Ordinal Independent Variables

Other Entries

SAGE Research Methods Foundations

By: Richard A. Williams
Published:2020
Length: 3,000 Words
DOI: http://dx.doi.org/10.4135/9781526421036
Methods: Ordinal Independent Variables
Online ISBN: 9781526421036
Disciplines: Anthropology, Business and Management, Criminology and Criminal Justice, Communication and Media Studies, Counseling and Psychotherapy, Economics, Education, Geography, Health, History, Marketing, Nursing, Political Science and International Relations, Psychology, Social Policy and Public Policy, Social Work, Sociology, Science, Technology, Computer Science, Engineering, Mathematics, Medicine
Access Date: October 30, 2020
Publishing Company: SAGE Publications Ltd
City: London

As Richard A. Williams and Christopher Quiroz (2019) noted, there are several procedures that can be used when the dependent variable in an analysis is ordinal. Possible methods include the ordered logit model, the generalized ordered logit model, interval regression, and stage models.

However, researchers also often want to use ordinal measures as independent/explanatory variables in their models. Rightly or wrongly, it is very common to treat such variables as continuous—or, more precisely, as having interval-level measurement with linear effects. When the items use a Likert-type scale (e.g., *strongly disagree*, *disagree*, *neutral*, *agree*, *strongly agree*), this may be a reasonable or at least not too controversial practice.

However, many ordinal items use categories that clearly are not equally spaced—for example, the options might be something like “daily,” “a few times a week,” “once a week,” “a few times a month,” “once a year,” and “never.” Some would argue that such variables should be treated as though they were categorical (i.e., nominal), not continuous. That is, just as might be done with a variable like religious affiliation where the categories are in no particular order (e.g., 1 = Catholic, 2 = Protestant, 3 = Jewish, 4 = Muslim, 5 = Other), the variable should be broken into indicator (dummy) variables (e.g., var1 = 1 if daily, 0 otherwise; var2 = 1 if a few times a week, 0 otherwise; var3 = 1 if once a week, 0 otherwise; and so on). Hence, if an ordinal variable has 5 categories, 4 indicator variables would be entered into the model instead. (One indicator variable, often referred to as the reference category, is left out because it would be perfectly correlated with the others, e.g., if var1, var2, var3, and var4 all equal 0 then var5 must equal 1. Conversely, if any of the first 4 variables equals 1 then var5 must equal 0.)

Researchers are generally reluctant to use this strategy unless it is clearly necessary. Treating ordinal variables as categorical rather than continuous ignores the fact that the categories are ordered, makes the model less parsimonious and harder to interpret, and becomes especially unwieldy if there are several ordinal independent variables in the model.

This entry discusses the arguments for treating ordinal variables as continuous versus treating them as categorical. It also discusses alternative strategies that recode ordinal variables in an effort to make them more like continuous measures that have linear effects.

Continuous Versus Categorical Approaches

David J. Pasta (2009) argued strongly for (usually) treating ordinal variables as continuous, even when the spacing is not equal across categories. He raised several points. First, continuous variables often raise the same concerns that ordinal variables do. Just because a variable is continuous does *not* mean that its effects are linear. The effect of a variable increasing from 1 to 2 need not be the same as increasing from 10 to 11 or from 10,000 to 10,001. A researcher may need to take logs, add squared terms, or estimate spline functions, for example, to adequately account for the nonlinear effects of a variable. Such issues are just a bit more obvious with ordinal variables because the number of possible values is limited, and it is often questionable to believe that the categories are equally spaced. But, nonlinear effects of continuous variables and unequally spaced intervals for ordinal variables can both potentially have similar effects on the correctness of a model.

Pasta (2009) further noted

> I am squarely in the camp that says “everything is linear to a first approximation” and therefore I am very cheerful about treating ordinal variables as continuous. Deviations from linearity can be important and should be considered once you have the basics of the model established, but it is very rare for an ordinal variable to be an important predictor and have it not be important when considered as a continuous variable. That would mean that the linear component of the relationship is negligible but the non-linear component is substantial. It is easy to create artificial examples of this situation, but they are very, very rare in practice. (p. 3)

Pasta (2009) ultimately concluded that “It turns out that it doesn’t matter that much in practice—the results are remarkably insensitive to the spacing of an ordinal variable except in the most extreme cases. It does, however, matter more when you consider the products of ordinal variables” (p. 2).

Pasta’s arguments are encouraging to those who would like to make their models as easy to understand as is possible. However, it must be remembered that these are Pasta’s personal positions, and not all agree with him. J Scott Long and Jeremy Freese (2006) agreed that ordinal variables are often treated as continuous. But, they added that

> The advantage of this approach is that interpretation is simpler, but to take advantage of this simplicity you must make the strong assumption that successive categories of the ordinal independent variable are equally spaced. For example, it implies that an increase from no publications by the mentor to a few publications involves an increase of the same amount of productivity as an increase from a few to some, from some to many, and from many to lots of publications. Accordingly, before treating an ordinal independent variable as if it were interval, you should test whether this leads to a loss of information about the association between the independent and dependent variable. (Long & Freese, 2006, p. 421)

For those who share Long and Freese’s concerns, there are multiple ways to formally test whether the assumption of linear effects with ordinal independent variables is justified.

Likelihood Ratio χ2 Tests and Bayesian Information Criterion/Akaike Information Criterion Tests

Williams and Quiroz (2019) described likelihood ratio, Wald, BIC, and AIC tests when estimating models with ordinal dependent variables. Variations of these approaches also work with ordinal independent variables. First, the researcher estimates two models. In the constrained model, the ordinal variable is treated as continuous, forcing its effects to be linear, and in the unconstrained model, it is treated as categorical, so that effects need not be linear. The researcher then uses a likelihood ratio χ^2 test (or a BIC test or AIC test) to decide whether use of the more parsimonious continuous measure is justified.

This point is illustrated here with an example from the European Social Survey (ESS, 2015), the same data set used by Williams and Quiroz (2019) in their discussion of ordinal dependent variables. The ESS is a cross-national study that has been conducted every 2 years across Europe since 2001. For this example, the 2012 ESS survey for Great Britain (ESS Round 6: European Social Survey Round 6 Data (2012)) is used. The study has 2,286 respondents. Although cases have unequal probabilities of selection, weighting had little effect on the results so to simplify the presentation they were not used. The dependent variable is voted (1 = *voted in last election*, 0 = *did not*). The ordinal independent variable is ipstrgv, which asks respondents how similar they are to a person who feels that “it is important that government is strong and ensures safety.” The possible responses are 1 = *Very much like me*, 2 = *Like me*, 3 = *Somewhat like me*, 4 = *A little like me*, 5 = *Not like me*, and 6 = *Not like me at all*.

Using the Stata Version 16.1 statistical package, we can proceed as follows.

```
. * Strong Government treated as continuous
. quietly logit voted c.ipstrgv, nolog
. est sto m1
. * Strong Government treated as categorical
. quietly logit voted i.ipstrgv, nolog baselevels
. est store m2
. * Test of Continuous vs Categorical treatment
. lrtest m1 m2, stats

Likelihood-ratio test                                 LR chi2(4)  =      4.09
(Assumption: m1 nested in m2)                         Prob > chi2 =    0.3936

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
          m1 |      2,132  -1260.418  -1257.693       2    2519.387   2530.717
          m2 |      2,132  -1260.418  -1255.647       6    2523.294   2557.283
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

The likelihood ratio, BIC, and AIC tests all agree that the more parsimonious model (Model m1) that treats belief in a strong government as a continuous variable is preferable. The χ^2 value for the differences between the models is far from being statistically significant. With BIC and AIC statistics, lower values are preferred when comparing models, and the BIC and AIC values are indeed both lower when belief in a strong government is treated as continuous. (Also, if ordinary least squares regression were being use for the analysis, *F* tests could also be used to contrast the models.)

Wald χ^2 Tests

Likelihood ratio tests are not always possible; for example, they generally cannot be done when surveyweighted data are used. Luckily, Wald tests are also possible. For a Wald test, only one model needs to be estimated. Both the continuous and categorical versions of the ordinal variable are included in that model. If the effects of the categorical variable are not statistically significant, then the continuous version alone is sufficient. In Stata, this can be done via use of the o. notation (o stands for omitted). (Note that, because we are including two versions of the ordinal variable, two categories of the ordinal variable must be excluded rather than the usual one.)

```
. * Wald test of whether Strong government shout be
. *    treated as continuous or categorical
. quietly logit voted c.ipstrgv o(1 2)ipstrgv, nolog baselevels
. testparm i.ipstrgv

 ( 1)  [voted]3.ipstrgv = 0
 ( 2)  [voted]4.ipstrgv = 0
 ( 3)  [voted]5.ipstrgv = 0
 ( 4)  [voted]6.ipstrgv = 0

           chi2(  4) =    4.00
         Prob > chi2 =    0.4067
```

Again, the results indicate that the continuous version of the variable is fine. None of the indicator variables for belief in a strong government either individually or as a group significantly improve model fit beyond what is obtained by simply treating the strong government variable as continuous.

Scoring Methods

Rather than choose between two flawed approaches, researchers sometimes try to determine whether coding of an ordinal measure can be made better. Powers and Xie (2008) noted that various scoring methods are sometimes used to assign values to ordinal variables. For example, suppose someone is asked how often they attend church, and the options are *daily* = 1, *once a week* = 2, *a few times a month* = 3, *once a year* = 4, and *never* = 5. The categories are ordered from most frequent to least frequent but the distance between categories is clearly not consistent. To make church attendance be more like a continuous variable, it might reasonably be recoded as *daily* = 1, *once a week* = 1/7, *a few times a month* = 1/14, *once a year* = 1/365.25, and *never* = 0.

Powers and Xie also noted that the midpoints of categories are sometimes used. For example, if an interval ranged between 1 and 10, the midpoint of 5.5 might be used; if another interval ranged between 11 and 25, the midpoint value of 18 could be used instead. While often done, midpoints can be poor estimates of the true values—for example, Powers and Xie say that for a category like “less than 12 years of schooling” a value like 5.5 would likely greatly underestimate years of schooling. There is also the problem of how to score an interval that does not contain an upper bound (e.g., an ordinal measure of annual income where the highest category is “income greater than $100,000”).

Powers and Xie also discussed more complicated scoring schemes which use normal transformations or require the use of auxiliary information. A good advanced scoring scheme may require a great deal of knowledge of the topic and measures as well as statistical sophistication by the researcher.

Another example, again using the ESS, shows that sometimes only minor recoding is required to allow an ordinal variable to be treated as continuous. The ESS asks respondents how much they agree or disagree with the statement "Gay men and lesbians should be free to live their own life as they wish." The dependent variable glfree is coded 1 if respondents agree or strongly agree, 0 if they are neutral, disagree, or strongly disagree. The ordinal independent variable, frequency of religious attendance, is coded 1 = *Every day*, 2 = *More than once a week*, 3 = *Once a week*, 4 = *At least once a month*, 5 = *Only on special holy days*, 6 = *Less often*, and 7 = *Never*.

Clearly, the categories are not evenly spaced, and when the religious attendance variable is used as is, the model that treats religious attendance as categorical fits much better than the model that treats it as continuous.

However, it is also clear that there is very little distance between Categories 5 (*Only on special holy days*) and 6 (*Less often*). When we combine Categories 5 and 6 to both be Category 5, and recode Category 7 (*Never*) to 6, the χ^2, BIC, and AIC statistics all prefer the model that treats religious attendance as continuous rather than categorical.

```
. * Treat religious attendance as continuous
. quietly logit glfree c.religatnd, nolog
. est sto m1
. * Treat religious attendance as categorical
. quietly logit glfree i.religatnd, nolog baselevels
. est store m2
. * Test continuous vs categorical
. lrtest m1 m2, stats

Likelihood-ratio test                                 LR chi2(4)  =      5.89
(Assumption: m1 nested in m2)                         Prob > chi2 =    0.2073

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
          m1 |      2,246  -1004.956  -958.7006       2    1921.401   1932.835
          m2 |      2,246  -1004.956   -955.754       6    1923.508   1957.809
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

Final Thoughts

Researchers should seriously consider treating ordinal independent variables as continuous and having linear effects. As Pasta (2009) noted, this practice is often no more questionable than treating continuous variables as if they had linear effects and often works very well in practice. Particularly, when there are several ordinal independent variables in a model, the parsimony and ease of interpretation gained by treating variables as continuous may outweigh the small increase in information that might be obtained by treating them as categorical.

However, there is no need to simply assume that the continuous approach is best. Formal tests can be conducted. In the examples presented here, those tests added further justification to treating the ordinal variable as continuous.

Also, it need not be assumed that the current coding of the variable (typically something like 1, 2, 3, 4, and 5) is best. There may be reasonable ways to recode the variable, so it better reflects a continuous measure. Sometimes even just combining two categories that are very close to each other may be sufficient.

Finally, when reasonable cases can be made for two different approaches, it often makes sense to try it both ways. Ideally, the researcher will find that it does not make much difference which way things are done, which will lessen possible concerns about the researcher's preferred approach. However, if the results do differ greatly depending on which approach is used, the researcher should consider carefully the best way to move forward.

References

ESS Round 6: European Social Survey Round 6 Data. (2012). *Data file edition 2.1*. Norwegian Social Science Data Services, Norway—Data Archive and Distributor of ESS data. Retrieved May 27, 2015, from http://www.europeansocialsurvey.org/

Long, J. S., & **Freese, J.** (2006). *Regression models for categorical dependent variables using Stata*. Stata Press.

Pasta, D. J. (2009). Learning when to be discrete: Continuous vs. categorical predictors. Paper 248-2009 in the SAS Global Forum 2009. https://support.sas.com/resources/papers/proceedings09/248-2009.pdf

Powers, D. A., & **Xie, Y.** (2008). *Statistical methods for categorical data analysis* (2nd ed.). Emerald.

Williams, R. A., & **Quiroz, C.** (2019). Ordinal regression models. In **P.Atkinson**, **S.Delamont**, **A.Cernat**, **J. W.Sakshaug**, & **R. A.Williams** (Eds.), *SAGE research methods foundations*. SAGE. http://dx.doi.org/10.4135/9781526421036885901