

Article

Parametric Frailty Analysis in Presence of Collinearity: An Application to Assessment of Infant Mortality

Olayan Albalawi ¹, Anu Sirohi ^{2,*}, Piyush Kant Rai ³ and Ayed R. A. Alanzi ⁴

¹ Department of Statistics, Faculty of Science, University of Tabuk, Tabuk 47512, Saudi Arabia; oalbalwi@ut.edu.sa

² Department of Mathematics, Sharda University, Greater Noida 201310, India

³ Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi 221005, India; raipiyush5@gmail.com

⁴ Department of Mathematics, College of Science and Arts in Gurayat, Jouf University, Gurayat 77454, Saudi Arabia; aralenzi@ju.edu.sa

* Correspondence: anu.sirohi@sharda.ac.in

Abstract: This paper analyzes the time to event data in the presence of collinearity. To address collinearity, the ridge regression estimator was applied in multiple and logistic regression as an alternative to the maximum likelihood estimator (MLE), among others. It has a smaller mean square error (MSE) and is therefore more precise. This paper generalizes the approach to address collinearity in the frailty model, which is a random effect model for the time variable. A simulation study is conducted to evaluate its performance. Furthermore, the proposed method is applied on real life data taken from the largest sample survey of India, i.e., national family health survey (2005–2006) data to evaluate the association of different determinants on infant mortality in India.

Keywords: frailty model; collinearity; maximum likelihood estimation; ridge regression; infant mortality

MSC: 62Nxx



Citation: Albalawi, O.; Sirohi, A.; Rai, P.K.; Alanzi, A.R.A. Parametric Frailty Analysis in Presence of Collinearity: An Application to Assessment of Infant Mortality. *Mathematics* **2022**, *10*, 2255. <https://doi.org/10.3390/math10132255>

Academic Editor: Gaorong Li

Received: 17 May 2022

Accepted: 19 June 2022

Published: 27 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In a survival analysis, the frailty model is a popular approach to model time to event data, which considers both observable and unobservable factors simultaneously [1]. A large number of studies have demonstrated that the consideration of unobservable variability improves the fitness of the model [2–5]. In the frailty model there is an assumption on the explanatory variables that the variables are independent to each other, while various practical situations do not satisfy this assumption. Several studies of infant survival considered factors such as place of residence, birth order, breastfeeding, mother education, tetanus toxoid, antenatal care, iron folic acid tablets, place of delivery, etc. [6–9]. Some of the variables, for example tetanus toxoid and antenatal care, are correlated to each other in the data set, which is analyzed in this study. The fact that parameter estimates become unstable when the explanatory variables are correlated in the case of the frailty model, means that the collinearity problem has seldom been considered.

To overcome the problem of collinearity ridge estimation is the alternative of the maximum likelihood estimation. Ridge estimator has smaller MSE than MLE. In the case of high multicollinearity, there is a considerable reduction in MSE. Ridge regression was originally proposed by Hoerl and Kennard (1970a,b) [10,11], and later on Schafer et al. (1984) [12] generalized this approach to logistic regression. A slightly different approach in the connection of the ridge type estimator is given by Daffy and Santner (1989) [13]. This approach is proposed in a standard linear regression, while Cessie and Houwelingen (1992) [14] generalized the same approach in logistic regression. However, this method is not adapted for frailty models.

This study developed a ridge regression estimator in the parametric frailty model and evaluated its performance through a simulation study. Further, the proposed approach is applied to examine the risk of infant mortality in India.

2. Ridge Estimator for Parametric Frailty Model

In this section, the approach of Cessie and Houwelingen (1992) [14] is extended to the parametric frailty model. The ridge estimator is derived as a restricted maximum likelihood estimator.

Frailty model is a mixture model in the survival analysis, where the risk of the individual death (hazard function) is a function of observable factors and a random effect term (unobserved frailty) [1,15]. The form of such a model is given as

$$\lambda(\tau_i | x_i, \alpha_i) = \alpha_i \lambda_0(\tau_i) \exp(\beta' x_i) \tag{1}$$

where $\lambda(\tau_i | x_i, \alpha_i)$ is the hazard of i -th individual at time $\tau_i, i = 1, 2, \dots, n$ with covariate vector $x'_i = (x_{i1}, x_{i2}, \dots, x_{iq})$ and unobserved frailty variable α_i . $\lambda_0(\tau_i)$ is the base line hazard function and $\beta' = (\beta_1, \beta_2, \dots, \beta_q)$ is the vector of regression coefficients. Here, q represent the number of regressors. It is assumed that the unobserved frailty follow a gamma distribution and is captured by its variance σ^2 [16,17]. This paper also assumes that the unobserved frailty follows gamma distribution. One assumption is also added on the response variable that it follows Weibull distribution with one parameter θ . On the basis of this assumption, $\lambda_0(\tau_i) = \theta \tau_i^{\theta-1}, \theta > 0$. The likelihood function of the frailty model (1) is given by,

$$L(\beta) = \prod_{i=1}^n (\alpha_i \lambda_0(\tau_i) e^{\beta' x_i})^{\delta_i} e^{-\alpha_i \gamma_0(\tau_i) e^{\beta' x_i}} \tag{2}$$

Using the assumption of the distribution (i.e., gamma distribution) of frailty, the random terms are resulted out, making the likelihood function,

$$L(\beta) = \prod_{i=1}^n \left(\frac{\lambda_0(\tau_i) e^{\beta' x_i}}{1 + \sigma^2 \gamma_0(\tau_i) e^{\beta' x_i}} \right)^{\delta_i} (1 + \sigma^2 \gamma_0(\tau_i) e^{\beta' x_i})^{-1/\sigma^2} \tag{3}$$

The log likelihood form of (3) is

$$l(\beta) = \sum_{i=1}^n [\delta_i \log(\lambda_0(\tau_i)) + \beta' x_i - \log(1 + \sigma^2 \gamma_0(\tau_i) e^{\beta' x_i}) - (1/\sigma^2) \log(1 + \sigma^2 \gamma_0(\tau_i) e^{\beta' x_i})] \tag{4}$$

The maximization of (4) gives the ordinary MLE $\hat{\beta}$ for β . Since (4) is non linear in β , m -th, the approximation is obtained by using the Newton–Raphson method, i.e.,

$$\hat{\beta}^{(m)} = \hat{\beta}^{(m-1)} + F(\hat{\beta}^{(m-1)})^{-1} U(\hat{\beta}^{(m-1)}) \tag{5}$$

where $U(\hat{\beta}^{(m-1)})$ is the derivative of $l(\beta)$ in respect to β using $\hat{\beta}^{(m-1)}$,

$$U(\hat{\beta}^{(m-1)}) = \sum_{i=1}^n [\delta_i x_i - \{(1 + \delta_i \sigma^2) \gamma_0(\tau_i) x_i e^{\beta^{(m-1)'} x_i}\} / \{1 + \sigma^2 \gamma_0(\tau_i) e^{\beta^{(m-1)'} x_i}\}] \tag{6}$$

and $F(\hat{\beta}^{(m-1)})$ is the negative of the matrix of the second derivative of $l(\beta)$ in respect to β using $\hat{\beta}^{(m-1)}$, represented as

$$F(\hat{\beta}^{(m-1)}) = X' V X \tag{7}$$

where, V is the $n \times n$ diagonal matrix, whose (i, i) -th element is $v_{ii} = \tau_i^\theta e^{\hat{\beta}^{(m-1)'} x_i} (1 + \delta_i \sigma^2) / (1 + \sigma^2 \tau_i^\theta e^{\hat{\beta}^{(m-1)'} x_i})$. The estimation in the frailty model is based on the iteration and continues until the small changes, which are decided as $\|\beta^{(m)} - \beta^{(m-1)}\| \leq 10^{-5}$. The Duffy and Santner (1989) [13] procedure of the ridge estimation is utilized to reduce the problem

of collinearity. On the basis of this procedure, the log likelihood function is modified with the addition of a penalty on β .

$$l^k(\beta) = l(\beta) - k\|\beta\|^2 \tag{8}$$

where $l^k(\beta)$ and $l(\beta)$ are the restricted and unrestricted log likelihood functions, respectively. Here, $\|\beta\| = (\sum_{j=1}^q \beta_j^2)^{1/2}$ is the norm of the parameter vector β . The maximization of (8) yields $\hat{\beta}^k$. The ridge parameter k controls the amount of shrinkage of the norm of β . When $k = 0$, the solution will remain the same as the ordinary MLE; however, if $k \rightarrow \infty$, then all β 's $\rightarrow 0$.

Collinearity among the covariates induces the unstable parameter estimates. Shrinking the β 's towards zero and allowing a little bias will stabilize the system and give estimates with smaller variance. Therefore, for a good choice of k , the restricted estimate ($\hat{\beta}^k$) is expected to be, on average, closer to true value of β than unrestricted MLE. The rule of selecting k is not fixed but the most common choice is to make a little increment in MSE. In the case of the logistic and Cox regression model, k is chosen as “ $1/\hat{\beta}'\hat{\beta}$, $q/\hat{\beta}'\hat{\beta}$ and $(q + 1)/\hat{\beta}'\hat{\beta}$ ” [12,18,19]. These three choices were investigated for parametric frailty model in our simulation studies.

Such as the unrestricted MLE, $\hat{\beta}^k$ is obtained by the Newton–Raphson method. The m -th order approximation for $\hat{\beta}^k$ is

$$\hat{\beta}^{k(m)} = \hat{\beta}^{k(m-1)} + F^k(\hat{\beta}^{k(m-1)})^{-1}U^k(\hat{\beta}^{k(m-1)}) \tag{9}$$

where $U^k(\hat{\beta}^{k(m-1)})$ is the derivative of the restricted log likelihood function (8) and $F^k(\hat{\beta}^{k(m-1)})$ is the negative of the matrix of the second derivative,

$$U^k(\hat{\beta}^{k(m-1)}) = U(\hat{\beta}^{k(m-1)}) - 2k\hat{\beta}^{k(m-1)}$$

$$F^k(\hat{\beta}^{k(m-1)}) = F(\hat{\beta}^{k(m-1)}) + 2kI$$

where I is a $q \times q$ identity matrix. Under certain regularity conditions, $\hat{\beta}$ is asymptotically unbiased, i.e., the $E(\hat{\beta}) = \beta$ with the covariance matrix $(X'VX)^{-1}$ [14,20]. The asymptotic bias of $\hat{\beta}^k$ becomes

$$E(\hat{\beta}^k - \beta) = -2k(X'VX + 2kI)^{-1}\beta$$

and the asymptotic variance of $\hat{\beta}^k$ is

$$(X'VX + 2kI)^{-1}X'VX(X'VX + 2kI)^{-1}$$

Similarly, asymptotic MSE can be obtained by taking the summation on the asymptotic variance and square of asymptotic bias.

3. Simulation Study

To evaluate the performance of estimators of β , they are compared in terms of their MSE. In this paper, the survival function is taken as

$$S(\tau_i) = \exp(-\gamma_0(\tau_i)\alpha_i e^{\beta'x_i})$$

Then, the cumulative distribution function is

$$F(\tau_i) = 1 - \exp(\gamma_0(\tau_i)\alpha_i e^{\beta'x_i})$$

The cumulative distribution function follows a uniform distribution with a range between 0 and 1. If $M \sim U(0, 1)$, then $1 - M \sim U(0, 1)$. Therefore, the survival function also follows a uniform distribution with the range [0–1].

$$M = \exp(-\gamma_0(\tau_i)\alpha_i e^{\beta'x_i}) \sim U(0, 1)$$

and

$$\tau = \gamma_0^{-1}(-\log(M)/\alpha_i e^{\beta' x_i})$$

In this study, it is assumed that the baseline distribution is the Weibull with one parameter $\theta = 0.5$, then $\gamma_0^{-1}(\cdot) = t^{1/\theta}$; presently, the failure time is

$$\tau = (-\log(M)/\alpha_i e^{\beta' x_i})^{1/\theta} \tag{10}$$

Using the expression (10), failure times are generated and the sample sizes are set as, 50, 200, and 2000. The frailty α_i is generated from the gamma distribution with scale and shape parameters 2 and 0.5, respectively, and M is generated randomly from $U(0, 1)$. This paper considered two covariates in the model and generated by the normal distribution, where the mean is taken as zero and the variance is one. The parameters are taken as 0.25 and 0.691. Pair wise correlation is taken in four categories, namely, small ($\rho = 0.4$), moderate ($\rho = 0.6$), high ($\rho = 0.8$) and very high ($\rho = 0.95$). It is assumed that the censoring times have a uniform distribution with a range from 0 to 4.5, which allows for censoring rates of approximately 39%. A total of 100 data sets were generated under each model and applied the parametric gamma frailty for each data set. The averaged MSE over q covariates for both ML and ridge estimators are computed and represented in Table 1.

Table 1. Comparison of MSE between the MLE and ridge estimators in simulation; $n = 50$.

ρ	Method	MSE (β)	% Reduction of MSE (β)
0.4	MLE	0.0801	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0579	27.72
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0444	44.57
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0356	55.56
0.6	MLE	0.1056	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0658	37.69
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0465	55.97
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0355	66.38
0.8	MLE	0.1871	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0772	55.74
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0457	75.57
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0319	82.95
0.95	MLE	0.6908	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0654	90.55
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0304	95.60
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0206	97.02

The findings of Tables 1–3 demonstrate that the ridge estimators have a smaller MSE than ML. Among the three categories of k , the ridge estimator performed better for $k = (q + 1)/\hat{\beta}'\hat{\beta}$. In the case of moderate correlation ($\rho = 0.6$), if k changes from $1/\hat{\beta}'\hat{\beta}$ to $(q + 1)/\hat{\beta}'\hat{\beta}$ then the percentage reduction will change from 37% to 66%. It is revealed from the Tables 1–3 that if ρ increases then the percentage reduction in MSE also increases. Figure 1 shows the graphical representation of MSE with respect to the deferent values of ρ .

Table 2. Comparison of MSE between the MLE and ridge estimators in simulation; $n = 200$.

ρ	Method	MSE (β)	% Reduction of MSE (β)
0.4	MLE	0.0197	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0181	8.12
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0167	15.23
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0155	21.32
0.6	MLE	0.0259	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0227	12.35
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0202	22
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0181	30.11
0.8	MLE	0.0459	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0350	23.74
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0279	39.22
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0230	49.89
0.95	MLE	0.1693	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.0643	62.02
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.0350	79.33
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.0227	86.59

Table 3. Comparison of MSE between the MLE and ridge estimators in simulation; $n = 2000$.

ρ	Method	MSE (β)	% Reduction of MSE (β)
0.4	MLE	0.00196	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.00194	1.02
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.00193	1.5
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.00192	2.04
0.6	MLE	0.00257	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.00254	1.16
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.00251	2.33
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.00247	3.89
0.8	MLE	0.00458	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.00444	3.05
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.00432	5.60
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.00421	8.07
0.95	MLE	0.01687	
	Ridge estimator for $k = 1/\hat{\beta}'\hat{\beta}$	0.01492	1.15
	Ridge estimator for $k = q/\hat{\beta}'\hat{\beta}$	0.01333	20.98
	Ridge estimator for $k = (q + 1)/\hat{\beta}'\hat{\beta}$	0.01200	40.58

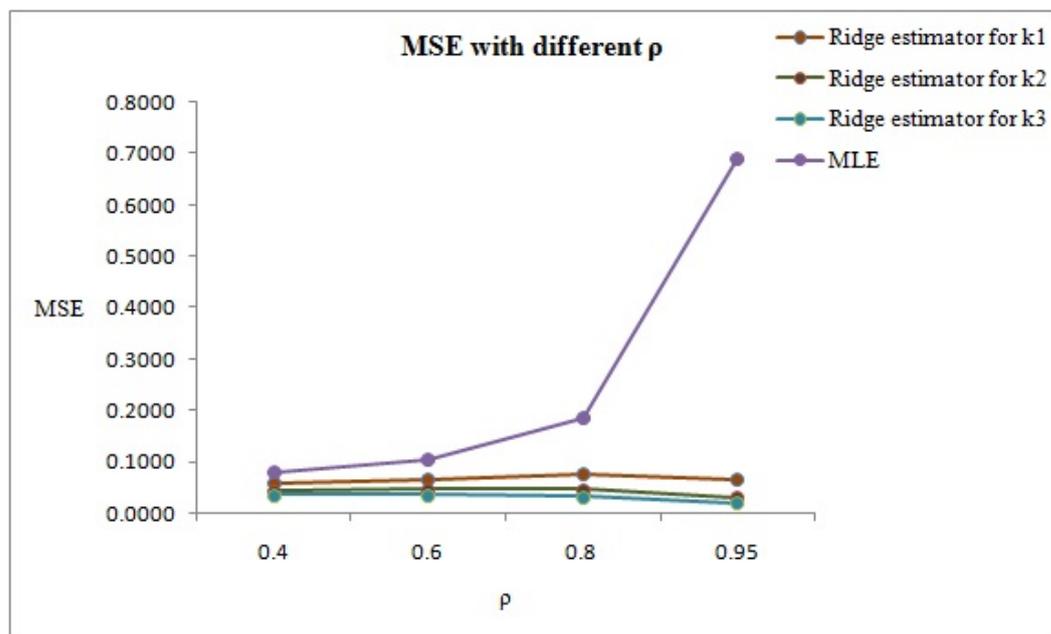


Figure 1. Comparison of MLE and ridge estimators of β in simulation study. Here, k_1 , k_2 and k_3 are ridge parameters such that; $k_1 = 1/\hat{\beta}'\hat{\beta}$, $k_2 = q/\hat{\beta}'\hat{\beta}$ and $k_3 = (q + 1)/\hat{\beta}'\hat{\beta}$.

4. Real Life Application

The proposed ridge regression estimator for the parametric frailty model is applied to a data set from the National Family Health Survey IIIrd Phase (NFHS III). The data set has 11,581 infants, out of which 10,448 infants are right censored. The response variable is the time of survival (in months) from the birth of the infant to the death of the infant. In this paper, the frailty model included 12 covariates of infant mortality, i.e., place of residence, breastfeeding, sex of infant, birth order, place of delivery, marital status of parents, tetanus toxoid (TT), antenatal care (ANC), age of mother, mother education, iron folic acid (IFA) and family size corresponding to each infant. Out of these covariates, TT and ANC are moderately correlated to each other and the correlation value is 0.6.

In some situations, the correlation cannot be considered as the standard measure of collinearity. The condition number is an alternative to measuring the collinearity. This study followed the Ozkale (2021) [21] approach to measure the condition number. On the basis of this approach, the condition number is calculated for the matrix $[(X^*)'(X^*)]$, where $X^* = V^{(1/2)}X$. The value of the condition number is 36.51, which is greater than 10. Therefore, collinearity is present in the data set [22,23].

Table 4 shows the distribution of births and deaths with different covariates in India 2005–2006. About 68% deaths occurred, where infants were not provided with breastfeeding. Approximately 53% infants were delivered at home, and the mothers were not educated in the same proportion. About 46% of births occurred when the size of the family was large, i.e., more than six members in the household. A total of 15%, 18% and 32% of mothers had not taken tetanus toxoid, antenatal care and IFA, respectively, during their pregnancy.

Table 4. Descriptive statistics of NFHS data.

Variables	Birth		Death	
	Number	Percentage	Number	Percentage
Place of residence				
Rural	7348	63.4	776	68.49
Urban	4233	36.6	357	31.51
Breastfeeding				
No	925	8.0	766	67.61
Yes	10,656	92.0	367	32.39
Sex				
Female	5644	48.7	523	46.16
Male	5937	51.3	610	53.84
Birth order				
1	3809	32.9	315	27.80
2	3259	28.1	268	23.65
3	1799	15.5	158	13.95
>=4	2714	23.4	392	34.60
Place of delivery				
Home	6110	52.8	699	61.69
Govt. hospital	2802	24.2	204	18.01
Prvt. Hospital	2648	22.9	224	19.77
Other	21	0.2	6	0.53
Marital status				
Unmarried	11	0.1	1	0.09
0–4 Years	4964	42.9	324	28.60
5–9 Years	3850	33.2	357	31.51
>=10 Years	2756	23.8	451	39.81
ANC				
No	2165	18.7	311	27.45
1–2	2687	23.2	303	26.74
>=3	6729	58.1	519	45.81
Age of mother				
< 18 Years	739	6.4	76	6.71
18–24 Years	5075	43.8	390	34.42
25–29 Years	3453	29.8	316	27.89
30–34 Years	1537	13.3	206	18.18
>=35 Years	777	6.7	145	12.80
Mother education				
No	4512	39	598	52.78
Primary	1687	14.6	196	17.30
Secondary	4436	38.3	309	27.27
Higher	946	8.2	30	2.65
IFA				
No	3708	32.0	484	42.72
Yes	7873	68.0	649	57.28
Family size				
1–3	1071	9.2	260	22.95
4–6	5214	45.0	514	45.37
>=7	5296	45.7	359	31.69

A frailty model is implemented on the data set and the results related to maximum likelihood and ridge estimates are summarized in Table 5. A hazard ratio ($\exp(\hat{\beta})$) less than one indicates a decreased death risk, while a greater than unity indicates an increased risk, and the interpretation of the hazard ratio is either the number of times the risk changes or the percentage change in the death risk.

Table 5. Estimation in frailty model.

Category	MLE		Ridge Estimator	
	$exp(\hat{\beta})$	S.E. ¹ ($\hat{\beta}$)	$exp(\hat{\beta}^k)$	S.E. ¹ ($\hat{\beta}^k$)
Place of residence				
Urban ^a	-	-	-	-
Rural	1.156	0.104	1.162	0.103
Breastfeeding				
No ^a	-	-	-	-
Yes	8.28×10^{-5} *	0.093	8.49×10^{-5} *	0.092
Sex				
Male ^a	-	-	-	-
Female	1.044	0.089	1.047	0.089
Birth order				
1	1.176	0.176	1.209	0.171
2	1.243	0.150	1.265	0.147
3	1.087	0.154	1.099	0.151
>=4 ^a	-	-	-	-
Place of delivery				
Home	1.149	0.134	1.162	0.133
Govt. hospital	0.935	0.145	0.947	0.143
Prvt. Hospital ^a	-	-	-	-
Other	1.350	0.688	1.336	0.614
Marital status				
Unmarried	0.507	1.652	0.674	0.985
0–4 Years ^a	-	-	-	-
5–9 Years	1.546 *	0.137	1.568 *	0.134
>=10 Years	1.939 *	0.192	1.988 *	0.187
TT				
No	1.246	0.153	1.253	0.151
1	1.077	0.162	1.079	0.160
>=2 ^a	-	-	-	-
ANC				
No	0.987	0.161	0.995	0.159
1–2	1.162	0.118	1.169	0.117
>=3 ^a	-	-	-	-
Age of mother				
<18 Years	1.459	0.193	1.461 *	0.191
18–24 Years ^a	-	-	-	-
25–29 Years	1.316 *	0.125	1.317 *	0.123
30–34 Years	1.723 *	0.170	1.717 *	0.167
>=35 Years	1.545 *	0.206	1.543 *	0.201
Mother education				
No	10.592 *	0.232	9.821 *	0.221
Primary	10.430 *	0.242	9.639 *	0.231
Secondary	7.560 *	0.223	7.036 *	0.213
Higher ^a	-	-	-	-
IFA				
No ^a	-	-	-	-
Yes	1.052	0.112	1.063	0.111
Family size				
1–3	2.175 *	0.138	2.175 *	0.137
4–6 ^a	-	-	-	-
>=7	0.715 *	0.099	0.720 *	0.098

^a Reference category; * statistical significance at 5% level of significance; ¹ S.E.: Standard Error.

Some significant results of the ridge regression are given as: Infants who were breast-feeding have a 99% lower death risk as compared to those who were not breastfeeding. Old married couples (more than nine years of marriage) have a 1.98 times higher risk of infant death as compared to newly married couples (0 to 4 years of marriage). If the age

of the mother is more than 34 years, then the risk of infant death is increased by 1.5 times as compared to the infant of the reference category (mothers aged 18 to 24 years). The uneducated mother has a 9.82 times higher risk of infant death as compared to the highly educated mother. The risk of infant death is twice if the size of family is small (1 to 3 members in the household) and this risk is 28% lower for a large family (more than 6 members in the household) as compared to the reference category (4 to 6 members in the household). The estimated value of the σ^2 is significantly non zero ($\sigma^2 = 2.727$), which means the unobserved variables have an impact on the survival times of the infants.

The results related to MSE corresponding to ridge estimators and the MLE are given in Table 6, which shows that the MSE is the lesser for the ridge estimator as compared to the MLE, and the benefits are more when the ridge parameter is $(q + 1) / \hat{\beta}'\hat{\beta}$.

Table 6. Comparison of MSE in NFHS data.

Method	MSE (β)	Percentage Reduction of MSE (β)
MLE	3.819	
Ridge estimator for $k = 1 / \hat{\beta}'\hat{\beta}$	3.553	6.97
Ridge estimator for $k = q / \hat{\beta}'\hat{\beta}$	2.110	44.75
Ridge estimator for $k = (q + 1) / \hat{\beta}'\hat{\beta}$	2.048	46.37

5. Discussion and Conclusions

This paper generalized the approach of ridge regression in the frailty model. The frailty model deals with the time to event data and considered both factors (i.e., measurable and unmeasurable) simultaneously. Here, the parametric form of the frailty model is utilized. A simulation study is conducted and explained that the ridge estimator is more accurate and precise, as compared to MLE. The benefits are higher if the ridge parameter k is taken as $(q + 1) / \hat{\beta}'\hat{\beta}$.

The method is then applied to assess the risk of infant death in India. Data are taken from the third phase of NFHS and revealed that the proposed method produced more precise estimates as compared to MLE. Ridge estimates demonstrate that breastfeeding and mothers' education are highly associated with infant mortality. The findings of this study demonstrates that maternal education is positively associated with infant mortality, i.e., the increment in the level of education reflects the decrement in the infant mortality; these results are significant. There is a high risk of infant death if the mother's age is less than 18 years, and the reason for this finding may be considered as the complications of delivery, premature birth, pregnancy and some other related issues of teenagers. A bigger family size has a positive impact on infant death.

6. Future Scope

This study utilized the ridge as an alternative of MLE in the presence of collinearity, while other alternatives such as lasso, elastic net, etc., are also available in the existence of collinearity. In future, these alternatives can also be applied in the frailty model and can be compared with the existing one.

Author Contributions: Data curation, O.A.; Formal analysis, P.K.R.; Methodology, A.S.; Software, A.R.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Wienke, A. *Frailty Models in Survival Analysis*; Chapman & Hall: London, UK, 2011; pp. 1–320.
2. Beard, R.E. Note on some mathematical mortality models. In *The Lifespan of Animals*; Wolstenholme, G.E.W., Conner, M.O., Eds.; Ciba Foundation Colloquium on Ageing, Little, Brown: Boston, MA, USA, 1959; pp. 302–311.
3. Blossfeld, H.P.; Hamerle, A. Unobserved heterogeneity in event history models. *Qual. Quant.* **1992**, *26*, 157–168. [[CrossRef](#)]

4. Heckman, J.; Hotz, J.; Walker, J. New evidence on the timing and spacing of births. *Am. Econ. Rev. Pap. Proc.* **1985**, *75*, 179–184.
5. Sastry, N. A nested frailty model for survival data, with an application to the study of child survival in Northeast Brazil. *J. Am. Stat. Assoc.* **1997**, *92*, 426–435. [[CrossRef](#)]
6. Magadi, M.A.; Diamond, I.; Rodrigues, R.N. The determinants of delivery care in Kenya. *Soc. Biol.* **2001**, *47*, 164–188. [[CrossRef](#)]
7. Omariba, D.; Beaujot, R.; Rajulton, F. Determinants of infant and child mortality in Kenya: An analysis controlling for frailty effects. *Popul. Res. Policy Rev.* **2007**, *26*, 299–321. [[CrossRef](#)]
8. Pebley, A.R.; Goldman, N.; Rodriguez, G. Prenatal and delivery care and childhood immunization in Guatemala: Do family and community matter? *Demography* **1996**, *33*, 231–247. [[CrossRef](#)]
9. Singh, K.K.; Pandey, N.; Gautam, A. Effect of Breastfeeding and Maternal Health Care programme on infant mortality. *Demogr. India* **2007**, *36*, 253–266.
10. Hoerl, A.E.; Kennard R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
11. Hoerl, A.E.; Kennard, R.W. Ridge regression: Some simulations. *Commun. Stat.* **1970**, *4*, 1105–1123. [[CrossRef](#)]
12. Schafer, R.L.; Roi, L.D.; Wolfe, R.A. A ridge logistic estimator. *Commun. Stat. Theory Methods* **1984**, *13*, 99–113. [[CrossRef](#)]
13. Duffy, D.E.; Santner, T.J. On the small sample properties of norm-restricted maximum likelihood estimators for logistic regression models. *Commun. Stat. Theory Meth.* **1989**, *18*, 959–980. [[CrossRef](#)]
14. Cessie, S.L.; Houwelingen, J.C.V. Ridge estimators in logistic regression. *Appl. Stat.* **1992**, *41*, 191–201. [[CrossRef](#)]
15. Vaupel, J.; Manton, K.; Stallard, E. Impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **1979**, *16*, 439–454. [[CrossRef](#)] [[PubMed](#)]
16. Oakes, D. A model for association in bivariate survival data. *J. R. Stat. Soc. (B)* **1982**, *44*, 414–422. [[CrossRef](#)]
17. Sastry, N. Community characteristics, individual and household attributes, and child survival in Brazil. *Demography* **1996**, *33*, 211–229. [[CrossRef](#)] [[PubMed](#)]
18. Smith, K.R.; Slattery, M.L.; French, T.K. Collinear nutrients and the risk of colon cancer. *J. Clin. Epidemiol.* **1991**, *44*, 715–723. [[CrossRef](#)]
19. Xue, X.; Kim, M.Y.; Roy, E.S. Cox regression analysis in presence of collinearity: an application to assessment of health risks associated with occupational radiation exposure. *Lifetime Data Anal.* **2007**, *12*, 333–350. [[CrossRef](#)] [[PubMed](#)]
20. Cox, D.R.; Hinkley, D.V. *Theoretical Statistics*; Chapman and Hall: London, UK, 1974.
21. Ozkale, M.R. The red indicator and corrected VIFs in generalized linear models. *Commun. Stat. Simul. Comput.* **2021**, *50*, 4144–4170. [[CrossRef](#)]
22. Mackinnon, M.J.; Puterman, A.L. Collinearity in generalized linear models. *Commun. Stat. Theory Methods* **1989**, *18*, 3463–3472. [[CrossRef](#)]
23. Weissfeld, L.A.; Sereika, S.M. A multicollinearity diagnostic for generalized linear models. *Commun. Stat. Theory Methods* **1991**, *20*, 1183–1198. [[CrossRef](#)]