Interactions and Sample Comparisons with Categorical Models:
Understanding and Applying the New ASR Standards

Professor Cyrus Schleifer - OU Sociology

New ASR Guidelines

- American Sociological Review (ASR) published guidelines for Quantitative Journal Submissions:
  1. P-values of where $p < 0.10$ will no longer be considered strong evidence
  2. That mediation and moderation language needs to be accompanied by appropriate test
  3. The language of multiple regression, multivariable regressions, multivariate regressions needs to be used accurately
  4. That sociologists need to take quantitative measurement issues more seriously
  5. Method sections need to be more organized and logically coherent

Interactions in Categorical Models

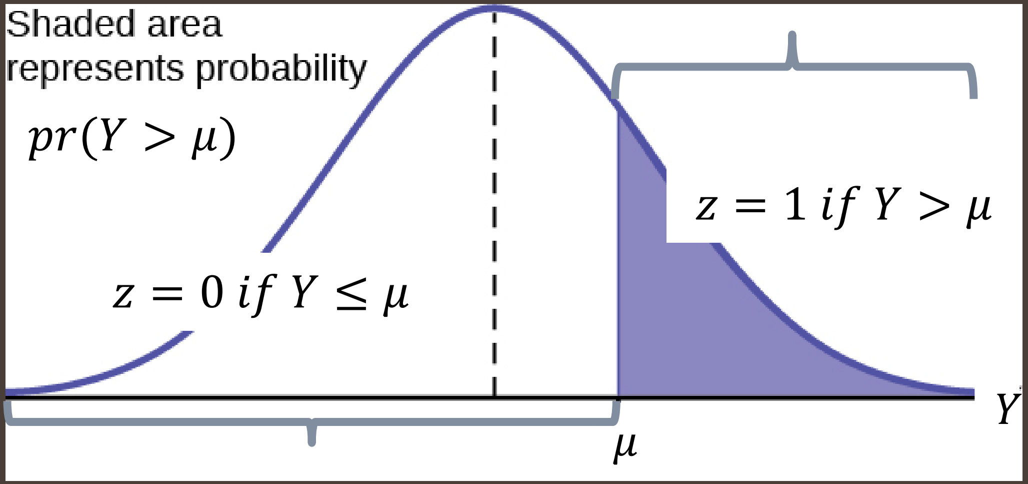
6. Comparing groups with interactions and/or comparing coefficients across models with categorical outcomes:

   - “Various problems have been raised with using z-statistics (and associated p-values) of the coefficients of a multiplicative term to test for a statistical interaction in nonlinear models with categorical dependent variables”
   - “The case is closed: don’t use the coefficient of the interaction term to draw conclusions about statistical interactions in models such as logit, probit, Poisson, and so on.”

The Problem

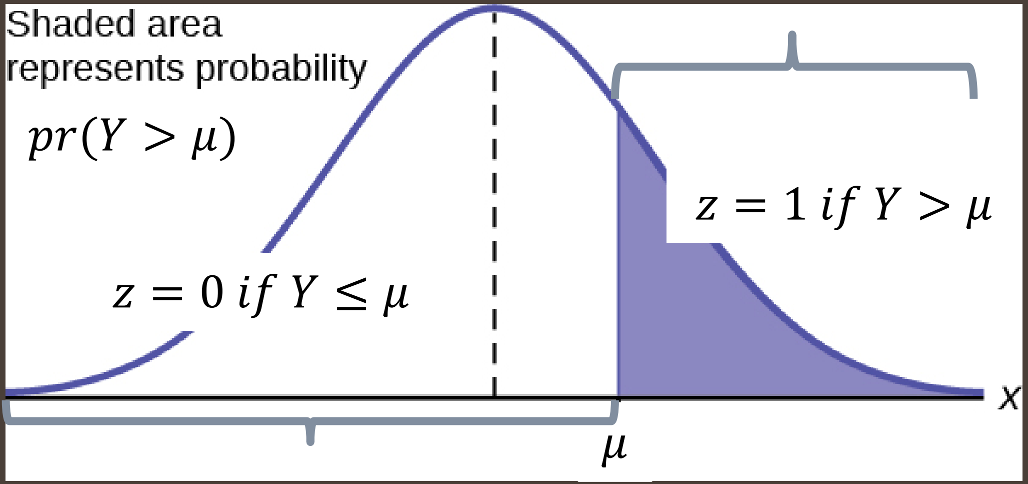
- The coefficients in binary regression models are confounded with residual variations
  - Unobserved Heterogeneity
- Differences in the degree of residual variations across groups can produce apparent differences in coefficients that "***are not indicative of true difference in causal effect***."

The Problem



- Latent Variable approach
  - We can imagine the binary variable z is equal to 1 if $Y > \mu$
  - and equal to 0 if $Y \leq \mu$
- We might model the latent propensity with a linear model

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \cdots + \alpha_j x_{ij} + \sigma \varepsilon_i$$



The Problem

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \cdots + \alpha_j x_{ij} + \sigma \varepsilon_i$$

- Here, the α's represent the parameters we are estimating (intercept and coefficients)
- The ε_i is the random disturbance (error) that is assumed to be independent of the x variables with a fixed variance.
- And σ is a parameter the fixes the ε_i

The Problem

- If we assume that the ε_i has a standard logistic distribution, it follows that the observed dichotomy z is governed by the logit model

$$g[pr(z_i = 1)] = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_j x_{ij}$$

  - where $g(p) = \ln(\frac{p}{1-p})$, which is the logit "link" function
  - Notice that here, we are estimating beta (β) parameters, as opposed to alphas (α)
  - Also notice, that for this regression we do not have an error term

The Problem

- We can relate the β coefficients from the logistic equation to the α coefficient from the linear regression:

$$\beta_0 = \frac{(\alpha_0 - \mu)}{\sigma} \; for\ the\ Intercept$$

$$\beta_j = \frac{\alpha_j}{\sigma} \; for\ j = 1, \ldots, j\ covariates$$

The Problem

- The problem is that the logistic coefficients are a product of the coefficient from imagined latent model of the propensities (α_j) and the parameter (σ) that we used to adjust the disturbance variance
- When we compare models using linear regressions, we compare:

$$H_0: \alpha_{jg1} = \alpha_{jg2}$$

- This includes comparing across models (through we need to adjust for different standard error - sample size - here)
- For interaction comparison it becomes:

$$H_0: \alpha_{main\,effect} = (\alpha_{main\,effect} + \alpha_{interactive\,effect})$$

The Problem

- However, when we compare models using categorical regressions, we compare:

$$H_0: \beta_{jg1} = \beta_{jg2} \therefore \frac{\alpha_{jg1}}{\sigma_{g1}} = \frac{\alpha_{jg2}}{\sigma_{g2}}$$

- Notice, that here for the comparison to work in the same way as for the linear regression model, we must assume variance parameter for group 1 (σ_{g1}) and the variance parameter for group 2 (σ_{g2}) are equal.

$$ie: \sigma_{g1} = \sigma_{g2} = \sigma$$

- This assumption is problematic and, when violated, could lead to spurious conclusions

Damn You Math

- What does this all mean:
  - Interaction coefficients from categorical regression models pose a problem
    - Cannot trust your stars in your tables
  - Care is needed when comparing coefficients across categorical models
    - Standard model comparison tests do not work under these conditions

Approaches to Address the Issue

- Luckily, several prominent social statisticians have set about developing approaches that avoid this issue and allow for comparison
  1. Adjusting for Unequal Residual Variation – Allison (1999)
  2. Choice Models – Richard Williams (2009)
  3. Comparing Probabilities – Long and Mustillo (2018)

Adjusting for Unequal Residual Variation

- Allison 1999 proposed a way forward
- Since the model is not identified if we allow both the coefficients and variance disturbance to differ across group, we can not model this issue directly
- If, however, we are willing to assume that one (or more) X variables has the same underlying coefficient (α), we can estimate the ratio of the disturbance variances

$$\frac{\beta_{2m}}{\beta_{2w}} = \frac{\alpha_{2m}/\sigma_m}{\alpha_{2w}/\sigma_w} = \frac{\sigma_m}{\sigma_w}$$

$$if\ we\ assume: \alpha_{2m} = \alpha_{2w}$$

Adjusting for Unequal Residual Variation

- If G_i is a binary grouping variable (here men and women) then our linear regression becomes:

$$y_i = \alpha_0 + \alpha_1 G_i + \sum_{j>1} \alpha_j x_{ij} + \sigma_i \varepsilon_i \quad \text{where } \sigma_i = \frac{1}{1 + \delta G_i}$$

- With a touch of mathematical fanciness, our logit regression becomes:

$$\ln\left(\frac{p}{1-p}\right) = (\alpha_0^* + \alpha_1 G_i + \sum_{j>1} \alpha_j x_{ij})(1 + \delta G_i)$$

Adjusting for Unequal Residual Variation

TABLE 2: Logit Regressions Predicting Promotion to Associate Professor for Male and Female Biochemists, Disturbance Variances Unconstrained

	All Coefficients Equal		*Articles Coefficient Unconstrained*	
Variable	*Coefficient*	SE	*Coefficient*	SE
Intercept	7.4913***	.6845	–7.3655***	.6818
Female	–0.93918**	.3624	–0.37819	.4833
Duration	1.9097***	.2147	1.8384***	.2143
Duration squared	–0.13970***	.0173	–0.13429***	.01749
Undergraduate selectivity	0.18195**	.0615	0.16997***	.04959
Number of articles	0.06354***	.0117	0.07199***	.01079
Job prestige	–0.4460***	.1098	–0.42046***	.09007
$\hat{\delta}$	–0.26084*	.1116	–0.16262	.1505
Articles × Female			–0.03064	.0173
Log likelihood	–836.28		–835.13	

*p < .05. **p < .01. ***p < .001.

Heterogeneous Choice Models

- The problem with Allison's approach, is that the choice of which coefficient to fix can effect the conclusions we draw
- Williams (2009) notes that Allison's proposed solution is part of a larger class of models know as the heterogeneous choice models
- He argues that the more generalized approach should be preferred because it does not require the same key assumptions

Heterogeneous Choice Models

- The Heterogeneous choice model take the following form:

$$\Pr(y_i = 1) = g\left(\frac{x_i\beta}{\exp(z_i\gamma)}\right) = g\left(\frac{x_i\beta}{\exp(\ln(\sigma_i))}\right) = g\left(\frac{x_i\beta}{\sigma_i}\right)$$

- Where g is the link function
- X is a vector of predictors
- Z is vector of the different error variance

Heterogeneous Choice Models

- There are a number of advantage of this class of models
  - Can yield information even when coefficients do not vary across groups
  - Estimate variance equation across multiple grouping variable
  - Variance estimate could be of substantive interest
  - Estimates on ordinal outcomes
  - General increase in model flexibility

Heterogeneous Choice Models

Table 3
Comparison of Allison (1999) and Heteroscedastic Logit Models

	Allison		Heteroscedastic Logit	
	Coefficient	*SE*	Coefficient	*SE*
Got promoted (1 = *yes*, 0 = *no*)				
Female	–0.939*	0.37	–0.939*	0.37
Duration	1.910***	0.20	1.910***	0.20
Duration squared	–0.140***	0.017	–0.140***	0.017
Undergraduate selectivity	0.182***	0.053	0.182***	0.053
Number of articles	0.0635***	0.010	0.0635***	0.010
Job prestige	–0.446***	0.097	–0.446***	0.097
Intercept	–7.491***	0.66	–7.491***	0.66
δ	–0.261*	0.11		
ln(σ)				
Female			0.302*	0.15
N	2,797		2,797	

*$p < .05$. ***$p < .001$.

Probabilities

- Scott Long and Sarah Mustillo (forthcoming – Online 2018) proposed that comparing the predicted probabilities - as opposed to the coefficient - is a way forward on this issue.

Probabilities

- "Tests of the equality of probabilities or marginal effects on the probability do not require additional assumptions since identical predictions are obtained using the α's and the β's"

- From Linear regression:

$$Group\ 0: Pr(Y=1|x_1,x_2) = Pr_0(\varepsilon \leq \alpha_0^0 + \alpha_1^0 x_1 + \alpha_2^0 x_2 | x_1, x_2)$$

$$Group\ 1: Pr(Y=1|x_1,x_2) = Pr_1(\varepsilon \leq \alpha_0^1 + \alpha_1^1 x_1 + \alpha_2^1 x_2 | x_1, x_2)$$

- From Logit regression

$$Group\ 0: Pr(Y=1|x_1,x_2) = Pr_0(\varepsilon \leq \beta_0^0 + \beta_1^0 x_1 + \beta_2^0 x_2 | x_1, x_2)$$

$$Group\ 1: Pr(Y=1|x_1,x_2) = Pr_1(\varepsilon \leq \beta_0^1 + \beta_1^1 x_1 + \beta_2^1 x_2 | x_1, x_2)$$

Probabilities

- This has several advantages, including:
  - Overcoming the issue of the error disturbance allowing for group comparison
  - Placing the tested effect back on the original scale of interest
  - Quick and easy approaches to graphically displaying the difference
  - Allows for similar model specifications, with straight forward post estimation tests of the probability difference across group

Other Approaches

- These three approaches are not an exhaustive list
  - Hard Coding Interactions
  - Y-standardization
  - Linear probability Models
  - ΔP approaches
- Whatever approach one chooses, given that sociology relies heavily on categorical comparisons both theoretically and empirically we need to stay abreast of the different approaches to this concern

Mood's (2009) Table

Table 6 Characteristics of estimated effects on binary dependent variables

	Capture nonlinearity	Comparable across groups, samples etc.	Comparable across models	Conditional effect estimate[a]
Measures based on odds and log-odds				
Odds ratio	Yes	No	No	Yes
Log-odds ratio	Yes	No	No	Yes
y-standardization	Yes	No	Yes	No
Allison's procedure	Yes	Yes[b]	No	Yes
Heterogeneous choice models	Yes	Yes[c]	No	Yes
Measures based on percentages				
Average marginal effect	No	Yes	Yes	No
Average partial effect	Yes[d]	Yes	Yes	No
Marginal effect	Yes[d]	No	No	Yes
ΔP	Yes[d]	No	No	Yes
Linear probability model	No	Yes	Yes	No[e]

[a] *In a multivariate model.*
[b] *If assumption that one variable has same effect in groups etc. is correct.*
[c] *If assumption about the functional form of the relationship is correct.*
[d] *If estimated at several places in the distribution.*
[e] *If the true relationship is nonlinear.*

STATA EXAMPLE

Comparing Probabilities