# A Comparative Analysis of Some Link Functions for Binomial Regression Models with Applications to Bioassay Data

Tinuade B. Dawotola<sup>1</sup>; Feridun Tasdan<sup>2</sup>

<sup>1</sup>Department of mathematics and philosophy, Western illinois university <sup>2</sup>Department of mathematics and philosophy, Western illinois university

Publication Date: 2025/01/16

#### Abstract

The response variable that represents the number of successes in a series of trials is too complex for classical linear regression to handle. On the other hand, binomial regression is seen to be more suitable, especially when dealing with bioassay data. In the context of the Generalised Linear Model (GLM), binomial regression is examined using certain link functions. The link functions logit, probit, complementary log-log (cloglog), Laplace, and Cauchy are frequently used for binomial regressions. While clog log, Laplace, and Cauchy are asymmetrical link functions, logit and probit are symmetrical. The study thereby aims to evaluate the performance of these link functions according to the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) indices. The study sheds light on how well these link functions work in the field of bioassays for modelling reactions with a binary result by expanding the analysis to the particular application of bioassay data. Additionally, the study incorporates the dose-response model that is frequently used in bioassay investigations. This approach examines the link between administered dosages and observed reactions during a set time period by exposing various groups to varying quantities of toxins or drugs.

**Keywords**: Classical Linear Regression, Binomial Regression, Generalized Linear Model (GLM), Logit, Probit, Complementary Log-Log (Clog log), Symmetrical link Functions, Asymmetrical link Function, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Bioassay, Dose-Response, Quantal Responses.

## I. INTRODUCTION

Modelling the relationship between a single response variable and several explanatory factors using linear models is known as classical linear regression [16, 17]. However, standard linear regression is insufficient when the response variable is the number of successes in repeated trials of a binomial experiment (a binomial random variable) [18, 19]. For the response variable with a binomial distribution, binomial regression is the proper regression model, and the binomial random variable has a binomial distribution [20, 21]. These models fall within the Generalised Linear Models (GLM) category. To determine the relationship between the explanatory variables and the response variable's expected value, the GLM framework uses a link function [1, 22, and 23].

In binomial regression, success is linked to explanatory factors but not explicitly predicted as a linear matrix of these variables (covariates) [24, 25]. Rather, binomial regression predicts the likelihood of success by using the inverse of a specified link function on a linear combination of variables [26, 27]. Any monotonically rising function mapping values from the range  $(-\infty, \infty)$  to (0, 1) can be the selected inverse link function [28]. In practice, the cumulative distribution functions (cdf) of known random distributions are frequently used to generate this inverse link function [29, 30]. For example, the logistic cdf is equivalent to the logit link function, while the normal distribution's cumulative distribution function is the source of the probit link function.

The logit link function is one of the link functions that are commonly used in binomial regression models. It is favoured because it easily interprets regression coefficients [31]. The logit model has a closed form and a linear equation for the natural or canonical parameter of the underlying exponential family [32]. Even while it makes odds ratios easier to understand, the logit link function does not guarantee the best fit for every

Dawotola, T. B., & Tasdan, F. (2025). A Comparative Analysis of Some Link Functions for Binomial Regression Models with Applications to Bioassay Data. *International Journal of Scientific Research and Modern Technology*, *3*(12), 106–116. https://doi.org/10.5281/zenodo.14670643 binomial regression model [2]. The cloglog connection is asymmetric in these three link functions, but the logit and probit links are symmetric. The constraints of symmetric link functions have been investigated in several works [33]. Symmetric link functions might not be appropriate when the binary or binomial response's probability converges to 0 at a rate different than its convergence to 1 (for covariates) [3].

Therefore, in binomial regression, they might not always produce the best fit for a particular dataset. Sometimes asymmetric (or skewed) links are considered to be better options. According to Chen et al. [3], an asymmetric link function is preferred over a symmetric one when working with data that is considerably unbalanced (where the number of 1s and 0s changes significantly). Misjudging the link function can lead to significant bias and a higher mean square error in estimations for regression parameters and success probability, according to Czado and Santner [4]. Additionally, Collet [5] showed that for modelling a certain binary response dataset, an asymmetric link function could be more appropriate than a symmetric one.



Fig 1 Cumulative Distribution Function Corresponding to the Logit, Probit, Cloglog, Laplace and Cauchy Link Functions.

The symmetric connections of logit and probit, the asymmetric links of cloglog, Laplace's asymmetric link, and Cauchy's asymmetric link are the three link functions represented graphically by the cumulative distribution function in Figure 1.

The solid red line shows the logistic distribution's cumulative distribution function (cdf), which corresponds to the logit link. The cdf of the standard normal distribution, connected to the probit connection, is represented by the dashed blue line. The cdf of the Gumbel distribution, connected to the cloglog connection, is also shown by the dotted green line. Additionally, the dotted orange line shows the cdf of the Cauchy distribution, which is connected to the Cauchy link, and the dashed purple line shows the cdf of the Laplace distribution, which is connected to the Laplace link function. Curves flowing through the point (0, 0.5)symmetrically in reverse show that the logit and probit link functions are symmetric, convergent to 0 at the same rate as they approach 1. The cloglog, Laplace, and Cauchy link functions, on the other hand, exhibit asymmetric curves and approach 1 faster than 0. The interpretation and precision of bioassay data can be greatly impacted by the link function selection, which may influence toxicity evaluations and drug development choices.

#### II. METHODS

We start by creating generic notation. Let Z be an independent dichotomous binary random variable from a Bernoulli experiment with

$$Z = \begin{cases} 1, & \text{if outcome is success} \\ 0, & \text{if outcome is failure} \end{cases}$$
(1)

If Zj is distributed as a Bernoulli with a chance of success  $P(Zj = 1) = \pi j$  and a probability of failure  $P(Zj = 0) = 1 - \pi j$ , and there are n independent binary random variables Z1,...,Zn, then we have

$$Y = \sum_{j=1}^{n} Z_j \tag{2}$$

for j = 1,...,n. which is a number of successes from n repeated trials, and Y is distributed as a binomial with Y ~ Binomial(n, $\pi$ ).

Let Pi =  $\frac{Y_i}{n_i}$  for i = 1,...,m be a proportion of successes from m independent binomial obser-vations, with E(Yi) = ni $\pi$ i and E(Pi) =  $\pi$ i.

Moreover, we consider xi to be p-dimensional covariate vector associated with Yi and  $\beta$  to be p-dimensional regression coefficients vector. Binomial regression models assume that

$$\pi \mathbf{i} = \mathbf{g}(\mathbf{x}_{\mathbf{i}}^{\mathrm{T}} \boldsymbol{\beta}), \tag{3}$$

Where g(.) represents specific known cumulative distribution functions (cdf), the structure of the model in equation (3) can be transformed into

$$g-1(\pi i) = x_i^T \beta \tag{4}$$

fori = 1,...,m.

The inverse of the function  $g-1(\bullet)$  is known as the link function. For example, the logit link function is produced if  $g(\bullet)$  is the logistic cumulative distribution

function (CDF). Likewise, the cloglog link is generated using the Gumbel CDF, and the probit link is derived when  $g(\bullet)$  is the standard normal CDF. For more information, see Agresti [6] and McCullagh and Nelder [1].

Where  $\Phi$  represents the cumulative distribution function (CDF) of a standard normal distribution, i.e.,  $\Phi$  is the CDF of Normal (0,1).

Link type	Regression Model	Tolerance Model
Logit	$\log(\frac{\pi_i}{1-\pi_i}) = x_i^T \beta$	$\pi_i = \frac{exp(x_i^T\beta)}{1 + exp(x_i^T\beta)}$
Probit	$\Phi^{-1}(\pi_i) = x_i^T \beta$	$\pi_i = \Phi(x_i^T \beta)$
Cloglog	$\log(-log(1-\pi_i)) = x_i^T \beta$	$\pi_i = 1 - exp(-exp(x_i^T\beta))$
Tanlana	$\operatorname{tlog}(2\pi_i) = x_i^T \beta$	$\pi_i = \frac{1}{2} exp(\frac{x_i^T \beta}{t}), x < 0$
Laplace	$-\mathrm{tlog}(2(1-\pi_i)) = x_i^T \beta$	$\pi_i = 1 - \tfrac{1}{2} exp(-\tfrac{x_i^T\beta}{t}), x > 0$
Cauchy	$t \tan((\pi_i - \frac{1}{2})\pi) = x_i^T \beta$	$\pi_i = \frac{1}{\pi} \arctan(\frac{x_i^T \beta}{t}) + \frac{1}{2}$

Fig 2 Link Functions with Corresponding Regression and Tolerance Distribution.

$$\Phi(x_i^T \beta) = \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz$$
(5)

The Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) are used to assess models with three distinct link functions. The AIC is calculated by subtracting twice the number of parameters and twice the maximised log-likelihood [7]. The dispersion for binomial families is fixed at one, and the number of parameters is equal to the number of coefficients. Schwarz [8] introduced the BIC, often called the Schwarz Information Criterion (SIC). The likelihood function is used to calculate both AIC and BIC. The best model is the one with the lowest AIC and BIC [34]. When comparing non-nested models, such as those with various link functions, AIC and BIC are especially helpful since they strike a compromise between model fit and complexity.

$$AIC = 2k - 2\ln(L) \tag{6}$$

$$BIC = \ln(n)k - 2\ln(L) \tag{7}$$

where L is the model's maximised likelihood function value, n is the number of observations or sample size, and k is the number of estimated parameters in the model.

# III. DISCUSSION

To make sure the chosen model fits well, we do a simulation study using a binomial response. The following approach is used to run the simulation:

- Step 1: Choose two explanatory variable binomial regression models and establish a vector of regression coefficients  $\beta = (\beta 0, \beta 1, \beta 2) = (-3, 1, -1)$ .
- Step 2: Create xi = (1, xi1, xi2) T after generating xi1 ~Normal (0, 1) and xi2 ~Normal (0, 1).
- Step 3: Use a link function to calculate πi so that πi = g (xiT β).
- Step 4: Generate ni~Poisson ( $\lambda$ ) with  $\lambda = 200$ , and Yi ~Binomial (ni,  $\pi$ i), representing the binomial distribution with  $\pi$ i as the probability of success, and ni as the sample size for the repeated Bernoulli trials.

Two cases—one symmetric and the other asymmetric—were investigated using two simulations. The logit tolerance model is used to calculate the likelihood of success  $\pi i$  in Simulation I (assuming symmetry) (see Table 1). The cloglog tolerance model is used in Simulation II to calculate the probability of success,  $\pi i$ , given the asymmetry assumption.

We create 1000 datasets for every simulation. We create m = 100 binomial response variables in each dataset, represented by Yi,j, where i = 1,...,100 and Yi,j~ Binomial(ni, $\pi$ i). To find the best-performing model in terms of AIC and BIC, we then fit the logit, probit, and cloglog models for each simulated dataset, respectively.

Table 1 Model Comparison: Average AIC and BIC for Different Link Functions in Simulations I and II

Simulation	Link Type	Average AIC	Average BIC
I	Logit model	488.47	496.29
I	Probit model	506.10	513.92
I	Cloglog model	499.86	507.68
П	Logit model	511.71	519.53
п	Probit model	565.22	573.03
П	Cloglog model	490.98	498.79

A comparison of the models using AIC and BIC for each fitted link function in the two simulations is shown in Tables 1 and 2. The results in Table 1 show that the logit model has the lowest average AIC and BIC when

the real link is the logit (which corresponds to Simulation I). Conversely, the cloglog model shows lower averages of AIC and BIC in Simulation II, when the real link is cloglog.

Table 2 Percentage of the Lowest AIC and BIC Associated with the Link Functions

Simulation	Link Type	% Lowest AIC	% Lowest BIC
Ι	Logit model	89.7%	89.7%
I	Probit model	1.2%	1.2%
I	Cloglog model	9.1%	9.1%
II	Logit model	5.6%	5.6%
II	Probit model	0%	0%
П	Cloglog model	94.4%	94.4%

The results in Table 2 support this observation, showing that when the real link is logit, the logit model fits the data better in 89.7% of 1000 datasets. Similarly, when the genuine link is cloglog, the cloglog model performs exceptionally well, with a success rate of 94.4%. The simulation study's empirical data emphasises

how crucial it is to choose the right link function because a poor decision might seriously impair the model's fit.

- Another Logistic Model Example
- Cauchy Link Function Model

Table 3 Coefficients of the Cauchy Link Function Model



Model Metrics	Value		
Degrees of Freedom	49 Total (Null); 47 Residual		
Null Deviance	69.31		
Residual Deviance	36.7		
AIC	42.7		

# $\succ$ Explanation

The results from the GLM using the Cauchy link function are summarized in the tables above.

- Coefficients Table: The coefficients of the Cauchy link function model show the estimated effect of each predictor (x1 and x2) on the response variable y.
- Model Metrics Table: The model fit metrics for the Cauchy link function model include degrees of freedom, null deviance, residual deviance, and AIC.

These measures aid in evaluating the model's complexity and goodness of fit.

The Cauchy link function model exhibits a residual deviance of 36.7 and an AIC of 42.7, indicating its fit to the data and model complexity. Further analysis and comparison with alternative models may provide additional insights into the suitability of this model for the given data.

Table 5 Coefficients of the Laplace Link Function Model

Model	(Intercept)	x1	x2
Laplace Link Function	0.4838	1.3904	1.8589

The data generation process in this analysis involves a sample size of 50 observations. It includes an intercept term and two random variables (x1 and x2), each generated from a standard normal distribution (N(0,1)).

This defines custom link functions based on the Cauchy and Laplace distributions, in addition to using

standard link functions like probit and logit. These custom links offer alternative modeling approaches within the generalized linear model (GLM) framework.

Laplace Link Function Model

Table 6 Coefficients of the Laplace Link Function Model

Model (Intercept)  $\mathbf{x2}$  $\mathbf{x1}$ Laplace Link Function 0.48381.39041.8589

Table 7 Model Fit Metrics of the Laplace Link Function Model

Model Metrics	Value		
Degrees of Freedom	49 Total (Null); 47 Residual		
Null Deviance	69.31		
Residual Deviance	35.5		
AIC	41.5		

#### ➤ Explanation

The results from the GLM using the Laplace link function are summarized in the tables above.

#### • Coefficients Table:

The coefficients of the Laplace link function model show the estimated effect of each predictor (x1 and x2) on the response variable.

#### • Model Metrics Table:

The model fit metrics for the Laplace link function model include degrees of freedom, null deviance, residual

deviance, and AIC. These metrics help assess the goodness of fit and complexity of the model.

The Laplace link function model exhibits a residual deviance of 35.5 and an AIC of 41.5, indicating its fit to the data and model complexity.

#### Comparison and Conclusion

Comparing the two models based on the AIC (Akaike Information Criterion) and residual deviance:

#### • *AIC*:

Lower values indicate a better balance between model fit and complexity. In this case, the Laplace link function model has a lower AIC (37.97) compared to the Cauchy link function model (38.74), suggesting that the Laplace model may be preferred in terms of model fit and complexity.

#### • Residual Deviance:

Lower deviance values indicate a better fit of the model to the data. The Laplace link function model also has a slightly lower residual deviance (31.97) compared to the Cauchy link function model (32.74), further supporting its better fit to the data.

Based on these metrics, the Laplace link function model appears to be the better option among the two custom link functions considered in this analysis.



Fig 3 AIC and Deviance Comparison1to the logit, Probit, Cloglog, Laplace and Cauchy link Functions.

# IV. APPLICATIONS TO BIOASSAY DATA

The drug development process is a complex and resource-intensive journey, often requiring precise methodologies for analyzing experimental data. Classical linear regression, traditionally employed for modeling continuous outcomes, faces limitations when applied to response variables representing the number of successes in a series of experiments [36]. In response to this challenge, binomial regression emerges as a more appropriate approach, particularly within the context of bioassay data. Bioassay experiments play a pivotal role in drug development, involving the exposure of multiple groups to varying levels of toxins or drugs. The assessment of responses within a fixed period forms the basis for understanding the effectiveness and toxicity of substances [37]. Classical linear regression's inadequacy in handling such discrete and binary outcomes necessitates the exploration of alternative statistical approaches.

The primary objective is to unravel the intricate relationship between dose (x) and the probability of

success ( $\pi$ ). It is essential to employ a regression model, g( $\pi$ ) =  $\beta$ 1 +  $\beta$ 2x, to characterize this relationship. The model aims to provide a quantitative framework for describing the probability of success as a function of the administered dose. This study holds significance in fields such as toxicology and pharmacology, where understanding dose-response relationships is crucial. By delving into the historical context and applying modern regression models, we anticipate contributing valuable insights that can enhance our comprehension of quantal responses in bioassay data.

Anticipated results from this research have the potential to advance our understanding of the complexities between dose levels and the likelihood of success in bioassay outcomes. This project serves as a valuable exploration of statistical modeling in the context of historical and contemporary bioassay data. For bioassay data, the probit model is one of the earliest models utilised. Many branches of the biological and social sciences employ probit models, which have natural interpretations. For instance,  $x = \mu$  is known as the median lethal dosage, or LD(50), as it is the amount that is predicted to kill half of the animals.

#### Beetle Mortality Rate

In entomological and toxicological investigations, the quantity of insecticide required to eliminate a specified proportion of the intended subjects (such as the median lethal dose, LD50) typically serves as a measure of the insecticide's effectiveness. The probit model, initially formulated by Bliss and refined by Finney, stands as a frequently employed technique for determining LD50 values in insect populations.

Number killed(ys)	Number of bettles(yf)	Dose(x)	prop
6	53	1.6907	0.1016949
13	47	1.7242	0.2166667
18	44	1.7552	0.2903226
28	28	1.7842	0.5000000
52	11	1.8113	0.8253968
53	6	1.8369	0.8983051
61	1	1.8610	0.9838710
60	0	1.8839	1.0000000

Table 8 Beetle Mortality Data.

The number of dead beetles following five hours of exposure to gaseous carbon disulphide at different concentrations is displayed in Table 8 (data from Bliss, 1935). These findings have important biological implications. The superior fit of the clog log model suggests that the mortality rate increases more rapidly at higher doses, which aligns with the known toxicological profile of carbon disulphide.

A comparison of observed beetle mortality numbers with fitted values from several dose-response models is also included in the analysis, along with graphs showing the Proportion of Success (prop = ys/yf) plotted against dosage (x). Deviance statistics are also provided to evaluate the performance of each model.

The analysis aimed to identify the best link function by comparing models using the AkaikeInformation Criterion (AIC) and Deviance (D) values. Among the tested models, the ExtremeValue/Cloglog link function exhibited the lowest Deviance (D = 3.45) and AIC (AIC = 33.64), indicating superior model fit and performance compared to Logit (D = 11.23, AIC = 41.43) and Probit (D = 10.12, AIC = 40.32) link functions.

Based on these metrics, we conclude that the ExtremeValue/Cloglog link function is the most suitable for modeling the dose-response relationship in the context of beetle mortality after exposure to gaseous carbon disulphide.

The superior performance of asymmetric link functions in this beetle mortality study reveals that researchers should consider these models when analyzing insecticide efficacy. This could lead to more accurate estimations of lethal doses, potentially reducing the amount of pesticide needed for effective pest control.

Probability of Success vs. x



(A) Proportion of Success (prop = ys/yf) plot-ted against dose (x).

	ObsiValues	Logit	Probit	ExtremeValue
	6	3.46	3.36	5.59
2	13	9.84	10.72	11.28
з	18	22.45	23.48	20.95
-	26	33.9	33.82	30.37
5	52	50.1	49.62	47.78
-6	53	63.29	53.32	54.14
7	61	69.22	59.66	61.11
-8	60	08.74	59.23	59.95
9	D	11.23	10.12	3.45
ro	AIC	41.43	40.32	33.64

(B) Comparison of observed beetle mortality numbers with fitted values from different dose-response models.

# ➤ Another Example

The observed data from experiments involving the effects of insecticide on adult flour beetles, including the analysis of responses over time, sex differences, and the application of mechanistic models to quantify the observed differences.

The data is sourced from Hewlett's work in 1974. It presents the daily mortality numbers of adult flour beetles (Triboliumcastaneum) after exposure to pyrethrum, an insecticide derived from plants. The pyrethrum was mixed with oil and applied at specified rates over small experimental areas, where the beetles were confined yet had freedom of movement. Food was also provided to minimize natural mortality.

- > Table Analysis:
- Columns: Doses of a substance (0.20, 0.32, 0.50, 0.80 mg/cm<sup>2</sup>)
- Rows: Time in days (1 to 13)

*Time (day)	0	.20	0.32		0.50		0.80	
	Male	Female	Male	Female	Male	Female	Male	Female
1	3	0	7	1	5	0	4	2
2	11	2	10	5	8	4	10	7
3	10	4	11	11	11	6	8	15
4	7	8	16	10	15	6	14	9
5	4	9	3	5	4	3	8	3
6	3	3	2	1	2	1	2	4
7	2	0	1	0	1	1	1	1
8	1	0	0	1	1	4	0	1
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	1	1
11	0	0	0	0	0	0	0	0
12	1	0	0	0	0	1	0	0
13	1	0	0	0	0	1	0	0
Amount Dead	43	26	50	34	47	27	48	43
Total	144	152	69	81	54	44	50	47

- Data Points: Number of mortalities for male (M) and female (F) beetles
- Total Mortalities: Summed for each dose and gender Observations:
- Mortality increases over time for both genders.
- Higher doses generally result in more mortalities.
- Female beetles show higher mortality than males at lower doses (0.20, 0.32 mg/cm<sup>2</sup>).
- At higher doses (0.50, 0.80 mg/cm<sup>2</sup>), mortalities are more similar between genders. Potential Insights:
- Gender may influence susceptibility at lower doses.
- Time and dose are critical factors in mortality rates.

• The data could inform dose-response models and toxicity thresholds.

# > Limitation

While this study provides valuable insights, it is important to note its limitations. The simulations, while extensive, may not capture all possible real-world scenarios. Additionally, the bioassay data used, while representative, may not encompass the full range of complexities encountered in all types of bioassay studies. Future research could address these limitations by expanding the range of simulated scenarios and applying these methods to a wider variety of bioassay datasets.



Fig 5 Time-Response Relationship of Adult Flour Beetles. Observed Data for Male and Female Flour Beetles at Different Doses. Data Source: Pack and Morgan.

#### V. CONCLUSION

Evaluating the binomial regression model requires careful consideration of the selected link functions. The comparison of different link functions in symmetric and asymmetric scenarios is the goal of this research. As illustrated in earlier simulations, it was evident that the logit link function exhibits superior performance in fitting symmetric scenarios, while the cloglog link function excels in asymmetric situations.

The AIC and BIC values offer valuable insights into the model's goodness of fit, and the per-centage metrics indicate the frequency with which each link function demonstrates optimal per-formance in terms of AIC and BIC across multiple simulations. Consequently, it is concluded that the meticulous selection of a link function significantly contributes to achieving an enhanced fit for the binomial regression model.

# REFERENCES

- [1]. McCullagh, P. and Nelder, J.A., 1989. *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.
- [2]. Li, J., 2014. *Choosing the Proper Link Function for Binary Data.* Ph.D. thesis, The University of Texas at Austin.
- [3]. Chen, M.H., Dey, D.K., and Shao, Q.M., 1999. A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, 94(448), pp.1172–1186.
- [4]. Czado, C. and Santner, T.J., 1992. The effect of link misspecification on binary regression inference. *Journal of Statistical Planning and Inference*, 33, pp.213–231.
- [5]. Collet, D., 2003. *Modelling Binary Data*. 2nd ed. Boca Raton, USA: Chapman & Hall/CRC.
- [6]. Agresti, A., 2013. *Categorical Data Analysis*. 3rd ed. New York: John Wiley and Sons Inc.
- [7]. Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), pp.716–723.
- [8]. Schwarz, G.E., 1978. Estimating the dimension of a model. *Annals of Statistics*, 6(2), pp.461–464.
- [9]. Bliss, C.I., 1935. The calculation of the dosemortality curve. *Annals of Applied Biology*, 22, pp.134–167.
- [10]. Finney, D.J., 1973. Statistical Methods in Bioassay. 2nd ed. New York: Hafner.
- [11]. Dobson, A.J. and Barnett, A.G., 2018. *An Introduction to Generalized Linear Models*. 4th ed. Chapman and Hall/CRC, pp.121–128.
- Peng, X., Lei, C., and Sun, X., 2021. Comparison of lethal doses calculated using logit/probitlog(dose) regressions with arbitrary slopes using R. *Journal of Economic Entomology*, 114, doi:10.1093/jee/toab044.
- [13]. Chen, D.G., 2007. Dose-time-response cumulative multinomial generalized linear model. *Journal of Biopharmaceutical Statistics*, 17, pp.173–185. <u>https://org/doi:10.1080/10543400601001543</u>.

- [14]. Morgan, B.J.T., 1992. *Analysis of Quantal Response Data*. London: Chapman and Hall.
- [15]. Pack, S.E. and Morgan, B.J.T., 1990. A mixture model for interval-censored time-to-response quantal assay data. *Biometrics*, 46, pp.749–757.
- [16]. Maulud, D. and Abdulazeez, A.M., 2020. A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*, *1*(2), pp.140-147.
- [17]. James, G., Witten, D., Hastie, T., Tibshirani, R. and Taylor, J., 2023. Linear regression. In An introduction to statistical learning: With applications in python (pp. 69-134). Cham: Springer International Publishing.
- [18]. Griffith, D.A., 2024. Comments on the Bernoulli Distribution and Hilbe's Implicit Extra-Dispersion. *Stats*, 7(1), pp.269-283.
- [19]. Bayes, C.L.S.A.I.G., Bazán, J.L. and Valdivieso, L., 2024. A robust regression model for bounded count health data. *Statistical Methods in Medical Research*, 33(8), pp.1392-1411.
- [20]. Zaidi, A. and Al Luhayb, A.S.M., 2023. Two statistical approaches to justify the use of the logistic function in binary logistic regression. *Mathematical Problems in Engineering*, 2023(1), p.5525675.
- [21]. Muschelli, J., Betz, J. and Varadhan, R., 2014. Binomial regression in R. In *Handbook of Statistics* (Vol. 32, pp. 257-308). Elsevier.
- [22]. Zareie, B., Poorolajal, J., Roshani, A. and Karami, M., 2023. Outbreak detection algorithms based on generalized linear model: a review with new practical examples. *BMC Medical Research Methodology*, 23(1), p.235.
- [23]. Korhonen, P., Nordhausen, K. and Taskinen, S., 2024. A Review of Generalized Linear Latent Variable Models and Related Computational Approaches. Wiley Interdisciplinary Reviews: Computational Statistics, 16(6), p.e70005.
- [24]. Hoffmann, J.P., 2016. *Regression models for categorical, count, and related variables: An applied approach*. Univ of California Press.
- [25]. Hilbe, J.M., 2011. *Negative binomial regression*. Cambridge University Press.
- [26]. Hafemeister, C. and Satija, R., 2019. Normalization and variance stabilization of singlecell RNA-seq data using regularized negative binomial regression. *Genome biology*, 20(1), p.296.
- [27]. Comulada, W.S. and Weiss, R.E., 2007. On models for binomial data with random numbers of trials. *Biometrics*, 63(2), pp.610-617.
- [28]. Dontchev, A.L. and Rockafellar, R.T., 2009. *Implicit functions and solution mappings* (Vol. 543). New York: Springer.
- [29]. Smithson, M. and Shou, Y., 2017. CDF-quantile distributions for modelling random variables on the unit interval. *British Journal of Mathematical and Statistical Psychology*, *70*(3), pp.412-438.
- [30]. Mathew, A., 2024. Quantile cumulative distribution function and its applications. *Communications in Statistics-Theory and Methods*, 53(11), pp.4194-4206.

- [31]. Breen, R., Karlson, K.B. and Holm, A., 2018. Interpreting and understanding logits, probits, and other nonlinear probability models. *annual review of sociology*, 44(1), pp.39-54.
- [32]. Efron, B., 2022. *Exponential families in theory and practice*. Cambridge University Press.
- [33]. Shim, H., Bonifay, W. and Wiedermann, W., 2023. Parsimonious asymmetric item response theory modeling with the complementary log-log link. *Behavior Research Methods*, 55(1), pp.200-219.
- [34]. Agiakloglou, C. and Tsimpanos, A., 2023. Evaluating the performance of AIC and BIC for selecting spatial econometric models. *Journal of Spatial Econometrics*, 4(1), p.2.
- [35]. van der Nest, G., Passos, V.L., Candel, M.J. and van Breukelen, G.J., 2020. An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. *Advances in Life Course Research*, 43, p.100323.
- [36]. Fox, J., 2015. *Applied regression analysis and generalized linear models*. Sage publications.
- [37]. Blomme, E.A. and Will, Y., 2016. Toxicology strategies for drug discovery: present and future. *Chemical research in toxicology*, 29(4), pp.473-504.