Binary contrasts for unordered polytomous regressors

Jeremy Freese
Stanford University
Stanford, CA
jfreese@stanford.edu

Sasha Johfre
Stanford University
Stanford, CA
sjohfre@stanford.edu

Abstract. In observational studies, regression coefficients for categorical regressors are overwhelmingly presented in terms of contrasts with a reference category. For unordered regressors with many categories, however, this approach often focuses on contrasting different pairs of categories with one another with little substantive rationale for foregrounding those comparisons. Mean contrasts, which compare categories to the overall mean, provide an alternative to the reference category, but the magnitude of mean contrasts is conflated with the relative sizes of the categories. Instead, binary contrasts compare a category to all the other categories, allowing the familiar interpretation for dichotomous regressors. Our program `binarycontrast` computes binary contrasts.

Keywords: `binarycontrast`, interpretation, contrast coding

With unordered polytomous explanatory variables, there is often no reason to consider any one category as being more substantively fundamental for interpreting results than the others. Nevertheless, the most common strategy for presenting coefficients for such variables involves assigning one category as the reference category, so that coefficients for other categories are identified as the contrasts between each of those categories and the reference. While contrasts between any pair of categories other than the reference can typically be obtained by simple arithmetic, this still takes for granted that the most useful quantities for interpretation are contrasts between pairs of categories.

An alternative approach is to present coefficients as the contrast of each category with the overall mean, which is sometimes called *weighted effect coding* (?), *deviation coding*, or (our preference) *mean contrasts* (?). Mean contrasts do not involve a reference category, but instead each category has its own coefficient. For a variable with k categories,

$$\sum_{m=1}^k p_m \beta_m = 0$$

where p_m is the proportion of observations belonging to category m and β_m is that category's coefficient. (In experimental research, one may simply be interested in results balanced across treatments and thus not weighting by the proportion, but our discussion here is directed to population-based observational studies for which estimates of population proportions are germane.)

While interpreting results with the reference category approach involves treating one category as different from the others, with mean contrasts the coefficients all have a uniform interpretation, as the difference between that category and the overall conditional mean. But this interpretation may also seem a bit peculiar given that the observations in the category in question also contribute to the overall mean, so that with mean contrasts a category is being contrasted partly with itself. One consequence is that, for mean contrasts, more frequent categories will tend to have coefficients closer to zero simply by virtue of being larger and hence more influential on the resulting mean. A further implication is that differences in the magnitudes of mean contrasts cannot be interpreted as differences in effect sizes.

If membership in category m was a binary variable, then the interpretation of β_m would be straightforward, along the lines of “Being in group m is associated with a β_m difference in the outcome relative to not being in group m .” The value of β_m here would also not depend on the relative frequency of m . Of course, if we were specifically interested in m vs. not- m , we could just fit our model with a binary measure instead of a polytomous one. Instead, we want to fit the model with all the categories of our polytomous measure, but we want still coefficients for each of the k categories to represent the *binary contrast*, that is, the contrast between those observations belonging to that category and those that do not.

In Stata, mean contrasts can be readily computed post-estimation in Stata using the `contrast` command, but binary contrasts cannot. Our package `binarycontrast` rectifies this by leveraging the option of the `contrast` command to specify custom contrasts. We will first present an example that illustrates binary contrasts. Then we explain how binary contrasts are computed and how `binarycontrast` is implemented in Stata.

1 Example

The General Social Survey (GSS) is a biennial probability-based survey of adults in the United States. In the GSS, subjective social standing is assessed on a ten-point scale, which we standardize for purposes here. Our explanatory variables for the example are the respondent’s region of residence (9 categories based on US Census regions) and race/ethnicity (4 categories, in which respondents who identify as Latino are coded as Latino and otherwise respondents are coded by self-identified race as “White,” “Black,” and “Other”).

We fit a linear regression model using the 2004-2018 surveys. Both region and race are unordered variables, and in this case, we have no especial reason to consider one category as more fundamental than the others for purposes of interpretation. Following ? [?], we fit the model in Stata using the reference categories that yield positively-signed estimates, which are the West North Central and Latino categories.

```
. regress rank_std ib4.region ib2.raceeth [pweight=wtssall], allbase noheader
(sum of wgt is 13,974.0596886215)
```

rank_std	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
region						
New England	.0644136	.0582653	1.11	0.269	-.0497942	.1786215
Middle Atl..	.1151715	.047418	2.43	0.015	.0222259	.2081172
East North..	.0016225	.0442593	0.04	0.971	-.0851317	.0883766
West North..	0 (base)					
South Atla..	.0704269	.0442648	1.59	0.112	-.0163381	.1571919
East South..	.0613814	.0540132	1.14	0.256	-.0444916	.1672544
West South..	.0547645	.0484734	1.13	0.259	-.0402499	.1497789
Mountain	.0356222	.0511767	0.70	0.486	-.064691	.1359354
Pacific	.0894345	.0466163	1.92	0.055	-.0019396	.1808086
raceeth						
Black	.170565	.0410977	4.15	0.000	.0900081	.2511219
Latino	0 (base)					
White	.251564	.0314132	8.01	0.000	.18999	.313138
Other	.1599367	.0538671	2.97	0.003	.0543499	.2655235
_cons	-.255715	.0483186	-5.29	0.000	-.350426	-.1610041

The region estimates indicate the difference in subjective status between respondents who live in each region and those who live in the West North Central states, adjusting for race. These 8 contrasts are only a modest fraction of the 36 possible contrasts between different regions of the country. When region is a covariate, it is often unclear how much we specifically care about any of these contrasts between pairs. Even if region was our main explanatory variable of interest, we might be less interested in specific differences between specific regions as opposed to how a region differs from the others as a whole.

From the above results, we know for sure that living in a West North Central state is negatively associated with subjective status and living in a Middle Atlantic state is positively associated with it. We can tell this because these regions have the lowest and highest coefficients, respectively. For categories without extreme coefficients, we can only surmise whether it is positively or negatively associated with the outcome by combining these coefficients with information on the relative frequencies of the different categories.

When we look at the coefficients for race/ethnicity, we can see that, net of region, being White is associated with higher subjective rank and being Latino is associated with lower subjective rank. For Blacks, we note that the coefficient is closer to Whites than it is for Latinos. Does this mean that Blacks have higher subjective social status than respondents who are not Black? There are more Whites than Latinos in the US, so the larger difference between Blacks and Latinos might be more than offset by the larger number of Whites and their disproportionate influence on the mean of non-Black respondents. One could thus reasonably look at the coefficients and not know if being Black is positively or negatively associated with the outcome.

Instead, we can compute results that contrast each category with the conditional mean:

Binary contrasts for unordered polytomous regressors

```
. contrast gw.region gw.raceeth, effects nowald
```

```
Contrasts of marginal linear predictions
```

```
Margins      : asbalanced
```

	Contrast	Std. Err.	t	P> t	[95% Conf. Interval]	
region						
(New England						
vs						
mean)	.006999	.0427947	0.16	0.870	-.0768844	.0908825
(Middle Atl..						
vs						
mean)	.057757	.0262122	2.20	0.028	.0063776	.1091363
(East Nort..						
vs						
mean)	-.0557921	.0204644	-2.73	0.006	-.0959051	-.0156791
(West Nort..						
vs						
mean)	-.0574146	.0373797	-1.54	0.125	-.1306837	.0158546
(South Atl..						
vs						
mean)	.0130123	.019226	0.68	0.499	-.0246732	.0506978
(East Sout..						
vs						
mean)	.0039668	.0368464	0.11	0.914	-.068257	.0761906
(West Sout..						
vs						
mean)	-.0026501	.0277165	-0.10	0.924	-.0569782	.0516779
(Mountain						
vs						
mean)	-.0217924	.0324544	-0.67	0.502	-.0854073	.0418226
(Pacific						
vs						
mean)	.03202	.023764	1.35	0.178	-.0145606	.0786005
raceeth						
(Black						
vs						
mean)	-.0277355	.0261565	-1.06	0.289	-.0790056	.0235347
(Latino						
vs						
mean)	-.1983005	.0264224	-7.51	0.000	-.2500918	-.1465091
(White						
vs						
mean)	.0532636	.0074297	7.17	0.000	.0387003	.0678268
(Other						
vs						
mean)	-.0383637	.0448507	-0.86	0.392	-.1262772	.0495497

From this, we can see that residents of Middle Atlantic states are the only group whose subjective social status is above the mean by the conventional $p < .05$ standard, while the East North Central region – *not* the West North Central region – is the only region significantly below the mean by this standard. The reason the East North Central region is statistically significantly different from the mean and the West North Central region is not (even though the coefficient for the latter is lower) is that the East North Central region is over twice as large as the West North Central region, so its standard

error is smaller.

We can also see that being Black is indeed negatively associated with lower subjective standing relative to the mean. The result for Latinos, meanwhile, is much farther from zero than any of the other race/ethnic groups, and nearly four times as large in magnitude as the difference for Whites. Then again, there are more Whites than Latinos in the GSS, so the reason that the difference is so much smaller for Whites could be largely or entirely due to there being many more Whites in the sample. From these results we cannot answer: is the difference between Latinos and non-Latinos actually bigger than the difference between Whites and non-Whites?

Furthermore, in interpreting the mean contrast of .053 for Whites, we might say something like ,“Adjusting for region, being white is associated with a .053 standard deviation increase in subjective social rank relative to the mean.” Notice, however, that this result is effectively comparing Whites to another “group” (i.e., the whole sample) that is mostly composed of Whites.

One might consider it more straightforward instead to compare non-overlapping groups, viz. to contrast those who are White with those who are not. This is the binary contrast. Here we use our `binarycontrast` command to compute it:

```
. binarycontrast region raceeth
```

	Contrast	Std. Err.	p	[95% Conf. Interval]	
region					
New England	.0073491	.044935	0.870	-.0807296	.0954277
Middle Atlantic	.0651206	.0295541	0.028	.0071907	.1230506
East North Central	-.0672075	.0246515	0.006	-.1155278	-.0188871
West North Central	-.0610512	.0397473	0.125	-.1389611	.0168588
South Atlantic	.0164922	.0243675	0.499	-.0312715	.0642558
East South Central	.0042331	.0393201	0.914	-.0728395	.0813057
West South Central	-.0029605	.0309627	0.924	-.0636516	.0577306
Mountain	-.0235829	.035121	0.502	-.0924247	.0452589
Pacific	.0379021	.0281295	0.178	-.0172354	.0930397
raceeth					
Black	-.0323593	.030517	0.289	-.0921767	.0274582
Latino	-.2328723	.0310289	0.000	-.2936931	-.1720516
White	.1572381	.0219331	0.000	.1142463	.2002299
Other	-.0402725	.0470823	0.392	-.1325601	.0520151

From these results, we can see that, adjusting for region, the mean subjective social rank of Latinos is .23 standard deviations lower than for non-Latinos, while the subjective rank of Whites is .16 standard deviations higher than for non-Whites.

When the only explanatory variable in the model is the categorical variable, then the estimated binary contrast for a category is the same as the estimated coefficient in a model in which the only regressor is that category as a dichotomy. For example, here we compute the binary contrast for a model in which race/ethnicity is the only regressor:

```
. quietly regress rank_std i.race
```

```
. binarycontrast race
```

	Contrast	Std. Err.	p	[95% Conf. Interval]	
raceeth					
Black	-.0388479	.0236871	0.101	-.0852778	.0075821
Latino	-.2094254	.0249532	0.000	-.2583369	-.1605139
White	.1393415	.018112	0.000	.1038394	.1748435
Other	-.0336227	.0418637	0.422	-.1156811	.0484357

The contrast for Latino above is the same as what we would observe in a model in which the only regressor was a dichotomous variable contrasting Latino and non-Latino respondents:

```
. regress rank_std 2.race, noheader
```

rank_std	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
raceeth						
Latino	-.2094254	.024963	-8.39	0.000	-.2583562	-.1604946
_cons	.0302493	.0091668	3.30	0.001	.0122812	.0482174

When there are more regressors in the model, the binary contrast will be similar to the result one would observe if the polytomous regressor was replaced by a dichotomous regressor, but not identical, as the estimates in the former case are affected by the covariances between the terms for the other categories and the other explanatory variables.

2 Computing binary contrasts

If y is our outcome of interest, we can define the mean and binary contrast for category m conditional on covariate values $\mathbf{x}$ as:

$$\begin{aligned} \text{Mean contrast}(m) &= E(y_m|\mathbf{x}) - E(y|\mathbf{x}) \\ \text{Binary contrast}(m) &= E(y_m|\mathbf{x}) - E(y_{-m}|\mathbf{x}) \end{aligned} \tag{1}$$

where y_{-m} refers to values of the outcome for observations in which our categorical regressor is not m . We use y_m to refer to the outcome value for observations in category m , and so the contrast between category m and any category can be written as $E(y_m) - E(y_i)$. A more elaborate way of writing the mean contrast, then, would be as each of

the k contrasts $E(y_m) - E(y_i)$ weighted by the frequency of i :

$$\begin{aligned}
 \text{Mean contrast}(m) &= E(y_m|\mathbf{x}) - E(y|\mathbf{x}) \\
 &= E(y_m|\mathbf{x}) - \sum_{i=1}^k p_i E(y_i|\mathbf{x}) \\
 &= \frac{\sum_{i=1}^k p_i (E(y_m|\mathbf{x}) - E(y_i|\mathbf{x}))}{\sum_{i=1}^k p_i}
 \end{aligned} \tag{2}$$

for which the sum of the relative frequencies for all categories $\sum_{i=1}^k p_i$ is 1. For the binary contrast, we want instead to compute the mean of the contrasts excluding the contrast $E(y_m) - E(y_m) = 0$. From the above formula for the mean contrast, we subtract $E(y_m) - E(y_m)$ from the numerator and p_m from the denominator.

$$\begin{aligned}
 \text{Binary contrast}(m) &= E(y_m|\mathbf{x}) - E(y_{-m}|\mathbf{x}) \\
 &= \frac{\left(\sum_{i=1}^k p_i (E(y_m|\mathbf{x}) - E(y_i|\mathbf{x}))\right) - p_m (E(y_m|\mathbf{x}) - E(y_m|\mathbf{x}))}{\left(\sum_{i=1}^k p_i\right) - p_m}
 \end{aligned} \tag{3}$$

Since $E(y_m) - E(y_m) = 0$, this term drops out, meaning that the resulting numerator for the binary contrast is the same as it was for the mean contrast. The denominator, on the other hand, is $1 - p_m$ instead of 1.

$$\text{Binary contrast}(m) = \frac{\sum_{i=1}^k p_i (E(y_m|\mathbf{x}) - E(y_i|\mathbf{x}))}{1 - p_m} \tag{4}$$

As written above, the only difference between the mean contrast and binary contrast is the denominator, meaning that:

$$\frac{\text{Binary contrast}(m)}{\text{Mean contrast}(m)} = \frac{1}{1 - p_m}$$

The binary contrast can therefore be obtained from the mean contrast for the same category by multiplying the mean contrast by $1/1 - p_m$. This implies that the mean and binary contrast for a given category will always be the same sign, and that the coefficients and standard errors for binary contrasts are always larger than those of the mean contrasts. For a given category, the proportional increase in the coefficient and standard error is the same (a factor of $1/1 - p_m$), implying that the p -value for the test of the contrast versus zero is identical for the mean contrast and binary contrast.

`binarycontrast` uses Stata's `contrast` command to automate the calculation of the binary contrast. Consider a regression model with polytomous regressor B that is specified with a reference category. We can use β_j for the estimated coefficient for the j th level of B , where $\beta_j = 0$ if the j th level of B is the reference category. We can specify a column vector β_B that contains $[\beta_j \dots \beta_k]$ for the k categories of B .

We can then transform β_B into various other contrasts of interest by specifying a contrast matrix $\mathbf{C}$ and computing $\mathbf{C}\beta_B$. For example, mean contrasts are given by defining $\mathbf{C}$ as a $k \times k$ matrix in which:

$$\mathbf{C}_{mn} = \begin{cases} 1 - p_m & \text{if } m = n \\ -p_n & \text{if } m \neq n \end{cases}$$

where p_m is the (weighted) proportion of observations in category m . For the binary contrast, we change this to:

$$\mathbf{C}_{mn} = \begin{cases} 1 & \text{if } m = n \\ -\frac{p_n}{1-p_m} & \text{if } m \neq n \end{cases}$$

In Stata, the mean contrast as defined above can be computed using the `contrast` command, using its `gw.` operator. `contrast` also allows the user to specify a custom contrast matrix. `binarycontrast` works by creating the custom contrast matrix and calling `contrast`. To construct the custom contrast matrix, `binarycontrast` requires the relative frequencies of the different categories, and it obtains these using Stata's `proportion` command. When calling `proportion`, `binarycontrast` uses the analytic sample and any weights that the user specified when fitting the model.

3 Syntax and stored results

```
binarycontrast varlist [ , mcompare(method) mean ]
```

for which *varlist* includes variables for which one wants contrasts. While polytomous variables should be specified when fitting the model using factor variable syntax (i.e., `i.region`), one should not re-use factor variable syntax with `binarycontrast` (i.e., `binarycontrast region`).

3.1 Options

`mean` computes mean contrasts instead of binary contrasts. The results are not different from using the `contrast` command directly.

`mcompare(method)` provides the method for correcting for multiple comparisons. This is passed along to the `contrast` command when it does its calculation. See [R] `contrast` for details.

3.2 Stored results

Scalars

`r(level)` level for confidence interval (as specified in estimation command)

Macros

`r(var)` variables for which results are provided
`r(contrast_type)` mean or binary
`r(var_values)` values of factor levels (in order indicated by `r(var)`)
`r(var_labels)` labels of factor levels (in varlist order)

Matrices

`r(table)` table of results
`r(proportions)` proportions of each category (via `proportion`)

4 Discussion

Although reference categories are the dominant way that coefficients for categorical regressors are presented, there is often little justification for unordered variables for foregrounding any particular category for interpretation. Basing interpretations on an arbitrary reference category may become increasingly strained for regressors with more categories. Both mean and binary contrasts provide coefficients for all categories that have a uniform interpretation, for which the sign of the coefficient corresponds to whether category membership is positively or negatively associated with the outcome.

The principal advantages of binary contrasts over mean contrasts are that the magnitude of the binary contrast for a category does not depend on the frequency of the category, and that its basic interpretation is the same as the familiar way of interpreting a coefficient for a dichotomous regressor. Two disadvantages of the approach bear emphasis. First, for mean contrasts, the contrast between any pair of categories remains available as a matter of simple subtraction, but this is not the case for binary contrasts. Instead, one would need to multiply the binary contrast by the proportion in the category—i.e., convert it to a mean contrast—to then be able to recover the contrast between a pair of categories. Second, while it is possible to extend binary contrasts to categorical interaction terms, the calculation is more cumbersome and the interpretation of the coefficient less clear, so `textttbinarycontrast` does not include support for interaction terms.