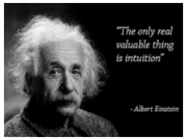
Multinomial Logistic Models

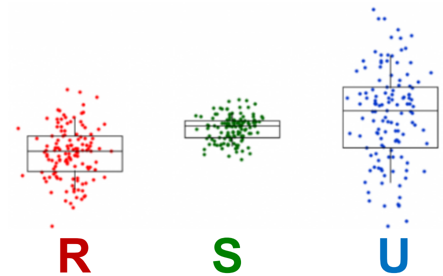
Kogod School of Business
American University • Washington, DC



KOGOD SCHOOL *of* BUSINESS
AMERICAN UNIVERSITY • WASHINGTON, DC

Multinomial Logistic: Intuition

- In **binomial logit** the response variable has **2 possible values** (e.g., "good mail" and "spam"), so we can easily quantify this variable with 2 dummy variables coded **0** (e.g., good mail) and **1** (e.g., spam)
- If the response variable has **more than 2 categories** we have a **multinomial logistic** situation
- For example, if we are trying to find the type of **home location** a person is likely to purchase (e.g., **rural**, **suburban**, **urban**).



- We can solve this problem **fitting 2 binomial models**.
- We first select the **"reference"** category (e.g., rural) that will be coded **0** in **both models**.
- We then fit **two binomial models**, one for **suburban vs. rural** and another for **urban vs. rural**.
- More generally, a response variable with **K categories** can be modeled with **K-1 binomial** logistic models. The category left out is called the **"reference"** category.
- **Multinomial logistic** regression **fits** all the binomials **together**

Multinomial Logistic Details

- A **multinomial** model can solved by **fitting binomial** logistic **models one by one**, but this yields separate fit statistics for model
- A **multinomial** logistic model **fits** all the binomials **together** and provides fit statistics for the model as a whole
- The response variable Y has K categories
  - We select one of them as the **"reference"** category k_R
  - We estimate logistic **coefficients** for the other $K\text{-}1$ categories
  - In **binomial** logistic, a coefficient for X_1 represents the increase in log odds of $Y = 1$, relative to $Y = 0$, when X_1 increases by 1 unit.
  - In **multinomial** logistic, the coefficient for a predictor X_{1k} for the k^{th} category represents the increase in log odds of $Y = k$, relative $Y = k_R$, when X_1 increases by 1 unit.

Multinomial Logistic Interpretation

For a **multinomial logit** model:

$$Logit(Y) = \beta_{01} + \beta_{11}(X_1) + \beta_{21}(X_2) + \beta_{31}(X_3) + \ldots\ldots$$
$$\beta_{02} + \beta_{12}(X_1) + \beta_{21}(X_2) + \beta_{31}(X_3) + \ldots\ldots$$
$$\ldots\ldots$$
$$\beta_{02} + \beta_{1K-1}(X_1) + \beta_{2K-1}(X_2) + \beta_{3K-1}(X_3) + \ldots\ldots + \varepsilon$$

- Y has K categories (e.g., Freshman, Sophomore, Junior, Senior withdrawing from school)
- $Logit(Y)$ is the **log odds** of the response variable being in category k (e.g., Junior), **relative** to the **reference** category k_R (e.g., Freshman)
- β_{ik} is the **effect** of variable X_i (e.g., GPA) on the log odds of $Y = k$
- But **odds** and log odds are always **relative** to something (e.g., leaving vs. staying)
- So, β_{ik} measures how much the log odds Y of being k, relative to the reference category k_R (e.g., Freshman), **change** when X_i (e.g., GPA) goes up by 1 unit.

Multinomial Logit: Fit Statistics

- The **vglm(){VGAM}** function in **R** reports the:
  - **Log-Likelihood**
  - **Deviance** (2LL) → $-2 * Log\ Likelihood$
  - **AIC** → calculated by adding $2 * Number\ of\ Variables$

Confusion Matrix:

$$Error\ Rate = \frac{Incorrect}{Total} = \frac{Off - Diagonal}{T}$$

$$Sensitivity_{Class} = \frac{Correct_{Class}}{Total_{Class}} = \frac{TP_A}{A_A}; \frac{TP_B}{A_B}; \frac{TP_C}{A_C}$$

$$Specificity_{Class} = \frac{Correct_{NotClass}}{Total_{NotClass}} =$$

$$\frac{TP_B+TPC}{A_B+AC}(for\ A); \frac{TP_A+TPC}{A_A+AC}(for\ B); \frac{TP_A+TPB}{A_A+AB}(for\ C)$$

Pred	Actual			Total
	A	B	C	
A	TP_A	$P_{A/B}$	$P_{A/C}$	P_A
B	$P_{B/A}$	TP_B	$P_{B/C}$	P_B
C	$P_{C/A}$	$P_{C/B}$	TP_C	P_C
Total	A_A	A_B	A_C	T

R Tips

`vglm(){VGAM}` → "Vector Generalized Linear Model" function in the `{VGAM}` "Vector Generalized and Additive Model" package is a function to fit multinomial logistic; "vector" refers to the fact that the response variable is no longer binary, but a vector of categorical values.

Note: there are other packages that can fit multinomial logit models, with their own strengths, difficulties and challenge for you to explore, e.g.: `multinom(){nnet}` and `glmnet(){glmnet}`

```
vglm.fit =
  vglm(y~x1+x2+etc.,family=multinomial(refLevel=1),
  data=dataName)
```

→ refLevel=1 indicates the first category in the response variable will be left out as a baseline

Like binary logistic with `glm()` → `logLik(vglm.fit)`; `-2*logLik(vglm.fit)`; `deviance(vglm.fit)`; and `AIC(vglm.fit)` work well with a `vglm()` object

Kogod School
of
Business