Simple Effect Measures for Interpreting
Models for Ordinal Categorical Data

Claudia Tarantola\textsuperscript{1} and Alan Agresti\textsuperscript{2}

\textsuperscript{1} Department of Economics and Business, University of Pavia, Pavia, Italy
\textsuperscript{2} Department of Statistics, University of Florida, Gainesville FL, USA

Address for correspondence: Claudia Tarantola, Department of Economics and Business, University of Pavia, Pavia, Italy.
E-mail: claudia.tarantola@unipv.it.
Phone: (+39) 0382 986213.
Fax: (+39) 0382 304226.

Abstract: We survey effect measures for models for ordinal categorical data that can be simpler to interpret than the model parameters. For describing the effect of an explanatory variable while adjusting for other explanatory variables, we present probability-based measures, including a measure of relative size and partial effect measures based on instantaneous rates of change. We also discuss summary measures of predictive power that are analogs of $R$-squared and multiple correlation for quantitative response variables. We illustrate the measures for an example and provide R code for implementing them.
Key words: Cumulative link models; Cumulative logits; Marginal effects; R-squared; stochastic ordering

1 Introduction

Popular models for ordinal categorical response variables, such as models that apply link functions to cumulative probabilities, are generalized linear models that employ non-linear link functions. As a consequence of the non-linearity, model parameters are not as simple to interpret as slopes and correlations for ordinary linear regression. The model effect parameters relate to measures, such as odds ratios and probits, not easily understood by methodologists. This article surveys simpler ways to interpret the effects of an explanatory variable and to summarize the model’s predictive power.

In Section 2, we present alternative summaries of the effect of an explanatory variable, while adjusting for other explanatory variables in the model. These include simple comparisons of the probability of extreme-response outcomes at extreme values of an explanatory variable, measures of average rates of change of the extreme-response probabilities, and group comparisons that result directly from latent variable models that induce standard ordinal models. In Section 3, we present measures of predictive power. A straightforward approach uses $R^2$ and multiple correlation measures that closely resemble those for ordinary linear models, possibly estimated for a corresponding latent variable linear model. Section 4 illustrates existing and proposed measures with an example and provides R code for the analyses. An Appendix at the journal website provides R code for new functions that we developed for these analyses.
2 Ordinal Effect Measures for Individual Explanatory Variables

For an ordinal response variable $y$ with $c$ categories, we consider models in which the explanatory variables may be a mixture of quantitative and categorical variables. We denote explanatory variable values by $\mathbf{x} = (x_1, \ldots, x_p)^T$. In describing ways of summarizing effects for a categorical explanatory variable, we refer also to a separate indicator variable $z$ that distinguishes between two groups.

Currently, the most popular ordinal models are special cases of the cumulative link model

$$\text{link}[P(y \leq j)] = \alpha_j - \beta z - \beta_1 x_1 - \cdots - \beta_p x_p, \quad j = 1, \ldots, c - 1, \quad (2.1)$$

for link functions such as the logit and probit. The nonlinear link function naturally produces effects on the link scale. For example, for cumulative logit models, $-\beta_1$ is the change in the cumulative logit per each 1-unit increase in $x_1$, adjusting for the other explanatory variables. This leads to odds ratios as natural effect measures. For instance, adjusting for the other variables, $\exp(\beta_1)$ is a multiplicative effect of each 1-unit increase in $x_1$ on the cumulative odds of response $> j$ vs. $\leq j$, for each $j$.

Such effect measures are not easy to interpret by scientists who need to understand the effects in more real-world terms. In addition, with nonlinear link functions, effects often behave in a way that is counterintuitive to those mainly familiar with ordinary linear models. For example, if an explanatory variable is added to the model that is uncorrelated with $x_1$, the partial effect of $x_1$ is typically different than in the model without the other explanatory variable, whereas it would be identical in an ordinary
linear model fitted by least squares. We next describe three types of interpretation that supplement estimated model-parameter effects with simpler effects reported on the probability scale rather than on the scale of the link function. Such effects are easier to understand and are typically more stable.

2.1 Extreme-category range-based probability summaries

In practice, special interest often focuses on the highest and lowest response categories, the most extreme outcomes. Those categories often represent a noteworthy state, such as the best or worst outcome (e.g., complete recovery vs. death). It is informative to report how probabilities in these extreme categories, \( P(y = 1) \) and \( P(y = c) \), change as explanatory variables change. As any \( x_k \) increases, cumulative link models that contain solely main effects imply monotonicity in the extreme-category probabilities but not in the other probabilities.

To summarize the effect of \( x_k \) on \( y \), it can be useful to report the difference between the model-fitted estimate of \( P(y = 1) \) and/or \( P(y = c) \), at the maximum and minimum values of \( x_k \), when other explanatory variables are set at particular values such as their means. For a binary variable \( z \), this is a comparison of the two groups on the extreme-category probabilities. For a continuous explanatory variable, a caveat for such measures is that their relevance depends on the plausibility of \( x_k \) taking extreme values when all other explanatory variables fall at their means.

2.2 Marginal effect measures

A second type of simple summary uses the rate of change in the probability of an extreme response category, as a function of \( x_k \). For this, we express the cumulative
Effect Measures for Ordinal Categorical Data

\[ F^{-1}[P(y \leq j)] = \alpha_j - \beta z - \mathbf{x}^T \beta, \quad j = 1, \ldots, c - 1, \quad (2.2) \]

where \( F^{-1} \) is the inverse of a standard cdf, \( \mathbf{x} \) is a column vector of explanatory variable values (excluding \( z \)), and \( \beta \) is a column vector of parameters for \( \mathbf{x} \). Let \( f(y) = \partial F(y)/\partial y \), which is the standard normal probability density function for probit models and the standard logistic probability density function for logistic models.

We first construct the effect for a quantitative explanatory variable. The rate of change in \( P(y = 1) \) at a particular value of \( x_k \), when other explanatory variables are fixed at certain values \( \mathbf{x}^* \), is the partial derivative of \( P(y = 1) \) with respect to \( x_k \),

\[ \partial P(y = 1|\mathbf{x} = \mathbf{x}^*)/\partial x_k. \]

Many sources, such as Greene (2008) and Long and Freese (2014), refer to such an instantaneous effect as a marginal effect. For the cumulative link model, the marginal effect of \( x_k \) on \( P(y \leq j) \), and hence on \( P(y = 1) \), is \( -f(\alpha_j - \beta z - \mathbf{x}^T \beta)\beta_k \). The marginal effect of \( x_k \) on \( P(y = c) \) is \( f(\alpha_j - \beta z - \mathbf{x}^T \beta)\beta_k \).

For the logit link, such an effect for \( x_k \) on \( P(y = 1) \) has the expression

\[ \partial P(y = 1|\mathbf{x} = \mathbf{x}^*)/\partial x_k = \beta_k P(y = 1|\mathbf{x} = \mathbf{x}^*)[1 - P(y = 1|\mathbf{x} = \mathbf{x}^*)]. \]

This takes values bounded above by its highest value of \( \beta_k/4 \) that occurs when \( P(y = 1|\mathbf{x} = \mathbf{x}^*) = 1/2 \). For cumulative probit models, the highest value of this instantaneous change is \( \beta_k/\sqrt{2\pi} \); also when \( P(y = 1|\mathbf{x} = \mathbf{x}^*) = 1/2 \). These maximum values need not be relevant, as \( P(y = 1) \) and \( P(y = c) \) need not be near 1/2 for most or all the data. Long and Freese (2014, pp. 242–246) summarized alternative versions of the marginal effect. One could compute the marginal effect with every
explanatory variable, including \( x_k \), set at its mean. This is called the *marginal effect at the mean* (MEM). The *marginal effect at representative values* (MER), is obtained by setting all explanatory variables at values considered to be of particular interest. The *average marginal effect* (AME), finds the marginal effect of \( x_k \) at each of the \( n \) sample values of the explanatory variables, and then average them. For a categorical explanatory variable, for each version one would instead use a *discrete change*, finding the change in \( P(Y = 1) \) (or \( P(y = c) \)) for a change in an indicator variable, holding all the other variables constant. For instance, for the \( n \) sample observations on \( x \), one could find the difference between \( P(y = 1) \) when \( z = 1 \) and when \( z = 0 \), and average the obtained values.

Greene (2008, pp. 775–785) showed how to obtain standard errors for the maximum likelihood estimators of marginal effect measures. Mood (2010) pointed out that the AME has behavior reminiscent of effects in ordinary linear models, in the sense that it is roughly stable when we add an explanatory variable to the model that is uncorrelated with the variable for which we are describing the effect. This behavior does not occur for the MEM or MER or the log odds ratio.

### 2.3 A probability summary for ordered comparison of groups

We next present an alternative way to summarize the effect of a categorical explanatory variable on an ordinal response \( y \), suggested by Agresti and Kateri (2017) and developed in a more general context by Thas et al. (2012). We discuss this in the context of comparing two groups (\( z = 0 \) and \( z = 1 \)).

It is often sensible to regard an ordinal categorical variable as crude measurement of an underlying continuous latent variable \( y^* \) that, if we could observe it, would
be the response variable in an ordinary linear model. In fact, Anderson and Philips (1981) showed that the cumulative link model (2.2) is implied by a model in which a latent response has conditional distribution with standard cdf given by the inverse of the link function. Let $y^*_1$ and $y^*_2$ denote independent underlying latent variables for the ordinal categorical response, representing the underlying distributions when $z = 1$ and when $z = 0$ respectively. At a particular setting $x$ for other explanatory variables, $P(y^*_1 > y^*_2; x)$ is a summary measure of relative size. This measure is most meaningful when the groups are stochastically ordered.

The normal latent variable model with $y^* \sim N(\beta z + \beta_1 x_1 + \cdots + \beta_p x_p, 1)$ implies the cumulative probit model

$$\Phi^{-1}[P(y \leq j)] = \alpha_j - \beta z - \beta_1 x_1 - \cdots - \beta_p x_p,$$

with $\{\alpha_j\}$ being cutpoints on the underlying scale and $\Phi$ being the standard normal cdf. For this model,

$$P(y^*_1 > y^*_2; x) = P\left(\frac{(y^*_1 - y^*_2) - \beta}{\sqrt{2}} > -\frac{\beta}{\sqrt{2}}\right) = \Phi\left(\frac{\beta}{\sqrt{2}}\right). \tag{2.3}$$

This is true regardless of the $x$ value, so we denote it by $P(y^*_1 > y^*_2)$. For the logit link, Agresti and Kateri (2017) showed that

$$P(y^*_1 > y^*_2) \approx \frac{\exp(\beta/\sqrt{2})}{1 + \exp(\beta/\sqrt{2})}, \tag{2.4}$$

for the $\beta$ coefficient of $z$ in the cumulative logit model. They also provided an expression for log-log links and proposed related measures for the observed $y$ scale that need not relate to latent variables.
3 Summary Measures of Predictive Power

Next we discuss ways to summarize how well we can predict \( y \) using the fit of the chosen ordinal model. Measures of predictive power can be useful for comparing different models. A model that is more complex than a working model need not be much more useful for predictions, regardless of whether its extra terms are statistically significant.

3.1 Concordance index

The concordance index (Harrell et al. 1996, sec. 5.5) is a measure of predictive power that strictly uses ordinality and is natural for models that imply stochastic orderings at various settings of the explanatory variables. Consider all pairs of observations that have different outcomes on \( y \). The concordance index estimates the probability that the predictions and the outcomes are concordant, that is, that the observation with the larger \( y \)-value also has a stochastically higher set of model-fitted probabilities. For cumulative link models, the stochastic ordering of the model-fitted probabilities is identical to the ordering of linear predictor values without the intercept. Of those pairs that are untied on \( y \) but tied on the linear predictor, half are treated as concordant and half as discordant, so that the concordance index has a null value of 1/2. The higher the value above 1/2, the better the predictive power. The concordance index is a linear transform to the \([0, 1]\) scale of a version of the ordinal measure of association called Somers’ \( d \).

Appealing features of the concordance index are its simple structure and its generality of potential application. Because it utilizes ranking information only, however, it
cannot distinguish between different link functions or linear predictors that yield the same stochastic orderings. With a single linear predictor in a cumulative link model, for instance, the concordance index assumes the same value for logit and complementary log-log link functions, even though the model fits can be quite different.

3.2 R-squared type measures

An alternative approach to summarizing predictive power adapts standard measures for quantitative response variables. For example, to mimic $R^2$ for ordinary linear models, we could assign ordered scores $\{v_j\}$ to the categories of $y$ and find the proportional reduction in variance in comparing the marginal variation to the conditional variation. That measure has the disadvantage of requiring response scores, which models such as cumulative link models do not require. A way to construct such a measure without assigning such scores is to estimate $R^2$ for the linear model for an underlying latent response variable. McKelvey and Zavoina (1975) suggested this measure for the cumulative probit model, for which the underlying latent variable model is the ordinary normal linear model. Let $y_i^*$ denote the value of the latent variable for subject $i$. The $R^2$ measure has the usual proportional reduction in variation form

$$R^2 = \frac{\sum (y_i^* - \bar{y}^*)^2 - \sum (\hat{y}_i^* - \hat{\bar{y}}^*)^2}{\sum (y_i^* - \bar{y}^*)^2} = \frac{\sum (\hat{y}_i^* - \bar{y}^*)^2}{\sum (y_i^* - \bar{y}^*)^2}.$$ 

This equals the estimated variance of $\hat{y}^*$ divided by the estimated variance of $y^*$. After fitting a cumulative link model we can estimate the variance of $\hat{y}^*$ by the variance of the linear predictor $\hat{y}^* = \hat{\beta}z + \hat{\beta}_1x_1 + \cdots + \hat{\beta}_p x_p$ without the intercepts. We cannot observe the latent variable or its sample variance, but we can estimate that variance by the estimated variance of $\hat{y}^*$ plus the variance of the latent variable distribution, which is 1 for the probit link and $\pi^2/3 = 3.29$ for the logit link (i.e., standard logistic
distribution).

An alternative proportional-reduction-in-variability approach uses a likelihood-based measure such as was proposed for binary data by McFadden (1974). We can express this in terms of deviance measures for the ungrouped data file. Denote the residual deviance by $D_M$ for the working model fit and denote the null deviance (i.e., for the model containing only intercept terms) by $D_0$. Denote the corresponding maximized log-likelihood values by $L_M$ and $L_0$. The pseudo $R$-squared measure

$$\frac{D_0 - D_M}{D_0} = 1 - \frac{L_M}{L_0}$$

equals 0 when the model provides no improvement in fit over the null model and it equals 1 when the model fits as well as the saturated model. A weakness of such a measure and related ones based on the log-likelihood is that the scale is not the same as for $y$. Interpreting the numerical value is difficult, other than in a comparative sense for different models.

For surveys of $R^2$ type measures in various contexts (but mainly for binary responses), see Liao and McGee (2003), Mittlböck and Schemper (1996), and Zheng and Agresti (2000).

### 3.3 Multiple correlation measures

Some statisticians prefer correlation measures over related $R^2$ measures, because of the appeal of working on the original scale and its proportionality to the effect size. For example, for the ordinary linear model, for fixed marginal standard deviations, doubling the slope also doubles the correlation.

For ordinal modeling, we could estimate the multiple correlation for the underlying
latent variable model, using the square root of the McKelvey and Zavoina (1975) $R^2$. An approach that does not require reference to a latent variable or assigning arbitrary scores to $y$ uses as scores the average cumulative proportions for the marginal distribution of $y$. For sample marginal proportions $\{p_j\}$, the average cumulative proportion in category $j$ is

$$v_j = \sum_{k=1}^{j-1} p_k + \left( \frac{1}{2} \right) p_j, \quad j = 1, 2, \ldots, c.$$ 

Such scores, which are linearly related to the midranks $\{r_j\}$ by

$$r_j = nv_j + 0.5, \quad v_j = (r_j - 0.5)/n,$$

are sometimes referred to as ridits. See Agresti (2010, Sec. 2.1) for discussion and examples of their use. In particular, (1) they satisfy $\sum_{j=1}^{c} p_j v_j = 0.50$, (2) if two adjacent categories of $y$ are combined, then the ridit score for the new category falls between the original two scores, with the other scores being unaffected, and (3) if the category ordering is reversed, the ridit score for category $j$ transforms from $v_j$ to $(1 - v_j)$. With such scores, one could construct the correlation for the $n$ sample observations between the observed outcome category score for a subject and the estimated mean score generated by the model-fitted probability values for the subject. With ridit or midrank scores, this is a multiple correlation version of the Spearman correlation.

## 4 Example

We illustrate the ordinal effect measures using data from a study of mental health (Agresti 2015, Section 6.3.3). The model relates a four-category ordinal response vari-
able measuring mental impairment (1 = well, 2 = mild symptom formation, 3 = moderate symptom formation, 4 = impaired) to a binary indicator of socioeconomic status (SES: 1 = high, 0 = low) and a quantitative life-events (LE) index taking values on the nonnegative integers between 0 and 9 with mean 4.3 and standard deviation 2.7. The $n = 40$ observations are available at www.stat.ufl.edu/~aa/glm/data.

Different packages in R permit fitting cumulative link models, but none of them implement many of the measures we have described. Separate functions are available for some measures, and we developed new functions based on existing ones. For fitting cumulative link models, we used the polr function of the R-package MASS. We illustrate with the cumulative logit model implied by the logistic latent variable linear model. The maximum likelihood fit is

$$\logit[\hat{P}(y \leq j)] = \hat{\alpha}_j + 1.111(\text{SES}) - 0.319(\text{LE}).$$

Table 1 shows model fitting and results, with edited output.

### 4.1 Effect measures for individual explanatory variables

For a quantitative variable, such as LE in the mental impairment data set, we can report the change in an extreme-category probability over its range, at the means of other explanatory variables or at particular categories of qualitative explanatory variables. Table 2 finds the estimated changes when LE changes from its minimum to its maximum value, separately for low SES and high SES subjects. For either SES group, as LE increases, the probability decreases substantially for the well category (by 0.389 for low SES and by 0.581 for high SES) and increases substantially for the impaired category (by 0.560 for low SES and by 0.354 for high SES). These changes
Table 1: R code and output (edited) for the cumulative logit model fitted to the mental impairment data.

```
> Mental <- read.table("http://www.stat.ufl.edu/~aa/glm/data/Mental.dat",header=T)
> head(Mental) # the first 4 of the 40 observations
     impair  ses  life
   1       1     1     1
   2       1     1     9
   3       1     1     0
   4       1     1     4
...
> attach(Mental)
> library(MASS)
> impair.f <- factor(impair) # polr requires response to be a factor
> logit.m <- polr(impair.f ~ ses + life, method="logistic")
> summary(logit.m)
Coefficients:
                Value Std. Error t value # reported t actually is a z Wald statistic
ses   -1.1112  0.6109   -1.819
life   0.3189  0.1210    2.635
```

characterize in a simple manner the very strong effect of LE on mental impairment.

Table 2 also shows the estimated changes between the SES levels, at the mean of LE, which is a discrete-change version of the marginal effect at the mean (MEM). For high SES compared to low SES at the mean of LE, the estimated probability is 0.208 higher for the well category and 0.176 lower for the impaired category.

We next consider marginal effects. The R-package erer of Sun (2016) has a function ocME that supplies marginal effects at the mean, using output from the polr function. Table 3 shows results, focusing again on the extreme response categories. At the mean of LE, the rate of change in the estimated probability per unit change in LE
Table 2: R code and output (edited) for extreme category probability changes in cumulative logit model for mental impairment. The changes compare the maximum and minimum life events values, at low SES and at high SES, and compare low and high SES at the mean for life events.

```r
> pred_max0 <- predict(logit.m, data.frame(ses=0, life=max(life)), type="probs")
> pred_max0
 1  2  3  4 # predicted outcome prob's
0.04102612 0.11914448 0.18048372 0.65934567
> pred_min0 <- predict(logit.m, data.frame(ses=0, life=min(life)), type="probs")
> pred_min0
 1  2  3  4
0.42998727 0.34080542 0.13029529 0.09891202
> pred_max0[c(1,4)] - pred_min0[c(1,4)]
 1  4
-0.3889612 0.5604337 # LE effect (at max - at min) in cat's 1 and 4 when SES=0.

> pred_max1 <- predict(logit.m, data.frame(ses=1, life=max(life)), type="probs")
> pred_min1 <- predict(logit.m, data.frame(ses=1, life=min(life)), type="probs")
> pred_max1[c(1,4)] - pred_min1[c(1,4)]
 1  4
-0.5811892 0.3542873 # LE effect (at max - at min) in cat's 1 and 4 when SES=1.

> pred1 <- predict(logit.m, data.frame(ses=1, life=mean(life)), type="probs")
> pred0 <- predict(logit.m, data.frame(ses=0, life=mean(life)), type="probs")
> pred1[c(1,4)] - pred0[c(1,4)]
 1  4
0.2078490 -0.1764812 # these are discrete marginal effects of SES at mean of LE
```

---
is $-0.062$ for the *well* outcome and $0.049$ for the *impaired* outcome. For categorical explanatory variables it reports the discrete change. When SES increases from 0 to 1 at the mean of LE, the estimated probability of the *well* outcome increases by $0.208$ and the estimated probability of the *impaired* outcome decreases by $0.176$. These are the same measures we just found and reported at the bottom of Table 2. The *ocME* function employs only logit and probit link functions. In the website Appendix, we present an extension of it that handles also log-log and complementary log-log link functions.

Table 3: R Code and output (edited) for marginal effect at the mean (MEM) and average marginal effect (AME) for the cumulative logit model fitted to the mental impairment data.

```
> library(erer)
> ocME(logit.m) # for marginal effects at the mean
  effect.1 effect.2 effect.3 effect.4
ses 0.208 0.053 -0.084 -0.176
life -0.062 -0.014 0.027 0.049

> ocAME(logit.m) # new function available at website Appendix

$ME.1 # category 1 (well)
   effect std.error z.value p.value
ses 0.198 0.104 1.913 0.056
life -0.057 0.019 -3.005 0.003

$ME.4 # category 4 (impaired)
   effect std.error z.value p.value
ses -0.171 0.094 -1.819 0.069
life 0.048 0.017 2.780 0.005
```
The `erer` package does not report average marginal effects, so we constructed a function called `ocAME` based on the `ocME` function. The function, available at the website Appendix, uses the discrete-change version when an explanatory variable is categorical. Table 3 also shows results of applying this function, for the extreme categories. At the 40 observed values for LE and SES, the rate of change in the estimated probability per unit change in LE averages to $-0.057$ for the well outcome and to $0.048$ for the impaired outcome. At the 40 observed values for LE, when SES increases from 0 to 1, the estimated probability of the well outcome increases by an average of 0.198 and the estimated probability of the impaired outcome decreases by an average of 0.171.

We next estimate the ordinal comparison measure introduced in Section 2.3, by comparing the SES groups while adjusting for LE. An exact estimate using formula (2.3) follows directly from the SES effect estimate in the cumulative probit model. Table 4 shows its value of 0.314 and its 95% profile likelihood confidence interval. For the cumulative logit model, we can use the approximate formula (2.4), which gives a similar result (here, 0.313) because of the similarity of logit and probit link functions. So, the estimated probability is 0.31 that mental impairment is worse at high SES than at low SES, adjusting for the LE index.

### 4.2 Measures of predictive power

The concordance index can be easily obtained in R as shown in Table 5, using an R package that has a function for Somers’ $d$. The command `logit.m$lp` provides the fitted values of the linear predictor without the intercepts. For the cumulative logit model, we estimate that for 70.5% of the untied pairs on mental impairment,
Table 4: R Code and output for stochastic superiority comparison of SES groups, using cumulative probit and logit models fitted to the mental impairment data.

```r
> probit.m <- polr(impair.f ~ ses + life, method = "probit")
> summary(probit.m) # we don't show intercept parameter estimates

Coefficients:
   Value  Std. Error  t value
ses  -0.6834   0.36411  -1.877
life  0.1954   0.06887   2.837

> pnorm(probit.m$coefficients[1]/sqrt(2)) # using formula (3)
  ses
0.3144742 # ML estimate of probit ordinal comparison measure for SES

> pnorm(confint(probit.m)[1,]/sqrt(2))
  2.5 %  97.5 % # profile likelihood CI
[1] 0.1608011 0.5074903

> exp(logit.m$coefficients[1]/sqrt(2))/(1+exp(logit.m$coefficients[1]/sqrt(2)))
  ses
0.31308 # formula (4) approx ordinal comparison for logit link
```

Effect Measures for Ordinal Categorical Data
the observation with the higher mental impairment also had a stochastically higher estimated distribution.

Table 5: R code and output (edited) for concordance index for cumulative logit model with mental impairment data.

```
> logit.m <- polr(impair.f ~ ses + life, method="logistic")
> library(DescTools)
> (SomersDelta(logit.m$lp, impair) + 1)/2
[1] 0.7047377
```

Table 6 shows how to find various $R^2$ measures. The McFadden pseudo $R^2$ measure is easily calculated using deviances or maximized log-likelihoods. The deviance for the model is 9% smaller than for the null model. For the logistic latent variable model, the multiple correlation is 0.473, with $R^2 = 0.224$. We estimate that for the underlying continuous measure of mental impairment, the conditional variability (given LE and SES) is 22.4% less than the marginal variability. For the 40 observations, with ridit scores (which are 0.15, 0.45, 0.6875, 0.8875), the multiple correlation is 0.479 and $R^2 = 0.230$. 


Table 6: R code and output (edited) for $R^2$ and multiple correlation measures for cumulative logit model with mental impairment data.

```r
> logit.m <- polr(impair.f ~ ses + life, method="logistic")
> logit.m0<- polr(impair.f ~ 1, method="logistic")

> R2 <- (logit.m0$deviance - logit.m$deviance)/logit.m0$deviance
> R2 # McFadden pseudo R-squared
[1] 0.09119561

1 - logLik(logit.m)/logLik(logit.m0)
'log Lik.' 0.09119561 (df=5)

> var(logit.m$lp)/(var(logit.m$lp) + (pi^2)/3)
[1] 0.2237609 # R-squared for latent var. model for cumulative logit
> sqrt(0.2237609)
[1] 0.4730337 # multiple correlation for latent variable model

> pred <- predict(logit.m, type = "probs")
> ridits <- (rank(impair) - 0.5)/40; ridits # rank fn. gives midrank scores
[1] 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500 0.1500
[12] 0.1500 0.4500 0.4500 0.4500 0.4500 0.4500 0.4500 0.4500 0.4500 0.4500 0.4500
[23] 0.4500 0.4500 0.6875 0.6875 0.6875 0.6875 0.6875 0.6875 0.6875 0.6875 0.8875
[34] 0.8875 0.8875 0.8875 0.8875 0.8875 0.8875 0.8875

> pred.ridit <- 0.15*pred[,1] + 0.45*pred[,2] + 0.6875*pred[,3] + 0.8875*pred[,4]
> cor(ridits, pred.ridit); cor(ridits, pred.ridit)^2
[1] 0.4793919 # Spearman multiple correlation analog

[1] 0.2298166 # an R-squared for rank scores
```

---
References


