A SAS/IML macro for GEE estimation of logistic regression parameters in correlated data with exchangeable correlation structure (optimized for large clusters).

(c) Bahjat F. Qaqish and Habib Moalem (1995)
Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7420, U.S.A.
bahjat_qaqish@unc.edu

1 Introduction

The file "blex.sas" implements estimation by generalized estimating equations (GEE) for logistic regression with correlated responses. The code is optimized (space and time) for large clusters. The response must be a 0 or 1 binary outcome. The only link function supported is the logit link. The only correlation structure supported is exchangeable (compound symmetry) correlation.

2 Syntax

The macro accepts many arguments, most of them optional. It is called as follows, with default values shown where appropriate:

```
%gee   (
  data=,              /* required: dataset name */
  yvar=,              /* required: response 0/1 */
  xvar=,              /* required: covariates */
  id=,                /* required: cluster id */
  offset=,            /* offset */
  epsilon=1.0e-4,     /* criterion for beta convergence */
  epsilon2=1.0e-2,    /* criterion for rho  convergence */
  monitor=no,         /* monitor convergence */
  beta=,              /* initial values */
  rho=0,              /* initial value */
  fixrho=no,          /* fixrho: no/yes */
  fixbeta=no,         /* fixbeta: no/yes */
  outest=,            /* estimtates and cov matrices dataset */
```

```
  outs=,                  /* fitted values and residuals dataset */
  conf=0.95,              /* confidence intervals */
  maxiter = 25   );       /* max. # iterations */
```

The arguments are as follows:

data: This argument specifies the name of the input SAS dataset (required).

yvar: This argument specifies the name of the binary (0 or 1) response variable (numeric, required).

xvar: This argument specifies the covariate(s) (numeric, required). If an intercept term is desired it must be included explicitly.

id: This argument specifies the cluster identifier (numeric, required).

offset: This argument specifies an offset variable (numeric, optional).

epsilon: This argument specifies the tolerance for convergence of the regression parameter β estimate (numeric, optional).

epsilon2: This argument specifies the tolerance for convergence of the correlation parameter ρ estimate (numeric, optional).

monitor: This argument controls the printing of intermediate estimates after each iteration. (optional, yes/no).

beta: This argument provides initial values of the regression parameters (numeric, optional).

rho: This argument provides initial value of the correlation parameter (numeric, optional).

fixrho: This argument controls updating of the correlation parameter (yes/no, optional). If fixrho=yes then ρ will be fixed at its initial value.

fixbeta: This argument controls updating of the regession parameters (yes/no, optional). If fixbeta=yes then β will be fixed at its initial value.

outest: This argument specifies the name of a dataset to which parameter estimates and estimated covariance matrices will be written (optional).

outs: This argument specifies the name of a dataset to which fitted values and residuals will be written (optional).

conf: This argument specifies the level of confidence intervals (numeric, optional).

maxiter: This argument specifies maximum number of iterations allowed (numeric, optional).

3 Example

The following program simulates a dataset then calls the macro to fit a model.

```
filename MACRO 'blex.sas';
%include MACRO;
**************************************************;
data A;         * generate a dataset;
  k = 20;       * # clusters;
```

```
  n = 5  ;      * cluster size;
  beta1 = 1;    * parameter values;
  beta2 =-1;
  beta3 = 0;
  beta4 =-1;
  beta5 = 1;
  sigma = 2;
  seed = 2141994;
  retain x1 1;  * intercept;
  array x (*) x1-x5;
  array beta (*) beta1-beta5;
  do id = 1 to k;       * one cluster;
    e = rannor(seed);
    x2 = rannor(seed);
    x3 = rannor(seed);
    do j = 1 to n;      * one obs.;
      x4 = rannor(seed);
      x5 = rannor(seed);
      lp = e * sigma;   * linear predictor;
      do l=1 to dim(beta);
        lp=lp+x(l)*beta(l);
      end;
      p = 1/(1+exp(-lp));
      y = (ranuni(seed) <= p);
      output;
    end;        * one obs.;
  end;  * one cluster;
  keep id y x1 - x5;
  run;
***************************************************;
%gee    (
  data=A,                  /* required: dataset name */
  yvar=y,                  /* required: response 0/1 */
  xvar=x1 x2 x3 x4 x5,     /* required: covariates */
  id=id,                   /* required: cluster id */
  );
**************************************************;
```

4 Output

```
Data set: A
Response: y
Cluster ID: id
Number of observations:      100

Number of covariates:        5

Number of clusters:       20

Cluster sizes: from         5  to         5

Converged, niter =         14

Number of estimable parameters =         5

Estimated intra-cluster correlation =  0.4724317

Summary of estimates and 0.95 confidence intervals:

Results using the robust variance estimator:

           Estimate  Robust SE         z      p-val Odds Ratio

X1        0.3081574  0.3365101  0.9157447  0.3598008  1.3609151
X2         -0.44014  0.3711703  -1.185818  0.2356942   0.643946
X3        -0.799785  0.3721224  -2.149252  0.0316144  0.4494256
X4        -0.238786  0.1537953  -1.552619  0.1205141  0.7875838
X5        0.7792223  0.3047162  2.5572064  0.0105517  2.1797764


          CI-lower   CI-upper

X1        0.703709  2.6318975
X2       0.3111063  1.3328773
X3       0.2167238  0.9319851
X4       0.5826207  1.0646517
X5       1.1996016  3.9608361

Results using the naive variance estimator:
```

	Estimate	Naive SE	z	p-val	Odds Ratio
X1	0.3081574	0.4191664	0.7351671	0.4622377	1.3609151
X2	-0.44014	0.3512052	-1.253228	0.2101226	0.643946
X3	-0.799785	0.4469432	-1.789456	0.0735414	0.4494256
X4	-0.238786	0.1831444	-1.30381	0.1922984	0.7875838
X5	0.7792223	0.2246308	3.4689018	0.0005226	2.1797764

	CI-lower	CI-upper
X1	0.5984611	3.094754
X2	0.3235214	1.281728
X3	0.1871626	1.0791864
X4	0.5500522	1.1276896
X5	1.4034787	3.385463