Multinomial GLMs for Multinomial Response with An Example of Brain Injury Recovery Stages

Frances Lin

Fall 2020

Multinomial Distribution

Binomial regression can be extended to multinomial regression for when response is multinomial (i.e. response has > 2 categories).

Here we assume that that the response (Y_{ij})

$$Y_{ij} \overset{iid}{\sim} Multinomial(n_i, \pi_{ij})$$

where π_{i1} is the probability of y_i falls into the first ($j = 1$) category, π_{i2} is the probability of y_i falls into the second ($j = 2$) category, etc.

Ordinal Regression for Ordinal Response

It is important to distinguish between nominal response (i.e. response that does not follow specific order) and ordinal response (i.e. response that has specific order) for multinomial response.

Some models (e.g. cumulative logit model) can only be applied to ordinal response, whereas some models (e.g. baseline-category logit model) can be applied to both nominal and ordinal response.

Data: GOS For Traumatic Brain Injury Assessment

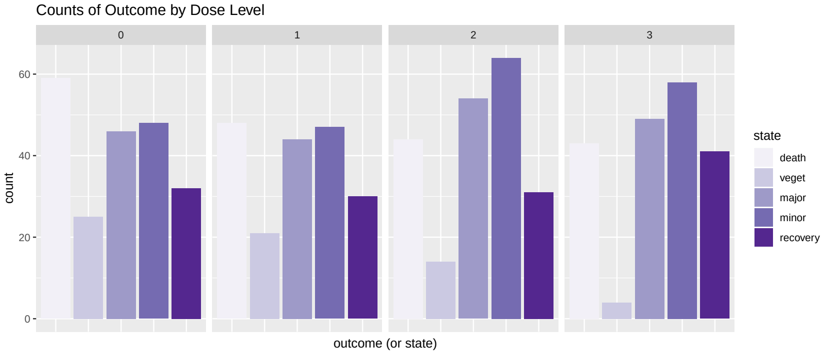
Data were obtained from the 1977 paper by Chuang-Stein and Agresti.The Glasgow Outcome Scale (GOS) is a scale that is commonly used for traumatic brain injury assessment.

GOS is the response, and outcome is ordinal with ordered categories: *death*, *vegetative state*, *major disability*, *minor disability*, and *good recovery*. Intravenous dose is the covariate, and dose is categorical with four levels: *placebo*, *low dose*, *medium dose*, and *high dose*.

dose	death	veget	major	minor	recovery
placebo	59	25	46	48	32
low	48	21	44	47	30
med	44	14	54	64	31
high	43	4	49	58	41

Initial Analysis

Initial analysis shows that the *placebo* group (title $= 0$) has a higher relative proportions of patients whose outcomes fall into the *death* category, and the *high dose* group (title $= 3$) has a higher relative proportions of patients whose outcomes fall into the *good recovery* category.



Model Fitting

Fitting the cumulative link model with different link function (e.g. `probit`, `cloglog`), we see that the model with the `logit` link performs the best, but the difference is minimal.

	logit	probit	cloglog
AIC	2475	2477	2478

Model Interpretation

Exponentiating the coefficients and the intercepts, we conclude that

1. all dose levels lead to favorable outcomes
2. the odds of a patient being in the more favorable categories increases as dose level increases
3. the estimated odds of outcome (y) for *low dose* (dose = 1) is 1.125 times the estimated odds for *placebo* (dose = 0), etc

factor(dose)1	factor(dose)2	factor(dose)3
1.125	1.374	1.683

Model Interpretation

4. for the reference level *placebo* (dose $= 0$), the estimated odds of outcome falls into category 1 (*death*) versus all other categories is 0.399, the estimated odds of outcome falls into category 1 (*death*) or category 2 (*vegetative state*) versus all other categories is 0.5956, etc

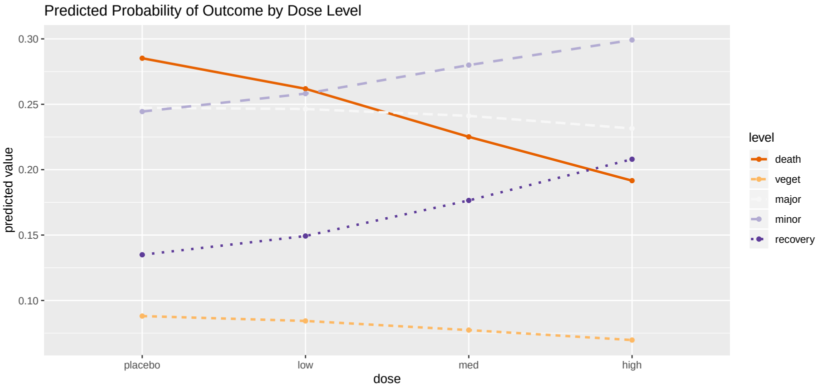
death|veget	veget|major	major|minor	minor|recovery
0.399	0.5956	1.636	6.41

The cumulative logit model is defined as

$$logit(P(Y \leq j)) = \log(\frac{P(Y \leq j)}{1 - P(Y \leq j)}) = \log(\frac{\pi_j}{1 - \pi_j}) = X\beta, j = 1, 2, ...j$$

Prediction (Predicted Probability)

dose	death	veget	major	minor	recovery
placebo	0.2852	0.08806	0.2474	0.2444	0.135
low	0.2619	0.08434	0.2464	0.2582	0.1493
med	0.2251	0.07735	0.2411	0.28	0.1765
high	0.1916	0.06973	0.2315	0.2992	0.208



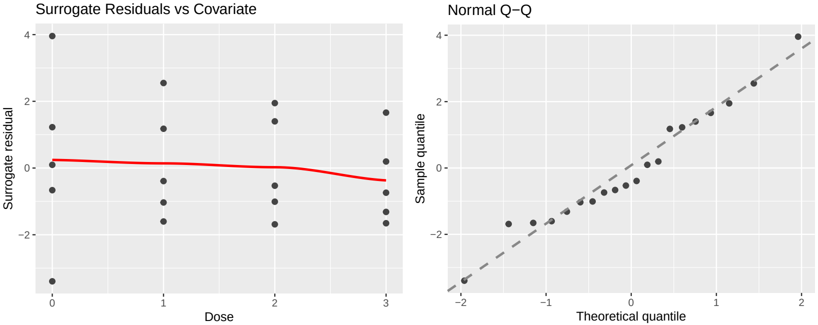
Collapsing Categories

Next, we collapse five (*death, veget, major, minor, recovery*) to three categories (*unfavorable, major, favorable*) and see that AIC value drops significantly. However, we should note that it is advisable to check literature or domain knowledge before collapsing categories, as it may sometimes result in loss of information or aggregation bias.

	logit	logit_collapse
AIC	2475	1712

Discussion

Major issues with model fit are perhaps not having enough data points and only including dose as the only covariate, but here concludes the demonstration of fitting an ordinal regression for ordinal response.



Thank you!

Multinomial GLMs for Multinomial Response with An Example of Brain Injury Recovery Stages

Frances Lin

MS student, Dept. of Statistics, Oregon State University