Journal of Biometrics & Biostatistics
ISSN: 2155-6180
Bohning and Viwatwongkasem, J Biom Biostat 2015, 6:2
DOI: 10.4172/2155-6180.1000231

Research Article | Open Access

A Miscellaneous Note on the Equivalence of Two Poisson Likelihoods

Bohning D[1]* and Viwatwongkasem C[2]

[1]*Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ, UK*
[2]*Department of Biostatistics, Faculty of Public Health Mahidol University, Bangkok, Thailand*

Abstract

This note shows that the concept of an offset, frequently introduced in Poisson regression models to cope with rate-type data, can be simply treated with a regular Poisson regression model. Hence Poisson regression models requiring an offset can be fitted with ordinary Poisson regression models. Some illustrations are provided and it is discussed how this result came about.

Keywords: Poisson likelihood; Offset; Log-link

Introduction

The paper at hand has anecdotal character and tells a story what can happen in the life of a university professor. An exam question was posted involving a Poisson regression model with an offset term. Most students found the solution that was also provided by the professor as the official solution (to the exam office), except one student who ignored the offset term and treated it as failures. The solution was marked initially as false as one would evidently not work along this pathway. However, as was correctly claimed by the student, the solution he/she provided was entirely identical with the solutions by his/her colleagues that were marked as correct. The professor sat down and worked out what was going on, and after two days, not without interruptions of the typical university daily life, it was found that also the unorthodox solution has a right to live as both likelihoods could be shown to identical. The purpose of this note is to share this experience.

We consider the data constellation as provided in Table. This is done entirely for illustrative reasons and does not provide any limitations on the generality of the major finding. Here a count (number of events, say) Y_i is considered for a binary exposure E_i and observation-time at risk T_i where i indicates the unit and goes from 1 to n. The person-time T_i is considered non-random

Whereas Y_i is a discrete random variable. This situation is commonly denoted as a rate-type problem with rate

$\lambda_i = E(Y_i)/T_i$.

This rate leads to a natural modelling as follows

$$\log E(Y_i) = \log T_i + \log \lambda_i \quad (1)$$

The term $\log T_i$ is a covariate with known coefficient and called offset whereas $\lambda_i = exp(\beta' z_i)$ contains the linear predictor with covariate vector z_i for the i[th] unit. For the situation of Table 1, with $z_i = (1, E_i)'$ the model would simply be

$$\log E(Y_i) = \log T_i + \beta_0 + \beta_1 E_i,$$

where β_1 is the log-risk ratio, usually the parameter of interest [1]. If one assumes that Y_i is Poisson with mean $E(Y_i)$ given by (1), then the associated log-likelihood is given by

$$log\ L_1 = \sum_i [Y_i log\ \lambda_i - T_i \lambda_i], \quad (2)$$

where we have ignored parameter independent terms. Note that the full log-likelihood (including parameter independent parts) is $\log L_1 + Y_i \log T_i - \log(Y_i!)$. Many packages exist that can fit offset models such as (1). For the data of Table 1 we use STATA (2013) to yield the output given in Figure 1 [2].

A Second Problem

Consider a second problem where we observe, for each unit i, a binary variate X_{ij} where j=1,…, T_i . In other words, X_{ij}=1 if the event of interest occurs and zero otherwise, for j=1, …, T_i. Let, for unit i, $Y_i = \sum_{j=1}^{T_i} X_{ij}$ denote the count of positives and $T_i - Y_i$ are the number of zeros. Let p_{ij} denote the probability for an event in the i[th] subject at the j[th] occasion. Then the clustered likelihood

$$\prod_{i=1}^{n} \prod_{j=1}^{T_i} p_{ij}^{X_{ij}} (1-p_{ij})^{1-X_{ij}} \quad 1\ j=1$$

arises, assuming independence over subjects and occasions. This simplifies further if the event probability is constant over occasion:

$$\prod_{i=1}^{n} p_i^{Y_i} (1-p_i)^{T_i - Y_i}.$$

Let us make the unconventional assumption, purely for the mathematical convenience of showing the equivalence, that X_{ij} follows a Poisson distribution with mean λ_i. In other words, we assume that

$$\log E(X_{ij}) = \log \lambda_i \quad (3)$$

unit *i*	*Yi*	*Ei*	*Ti*
1	2	1	3
2	4	1	5
3	6	1	7
4	3	0	4
5	4	0	5
6	1	0	3

Table 1: Hypothetical data for six units arising from a cohort study with response count Y, binary exposure E and person-time T.

***Corresponding author:** Bohning D, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ, UK, Tel: 00 44 (0) 78 7951 9129; Fax: 00 44 (0) 23 8059 3216; E-mail: d.a.bohning@soton.ac.uk

Received February 05, 2015; **Accepted** June 04, 2015; **Published** June 11, 2015

Citation: Bohning D, Viwatwongkasem C (2015) A Miscellaneous Note on the Equivalence of Two Poisson Likelihoods. J Biom Biostat 6: 231. doi:10.4172/2155-6180.1000231

```
. poisson Y E, exposure(T)

Iteration 0:   log likelihood =  -9.336025
Iteration 1:   log likelihood = -9.3360249

Poisson regression                                Number of obs   =          6
                                                  LR chi2(1)      =       0.16
                                                  Prob > chi2     =     0.6880
Log likelihood = -9.3360249                       Pseudo R2       =     0.0086

------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           E |   .1823215   .4564355     0.40   0.690    -.7122755    1.076919
       _cons |  -.4054651   .3535534    -1.15   0.251    -1.098417    .2874868
       ln(T) |          1  (exposure)
------------------------------------------------------------------------------
```

Figure 1: Regression output for model (1) using the package STATA.

```
. poisson X E_long [fweight = frequencies]

Iteration 0:   log likelihood = -25.921443
Iteration 1:   log likelihood = -25.921443

Poisson regression                                Number of obs   =         27
                                                  LR chi2(1)      =       0.16
                                                  Prob > chi2     =     0.6880
Log likelihood =  25.921443                       Pseudo R2       =     0.0031

------------------------------------------------------------------------------
           X |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      E_long |   .1823216   .4564355     0.40   0.690    -.7122755    1.076919
       _cons |  -.4054651   .3535534    -1.15   0.251    -1.098417    .2874868
------------------------------------------------------------------------------
```

Figure 2: Regression output based on model (3) using the package STATA.

unit *i*	*Ei*	*Xij*	*fi*
1	1	1	2
2	1	1	4
3	1	1	6
4	0	1	3
5	0	1	4
6	0	1	1
1	1	0	1
2	1	0	1
3	1	0	1
4	0	0	1
5	0	0	1
6	0	0	2

Table 2: The data of Table 1 reorganized: for the first set of n observations: let for each unit i f_i=Y_i=#{X_{ij}=1} denote the count of 1s in the binary set X_{i1}, · · · ,X_iT_i and for the second set of n observations: let for each unit i f_i=T_i-Y_i=#{X_{ij}=0} denote the number of 0s.

for i =, …, n and j=1, … , T_i , and X_{ij}~P o(λ_i). In particular, this implies P(X_{ij}=1)=p_i=exp(-λ_i)λ_i and P (X_{ij}=0)=1-p_i=exp(-λ_i). Then, the following associated likelihood (Corresponding also to the full likelihood) occurs:

$$L_2 = \prod_i [\exp(-\lambda_i)\lambda_i]^{Y_i} \times [\exp(-\lambda_i)]^{(T_i - Y_i)} \quad . \qquad (4)$$

Model (3) can be fitted by regressing a 2n vector of n 1s and n 0s, (1,…, 1, 0, … , 0)i , on the 2n vector(s) of associated covariates, for example $(E_1, \ldots, E_n, E_1, \ldots, E_n)'$ in the situation of Table 2, using as frequency weight vector the 2n-vector $f = (Y_1, \ldots, Y_n, T_1 - Y_1, \ldots, T_n - Y_n)'$. This can be done easily in STATA and yields the output in Figure 2. Note that the model (3) does not involve any offset. It is clear from Figures 1 and 2 that the results of the analysis for model (1) and (3) are identical. We show in the following that this is not accidental (as it might have only occurred in this special data set) but is far more general in nature.

Result

We have the following result:

Theorem 1:

log L_1=log L_2.

Proof: We start with the clustered log-likelihood (4) and show that it is identical to the log-likelihood (2):

$$\log L_2 = \sum_i Y_i \log \lambda_i - Y_i \lambda_i - (T_i - Y_i)\lambda_i$$

$$= \sum_i Y_i \log \lambda_i - T_i \lambda_i = \log L_1$$

which ends the proof.

Note that this result, although mathematical almost trivial, is quite general and does not depend in any way on the form of linear predictor. Note that the full (including all terms) log-likelihoods are not identical as also Figures 1 and 2 indicate. Typically, the full log-likelihood corresponding to model (1) will be larger, assuming $Y_i \le T_i$, since

$Y_i \log T_i \ge Y_i \log Y_i \ge \log(Y_i!)$

for i=1, … , n.

- We would like to mention that the result does not require Ti to be an integer. The log-likelihood remains well-defined even if T_i is of real value as it might occur with rate data or otherwise.

All that is required is that $T_i \geq Y_i$ which is only a question of scaling for T.

- The result does not require a special form of linear predictor. However, it is does not generalized beyond the log-link, typical for Poisson regression or log-linear modelling.

Discussion

The result might have more curiosity than impact, but is interesting in it. It means, for example, that we can fit offset models without paying attention to the special offset variate in (1), but simply use the conventional model

$$\log E(X_{ij}) = \log \lambda_i, \qquad (5)$$

Where, for each unit i, $X_{ij}=1$ exactly Y_i -times and T_i-Y_i times otherwise. The question arises how this result can be used. We see at least two applications:

- It can be used to fit Poisson-offset models in packages that do not provide the offset-option.
- It may be used to check the computational correctness of the offset option if the latter is available.

However, the most important take-home message might be that one needs to be careful when to decide about the correctness of a solution provided by an unorthodox thinking student. Finally, we wish to mention that the student received the full mark despite remaining doubts that he/she fully understands the depth of the equivalence.

References

1. Cameron AC, Trivedi PK (1998) Regression Analysis of Count Data. Cambridge, Cambridge University Press.
2. StataCorp (2013) Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.