COMMUN. STATIST.—SIMULA., 31(1), 61–73 (2002)

REGRESSION AND TIME SERIES

MEASURES OF EXPLAINED VARIATION IN GAMMA REGRESSION MODELS

Martina Mittlböck[1,*] and Harald Heinzl[1,2,†]

[1]Department of Medical Computer Sciences, University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria
[2]Division of Biostatistics, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

ABSTRACT

The common R^2 measure provides a useful means to quantify the degree to which variation in the dependent variable can be explained by the covariates in a linear regression model. Recently, there have been various attempts to apply the definition of the R^2 measure to generalized linear models. This paper studies two different R^2 measure definitions for the gamma regression model. These measures are related to deviance and sum-of-squares residuals. Depending on the sample size and the number of covariates fitted, so-called unadjusted R^2 measures may be substantially inflated, and the use of adjusted R^2 measures is then preferred. We study several known adjustments previously proposed for R^2 measures in regression models and illustrate the effect on the two unadjusted R^2 measures for the gamma regression model.

*E-mail: martina.mittlboeck@akh-wien.ac.at
†E-mail: harald.heinzl@akh-wien.ac.at

Comparing the resulting measures with underlying population values, we find the best adjustment via simulation.

Key Words: Adjusted R^2 measures; Deviance; Sum-of-squares; Shrinkage; Degrees of freedom; Predictive power; Explained variation

1. INTRODUCTION

A measure of explained variation, also called an R^2 measure, provides information about a regression model that is in addition to the information provided by parameter estimates and associated p-values. R^2 summarizes the predictive power of a regression model and gives the degree to which variation in the dependent variable can be explained by variation in covariates. R^2 measures are frequently used in linear regression and research has focused on their application in other familiar generalized linear models (1–7). Although R^2 measures are frequently called measures of the proportion of ‘‘explained variation’’, for generalized linear models the term variation is not as obvious as in least-squares-estimated linear regression, where it corresponds to variance. Several equivalent definitions exist for the common linear model R^2 measure, which give different results if they are extended and then applied in other modelling frameworks.

Regression models are often used to screen for prognostic factors, even in situations when the sample size is small compared to the number of covariates. By definition, the R^2 value increases monotonically if covariates are added to the model even if the added covariates are not correlated with the interesting outcome. That is, unadjusted R^2 measures can be substantially inflated, jeopardizing the ability to draw valid interpretations. R^2 values of more than 30% can easily be reached even when no association between independent and dependent variables exists. The use of adjusted R^2 measures which penalize for the number of fitted parameters is well-established in linear models. However, R^2 measures for other models lack an appropriate adjustment in general. For special cases, such as logistic and Poisson regression models, adjustments have been proposed in the literature. Some of these adjustments are based on an analogy to linear regression (1,4), some are based on methodical considerations (2,8), and others are based on computational evaluation (3).

In this paper we propose an adjustment of the deviance-based R^2 measure for generalized linear models, so that the expectation of the adjusted R^2 measure corresponds to the underlying population value.

Futhermore, we show that the resulting adjustment coincides with the adjusted R^2 measure in linear regression and with already proposed adjustments for R^2 in logistic and Poisson regression.

In Section 2, we introduce two R^2 measures for gamma regression models; one based on deviances as a measure of variation and the other based on the sum-of-squares. In Section 3, we introduce adjustments for the R^2 measures to reduce the bias of unadjusted R^2 measures. In Section 4, the results of a simulation study are shown comparing the unadjusted and the adjusted R^2 values to the underlying population value. Finally, some aspects of R^2 measures in general, and of the proposed adjustments in particular, are discussed in Section 5.

2. GAMMA REGRESSION MODELS AND UNADJUSTED R^2 MEASURES

It is common to find data where the variance increases with the mean (9). One special case applies if the coefficient of variation is assumed to be constant, the typical situation for the application of gamma regression models, where $E(Y) = \mu$, $\mathrm{var}(Y) = \mu^2/\nu$ and the coefficient of variation of Y is $\sqrt{\nu^{-1}}$. In this case we assume that all components of the dependent variable Y are independent and identically distributed according to the gamma distribution:

$$f_Y(y; \mu, \nu) = \frac{y^{\nu-1}\exp[-y/\mu]}{\mu^{\nu}\Gamma(\nu)}, \quad y \geq 0,\ \nu > 0,\ \mu > 0,$$

where $\Gamma(\nu)$ is the gamma function, μ is the scale parameter and ν determines the shape of the distribution.

When $\nu = 1$ the gamma distribution reduces to the exponential distribution, and if ν is an integer it corresponds to the Erlang distribution. For $\mu = 2$ and $\nu = \kappa/2$ a χ^2-distribution with κ degrees of freedom results and a normal limit is attained as $\nu \to \infty$.

Gamma regression is usually modelled using the reciprocal (canonical) link function $(\mu_i)^{-1} = \mathbf{x}_i \boldsymbol{\beta}$, $i = 1, \ldots, n$, where $\mathbf{x}_i$ is the vector of covariates and $\boldsymbol{\beta}$ is the parameter vector to be estimated with β^0 as the intercept and $\beta^1, \ldots, \beta^{\mathrm{k}}$ as the parameters for the k covariates.

The scaled deviance of gamma regression models is proportional to twice the difference between the log-likelihood achieved under the model of interest and the maximum attainable value in a saturated model. This is given by $D(y; \hat{\mu}, \nu) = \nu D(y; \hat{\mu})$, with $D(y; \hat{\mu}) = -2\sum_i(\log(y_i/\hat{\mu}_i) - (y_i - \hat{\mu}_i)/\hat{\mu}_i)$.

In linear regression the R^2 measure can be defined using sum-of-squares (10). A similar R^2 measure can also be constructed for gamma regression models:

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

where $\hat{\mu}_i$ is the estimated value of the i-th observation under the full model, when all covariates are fitted, and $\bar{\mu}$ is the mean of the dependent variable, that is the expected value under the null model, when only an intercept is fitted.

Another common definition for R^2 measures in generalized linear models is based on deviances (11), which is identical to R^2_{SS} in linear models:

$$R^2_{\mathrm{D}} = 1 - \frac{\nu D(y;\hat{\mu})}{\nu D(y;\bar{\mu})} = 1 - \frac{\sum_i [-\log(y_i/\hat{\mu}_i) + (y_i - \hat{\mu}_i)/\hat{\mu}_i]}{\sum_i [-\log(y_i/\bar{\mu}_i) + (y_i - \bar{\mu}_i)/\bar{\mu}_i]}$$

where $D(y;\hat{\mu})$ and $D(y;\bar{\mu})$ are the deviances under the full and under the null model respectively and ν is assumed to be constant for the full and the null model, so that it can be reduced. That is, ν is either known or it is estimated from the full model and this estimate is used for fitting the null model (1).

R^2_{SS} and R^2_{D} are related to the sum-of-squares and deviance residuals in generalized linear models, respectively. R^2_{D} increases monotonically if covariates are added to the model, independent of any prognostic value. This is not guaranteed for R^2_{SS} in generalized linear models, although this almost always seems to occur in practice (3).

3. ADJUSTED R^2 MEASURES

As already noted the inflation of R^2 measures can be considerable when the number of covariates is large relative to a given sample size. R^2 measures in linear models are adjusted by the degrees of freedom, so that their expectations are unbiased when the true correlation is zero (12). The same adjustment has also been suggested for logistic regression (2,13) and for Poisson regression models (4). The suitability of this adjustment for R^2_{SS} and R^2_{D} in gamma regression models is investigated here, with

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

$$R^2_{\mathrm{D},df} = 1 - \frac{(n-k-1)^{-1}D(y;\hat{\mu})}{(n-1)^{-1}D(y;\bar{\mu})}$$

Although the adjustment by the degrees of freedom is appropriate in linear regression models when using the sum-of-squares approach, it may only be approximately valid in gamma regression models and its behavior has to be investigated in detail.

Another R^2 adjustment results if shrinkage is applied to R^2 measures. Again in linear regression $R^2_{\mathrm{SS},df} = \gamma \times R^2_{\mathrm{SS}}$, where γ is the shrinkage factor (14–18) estimated by $\hat{\gamma} = (\text{model } \chi^2 - k)/(\text{model } \chi^2)$. Here, model χ^2 is the likelihood ratio χ^2 statistic for testing whether any of the k fitted covariates are associated with the response. The shrinkage factor γ reduces to an expression, well known in linear regression $\hat{\gamma} = (\mathrm{SSR} - k\sigma^2)/\mathrm{SSR}$, where SSR is the regression-sum-of-squares and σ^2 will in practice be estimated by the residual mean square. In gamma regression, model $\chi^2 = \nu(D(y;\bar{\mu}) - D(y;\hat{\mu}))$ and the resulting shrinkage-adjusted R^2 measures are denoted by $R^2_{\mathrm{SS},\gamma}$ and $R^2_{\mathrm{D},\gamma}$, respectively.

The shrinkage factor $\hat{\gamma}$ reaches its maximum value of nearly 1, indicating no shrinkage, if the model is highly significant and the model χ^2 achieves a high value. The lower the model χ^2 the closer is $\hat{\gamma}$ to zero, indicating that the calculated R^2 value is due to chance alone. Note that under the null hypothesis of no covariate effects $E(\text{model } \chi^2) = k$. If model $\chi^2 < k$ is estimated, then a negative shrinkage will result.

The resulting shrunken R^2_{D} measure can also be rewritten in the following form:

$$R^2_{\mathrm{D},\gamma} = \frac{\text{model } \chi^2 - k}{\text{model } \chi^2} \times R^2_{\mathrm{D}} = \frac{\nu D(y;\bar{\mu}) - \nu D(y;\hat{\mu}) - k}{\nu D(y;\bar{\mu})}$$
$$= 1 - \frac{D(y;\hat{\mu}) + k/\nu}{D(y;\bar{\mu})}$$

The likelihood-ratio statistic for testing all k explanatory covariates in regression models is $\nu[D(y;\bar{\mu}) - D(y;\hat{\mu})]$ which follows approximately a χ^2 distribution with k degrees of freedom under the null hypothesis of no covariate effects. Therefore, $\nu[D(y;\bar{\mu}) - D(y;\hat{\mu})]$ has an expectation of k and increasing $D(y;\hat{\mu})$ by k/ν results in $E(R^2_{\mathrm{D},\gamma}) = 0$ under the null hypothesis. Similar corrections for logistic and for Poisson regression models have been proposed (2,8).

The estimation of ν by maximum-likelihood is extremely sensitive to rounding errors in very small values of the outcome (19; pp. 295–296). Also, if the gamma assumption is false the coefficient of variation is not consistently estimated. For these reasons the more robust moment estimator $\hat{\nu}^{-1} = (\sum_i \{(y_i - \hat{\mu}_i)/\hat{\mu}_i\}^2)/(n-k)$ is preferred (19,20), which is consistent for ν^{-1}, provided of course that $\boldsymbol{\beta}$ has been consistently estimated. We therefore also calculate shrunken R^2 measures using the above robust estimation of ν, denoted by $R^2_{\mathrm{SS},\gamma^*}$ and $R^2_{\mathrm{D},\gamma^*}$, respectively.

In the following section all unadjusted and adjusted measures proposed here are compared with the population values and with each other.

4. SIMULATION STUDY AND RESULTS

The performance of the two unadjusted and the corresponding adjusted R^2 measures was compared under various conditions for a gamma regression model. Using the SAS procedure FACTEX (21), a factorial design was produced for sample sizes of 16, 32, 64 and 16 384 [$=2^{14}$] with 1, 3 and 5 completely balanced dichotomous covariates centered around zero with values −0.5 and 0.5. For simulations with five covariates and a sample size of 16 only a fractional factorial design was produced as there were too few observations for a factorial design with five covariates. With the SAS function RANGAM (21), gamma distributed random variables were generated with mean $\mu = [\beta^0 + \beta^1 x_1]^{-1}$, β^0 was set at the arbitrary value 0.1. β^1 was chosen so that R^2_{D} values of 0, 0.2, 0.4, 0.6 and 0.8 in a large sample with $n = 16\,384$ under different shape parameters ($\nu = 1$, 5, 10) were achieved. Only the first covariate x_1 was assumed to influence μ. The prognostic effect of all other covariates was eliminated by setting $\beta^2, \ldots, \beta^{\mathrm{k}} = 0$. The SAS procedure GENMOD (21) was used to fit a gamma regression to the data with inverse link, and a SAS macro was written to do all subsequent calculations. The number of repetitions was 1000 for the sample sizes of 16, 32 and 64. The results of R^2_{SS} and R^2_{D} of the large sample ($R^2_{\mathrm{SS},\ell}$ and $R^2_{\mathrm{D},\ell}$) with one covariate were taken to be the true values, respectively, because the estimations of the unadjusted and the adjusted R^2 measures are of essentially the same value.

The number of covariates (k), sample size (n), the population R^2 and the mean estimated R^2 values, are listed for R^2 measures based on deviance in Tables 1 and 2 for $\nu = 1$ and 10, respectively. For R^2 measures based on sum-of-squares the results are listed in Tables 3 and 4 for $\nu = 1$ and 10, respectively. For the sake of brevity, the results for $\nu = 5$ and $R^2 = 0.2$, 0.6 are discussed but not shown in the tables.

In the tables it can be seen that in most critical cases with very small sample size ($n = 16$) and many covariates ($k = 5$) both unadjusted R^2 measures increase dramatically, especially for smaller population R^2 values. The unadjusted R^2_{SS} and R^2_{D} measures can reach values of 30% or even

Table 1. Summary of Adjusted and Unadjusted R^2_{D} Values (in Percent) with Dichotomous Covariates, $\beta^0 = 0.1$ and $\nu = 1$

$R^2_{\mathrm{D},\ell}$	k	n	R^2_{D}	$R^2_{\mathrm{D},df}$	$R^2_{\mathrm{D},\gamma}$	$R^2_{\mathrm{D},\gamma^*}$
0	1	16	5	−1	0	0
		32	3	−1	0	0
		64	1	0	0	0
	3	16	17	−3	3	2
		32	8	−2	1	1
		64	4	−1	0	0
	5	16	28	−8	8	2
		32	13	−4	1	1
		64	7	−2	0	0
40	1	16	43	39	40	40
		32	41	39	40	40
		64	41	40	40	40
	3	16	51	39	43	42
		32	45	39	40	40
		64	43	40	40	40
	5	16	58	38	47	43
		32	48	38	41	41
		64	44	39	40	40
80	1	16	82	81	81	81
		32	81	80	80	80
		64	81	80	80	80
	3	16	84	80	82	81
		32	82	80	82	81
		64	81	80	80	80
	5	16	87	81	84	82
		32	83	80	81	81
		64	82	80	80	80

higher, when the model has no predictive capacity at all, that is the population R^2 equals zero. In all situations studied, R^2_{SS} gives smaller values than R^2_{D}. The higher the population R^2 and the smaller ν the bigger becomes the difference, e.g., $R^2_{\mathrm{D}} = 80\%$ and $R^2_{\mathrm{SS}} = 33\%$ for $\nu = 1$, whereas $R^2_{\mathrm{D}} = 80\%$ and $R^2_{\mathrm{SS}} = 72\%$ for $\nu = 10$. This is because R^2_{D} and R^2_{SS} are equivalent in linear regression and as ν increases, the gamma regression approaches a linear regression and consequently the difference between R^2_{D} and R^2_{SS} becomes smaller.

The observed bias of the adjusted R^2 measures presented here seems to be negligible compared to the bias of the unadjusted measure. They all

Table 2. Summary of Adjusted and Unadjusted R^2_{D} Values (in Percent) with Dichotomous Covariates, $\beta^0 = 0.1$ and $\nu = 10$

$R^2_{\text{D},\ell}$	k	n	R^2_{D}	$R^2_{\text{D},df}$	$R^2_{\text{D},\gamma}$	R^2_{D,γ^*}
0	1	16	6	0	1	0
		32	3	0	0	0
		64	2	0	0	0
	3	16	20	−1	5	0
		32	10	0	1	0
		64	5	0	0	0
	5	16	32	−2	11	1
		32	16	0	3	0
		64	8	0	1	0
40	1	16	44	40	40	40
		32	41	39	39	39
		64	41	40	40	40
	3	16	51	39	42	40
		32	45	39	40	39
		64	42	40	40	44
	5	16	59	38	46	39
		32	49	39	41	39
		64	44	40	40	40
80	1	16	82	81	81	81
		32	81	80	80	80
		64	80	80	80	80
	3	16	84	81	82	81
		32	82	80	80	80
		64	81	80	80	80
	5	16	87	80	83	80
		32	83	80	81	80
		64	81	80	80	80

represent an improvement over the unadjusted R^2 measures. However, because they generate values which are not always in agreement, the question remains, which measure behaves best when compared to the large sample estimate.

It can be easily seen that the adjustment based on shrinkage ($R^2_{\text{SS},\gamma}$ and $R^2_{\text{D},\gamma}$) does not perform well if the maximum-likelihood estimations for ν are used. $R^2_{\text{SS},\gamma}$ and $R^2_{\text{D},\gamma}$ are too high for small sample size and many covariates. The R^2 measures using the more robust estimation of ν, based on the Pearson's chi-squared statistic and the degrees of freedom, behave

Table 3. Summary of Adjusted and Unadjusted R^2_{SS} Values (in Percent) with Dichotomous Covariates, $\beta^0 = 0.1$ and $\nu = 1$. For the Comparability of Deviance-Based and Sum-of-Squares Measures $R^2_{\mathrm{D},\ell}$ Is Also Given

$R^2_{\mathrm{D},\ell}$	$R^2_{\mathrm{SS},\ell}$	k	n	R^2_{SS}	$R^2_{\mathrm{SS},df}$	$R^2_{\mathrm{SS},\gamma}$	$R^2_{\mathrm{SS},\gamma^*}$
0	0	1	16	6	−1	−1	0
			32	3	0	0	0
			64	1	0	0	0
		3	16	23	4	3	3
			32	10	1	1	1
			64	5	0	0	0
		5	16	46	19	11	3
			32	19	4	2	1
			64	9	4	0	0
40	26	1	16	34	29	31	31
			32	30	28	29	29
			64	28	27	28	28
		3	16	54	42	44	44
			32	40	33	36	36
			64	33	29	31	31
		5	16	81	72	64	58
			32	53	43	45	44
			64	38	33	35	35
80	33	1	16	43	39	43	43
			32	39	37	38	38
			64	36	35	36	36
		3	16	62	53	60	60
			32	47	42	46	46
			64	40	37	40	40
		5	16	88	82	85	83
			32	60	53	59	59
			64	45	41	45	45

much better. Therefore, $R^2_{\mathrm{SS},\gamma^*}$ and $R^2_{\mathrm{D},\gamma^*}$ are preferrable to $R^2_{\mathrm{SS},\gamma}$ and $R^2_{\mathrm{D},\gamma}$, respectively.

$R^2_{\mathrm{D},df}$ usually gives smaller values than $R^2_{\mathrm{D},\gamma}$ and $R^2_{\mathrm{D},\gamma^*}$, especially for smaller ν. $R^2_{\mathrm{D},df}$ tends to underestimate the true value, which may only be serious if R^2 and ν are small and the sample size is small relative to the number of covariates, e.g., for $R^2 = 20\%$, $\nu = 1$, $n = 16$ and $k = 5$ the mean value of $R^2_{\mathrm{D},df}$ is 15%. In conclusion, $R^2_{\mathrm{D},\gamma^*}$ behaves best of all adjusted R^2 measures based on the deviance.

Table 4. Summary of Adjusted and Unadjusted R^2_{SS} Values (in Percent) with Dichotomous Covariates, $\beta^0 = 0.1$ and $\nu = 10$. For the Comparability of Deviance-Based and Sum-of-Squares Measures $R^2_{D,\ell}$ Is Also Given

$R^2_{D,\ell}$	$R^2_{SS,\ell}$	k	n	R^2_{SS}	$R^2_{SS,df}$	$R^2_{SS,\gamma}$	R^2_{SS,γ^*}
0	0	1	16	7	0	1	0
			32	3	0	0	0
			64	2	0	0	0
		3	16	20	0	5	0
			32	10	0	1	0
			64	5	0	0	0
		5	16	34	0	11	1
			32	17	0	3	0
			64	8	0	1	0
40	38	1	16	42	38	39	39
			32	40	38	38	38
			64	39	38	38	38
		3	16	52	40	43	40
			32	45	39	40	39
			64	41	39	39	39
		5	16	62	43	48	41
			32	50	40	42	40
			64	44	39	40	39
80	72	1	16	75	73	74	74
			32	73	72	73	73
			64	72	72	73	72
		3	16	81	77	79	78
			32	76	74	75	75
			64	74	72	73	73
		5	16	88	82	84	81
			32	80	76	77	77
			64	75	73	74	74

For R^2_{SS} none of the adjustments is completely satisfactory, e.g., for $n = 16$ and $k = 5$ the achieved values for all ν are much higher than the large sample results. However if the sample size is moderate compared to the number of covariates fitted, e.g., for $k = 3$ and $n \geq 32$ or for $k = 5$ and $n \geq 64$, $R^2_{SS,df}$ and R^2_{SS,γ^*} are both acceptable. $R^2_{SS,\gamma}$ is not recommended because of the estimation of ν by maximum-likelihood.

5. DISCUSSION

In all regression models the use of adjusted R^2 measures are recommended since unadjusted R^2 values increase monotonically as the number of covariates increase, even if those added covariates have no prognostic value at all. This rather undesirable property of R^2 can result in artificially high values and may discourage investigators from searching for further prognostic factors.

We have reviewed the same adjustments for both types of R^2 measures, R^2_{SS} and R^2_D, as Cameron and Windmeijer (1) showed 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, the same correction may make sense for R^2 measures based on sum-of-squares and for R^2 measures based on deviance residuals. However, the difference between R^2 measures based on deviances and R^2 measures based on sum-of-squares may differ substantially. As such, the calculation and comparison of both measures is advised.

A general advantage of the deviance as a measure of variation is that it is optimized by the fitting process and it is additive for nested models if maximum likelihood estimates are used. For this, R^2 measures based on the deviances are favored by many researchers (1,4,6). The correction by degress of freedom is an ad-hoc method and accepted from linear models, whereas methodical arguments and the simulation study favor the shrinkage of R^2 measures when the robust estimation of ν is used.

Although sum-of-squares are not optimized by the fitting process and may not be additive for nested models, the use of the corresponding R^2 measures may be favored because of the more direct interpretability of sum-of-squares than of deviances (2,3,5,19,22). All proposed adjustments of R^2_{SS} are rather ad-hoc methods accepted from similarity to linear regression. These adjustments work well in situations where the sample size is not too small compared to the number of covariates fitted. An adequate bias reduction for other situations may only be achieved using computationally intensive methods, e.g., the use of a jack-knife estimator has been suggested (3) to correct the bias of a generalization of the multiple correlation coefficient which is similar to R^2_{SS}.

Usually the range of an adjusted measure of explained variation also includes negative values, as the adjustments are always based on expected values under the null hypothesis in which covariates exert no effect. In practice, a negative R^2 value has the same meaning as a value of zero, indicating that there is no explained variation at all.

But if the model fit is significant, R^2 measures should give positive values. Also, if a significant factor is added to a model, then the R^2 measure should increase. This is always true for R^2_{D}, $R^2_{\mathrm{D},\gamma}$ and $R^2_{\mathrm{D},\gamma^*}$ but not necessarily for the other R^2 measures, although in practice this will happen in most situations.

In conclusion, the use of R^2 measures give additional insight to the data, additional to parameter estimates and p-values. R^2 measures help to summarize the predictive power of a model. If investigators evaluate the effect of large numbers of covariates, the unadjusted R^2 measure may become a substantially misleading tool. The use of an adjusted R^2 measure is recommended in general.

ACKNOWLEDGMENT

The authors would like to thank an associate editor for helpful suggestions which led to improvements in the manuscript.

REFERENCES

1. Cameron, A.C.; Windmeijer, F.A.G. R^2 Measures for Count Data Regression Models with Applications to Health-Care Utilization. Journal of Business and Economic Statistics **1996**, *14*, 209–220.
2. Mittlböck, M.; Schemper, M. Explained Variation for Logistic Regression. Statistics in Medicine **1996**, *15*(19), 1987–1997.
3. Zheng, B.; Agresti, A. Summarizing the Predictive Power of a Generalized Linear Model. Statistics in Medicine **2000**, *19*, 1771–1781.
4. Waldhör, T.; Haidinger, G.; Schober, E. Comparison of R^2 Measures for Poisson Regression by Simulation. Journal of Epidemiology and Biostatistics **1998**, *3*(2), 209–215.
5. Ash, A.; Shwartz, M. R^2: A Useful Measure of Model Performance When Predicting a Dichotomous Outcome. Statistics in Medicine **1999**, *18*, 375–384.
6. Menard, S. Coefficients of Determination for Multiple Logistic Regression Analysis. The American Statistician **2000**, *54*(1), 17–24.
7. Zheng, B. Summarizing the Goodness of Fit of Generalized Linear Models for Longitudinal Data. Statistics in Medicine **2000**, *19*, 1265–1275.
8. Mittlböck, M.; Waldhör, T. Adjustments for R^2-Measures for Poisson Regression Models. Computational Statistics and Data Analysis **2000**, *34*(4), 461–472.

9. Bain, L. Gamma Distribution. In *Encyclopedia of Statistical Sciences—Volume 3*, Kotz, S., Johnson, N.L., Eds.; Wiley: New York, 1988; 292–298.
10. Kvålseth, T.O. Cautionary Note About R^2. The American Statistician **1985**, *39*, 279–285.
11. Korn, E.L.; Simon, R. Explained Residual Variation, Explained Risk, and Goodness of Fit. The American Statistician **1991**, *45*(3), 201–206.
12. Helland, I.S. On the Interpretation and Use of R^2 in Regression Analysis. Biometrics **1987**, *43*, 61–69.
13. Mittlböck, M.; Schemper, M. Computing Measures of Explained Variation for Logistic Regression Models. Computer Methods and Programs in Biomedicine **1999**, *58*, 17–24.
14. Copas, J.B. Regression, Prediction and Shrinkage (with Discussion). Journal of the Royal Statistical Society B **1983**, *45*, 311–354.
15. Copas, J.B. Cross-Validation Shrinkage of Regression Predictors. Journal of the Royal Statistical Society B **1987**, *49*, 175–183.
16. Copas, J.B. Using Regression Models for Prediction: Shrinkage and Regression to the Mean. Statistical Methods in Medical Research **1997**, *6*, 165–183.
17. Van Houwelingen, J.C.; Le Cessie, S. Predictive Value of Statistical Models. Statistics in Medicine **1990**, *9*, 1303–1325.
18. Harrell, F.E.; Lee, K.L.; Mark, D.B. Tutorial in Biostatistics: Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. Statistics in Medicine **1996**, *15*, 361–387.
19. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*. 2nd Ed.; Chapman and Hall Ltd: London, 1989.
20. Cook, R.J. Generalized Linear Model. In *Encyclopedia of Biostatistics*, Armitage, P., Colton, T., Eds., Wiley: New York, 1998; 1637–1650.
21. The SAS® System for Windows, Version 6.12. Cary, NC: SAS Institute Inc., 1996.
22. Buyse, M. R^2: A Useful Measure of Model Performance When Predicting a Dichotomous Outcome (Letter to the Editor). Statistics in Medicine **2000**, *19*, 271–274.