Computer Methods and Programs in Biomedicine 68 (2002) 205–214

Computer Methods and Programs in Biomedicine

www.elsevier.com/locate/cmpb

Calculating adjusted R^2 measures for Poisson regression models

Martina Mittlböck *

Section of Clinical Biometrics, Department of Medical Computer Sciences, University of Vienna, Spitalgasse 23, 1090 Vienna, Austria

Received 6 March 2000; received in revised form 2 April 2001; accepted 2 April 2001

Abstract

In regression models not only the parameter estimates and significances of explanatory variables are of interest, but also the degree to which variation in the dependent variable can be explained by covariates. In recent publications, an R^2 measure based on deviance was recommended for Poisson regression models, one of the most frequently used modelling tools in epidemiological studies. However, when sample size is small relative to the number of covariates in the model, simple R^2 measures may be seriously inflated and may need to be adjusted according to the number of covariates in the model. We present a SAS-macro that calculates adjustments for the R^2 measures in Poisson regression models based on log-likelihood and on sums of squares. The proposed measures are applied to real data sets and their performance is discussed.

Keywords: Adjusted R^2; Poisson regression; Deviance residuals; SAS-macro

1. Introduction

In recent years, several papers have dealt with R^2 measures for Poisson regression models [1–3]. The main interest has often been in the behaviour of several R^2 measures. Poisson regression models are often used to screen for prognostic factors in studies with small to moderate sample size and many covariates. In such situations, unadjusted R^2 measures may give substantially inflated values, jeopardizing the ability to draw valid interpretations. R^2 values of 30% or higher can easily be reached, even when no association between independent and dependent variables exists at all.

In linear regression models, the use of an adjusted R^2 measure is well established whereas, for Poisson regression models, some corrections have been recently proposed. We have written a SAS-macro [4] for calculating adjusted R^2 measures for Poisson regression suggested by Waldhör et al. [2] and Mittlböck and Waldhör [3]. All of the adjusted measures consider the number of fitted parameters in the model.

In Section 2, the R^2 measures and their adjusted versions are described, and in Section 3 two examples are given on how to use the SAS-macro

* Tel.: +43-1-404002276; Fax: +43-1-404002278.

E-mail address: martina.mittlboeck@akh-wien.ac.at (M. Mittlböck).

0169-2607/02/$ - see front matter
PII: S0169-2607(01)00173-0

and how to interpret the results of the output. The program code of the SAS-macro is given in Appendix A.

2. Description of measures

Cameron and Windmeijer [1] recommended the use of the following R^2 measure for Poisson regression models of the form $\log(\mu_i) = \boldsymbol{\beta}\mathbf{x}_i$, where μ_i is the expectation of a Poisson distributed variable, $\mathbf{x}_i$ is the vector of covariates for the ith observation and $\boldsymbol{\beta}$ is the parameter vector to be estimated with β^0 as the intercept and $\beta^1, \ldots, \beta^k$ as the parameters for the k covariates:

$$R^2_{\mathrm{DEV}} = 1 - \frac{\log L(y) - \log L(\hat{\mu})}{\log L(y) - \log L(\bar{y})}$$

$$= 1 - \frac{\sum_i [(y_i \log(y_i) - y_i)] - (y_i \log(\hat{\mu}_i) - \hat{\mu}_i)}{\sum_i [(y_i \log(y_i) - y_i) - (y_i \log(\bar{y}) - \bar{y})]}$$

where y_i denotes the observed value of the dependent variable, $\hat{\mu}_i$ the predicted value of the ith observation, $\bar{y}$ the mean of the dependent variable, $\log L(\mathrm{y})$ the log-likelihood of the saturated model, $\log L(\hat{\mu})$ the log-likelihood when all covariates are fitted in the model and $\log L(\bar{y})$ the log-likelihood when only an intercept is fitted.

The inflation of simple R^2 measures can be considerable when the number of covariates is large relative to a given sample size. Therefore, Waldhör et al. [2] suggested an adjustment in analogue to the R^2 adjustment in linear regression:

$$R^2_{\mathrm{DEV,df}} = 1 - \frac{(n-k-1)^{-1}[\log L(y) - \log L(\hat{\mu})]}{(n-1)^{-1}[\log L(y) - \log L(\bar{y})]}$$

Mittlböck and Waldhör [3] reviewed this and two alternative adjustments for R^2_{DEV}, and they described and compared the properties of the adjustments in detail. The first R^2_{DEV} adjustment is:

$$R^2_{\mathrm{DEV,adj,1}} = 1 - \frac{\log L(y) - [\log L(\hat{\mu}) - (k/2)]}{\log L(y) - \log L(\bar{y})}$$

$$= 1 - \frac{\log L(y) - \log L(\hat{\mu}) + (k/2)}{\log L(y) - \log L(\bar{y})}$$

$$= 1 - \frac{\sum_i [(y_i \log(y_i) - y_i) - (y_i \log(\hat{\mu}_i) - \hat{\mu}_i)] + (k/2)}{\sum_i [(y_i \log(y_i) - y_i) - (y_i \log(\bar{y}) - \bar{y})]}$$

The likelihood-ratio statistic for testing all k explanatory covariates in regression models is $2[\log L(\hat{\mu}) - \log L(\bar{y})]$, which follows approximately a χ^2 distribution with k degrees of freedom under the null hypothesis H_0: $\beta^1, \ldots, \beta^k = 0$. The expectation of the likelihood-ratio statistic under H_0 is therefore k and the bias of $\log L(\hat{\mu})$ is $k/2$ due to the estimation of the effect of k non-informative covariates. The proposal of Mittlböck and Waldhör [3] for the adjustment of the R^2 measure was therefore an increase of the numerator by $k/2$.

The second R^2_{DEV} adjustment is similar to a proposal of Mittlböck and Schemper [5,6] for logistic regression, which can also be applied to Poisson regression models. That is, the likelihood-ratio statistic $2[\log L(\hat{\mu}) - \log L(\beta^0)]$, where β^0 is the true intercept parameter, follows approximately a χ^2 distribution with $k+1$ degrees of freedom under the null hypothesis H_0: $\beta^1, \ldots, \beta^k = 0$. Therefore the bias of $\log L(\hat{\mu})$ is $(k+1)/2$. Similarly, the bias of $\log L(\bar{y})$ is $1/2$. This leads to the following adjusted measure:

$$R^2_{\mathrm{DEV,adj,2}} = 1 - \frac{\log L(y) - [\log L(\hat{\mu}) - (k+1)/2]}{\log L(y) - \left[\log L(\bar{y}) - \frac{1}{2}\right]}$$

$$= 1 - \frac{\log L(y) - \log L(\hat{\mu}) + (k+1)/2}{\log L(y) - \log L(\bar{y}) + \frac{1}{2}}$$

$$= 1 - \frac{\sum_i [(y_i \log(y_i) - y_i) - (y_i \log(\hat{\mu}_i) - \hat{\mu}_i)] + (k+1)/2}{\sum_i [(y_i \log(y_i) - y_i) - (y_i \log(\bar{y}) - \bar{y})] + \frac{1}{2}}$$

Obviously, $R^2_{\mathrm{DEV,adj,2}}$ always gives values closer to zero than $R^2_{\mathrm{DEV,adj,1}}$. Mittlböck and Waldhör [3] have shown in a simulation study that $R^2_{\mathrm{DEV,adj,1}}$ is preferable to $R^2_{\mathrm{DEV,adj,2}}$ and $R^2_{\mathrm{DEV,df}}$, except when the Poisson regression can be approximated by a linear regression, then $R^2_{\mathrm{DEV,df}}$ is preferable.

Other authors [5] prefer R^2 measures based on sums-of-squares, because of the more intuitive understanding which leads to the following R^2 measure for Poisson regression:

$$R^2_{\mathrm{SS}} = 1 - \frac{\sum_i (y_i - \hat{\mu}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

The adjustment based on the number of observations and on the fitted degrees of freedom is also analogue to the adjustment in linear regression models:

$$R^2_{\mathrm{SS,df}} = 1 - \frac{(n-k-1)^{-1}\sum_i (y_i - \hat{\mu}_i)^2}{(n-1)^{-1}\sum_i (y_i - \bar{y})^2}$$

3. Program description and examples

A SAS-macro has been written that gives the output of PROC GENMOD [7] for Poisson regression and unadjusted and adjusted R^2 measures of the model using PROC IML [8]. It requires the existence of a SAS input data set containing the dependent variable (the observed frequencies) and the independent covariates (categorical or continuous). Additionally the known total (or relative) exposure to risk for each distinct covariate pattern can be given. This may be an area, time, a volume, person-years lived or any other appropriate measure of size. The logarithm of the total exposure to risk is called offset. The following parameters may be given for the SAS-macro:

TITLE:	title for the listing of the analysis
DATA:	name of the data set
DEP:	name of the dependent variable–number of events
TOTAL:	total exposure to risk/exponential of the offset
CLASS:	names of independent categorical variables, separated by blanks
CONT:	names of independent continuous variables, separated by blanks

The program code of the SAS-macro R2POI is given in Appendix A and it is also available via world wide web at ⟨http://www/akh-wien.ac.at/imc/biometrie/r2poi.htm⟩.

The following two examples demonstrate the behaviour of the adjusted and unadjusted R^2 measures with real data. They also help to elucidate the role of explained variation measure in addition to the standard description of Poisson regression results by parameter estimates associated with explanatory factors and by corresponding confidence intervals or P values.

These examples are typical for Poisson regression. Usually the number of events (y_i) for each pattern of covariates is given. These frequencies are regressed on the covariates and, if appropriate, related to the total exposure to risk. In the second example, this is the person-years lived.

3.1. Example 1: assessing toxicity of pollutants in aquatic systems

This study assesses Nitrofen as pollutant in aquatic systems and is described by Bailer and Oris [9]. Nitrofen is both acutely toxic and reproductively toxic to aladoceran zooplankton. The reproductive toxicity test was conducted at measured nitrofen concentrations of 0, 80, 160, 235 and 310 μg/l. The observed numbers of offspring born in the first and the second broods of small freshwater invertebrate zooplankton (*Ceriodaphnia dubia*) in each concentration group are given.

Two models are fitted. The first uses the number of young produced in the first brood and the second one uses the number of young produced in the first two broods as dependent variable, respectively. The concentration (conc) was treated as the class variable. Only the first three (out of ten) measurements in each concentration group were used for this analysis to show the small sample effect.

To run the R2POI-macro, the following statements can be used:

```
%R2POI (DATA = toxicity,
TITLE = Assessing Toxicity of Pollutants in Aquatic Systems,
DEP = brood1, Class = conc);
```

The effect of Nitrofen on the first brood showed no significant effect (Fig. 1a). Nevertheless both unadjusted R^2 measures (R^2_{DEV},R^2_{SS}) are around 25–30%, indicating that nearly one-third of the variation in the number of young produced can be explained by the model. A more realistic view give the adjusted R^2 measures, adjusted for the number of covariates fitted. We can see that they are all close to zero or even negative. In case of negative results, one should use max(0, R^2_{adj}), as a negative R^2 value makes no sense in reality and would be equivalent to having no explained variation at all.

In Fig. 1b we see that the number of young produced in the first two broods decreases significantly with increasing concentration of Nitrofen. The difference between unadjusted (around 70%) and adjusted (around 57%) measures is not so dramatic than in the first model. As the number of observations is small ($n = 15$) compared with the number of covariates fitted ($k = 4$), the adjusted R^2 measures are still about 13% lower than the unadjusted measures.

3.2. Example 2: suicides among academics in Denmark 1970–1980

Based on data from the Danish census of 9 November 1970 and official mortality statistics for the preceding 10 years, the mortality of the total Danish labor force aged 20–64 years was studied. Anderson et al. [10] (pp. 17–18) give the number of suicides and the number of person-years lived (p_years) for academics specified by sex and age. Age is separated into nine age groups with an interval length of 5 years. Thus we have 18 observations (covariate patterns), nine age groups for males and females. In our model for this data set, we fit three covariates: sex, age and age-squared. There is no interaction between sex and age.

To run the SAS program for this example the following statements of the R2POI-macro can be used:

%R2POI (DATA = labor, TITLE = Suicides among academics in Denmark, DEP = suicides, TOTAL = p_years, CLASS = sex);

for fitting the categorical covariate sex only,

%R2POI (DATA = labor, TITLE = Suicides among academics in Denmark, DEP = suicides, TOTAL = p_years, CLASS = sex, CONT = age);

for fitting both sex and age, and

%R2POI (DATA = labor, TITLE = Suicides among academics in Denmark, DEP = suicides, TOTAL = p_years, CLASS = sex, CONT = age age2);

for fitting sex, age and age-squared.

For the sake of brevity, the results are summarized in Tables 1 and 2, presenting the results of the Poisson regressions and the corresponding estimated R^2 values, modelling (a) sex, (b) sex and age, and (c) sex, age and age-squared, respectively. In model (a), sex has no significant influence and all estimated R^2 measures are rather small. When added to the model, both age and age-squared prove to be significant factors. Suicide increases with increasing age until 45–49 years and decreases afterwards. In models (b) and (c), the adjusted R^2 measure results in about 10% lower values than the unadjusted measures. It is shown that age-squared contributes most to the reduction of unexplained variation, the adjusted values of model (c) are more than twice of the corresponding values of model (b). This could not be foreseen from parameter estimates or corresponding P values alone. About 45% of the uncertainty can be explained by sex, age and age-squared.

4. Conclusion

Cameron and Windmeijer [1] and Waldhör et al. [2] suggested the use of a deviance-based measure for explained variation as it is in accordance with the fitting process based on maximum-likelihood theory, whereas Mittlböck and Schemper [5] recommended to construct R^2 measures based on sums-of-squares because of the better intuitive understanding of sums-of-squares compared with the rather theoretical concept of the log-likelihood. Our macro calculates measures based on both concepts. Whatever intrinsic qualities each measure might have, the difference between the different measures is generally smaller than the difference between adjusted and unadjusted R^2 measures, if there are many covariates and small sample size.

Waldhör et al. [2] showed that $R^2_{\mathrm{DEV,df}}$ in Poisson regression models works well for larger μ, when the

```
Assessing Toxicity of Pollutants in Aquatic Systems

                          The GENMOD Procedure

                           Model Information

                 Description                    Value
                 Data Set                       WORK.TOXICITY
                 Distribution                   POISSON
                 Link Function                  LOG
                 Dependent Variable             BROOD1
                 Offset Variable                L_TOTAL
                 Observations Used              15

                        Class Level Information

                    Class     Levels  Values
                    CONC           5  0 80 160 235 310

                  Criteria For Assessing Goodness Of Fit

            Criterion              DF          Value        Value/DF
            Deviance               10         6.0138          0.6014
            Scaled Deviance        10         6.0138          0.6014
            Pearson Chi-Square     10         5.5086          0.5509
            Scaled Pearson X2      10         5.5086          0.5509
            Log Likelihood          .        48.3269               .

                     Analysis Of Parameter Estimates

      Parameter          DF     Estimate      Std Err    ChiSquare  Pr>Chi
      INTERCEPT           1       1.8458       0.2294      64.7344  0.0001
      CONC        0       1      -0.3054       0.3522       0.7517  0.3859
      CONC        80      1      -0.1112       0.3338       0.1110  0.7390
      CONC        160     1      -0.3054       0.3522       0.7517  0.3859
      CONC        235     1      -0.4595       0.3687       1.5531  0.2127
      CONC        310     0       0.0000       0.0000            .       .
      SCALE               0       1.0000       0.0000            .       .

NOTE:  The scale parameter was held fixed.


Calculation of adjusted and unadjusted R2 measures:
===================================================

number of observations =       15
number of fitted df including intercept =        5

R2 measures based on the log-likelihood:
----------------------------------------
  unadjusted             = 0.2501316
  adjusted with df       = 0 (a negative value of -0.049816 was calculated)
  adjusted with method 1 = 0 (a negative value of -0.248635 was calculated)
  adjusted with method 2 = 0 (a negative value of  -0.22107 was calculated)

R2 measures based on sums-of-squares:
-------------------------------------
  unadjusted       = 0.2938931
  adjusted with df = 0.0114504
```

(a)

Fig. 1a. Output of the SAS-macro R2POI for example 1 (assessing toxicity of pollutants in aquatic systems on the first brood).

```
Assessing Toxicity of Pollutants in Aquatic Systems

                    The GENMOD Procedure

                     Model Information

          Description                   Value
          Data Set                      WORK.TOXICITY
          Distribution                  POISSON
          Link Function                 LOG
          Dependent Variable            BROOD12
          Offset Variable               L_TOTAL
          Observations Used             15

                  Class Level Information

             Class     Levels  Values
             CONC           5  0 80 160 235 310

           Criteria For Assessing Goodness Of Fit

      Criterion              DF          Value       Value/DF
      Deviance               10         8.1728         0.8173
      Scaled Deviance        10         8.1728         0.8173
      Pearson Chi-Square     10         7.4379         0.7438
      Scaled Pearson X2      10         7.4379         0.7438
      Log Likelihood          .       346.2505              .

                Analysis Of Parameter Estimates

   Parameter          DF     Estimate      Std Err    ChiSquare  Pr>Chi
   INTERCEPT           1       1.8458       0.2294      64.7344  0.0001
   CONC        0       1       0.9874       0.2688      13.4958  0.0002
   CONC        80      1       0.9874       0.2688      13.4958  0.0002
   CONC        160     1       0.8842       0.2727      10.5124  0.0012
   CONC        235     1       0.7444       0.2786       7.1388  0.0075
   CONC        310     0       0.0000       0.0000            .       .
   SCALE               0       1.0000       0.0000            .       .

NOTE:  The scale parameter was held fixed.


Calculation of adjusted and unadjusted R2 measures:
===================================================

number of observations =        15
number of fitted df including intercept =        5

R2 measures based on the log-likelihood:
----------------------------------------
  unadjusted             = 0.7087485
  adjusted with df       = 0.5922479
  adjusted with method 1 = 0.5662019
  adjusted with method 2 = 0.5467187

R2 measures based on sums-of-squares:
-------------------------------------
  unadjusted       = 0.6944771
  adjusted with df = 0.5722679
```
(b)

Fig. 1b. Output of the SAS-macro R2POI for example 1 (assessing toxicity of pollutants in aquatic systems on the first two broods).

Table 1
Results of Poisson regression models of example 2 (suicides among academics)

Model	Source	β	S.E.	Degrees of freedom	Chi-square	P value
(a)	Intercept	−7.55	0.156	1		
	Sex	−0.25	0.176	1	1.99	0.1583
(b)	Intercept	−7.89	0.227	1		
	Sex	−0.26	0.176	1	2.10	0.1472
	Age	0.07	0.032	1	4.73	0.0297
(c)	Intercept	−8.88	0.444	1		
	Sex	−0.27	0.176	1	2.20	0.1379
	Age	0.53	0.174	1	10.10	0.0015
	Age-squared	−0.04	0.016	1	7.85	0.0051

Poisson regression is approximated by linear regression. Cameron and Windmeijer [1] stated that the concepts of deviance, maximum likelihood estimation and Kullback–Leibler distance are similar in function to the concept of residual sum of squares and least-squares estimation in linear models. Therefore, it is obvious that the same correction makes sense for R^2 measures based on sums-of-squares and for R^2 measures based on deviance residuals. However in Poisson regression situations with rare events, where the normal approximation is not appropriate, this correction can underestimate the true values substantially and an adjustment based on the expected optimism of the log-likelihood under the null-hypothesis seems to be more appropriate. Mittlböck and Waldhör [3] therefore suggested using the adjusted measure $R^2_{\mathrm{DEV,adj,1}}$. This adjustment is also in accordance with the basic character of R^2_{DEV}: it can be shown that it is positive if the model fit is significant and it increases if a significant variable is added to the model.

Usually the range of an adjusted measure of explained variation also includes negative values, as the correction is always based on expected values under the null hypothesis in which covariates exert no effect. But negative values for R^2_{adj} are not a problem, as in these situations the whole model is not significant and one is not interested in reporting and interpreting the proportion of explained variation. Normally R^2_{adj} should be set to zero.

In summary, $R^2_{\mathrm{DEV,adj,1}}$ behaves better than $R^2_{\mathrm{DEV,adj,2}}$ and $R^2_{\mathrm{DEV,df}}$ in typical situations where a Poisson regression is based on a small sample and/or many covariates.

In conclusion, we recommend routine evaluation of the adjusted proportion of explained variation in Poisson regression models. In any of these applications, investigators may easily be misled by highly significant P values or impressive parameter estimates for explanatory factors, while outcomes are far from being determined. R^2 measures offer a different and more accurate view. If investigators evaluate the effect of covariates, the unadjusted R^2 measure will increase with the number of covariates. This rather undesirable property of R^2 can result in artificially high values and may discourage investigators from searching for further prognostic factors. Therefore one should always use an adjusted R^2 measure that considers the number of evaluated covariates and gives a realistic view about the proportion of variation explained by covariates in the model.

Table 2
Estimated R^2 measures (%) of example 2 (suicides among academics)

Model	R^2_{DEV}	$R^2_{\mathrm{DEV,df}}$	$R^2_{\mathrm{DEV,adj,1}}$	$R^2_{\mathrm{DEV,adj,2}}$	R^2_{SS}	$R^2_{\mathrm{SS,df}}$
(a) Sex	7.4	1.7	3.5	3.4	3.0	0[a]
(b) Sex + age	25.9	16.1	18.1	17.4	25.9	16.0
(c) Sex + age + age-squared	56.7	47.5	45.0	43.3	50.4	39.8

[a] The negative value of −3.1% was set to zero.

Appendix A. SAS-macro R2POI for calculating unadjusted and adjusted R^2 measures for Poisson regression models

```
%MACRO R2POI (DATA=, DEP=, TOTAL=, TITLE=' ', CLASS= , CONT=  );

title "   &title";

data &data; set &data;
  %if &total=  %then %do;
    _total=1;
    %let total=_total;
  %end;
  l_total=log(&total);
run;
proc genmod data=&data;
  class &class;
  model &dep=&class &cont
        / dist=poisson link=log obstats offset=l_total;
  make 'ObStats' out=pred noprint;
  make 'ParmEst' out=pest;
run;
proc iml;
  file print;
  use &data var{&total};
    read all into total var{&total};
  close &data;
  use pest var{parm df estimate};
    read all into var var{df};
  close pest;
  use pred var{yvar1 pred};
    read all into y var{yvar1} where(pred^=.);
    read all into p var{pred} where(pred^=.);
  close pred;
  n=nrow(p);
  k=sum(var);
  mu_bar=sum(p)/sum(total);

  put;
  put "Calculation of adjusted and unadjusted R2 measures:";
  put "===================================================";
  put;
  put "number of observations = " n ;
  put "number of fitted df including intercept = " k ;

*****************************************************;
** R2 based on log-likelihood for Poisson model     **;
*****************************************************;
  d_sat=y[loc(y)];
  d_sat=sum(d_sat#log(d_sat)-d_sat);
  d_0=sum(y#log(mu_bar#total) - mu_bar#total);
  d_b=sum(y#log(p)-p);
  d0=d_sat-d_0;
  dbeta=d_sat-d_b;
  r2=1-dbeta/d0;
  r2_df=1-(dbeta/(n-k))/(d0/(n-1));
  r2_e1=1-(dbeta+(k-1)/2)/d0;
  r2_e2=1-(dbeta+k/2)/(d0+0.5);
  put;
```

```
  put "R2 measures based on the log-likelihood:";
  put "--------------------------------------";
  put "  unadjusted             = " r2 ;
  if r2_df<0 then do;
    put "  adjusted with df       = 0 (a negative value of " r2_df " was calculated)" ;
  end;
  else do;
    put "  adjusted with df       = " r2_df ;
  end;
  if r2_e1<0 then do;
    put "  adjusted with method 1 = 0 (a negative value of " r2_e1 " was calculated)" ;
  end;
  else do;
    put "  adjusted with method 1 = " r2_e1 ;
  end;
  if r2_e2<0 then do;
    put "  adjusted with method 2 = 0 (a negative value of " r2_e2 " was calculated)" ;
  end;
  else do;
    put "  adjusted with method 2 = " r2_e2 ;
  end;

*****************************************************;
** sums-of-squares for poisson model               **;
*****************************************************;
  sst=sum((y-mu_bar#total)##2);
  sse=sum((y-p)##2);
  r2_ss=1-sse/sst;
  r2_ssa=1-(sse/(n-k))/(sst/(n-1));
  put;
  put "R2 measures based on sums-of-squares:";
  put "-----------------------------------";
  put "  unadjusted       = " r2_ss ;
  if r2_ssa<0 then do;
    put "  adjusted with df = 0 (a negative value of " r2_ssa " was calculated)" ;
  end;
  else do;
    put "  adjusted with df = " r2_ssa ;
  end;
quit;

%MEND;
```

References

[1] A.C. Cameron, F.A.G. Windmeijer, R^2 measures for count data regression models with applications to health-care utilization, J. Business Econ. Stat. 14 (1996) 209–220.

[2] T. Waldhör, G. Haidinger, E. Schober, Comparison of R^2 measures for Poisson regression by simulation, J. Epidemiol. Biostat. 3 (1998) 209–215.

[3] M. Mittlböck, T. Waldhör, Adjustments for R^2-measures for Poisson regression models, Computat. Stat. Data Anal. 34 (2000) 461–472.

[4] SAS Institute Inc., The SAS® System for Windows, version 6.12, SAS Institute Inc., Cary, NC, 1996.
[5] M. Mittlböck, M. Schemper, Explained variation for logistic regression, Stati. Med. 15 (1996) 1987–1997.
[6] M. Mittlböck, M. Schemper, Computing measures of explained variation for logistic regression models, Comput. Methods Programs Biomed. 58 (1999) 17–24.
[7] SAS Institute Inc., SAS/STAT®, version 6.12, SAS Institute Inc., Cary, NC, 1996.
[8] SAS Institute Inc., SAS/IML®, version 6.12, SAS Institute Inc., Cary, NC, 1996.
[9] J.A. Bailer, J.T. Oris, in: N. Lange, L. Ryan, L. Billard, D. Brillinger, L. Conquest, J. Greenhouse (Eds.), Assessing Toxicity of Pollutants in Aquatic Systems, in Case Studies in Biometry, Wiley, New York, 1994 Chapter 2.
[10] P.K. Anderson, O. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, Springer-Verlag, New York, 1993.