

mixor: An R Package for Longitudinal and Clustered Ordinal Response Modeling

Kellie J. Archer

Virginia Commonwealth University

Donald Hedeker

University of Illinois Chicago

Rachel Nordgren

University of Illinois Chicago

Robert D. Gibbons

University of Chicago

Abstract

This paper describes an R package, **mixor**, that provides a function for fitting an ordinal response model when observations are either clustered or collected longitudinally. The function, **mixord** uses either adaptive (default) or non-adaptive Gauss-Hermite quadrature to numerically integrate over the distribution of random effects and Fisher scoring to obtain the likelihood solution, as described by [Hedeker and Gibbons \(1996\)](#). Generic methods for **mixor** including **summary**, **print**, **vcov**, **plot**, and **coef** can be applied to a **mixor** object. Examples of modeling longitudinal data, clustered data, grouped survival times, and weighted data are provided.

Keywords: ordinal response, longitudinal data, clustered data, random effects, R.

1. Introduction

Health status and outcomes, such as quality of life, functional status, and patient satisfaction, are frequently measured on an ordinal scale. In addition, most histopathological variables are ordinal, including scoring methods for liver biopsy specimens from patients with chronic hepatitis, such as the Knodell hepatic activity index, the Ishak score, and the METAVIR score. Often, outcomes collected are either clustered or collected longitudinally. For example, stage of breast cancer is derived using degree of tubule formation, nuclear pleomorphism, and mitotic count ([Ivshina, George, Senko, Mow, Putti, Smeds, Lindahl, Pawitan, Hall, Nordgren, Wong, Liu, Bergh, Kuznestsov, and Miller 2006](#)). Similarly, stage of hypopharyngeal cancer is derived using three ordinally scaled measures: tumor, node, and metastasis scores ([Cromer, Carles, Millon, Ganguli, Chalmel, Lemaire, Young, Demb  l  , Thibault, Muller, Poch, Abesassis, and Wasylyk 2004](#)). Multiple ordinally scaled variables give rise to clustered ordinal response data. In addition, some studies collect an ordinal response on each subject at multiple time-points. Currently only the **ordinal** package in the R programming environment provides a function for fitting cumulative link ordinal response random effects models ([Christensen 2013](#)). Although **ordinal** can fit both random intercept and random coefficient models, it currently does not implement quadrature methods for vector-valued random effects (including random coefficients models) nor can it handle nested random effects structures.

MIXOR, written in Fortran, is a stand-alone package for fitting cumulative link ordinal (and dichotomous) response models ([Hedeker and Gibbons 1996](#)). **MIXOR** supports probit, logit, log-log, and complementary log-log link functions and can fit models that include multiple

random effects. The stand-alone program requires the user to either specify all model instructions in an ASCII text file following a precisely defined format (Hedeker and Gibbons 1996) or specify the model options using a GUI interface which subsequently creates the batch-defined ASCII file. After submitting the job using the batch or interactive model, the empirical Bayes estimates, parameter estimates, and asymptotic variance-covariance matrix of the parameter estimates for the fitted model are written to different files (MIXOR.RES, MIXOR.EST, and MIXOR.VAR, respectively). To enhance the usability of the stand-alone program, we developed an R package **mixor** that interfaces to a Fortran dynamic-link library for fitting cumulative link ordinal response mixed effects models. Modeling results are returned within R providing a convenient environment for additional model fitting, testing, and plotting.

2. Ordinal Response Model

Herein we briefly describe the cumulative logit model for the traditional setting where data are neither clustered or collected longitudinally, to demonstrate the natural connections to the dichotomous setting. Let Y_i represent the ordinal response for observation i that can take on one of K ordinal levels. Denote the $N \times P$ covariate matrix as \mathbf{x} so that \mathbf{x}_i represents a $P \times 1$ vector for observation i and \mathbf{x}_p represents the $N \times 1$ vector for covariate p . For observations $i = 1, \dots, N$, the response Y_i can be reformatted as a response matrix consisting of N rows and K columns where

$$y_{ik} = \begin{cases} 1 & \text{if observation } i \text{ is class } k \\ 0 & \text{otherwise.} \end{cases}.$$

Therefore \mathbf{y}_k is an $N \times 1$ vector representing class k membership. Letting $\pi_k(\mathbf{x}_i)$ represent the probability that observation i with covariates \mathbf{x}_i belongs to class k , the likelihood for an ordinal response model with K ordinal levels can be expressed as

$$L = \prod_{i=1}^N \prod_{k=1}^K \pi_k(\mathbf{x}_i)^{y_{ik}}. \quad (1)$$

The cumulative logit model models $K - 1$ logits of the form

$$P(Y_i \leq k) = \frac{\exp(\alpha_k - \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_k - \mathbf{x}_i^T \boldsymbol{\beta})} \quad (2)$$

where α_k denotes the class-specific intercept or threshold and $\boldsymbol{\beta}$ is a $P \times 1$ vector of coefficients associated with explanatory variables \mathbf{x}_i (Agresti 2010). Equation 2 is formulated to subtract the $\mathbf{x}_i^T \boldsymbol{\beta}$ term from the thresholds as described in the seminal paper by McCullagh (1980), which provides an intuitive interpretation of the relationship between $\mathbf{x}_i^T \boldsymbol{\beta}$ and the probability of response; larger values of $\mathbf{x}_i^T \boldsymbol{\beta}$ correspond to higher probability of the response belonging to an ordinal class at the higher end of the scale. Other software packages fit cumulative link models using a plus sign in equation 2 so the **mixor** package flexibly permits the user to change the model parameterization. Note that the class-specific probabilities can be calculated by subtracting successive cumulative logits,

$$P(Y_i = k) = P(Y_i \leq k) - P(Y_i \leq k - 1).$$

Therefore for any class k , providing we let $-\infty = \alpha_0 < \alpha_1 < \dots < \alpha_{K-1} < \alpha_K = \infty$, we can express the class-specific probabilities by

$$\pi_k(\mathbf{x}_i) = \frac{\exp(\alpha_k - \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_k - \mathbf{x}_i^T \boldsymbol{\beta})} - \frac{\exp(\alpha_{k-1} - \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_{k-1} - \mathbf{x}_i^T \boldsymbol{\beta})}.$$

The function that links the probability to the linear predictor in equation 2 is the logit link,

$$\log \left(\frac{P(Y_i \leq k)}{1 - P(Y_i \leq k)} \right) = \alpha_k - \mathbf{x}_i^T \boldsymbol{\beta}.$$

Other link functions that can be used to link the cumulative probabilities to the linear predictor include the probit link,

$$\Phi^{-1}(P(Y_i \leq k))$$

where Φ^{-1} is the inverse of the cumulative standard normal distribution function; the log-log link,

$$\log(-\log(P(Y_i \leq k)))$$

and the complementary log-log link

$$\log(-\log(1 - P(Y_i = k))).$$

3. Longitudinal/Clustered Ordinal Response Models

Consider now the scenario where subjects $i = 1, \dots, N$ (level-2 units) are each observed $j = 1, \dots, n_i$ times (level-1 units) where the response at each j belongs to one of $k = 1, \dots, K$ ordered categories. Here j could index either clustered or longitudinal observations for unit i . We let p_{ijk} represent the probability that a subject i at j falls into class k . Therefore, the cumulative probability at j is $P(Y_{ij} \leq k) = \sum_{l=1}^k p_{ijl}$. The mixed-effects logistic regression model for the $K - 1$ cumulative logits is then given by

$$\log \left(\frac{P(Y_{ij} \leq k)}{1 - P(Y_{ij} \leq k)} \right) = \alpha_k - (\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{T} \boldsymbol{\theta}_i) \quad (3)$$

where the thresholds given by $(\alpha_1, \alpha_2, \dots, \alpha_{K-1})$ are strictly increasing, \mathbf{x}_{ij} is the covariate vector, and $\boldsymbol{\beta}$ is the vector of regression parameters. The unstandardized random effects $\boldsymbol{\nu}_i \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}_i})$ which are expressed by the standardized vector of the r random effects $\boldsymbol{\theta}_i$ and where \mathbf{T} is the Cholesky factorization of $\boldsymbol{\Sigma}_{\boldsymbol{\nu}}$ and \mathbf{z}_{ij} is the design vector for the r random effects. Letting the response vector be denoted by $\mathbf{y}_i = (Y_{ij1}, Y_{ij2}, \dots, Y_{ijK})$ where $y_{ijk} = 1$ if the response for subject i at j is in category k and 0 otherwise, so that $n_i = \sum_j \sum_{k=1}^K y_{ijk}$. The likelihood can be expressed as

$$l(\mathbf{y}_i | \boldsymbol{\theta}_i) = \prod_{j=1}^{n_i} \prod_{k=1}^K (P(y_{ij} \leq k) - P(y_{ij} \leq k-1))^{y_{ijk}} \quad (4)$$

Details on the maximum marginal likelihood estimation procedure which uses multidimensional quadrature to integrate over the distribution of random effects and Fisher's scoring for

obtaining the solution to the likelihood have been previously described (Hedeker and Gibbons 2006, 1996).

4. Implementation

The **mixor** package was written in the R programming environment (R Core Team 2013) and requires the **survival** package (Therneau 2013) when fitting discrete survival time models. The **mixord** function allows the user to specify a model formula, identify the level-2 identifier using the **id** parameter, and additionally specify whether any variables have a random slope (**which.random.slope**). The function also supports fitting non-proportional odds models by specifying the variables for which proportional odds is not assumed (KG) and can fit scaling models (KS). The function parses the user-specified parameters which are subsequently passed to a **MIXOR** Fortran dynamic linked library (Hedeker and Gibbons 1996) and results are returned to the fitted object in the form of a list. The default link is **link = "probit"**. Other allowable links include **logit**, **loglog**, and **cloglog**. The function uses adaptive quadrature with 11 quadrature points by default which can be changed by specifying **adaptive.quadrature = FALSE** to perform non-adaptive quadrature; the number of quadrature points to be used for each dimension of the integration can be changed by specifying a different integer value to **NQ1** in the function call. By default the quadrature distribution is normal (**quadrature.dist = "Normal"**) but the function supports usage of a uniform distribution (**quadrature.dist = "Uniform"**). Also, by default the random effects are assumed to be correlated; if independent random effects are assumed, then **indep.re = FALSE** should be specified. Generic methods for returning coefficient estimates, printing summaries, extracting variance-covariance estimates are available using **print**, **coef**, **summary**, **plot** and **vcov**.

5. Examples

The **mixor** package includes example datasets that are useful for demonstrating modeling a longitudinal ordinal response (**schizophrenia**), grouped survival data (**SmokeOnset**), frequency weighted data (**norcag**), and clustered ordinal response (**concen**).

5.1. Longitudinal Data

These data are from the National Institute of Mental Health Schizophrenia Collaborative Study and are stored in the **data.frame** **schizophrenia** (Gibbons and Hedeker 1994). Patients were randomized to receive one of four medications, either placebo or one of three different anti-psychotic drugs. The protocol indicated subjects were to then be evaluated at weeks 0, 1, 3, 6 to assess severity of illness; additionally some measurements were made at weeks 2, 4, and 5. The primary outcome is item 79 on the Inpatient Multidimensional Psychiatric Scale which indicates severity of illness. We will analyze **imps79o** which is an ordinal scaled version of the original variable **imps79** which has the following interpretation: 1 = normal, not ill at all; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill; and 7 = among the most extremely ill. The four category ordinal version (**imps79o**) grouped the responses as follows:

imps79	imps79o
1 & 2	1 (not ill or borderline)
3 & 4	2 (mildly or moderately)
5	3 (markedly)
6 & 7	4 (severely or most extremely ill)

Predictor variables of interest are `TxDrug` a dummy coded variable indicating treatment with drug (1) or placebo (0), the square root of the Week variable (`SqrtWeek`), and their interaction (`TxSWeek`).

Random Intercept Model

A random intercepts model can be fit as follows:

```
> library("mixor")
> data(schizophrenia)
> SCHIZ01.fit<-mixord(imps79o ~ TxDrug + SqrtWeek + TxSWeek, data=schizophrenia,
+ id=id, link="logit")
```

Note that the user supplies the model formula in the traditional way, specifies the data frame using `data`, the level-2 variable using `id`, and the link function using `link`. Methods such as `summary` and `print` can be applied to `mixor` model objects.

```
> summary(SCHIZ01.fit)
```

Call:

```
mixord(formula = imps79o ~ TxDrug + SqrtWeek + TxSWeek, data = schizophrenia,
      id = id, link = "logit")
```

```
Deviance =          3402.758
Log-likelihood =    -1701.379
RIDGEMAX =           0.2
AIC =              -1708.379
SBC =              -1722.659
```

	Estimate	SE	Z	P.value
(Intercept)	5.85924463	0.3428850	17.0880770	0.000000e+00
TxDrug	-0.05843032	0.3108636	-0.1879613	8.509070e-01
SqrtWeek	-0.76577287	0.1197460	-6.3949755	1.605731e-10
TxSWeek	-1.20615055	0.1331370	-9.0594681	0.000000e+00
Random.(Intercept)	3.77377698	0.4954283	7.6172012	2.597922e-14
Threshold2	3.03282074	0.1323782	22.9102787	0.000000e+00
Threshold3	5.15076669	0.1792538	28.7344869	0.000000e+00

While `print` simply prints a brief summary of the coefficients to the console, `coef` extracts the estimated parameters and returns them as a vector and `vcov` extracts and returns the asymptotic variance-covariance matrix of the parameter estimates (the order of the elements in this matrix follows the order listed by `coef`).

```
> print(SCHIZ01.fit)
```

Call:

```
mixord(formula = imps79o ~ TxDrug + SqrtWeek + TxSWeek, data = schizophrenia,
       id = id, link = "logit")
```

Coefficients:

(Intercept)	TxDrug	SqrtWeek
5.85924	-0.05843	-0.76577
TxSWeek	Random.(Intercept)	Threshold2
-1.20615	3.77378	3.03282
Threshold3		
5.15077		

```
> coef(SCHIZ01.fit)
```

(Intercept)	TxDrug	SqrtWeek
5.85924463	-0.05843032	-0.76577287
TxSWeek	Random.(Intercept)	Threshold2
-1.20615055	3.77377698	3.03282074
Threshold3		
5.15076669		

```
> vcov(SCHIZ01.fit)
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	0.11757010	-0.0748269661	-0.025627080	0.011013480	0.054763860
[2,]	-0.07482697	0.0966361929	0.019424570	-0.023174175	-0.014830605
[3,]	-0.02562708	0.0194245699	0.014339110	-0.012056227	-0.008162231
[4,]	0.01101348	-0.0231741746	-0.012056227	0.017725465	-0.011636936
[5,]	0.05476386	-0.0148306049	-0.008162231	-0.011636936	0.245449190
[6,]	0.02417834	0.0003265383	-0.002781150	-0.004547136	0.023080367
[7,]	0.03843122	-0.0029222399	-0.005920412	-0.006323587	0.047288216
	[,6]	[,7]			
[1,]	0.0241783364	0.038431220			
[2,]	0.0003265383	-0.002922240			
[3,]	-0.0027811500	-0.005920412			
[4,]	-0.0045471359	-0.006323587			
[5,]	0.0230803673	0.047288216			
[6,]	0.0175239795	0.019870754			
[7,]	0.0198707540	0.032131934			

By default the `mixord` function returns the `Intercept` and the `Threshold2` and `Threshold3` values which represent the ordinal departures from the intercept. If the $K - 1$ cutpoints are desired, they can be obtained using the `Contrasts` function.

```
> cm<-matrix(c(1,0,0,0,0, 0,0,
+             1,0,0,0,0,-1,0,
+             1,0,0,0,0, 0,-1),ncol=3)
> Contrasts(SCHIZ01.fit, contrast.matrix=cm)
```

```
$contrast.matrix
```

	1	2	3
(Intercept)	1	1	1
TxDrug	0	0	0
SqrtWeek	0	0	0
TxSWeek	0	0	0
Random.(Intercept)	0	0	0
Threshold2	0	-1	0
Threshold3	0	0	-1

```
$Contrasts
```

	Estimate	SE	Z	P.value
1	5.8592446	0.3428850	17.088077	0.000000000
2	2.8264239	0.2945121	9.596969	0.000000000
3	0.7084779	0.2698881	2.625080	0.008662856

The `plot` function produces a histogram and normal quantile-quantile plot of the empirical Bayes means for each random term.

Random Intercept and Trend

It may be of interest to account for subject heterogeneity through both the intercept and by time. A model that includes a random intercept and slope can be fit by additionally specifying the index corresponding to the variable(s) on the right-hand side (RHS) of the equation that should have a random coefficient(s) using the `which.random.slope` parameter. Note that multiple variables can be specified by the `which.random.slope` parameter. For example, `which.random.slope=c(1,3)` indicates that the first and third variables listed on the RHS of the model formula should be random coefficients. In this example, `SqrtWeek` is the second variable listed in the RHS of the model formula and allowing it to have a random coefficient is specified by `which.random.slope=2`.

```
> SCHIZ02.fit<-mixord(imps79o ~ TxDrug + SqrtWeek + TxSWeek, data=schizophrenia,
+ id=id, which.random.slope=2, link="logit")
> summary(SCHIZ02.fit)
```

```
Call:
```

```
mixord(formula = imps79o ~ TxDrug + SqrtWeek + TxSWeek, data = schizophrenia,
       id = id, which.random.slope = 2, link = "logit")
```

```
Deviance =          3325.486
Log-likelihood =    -1662.743
RIDGEMAX =           0
```

```
> plot(SCHIZ01.fit)
```

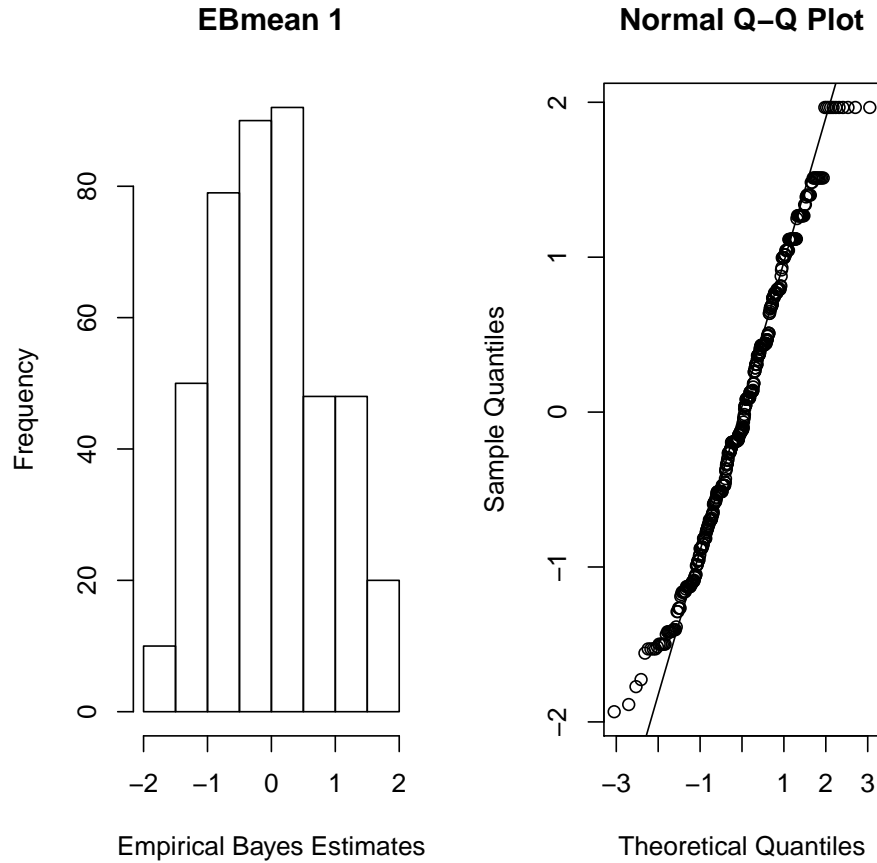


Figure 1: Histogram and normal quantile-quantile plot of the empirical Bayes means from a random intercepts cumulative logit model fit using `mixord` to the `schizophrenia` data.

AIC = -1671.743

SBC = -1690.103

	Estimate	SE	Z	P.value
(Intercept)	7.31883117	0.4807785	15.2228762	0.000000e+00
SqrtWeek	-0.88226097	0.2345675	-3.7612237	1.690841e-04
TxDrug	0.05791686	0.3991021	0.1451179	8.846178e-01
TxSWeek	-1.69486115	0.2681311	-6.3210165	2.598481e-10
(Intercept) (Intercept)	6.99764641	1.3692735	5.1104812	3.213392e-07
(Intercept) SqrtWeek	-1.50851397	0.5360233	-2.8142692	4.888828e-03
SqrtWeek SqrtWeek	2.00891623	0.4535871	4.4289538	9.469132e-06
Threshold2	3.90126044	0.2132568	18.2937239	0.000000e+00
Threshold3	6.50717220	0.2899909	22.4392320	0.000000e+00

By default, the model is fit assuming the random effect terms are correlated. The following components can be extracted from the fitted object as needed.

```
> names(SCHIZ02.fit)

[1] "call"           "Deviance"        "Quadrature.points"
[4] "Model"          "varcov"           "EBmean"
[7] "EBvar"          "RIDGEMAX"        "RLOGL"
[10] "SE"             "AIC"              "SBC"
[13] "AICD"           "SBCD"             "MU"
[16] "ALPHA"          "SIGMA"            "GAM"
[19] "TAU"
```

For example, we can extract the log-likelihoods from the random intercept and the random coefficient model.

```
> SCHIZ02.fit$RLOGL
```

```
[1] -1662.743
```

```
> SCHIZ01.fit$RLOGL
```

```
[1] -1701.379
```

To fit a model assuming independent random effects, the `indep.re` parameter should be `TRUE`.

```
> SCHIZ03.fit<-mixord(imps79o ~ TxDrug + SqrtWeek + TxSWeek, data=schizophrenia,
+   id=id, which.random.slope=2, indep.re=TRUE, link="logit", NQ1=11)
> summary(SCHIZ03.fit)
```

Call:

```
mixord(formula = imps79o ~ TxDrug + SqrtWeek + TxSWeek, data = schizophrenia,
      id = id, which.random.slope = 2, NQ1 = 11, link = "logit",
      indep.re = TRUE)
```

```
Deviance =      3338.65
Log-likelihood = -1669.325
RIDGEMAX =       0
AIC =          -1677.325
SBC =          -1693.645
```

	Estimate	SE	Z	P.value
(Intercept)	6.79432357	0.3999997	16.985820	0.000000e+00
SqrtWeek	-0.69871973	0.1952674	-3.578272	3.458736e-04
TxDrug	0.08663974	0.3150513	0.275002	7.833147e-01
TxSWeek	-1.66333486	0.2243166	-7.415122	1.214584e-13
Random.(Intercept)	4.10323238	0.7366603	5.570047	2.546705e-08
Random.SqrtWeek	1.24592463	0.2748986	4.532306	5.834328e-06
Threshold2	3.77716454	0.2015708	18.738647	0.000000e+00
Threshold3	6.19845237	0.2630027	23.568021	0.000000e+00

5.2. Grouped Survival Data

The `SmokeOnset` `data.frame` are data from the Television School and Family Smoking Prevention and Cessation Project, a study designed to increase knowledge of the effects of tobacco use in school-age children (Flay, Miller, Hedeker, Siddiqui, Brannon, Johnson, Hansen, Sussman, and Dent 1995). In this study students are nested within class and classes are nested within schools, so either `class` or `school` can be used as the level-2 variable. The primary outcome is time to smoking (`smkonset`) which was recorded post-intervention and at 1 and 2 year follow-up. Because not all students began smoking by the end of the study, the variable `sensor` indicates whether the student smoked (1) or did not smoke (0). Therefore this dataset represents grouped survival data which can be modeled using the `mixord` function using `link="cloglog"` to yield a proportional hazards survival model.

The first example represents `class` as the level-2 variable. The `IADD` parameter indicates whether the $\mathbf{x}^T\boldsymbol{\beta}$ term is added (`IADD=1`) or subtracted (`IADD=0`, default) from the thresholds in equation 2. In survival models, it is customary that positive regression coefficients represent increased hazard so `IADD=1` is specified here.

```
> data(SmokeOnset)
> require(survival)
> Surv.mixord<-mixord(Surv(smkonset,censor)~SexMale+cc+tv, data=SmokeOnset,
+ id=class, link="cloglog", NQ1=20, IADD=1)
> summary(Surv.mixord)
```

Call:

```
mixord(formula = Surv(smkonset, censor) ~ SexMale + cc + tv,
      data = SmokeOnset, id = class, NQ1 = 20, link = "cloglog",
      IADD = 1)
```

```
Deviance =          3185.551
Log-likelihood =    -1592.776
RIDGEMAX =           0
AIC =              -1599.776
SBC =              -1609.918
```

	Estimate	SE	Z	P.value
(Intercept)	-1.66845209	0.10185552	-16.3805759	0.0000000
SexMale	0.06092290	0.08377605	0.7272115	0.4670964
cc	0.05242881	0.09053740	0.5790846	0.5625321
tv	0.01272458	0.09220132	0.1380086	0.8902336
Random.(Intercept)	0.03594543	0.03559023	1.0099802	0.3125048
Threshold3	0.71716030	0.04671548	15.3516619	0.0000000
Threshold4	1.23410592	0.05492814	22.4676441	0.0000000

Alternatively, we can perform a students in schools analysis.

```
> School.mixord<-mixord(Surv(smkonset,censor)~SexMale+cc+tv, data=SmokeOnset,
+ id=school, link="cloglog", NQ1=20, IADD=1)
> summary(School.mixord)
```

Call:

```
mixord(formula = Surv(smkonset, censor) ~ SexMale + cc + tv,
      data = SmokeOnset, id = school, NQ1 = 20, link = "cloglog",
      IADD = 1)
```

```
Deviance =          3187.388
Log-likelihood =    -1593.694
RIDGEMAX =           0.6
AIC =              -1600.694
SBC =              -1605.357
```

	Estimate	SE	Z	P.value
(Intercept)	-1.656099102	0.10731853	-15.4316229	0.0000000
SexMale	0.057262507	0.12365250	0.4630922	0.6432983
cc	0.044667017	0.10417986	0.4287491	0.6681059
tv	0.021328858	0.09372506	0.2275684	0.8199818
Random.(Intercept)	0.002654578	0.01655541	0.1603450	0.8726093
Threshold3	0.713333149	0.04866210	14.6589071	0.0000000
Threshold4	1.225077542	0.05230455	23.4220075	0.0000000

The last example demonstrates a students in classrooms analysis with varying sex effect across time intervals. Further, this model does not assume proportional hazards for the **SexMale** variable. Recall that when assuming proportional hazards, there is no k class subscript on the parameter estimates but rather the estimated coefficient for a given variable reflects the explanatory relationship between that covariate and the cumulative link. Partial proportional odds models are less restrictive and are fit by including $K - 1$ parameter estimates for those variables for which the proportional odds (PO) is not assumed ([Peterson and Harrell 1990](#)). This method is applicable for fitting partial proportional hazards models. In `mixord`, `KG=N` (where `N` is an integer) indicates not to assume proportional hazards (or odds for PO models) for the first `N` variables on the RHS of the equation. When using `KG`, the order of the variables on the RHS is important because the integer value passed to `KG` represents the number of variables, starting with the first, for non-proportional hazards (or odds for PO models) estimation. In this example we specify `KG=1` as we do not want to assume proportional hazards for **SexMale**.

```
> students.mixord<-mixord(Surv(smkonset,censor)~SexMale+cc+tv, data=SmokeOnset,
+   id=class, link="cloglog", KG=1, NQ1=20, IADD=1)
> students.mixord
```

Call:

```
mixord(formula = Surv(smkonset, censor) ~ SexMale + cc + tv,
      data = SmokeOnset, id = class, NQ1 = 20, link = "cloglog",
      KG = 1, IADD = 1)
```

Coefficients:

(Intercept)	SexMale	cc
-1.79603	0.30837	0.05421

	tv	Random.(Intercept)	Threshold3
	0.01230	0.03516	0.83527
Threshold4	Threshold3SexMale	Threshold4SexMale	
1.39718	-0.22916	-0.32141	

Again, if the estimates of the ordinal cutpoints are desired, the `Contrasts` function can be applied after structuring a suitable contrast matrix.

```
> cm<-matrix(c(1,1,0,0,
+             0,0,1,1,
+             0,0,0,0,
+             0,0,0,0,
+             0,0,0,0,
+             1,0,0,0,
+             0,1,0,0,
+             0,0,1,0,
+             0,0,0,1),byrow=TRUE,ncol=4)
> Contrasts(students.mixord, contrast.matrix=cm)
```

```
$contrast.matrix
      1 2 3 4
(Intercept)  1 1 0 0
SexMale      0 0 1 1
cc           0 0 0 0
tv           0 0 0 0
Random.(Intercept) 0 0 0 0
Threshold3    1 0 0 0
Threshold4    0 1 0 0
Threshold3SexMale 0 0 1 0
Threshold4SexMale 0 0 0 1
```

```
$Contrasts
      Estimate      SE      Z      P.value
1 -0.96076040 0.08991737 -10.6849254 0.000000e+00
2 -0.39884946 0.08728111 -4.5697113 4.883966e-06
3  0.07920575 0.09744029  0.8128645 4.162958e-01
4 -0.01304414 0.09304298 -0.1401947 8.885061e-01
```

5.3. Frequency Weighted Data

The `mixord` function also accomodates weighted data where the data are stored as the number of level-2 observations observed (frequency weight) for a unique response pattern and covariate vector. In this example, the `norcag` data pertain to attitudes towards sex as measured in the 1989 General Social Survey ([Agresti and Lang 1993](#)). Each subject provided ordinal responses on three items concerning their opinion on early teens (age 14-16) having sex before marriage (Item1), a man and a woman having sex before marriage (Item2), and a married

person having sex with someone other than their spouse (Item3). However, the data are provided as frequencies (`freq`) by unique response pattern (ID) where the differences in item responses were stored as `Item2vs1` (attitude towards premarital vs teenage sex) and `Item3vs1` (attitude towards extramarital vs teenage sex). To fit a random intercepts model assuming proportional odds for differences in item responses we specify our model as before except we need to specify `weights=freq`.

```
> data("norcag")
> Fitted.norcag<-mixord(SexItems~Item2vs1+Item3vs1, data=norcag, id=ID,
+   weights=freq, link="logit", NQ1=20)
> summary(Fitted.norcag)
```

Call:

```
mixord(formula = SexItems ~ Item2vs1 + Item3vs1, data = norcag,
      id = ID, weights = freq, NQ1 = 20, link = "logit")
```

```
Deviance =          2436.854
Log-likelihood =    -1218.427
RIDGEMAX =           0.1
AIC =              -1224.427
SBC =              -1236.917
```

	Estimate	SE	Z	P.value
(Intercept)	-2.0811012	0.20284908	-10.259358	0.000000e+00
Item2vs1	3.8075990	0.26884433	14.162839	0.000000e+00
Item3vs1	-0.5709183	0.19782511	-2.885975	3.902033e-03
Random.(Intercept)	5.1395920	0.95909021	5.358820	8.376714e-08
Threshold2	1.1342293	0.09838877	11.528037	0.000000e+00
Threshold3	2.7832343	0.18658964	14.916339	0.000000e+00

To fit this model without the proportional odds assumption, the `KG` parameter is used. Here `KG=2` indicates not to assume proportional odds for the first 2 variables on the RHS of the equation. Again recall that when using `KG`, the order of the variables on the RHS is important because the integer value passed to `KG` represents the number of variables, starting with the first, for non-proportional odds estimation.

```
> Fitted.norcag.np<-mixord(SexItems~Item2vs1+Item3vs1, data=norcag, id=ID,
+   weights=freq, link="logit", NQ1=10, KG=2)
> summary(Fitted.norcag.np)
```

Call:

```
mixord(formula = SexItems ~ Item2vs1 + Item3vs1, data = norcag,
      id = ID, weights = freq, NQ1 = 10, link = "logit", KG = 2)
```

```
Deviance =          2401.609
Log-likelihood =    -1200.805
```

```
RIDGEMAX =      0.1
AIC =          -1210.805
SBC =          -1231.621
```

	Estimate	SE	Z	P.value
(Intercept)	-1.8190535	0.1927003	-9.4398040	0.000000e+00
Item2vs1	3.1554179	0.2673298	11.8034664	0.000000e+00
Item3vs1	-0.5996560	0.1990526	-3.0125497	2.590630e-03
Random.(Intercept)	4.4001069	0.8098931	5.4329477	5.543065e-08
Threshold2	1.5922660	0.1851041	8.6020038	0.000000e+00
Threshold3	3.2119220	0.3398884	9.4499318	0.000000e+00
Threshold2Item2vs1	0.9606217	0.2061822	4.6590915	3.176080e-06
Threshold3Item2vs1	1.0646810	0.3649120	2.9176375	3.526942e-03
Threshold2Item3vs1	0.2631802	0.2661222	0.9889450	3.226901e-01
Threshold3Item3vs1	-0.6172965	0.6539133	-0.9440036	3.451678e-01

The `mixord` function also allows the user to specify covariates that influence the scale through the `KS` parameter. For the $K - 1$ logits the model is

$$\log \left(\frac{P(Y_{ij} \leq k)}{1 - P(Y_{ij} \leq k)} \right) = \frac{\alpha_k - (\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{T} \boldsymbol{\theta}_i)}{\exp(\mathbf{w}_{ij}^T \boldsymbol{\tau})} \quad (5)$$

where \mathbf{w}_{ij} is the design matrix for the covariates that influence scale and $\boldsymbol{\tau}$ are their effects (Hedeker, Berbaum, and Mermelstein 2006). The `KS` parameter is specified in the same way that `KG` is; the order of the variables on the RHS is important because the integer value passed to `KS` represents the number of variables, starting with the first, for scaling.

```
> Fitted.norcag.scale<-mixord(SexItems~Item2vs1+Item3vs1, data=norcag, id=ID,
+ weights=freq, link="logit", NQ1=10, KS=2)
> summary(Fitted.norcag.scale)
```

Call:

```
mixord(formula = SexItems ~ Item2vs1 + Item3vs1, data = norcag,
       id = ID, weights = freq, NQ1 = 10, link = "logit", KS = 2)
```

```
Deviance =      2418.607
Log-likelihood = -1209.303
RIDGEMAX =      0.2
AIC =          -1217.303
SBC =          -1233.957
```

	Estimate	SE	Z	P.value
(Intercept)	-2.22115989	0.3359960	-6.61067306	3.825762e-11
Item2vs1	4.50668146	0.6914749	6.51749136	7.149303e-11
Item3vs1	-0.62054126	0.3698631	-1.67775923	9.339411e-02
Scale.Item2vs1	0.63125160	0.1966543	3.20995548	1.327555e-03
Scale.Item3vs1	-0.01540021	0.3165635	-0.04864809	9.611997e-01

```
Random.(Intercept)  6.78760830  2.4703548  2.74762488  6.002864e-03
Threshold2          1.39174094  0.2349246  5.92420391  3.138139e-09
Threshold3          3.61363106  0.5793038  6.23788650  4.435223e-10
```

5.4. Clustered Data

An example of naturally clustered data is twins clustered within twin pair. The outcome in the `concen` data reflects trouble concentrating (`TConcen`) which was recorded for both monozygotic and dizygotic twins (Ramakrishnan, Goldberg, Henderson, Eisen, True, Lyons, and Tsuang 1992). Each twin pair is uniquely identified by `ID` and type of twin is recorded using two dummy variables: `Mz`, an indicator variable representing MZ twins (1 = MZ, 0 = DZ) and `Dz`, an indicator variable representing DZ twins (1 = DZ, 0 = MZ). These data are also frequency weighted such that `freq` represents the frequency of the pattern.

The `mixord` function requires the data to be sorted by the `id` variable, which is performed prior to the function call.

```
> data(concen)
> concen<-concen[order(concen$ID),] # sort the data by twin pair ID
```

A common ICC probit model can be fit using

```
> Common.ICC<-mixord(TConcen~Mz, data=concen, id=ID, weights=freq,
+   link="probit", NQ1=10, random.effect.mean=FALSE)
> summary(Common.ICC)
```

Call:

```
mixord(formula = TConcen ~ Mz, data = concen, id = ID, weights = freq,
      NQ1 = 10, link = "probit", random.effect.mean = FALSE)
```

```
Deviance =          8549.288
Log-likelihood =    -4274.644
RIDGEMAX =           0
AIC =              -4277.644
SBC =              -4287.015
```

	Estimate	SE	Z	P.value
(Intercept)	0.74161872	0.03195820	23.205896	0.000000e+00
Mz	0.06257477	0.03948314	1.584848	1.130009e-01
Random.(Intercept)	0.36453782	0.05183966	7.032026	2.035483e-12

A Varying ICC probit model can be fit using

```
> Varying.ICC<-mixord(TConcen~Mz+Dz, data=concen, id=ID, weights=freq,
+   which.random.slope=1:2, exclude.fixed.effect=2, link="probit",
+   NQ1=20, random.effect.mean=FALSE, UNID=1)
> summary(Varying.ICC)
```

Call:

```
mixord(formula = TConcen ~ Mz + Dz, data = concen, id = ID, which.random.slope = 1:2,
       weights = freq, exclude.fixed.effect = 2, NQ1 = 20, link = "probit",
       random.effect.mean = FALSE, UNID = 1)
```

```
Deviance =          8543.072
Log-likelihood =    -4271.536
RIDGEMAX =           0
AIC =              -4275.536
SBC =              -4288.03
```

	Estimate	SE	Z	P.value
(Intercept)	0.7074312	0.03233795	21.876194	0.000000e+00
Mz	0.1314556	0.04787580	2.745764	6.037018e-03
Random.Mz	0.4917497	0.08139710	6.041366	1.528154e-09
Random.Dz	0.2346398	0.06469088	3.627093	2.866297e-04

Note that UNID is an indicator variable where UNID=0 (default) reflects that the random effects are multi-dimensional and UNID=1 reflects that the random effects are variables related to a uni-dimensional random effect (e.g., item indicators of a latent variable).

Summary

Herein we have described the **mixor** package which works in conjunction with the **MIXOR** Fortran stand-alone program in the R programming environment. The package provides a function for fitting cumulative link mixed-effects ordinal response models using either a probit, logit, log-log, or complementary log-log link function.

Acknowledgments

Research reported in this publication was supported by the National Library Of Medicine of the National Institutes of Health under Award Number R01LM011169. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Agresti A (2010). *Analysis of Ordinal Categorical Data*. John Wiley & Sons, Hoboken, NJ.
- Agresti A, Lang JB (1993). "A Proportional Odds Model with Subject-Specific Effects for Repeated Ordered Categorical Responses." *Biometrika*, **80**, 527–534.
- Christensen RHB (2013). "ordinal: Regression Models for Ordinal Data." URL <http://www.cran.r-project.org/package=ordinal/>.

- Cromer A, Carles A, Millon R, Ganguli G, Chalmel F, Lemaire F, Young J, Dembélé D, Thibault C, Muller D, Poch O, Abesassis J, Wasylyk B (2004). “Identification of Genes Associated with Tumorigenesis and Metastatic Potential of Hypopharyngeal Cancer by Microarray Analysis.” *Oncogene*, **23**, 2484–2498.
- Flay BR, Miller TQ, Hedeker D, Siddiqui O, Brannon BR, Johnson CA, Hansen WB, Sussman S, Dent C (1995). “The Television, School and Family Smoking Prevention and Cessation Project: VIII. Student Outcomes and Mediating Variables.” *Preventive Medicine*, **24**, 29–40.
- Gibbons RD, Hedeker D (1994). “Application of Random-Effects Probit Regression Models.” *Journal of Consulting and Clinical Psychology*, **62**, 285–296.
- Hedeker D, Berbaum M, Mermelstein RJ (2006). “Location-Scale Models for Multilevel Ordinal Data: Between- and Within-Subjects Variance Modeling.” *Journal of Probability and Statistical Science*, **4**, 1–20.
- Hedeker D, Gibbons RD (1996). “MIXOR: a Computer Program for Mixed-Effects Ordinal Regression Analysis.” *Computer Methods and Programs in Biomedicine*, **49**, 157–176.
- Hedeker D, Gibbons RD (2006). *Longitudinal Data Analysis*. Wiley-Interscience, Hoboken, NJ.
- Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, Lindahl T, Pawitan Y, Hall P, Nordgren H, Wong JEL, Liu ET, Bergh J, Kuznestsov VA, Miller L (2006). “Genetic Reclassification of Histologic Grade Delineates New Clinical Subtypes of Breast Cancer.” *Cancer Research*, **66**, 10292–10301.
- McCullagh P (1980). “Regression Models for Ordinal Data.” *Journal of the Royal Statistical Society, B*, **42**, 109–142.
- Peterson B, Harrell FE (1990). “Partial Proportional Odds Models for Ordinal Response Variables.” *Applied Statistics*, **39**, 205–217.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Ramakrishnan V, Goldberg J, Henderson WG, Eisen SA, True W, Lyons MJ, Tsuang MT (1992). “Elementary Methods for the Analysis of Dichotomous Outcomes in Unselected Samples of Twins.” *Genetic Epidemiology*, **9**, 273–287.
- Therneau T (2013). “A Package for Survival Analysis in S. R package version 2.37-4.” URL <http://CRAN.R-project.org/package=survival>.

Affiliation:

Kellie J. Archer
Department of Biostatistics
Virginia Commonwealth University

Box 980032
Richmond, VA 23298-0032
E-mail: kjarcher@vcu.edu
URL: <http://www.people.vcu.edu/~kjarcher/>

Donald Hedeker
Division of Epidemiology and Biostatistics
School of Public Health
University of Illinois at Chicago
1603 West Taylor Street
Chicago, IL 60612-4394
E-mail: hedeker@uic.edu
URL: <http://tigger.uic.edu/~hedeker/>

Rachel Nordgren
Division of Epidemiology and Biostatistics
School of Public Health
University of Illinois at Chicago
1603 West Taylor Street
Chicago, IL 60612-4394
E-mail: rknordgren@gmail.com

Robert D. Gibbons
Professor of Medicine and Health Studies
Director, Center for Health Statistics
University of Chicago
5841 S. Maryland Avenue
MC 2007 office W260
Chicago IL 60637
E-mail: rdg@uchicago.edu