Part II: Generalized Linear Models

Chapter II.4

Models for Nominal/Ordinal Response

Topics

To be discussed...

- Multinomial Response
- Generalized Odds
- Logistic Regression for Nominal Response
- Multivariate GLM
- Logistic Regression for Ordinal Response

Models for Nominal/Ordinal Response

The response variable Y is categorical with J $(J > 2)$ response categories.

II.4.1 Remark (multinomial trials)

For subject i $(i = 1, \dots, n)$, let π_{ij} denote the probability of response in category j $(j = 1, \dots, J)$. Then it holds $\sum_{j=1}^{J} \pi_{ij} = 1$ and

$$\boldsymbol{Y}_i \sim \mathcal{M}(1, \boldsymbol{\pi}_i) \quad i = 1, \dots, n,$$

where $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{iJ})$ is the multinomial trial of subject i with $Y_{ij} = 1$ when the response is in category j and $Y_{ij} = 0$ otherwise, and $\boldsymbol{\pi}_i = (\pi_{i1}, \dots, \pi_{iJ})$ is the associated probability vector with $\boldsymbol{\pi}_i \in [0,1]^J$. It holds $\sum_{j=1}^{J} Y_{ij} = 1$ and

$$P(\boldsymbol{Y}_i = \boldsymbol{y}_i) = \prod_{j=1}^{J} \pi_{ij}^{y_{ij}} \ .$$

☞ $\dim(\boldsymbol{\Theta}) = J - 1$ (s. Remark II.1.4: Multinomial Distribution with $n = 1$, $m = J - 1$).

Remark

We considered above ungrouped data. Multinomial GLMs can be expressed in terms of grouped or ungrouped data (similarly to binary GLMs). In case all explanatory variables are categorical, GoF statistics and residuals are considered for the grouped data.

Multinomial Logistic Regression

☞ We shall consider models separately for *nominal* and *ordinal* responses.

II.4.2 Set-up for nominal responses

- Let $\tilde{Y}_i$ be the *nominal* response category for subject i (out of J categories).
  $\tilde{Y}_i = j$ means that $Y_{ij} = 1$ and $Y_{i\ell} = 0$ for $\ell \neq j$, for the J multinomial indicators in $\boldsymbol{Y}_i$.
  ☞ $\mathsf{P}(\tilde{Y}_i = j) = \pi_{ij}$, $j = 1, \ldots, J$
- Model the dependence of $\mathsf{P}(\tilde{Y}_i = j)$, $j = 1, \ldots, J$, for subject i, on the explanatory variables $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})$, which can be categorical and/or continuous.
- Observations on $\boldsymbol{Y}_i$ for fixed $\boldsymbol{X}_i = \boldsymbol{x}_i$ are *multinomial distributed*.

☞ Multicategory logistic models simultaneously describe the log odds (logits) for all $J(J-1)/2$ pairs of categories.

- Given a certain choice of $J-1$ logits (for certain reference category or for adjacent categories), the rest are redundant. Such a set is known as *minimal set* of logits.

Generalized Odds for Categorical Responses

II.4.3 Definition (baseline–category logits)

Each response category is paired with a baseline-category (category 1 or J). The logits with baseline (reference) category the last (J) one are defined by

$$\log\frac{\pi_{i1}}{\pi_{iJ}},\ \log\frac{\pi_{i2}}{\pi_{iJ}},\ldots,\ \log\frac{\pi_{i,J-1}}{\pi_{iJ}}.$$

II.4.4 Remark

The j-th baseline-category logit, $\log\frac{\pi_{ij}}{\pi_{iJ}}$, is the logit of the <u>conditional</u> probability

$$\text{logit}[P(Y_{ij}=1|Y_{ij}=1 \text{ or } Y_{iJ}=1)] = \log\frac{P(Y_{ij}=1|Y_{ij}=1 \text{ or } Y_{iJ}=1)}{1-P(Y_{ij}=1|Y_{ij}=1 \text{ or } Y_{iJ}=1)} = \log\frac{\pi_{ij}}{\pi_{iJ}}$$

Let $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})$ denote the explanatory variable values for subject i, and let $\boldsymbol{\beta}_j = (\beta_{j1}, \ldots, \beta_{jp})'$ be the parameters for the j-th logit.
Then, the baseline-category logits are modeled as

$$\log \frac{\pi_{ij}}{\pi_{iJ}} = \boldsymbol{x}_i \boldsymbol{\beta}_j = \sum_{k=1}^{p} \beta_{jk} x_{ik} = \beta_{j1} + \sum_{k=2}^{p} \beta_{j,k} x_{i,k}, \quad j = 1, \ldots, J-1.$$

Typically $x_{i1} = 1$ for the intercept term β_{j1}, which also differs for each logit.
☞ This model describes simultaneously the effects of $\mathbf{x}$ on the $J-1$ logits.

II.4.5 Remark

1. These $J-1$ equations determine equations for logits for any other pairs of response categories, since
$$\log \frac{\pi_{ia}}{\pi_{ib}} = \log \frac{\pi_{ia}}{\pi_{iJ}} - \log \frac{\pi_{ib}}{\pi_{iJ}} = \boldsymbol{x}_i(\boldsymbol{\beta}_a - \boldsymbol{\beta}_b) \ .$$

2. The model treats the response variable as nominal in the sense that if the model holds and the outcome categories are permuted in any way, the model still holds with the corresponding permutation of the effects.

II.4.6 Definition (Baseline-Category Logit Model)

The multinomial (or baseline-category) logit model can be expressed in terms of the response probabilities as

$$\pi_{ij} = \frac{\exp(\boldsymbol{x}_i\boldsymbol{\beta}_j)}{1 + \sum_{\ell=1}^{J-1} \exp(\boldsymbol{x}_i\boldsymbol{\beta}_\ell)}, \quad j = 1, \dots, J-1, \tag{II.7}$$

for $i = 1, \dots, n$, with $\boldsymbol{\beta}_J = \mathbf{0}$ (for identifiability reasons).

II.4.7 Remark

1. It holds $\sum_{j=1}^{J} \pi_{ij} = 1$ for $i = 1, \dots, n$.
2. For $J = 2$, model (II.7) simplifies to the binary logistic regression model.
3. The interpretation of the effects conditional on response in category j or J is straightforward while an 'overall' interpretation needs caution: all $\boldsymbol{\beta}_\ell$ contribute to π_{ij} (not just $\boldsymbol{\beta}_j$):
   $\hookrightarrow$ Binary logistic regression: $\partial\pi_i/\partial x_{ik} = \beta_k\pi_i(1-\pi_i)$
   $\hookrightarrow$ Multinomial logistic regression: $\partial\pi_{ij}/\partial x_{ik} = \pi_{ij}\left(\beta_{jk} - \sum_{j'} \pi_{ij'}\beta_{j'k}\right)$.
4. The multinomial logit model is a *multivariate GLM*.

Multivariate GLM

A multivariate GLM applies to random components that have distribution in a multivariate generalization of the EDF.

II.4.8 Definition (Multivariate Exponential Dispersion Family)

For a GLM having a multivariate response, the random response vector $\boldsymbol{Y}_i \in \mathbb{R}^q$ of every subject i out of a random sample of size n $(i = 1, \dots, n)$, has a distribution in the *multivariate exponential dispersion family* (MEDF), with probability density function (pdf) or probability mass function (pmf) of the form

$$f(\boldsymbol{y}_i;\ \boldsymbol{\vartheta}_i, \phi) = \exp\left\{\frac{\boldsymbol{y}_i'\boldsymbol{\vartheta}_i - b(\boldsymbol{\vartheta}_i)}{a(\phi)} + c(\boldsymbol{y}_i, \phi)\right\}, \tag{II.8}$$

where $\boldsymbol{\vartheta}_i$ is the natural parameter and for some specific functions $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$, giving rise to different distributions in the family.

☞ Compare to Definition II.4.8 of (univariate) EDF.

Multivariate GLM

II.4.9 Multivariate GLM

Multivariate response: $\boldsymbol{Y}$, $\boldsymbol{Y} \in \mathbb{R}^q$ – Explanatory variables: $X_2, \ldots, X_p$

($X_1 = 1$ ☞ coefficient of the intercept)

1. Random component: $\boldsymbol{Y} \sim f(\boldsymbol{y}_i;\ \boldsymbol{\vartheta}_i, \phi)$ ☞ belongs to the multivariate EDF
   $\boldsymbol{\mu}_i = E(\boldsymbol{Y}_i)$ for subject i, $i = 1, \ldots, n$.
2. Systematic component (linear predictor): $\mathbf{X}_i \boldsymbol{\beta}$,
   where $\mathbf{X}_i$ is the model matrix for observation i.
3. Link function g relating linear predictor to $\boldsymbol{\mu}_i$: $\mathbf{g}(\boldsymbol{\mu}_i) = \mathbf{X}_i \boldsymbol{\beta}$,
   with $\mathbf{g} = (g_1, \ldots, g_k)$.
   ☞ canonical link: $\mathbf{g}(\boldsymbol{\mu}_i) = \boldsymbol{\vartheta}_i$

II.4.10 Example

The multinomial distribution is a member of the MEDF (s. Example II.2.5: binomial distribution). Consider $\boldsymbol{Y} = (Y_1, \ldots, Y_J) \sim \mathcal{M}(m, \boldsymbol{\pi})/m$ with $\sum_{j=1}^{J} \pi_j = 1$. Then the *pmf* is

$$f(\boldsymbol{y}; \boldsymbol{\vartheta}, \phi) = \exp\left\{ \log \binom{m}{my_1 \ldots, my_J} + \sum_{j=1}^{J-1} my_j \log(\pi_j) + (m - \sum_{j=1}^{J-1} my_j) \log(\pi_J) \right\},$$

which trasforms to

$$f(\boldsymbol{y}; \boldsymbol{\vartheta}, \phi) = \exp\left\{ \sum_{j=1}^{J-1} my_j \log\left(\frac{\pi_j}{\pi_J}\right) - (-m \log(\pi_J)) + c(\boldsymbol{y}, \phi) \right\}$$

☞ natural parameter: $\boldsymbol{\vartheta} = \left(\log(\frac{\pi_1}{\pi_J}), \ldots, \log(\frac{\pi_{J-1}}{\pi_J})\right)$ – dispersion: $\phi = 1$ – weight: $w = m$

☞ In our set–up, we have n multinomial distributions $\boldsymbol{Y}_i \sim \mathcal{M}(m_i, \boldsymbol{\pi}_i)/m_i$, $i = 1, \ldots, n$, one for each subject in the sample (for ungrouped data $m_i = 1$).

Multinomial GLMs

II.4.11 Multinomial logistic regression as a GLM

Categorical response: Y, $Y \in \{1, \ldots, J\}$ – Explanatory variables: $X_2, \ldots, X_p$
($X_1 = 1$ ☞ coefficient of the intercept)

1. Random component: $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i,J-1})' \sim \mathcal{M}(1, \boldsymbol{\pi}_i)$ ☞ belongs to the multivariate EDF $\boldsymbol{\mu}_i = E(\mathbf{Y}_i) = (\pi_{i1}, \ldots, \pi_{i,J-1})'$ for subject i, $i = 1, \ldots, n$.

2. Systematic component (linear predictor): $\mathbf{X}_i\boldsymbol{\beta}$
Model matrix for observation i: $\mathbf{X}_i \in \mathbb{R}^{(J-1)\times p(J-1)}$ is block diagonal having on the diagonal the row vector $\mathbf{x}_i \in \mathbb{R}^{1\times p}$.
Parameter vector: $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{J-1})' \in \mathbb{R}^{p(J-1)}$
(i.e. consists of one vector per every response level).

3. Canonical link function $\mathbf{g}$ relating linear predictor to $\boldsymbol{\mu}_i$: $\mathbf{g}(\boldsymbol{\mu}_i) = \mathbf{X}_i\boldsymbol{\beta} = \boldsymbol{\vartheta}_i$
with $g_j(\boldsymbol{\mu}_i) = \log\left(\frac{\mu_{ij}}{1-\sum_{\ell=1}^{J-1}\mu_{i\ell}}\right) = \vartheta_{ij}, \; j = 1, \ldots, J-1$.

II.4.12 Inference for the baseline–category logit model

- Maximum likelihood fitting of the baseline–category logit model maximizes the *multinomial loglikelihood* (s. Example II.4.10) subject to $\boldsymbol{\pi} = (\pi_{ij})$ that satisfy the $J-1$ equations of the model (i.e. express (π_{ij}) in terms of $\boldsymbol{\beta}$ according to model II.4.11(3) and maximize the loglikelihood in terms of $\boldsymbol{\beta}$).
- All results and discussions on estimation, model selection and goodness of fit testing for the univariate GLMs extend straightforwardly to multivariate GLMs.

II.4.13 Example (Alligators)

Agresti (2013, CDA) Sample of 219 alligators in four Florida lakes ☞ Data file: alligators.dat
Response: Primary Food Choice of Alligators (Fish, Invertebrate, Reptile, Bird, Other)

Explanatory variables:
Lake (Hancock, Ocklawaha, Trafford, George)
Size of the Alligator (≤ 2.3m, > 2.3m)

size= **0**	food				
lake	**1**	**2**	**3**	**4**	**5**
1	7	0	1	3	5
2	13	8	6	1	0
3	8	7	6	3	5
4	17	1	0	1	3

size=**1**	food				
lake	**1**	**2**	**3**	**4**	**5**
1	23	4	2	2	8
2	5	11	1	0	3
3	5	11	2	1	5
4	16	19	1	2	3

```
> alligators <- read.table("alligators.dat", header=T)
> alligators[1:2,] # each response category in a separate column!
    lake   size   y1   y2   y3   y4   y5
1      1      1   23    4    2    2    8
2      1      0    7    0    1    3    5
```

Baseline-Category Logit Model in R

```
> library(VGAM)   # vglm() function (reference category is the last):
> fit <- vglm(formula = cbind(y2,y3,y4,y5,y1) ~ size + factor(lake), family= multinomial,
      data=alligators)   # fish=1 is baseline category (last in cbind() above)
> summary(fit)    # part of the output follows
Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept):1 -3.2074 0.6387 -5.021 5.13e-07 ***
(Intercept):2 -2.0718 0.7067 -2.931 0.003373 **
(Intercept):3 -1.3980 0.6085 -2.297 0.021601 *
(Intercept):4 -1.0781 0.4709 -2.289 0.022061 *
size:1 1.4582 0.3959 3.683 0.000231 ***
size:2 -0.3513 0.5800 -0.606 0.544786
⋮
Names of linear predictors: log(mu[,1]/mu[,5]), log(mu[,2]/mu[,5]),
log(mu[,3]/mu[,5]), log(mu[,4]/mu[,5])

Residual deviance: 17.0798 on 12 degrees of freedom

Warning: Hauck-Donner effect detected in the following estimate(s): '(Intercept):1'

# p-value:
> pchisq(17.07983,df=12, lower.tail=FALSE)
[1] 0.146619
```

Hauck-Donner effect (HDE)

Attention is required with the interpretation of the results when a HDE occurs. HDE leads to an upward biased p-value and loss of power.[a]

[a]HDE: a Wald test statistic is not monotonic increasing as a function of increasing distance between the parameter estimate and the null value. This may be due to separability effects (see: T.W.Yee (2020): https://arxiv.org/abs/2001.08431).

II.4.14 Interpretation of Output (Example II.4.13)

Estimated odds ratio for food choice (adjusting for lake)

The odds that the primary food choice is `y2` (ivertebrates) instead of `y1` (fish) for small alligators is estimated to be 4.3 times higher than the corresponding odds for large alligators in the same lake:

```
> food <- coefficients(fit)[5:8]; exp(food[1])
size:1
4.298236
```

☞ Since the model has no interactions, the above odds ratio is the same for all lakes.

II.4.15 $(1-)100\%$ Wald CI for odds ratio (OR)

The following function can be used to derive asymptotic $(1-)100\%$ Wald CI for an OR, when the input is the $\log$(OR):

```
> CI.or <- function(x, se, conf.level=0.95) {
Za2 <- qnorm(0.5*(1+conf.level)); CI <- exp(x+c(-1,1)*Za2*se); return(CI) }
```

Implemented for the specific OR in Remark II.4.14 above:

```
> se <- c(0.39595, 0.58003, 0.64248, 0.44825)  # standard errors for size
> CI.or(food[1], se[1])
[1] 1.978140 9.339497
> CI.or(food[1], se[1], conf.level=0.99)
[1] 1.550078 11.918648
```

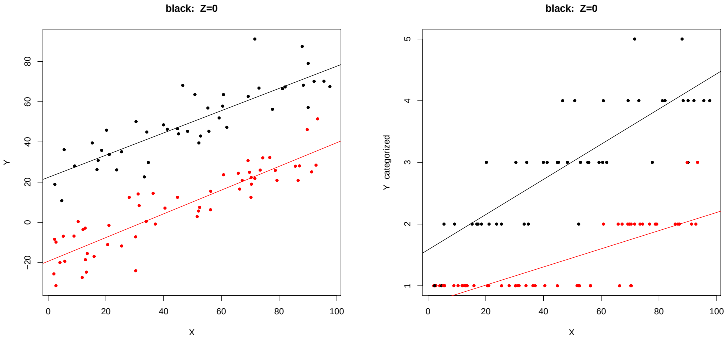
Why special methods for treating ordinal variables?

Why not just assign scores to the ordered categories and use standard regression methods (as many do)?

- Multinomial distribution is assumed for the response variable Y of J categories (no more normal).
- Estimated mean from the model fit may fall below the lowest or above the highest score (it should be: $1 \leq \hat{\boldsymbol{\mu}} = g^{-1}(\mathbf{x}\hat{\boldsymbol{\beta}}) \leq J$).
- For categorical data, the interest lies more on estimating category probabilities than means.
- Regardless of fitting method or distributional assumption, *ceiling effects* and *floor effects* can cause bias in results.

II.4.16 Example (floor effect)

Agresti (2010, *Analysis of Ordinal Categorical Data*, Section 1.3.3)



$$X \sim \mathcal{U}(0,\ 100)\ ,\quad P(Z=0) = P(Z=1) = 0.5\ ,$$

$$Y = 20 + 0.6X - 40Z + \epsilon\ ,\quad \epsilon \sim \mathcal{N}(0,\ 100),$$

$$Y_c = \begin{cases} 1\ , & Y \leq 20 \\ 2\ , & 20 < Y \leq 40 \\ 3\ , & 40 < Y \leq 60 \\ 4\ , & 60 < Y \leq 80 \\ 5\ , & Y > 80 \end{cases}$$

Logistic Regression for Ordinal Response

II.4.17 Set-up for ordinal responses

- Let $\tilde{Y}_i$ be the *ordinal* response category for subject i (out of J categories).
  $\tilde{Y}_i = j$ means that $Y_{ij} = 1$ and $Y_{i\ell} = 0$ for $\ell \neq j$, for the J multinomial indicators in $\boldsymbol{Y}_i$.
  Thus, $\mathsf{P}(\tilde{Y}_i = j) = \pi_{ij}$ and $\mathsf{P}(\tilde{Y}_i \leq j) = \pi_{i1} + \ldots + \pi_{ij}$, $j = 1, \ldots, J$
- Model the dependence of $\mathsf{P}(\tilde{Y}_i = j)$ or $\mathsf{P}(\tilde{Y}_i \leq j)$, $j = 1, \ldots, J$, for subject i, on its explanatory variables values $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})$, which can be categorical and/or continuous.

II.4.18 Definition (logits for ordinal responses)

For an ordinal response variable Y with J levels, the following logits can be defined (among others) for $j = 1, \ldots J - 1$:

- cumulative logit: $\log\left(\frac{\mathsf{P}(Y \leq j)}{1 - \mathsf{P}(Y \leq j)}\right) = \log\left(\mathsf{P}(Y \leq j)/\mathsf{P}(Y > j)\right)$
- adjacent categories logit: $\log\left(\mathsf{P}(Y = j)/\mathsf{P}(Y = j + 1)\right)$
- continuation-ratio logit: $\log\left(\mathsf{P}(Y = j)/\mathsf{P}(Y > j)\right)$

☞ Logistic regression models can be defined for any type of logit above.

II.4.19 Definition (Cumulative Logit Model)

The cumulative logit model (CLM) is defined as

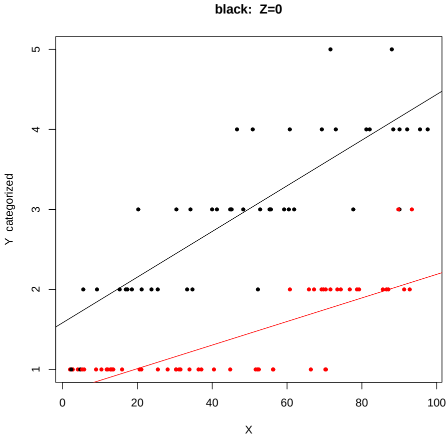
$$\text{logit}[\mathrm{P}(\tilde{Y}_i \le j)] = \alpha_j + \boldsymbol{x}_i\boldsymbol{\beta}, \qquad j = 1, \ldots, J-1. \tag{II.9}$$

for $i = 1, \ldots, n$.

II.4.20 Remark

1. A CLM for an ordinal response of J categories, consists of $J-1$ equations, one for each cumulative logit.
2. Each cumulative logit has its own intercept.
3. The $\{\alpha_j\}$ are increasing in j (☞ for fixed $\boldsymbol{x}_i$, the cumulative probability $\mathrm{P}(\tilde{Y}_i \le j)$ increases in j and the logit is increasing in $\mathrm{P}(\tilde{Y}_i \le j)$).
4. Intercepts depend on j but the other effects do not!
5. CLM treats the response variable as ordinal:
if the order of the response categories is reversed, the model continues to hold with a change in the sign of its parameters.

II.4.21 Example II.4.16 (continued)



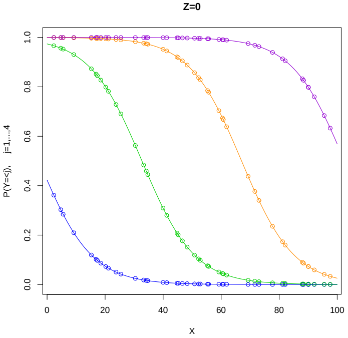
$$Z = 0\text{:}\ \ Y = 1, \ldots, 5$$

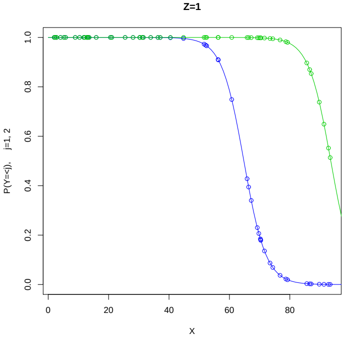
$$\hookrightarrow \ \frac{\mathsf{P}(Y \leq j)}{\mathsf{P}(Y > j)} \ , \quad j = 1, \ldots, 4$$

$$Z = 1\text{:}\ \ Y = 1, 2, 3$$

$$\hookrightarrow \ \frac{\mathsf{P}(Y \leq j)}{\mathsf{P}(Y > j)} \ , \quad j = 1, 2$$

Cumulative Logit Model with Proportional Odds: $P(Y \leq j) = \frac{\exp(\alpha_j + \beta X)}{1+\exp(\alpha_j+\beta X)}, \quad j = 1, \dots, J-1.$





$$\hat{\alpha}_j = \begin{cases} -0.31\ , & j = 1 \\ 3.61\ , & j = 2 \\ 7.41\ , & j = 3 \\ 11.32\ , & j = 4 \end{cases}$$

$$\hat{\alpha}_j = \begin{cases} 17.60\ , & j = 1 \\ 25.42\ , & j = 2 \end{cases}$$

$\hat{\beta} = -0.11,$

$\hat{\beta} = -0.27$

Cumulative Logit Model (CLM)

II.4.22 Definition (Cumulative OR)

An odds ratio of cumulative probabilities is called a cumulative odds ratio:

$$COR_{ij}(\boldsymbol{a},\boldsymbol{b}) = \frac{\mathsf{P}(\tilde{Y}_i \le j|\boldsymbol{x}_i = \boldsymbol{a})/\mathsf{P}(\tilde{Y}_i > j|\boldsymbol{x}_i = \boldsymbol{a})}{\mathsf{P}(\tilde{Y}_i \le j|\boldsymbol{x}_i = \boldsymbol{b})/\mathsf{P}(\tilde{Y}_i > j|\boldsymbol{x}_i = \boldsymbol{b})}$$

II.4.23 Remark ('Proportional Odds' Property)

A CLM satisfies:

$$\log[COR_{ij}(\boldsymbol{a},\boldsymbol{b})] = \mathsf{logit}[\mathsf{P}(\tilde{Y}_i \le j|\boldsymbol{x}_i = \boldsymbol{a})] - \mathsf{logit}[\mathsf{P}(\tilde{Y}_i \le j|\boldsymbol{x}_i = \boldsymbol{b})] = (\boldsymbol{a}-\boldsymbol{b})\boldsymbol{\beta}\ ,$$

meaning that the odds that the response $\le j$ at $\boldsymbol{x}_i = \boldsymbol{a}$ are $\exp[(\boldsymbol{a}-\boldsymbol{b})\boldsymbol{\beta}]$ times the odds at $\boldsymbol{x}_i = \boldsymbol{b}$, for every j.

- The log cumulative OR is proportional to the distance between $\boldsymbol{a}$ and $\boldsymbol{b}$.
- For each j, the odds that $\tilde{Y}_i \le j$ multiply by $\exp(\beta_k)$ per 1-unit increase in x_{ik}, adjusting for all the other explanatory variables.
- The same proportionality constant applies to all $J-1$ cumulative logits; that is, the effect (parameter $\boldsymbol{\beta}$) does not depend on j.

II.4.24 Example (detecting trend in dose response)

Effect of intravenous medication doses on patients with subarachnoid hemorrhage trauma

Response: Glasgow Outcome Scale
(Death, Veget. State, Major Disab., Minor Disab., Good Recov.)

Explanatory variable: Treatment (Placebo, Low dose, Med dose, High dose).

	response				
dose	1	2	3	4	5
1	59	25	46	48	32
2	48	21	44	47	30
3	44	14	54	64	31
4	43	4	49	58	41

II.4.25 Example II.4.24 in `R`

The cumulative logit model (II.9) can be fitted in the packages `VGAM` and `MASS`. The related commented code can be found in file `Cumulative_Logit_Model (trauma).pdf`. The data are provided in files '`trauma.dat`' (in format for `VGAM`) and '`trauma2.dat`' (in format for `MASS`).

Other Cumulative Link Models

II.4.26 Definition (Cumulative Link Model)

Let G^{-1} be a link function that is the inverse of a continuous cdf G. The *cumulative link model* is defined as

$$G^{-1}[\mathsf{P}(\tilde{Y}_i \leq j)] = \alpha_j + \boldsymbol{x}_i\boldsymbol{\beta}, \qquad j = 1, \dots, J-1.$$

for $i = 1, \dots, n$.

II.4.27 Remark

- As in the cumulative logit model with proportional odds, effects are the same for each cumulative probability (independent of j).
  This holds, when a latent variable Y^* satisfies a linear model with the cdf of the error term being G (compare to Remark II.3.6).
- If G is the standard normal cdf Φ, then the the cumulative link model becomes the *cumulative probit model*, which is a generalization of the binary probit model for ordinal responses.