Iraqi Journal of Statistical Sciences, Vol. 23, No. 1, 2026, Pp (141-148)

Iraqi Journal of Statistical Sciences

http://stats.uomosul.edu.iq

Survival Analysis, Mortality Rate and Generalized Linear Models (GLM) with Failures (Application Study)

Mahdi Saber Raza Shalee[*]

Department Of Software Engineering, University of Salahaddin Erbil, College of Engineering, Kurdistan Region, Iraq.

Article information

Article history:
Received: September 1, 2025
Revised: October 11,2025
Accepted: December 30,2025
Available online: June 1,2026

Keyword:
Survival analysis, Mortality rate, lost data, Generalized linear models; complementary log-log, Latin square and Failure

Correspondence:
Mahdi Saber Raza Shalee[*]
mahdi.raza@su.edu.krd

Abstract

The number of failures of the first type is counted for each of a series of n time intervals, whereas the number of failures of the second type is only determined for the entire period. This theoretical framework is offered for the analysis of survival data when two forms of failure occur. First-type failure rates are associated with experimental and explanatory factors, while second-type failure rates are regarded as nuisance characteristics. A Latin square experiment is used to describe two models that are based on a precise and an approximate method. As long as there is a 2^{-n} fraction of experimental units until the experiment's conclusion, the approximate model outperforms the exact model. The study presents a theoretical framework for analyzing survival data in which two types of failures occur. The first type's failures are tallied for each of n time intervals, whereas the second type's failures are only counted for the entire period. First-type failure rates are associated with experimental and explanatory factors, while second-type failure rates are regarded as nuisance characteristics.

DOI: 10.33899/iqjoss.v23i1.62133 , ©Authers, 2026, College of Computer Science and Mathematics University of Mosul.
This is an open access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

A lot of interest has been shown in survival analysis with various kinds of failure, including failure and censoring, as well as in connecting survival times to experimental and explanatory parameters. The techniques are especially applicable to dependability testing in industrial applications where survival times can be determined and individuals may be identified, as well as to the medical sciences. Instead of identifying the individuals, the number of failures within a specified time period is counted.

Two different kinds of failures are addressed in this unique situation. While the number of failures of the second type is only recorded for the entire period, the number of failures of the first type is recorded for each successive sub-interval of the overall time period. The following illustrates the potential relationships between explanatory variables and experimental factors and failure rates. An experiment that uses a Latin square arrangement of water velocities to evaluate the downstream distribution of juvenile trout is described in part 2. A theoretical basis for the study is presented in part 3, while parts 4 and 5 provide an approximation and a more precise method of estimating parameters. According to Thompson and Baker (1981), the precise estimation process is an illustration of a general composite link function. More specifically, because there is no single link function and nuisance parameters that are independent of explanatory variables are included in the model, the process fits into the overall framework that Cox (1984) defined. In Section 6, the data analysis is discussed and the parameter estimations from the two methods for the experiment in part 2 are compared. As long as the losses aren't so large that the experimental supply is severely exhausted, the two methods produce estimates that are comparable.

2. Materials and methods

2.1 The application of the Study

Salmon farming projects in the Kurdistan region of Iraq are increasing due to the availability of suitable water and climate conditions, with 12 projects producing around two thousand tons of fish annually. On the other hand, these projects need government support, as there are very suitable areas for salmon farming and investment in the provinces of Duhok, Erbil, and Sulaymaniyah. On the way to the salmon ponds in the mountainous areas near the border with Turkey next to the Great Zai River, the water here does not stop flowing, in the rivers and springs, while rainfall hardly ceases in the Dieralok area it established numerous salmon ponds that surpass an annual production of 60-80 tons of fish Baban, S. (2015) & Fahad, H (2024) .

The trial is part of a series that examines how susceptible juvenile (Shabot and Baleot) salmonid fish in channel Dukan are to downstream dispersal due to changes in water velocity brought on by reservoir releases. Young fish are placed in four water channels at the beginning of each experiment, and they are given time to settle in. Four water velocity treatments are applied in a Latin square configuration over the course of four days. At the end of the trial, all of the fish that are still in the channels in Duhok city were recovered by electrofishing and tallied. The displaced fish are gathered at the channel Duhok outflows in nets that are emptied every day. Natural mortality is the cause of the disparity between the initial stock size and the number of fish that were relocated or recovered. In the spring of 2023, six similar tests were conducted in four experimental channels Authority's Dieralok, three of which involved Shabot (Salmo salar L.) and three of which involved Baleot (Salmo trutta L.) Harlıoğlu, M., et.al (2023). In order to examine three distinct developmental periods, ranging from 8 to 55 days after emerging from the gravel, the trials switched between species. Low (5.5 cm/time (s)(means that the fish travels 5.5 Centimeters in 1 Second) , medium-low (20 cm s-1), medium-high (35 cm s-1), and high flow (65 cm s-1) were the four water velocity treatments Giblin, P. (2017). Originally, the channels were stocked with roughly 150 young fish, with the exception of the third trout experiment, where the initial population was only about 65 due to a lack of experimental material. This is the experiment that is shown in (Table 1).

Experiment data relating young trout's susceptibility to downstream dispersal by varying water velocities Raju, Arumugam et.al (2021). Over the course of four days, four channels get four treatments: low A, medium low (B), and medium high (C). Stock size, numbers recovered during the experiment, and numbers lost due to death and emigration are provided. Fitted values and residuals, respectively, for the precise model outlined in part 6 are indicated by values enclosed in brackets.

Table1: Shows the period of four days, recovery and mortality.

Channel	**First**	**Second**	**Third**	**Four**
Initial stock	**65**	**63**	**63**	**61**
Emigration				
Day 1	*A* 11 (14.0, -0.80)	*D* 41 (39.0, 0.32)	*C* 21 (22.5, -0.31)	*B* 8 (6.1, 0.79)
Day 2	*B* 3 (3.1, -0.06)	*C* 4(5.7, -0.70)	*D* 17 (16.0.15)	*A* 7 (507, 0.55)
Day 3	*C* 14 (10.4, 1.12)	*B* 0 (1.6, -1.27)	*A* 6 (4.2, 0.90)	*D* 10 (13.9, -1.05)
Day 4	*D* 1 (0.6, 0.44)	*A* 0 (0.1, -0.35)	*B* 0 (0.1, -0.27)	*C* 0 (0.2, -0.41)
Recovery	21 (22.3, -0.27)	14 (12.6, 0.40)	16 (16.9, -0.23)	13 (11.7, 0.39)
Mortality	15 (14, 0.11)	4 (4.0, -0.01)	3 (3.0, 0.01)	23 (23.5, -0.10)

Although the two species may have different preferences for velocity, young salmon and brown trout often inhabit first-flowing rivers and streams. According to one explanation, the fish eat the animal debris that floats in the river. As a result, water velocity may have an impact on the rate at which food is supplied as well as the energy that fish use to feed and maintain their place in the flow. The fish may migrate downstream to a more suitable habitat if the water velocity is too high for it. Nevertheless, very little is understood about the dynamics at play here, the reason behind the exodus, or whether it was voluntarily or not. At this point, the channel studies are not intended to explicitly clarify the mechanisms at play, but rather to establish a relationship between emigration rate and water velocity treatment Jing Ouyang., et.al (2023).

2.2. Conceptual Structure

Examine an observational research or experiment where n time-sequences $(t_0, t_m]$ are split into m successive time periods $t_0 < t_1 < t_2 < t_3 < \ldots < t_m$. Wij individuals emigrate K_{ij}, individuals die, and L_{ij} individuals survive to start time period $i+1$ from Nj individuals at time t_0, who start time-sequence j ($j = 1, 2, \ldots, n$), in each time period i ($i = 1, 2, \ldots, m$). L_{mj} individuals have survived the entire time sequence as of time t_m. W_{ij} ($i = 1, 2, \ldots, m$), L_{mj}, and the total number of deaths X_j, which is determined by $X_j = \sum K_{ij}$, are the observed counts for time-sequence j. The notations are summarised below for clarity.

Table2: The survival time to emigration in the absence mortality time-sequence *j*:

period	1	2	...	m	Total
time interval	$(t_0, t_1]$	$(t_1, t_2]$	…	$(t_{m-1}, t_m]$	
emigrations	W_{1j}	W_{2j}	...	W_{mj}	.
mortalities	K_{1j}	K_{2j}	...	K_{mj}	X_j
survivals	L_{1j}	L_{2j}	...	L_{mj}	.
total	C_j	L_{1j}	...	L_{m-1}, j	

With interval-specific rates, survival time to death in the absence of emigration σ_{ij} and survival time to emigration in the absence of mortality β_{ij}, assume that the survival time to emigration in the absence of mortality and the survival time to death in the absence of emigration are exponentially distributed for period i of sequence j. If the survival time to death is μ, the survival time to emigration is S, and the survival time to either death or emigration is E, the survival probability is given as Raza& Broom (2023) and Jiahui Zhang (2018)

$$\Pr(E > t_i \,|E > t_{i-1}) = e_{ij} = \exp\left\{-\left(\sigma_{ij} + \beta_{ij}\right)(t_i - t_{i-1})\right\}, \tag{1}$$

$$P_r(\, t_{i-1} < S \le t_i \,|E > t_{i-1}) = p_{ij} = \frac{\sigma_{ij}}{\sigma_{ij} + \beta_{ij}}\left(1 - e_{ij}\right), \tag{2}$$

$$P_r(\, t_{i-1} < \mu \le t_i \,|E > t_{i-1}) = q_{ij} = \frac{\beta_{ij}}{\sigma_{ij} + \beta_{ij}}\left(1 - e_{ij}\right). \tag{3}$$

Expectations are then given by, while the C_j is individuals at time t_0.

$$E\left(K_{1j}\right) = C_{jq1j}, \qquad E\left(K_{ij}\right) = C_{jqij} \prod_{k=1}^{i-1} e_{kj}\,, \quad i > 1 \tag{4}$$

$$E\left(W_{1j}\right) = C_{jp1j}, \qquad E\left(W_{ij}\right) = C_{jpij} \prod_{k=1}^{i-1} e_{kj}\,, \quad i > 1 \tag{5}$$

$$E\left(L_{mj}\right) = C_j \prod_{k=1}^{m} e_{kj}\,, \qquad E\left(X_j\right) = C_j\{q_{1j} + \textstyle\sum_{i=2}^{m}(q_{ij} \prod_{k=1}^{i-1} e_{kj})\}. \tag{6}$$

Without loss generally, all time intervals are assumed to have unit lengths. Now, two more presumptions are introduced to facilitate the construction of the basic theory. These must be justified in any application, including the trout experiment. Emigration rates are thought to have a log-linear relationship with both experimental and explanatory factors. Hence $\log \sigma_{ij} = \eta_{ij}$ where η is a linear function of the explanatory variables. With a sequence-specific rate of $\beta_{ij} = \beta_j$, it is also assumed that mortality remains constant across a sequence. For observations w_{ij}, x_j, l_{mj} of random variables W_{ij}, X_j, L_{mj} ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$) estimates of mortality rates β_j ($j = 1, 2, \ldots, n$) and of the linear model parameters are required.

2.3. The Estimate Parameter in the Model.

Assumption that all mortality in the time interval $(t_{i-1}, t_i]$ occurs instantly at time ti and has anticipated value may serve as the foundation for an approximation.

$$E\left(K_{ij}\right) = L_{i-1,j}\{1 - \exp\left(-\beta_j\right)\}. \tag{7}$$

In the absence of emigration, this is the number that is anticipated to perish. The principles L_{ij} can be calculated using

$$\tilde{L}_{0j} = C_j, \tilde{L}_{ij} = \tilde{L}_{i-1,j} - w_{ij} - E\left(K_{ij}\right), (i = 1, 2, \ldots, m). \tag{8}$$

It is possible to select the mortality rate in such a way that the observed and expected total mortality in sequence j are equal. An easy search or one solution to the equation will yield the estimated mortality, $\tilde{\beta}_j$, with this quality.

$$C_j \exp\left(-l\tilde{\beta}_j\right) - \sum_{i=1}^{m} w_{ij} \exp\left\{-(l-i)\tilde{\beta}_j\right\} = l_{mj.} \quad (9)$$

It is thus believed that emigration will act with constant risk during that time. When mortality is absent, p_{ij}=1- exp (-σ_{ij}) , which yields, where $\eta_{ij} = \log\left(-\log\left(1 - p_{ij}\right)\right.$ is the proposed linear function of the emigration explanatory factors. The generalized linear interactive modeling (GLIM) tool can be used to fit a log linear model to the emigration rates using some simplifying assumptions (Baker and Nelder, 1978). A supplementary log-log function connects the linear model to the binomial expectation, and the numbers emigrating, wij, are assumed to be binomially distributed with denominators $\check{L}_{i-1,j}$. Because losses are lower when both forces of removal occur simultaneously than when each is considered to act independently, this method will typically understate both emigration and mortality rates.

2.4. The Practical Method

It is assumed in the precise approach that mortality and emigration factors operate simultaneously. Equation (1) provides the predicted values for the multinomially distributed individuals that bring a sequence into $l+2$ classes. Under the condition that they add up to the multinomial denominator, Birch (1963) shows that the probability of the $l+2$ multinomial classes is proportional t to that of the $l+2$ Poisson observations. Therefore, maximising a Poisson rather than a multinomial likelihood may yield maximum likelihood estimates. Equation (1) provides the relationship between Poisson expectation and linear function parameters and mortality rates. This relationship is the kind of broad composite link function that Thompson and Baker (1981) defined. Iterative weighted least squares can be used to fit the parameters using maximum likelihood (Appendix).

According to Cox (1984), there is frequently no single link function in estimating using exponential family models. Since each component of the matrix of derivatives needed by the iterative technique needs to be calculated independently, the precise model presented here is simply one example of this type. The inclusion of nuisance factors that are independent of predictor variables is made possible by Cox's more general approach, and mortality rates can be conceptualised in this manner in the precise model. On the other hand, Maximum likelihood estimation for the accurate approach, If α is the vector of linear model parameters and γ the vector of log mortality rates was $\gamma_i = \log\beta_i$ After that the absolute vector of model parameters λ , of length l, is given by $\lambda' = (\alpha', \gamma')$, where the superscript denotes the vector transposed Raza & Broom (2016). The full design matrix of model A is given by

$$A = \begin{pmatrix} L & 0 \\ 0 & I \end{pmatrix}, \quad (10)$$

Where L is the Latin square design matrix and I is the identity matrix for the log mortality rates. The vector of emigration rates σ and the mortality rates β is concatenated to give the vector

$$\tau' = (\sigma', \beta') \text{ and } \log\tau = \eta = A\beta. \quad (11)$$

Each element of the vector of predictable values μ is related to the elements of τ by the functions given in equation (1). At each iteration the next working dependent variable u is considered from the vector of observations x and the earlier fitted means $\mu^*$by $u = D_\eta + (x - \mu^*)$, where D is the matrix of derivatives (d_{rs}) and

$$d_{rs} = \frac{\partial\mu_r}{\partial\eta_s} = \frac{\partial\mu_r}{\partial\tau_s}\,\tau_s. \quad (12)$$

The next set of independent variables, vectors $U_1, \ldots, U_l$, are obtained from

$$V = (U_1, \ldots, U_l) = DA. \quad (13)$$

The vectors of weights w at each iteration are those appropriate to the Poisson variance-mean relationship, namely $w_j = {}^{1}\!/_{\mu_j^*}$.

2.5. **The Analysis of the study**

An extensive calibration exercise was conducted to examine the physical differences between the water channels prior to the start of the series of experiments to explore salmonid reaction to varied water velocity. This exercise found that although the channels varied in various ways, they were appropriate for the suggested tests. The four channels are specifically coupled so that each pair of channels shares a feed pipe from the header tank. One channel's needs take precedence over those of its partner when the values from the two channels are opened to control the water discharge. Additionally, the water flow varies from one channel to another along its width and length. As a result, each treatment's flow is a target average flow, requiring intricate valve adjustments. Depending on the channel, real treatments may deviate from intended treatments. Additionally, the header tank supply is insufficient to sustain high flows in every channel at the same time. Despite its weak design, a previous pilot study indicated that the impact of prior treatment on current time-period results was minimal. Since the treatments' standard errors differed in the pilot study analysis, one goal of further research was to estimate treatment effects with a similar level of accuracy. As a result, the study was designed and analysed as a sequence of Latin square experiments, with a term for the impact of channels on emigration rates included in the analysis from the outset. The distribution and migrations of young salmonids in their natural habitat are probably related to the predicted relative treatment effects. Nevertheless, young fish in the wild cannot be affected by the anticipated mortality rates. Predatory fish like the bullhead (Cottus gobio L.) are admitted into the channel inflows, and the quantity of these fish may differ from one channel to another. There may be channel-specific differences in the species composition and population density of other predators, such as bottom-dwelling invertebrates (predatory Plecoptera). The fish in a single channel share a habitat where they are equally protected from predators and at risk of them. Mortality will not be correlated with treatment because it is improbable that any physical harm to the fish at the flows under consideration will result in mortality. Therefore, it makes sense to assume channel-specific mortality in this experiment. Because mortality rates in some channels may be intentionally elevated and hence have no intrinsic value, it may be vital to accurately assess them to lessen bias in treatment effect estimations Younis S., et.al (2020) & Raza& Broom (2023) . The above-mentioned exact and approximate approaches have been used to analyse the data from Table 1. The data from a single channel are used to identify each of the four time sequences. The Latin square design that incorporates the effects of day (d_i), channel (c_j), and treatment (U_k) is suitable for the proposed linear model for log-emigration rate, specifically

$$\log \sigma_{ij(k)} = \eta_{ij(k)} = Q + d_i + c_j + U_k \ , (i, j, k = 1, 2, 3, 4). \tag{14}$$

Table 3 compares the parameter estimates for the two methods. When compared to the exact model, the approximate model performs well in terms of estimate precision. For the analysis of the data in Table 1 using approximate and precise models, parameter estimates and bracketed approximate standard errors are used. Additionally, the models' deviations are provided.

Table 3: Estimated parameters and bracketed approximate standard errors are used

		Approximate model	*Exact model*
Mean	Q	-1.42 (0.25)	-1.36 (0.25)
Days	d_2	-0.48 (0.23)	-0.49 (0.23)
	d_3	-0.25 (0.27)	-0.25 (0.27)
	d_4	-3.42 (1.02)	-3.44 (1.02)
Channels	c_2	0.20 (0.34)	0.16 (0.35)
	c_3	0.10 (0.30)	0.06 (0.31)
	c_4	-0.11 (0.31)	-0.05 (0.32)
Treatments	U_2	-0.74 (0.38)	-0.75 (0.38)
	U_3	0.50 (0.30)	0.50 (0.30)
	U_4	1.18 (0.31)	1.20 (0.31)
Mortality	$\log \beta_1$	-2.44	-2.34 (0.26)

	$\log\beta_2$	-3.32	-3.05 (0.50)
	$\log\beta_3$	-3.85	-3.67 (0.58)
	$\log\beta_4$	-1.79	-1.68 (0.21)
Deviance		9.88 (6 df)	9.93 (6 df)

As anticipated, the approximate method understates the general mean in the linear model, which includes the average emigration rate and mortality. Relative day and treatment impact estimates, however, are strikingly similar. Since the treatment effects are the main parameters of interest in this experiment, this is a pleasing outcome.

The goal of estimating treatment effects with comparable precision has been accomplished because the standard errors of the treatments are comparable. The two models have different channel effects, which quantify the impact of divergence from target flow on emigration along and across the channel. The link between channel effects and channel-specific mortality rates which gauge a channel's level of protection and risk from predation could be the explanation for this. This is not the case, as can be seen by looking at the correlation matrix of the parameters for the exact model (Table 4). All of the linear model's parameters have no correlation with the estimated log-mortality rates. In contrast, the linear model parameters exhibit a high degree of association with one another. The factors are orthogonal and the corresponding submatrices of the correlation matrix, such as between treatment and times, are zero in a Latin square designed experiment where observations are considered to be normally distributed and effects are additive. of contrast, the effects of the Latin square model fitted here are multiplicative. The explanation for the good estimates by the approximate model is simple. If, for N individuals starting a unit time w emigrates and k die then the maximum likelihood estimates σ and $\widehat{\sigma}'$of the emigration rate for the approximate and exact model are given by the following equation. The approximation model's good estimations have a straightforward explanation. The maximum likelihood estimates σ and ($\widehat{\sigma'}$)of the emigration rate for the approximate and exact models are provided by assuming that N people begin a unit time w emigrate and k die.

$$\hat{\sigma} = \log\left(1 - \frac{w}{C}\right) = \log\,(w, C), \qquad \widehat{\sigma'} = \frac{w}{w+k}\, log\left(1 - \frac{w+k}{C}\right) = \log h(w, k, C). \tag{15}$$

Table 4: Parameter estimate correlation matrix for the specific model.

Q	1.00													
d_2	-0.23	1.00												
d_3	-0.30	0.28	1.00											
d_4	-0.11	0.07	0.13	1.00										
c_2	-0.42	0.14	0.17	0.17	1.00									
c_3	-0.40	-0.11	0.14	0.14	0.69	1.00								
c_4	-0.37	-0.05	0.07	0.07	0.50	0.44	1.00							
v_2	-0.34	-0.04	0.03	-0.02	-0.16	-0.06	-0.27	1.00						
v_3	-0.41	0.06	-0.30	-0.07	-0.35	-0.45	-0.02	0.34	1.00					
v_4	-0.27	-0.09	-0.31	-0.12	-0.64	-0.48	-0.28	0.44	0.69	1.00				
$\log\beta_1$	0.01	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.00	1.00			
$\log\beta_2$	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	1.00		
$\log\beta_3$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	
$\log\beta_4$	0.00	0.00	0.01	0.00	0.00	0.00	0.04	-0.01	0.00	0.00	0.00	0.00	0.00	1.00
	Q	d_2	d_3	d_4	c_2	c_3	c_4	v_2	v_3	v_4	$\log\beta_1$	$\log\beta_2$	$\log\beta_3$	$\log\beta_4$

As long as the ratios w / C and k / C are not excessively big, the function *g(w,C)* approximates the function *h(w,k,C)* quite well. The former function is within 7% of the latter function if *1-(w+k)/C* does not exceed 0.5. In an experiment with four consecutive time periods, a minor recovery of one- sixteenth of the beginning stock will guarantee that the two analysis approaches produce results that are comparable. To lessen bias in the estimate of the linear model's parameters, a higher computational effort is needed for the precise analysis if recovery is low. The square of the linear predictor was added to the linear model using the approximation approach in order to test the hypothesis that emigration rates are correlated with experimental factors in a log linear fashion. After decreasing from 9.88 (6 df) to 9.80 (5 df), the deviation showed that the multiplicative scale is sufficient. For the precise model of Table 3, Table 1 provides the observed (*O*), fitted (*F*) value, and Pearson results computed as$(O - F)/\sqrt{F}$. Given the size of the residuals, the model's fit to the data is sufficient. 9.93 on 6 df is well within the accepted 95% margins, assuming that the chi-squared approximation to the deviation is satisfactory. Plotting residuals versus previously administered treatment will show inadequacies with respect to carry-over effects, which were formerly thought to be insignificant. There is no discernible pattern in the residuals plotted against the previous treatment and against whether the previous treatment was higher or lower than the current treatment; however, if the treatments are rearranged in ascending order *B, A, C,* and *D* based on the magnitude of their estimated parameters, and the residuals are plotted against whether the previous treatment was higher or lower, then a pattern may be suggested. Five residuals are negative and one is positive for treatments with higher velocity in the past. However, of these five, two (-0.27) and (-0.41) are derived from zero observations with modest expected values (0.1 and 0.2), and one residual is extremely small (-0.06). Two negative and one positive residual are left after ignoring these as proof of lack of fit. A dummy variable that takes a unit value when the prior treatment is higher in the order *B, A, C,* and *D* and zero otherwise was added to the linear model in order to test for lack of fit. The deviation decreased from 9.93 (6 df) to 9.01 (5 df), and the estimated coefficient was -0.54 with a standard error of 0.57. The presumption that the impact of prior therapy is insignificant is maintained notwithstanding the lack of compelling evidence demonstrating a lack of fit.

While the calculated treatment parameters are not monotonically ordered, the treatment levels are. According to the multiplicative treatment effects of 1.0 (*A*), 0.47 (*B*), 1.65 (*C*), and 3.32 (*D*), young trout prefer a medium-low flow of roughly 25 cm *s-1*, below which emigration will be higher. The three trout trials all show a similar pattern of a *J* shaped reaction, which suggests that the pattern of response is real and not the result of a single experiment. The linear model's channel effects were negligible, suggesting that the emigration rate is not much impacted by physical channel variations. This supports the notion that repeated results are achievable and that the channels are an appropriate experimental medium for such investigations.

3. Conclusion

The assumption of constant mortality within a series was made in the creation of the approximate and exact models. This assumption is reasonable when using the model for the trout experiment, but in other contexts, it could be more reasonable to assume that mortality is correlated with treatment or linearly correlated with additional explanatory variables. Depending on how complex the mortality model is, the approximation model might be used in this situation. It would be challenging to estimate mortality initially and later eliminate it, and the approximation model would become less useful. The linear model for mortality can be included in the design matrix, and computing the matrix of derivatives would be no more challenging than under the simple assumption.

Acknowledgments

The data were obtained from a personal project by the manager Mr Biwar Haji in Dieralok- Duhok, supported by the Kurdistan Region Government. Thanks also go to the Director of the Fisheries Department in Kurdistan Region-Iraq.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Ethical Approval

Ethical approval was not required for this study as it did not involve human participants, personal data.

References

1. Baban, S. (2015). "Revitalising Agriculture and Water Sectors in the Kurdistan Region, Iraq", Athens: ATINER'S Conference Paper Series, No: GEO2015-1618.

2. Baker. R. J. and Nelder, J. A. (1978*) General Linear Interactive Modelling (GLIM). Release 3.* Oxford: Numerical Algorithms Group.
3. Birch, M. W. (1963) Maximum likelihood in three-way contingency tables. *J. R. Statist. Soc.* B, 25, 230-233.
4. Cox, C. (1984) Generalized linear models-the missing link. *Appl. Statist*., 33, 18-24.
5. Thompson, R, and Baker, R. J. (1981) Composite link functions in generalized linear models. *Appl. Statist*., 30, 125-131.
6. Farhad, Hasan Aziz (2024). Climate and Water Resources of Iraq and Kurdistan Region 2022.University of Slemani.
7. Frank J. Hall and Mirosliv Rozloznik (2016) G-matrices, J-orthogonal matrices, and their sign patterns Czechoslovak Mathematical Journal, 66 (141) (2016), 653–670.
8. Giblin, P. (2017). [Review of *Latin squares and their applications (2nd edn.)*, by D. Keedwell & J. Dénes]. *The Mathematical Gazette*, Vol. 101, No. 552, pp. 571-573. https://www.jstor.org/stable/26538967
9. Harlıoğlu, M., Omar Mustafa Mustafa, S., and Batool, Z. (2023). The Present Situation of the Fisheries Sector in Iraq: A Critical Review. Çanakkale Onsekiz Mart University Journal of Marine Sciences and Fisheries, 6(1), 70-75. https://doi.org/10.46384/jmsf.1216078
10. Jiahui Zhang (2018). Maximum Likelihood (ML) Extraction in Exploratory Factor Analysis (EFA) with SPSS. January, 23, 2018/ Applied Statistics Journal. https://jiahuizzz.wordpress.com/2018/01/23/maximum-likelihood-ml-extraction-in-exploratory-factor-analysis-efa-with-spss/
11. Jing Ouyang, Kean Ming Tan and Gongjun Xu (2023). High-Dimensional Inference for Generalized Linear Models with Hidden Confounding. Journal of Machine Learning Research. Volume 24, pp. 1-61. https://www.jmlr.org/papers/volume24/22-0834/22-0834.pdf.
12. Raju, Arumugam , S.Safrin Preethi and Rajathi .M (2021). Application of latin sequare design in the analysis of the effectiveness of four different traffic violation on road accidents study. Compliance Engineering Volume.12, No. 2, pp.18-27. ISSN NO: 0898-3577
13. Raza, M.S. and Broom, M. (2016). Survival Analysis Modeling With Hidden Censoring. Journal of Statistical Theory and Practice, Vol. 10, No. 2, pp. 375-388. https://doi.org/10.1080/15598608.2016.1152205
14. Raza, M. S. and Broom, M(2023). The use of survival analysis modelling with incomplete data with application to breast cancer, Asian Journal of Probability and Statistics, Vol. 25,ISSUE, 3., pp. 45-69.DOI: 10.9734/ajpas/2023/v25i3563 .
15. Younis S. Abdullah, Shamall M. A. Abdullah, & Ridha H. Hussein. (2020). Biodiversity of Fishes in Sulaimani Province in Kurdistan Region, Iraq. *Zanco Journal of Pure and Applied Sciences*, *32*(1), 39–44. https://doi.org/10.21271/ZJPAS.32.1.5

تحليل البقاء على قيد الحياة ومعدل الوفيات والنماذج الخطية المعممة (GLM) مع حالات الفشل (دراسة تطبيقية)

مهدي صابر رضا شالي
قسم هندسة البرمجيات ،كلية الهندسة ،جامعة صلاح الدين، اربيل.

الخلاصة: يحسب عدد حالات الفشل من النوع الأول لكل سلسلة من فترات زمنية من ***n***، بينما يُحدد عدد حالات الفشل من النوع الثاني للفترة بأكملها فقط. يُقدم هذا الإطار النظري لتحليل بيانات البقاء عند حدوث نوعين من الفشل. ترتبط معدلات فشل النوع الأول بالعوامل التجريبية والتفسيرية، بينما تُعتبر معدلات فشل النوع الثاني خصائص غير المتوقعة. تُستخدم تجربة المربع اللاتيني لوصف نموذجين يعتمدان على طريقة دقيقة وأخرى تقريبية. طالما كان هناك جزء 2^(–n) من الوحدات التجريبية حتى انتهاء التجربة، فإن النموذج التقريبي يتفوق في الأداء على النموذج الدقيق. تقدم الدراسة إطارًا نظريًا لتحليل بيانات البقاء التي يحدث فيها نوعان من الفشل من الاسماك. تُحسب حالات فشل النوع الأول لكل فترة زمنية من ***n***، بينما تُحسب حالات فشل النوع الثاني للفترة بأكملها فقط. ترتبط معدلات فشل النوع الأول بالعوامل التجريبية والتفسيرية، بينما تُعتبر معدلات فشل النوع الثاني خصائص للخصائص غير المتوقعة

الكلمات المفتاحية : تحليل البقاء على قيد الحياة، معدل الوفيات، البيانات المفقودة، النماذج الخطية المعلمة؛ اللوغاريتم التكميلي، المربع اللاتيني والفشل.