Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2024, 2024:54
https://doi.org/10.28919/cmbn/8503
ISSN: 2052-2541

GENERALIZED LINEAR MODEL WITH BAYES ESTIMATOR: MODELING THE NUMBER OF CHILDREN IN MARRIED COUPLES

DEBRINA FEREZAGIA*
Vocational Program, Universitas Indonesia, Depok, Indonesia

Abstract: The number of children in a married couple is related to the welfare and resilience of the family. Modeling the number of children involves a binary response variable. This research aims to compare GLM modeling with Bayes estimators for response variables that follow the Poisson and Binomial distributions. The research results show that the Bayes parameter estimator in the GLM model with binary response variables in the case of the number of children in Depok City in 2020 is greatly influenced by the prior distribution used. The best model in this case is GLM of binary response variables with a prior distribution following the Cauchy distribution with a scale parameter of 10, compared to other models because it has the smallest AIC value, 273.1. Meanwhile, the Bayes estimator in the GLM model with count data variables (assumed to follow the Poisson distribution), namely the number of children, both with prior distributions following the Normal Distribution and Cauchy Distribution, have almost the same estimator values and nearly the same AIC model values. This research theoretically contributes to the Bayes estimation method, with the result that for binary response variables, the Cauchy prior distribution is more appropriate to use than using the Normal Distribution as a prior distribution in the case of a number of children. Apart from that, in real terms, this research helps know the factors that influence the number of children.

Keywords: Poisson; binomial; Bayes; Cauchy; number of children.
2020 AMS Subject Classification: 92D25.

---

*Corresponding author
E-mail address: debrinaferezagiaa@gmail.com
Received February 03, 2024

1. INTRODUCTION

Married couples generally want children in their family. In ancient times, eastern culture wanted a large number of children. But every married couple has the right to decide the number of children they want. There are many circumstances in which parents may find themselves unable to fulfill their obligations, resulting in their children not receiving proper parental care and protection. Such situations can lead to a variety of stress-inducing problems, including poverty, health problems, domestic or community violence, stigma, emergencies, or drug abuse [1]–[3].

Currently, in several developed countries, married couples decide to have few children or be childfree. Some of the reasons why married couples choose to have a small number of children are maintaining the quality of the marriage, obtaining quality offspring, giving children the right to a "standard of living" that is adequate for the child's physical, mental, spiritual and social development, reducing the risk of defects in children [4] –[7]. Also, married couples who do not have children will find it easy to mobilize geographically to the desired area, reducing the stress of raising children and maintaining mental health [8]–[10].

The results of the Statistics Indonesia, population projections show that 30.1 percent or 79.55 million of Indonesia's population are children aged 0-17. This means that one in three Indonesian residents are children. In the future, it is projected that the number of children in Indonesia will not experience significant changes [11]. Indonesian family life is vulnerable to economic pressure, stress symptoms, food security, and psychological well-being [12]. Thus, the socio-economic conditions of parents, welfare, and family resilience can influence the child's condition [13], [14].

The problem faced when analyzing the number of children is that the variable for the number of children does not follow the Normal Distribution because the Normal Distribution has a range of all real numbers, $y \in (-\infty\infty)$. Meanwhile, the number of children is an integer greater than zero, $y \in (0, 1, 2, \dots, \infty)$. So, linear regression analysis will be hampered due to violations of the normality assumption [15]–[17]. For count data, you can use the Generalized Linear Model (GLM) [18], [19]. The analysis that can be carried out to deal with this problem assumes that the number of children follows a Poisson distribution because the number of children is in the form of count data. So, the analysis used is the Poisson Generalized Linear Model (GLM) [20], [21]. In addition, the number of children can be categorized into two categories so that the number of children will follow the Binomial Distribution [22], [23].

In parameter estimation, GLM can use Maximum Likelihood (ML), Penalized Quasi-Likelihood (PQL), or Bayes Estimator. Bayes estimator has the smallest Root Mean Squared Error of Prediction value compared to other estimators, slight bias and variance, and can be used in liner models with inequality constraints [24], [25].

The research aims to compare GLM Poisson modeling with GLM Binomial in the case of data on the number of children in Depok City, Indonesia. The city of Depok was chosen because Depok is a satellite city of the Indonesian capital, which has quite a large population, 2,056,335 people, with an area of 200.29 km^2. The Bayes estimator is used in this analysis as an alternative to the commonly used maximum likelihood estimator.

2. Materials and Methods

The data used is the 2020 National Socio-Economic Survey data in Depok City, with the variable of interest being the number of children. The total sample was 987 households in Depok City; 580 households were taken where one of the household members had given birth. The independent variables used in this research are [4], [5], [26], [27].

Table 1. Independent Variables Used

	Independent Variables	Data type
	(1)	*(2)*
x_1	mother's age	Numerical
x_2	age of mother's first marriage	Numerical
x_3	mother's educational level	Categorical
x_4	mother's working hours	Numerical
x_5	there are household members who travel long distances	Categorical

	Independent Variables (1)	Data type (2)
x_6	there are household members who are victims of crime	Categorical
x_7	the household has JKN/Jamkesda	Categorical
x_8	age at the time of the mother's first pregnancy	Numerical
x_9	age at the mother's first birth	Numerical
x_{10}	Mother has used or is currently using birth control	Categorical
x_{11}	experiencing economic worries	Categorical
x_{12}	many families live together at home	Numerical
x_{13}	status of residence	Categorical
x_{14}	living area	Numerical
x_{15}	possession of LPG 5.5 kg or more	Categorical
x_{16}	ownership of a computer/laptop	Categorical
x_{17}	gold ownership of 10g or more	Categorical
x_{18}	car ownership	Categorical
x_{19}	possession of a prosperous family card	Categorical
x_{20}	recipients of the Family Hope Program (PKH)	Categorical
x_{21}	food aid recipients	Categorical
x_{22}	recipients of social assistance/local government subsidies	Categorical
x_{23}	many household members live together	Numerical
x_{24}	expenditure on food, drinks, and cigarettes for a week	Numerical
x_{25}	calorie consumption	Numerical
x_{26}	consume protein	Numerical
x_{27}	fat consumption	Numerical
x_{28}	consume carbohydrates	Numerical

The parameter estimation method used for the analysis is the Bayes estimator in the Generalized Linear Model (GLM).

Generalized Linear Model (GLM)

GLM is a development of the classical linear model where the random variable Y is an independent component with a mean value (μ_i). There are three main components in GLM [19] :

a) The random component, namely the component Y_i that, is free and spreads with expected value $E(Y_i) = \mu_i$ and variety $Var(Y_i) = \sigma_i^2$
b) The systematic component, namely X_i the $i = 1, 2, 3 \dots, p$ one that produces the linear estimator η, where $\eta = X\beta$
c) link function, $\eta = g(\mu)$

In the classical linear model, component (i) is normally distributed, and component (iii) is an identity function. Meanwhile, in GLM, component (i) may come from one of the other members of the exponential distribution family, and component (iii) is another monotone function [19]. Thus, GLM can be modeled by:

$$g\big(E(Y_i|x)\big) = g(\mu) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p = \eta(x)$$

Where the variance Y_i is a function of the mean value Y_i, $\mathrm{Var}(Y_i) = \phi\ \mathrm{Var}(\mu_i)$.

Bayesian estimator

According to Gelman et al. (2004), the posterior density function in Bayes' rule is proportional to the product of the prior distribution and the data distribution. The following is Bayes' rule for forming the posterior distribution of data:

$$p(\theta|y,x) = \frac{p(\theta,y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)}$$

The approximate form for the equation above is obtained by removing $p(y)$ since it does not depend on θ; it is, therefore p(y), considered constant. The equation form becomes

$$p(\theta|y,x) \propto p(\theta)p(y|\theta)$$

which is the equation: Posterior ∝ Prior. Likelihood

Akaike Information Criterion

The selection of the best model found by Akaike and Schwarz is based on the maximum likelihood estimation (MLE) method. According to the AIC method, the best model is the model that has the smallest AIC value [28].

3. Results and Discussion

Data exploration is carried out before modeling is carried out; first, the Spearman correlation value will be calculated. The correlation between independent variables is shown in Figure 1, and it can be seen that only a few independent variables influence the response variable (number of children). The independent variables are x_{10} (using birth control), x_{13} (residence status), x_{14} (residence size), x_{16} (ownership of a computer/laptop), x_{23} (many household members living together), x_{24} (expenditure on food, drinks and cigarettes during the week), x_{25} (calorie consumption), x_{26} (protein consumption), x_{27} (fat consumption), x_{28} (carbohydrate consumption).

	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	x20	x21	x22	x23	x24	x25	x26	x27	x28
y	1.00																												
x1	0.08	1.00																											
x2	-0.04	0.00	1.00																										
x3	0.07	-0.14	0.43	1.00																									
x4	-0.01	0.16	0.05	0.04	1.00																								
x5	-0.04	-0.10	0.12	0.13	0.05	1.00																							
x6	0.09	0.01	-0.04	0.06	0.06	0.03	1.00																						
x7	0.04	0.08	0.07	0.05	-0.03	0.03	-0.01	1.00																					
x8	-0.07	0.02	0.94	0.39	0.06	0.11	-0.04	0.04	1.00																				
x9	-0.08	0.04	0.93	0.38	0.06	0.12	-0.04	0.04	0.99	1.00																			
x10	0.23	-0.14	-0.03	0.01	0.02	0.00	0.03	0.03	-0.04	-0.04	1.00																		
x11	0.00	0.02	-0.10	-0.11	0.10	-0.07	-0.03	0.03	-0.12	-0.11	0.05	1.00																	
x12	0.03	0.11	-0.08	-0.05	-0.03	0.01	-0.01	0.04	-0.05	-0.06	-0.03	-0.02	1.00																
x13	-0.15	-0.24	-0.01	-0.11	-0.03	-0.07	0.01	-0.09	-0.01	-0.03	0.02	0.01	-0.06	1.00															
x14	0.19	0.32	0.11	0.28	0.00	0.05	0.06	0.10	0.09	0.10	-0.06	-0.11	0.17	-0.52	1.00														
x15	0.07	0.19	0.22	0.31	0.09	0.19	0.06	0.04	0.21	0.22	-0.08	-0.07	0.05	-0.15	0.39	1.00													
x16	0.15	0.16	0.24	0.41	0.04	0.15	-0.01	0.07	0.23	0.23	-0.10	-0.18	0.04	-0.24	0.43	0.46	1.00												
x17	-0.01	0.05	0.14	0.32	-0.01	0.20	-0.01	0.03	0.14	0.15	-0.17	-0.20	-0.01	-0.20	0.36	0.34	0.40	1.00											
x18	0.10	0.13	0.22	0.36	0.10	0.17	0.02	0.02	0.21	0.21	-0.06	-0.11	0.04	-0.26	0.48	0.45	0.42	0.43	1.00										
x19	0.06	-0.04	0.06	0.22	-0.03	0.19	-0.01	0.01	0.06	0.06	-0.06	-0.16	-0.01	-0.20	0.27	0.25	0.40	0.40	0.31	1.00									
x20	0.14	0.09	-0.12	-0.18	0.07	-0.04	-0.04	0.09	-0.11	-0.12	0.06	0.18	0.09	0.05	-0.10	-0.12	-0.18	-0.13	-0.13	-0.17	1.00								
x21	0.11	0.13	-0.08	-0.16	0.06	-0.12	-0.04	0.10	-0.07	-0.08	0.05	0.07	0.04	-0.02	-0.06	-0.11	-0.16	-0.12	-0.12	-0.16	0.65	1.00							
x22	0.16	0.06	-0.05	-0.09	0.06	-0.04	-0.03	0.08	-0.07	-0.06	0.05	0.02	0.01	0.00	-0.09	-0.07	-0.11	-0.09	-0.08	-0.15	0.35	0.31	1.00						
x23	0.90	0.08	-0.06	0.05	0.01	-0.02	0.09	0.05	-0.09	-0.10	0.19	0.00	0.30	-0.15	0.25	0.09	0.13	0.01	0.12	0.04	0.15	0.11	0.14	1.00					
x24	0.25	0.19	0.08	0.30	0.06	0.15	0.08	0.12	0.05	0.05	-0.03	-0.06	0.14	-0.22	0.47	0.38	0.44	0.34	0.44	0.40	-0.08	-0.09	-0.10	0.30	1.00				
x25	0.58	0.16	-0.08	0.05	0.01	0.07	0.11	0.10	-0.10	-0.10	0.14	0.00	0.21	-0.20	0.29	0.15	0.16	0.08	0.18	0.16	0.10	0.06	0.07	0.67	0.59	1.00			
x26	0.53	0.16	-0.04	0.09	0.00	0.07	0.13	0.12	-0.05	-0.04	0.09	0.01	0.23	-0.20	0.32	0.20	0.22	0.15	0.24	0.21	0.08	0.04	0.05	0.60	0.68	0.90	1.00		
x27	0.48	0.10	-0.05	0.10	-0.02	0.09	0.09	0.07	-0.06	-0.07	0.05	-0.07	0.18	-0.19	0.29	0.16	0.20	0.09	0.23	0.20	-0.01	0.00	0.04	0.55	0.61	0.87	0.80	1.00	
x28	0.57	0.20	-0.11	-0.01	0.03	0.03	0.11	0.09	-0.12	-0.12	0.17	0.04	0.18	-0.15	0.23	0.13	0.12	0.05	0.11	0.08	0.13	0.09	0.08	0.65	0.49	0.92	0.82	0.67	1.00

Figure 1. Spearman correlation between variables

The model proposed in this research is a GLM with the distribution of response variables following the Poisson distribution, with the link function being the log function (from now on referred to as GLM Poisson). Another model is GLM, with the distribution of response variables following the Binomial Distribution with a link function, namely the logit function (from now on referred to as Binomial GLM). In both models, it will be assumed that the prior distribution follows the normal distribution and the Cauchy distribution. The Cauchy distribution was chosen as the prior distribution because the Cauchy distribution was indicated to have superior capabilities compared to the normal distribution in logistic regression [29].

From Figure 2, you can see the Normal and Cauchy distributions, which will be used as the prior distribution in Bayes estimation.

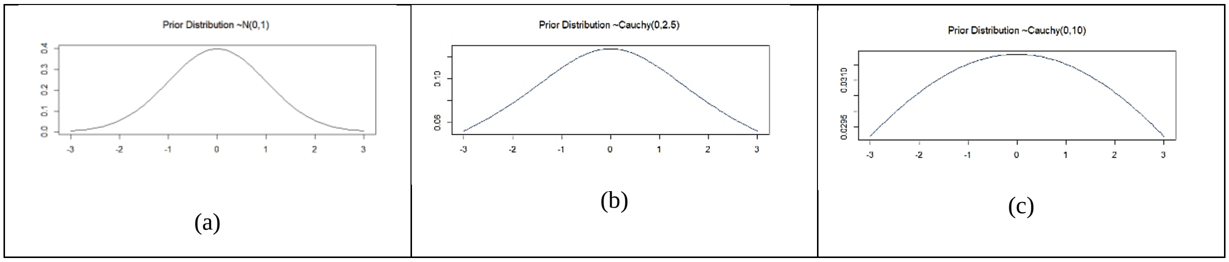

Figure 2. a) Standard Normal Distribution (0.1); b) Cauchy distribution (0, 2.5); and c) Cauchy distribution (0.10)

The proposed models can be seen in Table 2. The Poisson GLM uses two prior distributions, namely the Normal Distribution (0, 1) and the Cauchy Distribution (0, 2.5). Meanwhile, Binomial GLM uses three prior distributions, namely Normal Distribution (0, 1), Cauchy Distribution (0, 2.5), and Cauchy Distribution (0, 10).

Table 2. Some Proposed Modeling

No	Model	Information
(1)	*(2)*	*(3)*
1. a	$\log(\lambda_i) = \eta = X\beta$	y ~ Poisson distribution, with Normal prior (0, 1)
1. b	$\log(\lambda_i) = \eta = X\beta$	y ~ Poisson distribution, with Cauchy prior (centered at 0, and scale parameter 2.5)
2. a	$logit(\pi_i) = \eta = X\beta$	y ~ Binomial distribution, with Normal prior (0, 1)
2. b	$logit(\pi_i) = \eta = X\beta$	y ~ Binomial distribution, with Cauchy prior (centered at 0, and scale parameter 2.5)
2. c	$logit(\pi_i) = \eta = X\beta$	y ~ Binomial distribution, with Cauchy prior (centered at 0, and scale parameter 10)

Parameter estimation was carried out for the two models in Table 2 using the Bayes estimator, where the Poisson GLM was the first model and the Binomial GLM was the second model. A comparison of the GLM Poisson model was carried out with the Prior Distribution of parameters assumed to follow the Normal Distribution (Model 1. a) and when the Prior Distribution was assumed to follow the Cauchy Distribution (Model 1. b)

The second model is a Binomial GLM, and several comparisons are made of prior distribution assumptions. Model 2. A assumes a Binomial GLM with the distribution of prior parameters following the Normal Distribution (0, 1); Model 2. b assumes a Binomial GLM with a distribution of prior parameters following the Cauchy Distribution (0, 2.5); and Model 2. c assumes a Binomial GLM with a prior distribution that also follows the Cauchy Distribution but with a scale parameter of 10, Cauchy Distribution (0, 10).

Poisson GLM

In Table 3, you can see the parameter estimates for the model coefficients using the Bayes estimator. The independent variables that have a significant effect (absolute level $\alpha = 0.05$) on the number of children are the variable x_{12} (the number of families living together at home), which has a negative effect, and x_{23} (the number of household members living together,) which has a positive impact. Model 1 GLM Poisson uses a log link function, so it must be exponential first to interpret it. So, it can be interpreted that every additional family in the household (x_{12}) will reduce the average number of children 0.54 times from both Model 1. a and Model 1. b, assuming other variables are constant. Meanwhile, adding one person to the household will increase the average number of children by 1.47 times.

Table 3. Bayes Estimator for GLM Poisson Parameters (Model 1)

	Model 1. a Prior Normal (0, 1)		Model 1. b Prior Cauchy(0, 2.5)	
Parameter Est.	Estimate	Std. Error	Estimate	Std. Error
(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	-0.5461	0.3358	-0.5524	0.3374
x1	0.0004	0.0048	0.0005	0.0048
x2	0.0047	0.0209	0.0045	0.0212
x34	0.1549	0.3930	0.1763	0.4130
x35	-0.0711	0.6999	-0.1064	0.8817
x37	0.0832	0.1073	0.0876	0.1088
x38	-0.0694	0.5767	-0.0852	0.6648
x311	0.0564	0.0955	0.0605	0.0972
x312	0.1400	0.3824	0.1589	0.4019
x313	0.0717	0.1199	0.0758	0.1217
x315	0.1304	0.2256	0.1387	0.2301
x316	0.1790	0.1674	0.1867	0.1700
x318	0.1117	0.1258	0.1165	0.1279
x320	-0.0251	0.2240	-0.0200	0.2290
x4	-0.0010	0.0013	-0.0010	0.0013
x51	-0.0221	0.0647	-0.0220	0.0649
x61	0.0387	0.1654	0.0380	0.1672
x71	0.0344	0.0769	0.0342	0.0771
x8	0.0015	0.0458	0.0019	0.0508
x9	-0.0050	0.0439	-0.0052	0.0486
x101	0.0729	0.0700	0.0724	0.0701
x111	-0.0018	0.0996	-0.0024	0.1001
x12	-0.6061	0.1180	-0.6094	0.1185
x132	-0.0544	0.0819	-0.0547	0.0823
x133	-0.0518	0.1203	-0.0530	0.1211
x134	-0.0066	0.2494	-0.0091	0.2560
x14	-0.0005	0.0007	-0.0005	0.0007
x151	-0.0394	0.0927	-0.0398	0.0931
x161	0.0666	0.0808	0.0662	0.0811

	Model 1. a Prior Normal (0, 1)		Model 1. b Prior Cauchy(0, 2.5)	
Parameter Est.	Estimate	Std. Error	Estimate	Std. Error
(1)	*(2)*	*(3)*	*(4)*	*(5)*
x171	-0.0334	0.0794	-0.0340	0.0797
x181	0.0148	0.0912	0.0142	0.0916
x191	0.0491	0.0747	0.0504	0.0750
x201	0.1137	0.1668	0.1176	0.1693
x211	-0.0590	0.1727	-0.0628	0.1752
x221	0.0068	0.1925	0.0048	0.1953
x23	0.3874	0.0369	0.3897	0.0371
x24	0.0000	0.0000	0.0000	0.0000
x25	0.0000	0.0000	0.0000	0.0000
x26	0.0000	0.0001	0.0000	0.0001
x27	0.0000	0.0001	0.0001	0.0001
x28	0.0000	0.0000	0.0000	0.0000

The estimated parameter values in Table 3 show that the resulting estimated parameter values are not much different for both Model 1. a and Model 1. b. The AIC values for Model 1. a and Model 1. b are 1577.7 and 1577.6, respectively, with very little difference, which can be seen in Table 4.

Table 4. Comparison of GLM Poisson in Model 1. a and Model 1. b

Model	AIC	Information
(1)	*(2)*	*(3)*
1. a	1577.7	Poisson GLM with Normal Prior (0, 1)
1.b	1577.6	Poisson GLM with Cauchy Prior (0, 2.5)

It can be concluded that the Poisson GLM using a prior distribution following the Normal Distribution and the Cauchy Distribution produces almost exactly good modeling, as can be seen from the AIC values, which are nearly the same as the residual plot in Figure 3.

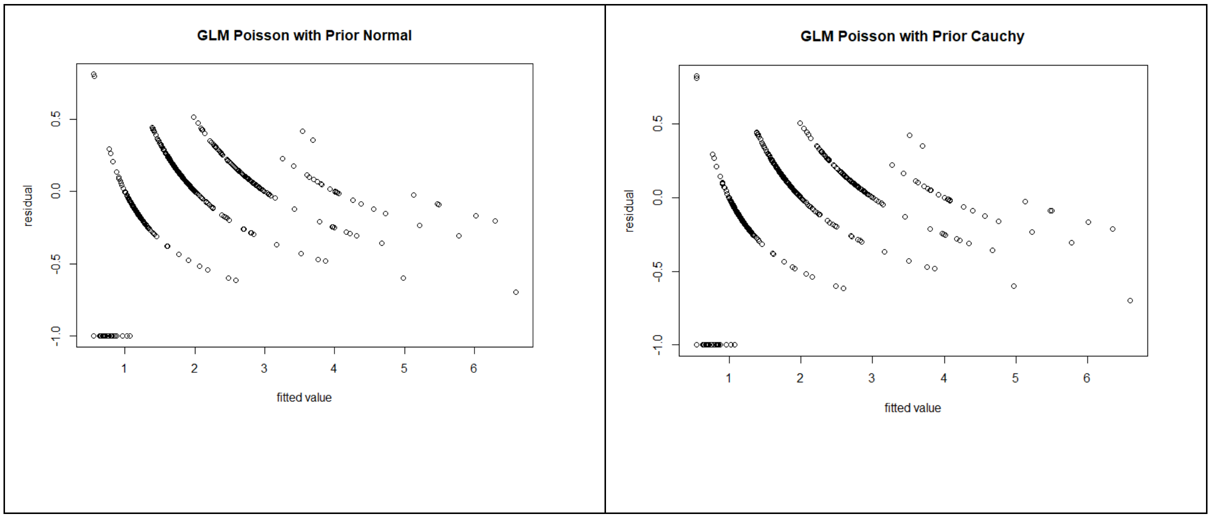

Figure 3. GLM Poisson residuals

Binomial GLM

In Binomial GLM modeling (Model 2), the number of children per household is dichotomized where the number of children is less than equal to two, $y \in (0, 1, 2)$ will be categorized as "1" (according to recommendations from the Government/National Population and Family Planning Agency (BKKBN)) and category "0" for households with more than three children, $y \in (3, 4, 5, \ldots, \infty)$.

Parameter estimation was carried out using a Bayes estimator with a prior distribution assumed to have a Normal Distribution (0.1), a Cauchy Distribution (0, 2.5), and a Cauchy Distribution (0, 10). The results of parameter estimation using the Bayes estimator can be seen in Table 5. The estimator values from Model 2. a, Model 2. b, and Model 2. c and the significant variables from the three models are quite different from each other. From this, it can be concluded that determining the prior distribution in the Bayes estimator greatly influences the estimation results.

Table 5. Bayes Estimator for Binomial GLM Parameters (Model 2)

Parameter Est.	Model 2. a Prior Normal (0,1)		Model 2. b Cauchy Priors (0, 2.5)		Model 2. c Cauchy Priors (0, 10)	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Intercept	9.9640	1.6490	14.6400	2,3700	15.9200	2.6080
x1	-0.0258	0.0226	-0.0371	0.0307	-0.0471	0.0344
x2	-0.0872	0.0825	-0.1984	0.1577	-0.3647	0.2448
x34	0.1453	0.9392	0.2140	2.1120	0.7338	6.3880
x35	0.0041	0.9979	0.0013	2.4960	0.0067	9.8960
x37	-0.0546	0.4255	-0.3984	0.5939	-0.6620	0.6929
x38	0.0158	0.9923	0.0076	2,4800	0.0490	9.4840
x311	0.1552	0.3760	-0.0497	0.5256	-0.2743	0.6200
x312	-0.2058	0.8751	-0.6584	1.6140	-1.4650	2.5380
x313	-0.1314	0.4674	-0.4041	0.6676	-0.6577	0.7887
x315	-0.1361	0.6859	-0.2926	1.0230	-0.5930	1.2740
x316	-0.5520	0.6037	-1.1200	0.9293	-1.7050	1.1270
x318	-0.0856	0.4704	-0.1643	0.6637	-0.3946	0.7852
x320	-0.4326	0.7248	-1.2520	1.1270	-2.0260	1.3580
x4	0.0036	0.0059	0.0052	0.0078	0.0051	0.0085
x51	0.3748	0.3030	0.4680	0.4077	0.5760	0.4469
x61	-0.3463	0.6668	-0.4179	1.0200	-0.5797	1.2620
x71	0.2730	0.3468	0.3154	0.4576	0.3763	0.4988
x8	0.0061	0.0916	0.0018	0.1849	-0.0781	0.3855
x9	0.0701	0.0888	0.1756	0.1797	0.4152	0.3453
x101	-0.3308	0.3295	-0.4918	0.4354	-0.6297	0.4721
x111	0.1439	0.4310	0.1944	0.5821	0.2444	0.6513
x12	4.2590	0.6010	7.1650	0.9365	7.6610	1.0240
x132	-0.1806	0.3701	-0.1094	0.4952	-0.0345	0.5419
x133	-0.3552	0.5143	-0.4623	0.7168	-0.3768	0.8179
x134	-0.1551	0.7386	-0.7853	1.0950	-0.9561	1.3620
x14	0.0008	0.0030	0.0035	0.0039	0.0043	0.0042
x151	-0.0609	0.4020	-0.0481	0.5528	0.0093	0.6171

Parameter Est.	Model 2. a Prior Normal (0,1)		Model 2. b Cauchy Priors (0, 2.5)		Model 2. c Cauchy Priors (0, 10)	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
x161	-0.5855	0.3610	-0.8976	0.4908	-1.0260	0.5377
x171	0.1226	0.3665	0.0274	0.4939	-0.0159	0.5383
x181	-0.2407	0.4028	-0.1315	0.5540	-0.0778	0.6147
x191	-0.1142	0.3458	-0.1787	0.4627	-0.2722	0.5054
x201	-0.6347	0.6088	-1.4190	0.9160	-1.9810	1.0750
x211	0.4275	0.6385	1.0900	1.0100	1.9310	1.3190
x221	-0.8412	0.6780	-1.2300	1.1300	-1.9240	1.6220
x23	-2.6320	0.2219	-4.2690	0.4096	-4.5440	0.4594
x24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
x25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
x26	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004
x27	-0.0003	0.0002	-0.0005	0.0003	-0.0008	0.0004
x28	0.0000	0.0001	0.0000	0.0001	-0.0001	0.0002

From Table 6, it can be concluded that Model 2. c, namely Binomial GLM with a prior distribution following the Cauchy distribution (0, 10), is the best model because it has the smallest AIC value compared to Model 2. a and Model 2. b.

The best Binomial GLM model (Model 2. c) can be written as follows:

$$logit(\hat{\pi}_i) = 15.92 - 0.0471x_1 - 0.3647x_2 + \cdots - 0.0008x_{27} - 0.0001x_{28}$$

$$\log\left(\frac{\hat{\pi}_i}{1-\hat{\pi}_i}\right) = 15.92 - 0.0471x_1 - 0.3647x_2 + \cdots - 0.0008x_{27} - 0.0001x_{28},$$

With significant variables being x_{12} (many families living together at home), x_{16} category 1 (having a computer/laptop at home), x_{20} category 1 (receiving assistance from the family hope program/PKH), x_{23} (many House Members Living together) and x_{27} (fat consumption). It should be remembered that category "1" in the response variable has a number of children less than or equal to two. So, category "0" indicates households that have many children (more than two children).

From the significant variables in Model 2. c, it can be seen that the coefficient for many families is positive (x_{12}). Meanwhile, other important variables such as computer/laptop ownership, PKH assistance recipients, number of household members, and fat consumption have negative coefficients. The model is in the form of a logit $\hat{\pi}_i$so. It needs to be exponentiated first, and the interpretation is in the form of an odds ratio value in the form of a trend. In other words, increasing the number of families in a household will make that household more likely to have less than or equal to two children. Meanwhile, owning a computer/laptop, receiving PKH assistance, increasing the number of household members, and increasing fat consumption will make households tend to have more than two children. With the Binomial GLM residual plot in Figure 4.

Table 6. Comparison of Binomial GLM in Model 2. a, Model 2. b and Model 2. c

Model	AIC	Information
(1)	*(2)*	*(3)*
2. a	306.3	Binomial GLM with Normal Prior (0, 1)
2. b	276.6	Binomial GLM with Cauchy Prior (0, 2.5)
2. c	273.1	Binomial GLM with Cauchy Prior (0, 10)

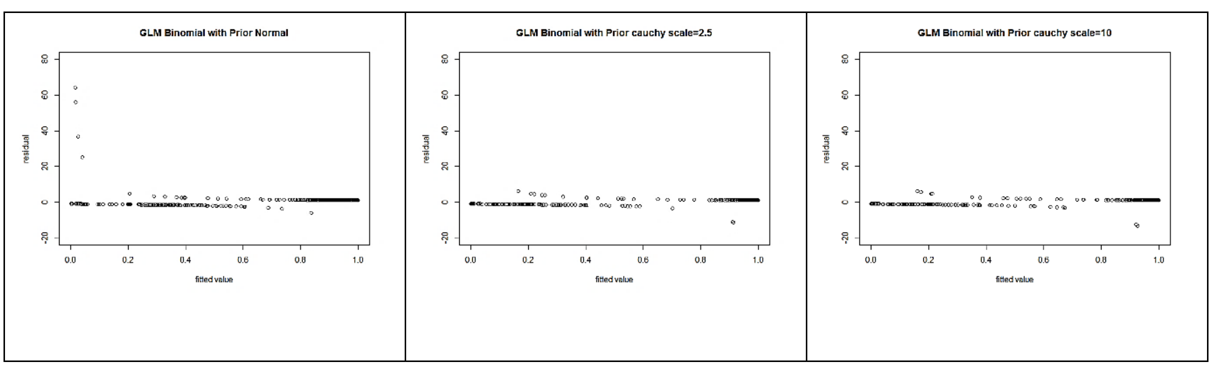

Figure 4. Binomial GLM residuals

4. CONCLUSIONS

Bayes estimator can be used as an alternative in estimating parameters; in GLM modeling, the best model is Binomial GLM (binary response variable) with Prior following the Cauchy Distribution with a scale parameter of 10. This can be seen from the lowest AIC value, namely 273.1. The choice of prior distribution in the Bayes estimator with Binomial GLM greatly determines the resulting parameter estimator.

In contrast to Poisson GLM (counted data response variables), the use of prior distributions, both Normal Distribution and Cauchy Distribution, produces almost the exact estimators. Nearly the same in terms of the estimated values of the model coefficient parameters, significant independent variables, and the AIC value of the model.

The independent variables that are significant for the number of children in the Binomial GLM model are the number of families living together at home, ownership of computers/laptops at home, recipients of the Family Hope Program/PKH assistance), many household members living together, and household fat consumption.

ACKNOWLEDGMENTS

The author would like to thank the research and community service unit of the UI Vocational Education Program, for their support for the funding.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] P.B. Rothenberg, P.E. Varga, The relationship between age of mother and child health and development., Amer. J. Public Health 71 (1981), 810-817. https://doi.org/10.2105/ajph.71.8.810.

[2] M. Ungar, Overprotective parenting: helping parents provide children the right amount of risk and responsibility, Amer. J. Family Therapy. 37 (2009), 258-271. https://doi.org/10.1080/01926180802534247.

[3] W. Ghaffar, M. Manby, T. Race, Exploring the experiences of parents and carers whose children have been

subject to child protection plans, Br. J. Soc. Work 42 (2011), 887-905. https://doi.org/10.1093/bjsw/bcr132.

[4] J. Blake, Family size and the quality of children, Demography. 18 (1981), 421-442. https://doi.org/10.2307/2060941.

[5] N. Petrowski, C. Cappa, P. Gross, Estimating the number of children in formal alternative care: Challenges and results, Child Abuse Neglect. 70 (2017), 388-398. https://doi.org/10.1016/j.chiabu.2016.11.026.

[6] N.D. Glenn, S. McLanahan, Children and marital happiness: A further specification of the relationship, J. Marriage Family. 44 (1982), 63. https://doi.org/10.2307/351263.

[7] F.N. Ilhomovna, Responsibility of parents before the offspring, in: International Conference on Advance Research in Humanities, Applied Sciences and Education, pp. 441-446, 2022.

[8] L.H. Long, The influence of number and ages of children on residential mobility, Demography. 9 (1972), 371-382. https://doi.org/10.2307/2060860.

[9] E.S. Jami, A.R. Hammerschlag, M. Bartels, et al. Parental characteristics and offspring mental health and related outcomes: a systematic review of genetically informative literature, Transl. Psych. 11 (2021), 197. https://doi.org/10.1038/s41398-021-01300-2.

[10] S. Marsh, R. Dobson, R. Maddison, The relationship between household chaos and child, parent, and family outcomes: a systematic scoping review, BMC Public Health. 20 (2020), 513. https://doi.org/10.1186/s12889-020-08587-8.

[11] Statistics Indonesia, 2020 national socio-economic survey, Jakarta, Indonesia, 2020.

[12] E. Sunarti, A.F. Fithriyah, N. Khoiriyah, et al. Portrait of Indonesian family during one year the COVID-19 pandemic: Analysis of factors influencing family welfare and resilience, J. Disaster Res. 17 (2022), 31-42. https://doi.org/10.20965/jdr.2022.p0031.

[13] A. Kalil, Family resilience and good child outcomes: a review of the literature, Centre for Social Development, Ministry of Social Development, Wellington, 2003.

[14] M.I. Armstrong, S. Birnie-Lefcovitch, M.T. Ungar, Pathways between social support, family well being, quality of parenting, and child resilience: what we know, J. Child. Fam. Stud. 14 (2005), 269-281. https://doi.org/10.1007/s10826-005-5054-4.

[15] G.K. Uyanık, N. Güler, A study on multiple linear regression analysis, Procedia - Soc. Behav. Sci. 106 (2013), 234-240. https://doi.org/10.1016/j.sbspro.2013.12.027.

[16] A.F. Schmidt, C. Finan, Linear regression and the normality assumption, J. Clinic. Epidemiol. 98 (2018), 146-151. https://doi.org/10.1016/j.jclinepi.2017.12.006.

[17] G.E.P. Box, G.S. Watson, Corrigenda, Biometrika. 52 (1965), 669-669. https://doi.org/10.1093/biomet/52.3-

4.669-a.

[18] M.A.U. Anietie, E.R. Surajudeen, T.B.L. Kemi, An empirical study of generalized linear model for count data, J. Appl. Comput. Math. 04 (2015), 1000253. https://doi.org/10.4172/2168-9679.1000253.

[19] P. McCullagh, J.A. Nelder, Generalized linear models, Springer, Boston, MA, 1989. https://doi.org/10.1007/978-1-4899-3242-6.

[20] G.M. Abdella, J. Kim, K.N. Al-Khalifa, et al. Penalized Conway-Maxwell-Poisson regression for modelling dispersed discrete data: The case study of motor vehicle crash frequency, Safe. Sci. 120 (2019), 157-163. https://doi.org/10.1016/j.ssci.2019.06.036.

[21] J.W. Hardin, J.M. Hilbe, Generalized linear models and extensions, Stata Press, College Station, TX, 2001.

[22] R.S. de Souza, E. Cameron, M. Killedar, et al. The overlooked potential of generalized linear models in astronomy, I: Binomial regression, Astron. Comput. 12 (2015), 21-32. https://doi.org/10.1016/j.ascom.2015.04.002.

[23] S. Paul, K.K. Saha, The generalized linear model and extensions: a review and some biological and environmental applications, Environmetrics. 18 (2007), 421-443. https://doi.org/10.1002/env.849.

[24] E. Sunandi, K.A. Notodiputro, B. Sartono, A study of generalized linear mixed model for count data using hierarchical Bayes method, Medstat. 14 (2021), 194-205. https://doi.org/10.14710/medstat.14.2.194-205.

[25] R. Ghosal, S.K. Ghosh, Bayesian inference for generalized linear model with linear inequality constraints, Comput. Stat. Data Anal. 166 (2022), 107335. https://doi.org/10.1016/j.csda.2021.107335.

[26] A.R.E. Ghannam, An examination of factors affecting fertility rate differentials as compared among women in less and more developed countries, J. Human Ecol. 18 (2005), 181-192. https://doi.org/10.1080/09709274.2005.11905828.

[27] A. Awad, I. Yussof, Factors affecting fertility - new evidence from Malaysia, Bull. Geography. Soc.-Econ. Ser. 36 (2017), 7-20. https://doi.org/10.1515/bog-2017-0011.

[28] A.A. Grasa, Econometric model selection procedures: A survey, Adv. Stud. Theor. Appl. Econ. (1989), 57-91. https://doi.org/10.1007/978-94-017-1358-0_4.

[29] A. Gelman, A. Jakulin, M.G. Pittau, et al. A weakly informative default prior distribution for logistic and other regression models, Ann. Appl. Stat. 2 (2008), 1360-1383. https://doi.org/10.1214/08-aoas191.