Zagazig Veterinary Journal, ©Faculty of Veterinary Medicine,
Zagazig University, 44511, Egypt.
Volume 50, Number 2, p. 177-185, June 2022
DOI: 10.21608/zvjz.2022.140572.1181

RESEARCH ARTICLE

Log-linear Model for Describing the Relationship between Chromosomal Aberrations and Infertility Problems in Holstein-Friesian Cows

Fatma D. M. Abdallah
Department of Animal Wealth Development, Faculty of Veterinary Medicine, Zagazig University, 44511, Egypt.
* Corresponding author: moibho_stat2021@yahoo.com

Article History: Received: 24/05/2022 Received in revised form: 20/06/2022 Accepted: 30/06/2022

Abstract

Log-linear analysis is widely applied in different scientific research areas, such as its use in veterinary medicine. The objective of this study is to model the relationship between chromosomal aberrations and some diseases in different groups of Holstein-Friesian cows arranged in a contingency table using log linear model. The variables under study were chromosomal aberrations (structural and numerical) and groups of animals with normal (control) and abnormal states. The SPSS statistical package was used for analyzing the data. The results showed that the saturated model significantly fitted the data. The likelihood ratio statistic was 421.023 with a *P*-value of 0.000, indicating that two-way interactions (group of animals × chromosomal aberrations; group of animals × disease status; and chromosomal aberrations × disease status) have a highly significant effect and are good predictors in the model. The three-way interaction (group of animals, chromosomal aberrations, and disease status) was not significant (*P*-value = 0.858), so it was eliminated. After backward elimination statistics, it is found that all two-way interactions (group of animals × chromosomal aberrations, group of animals × disease status and chromosomal aberrations × disease status interactions) should not be deleted from the model to avoid model distortion. The total diseased animals compared to total non-diseased ones are both more likely to be grouped where odds ratio = 1.45 with 95% CI (1.549 - 1.360) and be supposed to have chromosomal aberrations. This model was the best fit model because it showed all possible effects, including main effects, interaction effects between each two variables, and interaction effects between the three variables.

Keywords: Chromosomal aberrations, contingency tables, log-linear models, odds ratio, and saturated model.

Introduction

It is known that the relationship between chromosomal aberrations and reproductive disorders in livestock is of great importance, so it has been studied for many years [1]. Studying and modelling these relationships depends on using different statistical methods of categorical data analysis.

It is known that the analysis of categorical data is a very important part of statistics. Classical statistical methods for categorical data analysis such as the chi-square test are not adequate in the case of many qualitative variables (more than two) and it cannot determine the interaction among them [2].

A Log-linear model is a discrete multivariate statistical method used for analyzing and modelling qualitative factors data that doesn't require a response [3].It is not important to divide the variables into dependent and independent where all the variables are at the same level.

The chi-square test used two qualitative factors only. The log-linear technique can detect different interactions in multidimensional contingency tables with more than two categorical variables. Log-linear models are similar to analysis of variance (ANOVA) with multifactorial designs, so it is known as ANOVA of qualitative data, which uses the likelihood ratio chi-square statistic [4].

This model is very useful for the analysis of count data by calculating the maximum likelihood estimates, and it can be suitable when the data is continuous [5]. The count variable characterized by its distribution is not normal and its variance is not homogeneous, so a log-linear model and Poisson distribution are suggested [6].

The difference between this technique and other qualitative data analysis methods is the application of models. In the case of the log-linear technique, a distribution is assumed for the data, a model is assumed, and estimations are obtained to be compared with the observed frequencies for model evaluation [7, 8].

Log-linear models have two types: association models and logit models. Association models aim to show the independence or relationship between variables using the independence test of chi-square. Logit models aim to detect whether values of specific categories of the first variable differ in response to another variable by a chi-square homogeneity test where the first variable is the explanatory, and the second is the response or logit [9].

This study aimed to examine the association between chromosomal aberrations and some reproductive disorders in Holstein-Friesian cows depending on suggesting a suitable model for the prediction process. This model also shows the effect of main variables and their interactions (two- way and three-way interactions).

Materials and methods

Source of data

The data set is obtained from a study conducted in animal wealth development department; Faculty of veterinary medicine, Zagazig University, to examine the relationship between chromosomal aberrations and some reproductive disorders in Holstein-Friesian cows [10].The author selected samples randomly and used chi-square test for analysis.

These variables were arranged as follows:

Chromosomal aberrations are divided into numerical (polyploidy and aneuploidy) and structural (gaps, breaks, deletion, fragments, ring chromosomes and centromeric attenuation).

Animals were divided into two groups: normal (no disease or control) and abnormal. The abnormal state is related to different infertility problems (repeat breeder, anestrum, retained placenta, free-martin, vaginal prolapse, uterine prolapse, uterine torsion, and habitual abortion) and the total presence of disease (total yes, total no). These variables are qualitative (nominal) and divided into different categories.

A sample of 60 Holstein-Friesian cows (12 control, 8 repeat breeding animals, 6 anestruous, 4 freemartin, 8 retained placenta, 5 vaginal prolapse, 6 uterine prolapse, 5 animals with habitual abortion, and 6 uterine torsion) from different private farms located at Sharkia Governorate. The animals which subjected to this study were classified into nine different groups.

Animals were classified according to their reproductive histories from farm records, clinical signs, and rectal palpation.

Blood samples were collected, whole blood cultures were set up, and chromosomal preparations were made. Three thousands metaphases were examined for different groups of animals.

Variables are arranged in a contingency table. Associations and interactions between variables and among sub-categories were examined with log-linear models. Statistical calculations were applied by using SPSS 25[11].

Model characteristics

It is known that is $y = b_0 + b_1x_1 \ldots + b_jx_j + e$ is the general linear model. This part ($b_0 + b_1x_1 \ldots + b_jx_j$) determines the predicted value of $y / x_1 \ldots x_j$. [12].

1. The log-linear model is a general linear model with a logarithm function as a link function $\eta = \log(\mu)$.
2. The inverse link function is an exponential function. The mean is $\mu = E(y) = e^{\eta}$.
3. This model is the logarithm of the expected frequency as a linear combination of the main effects and interactions. It uses odds

ratios which can be defined as $\pi/(1-\pi)$, and it is the ratio of event occurrence (π) to that of not occurrence ($1-\pi$). The odds ratio with symbol ψ is as follows:

$$\psi = \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]}$$

4. Mu is the predicted value of y, then f (mu) = $b_0 + b_1x_1 ... + b_jx_j$. Then the model is log (mu) = $b_0 + b_1x_1 ... + b_jx_j$.

When the parameter values $b_0 ... b_j$ are important for calculating a predicted value for y, then Mu = exp ($b_0 + b_1x_1 ... + b_jx_j$).

5. The Poisson variable mean and variance are equal to λ.

$\mu = \text{var}(y) = \lambda$ and $\mu_i = \lambda_i = e^{x_i\beta}$.

Where e is the base of the natural logarithm (e = 2.718).

6. The log-linear model of a Poisson variable is Log (μ_i) = log (λ_i) = $X_i\beta$.

Where, X_i is a linear combination of the vector of explanatory factors X_i and the corresponding vector of parameters β.

Mathematical model

There are many types of this model as follows:

a. *The saturated model* which includes all interactions [13].

In case of contingency table with three categorical variables A, B, and C with indices a, b, and c. Log of m_{abc} is the predicted frequency. The saturated log-linear model for the three-way contingency table is as in the following equation:

$$\log m_{abc} = \lambda + \lambda_a^A + \lambda_b^B + \lambda_c^C + \lambda_{ab}^{AB} + \lambda_{ac}^{AC} + \lambda_{bc}^{BC} + \lambda_{abc}^{ABC}$$

$$m_{abc} = \exp\left(\lambda + \lambda_a^A + \lambda_b^B + \lambda_c^C + \lambda_{ab}^{AB} + \lambda_{ac}^{AC} + \lambda_{bc}^{BC} + \lambda_{abc}^{ABC}\right)$$

Any group of parameters sums to zero in what is called effect coding; the λ term denotes the grand mean of log m_{abc}. The single-factor estimates are λ_a^A, λ_b^B and λ_c^C. The two-factor interaction λ_{ab}^{AB}, λ_{ac}^{AC} and λ_{bc}^{BC} show the partial association between A and B, A and C, and B and C, respectively. The three-variable interaction parameters λ_{abc}^{ABC} show the conditional two-factor interactions which is different from one another within the categories of the third variable [14].

$$\sum_a \lambda_a^A = \sum_b \lambda_b^B = \sum_c \lambda_c^C = 0,$$

$$\sum_a \lambda_{ab}^{AB} = \sum_b \lambda_{ab}^{AB} = \sum_a \lambda_{ac}^{AC} = \sum_c \lambda_{ac}^{AC} = \sum_b \lambda_{bc}^{BC} = 0,$$

$$\sum_a \lambda_{abc}^{ABC} = \sum_b \lambda_{abc}^{ABC} = \sum_c \lambda_{abc}^{ABC} = 0.$$

b. Non-saturated models

Theus and Lauer [15] showed different types from non-saturated as follows:

1. Mutual independence model → (λ^X, λ^Y, and λ^Z only).
2. Partial independence → (additional presence of one λ^{AB}, A, B ∈ {X, Y, Z}, A ≠ B).

$$\text{Log}(m_{ijk}) = \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z + \lambda_{jk}^{yz}.$$

3. Conditional independence → (all parameters except λ^{XYZ} and one λ^{AB}, A, B ∈ {X, Y, Z}, A ≠ B). Log (m_{ijk}) $= \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z + \lambda_{ik}^{xz} + \lambda_{jk}^{yz}$.
4. No three-way interaction→ (all parameters found except λ^{XYZ}).

Log (m_{ijk}) $= \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z + \lambda_{ij}^{xy} + \lambda_{ik}^{xz} + \lambda_{jk}^{yz}$.

5. Three-way interaction (includes all parameters; the saturated model).

The expected values e_{ijk} similar to the observed ones o_{ijk} exactly. In this case χ^2 and G^2 values are equal zero as goodness of fit tests [16].

Hierarchical log-linear model is a non-saturated (trivariate) independence model $\log m_{abc} = \lambda + \lambda_a^A + \lambda_b^B + \lambda_c^C$.

The model assumptions

Independence of the observations with random selection of subjects, the sample should be sufficient enough, the expected frequencies ≥ 5 for ≥ 80% of the categories and all expected frequencies ≥ 1 to avoid reducing power. Low power can be avoided by increasing data, combining variable categories or removing variable [17].

Hypothesis is tested using Pearson's chi square and the likelihood ratio statistic. It is

that the expected frequencies differ from the observed ones.

The likelihood-ratio chi-square (G^2) = $2\sum[O_i \ln(O_i/E_i)]$ where O_i is the observed and E_i is the expected frequency.

The model is a good fit in case of the observed and expected frequencies are very like (not have a significant difference). A significant result indicated that the model was significantly unlike the data (bad fit model). The likelihood ratio statistic is good in small samples [18].

Results

Likelihood ratio and Pearson chi square values were not calculated and their values were 0 as shown in Table (1).

Table 1: **Goodness-of-fit tests for initial examination of the model.**

	Chi-square	Df	Significance
Likelihood ratio	0.00	0	.
Pearson	0.00	0	.

-df are the degrees of freedom.

As illustrated in Table (2) that the likelihood ratio and Pearson chi-square statistics when K = 1 were 5024.815 and 4903.707, respectively and their P-value were 0.000** which indicated a highly significant effect and the main effects terms were not removed from the model.

When (K = 2) the two-way interactions (group of animals×chromosomal aberrations, group of animals × disease status and chromosomal aberrations × disease status interactions).

Likelihood ratio and Pearson chi-square statistics when K = 2 were 421.055 and 393.683, respectively and their *P*-value were 0.000** which indicated a highly significant effect of two-way interactions and these interactions were not removed from the model. Removing this significant interaction from the model means a bad effect on it.

For (K = 3) means the three-way interactions (group of animals × chromosomal aberrations × disease status interactions).

The likelihood ratio and Pearson chi-square statistics for K = 3 were 0.032 and 0.032, respectively, and their *P* value was 0.858 which indicated a non-significant effect of three-way interactions. Therefore, it was removed from the model without affecting the model fitness.

Also, Table (2) for K-way effects without higher order for K = 1, 2, and 3 gives the same meaning of K-way effects with higher order.

Table 2: K-way and higher-order effects for detecting the significance of main and interaction effects.

	K	df	Likelihood ratio		Pearson		Number of iterations
			Chi-square	Sig.	Chi-square	Sig.	
K-way and higher order effects	1	7	5024.815	0.00	4903.707	0.00	0
	2	4	421.055	0.00	393.683	0.00	2
	3	1	0.032	0.858	0.032	0.858	3
K-way effects	1	3	4603.760	0.00	4510.024	0.00	0
	2	3	421.023	0.00	393.651	0.00	0
	3	1	0.032	0.858	0.032	0.858	0

-df are the degrees of freedom. – sig. is the P value.

Partial chi-square test for two-way interaction (group of animals*chromosomal aberrations, group of animals *disease status and chromosomal aberrations *disease status) showed significant parameters with *P* value = 0.014**, 0.00**, 0.00*, respectively. Main effects (group of animals and disease status) were significant where their *P* values were 0.00**. Chromosomal aberrations main effect was non-significant (*P* value = 1.000) as shown in Table (3).

Table 3: Partial associations for deleting the non-significant effect from the model.

Effect	df	Partial chi-square	Sig.	Number of iterations
Group of animals *Chromosomal aberrations	1	6.092	0.014	2
Group of animals *Disease status	1	210.826	0.000	2
Chromosomal aberrations *Disease status	1	216.289	0.000	2
Group of animals	1	2312.937	0.000	2
Chromosomal aberrations	1	0.000	1.000	2
Disease status	1	2290.823	0.000	2

The backward elimination for choosing the best-fit model is shown in Table (4). The three-way interaction terms (group of animals *chromosomal aberrations* disease status interaction) were deleted one at a time as the first steps (*P*-value more than 0.05). Non-significant interaction terms were deleted one at a time until all those left were significant (Group of animals * chromosomal aberrations, Group of animals *disease status and chromosomal aberrations *disease status interactions). The process then stopped and the best-fit model for this data is

$$\log m_{abc} = \lambda + \lambda_a^A + \lambda_b^B + \lambda_c^C + \lambda_{ab}^{AB} + \lambda_{ac}^{AC} + \lambda_{bc}^{BC} .$$

Table (4): Backward elimination statistics for detecting the deleted effect step by step.

Step			Effects	Chi square	df	Sig.	Number iterations
0	Generating class		Group of animals *Chromosomal aberrations* Disease status.	0.00	0	.	
	Deleted effect	1	Group of animals *Chromosomal aberrations* Disease status.	0.032	1	0.858	3
1	Generating class		Group of animals *Chromosomal aberrations. Group of animals *Disease status. Chromosomal aberrations *Disease status.	0.032	1	0.858	
	Deleted effect	1	Group of animals *Chromosomal aberrations.	6.092	1	0.014	2
		2	Group of animals *Disease status.	210.826	1	0.00	2
		3	Chromosomal aberrations *Disease status.	210.826	1	0.00	2
2	Generating class		Group of animals *Chromosomal aberrations.	0.032	1	0.858	3
			Group of animals *Disease status.				
			Chromosomal aberrations *Disease status.				

-At each step, the effect with the largest significance level for the Likelihood ratio change is deleted, provided the significance level is larger than .050.
- Statistics are displayed for the best model at each step after step 0.
- For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

The Parameter estimates table gave single estimates (λ) for each effect, z-score rather than chi-square test to obtain confidence intervals. The value of z is the important value for comparison between effects as shown in Table (5).

Table (5): Parameter estimates for detecting the most important effect.

Effect	Parameter	Estimate (λ)	SE	Z	Sig.	95% CI	
						Lower Bound	Upper Bound
Group of animals *Chromosomal aberrations* Disease status	1	-0.004	0.033	-0.130	0.896	-0.069	0.061
Group of animals *Chromosomal aberrations	1	-0.044	0.033	-1.337	0.181	-0.109	0.021
Group of animals *Disease status	1	-0.373	0.033	-11.262	0.000	-0.438	-0.308
Chromosomal aberrations *Disease status	1	-0.251	0.033	-7.562	0.000	-0.316	-0.186
Group of animals	1	-0.964	0.033	-29.080	0.000	-1.029	-0.899
Chromosomal aberrations	1	-0.179	0.033	-5.386	0.000	-0.243	-0.114
Disease status	1	-0.995	0.033	-30.032	0.000	-1.060	-0.930

-SE is the standard error. – λ is the mean of the model. – CI is the confidence interval.

For prediction process the model will be

Log m_{abc}= (λ) + (λ) for group of animals + (λ) for chromosomal aberrations + (λ) for disease status + (λ) for group of animals * chromosomal aberrations + (λ) for group of animals *disease status + (λ) for chromosomal aberrations *disease status.

The main effect (disease status of animals) is the most important effect in the model (z = -30.032) as it had the highest value among the main effects. This means that the most important factor for detecting frequencies in the contingency table is disease status followed by the group of animals (z = -29.080) and then the group of animals*disease status interaction (z = -11.262) followed by chromosomal aberrations *disease status where (z = -7.562) and their *P* values were 0.00**. The most important part of variables interactions (two-way) are group of animals*disease status and chromosomal aberrations *disease status. Group of animal*Chromosomal aberrations interaction was non-significant with *P* value = 0.181.

Odds ratios are calculated to reflect the nature of the interaction. The exponential (exp) of the estimate (λ) gives the odds ratio as presented in Table (5). The parameter estimates of the interaction part (group of animals*disease status is -0.373 and exp (0.373) = 1.45 with a *P*-value of 0.00.

The parameter estimates of the interaction part (chromosomal aberrations *disease status is -0.251 and exp (0.251) = 1.28 with a *P*-value of 0.00.

Finally, the final model is evaluated with the likelihood ratio and chi-square test. The results showed a test with non-significant *P*-value of 0.858 as shown in Table (6). This means that the predicted values obtained by the model were not significantly different from the observed data and that the non- saturated without three-way interaction model fitted the data adequately.

Table 6: Goodness-of-fit tests for final evaluation of the model.

	Chi-square	df	significance
Likelihood ratio	0.032	1	0.858
Pearson	0.032	1	0.858

-df the degrees of freedom.

Discussion

Goodness of fit tests are two tests of the null hypothesis and their values were an indicator of fitting the saturated log-linear model to the data because chi-square value was not calculated [16, 18].

The components and the interactions between variables and which of them will be deleted from the model is studied. As a general rule, any effect (main or interaction) with *P*-value > 0.05 will be deleted from the model as it is non-significant.

Likelihood ratio and Pearson chi-square statistics with their *P*-values are calculated when K = 1, 2 and 3. When (K =1) this belongs the one-way effects (the main effects of group of animals, chromosomal aberrations and disease status) and any higher-order effects have an important role in affecting the model significance. There are many higher-order effects (two-way interactions and the three-way interaction).

The main effects (group of animals, chromosomal aberrations and disease status) and the two-way interaction effects (group of animals × chromosomal aberrations, group of animals × disease status and chromosomal aberrations × disease status interactions) remained in the model as a good predictors. Any one of two-way interaction will remain or will be deleted from the model is not known until this step.

Finally, the model with the first and second order effects would be good for representing data.

Partial associations for examining which effects are significant using partial chi-square statistics. Partial chi-square test for two-way interaction (group of animals*chromosomal aberrations, group of animals *disease status and chromosomal aberrations *disease status) showed

significant effect. Main effects (group of animals and disease status) were significant and chromosomal aberrations main effect was non-significant.

After determining which parameters were significant and which model fitted the data well. The backward elimination process is used for detecting the significance of the elements in a saturated model. It started with all the components of the saturated model and removed the effects one at a time from the highest to the lowest order. The "deleted effect" is the alteration in the chi-square after the effect is removed from the model. At each step, the effect with the highest significance level for the likelihood ratio change is removed (*P*-value greater than 0.05). This process started with the saturated model [19].

The estimates (λ) used to estimate the cell frequencies. These coefficients examine the dependency of associated categories of the variables. These coefficients can be standardized by its division by its standard errors. The standardized parameters are known as "Z-value". The estimation of the standardized parameter (Z) shows which categories association is the most important with neglecting the plus or minus sign. The highest Z value was the most effective in the model, as the results showed that the disease status of animals was the most effective.

Two significant relationships were found (group of animals*disease status interactions). This means that total diseased animals compared to total non-diseased ones are both more likely to be grouped (OR = exp (0.373) = 1.45 with 95% CI exp (-0.438) =1.549 to exp (-0.308) = 1.360 and be supposed to chromosomal aberrations (OR = exp (0.251) = 1.28 with 95% CI exp (-0.316) =1.371 to exp (-0.186) = 1.20. The standardized form (Z) for both interactions is of near sizes (-11.262 and -7.562). This indicated that both relationships are equally important to explain the studied data.

Finally, the expected and observed frequency was not different and chi-square value calculated with non-significant *P*-value.

Conclusions

A log-linear model is a type of general linear model used to study relationships between the qualitative factors in a contingency table. This is used to show the best fit model for investigating the relationship between chromosomal aberrations and some reproductive disorders in Holstein- Friesian cows. It is better than the chi-square test in studying the relationship between more than two categorical variables and giving the best fit model for prediction methods.

From the results of this study, the conclusion is that the best fit model was stopped at the two way interactions by deleting the effects of three factors. The model is important for predicting the effect of each category on each factor in cows depending on odds ratio and this improves the future production by avoiding these problems.

Conflict of interests

The author has no conflict of interest to disclose.

References

[1] Holecková, B.; Schwarzbacherová Galdíková, M.; Koleničová, S.; Halušková, J.; Staničová, J.; Verebová, V. and Jutková, A. (2021): Chromosomal aberrations in cattle. Genes, 12:9-1330. https:// doi.org /10.3390/ genes12091330. https:// www. mdpi. com/journal/genes.

[2] Agresti, A. (1996): An Introduction to categorical data analysis, 1st edition. USA: John Wiley & Sons145-150 p.

[3] Goodman, L.A. (1971): The analysis of multidimensional contingency tables: Stepwise procedures and direct estimation methods for building models for multiple classifications. Technometrics, 13:33.

[4] Allen, M. (2017): Log-linear analysis. The SAGE Encyclopedia of Communication Research Methods.

https:// dx.doi. org/10. 4135/ 978148 3381411.

[5] Berg, P.T. (1980): On the log linear Poisson and gamma model. Astm Bulletin, 11:35 -40.

[6] Kaps, M. and Lamberson, W.R. (2004): Biostatistics for animal science. CABI Publishing 412 p.

[7] Hoffman, D.C. and Franke, G.R. (1986): Correspondence analysis: Graphical representation of categorical data in marketing research. J. of Market. Res., 23:213-27.

[8]Van der Heijden, P.G.M.; de Falguerolles, A. and de Leeuw, J. (1989): A Combined approach to contingency table analysis using correspondence analysis and loglinear analysis. Applied Statistics, 38: 249-92.

[9] Hinkle, D.E and McLaughlin, G.W. (1984): Selection of models in contingencytables: Reexamination. Agathon Press Inc, 21: 4-415.

[10] Awad Allah, A. W. Z. (2011): Cytogenic studies related to some infertility problems in cattle and buffloes. Master Thesis, Fac. of Vet. Med., Zagazig Univ., Egypt.

[11] SPSS V25. IBM Corp. Released (2017): IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

[12] Seavy, N. E.; Quader, S.; Alexander, J.D. and Ralph. C.J. (2005): Generalizedlinear models and point count data: Statistical considerations for the design and analysis of monitoring studies. USDA Forest Service Gen, Tech Rep PSW GTR 191.

[13] Ripatti, P. (2000): Use of log-linear models in forecasting structural changes in Finnishnon-industrial private forest ownership. Silva Fennica, 3: 275–283.

[14]Vermunt, J. K. (2005): Log-linear models. Encyclopediaof statistics in behavioral science. UK: Wiley, Chichester1082-1093 p.

[15] Theus, M. and Lauer, S.R.W. (1999): Visualizing log linear model. American Statistical Association, Institute of Mathematical Statistics, and Interface Foundation of North America Journal of Computational and Graphical Statistics, 3: 396 - 412.

[16] Chan, Y. H. (2005): Biostatistics 306. Log-linear models: Poisson regression. Singapore Med J,46: 8- 377.

[17] Field, A.(2005): Discovering statistics using SPSS. 2nd edition. Thousand Oaks. CA: Sage Publications695–718 p.

[18] Field, A.(2018): Discovering statistics using IBM SPSS statistics. 5thedition. Ltd (UK): Sage Publications.

[19] Olmus, H. and Erbas, S. (2012): Analysis of traffic accidents caused by drivers by using log-linear models. Promet Traffic&Transportation. 6: 495-504.

الملخص العربي

النموذج الخطي ا للوغاريتمي لوصف العلاقة بين الانحرافات الصبغية ومشاكل العقم في أبقار هولشتاين فريزيان

فاطمة دسوقي محمد عبدالله

قسم تنمية الثروة الحيوانية- كلية الطب البيطري-جامعة الزقازيق

يتم تطبيق النموذج الخطي اللوغاريتمي على نطاق واسع في مجالات البحث العلمي المختلفة مثل استخدامه في الطب البيطري. حيث يهدف هذا البحث الي نمذجة البيانات العددية باستخدام النموذج اللوغاريتمي الخطي في تحديد طبيعة العلاقة بين الانحرافات الصبغية وبعض الأمراض في مجموعات حيوانية مختلفة من أبقار هولشتاين- فريزيان المرتبة في جدول مقارنة. وكانت المتغيرات قيد الدراسة هي الانحرافات الصبغية (التركيبية والعددية) ومجموعات الحيوانات (حيوانات ذات تاريخ مرضي وحيوانات لا تعاني من اي امراض) وتم استخدام الحزمة الاحصائية اس بي اس اس لتحليل البيانات لتظهر النتائج التالية. والتي كانت تساوي 421.023 وكانتLikelihood value وجد أن النموذج المشبع كان اكثر ملائمة للبيانات اعتمادا علي ذات دلالة معنوية عالية جدا مما يدل علي وجود تأثيركبير للنموذج الثنائي (مجموعةالحيوانات × =P) (**0.000قيمة الانحرافات الصبغية،مجموعةالحيوانات × حالةالمرضوالانحرافاتالصبغية × حالةالمرض) في عملية التبنؤ. وهي غير .0 =P)858)كما وجد عدم معنوية النموذج الثلاثي (مجموعةالحيوانات × الانحرافاتالصبغية × حالةالمرض) لأن قيمة معنوية مما يرجح حذف النموذج الثلاثي لأنه غير ملائم لتمتيل البيانات وغير ملائم لعملية التنبؤ. وجد أيضا جميع أجزاء النموذج الثنائي يجب الأ تحذف حتي ولو لم يتحدد ايها ذات دلالة معنوية او لا حتي لا يتشوه نموذج التنبؤ. كما استخدمت نسبة الارجحية في عملية المقارنة حيث كانت نسبة الحيوانات المريضة مقارنة بالحيوانات السليمة والتي من الممكن ان توضع في مجموعات وتكون اكثر عرضة للانحرافات الصبغية بنسبة 1.45 وحدود ثقة (1.549-1.360). الخلاصة ان هذا النموذج اكثر ملائمة للبيانات حيث انه يظهر التاثيرات الاساسية للمتغيرات والتفاعلات الثنائية والثلاثية بينهم للتعبير عن البيانات بصورة جيدة