GENERALIZED LINEAR MODELS
Ordinal Logistic Regression

Γιώτα Τουλούμη
Καθηγήτρια Βιοστατιστικής και Επιδημιολογίας
Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής
Ιατρική Σχολή Πανεπιστημίου Αθήνας

gtouloum@med.uoa.gr

Ordinal scales

- Ordinal scales occur very frequently (examples: food-testing, classification of radiographs, determination of physical or mental well-being)
- The choice and the definition of the response categories: **arbitrary** or **subjective**
- If a new category is formed by combining adjacent categories of the old scale, the form of the conclusions should be unaffected
- The above lead to models based on cumulative response probabilities rather than category probabilities

$$\gamma_j = \Pr(Y \le j) \quad \text{rather than} \quad \pi_j = \Pr(Y = j)$$

- The two sets of probabilities are equivalent, but cumulative probabilities are likely to have better properties
- GLM models with link functions:

$$\text{logistic scale} : \log\{\gamma_j /(1-\gamma_j)\}$$

$$\text{comlementary log}-\text{log scale} : \log\{-\log(1-\gamma_j)\}$$

Models for ordinal scales

$$\log\{\gamma_j(x)/(1-\gamma_j(x))\} = \kappa_j - \beta^T x \quad , j = 1,\ldots,I-1 \qquad I = \#\,of\ \text{categories} \qquad (1)$$

where $\gamma_j = pr(Y \le j|\mathrm{x})$ cumulative probability up to and including category j

Model 1 is known as the proportional-odds model because the ratio of the odds of the event Y≤j at $x=x_1$ and $x=x_2$ is:

$$\frac{\gamma_j(x_1)/(1-\gamma_j(x_1))}{\gamma_j(x_2)/(1-\gamma_j(x_2))} = \exp\{-\beta^T(x_1-x_2)\}$$

which is independent of the choice of the category (j).

The negative sign in (1) is a convention ensuring that large values of $\beta^T x$ lead to an increase of probability in the higher-numbered categories.

$\boldsymbol{\kappa}$ must satisfy $\kappa_1 \le \kappa_2 \le \ldots \le \kappa_{I-1}$

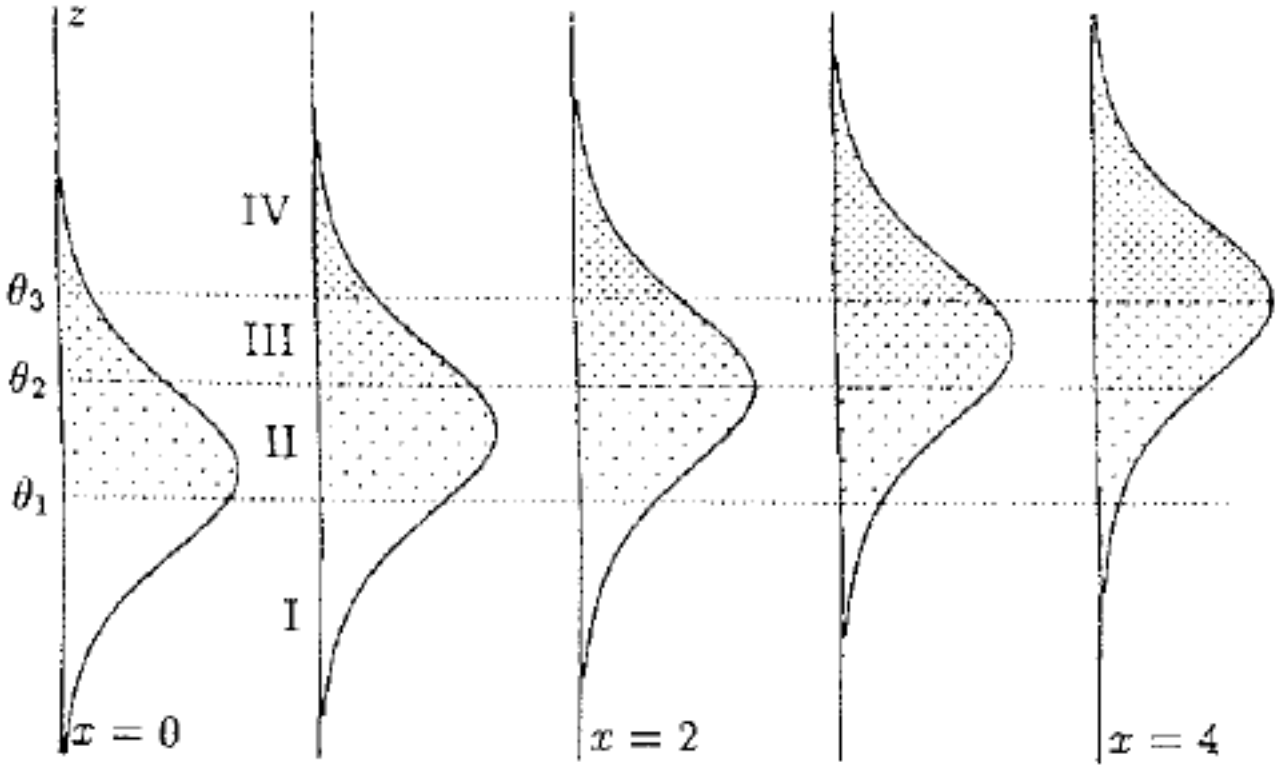


Fig. 5.1a. *Diagram showing how the response probabilities for the logistic model* (5.1) *vary with* x *when* $\beta > 0$. *Response categories are represented as four contiguous intervals of the* z*-axis. Higher-numbered categories have greater shade density.*

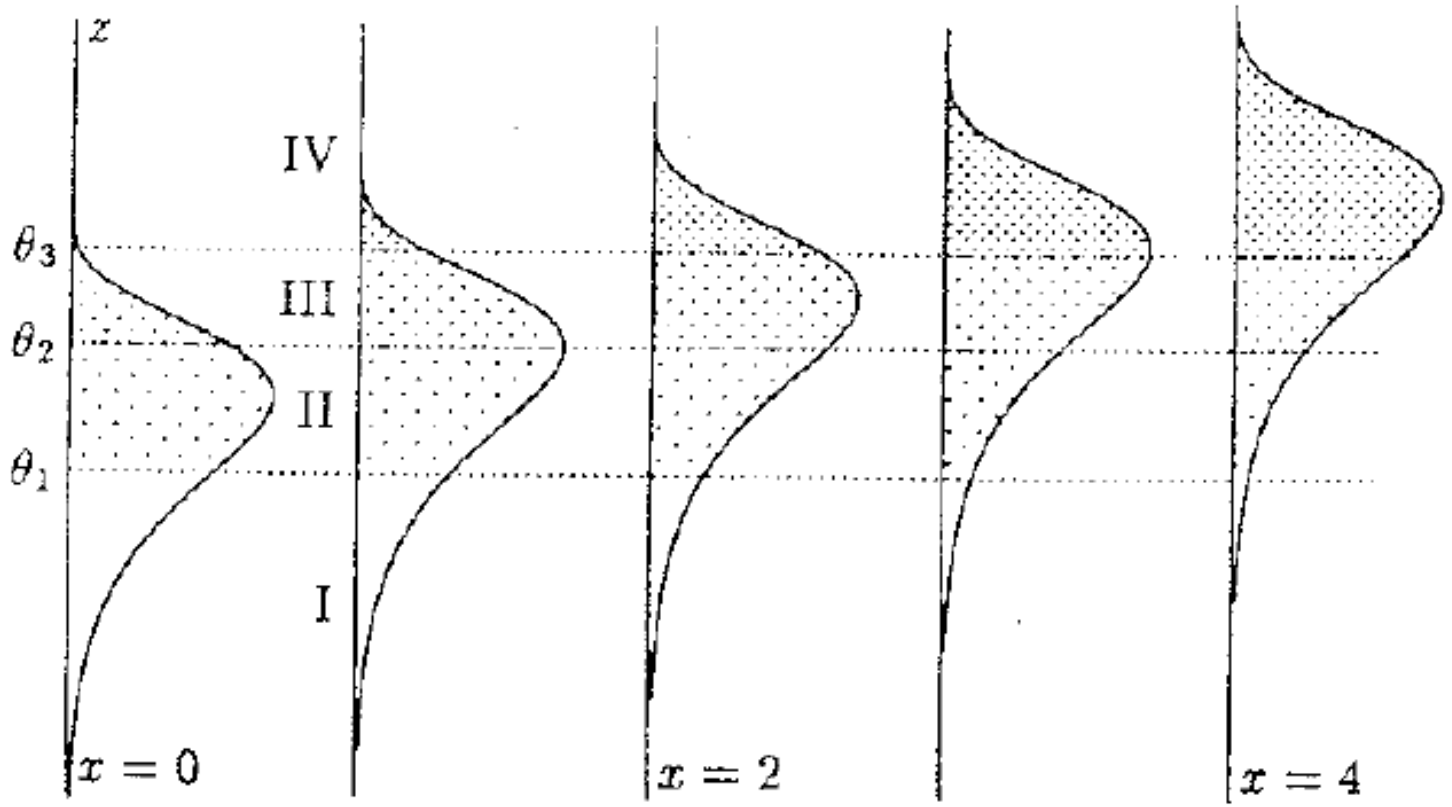


Fig. 5.1b. *Diagram showing how the probabilities for the four response categories in the complementary-log-log model* (5.3) *vary with* x *when* $\beta > 0$. $\pi_1(x)$ *and* $\pi_4(x)$ *each change by a factor of* 10 *or more, whereas* $\pi_3(x)$ *is almost constant over* $1 \leq x \leq 4$.

Logistic regression: a latent-variable approach

Imagine a continuous measure U, which is related with the outcome so that the probability of "success" $\pi = P(Y=1) = P(U > \kappa)$. Conversely, $1-\pi = P(Y=0) = P(U \le \kappa)$. In other words, "success" occurs when the underlying (latent) variable attains measurements above a threshold κ.

The logit model can be derived as follows: The continuous variable U is related to a set of explanatory variables $X_1, X_2, \cdots, X_p$ by a linear model

$$U = \beta_1^* X_1 + \cdots + \beta_p^* X_p + \varepsilon = \boldsymbol{\beta}^{*\prime}\mathbf{X} + \varepsilon$$

where ε is distributed according to distribution $F(\varepsilon)$. Notice the absence of an intercept in this formulation (in some references, the intercept is retained and is grouped later with the threshold κ).

The latent-variable approach*

From above,

$$\pi = P(U > \kappa)$$

$$= P(U - \boldsymbol{\beta}^{*\prime}\mathbf{X} > \kappa - \boldsymbol{\beta}^{*\prime}\mathbf{X}) \quad .$$

$$= P\left[\varepsilon > \kappa - \boldsymbol{\beta}^{*\prime}\mathbf{X}\right] = 1 - F(\kappa - \boldsymbol{\beta}^{*\prime}\mathbf{X})$$

Thus,

$$1 - \pi = F(\kappa - \boldsymbol{\beta}^{*\prime}\mathbf{X})$$

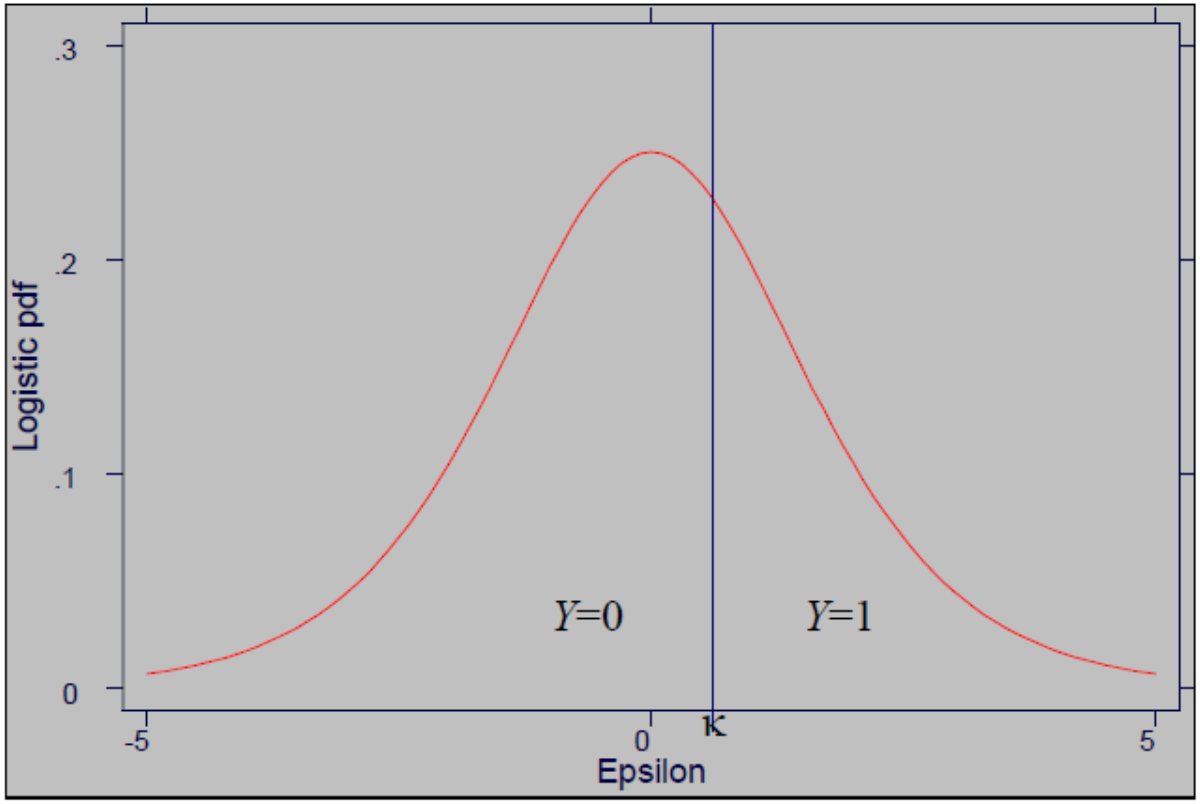
$$F^{-1}(1-\pi) = \kappa - \boldsymbol{\beta}^{*\prime}\mathbf{X} = \eta^*.$$

where $\eta^* = \kappa - \beta_1^* X_1 - \cdots - \beta_p^* X_p$. This is a generalized linear model with link $\eta = g(\pi) = F^{-1}(1-\pi)$ for some distribution F.

Notice that we defined $\pi = P(Y=1) = P(U > \kappa)$ to remain consistent with the results of the STATA output. Although this resulted in deriving the model in terms of the probability of "failure" it will not affect the results.

Latent-variable approach

The continuous measure U, which is related with the outcome so that the probability of "success" $\pi = P(Y=1) = P(U > \kappa)$ is shown below:



The latent-variable approach

The advantage of the latent-variable formulation is

- The model arises frequently in this manner in the real world
- This formulation aids in the understanding of other possibilities of distributions for ε (e.g., in probit or complementary log-log regression)
- This formulation leads to a more natural generalization to ordinal outcome variables

Logistic distribution

The logistic **random variable** X with mean μ and variance σ^2 has a cumulative distribution function

$$F(x \mid \mu, \sigma) = \{1 + \exp[-\pi(x-\mu)/(\sigma\sqrt{3})]\}^{-1}$$

$$-\infty < x < \infty$$

$$-\infty < \mu < \infty \qquad \sigma > 0, \qquad X \sim \text{logistic}(\mu, \sigma^2)$$

$$E(X) = \mu; \quad Var(X) = \frac{\pi^2\sigma^2}{3}$$

and density function

$$f(x \mid \mu, \sigma) = \frac{\pi}{\sigma\sqrt{3}} F(x \mid \mu, \sigma)[1 - F(x \mid \mu, \sigma)]$$

Consider the "*canonical form*" (μ=0, σ=1) which corresponds to the random variable Z with mean μ=0 and variance $\pi^2/3$. It has cumulative distribution and density functions:

$$F(Z) = \frac{1}{1+e^{-z}}$$

$$f(Z) = F(Z)[1 - F(Z)]$$

Logistic regression

If F(ε) is the logistic distribution then, $1-\pi = F(\eta^*) = \dfrac{1}{\left[1+e^{-\eta^*}\right]} \Leftrightarrow \pi = \dfrac{e^{-\eta^*}}{\left[1+e^{-\eta^*}\right]}$. Substituting n^* we have,

$$\pi = \frac{e^{(\beta_o+\beta_1 X_1+\cdots+\beta_p X_p)}}{\left[1+e^{(\beta_o+\beta_1 X_1+\cdots+\beta_p X_p)}\right]}, \text{where } \beta_o = -\kappa \text{ and } \beta_j = \beta_j^*$$

This is of course the *ordinary logistic regression* model. This can also be extended by setting $F \equiv \Phi$, the standard normal distribution function. This is the *probit* regression model $\eta = \Phi^{-1}(1-\pi)$

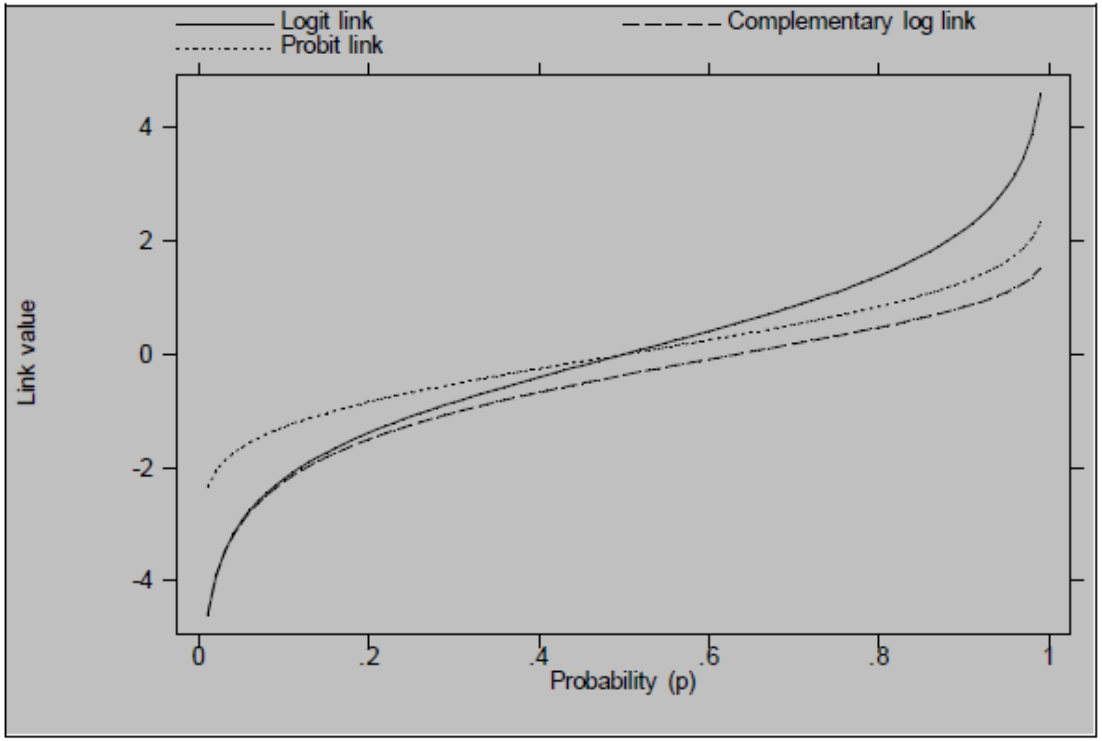
$$\pi = 1-\Phi(n^*) = P(Y=1) = \int_{\eta^*}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx$$

A less frequently used analysis involves the "complementary log-log" (or c-log-log) link

$$\eta = \log\left[-\log(1-\pi)\right]$$

which corresponds to the extreme-value distribution $\pi = F(\eta) = 1-e^{-e^{\eta}}$.

The logistic, probit and c-log-log links



The logistic regression and probit regressions tend to give similar results, although, since the logistic distribution ($\mu=0$, $\sigma=1$) has standard deviation $\pi/\sqrt{3}$, logistic-regression coefficients should be divided by $\pi/\sqrt{3}$ before compared to probit-regression coefficients.

Logistic regression analysis

We repeat here the analysis of the contraceptive use by fitting the two-factor additive model:

```
. char more[omit] 0

. xi: logit cuse i.more i.age [freq=N], nolog
i.more                Imore_0-1     (naturally coded; Imore_0 omitted)
i.age                 Iage_1-4      (naturally coded; Iage_1 omitted)

Logit estimates                                   Number of obs   =       1607
                                                  LR chi2(4)      =     128.88
                                                  Prob > chi2     =     0.0000
Log likelihood = -937.40449                       Pseudo R2       =     0.0643

------------------------------------------------------------------------------
       cuse |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
------------+-----------------------------------------------------------------
     Iage_2 |   .3678306   .1753673      2.097   0.036        .024117    .7115443
     Iage_3 |   .8077888   .1597533      5.056   0.000        .494678      1.1209
     Iage_4 |   1.022618   .2039337      5.014   0.000       .6229158    1.422321
    Imore_1 |  -.824092    .1171128     -7.037   0.000      -1.053629   -.5945552
      _cons |  -.8698414   .1571298     -5.536   0.000      -1.17781    -.5618727
------------------------------------------------------------------------------
```

Analysis as an ordinal logistic regression

The former model can be analyzed as an ordinal regression model. We can imagine some latent measure U such that $P(U>\kappa)=P(Y=1)$. The output of the command `ologit` is as follows:

```
. xi: ologit cuse i.more i.age [freq=N], nolog table
i.more                Imore_0-1     (naturally coded; Imore_0 omitted)
i.age                 Iage_1-4      (naturally coded; Iage_1 omitted)

Ordered logit estimates                           Number of obs   =       1607
                                                  LR chi2(4)      =     128.88
                                                  Prob > chi2     =     0.0000
Log likelihood = -937.40449                       Pseudo R2       =     0.0643
------------------------------------------------------------------------------
       cuse |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
------------+-----------------------------------------------------------------
     Iage_2 |   .3678306   .1753673      2.097   0.036        .024117    .7115443
     Iage_3 |   .8077888   .1597533      5.056   0.000        .494678     1.1209
     Iage_4 |   1.022618   .2039337      5.014   0.000       .6229158    1.422321
    Imore_1 |   -.824092   .1171128     -7.037   0.000      -1.053629   -.5945552
------------+-----------------------------------------------------------------
      _cut1 |   .8698414   .1571298             (Ancillary parameter)
------------------------------------------------------------------------------

       cuse |      Probability              Observed
------------|-----------------------------------
         No |   Pr(       xb+u<_cut1)         0.6845
        Yes |   Pr(_cut1<xb+u)                0.3155
```

No contraceptive use
Yes, contraceptive use

Interpretation of results

The model fitted by STATA is $P(Y=1|X_1,\cdots,X_4)=P(U>\kappa|X_1,\cdots,X_4)=1-\gamma(\mathbf{X})=\frac{\exp(\boldsymbol{\beta}\mathbf{X}-\kappa)}{1+\exp(\boldsymbol{\beta}\mathbf{X}-\kappa)}$,

where $\gamma=P(U\le\kappa)$. The estimates $\beta_1,\ldots,\beta_4$ measure the log-odds ratio of each factor as before. Notice that the estimates of the intercept and the estimate of the cutoff point κ are equal but have opposite signs, that is, $\kappa=-\beta_o$, the intercept estimated by an ordinary logistic regression.

It is very important to notice that the odds ratio does not depend on κ. We see this by comparing two 25-29 years of age, one that desires more children and one that does not. The odds ratio of these two women in terms of using contraceptive methods is

$$\Psi=\frac{P(Y=1|X_1=1,X_4=1)/P(Y=0|X_1=1,X_4=1)}{P(Y=1|X_1=1,X_4=0)/P(Y=0|X_1=1,X_4=0)}=\frac{1-\gamma(X_1=1,X_4=1)/\gamma(X_1=1,X_4=1)}{1-\gamma(X_1=1,X_4=0)/\gamma(X_1=1,X_4=0)}$$

$$=\frac{\frac{\exp(\beta_1+\beta_4-\kappa)}{1+\exp(\beta_1+\beta_4-\kappa)}\Big/\frac{1}{1+\exp(\beta_1+\beta_4-\kappa)}}{\frac{\exp(\beta_1-\kappa)}{1+\exp(\beta_1-\kappa)}\Big/\frac{1}{1+\exp(\beta_1-\kappa)}}=\frac{\exp(\beta_1+\beta_4-\kappa)}{\exp(\beta_1-\kappa)}=\exp(\beta_4)$$

Interpretation of the results (continued)

Finally, consider the following table of the overall use of contraceptive methods in the sample:

```
. tab cuse [freq=N]

Contracepti |
     ve use |
   (Yes/No) |      Freq.     Percent        Cum.
------------+-----------------------------------
         No |       1100       68.45       68.45
        Yes |        507       31.55      100.00
------------+-----------------------------------
      Total |       1607      100.00
```

Notice that the percentage of contraception use is 0.3155 and of no use is 0.6845. These are exactly the estimates of the probabilities listed by the `table` option in the `ologit` command.

Predicted values

The predicted probabilities of contraceptive use are produced as follows:

The probabilities $P(Y=1|\mathbf{x})=1-\gamma(X_1=x_1,\cdots,X_4=x_4)=\frac{\exp(\beta_1 x_1+\cdots+\beta_4 x_4-\kappa)}{1+\exp(\beta_1 x_1+\cdots+\beta_2 x_4-\kappa)}$. For example, for a

woman 25-29 years old that desires more children, we have:

$$\pi=P(Y=1)=1-\gamma(1,0,0,1)=\frac{\exp[-.824092+0.3678306-.8698414]}{1+\exp[-.824092+0.3678306-.8698414\ ]}=0.20980474$$

We see this from the following output:

```
. predict p0 p1
(option p assumed; predicted probabilities)

.  list age educat more cuse  p0 p1 if age==2 & more==1

          age     educat       more       cuse         p0         p1
  9.    25-29       High        Yes        Yes   .7901953   .2098047
 10.    25-29        Low        Yes         No   .7901953   .2098047
 11.    25-29       High        Yes         No   .7901953   .2098047
 16.    25-29        Low        Yes        Yes   .7901953   .2098047
```

where `p1` is the predicted probability of contraceptive use and `p0` of non-use.

Probit regression analysis

```
. xi: probit cuse i.age i.more [freq=N], nolog
i.age                 Iage_1-4      (naturally coded; Iage_1 omitted)
i.more                Imore_0-1     (naturally coded; Imore_0 omitted)

Probit estimates                                  Number of obs   =       1607
                                                  LR chi2(4)      =     127.51
                                                  Prob > chi2     =     0.0000
Log likelihood = -938.09112                       Pseudo R2       =     0.0636

------------------------------------------------------------------------------
       cuse |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
------------+-----------------------------------------------------------------
    Iage_2 |   .2086109   .1003457     2.079   0.038        .0119369     .405285
    Iage_3 |   .4685637   .0928326     5.047   0.000        .2866152    .6505122
    Iage_4 |   .6048679   .1226446     4.932   0.000        .3644889    .8452469
   Imore_1 |  -.4964618   .0714451    -6.949   0.000       -.6364916   -.3564319
     _cons |   -.515345   .0922618    -5.586   0.000       -.6961748   -.3345152
```

We notice that the results are almost identical to the logistic-regression analysis. Recall however that the logit coefficients are not standardized but must be divided $\pi/\sqrt{3}$. Having done this, the estimates are extremely close.

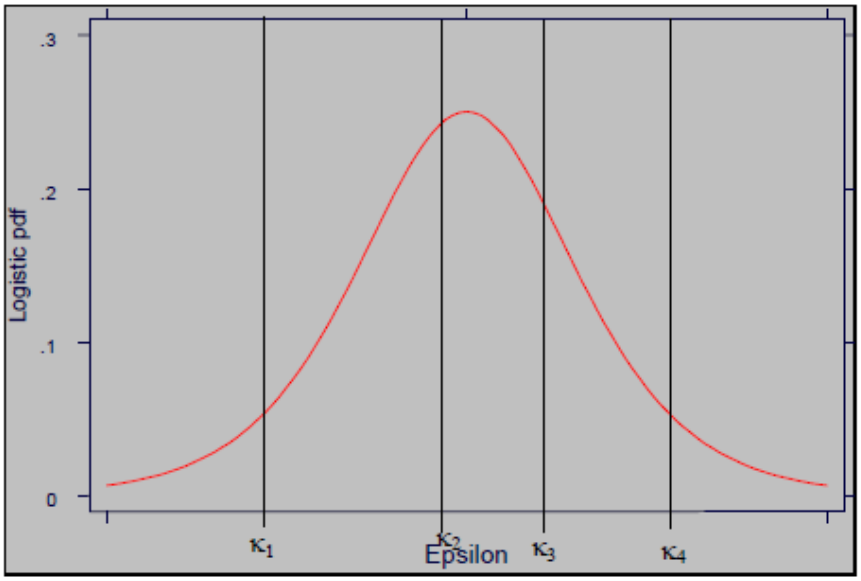
$$\pi / \sqrt{3} \approx 1.81$$

Ordinal regression

If Y is the outcome variable taking I response categories $y_1 < y_2 \cdots < y_{I-1} < y_I$. Then if U is distributed according to a continuous distribution F we define $\gamma_j = P(U \le \kappa_j)$, $j = 1, \cdots, I$ with $\kappa_I = \infty$ so $\gamma_I = P(U \le \kappa_I) = 1$. The probability of the first outcome is $\pi_1 = \gamma_1$ and for each j^{th} outcome it is,

$$\boxed{\pi_j = \gamma_j - \gamma_{j-1},\ j = 1,\ldots,I-1 \text{ and } \pi_I = 1 - \gamma_{I-1}}$$

The situation is shown graphically, for the logistic-distribution situation and I=5:



Treatment of lung cancer (Holtbrugge and Schumacher, App Stat, 1991)

Consider the following example: Lung-cancer patients were randomized to receive two different kinds of chemotherapy (sequential therapy and alternating therapy). The outcome was classified as "progressive disease", "no change", "partial remission" and "complete remission". The goal of this analysis is to compare the two therapies in terms of patient outcome. The data are presented below:

Therapy	Sex	Progressive Disease	No Change	Partial remission	Complete remission
Sequential	Male	28	45	29	26
	Female	4	12	5	2
Alternative	Male	41	44	20	20
	Female	12	7	3	1

We will analyze these data as an ordinal-regression model with four categories (i.e., I=4).

The proportional-odds model

Since $\gamma_j = P(U \le \kappa_j) = P(U - \boldsymbol{\beta}'\mathbf{X} \le \kappa_j - \boldsymbol{\beta}'\mathbf{X}) = P[\varepsilon \le \kappa_j - \boldsymbol{\beta}'\mathbf{X}] = F(\kappa_j - \boldsymbol{\beta}'\mathbf{X})$. Thus,

$$\gamma_j = \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{X} - \kappa_j)}$$

Thus, the log-odds of the response variable being less than or equal to κ_j is $\log\left[\frac{\gamma_j}{1-\gamma_j}\right] = -(\boldsymbol{\beta}'\mathbf{X} - \kappa_j)$ and thus the log-odds ratio for two different values of $\mathbf{x}_1$ and $\mathbf{x}_2$ is $-\boldsymbol{\beta}'(\mathbf{x}_1 - \mathbf{x}_2)$, which is independent of j. This is called the *proportional-odds model* (alternatively: $\log(C_i) = \log(\frac{C_i}{1-C_i}) = \log[\frac{\Pr(Y_i > y_i | X)}{\Pr(Y_i \le y_i | X)}]$).

The proportional-odds model implies that the relationship between each of the covariates X and Y is independent of i, i.e., log odds ratio across response categories identical. It is invariant when codes of the response Y are reverse (i.e., y_1=complete remission, etc) and under collapsibility of the categories (i.e., categories 1 and 2 combined).

Ordinal data analysis

Analysis proceeds with STATA's command `ologit`. The output is as follows:

```
. char sex[omit] 2

. xi: ologit outc i.sex i.therapy, nolog tab
i.sex                 Isex_1-2     (naturally coded; Isex_2 omitted)
i.therapy             Ithera_0-1   (naturally coded; Ithera_0 omitted)

Ordered logit estimates                           Number of obs   =        299
                                                  LR chi2(2)      =      10.91
                                                  Prob > chi2     =     0.0043
Log likelihood = -394.52832                       Pseudo R2       =     0.0136

------------------------------------------------------------------------------
        outc |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
      Isex_1 |   .5413938   .2871816      1.885   0.059      -.0214717    1.104259
    Ithera_1 |   -.580685   .2121478     -2.737   0.006      -.9964871    -.164883
-------------+----------------------------------------------------------------
       _cut1 |  -.7766492   .2880856             (Ancillary parameters)
       _cut2 |   .7906273   .2866223
       _cut3 |    1.84145   .3056123
------------------------------------------------------------------------------

        outc |      Probability            Observed
-------------|----------------------------------------
    Progress |   Pr(      xb+u<_cut1)         0.2843
    No chang |   Pr(_cut1<xb+u<_cut2)         0.3612
    Partial  |   Pr(_cut2<xb+u<_cut3)         0.1906
    Complete |   Pr(_cut3<xb+u)               0.1639
```

85/299 progressive disease
108/299 no change
67/299 partial remission
49/299 complete remission

Interpretation of the coefficients

The estimated coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ are respectively the log-odds ratios of having each of the four or more outcomes versus having any of the previous ones, between males and females and those receiving alternating versus sequential therapy.

THE PROPORTIONAL-ODDS MODEL SPECIFIES THAT THERE IS NO DIFFERENCE IN THE ODDS RATIO REGARDLESS OF THE DICHOTOMIZATION OF THE RESPONSE

For example, two male subjects receiving different therapies have log-odds ratio of experiencing outcome j or higher versus any of the outcomes $1,\ldots,j-1$ equal to

$$\log(\Psi)=\log\left[\frac{P(Y>j|X_1=1,X_2=1)/P(Y\le j|X_1=1,X_2=1)}{P(Y>j|X_1=0,X_2=1)/P(Y\le j|X_1=0,X_2=1)}\right]=\log\left[\frac{\frac{\exp(\beta_1+\beta_2-\kappa_j)}{1+\exp(\beta_1+\beta_2-\kappa_j)}\Big/\frac{1}{1+\exp(\beta_1+\beta_2-\kappa_j)}}{\frac{\exp(\beta_2-\kappa_j)}{1+\exp(\beta_2-\kappa_j)}\Big/\frac{1}{1+\exp(\beta_2-\kappa_j)}}\right]=\beta_1$$

This does not depend on which outcome we refer to. A similar argument holds for β_2.

Interpretation of the coefficients (continued)

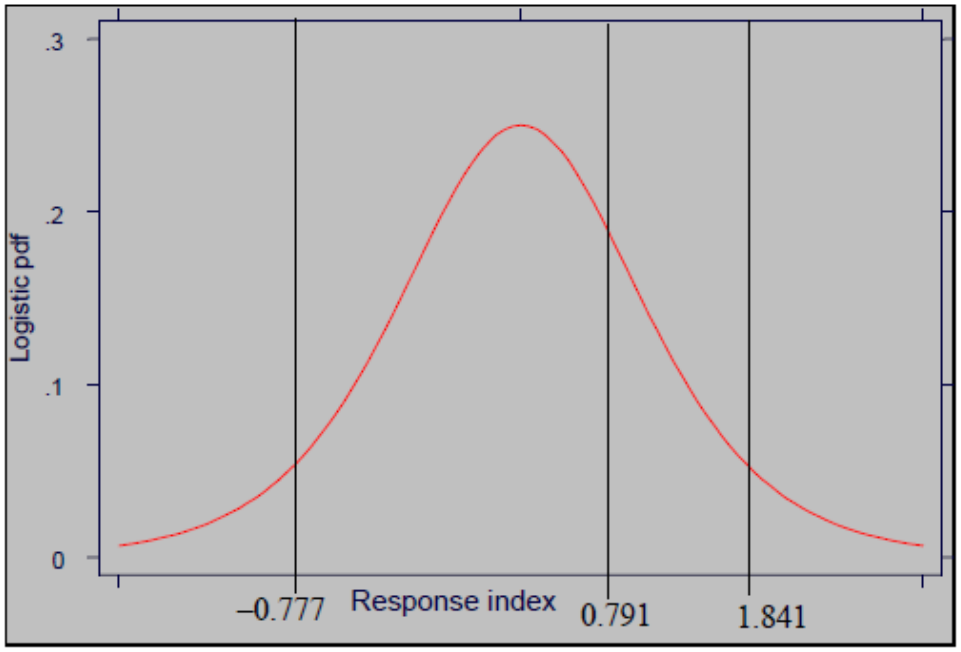
- `sex`: The estimate $\hat{\beta}_1 = 0.541$, which is translated to an estimated odds ratio of $e^{0.541} \approx 1.72$. Males receiving the same therapy are 72% more likely than females to experience full remission versus at best partial remission (or partial or full remission versus at best no change and so on). The gender effect is significant at the 10% but not the 5% level.
- `therapy`: In this case, we have $\hat{\beta}_2 = -0.581$. This means that the odds ratio of two subjects of the same gender is $e^{-0.581} \approx 0.56$. Subjects receiving alternating therapy (X_2=1) are almost half as likely as subjects receiving sequential therapy to experience full remission versus at most a partial remission. Equivalently (by the assumptions of the proportional-odds model), the odds ratio are the same for these subjects in terms of the experience of partial or complete remission versus no change or progressive disease.

Interpretation of the ordinal model

The underlying logistic distribution is:

$$f(\varepsilon) = F(\varepsilon) * [1 - F(\varepsilon)] = \frac{1}{1+e^{-\varepsilon}} * (1 - \frac{1}{1+e^{-\varepsilon}}) = \frac{e^{-\varepsilon}}{\left(1+e^{-\varepsilon}\right)^2}$$

according to an index ε that is related to patient outcome (e.g., percent change of tumor size). For example, the logistic pdf for male subjects receiving alternating therapy is shown below:



Interpretation of the model (continued)

The probabilities $\gamma_j(X_1,X_2)=\frac{1}{1+\exp(\beta_1 X_1+\beta_2 X_2-\kappa_j)}$, $j=1,2,3$. For the previous example of male subjects receiving alternating therapy (i.e., $X_1=1, X_2=1$) we have:

$$\gamma_1(1,1)=\frac{1}{1+\exp\left[0.5413938+(-0.580685)-(-0.7766492)\right]}=0.3235821$$

$$\gamma_2(1,1)=\frac{1}{1+\exp\left[0.5413938+(-0.580685)-0.7906273\right]}=0.6963377$$

$$\gamma_3(1,1)=\frac{1}{1+\exp\left[0.5413938+(-0.580685)-1.84145\right]}=0.86769624$$

Thus, the probabilities in this group $\pi_1(1,1)=\gamma_1(1,1)=0.323$, $\pi_2(1,1)=\gamma_2(1,1)-\gamma_1(1,1)=0.373$, $\pi_3(1,1)=\gamma_3(1,1)-\gamma_2(1,1)=0.171$ and $\pi_4(1,1)=1-\gamma_3(1,1)=0.132$.

Male, alternating therapy: 125
Progressive disease: 41
41/125=0.323

The predicted probabilities are produced by the command `predict`

Probit analysis

The previous example can be easily analyzed via probit analysis. The STATA output follows:

```
. xi: oprobit outc i.sex i.therapy, nolog
i.sex                 Isex_1-2     (naturally coded; Isex_2 omitted)
i.therapy             Ithera_0-1   (naturally coded; Ithera_0 omitted)

Ordered probit estimates                          Number of obs   =        299
                                                  LR chi2(2)      =      10.79
                                                  Prob > chi2     =     0.0045
Log likelihood =  -394.5871                       Pseudo R2       =     0.0135

------------------------------------------------------------------------------
        outc |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
      Isex_1 |   .3401406    .174902      1.945   0.052       -.002661    .6829422
    Ithera_1 |  -.3344764    .125435     -2.667   0.008      -.5803245   -.0886282
-------------+----------------------------------------------------------------
       _cut1 |   -.459358    .176613            (Ancillary parameters)
       _cut2 |   .5050695   .1760197
       _cut3 |   1.122025   .1836877
```

The coefficients of the probit regression are similar in sign and size to those of the logistic regression, if the latter are divided by $\pi/\sqrt{3}$.

A test for the validity of the proportional-odds assumption

```
. xi: omodel logit outc  i.therapy i.sex
i.therapy             Ithera_0-1    (naturally coded; Ithera_0 omitted)
i.sex                 Isex_1-2      (naturally coded; Isex_2 omitted)

Iteration 0:   log likelihood = -399.98398
Iteration 1:   log likelihood = -394.53988
Iteration 2:   log likelihood = -394.52832
Iteration 3:   log likelihood = -394.52832

Ordered logit estimates                           Number of obs   =        299
                                                  LR chi2(2)      =      10.91
                                                  Prob > chi2     =     0.0043
Log likelihood = -394.52832                       Pseudo R2       =     0.0136

------------------------------------------------------------------------------
        outc |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
-------------+----------------------------------------------------------------
    Ithera_1 |   -.580685   .2121478     -2.737   0.006      -.9964871    -.164883
      Isex_1 |   .5413938   .2871816      1.885   0.059      -.0214717    1.104259
-------------+----------------------------------------------------------------
       _cut1 |  -.7766492   .2880856            (Ancillary parameters)
       _cut2 |   .7906273   .2866223
       _cut3 |    1.84145   .3056123
------------------------------------------------------------------------------

Approximate likelihood-ratio test of proportionality of odds
across response categories:
         chi2(4) =      3.26
       Prob > chi2 =    0.5147
```

DF: (3-1)*2=4

Checking the proportional-odds assumption

The approximate likelihood-ratio test reported by the `omodel` command tests the hypothesis that the coefficients in models arise from arbitrary dichotomization of the outcome variable (e.g., complete remission versus up to partial remission, partial and complete remission versus no change or progressive disease and no change or better versus progressive disease).

The test has four degrees of freedom equal to the number of pair-wise comparisons between the dichotomizations of the outcome variable minus one times the number of the coefficients of the model. The approximate value is 3.26, which, compared to a chi-square with 4 degrees of freedom results in a non-significant p-value 0.5147.

There is no evidence that the proportional-odds assumption is violated in this model.

Constrained multinomial models*

The **proportional-odds** or cumulative-logit model is a multinomial model where some constrains have been imposed on the coefficients. In this case, and all the other models we consider, we allow for different intercepts, but force the slope coefficients to be equal for the various dichotomizations of the outcome variable. The proportional-odds model is given as follows:

$$\log\left(\frac{\gamma_j}{1-\gamma_j}\right)=\alpha_j-\mathbf{x}'\boldsymbol{\beta},\ \text{where}\ \gamma_j=P(U\le\kappa_j),\ j=1,\ldots,I-1$$

The **adjacent-categories** model specifies that the effect of the explanatory variables is constant across adjacent categories. Different intercepts are allowed here as well. The model is given by

$$\log\left(\frac{p_j}{p_{j-1}}\right)=\alpha_j+\mathbf{x}'\boldsymbol{\beta},\ j=2,\ldots,I$$

A specialized model is the **continuation-ratio** model. It specifies that, to reach each subsequent category, one has to go through the previous ones. This model is applied mostly to education research. It is given from the following equation:

$$\log\left(\frac{\Pr(Y=j|\mathbf{x})}{\Pr(Y<j|\mathbf{x})}\right)=\alpha_j+\mathbf{x}'\boldsymbol{\beta}_j,\ j=2,\ldots,I$$