



جامعة
بنغازي الحديثة



**مجلة جامعة بنغازي الحديثة للعلوم
والدراسات الإنسانية
مجلة علمية إلكترونية محكمة**

العدد الخامس

لسنة 2019

حقوق الطبع محفوظة

شروط كتابة البحث العلمي في مجلة جامعة بنغازي الحديثة للعلوم والدراسات الإنسانية

- 1- الملخص باللغة العربية وباللغة الانجليزية (150 كلمة).
- 2- المقدمة، وتشمل التالي:
 - ❖ نبذة عن موضوع الدراسة (مدخل).
 - ❖ مشكلة الدراسة.
 - ❖ أهمية الدراسة.
 - ❖ أهداف الدراسة.
 - ❖ المنهج العلمي المتبع في الدراسة.
- 3- الخاتمة. (أهم نتائج البحث - التوصيات).
- 4- قائمة المصادر والمراجع.
- 5- عدد صفحات البحث لا تزيد عن (25) صفحة متضمنة الملاحق وقائمة المصادر والمراجع.

القواعد العامة لقبول النشر

1. تقبل المجلة نشر البحوث باللغتين العربية والانجليزية؛ والتي تتوافر فيها الشروط الآتية:
 - أن يكون البحث أصيلاً، وتتوافر فيه شروط البحث العلمي المعتمد على الأصول العلمية والمنهجية المتعارف عليها من حيث الإحاطة والاستقصاء والإضافة المعرفية (النتائج) والمنهجية والتوثيق وسلامة اللغة ودقة التعبير.
 - ألا يكون البحث قد سبق نشره أو قُدم للنشر في أي جهة أخرى أو مستل من رسالة أو اطروحة علمية.
 - أن يكون البحث مراعيًا لقواعد الضبط ودقة الرسوم والأشكال - إن وجدت - ومطبوعاً على ملف وورد، حجم الخط (14) وبخط (Arial 'Body') للغة العربية. وحجم الخط (12) بخط (Times New Roman) للغة الإنجليزية.
 - أن تكون الجداول والأشكال مدرجة في أماكنها الصحيحة، وأن تشمل العناوين والبيانات الإيضاحية.
 - أن يكون البحث ملتزماً بدقة التوثيق حسب دليل جمعية علم النفس الأمريكية (APA) وتثبيت هوامش البحث في نفس الصفحة والمصادر والمراجع في نهاية البحث على النحو الآتي:
 - أن تُثبت المراجع بذكر اسم المؤلف، ثم يوضع تاريخ نشره بين حاصرتين، يلي ذلك عنوان المصدر، متبوعاً باسم المحقق أو المترجم، ودار النشر، ومكان النشر، ورقم الجزء، ورقم الصفحة.
 - عند استخدام الدوريات (المجلات، المؤتمرات العلمية، الندوات) بوصفها مراجع للبحث: يُذكر اسم صاحب المقالة كاملاً، ثم تاريخ النشر بين حاصرتين، ثم عنوان المقالة، ثم ذكر اسم المجلة، ثم رقم المجلد، ثم رقم العدد، ودار النشر، ومكان النشر، ورقم الصفحة.
2. يقدم الباحث ملخص باللغتين العربية والانجليزية في حدود (150 كلمة) بحيث يتضمن مشكلة الدراسة، والهدف الرئيسي للدراسة، ومنهجية الدراسة، ونتائج الدراسة. ووضع الكلمات الرئيسية في نهاية الملخص (خمس كلمات).

3. تحتفظ مجلة جامعة بنغازي الحديثة بحقها في أسلوب إخراج البحث النهائي عند النشر.

إجراءات النشر

ترسل جميع المواد عبر البريد الإلكتروني الخاص بالمجلة جامعة بنغازي الحديثة وهو كالتالي:

- ✓ يرسل البحث إلكترونياً (Word + Pdf) إلى عنوان المجلة info.jmbush@bmu.edu.ly او نسخة على CD بحيث يظهر في البحث اسم الباحث ولقبة العلمي، ومكان عمله، ومجاله.
- ✓ يرفق مع البحث نموذج تقديم ورقة بحثية للنشر (موجود على موقع المجلة) وكذلك ارفاق موجز للسيرة الذاتية للباحث إلكترونياً.
- ✓ لا يقبل استلام الورقة العلمية الا بشروط وفورمات مجلة جامعة بنغازي الحديثة.
- ✓ في حالة قبول البحث مبدئياً يتم عرضة على مُحكمين من ذوي الاختصاص في مجال البحث، ويتم اختيارهم بسرية تامة، ولا يُعرض عليهم اسم الباحث أو بياناته، وذلك لإبداء آرائهم حول مدى أصالة البحث، وقيمتها العلمية، ومدى التزام الباحث بالمنهجية المتعارف عليها، ويطلب من المحكم تحديد مدى صلاحية البحث للنشر في المجلة من عدمها.
- ✓ يُخطر الباحث بقرار صلاحية بحثه للنشر من عدمها خلال شهرين من تاريخ الاستلام للبحث، وبموعد النشر، ورقم العدد الذي سينشر فيه البحث.
- ✓ في حالة ورود ملاحظات من المحكمين، تُرسل تلك الملاحظات إلى الباحث لإجراء التعديلات اللازمة بموجبها، على أن تعاد للمجلة خلال مدة أقصاها عشرة أيام.
- ✓ الأبحاث التي لم تتم الموافقة على نشرها لا تعاد إلى الباحثين.
- ✓ الأفكار الواردة فيما ينشر من دراسات وبحوث وعروض تعبر عن آراء أصحابها.
- ✓ لا يجوز نشر إي من المواد المنشورة في المجلة مرة أخرى.
- ✓ يدفع الراغب في نشر بحثه مبلغ قدره (400 دل) دينار لبيي إذا كان الباحث من داخل ليبيا، و (200 \$) دولار أمريكي إذا كان الباحث من خارج ليبيا. علماً بأن حسابنا القابل للتحويل هو: (بنغازي - ليبيا - مصرف التجارة والتنمية، الفرع الرئيسي - بنغازي، رقم 001-225540-0011. الاسم (صلاح الأمين عبدالله محمد).
- ✓ جميع المواد المنشورة في المجلة تخضع لقانون حقوق الملكية الفكرية للمجلة.

info.jmbush@bmu.edu.ly

00218913262838

د. صلاح الأمين عبدالله
رئيس تحرير مجلة جامعة بنغازي الحديثة
Dr.salahshalufi@bmu.edu.ly

Impact of Overdispersion on Log-linear Models for Contingency Tables

Dr. Jalal A. Moaiti,* Radi A. Othman, Ali M. Moftah*****

(Members of Teaching Staff, Department of Statistics, Faculty of Arts and Sciences, University of Benghazi, Libya)

Abstract:

Most of the practical studies in social, medical and many other sciences have utilized the use of chi square to study the relationship between many types of categorical data, and it has been widely used. However, the use of chi square test of independence can be affected by overdispersed counts of a contingency table. In this work, the effect of the overdispersion on the log-linear models using Poisson distribution and negative binomial distribution is investigated. These models were applied on (2x2), (3x3) and (4x4) generated contingency tables of overdispersed and Poisson distributions. A simulation study was applied on different conditions of counts average and degree of dispersion. The Poisson distribution was severely affected by the overdispersion of the data. As a solution to this problem, the negative binomial distribution was used and it found that, it performed well to alleviate of this problem in some cases of the simulation stages. However, using the Poisson model on overdispersed counts of a contingency table, can inflate the Type one error of the deviance and the Pearson chi square as well.

Keywords: Poisson regression model, Negative binomial regression model, Log-linear Models, Overdispersion.

تأثير التشتت الزائد على النماذج الخطية اللوغاريتمية لجدول التوافق

* د. جلال عبدالله امعيطي، ** أ. راضي عبدالرحيم عثمان، *** أ. علي محبوب مفتاح

(أعضاء هيئة تدريس بقسم الاحصاء – كلية الآداب والعلوم المرج – جامعة بنغازي - ليبيا)

المخلص :

معظم الدراسات في العلوم الاجتماعية والطبية والعديد من العلوم الأخرى تستخدم على نطاق واسع اختبار مربع كاي (chi square test) لدراسة العلاقة بين البيانات الوصفية، ولكن يمكن أن يتأثر استخدام اختبار مربع كاي للاستقلالية بالتشتت الزائد في جداول التوافق. في هذا الدراسة تم التحقق من تأثير التشتت الزائد (Overdispersion) على النماذج الخطية اللوغاريتمية باستخدام توزيع بواسون وتوزيع ذي الحدين السالب، حيث تم تطبيق هذه النماذج بتوليد جداول التوافق (2×2) و (3×3) و (4×4) في حالتين، حالة تمثل التشتت الزائد والأخرى تمثل توزيعات بواسون. وطبقت دراسة محاكاة على ظروف مختلفة من متوسط التكرارات (counts average) ودرجة التشتت. واتضح أن توزيع بواسون تأثر بشدة من التشتت الزائد بالبيانات. وكحل لهذه المشكلة، تم استخدام توزيع ذي الحدين السلبى واتضح أنه كان جيداً للحد من تأثير مشكلة التشتت الزائد في بعض حالات مراحل المحاكاة. وبالتالي فإن استخدام نموذج بواسون عند وجود التشتت الزائد في لجدول التوافق يمكن أن يؤدي إلى تضخيم خطأ النوع الأول في الانحراف (deviance) وبيرسون مربع كاي (Pearson chi square) أيضاً.

الكلمات المفتاحية: نموذج انحدار بواسون، النماذج الخطية اللوغاريتمية، نموذج انحدار ذي الحدين السالب، التشتت الزائد.

1.Introduction:

Loglinear models (LLM) are using to study the relationships among two or more discrete variables. Often referred to multiway frequency analysis and it is an extension of the familiar chi-square test for independence in two-way contingency tables. LLM may be used to analyze surveys and questionnaires which have complex interrelationships among several questions. Although questionnaires are often analyzed by considering only two questions at a time, this ignores the important of three-way and multiway relationships among the questions. The use of LLM on this type of data is analogous to the use of multiple Poisson regression rather than simple correlations of data.

Brief History: Until the late 1960's, contingency tables (two-way tables formed by cross classifying categorical variables) were typically analyzed by calculating chi-square values to test the hypothesis of independence. When tables consisted of more than two variables, researchers would compute the chi-squares for two-way tables and then again for multiple sub-tables formed from them, in order to determine if associations and/or interactions were taking place among the variables. In the 1970's the analysis of cross-classified data changed quite dramatically with the publication of a series of papers on loglinear models by L.A. Goodman. Many other books appeared around that time building on Goodman's work (Bishop, Finberg & Holland 1975; Haberman 1974). Now researchers were introduced to a wide variety of models that could be fitted to cross-classified data. Thus, the introduction of the loglinear model provided them with a formal and rigorous method for selecting a model or models for describing associations between variables.

Overview: The loglinear model is one of the specialized classes of generalized linear models when the distribution of data is a Poisson-distribution. Loglinear models class is an extension of the two-way contingency table where the conditional relationship between two or more discrete categorical variables is analyzed by taking the natural logarithm of the cell frequencies of a contingency table. Although loglinear models can be used to analyze the relationship between two categorical variables (two-way contingency tables), they are more commonly used to evaluate multiway contingency tables that involve three or more variables. The variables analyzed by log linear models are a response variable which represents the counts, and other explanatory variables. Therefore, loglinear models are demonstrate the individual and interaction effects of the explanatory variables on a response variable.

Basic Strategy and Key Concepts: The basic strategy in loglinear modeling involves fitting models to the observed frequencies in the cross-tabulation of categorical variables. The models can then be represented by a set of expected frequencies that may or may not resemble the observed frequencies. Models will vary in terms of the marginal they fit, and can be described in terms of the constraints they place on the associations or interactions that are present in the data. The pattern of association among variables can be described by a set of odds and by one or more odds ratios derived from them. Once expected frequencies are obtained, we then compare models that are hierarchical to one another and choose a preferred model, which is the most parsimonious model that fits the data. It's important to note that a model is not chosen if it bears no resemblance to the observed data. The choice of a preferred model is

typically based on a formal comparison of goodness-of-fit statistics associated with models that are related hierarchically (models containing higher order terms also implicitly include all lower order terms). Ultimately, the preferred model should distinguish between the pattern of the variables in the data and sampling variability, thus providing a defensible interpretation. Modeling count variables is a common task in economics and the social sciences. The classical Poisson regression model for count data is often of limited use since empirical count data sets usually exhibit overdispersion with the counts. Another more formal way can be used as a solution to this problem, is using Negative Binomial model instead. In this study, the impact of overdispersion on the Poisson model along with the Negative Binomial model will be investigated. SAS Monte Carlo simulation procedure will be used to generate counts of light-tailed (Uniform) and Poisson distributions, and then these models will be fitted to some conditions of contingency tables. The focus will be on the type-I error rate of the deviance and ordinary Pearson chi square goodness of fit of these models.

2. The Problem:

When applying generalized linear models with a known scale as is certainly the case for the binomial and Poisson distributions where ($\emptyset = 1$), subject to certain asymptotic conditions for a well fitting model we would expect:

Residual Deviance \approx Residual degrees of freedom (df).

What if the Residual Deviance \gg Residual df ?

There are two possible scenarios we need to consider:

(i) We may simple have a badly fitting model for one of a number of reasons such as:

- Omitted terms or variables in the linear predictor;
- Incorrect relationship between mean and explanatory variables, *i.e.*

We may have the wrong link function or need to transform one or more explanatory variables;

- Outliers.

Standard model diagnostics allow us to explore these aspects, see Mc-Cullagh and Nelder (1989), Chapter 12.

(ii) The variation may simply be greater than that predicted by model and it is this phenomenon that is described as overdispersion then we have:

- Count data with $Var(y_i) > \mu_i$.
- Proportion data with $Var(y_i) > \mu_i \pi_i (1 - \pi_i)$.

Causes of Overdispersion: There are many different possible causes of overdispersion and in any modeling situation a number of these could be involved.

Some possibilities are:

- Variability of experimental material: this can be thought of as individual variability of the experimental units and may give an additional component of variability which is not accounted for by the basic model.
- Correlation between individual responses.
- Cluster sampling.
- Aggregate level data: the aggregation process can lead to compound distributions.
- Omitted unobserved variables: in some sense the other categories are all special cases of this, but generally in a rather complex way.

In some circumstances the cause of the overdispersion may be apparent from the nature of the data's collection process. Although it should be noted that different explanations of the overdispersion process can lead to the same model, so in general it

is difficult to infer the precise cause, or underlying process which leading to the overdispersion.

Consequences of Overdispersion: When we identify the possible presence of overdispersion, what are the consequences of failing to take it into account? Firstly, the standard errors obtained from the model will be incorrect and may be seriously underestimated and consequently we may incorrectly assess the significance of individual regression parameters. Also, changes in deviance associated with model terms will also be too large and this will lead to the selection of overly complex models. Finally, interpretation of the model will be incorrect and any predictions will be too precise.

It turns out that overdispersion is very common for count data and it typically underestimates variability, standard errors and thus inflated p-values. There are a number of ways of overcoming this limitation, the effectiveness of which depend on the causes of overdispersion. The negative binomial model is useful for accommodating overdispersion, when it is likely caused by clumping (due to the influence of other unmeasured factors) within the response.

3. Relevant Literature:

The negative binomial distribution has been suggested by some as an alternative to the Poisson when there is evidence of “overdispersion” (Paternoster and Brame 1997). Stated loosely for the moment, “overdispersion” implies that there is more variability around the model’s fitted values than is consistent with a Poisson formulation. The negative binomial is proposed as a means to correct for this problem, and some go so far to say that it automatically does so (Osgood 2000). There is a parameter whose estimated value inflates the Poisson dispersion as needed. Regression modeling, broadly construed, has been skeptically examined before by a large number of statisticians and social scientists. For example, in a recent book written for social scientists, (Berk, 2003) unpacks what regression models require and argues that in general they are best suited for descriptive purposes only. (Freedman, 2005) provides a more technical discussion that is no less critical. Morgan and (Winship, 2007) make a case for abandoning conventional regression modeling altogether in social research and suggest a counterfactual approach relying on various kinds of matching strategies.

Logistic regression model with overdispersion will has unbiased maximum-likelihood estimates. However, their standard error will be underestimated and confidence interval becomes narrower than the actual one. As a result, significant test for regression parameters may be unreliable, as it tends to reject null hypothesis of no parameter effect which lead to misleading conclusion. The approach in handling overdispersion was introduced firstly by (Williams, 1982). Williams equates the value of Pearson’s chi-square statistic of the model to its expected value to obtain an optimal value of inflation factor inserted in the weighting matrix of parameter estimation. Previous study on logistic regression with overdispersion carried out by (Kurnia, Saefuddin and Sutisna, 2002) and (Saefuddin, Setiabudi and Achsani, 2011) showed that the parameter estimates were unbiased with small standard errors on

standard logistic regression model. Applying Williams approach to the models, produced different conclusion of parameter effects due to correction to the standard errors of estimates.

Score tests for zero-inflation in Poisson and binomial regression models have been investigated by van den (Broek, 1995) and (Deng & Paul, 2000).

Modeling using extended Poisson process models (EPPMs) was originally developed by (Faddy, 1997), where the construction of discrete probability distributions having very general dispersion properties was described. The Poisson and negative binomial distributions are special cases of this modeling which includes both underdispersion and overdispersion relative to the Poisson, with the negative binomial having the most extreme level of overdispersion within the EPPM family. (Faddy and Smith, 2008) incorporated covariate dependence in the mean via a reparameterization using an approximate form of the mean; and (Faddy and Smith, 2011) extended this to incorporate covariate dependence in the dispersion, this being achieved by a reparameterization using an approximate form of the variance. The supplementary material for (Faddy and Smith, 2011) contained R code illustrating fitting these models. This R code has been extended and generalized to have inputs and outputs more akin to those of a generalized linear model (GLM) as in the R function `glm` and the R function `betareg` (Cribari-Neto and Zeileis 2010, Grün, Kosmidis, and Zeileis 2012).

There are many examples of overdispersed count models in ecology, with important applications ranging from species richness to spatial distributions to parasitism. (O'Hara, 2005) noted the differences between Poisson and negative binomial distributions for species richness, with each being appropriate only when data were simulated from the correct model. (Alexander et al., 2000) used a negative binomial distribution with a spatial model of parasitism. (White and Bennetts, 1996) modeled bird counts with a negative binomial distribution. For trend and abundance estimation for harbor seals, (Frost et al., 1999) and (Small et al., 2003) and (Mathews and Pendleton, 2006) used Poisson regression, (Ver Hoef and Frost, 2003) used an overdispersed Poisson regression, and (Boveng et al., 2003) used negative binomial regression.

Because overdispersion is so common, several models have been developed for these data, including the negative binomial, quasi-Poisson (Wedderburn, 1974), generalized Poisson (Consul, 1989) and zero-inflated (Lambert, 1995) models. Relationships among some of the distributions can be found in (Joe and Zhu, 2005) and (Lord et al., 2005). Despite these developments, the quasi-Poisson and negative binomial models are used most often, largely because they are widely available in software and they generalize easily to the regression case.

4. Methods:

In this Section, brief theoretical review of log-linear models using Poisson and negative binomial distributions will be presented, and these methods will be utilized to generate the data for this study.

4.1. Log-linear model:

Loglinear models (LLM) study the relationships between a discrete response variable (counts) and other explanatory variables. Often it refers to analysis of multiway frequency tables, and it can be considered as an extension of the familiar chi-square test for independence in two-way contingency tables.

Limitations and Assumptions: Since the use of LLM requires few assumptions about population distributions, they are remarkably free of limitations. They may be applied to almost any circumstance in which the variables are (or can be made) discrete. It can even be used to analyze continuous variables which fail to meet distributional assumptions (by collapsing the continuous variables into a few categories).

Three basic assumptions should be considered when using LLM.

1. *Observations are independent from each other.* In practice, this means that each observation comes from different subjects.
2. *All observations are identically distributed.* This means that they are obtained in the same way.
3. *The number of observations is large.* Since LLM make use of large sample approximations, they require large samples. The LLM algorithm begins by taking the natural logarithm of each of the cell frequencies, so empty cells (those with frequencies of zero) are not allowed. LLM appear to be less restrictive than traditional chi-square contingency tests, so rules that are used for those tests may be used for LLM analysis as well.

Fundamental Approach: LLM analysis requires two steps. It is easy to become lost in details of each of these steps, but it is important to keep in mind the overall purpose of each task,

1. *Selecting an appropriate model.* The first step is to find an appropriate model of the data. Several techniques may be used to find an appropriate LLM. One of the most popular is the step-down technique in which complex terms are removed until all terms remaining are significant.
2. *Interpreting the selected model.* Once a model is selected, it must be interpreted. This is the step in which you determine what your data are telling you.

The Notation of Loglinear Models: Consider a two-way table in which the row-variable A has categories (levels) ($i=1, \dots, I$) and the column-variable B has categories ($j=1, \dots, J$).

A multiplicative model that reproduces the cell frequencies (f_{ij}) exactly is

$$m_{ij} = N\alpha_i\beta_j\gamma_{ij}$$

Where $[m_{ij} = E(f_{ij})]$ is the expected frequency of the (i^{th}) row and the (j^{th}) column.

When the (m_{ij}) are estimated using maximum likelihood, the results are denoted (\hat{m}_{ij}).

Also note that:

$$N = \sum_{ij} f_{ij}$$

One aspect of the table that is of interest is whether A and B are independent. This is often tested using the familiar chi-square test. Independence would be established if all (γ_{ij}) were equal to one.

Because of its multiplicative form, the above formula is difficult to work with. However, if we take the logarithm of both sides, we can rewrite it as

$$\ln(m_{ij}) = \theta + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

The $(\lambda's)$ are called *effects*. The superscript indicates the variable(s) and the subscripts refer to the individual categories of those variables. The *order* of an effect is equal to the number of variables in the superscript.

Because this formulation is additive, it is called a *loglinear* model. Because of the logarithms, this model has the added constraint that none of the (m_{ij}) are zero.

Notice that the total number of $\lambda's$ in this model is $[I+I+J+(I \times J)]$ which is greater than the number of cell frequencies (which is $I \times J$). When the number of parameters is greater than or equal to the number of cells, we say the model is *saturated*. A saturated model reproduces the observed frequencies exactly.

By testing whether certain of the $(\lambda's)$ are zero, you can test various interrelationships. For example, to test whether all of the frequencies are equal, you would test whether all first-order and second-order effects (the $\lambda^{A'}$'s, $\lambda^{B'}$'s, and $\lambda^{AB'}$'s) are zero. Testing whether the $(\lambda^{AB'}$'s) are zero would test whether variables A and B are independent (no interaction).

Testing whether the $(\lambda^{A'}$) were zero would test whether the probabilities of the categories of A are equal. As you can see, this model will let you answer many interesting questions about factors A and B .

Goodness of Fit: When dealing with several competing models, the relative quality of each model must be considered. The quality of a model, as measured by its goodness of fit to the data, may be tested using either of two chi-square statistics:

- *The Pearson chi-square statistic*

$$\chi^2 = 2 \sum_{i,j,k} \frac{(f_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}}$$

- *The likelihood-ratio statistic*

$$G^2 = 2 \sum_{i,j,k} f_{ijk} \ln \left(\frac{f_{ijk}}{\hat{m}_{ijk}} \right)$$

Both of these statistics are distributed as a chi-square random variable when N is large and none of the (\hat{m}_{ijk}) are small. If a few of the (\hat{m}_{ijk}) are small, the chi-square approximation is still fairly close. Both of these statistics have $n-p$ degrees of freedom where (n) is the number of cells in the table and (p) is the number of parameters in the model on which the (\hat{m}_{ijk}) are based.

You should understand exactly what these two chi-square statistics are testing. They test whether the terms in the saturated model that are not included in the current model are significantly different from zero.

A word of Caution: the difference between the two (G^2) is distributed as a chi-square only when the more complete model fits the data adequately. That means that the (G^2) of the larger model should be nonsignificant. Because of the additively property of (G^2), it is very popular in LLM.

Again, this additively property does not hold for the Pearson chi-square statistic. Why do we even compute this value? Why not just use the likelihood ratio statistic?

First, some studies indicate that the Pearson goodness of fit test may be more accurate. Second, since both of these are asymptotic tests, you can be more comfortable with small sample results when both tests lead to the same conclusion.

Model Selection Techniques: Since your first task in the analysis is to find a well-fitting model with as few terms as possible, you must adopt some method to limit the number of models you consider. The program provides several possible model selection methods. The final model will result from applying several of these techniques to your data.

4.2. The Poisson Model:

Poisson regression is similar to regular multiple regressions models except that the dependent (Y) variable is an observed count that follows the Poisson distribution. Thus, the possible values of Y are the nonnegative integers: 0, 1, 2, 3, and so on. It is assumed that large counts are rare. Hence, Poisson regression is similar to logistic regression, which also has a discrete response variable. However, the response is not limited to specific values as it is in logistic regression ($y=0$ or 1).

Most books on regression analysis briefly discuss Poisson regression, such as (Cameron and Trivedi, 1998).

The Poisson distribution: The Poisson distribution can be written,

$$\Pr(Y = y|\mu) = \frac{e^{-\mu}\mu^y}{y!} \quad (y = 0, 1, 2, \dots)$$

Notice that the Poisson distribution is specified with a single parameter(μ). This is the mean incidence rate of a rare event per unit of *exposure*. Exposure may be time, space, distance, area, volume, or population size. Because exposure is often a period of time, we use the symbol (t) to represent the exposure. When no exposure value is given, it is assumed to be one.

The parameter (μ) may be interpreted as the risk of a new occurrence of the event during a specified exposure period, (t). The probability of (y) events is then given by

$$\Pr(Y = y|\mu, t) = \frac{e^{-\mu t}\mu t^y}{y!} \quad (y = 0, 1, 2, \dots)$$

The Poisson distribution has the property that its mean and variance are equal.

The Poisson Regression Model: In Poisson regression, we suppose that the Poisson incidence rate (μ) is determined by a set of (k) explanatory variables (the X 's). The expression relating these quantities is

$$\mu = \text{texp}(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$$

Note that often [$x_1 = 1$], and (β_1) is called the *intercept*. The regression coefficients [$\beta_1, \beta_2, \dots, \beta_k$] are unknown parameters that are estimated from a set of data. Their estimates are labeled [b_1, b_2, \dots, b_k] Using this notation, the fundamental Poisson regression model for an observation (t) is written as

$$\Pr(Y = y | \mu_i, t_i) = \frac{e^{-\mu_i t_i} (\mu_i t_i)^y}{y_i!}$$

Where

$$\begin{aligned} \mu_i &= t_i \mu(X_i' \beta) \\ &= t_i \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}) \end{aligned}$$

That is, for a given set of values of the explanatory variables, the outcome follows the Poisson distribution.

Solution by Maximum Likelihood Estimation: The regression coefficients are estimated using the method of maximum likelihood. The logarithm of the likelihood function is,

$$\ln[\mathcal{L}(y, \beta)] = \sum_{i=1}^n y_i \ln[t_i \mu(X_i' \beta)] - \sum_{i=1}^n t_i \mu(X_i' \beta) - \sum_{i=1}^n \ln(y_i!)$$

Note that some statistical packages ignore the last term since it does not involve the regression parameters. This will make their calculated log-likelihoods different from ours.

The likelihood equations may be formed by taking the derivatives with respect to each regression coefficient and setting the result equal to zero. Doing this leads to a set of nonlinear equations that admits no closed-form solution. Thus, an iterative algorithm must be used to find the set of regression coefficients that maximum the loglikelihood. Using the method of iteratively reweighted least squares, a solution may be found in five or six iterations.

Distribution of the MLE's: Applying the usual maximum likelihood theory, the asymptotic distribution of the maximum likelihood estimates (MLE's) is multivariate normal. That is,

$$\hat{\beta} \sim N(\beta, \hat{\beta} V_{\hat{\beta}}), \text{ where } V_{\hat{\beta}} = (\sum_{i=1}^n \mu_i x_i x_i')^{-1}$$

Remember that in the Poisson model the mean and the variance are equal. In practice, the data almost always reject this restriction. Usually, the variance is greater than the mean and this situation called *overdispersion*. The increase in variance is represented in the model by a constant multiple of the variance-covariance matrix. That is, we use,

$$V_{\hat{\beta}} = \phi (\sum_{i=1}^n \mu_i x_i x_i')^{-1}, \text{ where } (\phi) \text{ is estimated using}$$

$$\hat{\phi} = \frac{1}{n-k} \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$$

Goodness of Fit Tests: Overall performance of the model is measured by two chi-square tests,

$$P_P = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, \text{ and the deviance, or } G^2 \text{ statistic,}$$

$$D_P = \sum_{i=1}^n \left\{ y_i \ln \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right\}$$

Both of these statistics are approximately chi-square distributed with $n - k$ degrees of freedom. When a test is rejected, there is a significant lack of fit. When a test is not rejected, there is no evidence of lack of fit.

The Pearson statistic is only chi-square distributed when you are analyzing grouped data, so if you are not using a frequency variable, you should not use the Pearson statistic as a goodness of fit test. The Pearson statistic may be used as a test of overdispersion.

Deviance: The deviance is twice the difference between the maximum achievable log-likelihood and the log-likelihood of the fitted model. In multiple regression models and under the assumption of normality, the deviance is the residual sum of squares. In the case of Poisson regression, the deviance is a generalization of the sum of squares and it is as same as the Pearson chi square and G^2 tests. The formula for the deviance is

$$D(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2\{LL_{\mathbf{y}} - LL_{\hat{\boldsymbol{\mu}}}\}$$

4.3. Negative binomial model:

The Negative Binomial Distribution: The Poisson distribution may be generalized by including a gamma noise variable which has a mean of (1) and a scale parameter of (ν). The Poisson-gamma mixture (negative binomial) distribution that results is

$$\Pr(\mathbf{Y} = \mathbf{y}_i | \boldsymbol{\mu}_i, \boldsymbol{\alpha}) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i} \right)^{y_i}$$

Where

$$\boldsymbol{\mu}_i = \mathbf{t}_i \boldsymbol{\mu}, \quad \boldsymbol{\alpha} = \frac{1}{\nu}$$

The parameter (μ) is the mean incidence rate of (y) per unit of exposure. Exposure may be time, space, distance, area, volume, or population size. Because exposure is often a period of time we use the symbol (t_i) to represent the exposure for a particular observation, and when no exposure given it is assumed to be one.

The parameter (μ) may be interpreted as the risk of a new occurrence of the event during a specified exposure period, (t).

The results below make use of the following relationship derived from the definition of the gamma function

$$\ln \left(\frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})} \right) = \sum_{j=0}^{y_i-1} \ln(j + \alpha^{-1})$$

The Negative Binomial Regression Model: In negative binomial regression, the mean of (y) is determined by the exposure time (t) and a set of (k) regressor variables (the x 's). The expression relating these quantities is

$$\mu_i = \exp(\ln(t_i) + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$$

Often $\beta_1 \equiv 1$, in which case (β_1) is called the *intercept*. The regression coefficients ($\beta_1, \beta_2, \dots, \beta_k$) are unknown parameters that are estimated from a set of data. Their estimates are symbolized as (b_1, b_2, \dots, b_k).

Using this notation, the fundamental negative binomial regression model for an observation (i) is written as

$$\Pr(Y = y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha\mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i}$$

Solution by Maximum Likelihood Estimation: The regression coefficients are estimated using the method of maximum likelihood. (Cameron 2013, page 81) gives the logarithm of the likelihood function as

$$\mathcal{L} = \sum_{i=1}^n \{ \ln[\Gamma(y_i + \alpha^{-1})] - \ln[\Gamma(\alpha^{-1})] - \ln[\Gamma(y_i + 1)] - \alpha^{-1} \ln[1 + \alpha\mu_i] - y_i \ln[1 + \alpha\mu_i] + y_i \ln(\alpha) + y_i \ln(\mu_i) \}$$

Rearranging gives

$$\mathcal{L} = \sum_{i=1}^n \left\{ \left(\sum_{j=0}^{y_i-1} \ln(j + \alpha^{-1}) \right) - (\ln(\Gamma(y_i + 1))) - (y_i + \alpha^{-1}) \ln(1 + \alpha\mu_i) + y_i \ln(\mu_i) + y_i \ln(\alpha) \right\}$$

The first derivatives of \mathcal{L} were given by (Cameron, 2013) and (Lawless, 1987) as

$$\frac{\partial \mathcal{L}}{\partial \beta_j} = \sum_{i=1}^n \frac{x_{ij}(y_i + \mu_i)}{1 + \alpha\mu_i}, \quad j = 1, 2, \dots, k$$

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \sum_{i=1}^n \left\{ \alpha^{-2} \left(\ln(1 + \alpha\mu_i) - \sum_{j=0}^{y_i-1} \frac{1}{j + \alpha^{-1}} \right) + \frac{y_i - \mu_i}{\alpha(1 + \alpha\mu_i)} \right\}$$

$$\frac{-\partial^2 \mathcal{L}}{\partial \beta_r \partial \beta_s} = \sum_{i=1}^n \frac{\mu_i(1 + \alpha y_i) x_{ir} x_{is}}{\alpha(1 + \alpha\mu_i)^2}, \quad r, s = 1, 2, \dots, k$$

$$\frac{-\partial^2 \mathcal{L}}{\partial \beta_r \partial \alpha} = \sum_{i=1}^n \frac{\mu_i(y_i + \mu_i) x_{ir}}{(1 + \alpha\mu_i)^2}, \quad r = 1, 2, \dots, k$$

$$\frac{-\partial^2 \mathcal{L}}{\partial \alpha^2} = \sum_{i=1}^n \left\{ \sum_{j=0}^{y_i-1} \left(\frac{j}{j + \alpha^{-1}} \right)^2 + 2\alpha^{-3} \ln(1 + \alpha\mu_i) - \frac{2\alpha^2 \mu_i}{1 + \alpha\mu_i} - \frac{(y_i + \alpha^{-1}) \mu_i^2}{(1 + \alpha\mu_i)^2} \right\}$$

Equating the gradients to zero gives the following set of likelihood equations

$$\sum_{i=1}^n \frac{x_{ij}(y_i - \mu_i)}{1 + \alpha\mu_i} = 0, \quad j = 1, 2, \dots, k$$

$$\sum_{i=1}^n \left\{ \alpha^{-2} (\ln(1 + \alpha\mu_i) - \sum_{j=0}^{y_i-1} \frac{1}{j + \alpha^{-1}}) + \frac{(y_i - \mu_i)}{\alpha(1 + \alpha\mu_i)} \right\} = 0$$

Distribution of the MLE's: (Cameron, 2013) gives the asymptotic distribution of the maximum likelihood estimates as multivariate normal as follows,

$$\begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix} \sim N \left[\begin{bmatrix} \beta \\ \alpha \end{bmatrix}, \begin{bmatrix} V(\hat{\beta}) & Cov(\hat{\beta}, \hat{\alpha}) \\ Cov(\hat{\beta}, \hat{\alpha}) & V(\hat{\alpha}) \end{bmatrix} \right]$$

Where

$$V(\hat{\beta}) = \left[\sum_{i=1}^n \frac{\mu_i}{1+\alpha\mu_i} x_i x_i' \right]^{-1}$$

$$V(\hat{\alpha}) = \sum_{i=1}^n \left\{ \alpha^{-4} (\ln(1 + \alpha\mu_i) - \sum_{j=0}^{y_i-1} \frac{1}{j+\alpha-1})^2 + \frac{\mu_i}{\alpha^2(1+\alpha\mu_i)} \right\}^{-1}$$

$$Cov(\hat{\beta}, \hat{\alpha}) = \mathbf{0}$$

Deviance: The deviance is twice the difference between the maximum achievable log-likelihood and the log-likelihood of the fitted model. In multiple regression models under the assumption of normality, the deviance is the residual sum of squares. In the case of negative binomial regression, the deviance is a generalization of the sum of squares. The maximum possible log likelihood is computed by replacing (μ_i) with (y_i) in the likelihood formula. Thus, we have

$$D = 2[\mathcal{L}(y_i) - \mathcal{L}(\mu_i)]$$

$$= 2 \sum_{i=1}^n \left\{ y_i \ln\left(\frac{y_i}{\mu_i}\right) - (y_i + \alpha^{-1}) \ln\left(\frac{1+\alpha y_i^2}{1+\alpha\mu_i}\right) \right\}$$

4.4. Procedures:

Our data in this work are generated using (SAS) program, and they represent count observations of two different discrete distributions. SAS Monte Carlo simulation will be used over large number of replications (10,000). In each replication the program will generate a contingency tables of (2x2), (3x3) and (4x4) counts. The counts will be randomly generated such that the rows and the columns will be independent. The (10000) replications will be applied on some selected distributions, counts of Poisson distribution with specified average and counts of discrete uniform distribution with specified average. Poisson distribution will represent the regular case and the uniform distribution will represent the overdispersion case (light tailed distribution). The process will be replicated on some selected count averages of both distributions (Count averages: 5, 50, 100, 200 and 400).

In each replication, the program will fit a log-linear model such that the rows and the columns are independent. The empirical p-values (type I error) of the model fit's test (from each replication) will be collected and changed to indicator variable (0) or (1). The program will record the value (1) in case of rejecting the null hypothesis (the model is fit) and (0) in case of acceptance. These values will be collected in a data set during (10000) replications stage. Finally, the proportion of rejection will be calculated over all (10000) replications. The calculated proportion will represent the probability of type one error since our generated counts are random in each replication. This process will be repeated for all distribution cases and count average cases.

We will apply the log-linear model using both Poisson distribution and the negative binomial distribution on the generated counts from Poisson and Uniform distribution. The performance of each model will be evaluated according to the probability of type I error for all simulation stages.

5. Results:

The following tables shows the probabilities of type one error for each model over all cases:

Empirical Type-I error for each case over 10000 replications (2x2 Contingency Tables)

Count average	Counts generated from Poisson distribution (Regular case)		Counts generated from uniform distribution (Overdispersion case)	
	Poisson Model	Neg.bin Model	Poisson Model	Neg.bin Model
5	0.0600	0.0259	0.0270	0.0120
50	0.0510	0.0530	0.5580	0.3110
100	0.0520	0.0490	0.6920	0.4787
200	0.0480	0.0390	0.7870	0.3790
400	0.0550	0.0460	0.8531	0.4490

In case of counts generated from Poisson distribution the Poisson model gave probabilities of type one error close to 0.05 for most count averages. as well in case negative binomial model gave probabilities of type one error close to to 0.05. except one case (count average = 5) the probability of type one error equal to 0.02

In case of counts generated from light tailed distribution (Uniform), turns out that both models gave large probabilities of type one error for all count averages. But whenever the sample increased then the negative binomial model has less probability of type one error compare to Poisson model.

Empirical Type-I error for each case over 10000 replications (3x3 Contingency Tables)

Count average	Counts generated from Poisson distribution (Regular case)		Counts generated from uniform distribution (Overdispersion case)	
	Poisson Model	Neg.bin Model	Poisson Model	Neg.bin Model
5	0.0399	0.0180	0.0010	0.0002
50	0.0487	0.0010	0.9290	0.1220
100	0.0499	0.0220	0.9800	0.0780
200	0.052	0.0302	0.9980	0.0430
400	0.0469	0.0620	1.0000	0.0320

In case of counts generated from Poisson distribution the Poisson model gave probabilities of type one error close to 0.05 for all count averages which is reasonable result. By comparing the results (Poisson model and negative binomial model),The negative binomial model gave small probability of type one for small samples size (0.0181for count average = 5), (0.0010, for count average = 50), and(0.0220 for

count average = 100) and it became close to the Poisson model results with increasing sample size.

in case of counts generated from light tailed distribution (Uniform), the Poisson model gave larger probabilities of type one error in all cases, Increases by increasing sample size (large count average). And The negative binomial model gave large probability only in case of sample (count average = 50), and It gave a small and satisfying probabilities with increase the sample size (large count averages).

***Probability of Type-I error for each case over 10000 replications
(4x4 Contingency Tables)***

Count average	Counts generated from Poisson distribution (Regular case)		Counts generated from uniform distribution (Overdispersion case)	
	Poisson Model	Neg.bin Model	Poisson Model	Neg.bin Model
5	0.0511	0.3610	0.6561	0.4923
50	0.0466	0.2039	0.9998	0.0123
100	0.0572	0.1513	1.0000	0.0042
200	0.0497	0.0837	1.0000	0.0040
400	0.0500	0.0348	1.0000	0.0036

In case of counts generated from Poisson distribution the Poisson model gave probabilities of type one error close to 0.05 for all count averages which is reasonable result. The negative binomial model gave large probability of type one for small sample size (0.361 for count average = 5) , (0.2039, for count average = 50) and (0.1513 for count average = 100) and it became smaller as the sample size increase.

In case of counts generated from light tailed distribution (Uniform), the Poisson model gave larger probabilities of type one error in all cases, especially when the sample size is large (large count average). The negative binomial model gave large probability only in case of small sample (count average = 5), and it gave smaller probabilities as well as the sample increased (large count averages).

6. Summary and Conclusions:

This paper introduced the Poisson and negative binomial models as appropriate techniques to describe two-Way Contingency Tables (or a count data) of a response variable.

Through the above results, the study concluded that the Poisson model is severely affected by the existence of overdispersion problem, and gave large probability of type one error even if the count average is small. The negative binomial model is a good To reduce the problem of overdispersion in data of contingency tables, especially when the count average is large enough.

So for an overdispersion data, the negative binomial model is better than the Poisson model.

References:

- 1- Alexander, N., R. Moyeed, & J. Stander. 2000. *Spatial modeling of individual-level parasite counts using the negative binomial distribution*. *Biostatistics* 1:453–463.
- 2- Berk, R. A. (2003) *Regression Analysis: A Constructive Critique*. Newbury Park, CA.: Sage Publications.
- 3- Bishop, Y. M. M., S. E. Fienberg, & P. W. Holland (1975). *Discrete Multivariate Analysis*. Cambridge, MA: MIT Press.
- 4- Boveng, P.L., J. L. BENGTON, D.E. WITHROW, J.C. CESARONE, M.A. SIMPKINS, K.J. FROST AND J. BURNS. 2003. *The abundance of harbor seals in the Gulf of Alaska*. *Marine Mammal Science* 19:111–127.
- 5- Broek, V. J. (1995). *A score test for zero inflation in a Poisson distribution*. *Biometrics*, 51, 738–743.
- 6- Cameron, C., and P. Trivedi. *Regression Analysis of Count Data*. New York: Cambridge University Press, 1998.
- 7- Cameron AC, Trivedi PK (2013). *Regression Analysis of Count Data*. 2nd edition. Cambridge University Press.
- 8- Consul, 1989. *Generalized Poisson Distribution: Properties and Application*. New York Marcel Dekker.
- 9- Cook RD, Weisberg S (1999) *Applied regression including computing and graphics*. Wiley, New York, NY.
- 10- Cribari-Neto F, Zeileis A (2010). “Beta Regression in R.” *Journal of Statistical Software*, 34 (2), 1–24. doi:10.18637/jss.v034.i02.
- 11- D. Deng & S. R. Paul (2000). *Score tests for zero-inflation in generalized linear models*. *The Canadian Journal of Statistics*, 28, 563–570.
- 12- Faddy MJ (1997). *Extended Poisson Process Modeling and Analysis of Count Data*. *Biometrical Journal*, 39(4), 431–440. doi:10.1002/bimj.4710390405.
- 13- Faddy MJ, Smith DM (2008). *Extended Poisson Process Modeling of Dilution Series*. *Journal of Applied Statistics*, 35(4), 461–471. doi:10.1111/bimj.1467-9876.2008000622.
- 14- Faddy MJ, Smith DM (2011). *Analysis of Count Data with Covariate Dependence in Both Mean and Variance*. *Journal of Applied Statistics*, 38(12), 2683–2694. doi:10.1080/02664763.2011.567250
- 15- Freedman DA (2005). *Statistical models: theory and practice*. Cambridge University Press, Cambridge.
- 16- Freund JE (1971) *Mathematical statistics*, 2nd edn. Prentice Hall, New York.
- 17- Greene WH (2003) *Econometric analysis*, 5th edn. Prentice Hall, New York.
- 18- Grün B, Kosmidis I, Zeileis A (2012). *Extended Beta Regression in R: Shaken, Stirred, Mixed, and Partitioned*. *Journal of Statistical Software*, 48(11), 1–25. doi: 10.18637/jss.v048.i11.
- 19- Haberman, S. J. 1974b. *Log-linear models for frequency tables with ordered classifications*. *Biometrics* 36: 589–600.
- 20- Joe, H., Zhu, R. 2005. *Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution*. *Biometrical Journal*. 47: 219–229.
- 21- Kurnia, A., A. Saefuddin and E. Sutisna. *Overdispersi dalam Regresi Logistic*. Proc. of Seminar Nasional Statistika, Bogor. 2002: 11–17.
- 22- Lambert, D. 1992. *Zero-inflated Poisson regression, with an application to defects in manufacturing*. *Technometrics* 34: 1–14.

- 23- Lambert, D. & K. Roeder (1995). *Overdispersion diagnostics for general-ized linear models*. *Journal of the American Statistical Association*, 95,1225{1237.
- 24- Lawless, J.E (1987). *Regression methods for Poisson process data*. J. Arne,: Statist. Assoc., Sept. 1987.
- 25- Lord, D., Guikema, S.D., Geedipally, S.R. 2008. *Application of the Conway–Maxwell Poisson generalized linear model for analyzing motor vehicle crashes*. *Accident Analysis & Prevention*. 40(3): 1123–1134.
- 26- Manski CF (1990) *Nonparametric bounds on treatment effects*. *Am Econ Rev Papers Proc* 80:319–323.
- 27- Mathews, E.A., & Pendleton, G.W., 2006. *Declining trends in harbor seal (Phoca vitulina) numbers at glacial ice and terrestrial haulouts in Glacier Bay National Park*. 1992– 2002: *Marine Mammal Science*, v. 22, no. 1, p. 167–189.
- 28- McCullagh P, Nelder JA (1989) *Generalized linear models*, 2nd edn. Chapman and Hall, New York
- 29- Morgan SL, Winship C (2007). *Counterfactuals and causal inference: methods and principle for social research*. Cambridge University Press, Cambridge.
- 30- Nagin DS (2005). *Group-based modeling of development*. Harvard University Press, Cambridge, MA.
- 31- O'Hara, R.B. (2005). *Species richness estimators: how many species can dance on the head of a pin?* *Jornal Animal Ecology* 74, 375–386.
- 32- Osgood W (2000). *Poisson-based regression analysis of aggregate crime rates*. *J Quant Criminol* 16:21–43.
- 33- Osgood DW, Chambers JM (2000). *Social disorganization outside the metropolis: an analysis of rural youth violence*. *Criminology* 38:81–115.
- 34- Paternoster R, Brame R, Bachman R, Sherman L (1997) Do fair procedures matter? *The effect of procedural justice on spouse assault*. *Law Soc Rev* 31:163–204.
- 35- Saefuddin, N.A. Setabudi & N.A. Achsani. *The Effect of Overdispersion on Regression Based Decision With Application to Chum Analysis on Indonesian Mobile Phone*, Vol.60 No.4,602-610, 2011.
- 36- Small, R.J., Pendleton, G.W., & Pitcher, K.W., 2003, Trends in abundance of Alaska harbor seals, 1983-2002: *Marine Mammal Science*, v. 19, p. 344–362.
- 37- Ver Hoef, J. M., & P. L. Boveng. 2007. *Quasi-Poisson vs. negative binomial regression: How should we model overdispersed count data?* *Ecology* 88:2766–2772.
- 38- Wedderburn, R. (1974). *Quasi-likelihood functions, generalized linear models and the gauss-newton method*. *Biometrika*, 61, 439{47.
- 39- White, G. C., and R. E. Bennetts. 1996. *Analysis of frequency count data using the negative binomial distribution*. *Ecology* 77:2549–2557.
- 40- Williams, D. (1982). Extra-binomial variation in logistic linear models. *Applied Statistics*, 31, 144{148.
- 41- Williams, D. (1996). *Overdispersion in logistic-linear models*. In B. Morgan (Ed.), *Statistics in Toxicology*. Oxford: Clarendon Press.