128, −1.872]
betaC[1]	0.500	0.427	[0.180 , 0.677]	0.551	[0.324, 0.776]
betaC[2]	0.600	0.635	[0.380 , 0.900]	0.569	[0.569, 0.080]
σ	1.500	1.527	[1.342, 1.736]	1.475	[1.321, 1.647]
α	1.500	–	–	1.470	[1.237, 1.732]]
Parameter	True Value	WBRP	HPD 95% Interval	WBFG	HPD 95% Interval
betaU[1]	1.000	1.048	[0.868, 1.234]	0.978	[0.874, 1.139]
betaU[2]	−1.000	−0.981	[−1.124 , −0.837]	−0.989	[−1.117, −0.867]
betaU[3]	−2.000	−2.018	[−2.191 , −1.852]	−1.961	[−2.102, −1.825]
betaC[1]	0.500	0.513	[0.287 , 0.747]	0.863	[0.385, 1.557]
betaC[2]	0.600	0.580	[0.355 , 0.817]	0.990	[0.450, 1.754]
σ	1.500	1.473	[1.271, 1.715]	1.475	[1.263, 1.595]
α	2.000	–	–	1.181	[0.456, 3.267]
λ	1.500	1.483	–	0.904	[0.275, 2.112]

and robustness of our modeling framework. Across all settings, the recovery of true parameter values remained consistently accurate, with most estimates closely matching the true values and falling within the 95% HPD intervals. Moreover, Bayesian model assessment criteria indicated that the fitted models using alternative link functions performed comparably well to the true model, suggesting that the approach is not overly sensitive to modest link function misspecification. Full results from this extensive simulation study are provided in the supplementary material.

7 Application to Iowa recidivism data

We apply the twelve mixture cure rate models—combinations of three AFT models (WB, LN, and LLG) with four link functions (L, S, RP, and FG)—to predict the short-term recidivism. The data consists of the records of 5010 individuals leaving prison and reentering the community of Iowa in 2018.

Using stan, each model configuration was sampled across four independent chains with 2,000 iterations per chain, half of which were allocated to warm-up phases. The convergence of these models was rigorously evaluated through standard diagnostic checks including trace plots and assessments of the potential scale reduction factor (R-hat). All models achieved satisfactory convergence as indicated by R-hat values close to 1.0, and there were no divergent transitions nor any instances where the maximum tree depth was saturated.

7.1 Comparison

Table 5 compares model fit across Weibull, Lognormal, and Loglogistic mixture cure rate models using LOOIC, WAIC, and DIC, where smaller values indicate better fit. The LLGRP consistently

Table 3: Lognormal AFT mixture cure rate model: Comparison of True vs. Estimated Parameter Values.

Parameter	True Value	LNL	HPD 95% Interval	LNS	HPD 95% Interval
betaU[1]	1.000	0.992	[0.707, 1.303]	0.989	[0.698, 1.309]
betaU[2]	−1.000	−0.983	[−1.239, −0.732]	−0.990	[−1.211, −0.774]
betaU[3]	−2.000	−1.968	[−2.231, −1.719]	−1.972	[−2.220, −1.736]
betaC[1]	0.500	0.490	[0.217 , 0.777]	0.506	[0.229, 0.789]
betaC[2]	0.600	0.602	[0.323 , 0.897]	0.611	[0.332, 0.905]
σ	1.500	1.506	[1.316, 1.729]	1.508	[1.326, 1.722]
α	1.500	–	–	1.520	[1.206, 1.919]
Parameter	True Value	LNRP	HPD 95% Interval	LNFG	HPD 95% Interval
betaU[1]	1.000	0.984	[0.589, 1.430]	0.970	[0.668, 1.300]
betaU[2]	−1.000	−0.963	[−1.264 , −0.670]	−0.978	[−1.211, −0.749]
betaU[3]	−2.000	−1.955	[−2.276 , −1.653]	−1.960	[−2.212, −1.717]
betaC[1]	0.500	0.507	[0.260 , 0.772]	0.814	[0.335, 1.530]
betaC[2]	0.600	0.611	[0.350 , 0.899]	0.974	[0.425, 1.779]
σ	1.500	1.515	[1.277, 1.808]	1.501	[1.315, 1.722]
λ	1.500	1.503	[1.209, 1.807]	0.845	[0.235, 2.054]
α	2.000	–	–	1.182	[0.439, 3.266]

Table 4: Loglogistic AFT mixture cure rate model: Comparison of True vs. Estimated Parameter Values.

Parameter	True Value	LLGL	HPD 95% Interval	LLGS	HPD 95% Interval
betaU[1]	1.000	0.993	[0.771, 1.236]	0.989	[0.778, 1.216]
betaU[2]	−1.000	−0.989	[−1.190, −0.793]	−0.986	[−1.163, −0.813]
betaU[3]	−2.000	−1.989	[−2.209, −1.782]	−1.977	[−2.171, −1.789]
betaC[1]	0.500	0.498	[0.234 , 0.774]	0.508	[0.261, 0.760]
betaC[2]	0.600	0.604	[0.329 , 0.894]	0.608	[0.349, 0.876]
σ	1.500	1.491	[1.281, 1.733]	1.490	[1.300, 1.707]
α	1.500	–	–	1.518	[1.242, 1.848]
Parameter	True Value	LLGRP	HPD 95% Interval	LLGFG	HPD 95% Interval
betaU[1]	1.000	0.988	[0.697, 1.303]	0.980	[0.756, 1.221]
betaU[2]	−1.000	−0.977	[−1.216 , −0.746]	−0.985	[−1.170, −0.804]
betaU[3]	−2.000	−1.974	[−2.236 , −1.726]	−1.973	[−2.181, −1.773]
betaC[1]	0.500	0.505	[0.263 , 0.760]	0.817	[0.341, 1.529]
betaC[2]	0.600	0.609	[0.356 , 0.881]	0.976	[0.430, 1.777]
σ	1.500	1.478	[1.230, 1.778]	1.491	[1.290, 1.720]
λ	1.500	1.506	[1.240, 1.787]	0.839	[0.235, 2.034]
α	2.000	–	–	1.166	[0.442, 3.180]

Table 5: Bayesian assessment values with smaller values indicating better fit.

LOOIC	Weibull	Lognormal	Loglogistic
logit	32464.5	32498.6	32439.1
slogit	32466.8	32499.1	32438.8
rplogit	32464.3	32495.8	32438.3
fglogit	32467.8	32494.8	32440.2
WAIC	Weibull	Lognormal	Loglogistic
logit	32464.0	32498.0	32439.0
slogit	32466.5	32498.2	32439.3
rplogit	32464.3	32494.6	32439.1
fglogit	32468.7	32493.7	32441.3
DIC	Weibull	Lognormal	Loglogistic
logit	32403.8	32434.7	32378.8
slogit	32405.8	32434.4	32378.7
rplogit	32404.1	32432.7	**32378.0**
fglogit	32407.1	32432.8	32379.7

achieves the lowest LOOIC, WAIC, and DIC, making it the best-performing model. The fglogit model follows closely but with slightly higher values. In contrast, Weibull and Lognormal models show relatively comparable but worse performance, with the WBL performing best within the Weibull category and the LNFG performing best within the Lognormal category.

The selected models—WBL, LNFG, and LLGRP—represent the best-performing specifications within their respective AFT families, as determined by the lowest values of LOOIC, DIC, and WAIC. Among these, the LLGRP model introduces additional flexibility in modeling the cure fraction, which outperformed the conventional logistic link in WBL and the more generalized flexible generalized logit link in LNFG. The estimates for the shape parameters λ and α in the LNFG and LLGRP models provide additional flexibility in the cure probability function, allowing it to better capture nonlinearities in the data. The LLGRP model's superior performance suggests that incorporating this additional shape flexibility improves fit and better represents the underlying data patterns.

Across the WBL, LNFG, and LLGRP models, the posterior mean estimates for both the latency and cure components consistently showed the same direction of effect for corresponding covariates in Tables 6, 7, and 8. Regardless of the baseline distribution or link function, covariates with positive or negative effects in one model tended to show similar effects in others, suggesting that the models are capturing robust and stable associations.

While the tables highlight the top-performing models among twelve candidates based on Bayesian model assessment criteria, similar patterns of statistical significance and effect direction were observed across all model specifications. The persistence of key covariates as significant predictors across varying model structures underscores the robustness of the findings and supports the reliability and interpretability of the proposed framework across a range of plausible modeling scenarios.

Inference in survival component

Age. Age demonstrated a consistently positive association with latency across all models, indicating that older individuals tend to experience longer times to failure events such as recidivism.

Within the AFT framework, this implies a decelerating effect of age on the hazard rate, suggesting that aging may exert a protective influence and slow the progression toward adverse outcomes.

Race. Race was also found to significantly influence survival time. In particular, Hispanic and White individuals exhibited longer survival durations relative to the baseline group. These findings point to potential disparities in survival dynamics that may be shaped by underlying sociodemographic or systemic differences.

Supervision components. Facility-specific effects emerged as important determinants of latency. Individuals supervised at the Clarinda Correctional Facility, Iowa Medical Classification Center, and Iowa State Penitentiary exhibited significantly prolonged times to recidivism, as reflected by positive coefficients. These differences may reflect varying institutional environments, rehabilitative programs, or offender compositions across facilities. Additionally, individuals classified with a moderate risk ranking were associated with longer latency durations, further underscoring the role of institutional and individual-level characteristics in shaping survival outcomes.

Inference in incidence probability component

Months supervised. The duration of supervision showed a positive association with the probability of being uncured, indicating that individuals under longer supervision were more likely to reoffend. This suggests that extended supervision may reflect underlying risk factors or behavioral patterns associated with a higher likelihood of recidivism.

Supervision end reason. The reason for ending supervision was a strong predictor of incidence probability. In particular, individuals paroled to an in-state detainer (INS) were significantly more likely to be classified as cured, as indicated by the negative coefficient. This implies a lower likelihood of recidivism in this group. Notably, this covariate also appeared with a negative coefficient in the latency component, suggesting that these individuals not only have a higher probability of being cured but also tend to experience shorter times to failure if they do reoffend, indicating a complex dual effect.

Offense type. Offense type had a differential impact on both the cure probability and time to event. Individuals convicted of Public Order offenses had higher probabilities of being uncured and shorter survival times, suggesting a strong association with rapid and frequent recidivism. Conversely, those categorized under Other offense types were more likely to be cured but exhibited shorter latency when recidivism occurred. These findings underscore the importance of offense characteristics in influencing post-release outcomes.

8 Conclusion

The analysis revealed that several variables including race and supervision type were significant predictors of recidivism across most models, highlighting the complex interplay between demographic factors and institutional effects on recidivism rates. The application of the mixture cure rate model, incorporating different link functions in predicting recidivism provides valuable insights into the factors influencing the likelihood and timing of reoffending. Among the twelve models, which combined three AFT models: Weibull, lognormal, and loglogistic, and four link functions: logit, skewed logit, reversed power logit, and flexible generalized logit, the cure rate model with log-logistic AFT combined with reversed power logit link turned out to be the best model in terms of information criterion values. This model's capability to handle skewed data distributions and account for asymmetric relationships between covariates and outcomes

Table 6: WBL model: Posterior means and 95% HPD intervals for latency and cure rate variables.

Parameter	**Mean**	**HPD 95% Interval**
Latency Variables		
Age	4.39	[3.604, 5.289]
Race		
Hispanic	0.384	[0.135, 0.625]
White	0.365	[0.058, 0.668]
Supervising Unit		
Clarinda Correctional Facility	1.672	[0.779, 2.424]
Iowa Corr. Institution for Women	0.403	[0.181, 0.617]
Iowa Medical Classification Center	2.104	[1.219, 2.886]
Iowa State Penitentiary	0.564	[0.348, 0.759]
Newton Correctional Facility	0.340	[0.119, 0.551]
Supervision Type		
Work Release	0.288	[0.054, 0.508]
Supervision End Reason		
Paroled to Detainer (INS)	−0.563	[−0.741, −0.399]
Paroled to Detainer (Iowa)	1.806	[0.245, 3.127]
Paroled to Detainer (Out-of-State)	−0.675	[−1.021, −0.278]
Paroled w/ Immediate Discharge	1.055	[0.011, 1.930]
Supervision Offense Type		
Other	−0.755	[−0.983, −0.533]
Public Order	−0.112	[−0.218, −0.007]
Risk Ranking		
Moderate	0.131	[0.033, 0.232]
σ_{wb}	1.499	[1.435, 1.565]
Incidence Variables		
Months Supervised	0.268	[0.194, 0.343]
Supervising Unit		
Iowa State Penitentiary	−1.873	[−2.327, −1.459]
Supervision End Reason		
Paroled to Detainer (INS)	−0.562	[−0.795, −0.304]
Offense Type		
Other	−1.019	[−1.402, −0.654]
Public Order	0.308	[0.117, 0.492]
Violent	0.647	[0.401, 0.897]

Table 7: LNFG model: Posterior means and 95% HPD intervals for latency and cure rate variables.

Parameter	**Mean**	**HPD 95% Interval**
Latency Variables		
Age	4.335	[3.293, 5.362]
Race		
Hispanic	0.673	[0.337, 1.018]
White	0.674	[0.282, 1.086]
Supervising Unit		
Clarinda Correctional Facility	1.534	[0.547, 2.566]
Iowa Corr. Institution for Women	0.484	[0.162, 0.792]
Iowa Medical Classification Center	1.946	[0.942, 2.982]
Iowa State Penitentiary	0.425	[0.127, 0.730]
Newton Correctional Facility	0.418	[0.105, 0.733]
Supervision Type		
Work Release	0.346	[0.007, 0.688]
Supervision End Reason		
Paroled to Detainer (INS)	−0.743	[−0.926, −0.568]
Paroled to Detainer (Iowa)	1.990	[0.975, 3.137]
Paroled to Detainer (Out-of-State)	−1.173	[−1.646, −0.720]
Immediate Discharge	0.976	[0.135, 1.668]
Offense Type		
Other	−1.179	[−1.454, −0.913]
Public Order	−0.189	[−0.338, −0.036]
Risk Ranking		
Moderate	0.255	[0.132, 0.380]
σ_{ln}	0.996	[0.956, 1.051]
Incidence Variables		
Months Supervised	0.446	[0.269, 0.679]
Supervising Unit		
Iowa State Penitentiary	−2.705	[−3.787, −1.815]
Supervision End Reason		
Paroled to Detainer (INS)	−0.555	[−0.985, −0.169]
Offense Type		
Other	−0.966	[−1.642, −0.390]
Public Order	0.637	[0.299, 1.067]
Violent	1.178	[0.650, 1.814]
λ_{fg}	0.566	[0.303, 1.007]
α_{fg}	0.987	[0.519, 1.908]

Table 8: LLGRP model: Posterior means and 95% HPD intervals for latency and cure rate variables.

Parameter	**Mean**	**HPD 95% Interval**
Latency Variables		
Age	4.34	[3.319, 5.336]
Race		
Hispanic	0.545	[0.228, 0.867]
White	0.549	[0.166, 0.920]
Supervising Unit		
Clarinda Correctional Facility	1.511	[0.553, 2.527]
Iowa Corr. Institution for Women	0.465	[0.191, 0.733]
Iowa Medical Classification Center	1.954	[0.988, 2.963]
Iowa State Penitentiary	0.512	[0.260, 0.765]
Newton Correctional Facility	0.389	[0.122, 0.654]
Supervision Type		
Work Release	0.325	[0.035, 0.604]
Supervision End Reason		
Paroled to Detainer (INS)	−0.718	[−0.885, −0.554]
Paroled to Detainer (Iowa)	1.926	[0.896, 3.148]
Paroled to Detainer (Out-of-State)	−1.106	[−1.573, −0.631]
Immediate Discharge	0.967	[0.052, 1.700]
Offense Type		
Other	−1.120	[−1.357, −0.872]
Public Order	−0.146	[−0.274, −0.018]
Risk Ranking		
Moderate	0.218	[0.102, 0.334]
σ_{llg}	1.853	[1.756, 1.954]
Incidence Variables		
Months Supervised	0.380	[0.237, 0.564]
Supervising Unit		
Iowa State Penitentiary	−2.557	[−3.355, −1.912]
Supervision End Reason		
Paroled to Detainer (INS)	−0.578	[−0.960, −0.227]
Offense Type		
Other	−1.021	[−1.605, −0.484]
Public Order	0.507	[0.213, 0.863]
Violent	0.952	[0.507, 1.505]
λ_{rp}	0.623	[0.362, 1.105]

marks advancement over traditional logit-based approaches. The results indicate that demographic variables such as age, race, and sex, as well as supervision-related factors, play critical roles in recidivism.

These results underscore the effectiveness of using a flexible link approach within the Bayesian framework for predicting recidivism, providing valuable insights that can help tailor intervention programs and policies to mitigate risks associated with repeating offenses. The use of the cure rate model with asymmetric link functions represents a significant methodological advancement in the study of recidivism. By addressing the limitations of symmetric link functions and incorporating flexible modeling approaches, this study contributes to a deeper understanding of recidivism and offers a robust framework for future research and policy development.

Future research could further enhance the predictive power and flexibility of cure rate models by incorporating advanced modeling techniques. For instance, Li et al. (2016) proposed a flexible cure rate model for spatially correlated survival data using generalized extreme value distribution and Gaussian process priors. Incorporating Gaussian processes into the cure rate model could account for spatial dependencies and provide more precise predictions, particularly in contexts where geographic and environmental factors influence recidivism rates. For the incidence part, where we use logistic regression or other link functions to model the probability of being cured, replacing the linear predictor with a Gaussian process allows for a flexible representation for the relationship between covariates and the cure probability. This approach could lead to the development of more sophisticated models that better capture the complex dynamics of survival data with cure fraction and inform targeted policies and interventions.

Supplementary Material

The supplementary material includes the tables of comprehensive assessment where each link function (logit, slogit, rplogit, and fglogit) was used to fit data generated from all other link function settings under Weibull, lognormal, and loglogistic latency models—resulting in twelve distinct true model scenarios.

Data availability

The data used in this paper are openly available in the Iowa government data source for the correctional system at https://data.iowa.gov/Correctional-System/Iowa-Prison-Recidivism-and-Change-by-dnzw-paxg/about_data.

References

Bazán J, Torres-Avilés F, Suzuki A, Louzada F (2017). Power and reversal power links for binary regressions: An application for motor insurance policyholders. *Applied Stochastic Models in Business and Industry*, 33: 22–34.

Berkson J, Gage RP (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47(259): 501–515.

Boag JW (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(1): 15–53.

Chen MH, Dey DK, Shao QM (1999). A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, 94(447): 909–919.

De la Cruz R, Fuentes C, Padilla O (2022). A bayesian mixture cure rate model for estimating short-term and long-term recidivism. *Entropy (Basel)*, 25(1): 56.

Farewell VT (1977). A model for a binary variable with time-censored observations. *Biometrika*, 64(1): 43–46.

Farewell VT (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, 38(3): 1041–1046.

Ghosh I, Alzaatreh A (2018). A new class of generalized logistic distribution. *Commun. Stat.*, 47(14): 3459–3473.

Gupta RD, Kundu D (2001). Exponentiated exponential family: An alternative to gamma and weibull distributions. *Biom J.*, 43: 117–130.

Kim S, Chen MH, Dey DK (2008). Flexible generalized t-link models for binary response data. *Biometrika*, 95(3): 535–548.

Kuk A, Chen C (1992). A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79: 531–541.

Li CS, Taylor J (2002). A semi-parametric accelerated failure time cure model. *Statistics in Medicine*, 21: 3235–3247.

Li D, Wang X, Dey DK (2016). A flexible cure rate model for spatially correlated survival data based on generalized extreme value distribution and gaussian process priors. *Biometrical Journal*, 58(5): 1178–1197.

Maller RA, Zhou X (1996). *Survival analysis with long-term survivors.*

Maltz MD (1984). *Recidivism.* Academic Press, Florida.

Nagler J (1994). Scobit: An alternative estimator to logit and probit. *American Journal of Political Science*, 38(1): 230–255.

Partanen J (1969). On waiting time distributions. *Acta Sociologica*, 12(3): 132–143.

Prasetyo RB, Kuswanto H, Iriawan N, Ulama BSS (2020). Binomial regression models with a flexible generalized logit link function. *Symmetry*, 12(2): 221.

Schmidt P, Witte AD (1989). Predicting criminal recidivism using 'split population' survival time models. *Journal of Econometrics*, 40(1): 141–159.

Schmidt P, Witte AD (2012). *Predicting recidivism using survival models.* Springer Science & Business Media.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64: 583–639.

Sy J, Taylor J (2000). Estimation in a cox proportional hazards cure model. *Biometrics*, 56: 227–236.

Vehtari A, Gelman A, Gabry J (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5): 1413–1432.

Watanabe S (2010). Asymptotic equivalence of bayes cross-validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11: 3571–3594.

Yamaguchi K (1992). Accelerated failure-time regression models with a regression model of surviving fraction: An application to the analysis of 'permanent employment' in japan. *Journal of the American Statistical Association*, 87: 284–292.

Zhang J, Peng Y (2007). A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine*, 26(16): 3157–3171.