

# MCMCGuide10

*Silvan Hüsler*

*18 Januar 2017*

## Contents

10.1 Simple logistic regression model . . . . .	1
10.2 Random effects logistic regression model . . . . .	3
10.3 Random coefficients for area type . . . . .	5
10.4 Probit regression . . . . .	6
10.5 Running a probit regression in MLwiN . . . . .	6

```
data(bang1, package = "R2MLwiN")
```

## 10.1 Simple logistic regression model

```
(mymodel1 <-  
  runMLwiN(  
    logit(use) ~ 1 + age,  
    D = "Binomial",  
    estoptions = list(  
      EstM = 1),  
    data = bang1))  
  
## MLwiN is running, please wait.....  
  
##  
## -----  
## MLwiN (version: 2.36) multilevel model (Binomial)  
## Estimation algorithm: MCMC      Elapsed time : 38.53s  
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000  Chains: 1  Burn-in: 500  
## Bayesian Deviance Information Criterion (DIC)  
## Dbar      D(thetabar)    pD      DIC  
## 2591.282   2589.292     1.990    2593.272  
## -----  
## The model formula:  
## logit(use) ~ 1 + age  
## Level 1: llid  
## -----  
## The fixed part estimates:  
##           Coef.   Std. Err.      z    Pr(>|z|)      [95% Cred.   Interval]    ESS  
## Intercept  -0.43703   0.04725   -9.25  2.264e-20 ***   -0.52609   -0.34091   1095  
## age         0.00633   0.00506    1.25   0.2112        -0.00356    0.01637   1058  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
## -----  
## The random part estimates at the llid level:  
##           Coef.   Std. Err.   [95% Cred.   Interval]    ESS  
## var_bcons_1  1.00000     1e-05     1.00000     1.00000   5000  
## -----  
  
summary(mymodel1@chains[, "FP_age"])
```

```
##
## Iterations = 1:5000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 5000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean           SD      Naive SE Time-series SE
##      6.327e-03    5.061e-03    7.157e-05    1.556e-04
##
## 2. Quantiles for each variable:
##
##      2.5%      25%      50%      75%      97.5%
## -0.003556  0.002927  0.006276  0.009770  0.016368

sixway(mymodel1@chains[, "FP_age", drop = FALSE], "beta_1")
```

15,000 iterations

```
(mymodel2 <-
  runMLwiN(
    logit(use) ~ 1 + age,
    D = "Binomial",
    estoptions = list(
      EstM = 1,
      mcmcMeth = list(iterations = 15000)),
    data = bang1))

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
## Estimation algorithm: MCMC      Elapsed time : 23.38s
## Number of obs: 1934 (from total 1934)      Number of iter.: 15000  Chains: 1  Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)    pD      DIC
## 2591.284   2589.290    1.994    2593.278
## -----
## The model formula:
## logit(use) ~ 1 + age
## Level 1: llid
## -----
## The fixed part estimates:
##           Coef.    Std. Err.      z    Pr(>|z|)    [95% Cred. Interval]    ESS
## Intercept  -0.43781    0.04666   -9.38  6.45e-21 ***   -0.52870    -0.34622    3402
## age         0.00656    0.00514    1.28  0.2015         -0.00352     0.01686    3311
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the llid level:
##           Coef.    Std. Err.    [95% Cred. Interval]    ESS
## var_bcons_1  1.00000    1e-05      1.00000      1.00000    15000
## -----
```

```
sixway(mymodel1@chains[, "FP_age", drop = FALSE], "beta_1")
```

Change to 5000 iterations by default

```
(mymodel3 <-
  runMLwiN(
    logit(use) ~ 1 + age + lc,
    D = "Binomial",
    estoptions = list(EstM = 1),
    data = bang1))

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
## Estimation algorithm: MCMC      Elapsed time : 23.33s
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000  Chains: 1  Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)    pD      DIC
## 2519.976   2515.099    4.877    2524.852
## -----
## The model formula:
## logit(use) ~ 1 + age + lc
## Level 1: l1id
## -----
## The fixed part estimates:
##              Coef.    Std. Err.      z    Pr(>|z|)    [95% Cred.    Interval]    ESS
## Intercept    -1.26579    0.11517   -10.99   4.225e-28   ***    -1.48935    -1.04721    102
## age          -0.02186    0.00725    -3.01   0.002587   **     -0.03700    -0.00816    244
## lcOne_child   1.03150    0.14341     7.19   6.344e-13   ***     0.75458     1.29899    303
## lcTwo_children 1.18475    0.16572     7.15   8.725e-13   ***     0.86284     1.50298    201
## lcThree_plus  1.11811    0.16513     6.77   1.277e-11   ***     0.82670     1.45181    114
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the l1id level:
##              Coef.    Std. Err.    [95% Cred.    Interval]    ESS
## var_bcons_1  1.00000    1e-05      1.00000    1.00000    5000
## -----
```

## 10.2 Random effects logistic regression model

```
(mymodel4 <-
  runMLwiN(
    logit(use) ~ 1 + age + lc + (1 | district),
    D = "Binomial",
    estoptions = list(EstM = 1),
    data = bang1))

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
```

```

##          N min      mean max N_complete min_complete mean_complete
## district 60    2 32.23333 118          60          2      32.23333
##          max_complete
## district      118
## Estimation algorithm: MCMC      Elapsed time : 20.94s
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000  Chains: 1  Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)      pD      DIC
## 2395.915    2354.503    41.412    2437.328
## -----
## The model formula:
## logit(use) ~ 1 + age + lc + (1 | district)
## Level 2: district      Level 1: llid
## -----
## The fixed part estimates:
##          Coef.      Std. Err.      z      Pr(>|z|)      [95% Cred.      Interval]      ESS
## Intercept      -1.46065      0.12898     -11.32     9.878e-30     ***      -1.71772      -1.21598     108
## age             -0.02473      0.00747      -3.31     0.0009365     ***      -0.04016      -0.01094     377
## lcOne_child      1.08142      0.14487      7.46     8.349e-14     ***      0.78755      1.36517     298
## lcTwo_children    1.29737      0.15510      8.36     6.027e-17     ***      1.00176      1.59121     201
## lcThree_plus      1.27047      0.15697      8.09     5.795e-16     ***      0.96509      1.58183     139
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the district level:
##          Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_Intercept    0.30935      0.10079      0.16078      0.55109     417
## -----
## The random part estimates at the llid level:
##          Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_bcons_1      1.00000      1e-05      1.00000      1.00000     5000
## -----
summary(mymodel4@chains[, "RP2_var_Intercept"])

##
## Iterations = 1:5000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 5000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##          Mean          SD          Naive SE Time-series SE
##          0.309348      0.100789      0.001425      0.004937
##
## 2. Quantiles for each variable:
##
##          2.5%      25%      50%      75%      97.5%
##          0.1608 0.2379 0.2930 0.3612 0.5511

sixway(mymodel4@chains[, "RP2_var_Intercept", drop = FALSE], "sigma2u0")

```

## 10.3 Random coefficients for area type

```
(mymodel5 <-
  runMLwiN(
    logit(use) ~ 1 + age + lc + urban + (1 | district),
    D = "Binomial", estoptions = list(EstM = 1),
    data = bang1))

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
##      N min      mean max N_complete min_complete mean_complete
## district 60    2 32.23333 118          60              2      32.23333
##      max_complete
## district      118
## Estimation algorithm: MCMC      Elapsed time : 25.06s
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000 Chains: 1 Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)      pD      DIC
## 2369.363    2330.577    38.786    2408.149
## -----
## The model formula:
## logit(use) ~ 1 + age + lc + urban + (1 | district)
## Level 2: district      Level 1: l1id
## -----
## The fixed part estimates:
##      Coef.      Std. Err.      z      Pr(>|z|)      [95% Cred.      Interval]      ESS
## Intercept      -1.71506      0.16280     -10.53    5.991e-26    ***      -2.04311      -1.40106      70
## age             -0.02679      0.00805      -3.33    0.0008673    ***      -0.04202      -0.01015     189
## lcOne_child      1.13969      0.15923      7.16     8.223e-13    ***      0.84281       1.45565     253
## lcTwo_children   1.39124      0.18170      7.66     1.903e-14    ***      1.03221       1.73964     152
## lcThree_plus     1.36047      0.18416      7.39     1.498e-13    ***      0.99559       1.73294      93
## urbanUrban       0.73553      0.12339      5.96     2.504e-09    ***      0.49766       0.97904     388
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the district level:
##      Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_Intercept    0.23668      0.08368      0.10751      0.43849      330
## -----
## The random part estimates at the l1id level:
##      Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_bcons_1      1.00000      1e-05      1.00000      1.00000      5000
## -----
## -----
(mymodel6 <-
  runMLwiN(
    logit(use) ~ 1 + age + lc + urban + (1 + urban | district),
    D = "Binomial",
    estoptions = list(EstM = 1),
    data = bang1))

## MLwiN is running, please wait.....

##
```

```

## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
##           N min      mean max N_complete min_complete mean_complete
## district 60    2 32.23333 118           60           2       32.23333
##           max_complete
## district      118
## Estimation algorithm: MCMC      Elapsed time : 32.59s
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000  Chains: 1  Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)      pD      DIC
## 2328.933   2272.311   56.622   2385.556
## -----
## The model formula:
## logit(use) ~ 1 + age + lc + urban + (1 + urban | district)
## Level 2: district      Level 1: llid
## -----
## The fixed part estimates:
##           Coef.      Std. Err.      z      Pr(>|z|)      [95% Cred.      Interval]      ESS
## Intercept      -1.72354      0.15906     -10.84     2.332e-27     ***      -2.03039      -1.39442      61
## age             -0.02727      0.00786      -3.47     0.0005197     ***      -0.04282      -0.01122     203
## lcOne_child      1.15731      0.15738      7.35     1.928e-13     ***      0.83848       1.46651     229
## lcTwo_children   1.37595      0.17441      7.89     3.039e-15     ***      1.02279       1.70107     186
## lcThree_plus     1.38420      0.17993      7.69     1.439e-14     ***      1.00169       1.73599     102
## urbanUrban       0.80537      0.18896      4.26     2.025e-05     ***      0.42332       1.18309     110
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the district level:
##           Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_Intercept      0.41802      0.13682      0.20269      0.73076      206
## cov_Intercept_urbanUrban -0.43246      0.17573     -0.83467     -0.15496     127
## var_urbanUrban      0.73806      0.30328      0.29716      1.45696     142
## -----
## The random part estimates at the llid level:
##           Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_bcons_1      1.00000      1e-05      1.00000      1.00000     5000
## -----

```

## 10.4 Probit regression

## 10.5 Running a probit regression in MLwiN

### Gibbs

```

(mymodel7 <-
  runMLwiN(
    probit(use) ~ 1 + age + lc + urban + (1 + urban | district),
    D = "Binomial",
    estoptions = list(
      EstM = 1,
      mcmcMeth = list(
        fixM = 1,
        residM = 1)),
    data = bang1))

```

```

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
##           N min      mean max N_complete min_complete mean_complete
## district 60    2 32.23333 118           60           2       32.23333
##           max_complete
## district           118
## Estimation algorithm: MCMC      Elapsed time : 20.14s
## Number of obs: 1934 (from total 1934)      Number of iter.: 5000  Chains: 1  Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar      D(thetabar)      pD      DIC
## 2328.195    1007.337    1320.858    3649.052
## -----
## The model formula:
## probit(use) ~ 1 + age + lc + urban + (1 + urban | district)
## Level 2: district      Level 1: l1id
## -----
## The fixed part estimates:
##           Coef.      Std. Err.      z      Pr(>|z|)      [95% Cred.      Interval]      ESS
## Intercept      -1.03853      0.09948     -10.44     1.63e-25     ***      -1.24024      -0.84696      564
## age             -0.01616      0.00490      -3.30     0.0009643     ***      -0.02572      -0.00659     1731
## lcOne_child      0.68254      0.09656      7.07     1.566e-12     ***      0.49400      0.86843     1770
## lcTwo_children    0.82280      0.10595      7.77     8.12e-15     ***      0.61187      1.02849     1761
## lcThree_plus     0.81935      0.11088      7.39     1.476e-13     ***      0.60471      1.03509     1347
## urbanUrban       0.50127      0.10917      4.59     4.399e-06     ***      0.28980      0.72040      601
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the district level:
##           Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_Intercept      0.15505      0.04953      0.07702      0.26990      740
## cov_Intercept_urbanUrban -0.16011      0.06674      -0.32193      -0.05679     521
## var_urbanUrban      0.26745      0.11815      0.09442      0.55287     399
## -----
## The random part estimates at the l1id level:
##           Coef.      Std. Err.      [95% Cred.      Interval]      ESS
## var_bcons_1      1.00000      1e-05      1.00000      1.00000     5000
## -----

```

## Univariate MH by default

```

(mymodel8 <-
  runMLwiN(
    probit(use) ~ 1 + age + lc + urban + (1 + urban | district),
    D = "Binomial",
    estoptions = list(EstM = 1),
    data = bang1))

## MLwiN is running, please wait.....

##
## -----
## MLwiN (version: 2.36) multilevel model (Binomial)
##           N min      mean max N_complete min_complete mean_complete

```

```

## district 60 2 32.23333 118 60 2 32.23333
## max_complete
## district 118
## Estimation algorithm: MCMC Elapsed time : 44.74s
## Number of obs: 1934 (from total 1934) Number of iter.: 5000 Chains: 1 Burn-in: 500
## Bayesian Deviance Information Criterion (DIC)
## Dbar D(thetabar) pD DIC
## 2327.162 2271.196 55.966 2383.128
## -----
## The model formula:
## probit(use) ~ 1 + age + lc + urban + (1 + urban | district)
## Level 2: district Level 1: llid
## -----
## The fixed part estimates:
## Coef. Std. Err. z Pr(>|z|) [95% Cred. Interval] ESS
## Intercept -1.04763 0.09898 -10.58 3.511e-26 *** -1.24631 -0.86649 63
## age -0.01640 0.00466 -3.52 0.0004287 *** -0.02549 -0.00736 241
## lcOne_child 0.69214 0.10000 6.92 4.471e-12 *** 0.50055 0.88713 281
## lcTwo_children 0.82995 0.10818 7.67 1.695e-14 *** 0.61289 1.04145 205
## lcThree_plus 0.82927 0.10760 7.71 1.289e-14 *** 0.62619 1.04235 109
## urbanUrban 0.50550 0.10936 4.62 3.791e-06 *** 0.30074 0.74285 92
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## -----
## The random part estimates at the district level:
## Coef. Std. Err. [95% Cred. Interval] ESS
## var_Intercept 0.15931 0.05075 0.07825 0.27472 136
## cov_Intercept_urbanUrban -0.17052 0.06553 -0.32494 -0.07309 123
## var_urbanUrban 0.28121 0.11114 0.11954 0.54435 129
## -----
## The random part estimates at the llid level:
## Coef. Std. Err. [95% Cred. Interval] ESS
## var_bcons_1 1.00000 1e-05 1.00000 1.00000 5000
## -----
library(texreg)

## Version: 1.36.18
## Date: 2016-10-22
## Author: Philip Leifeld (University of Glasgow)
##
## Please cite the JSS article in your publications -- see citation("texreg").
##
## Attaching package: 'texreg'
##
## The following object is masked from 'package:magrittr':
##
## extract
##
## The following object is masked from 'package:tidyr':
##
## extract
screenreg(
  list(mymodel7, mymodel8),
  custom.model.names=c("Gibbs", "Metropolis"),
  groups = list(

```



```

    "Fixed Part" = 1:6,
    "Level-2"    = 7:9,
    "Level-1"    = 10:10),
  stars = numeric(0),
  include.nobs=FALSE,
  include.loglik=FALSE,
  include.deviance=FALSE,
  include.dbar=FALSE,
  include.dthetabar=FALSE,
  include.pd=FALSE,
  include.dic=FALSE)

##
## =====
##                               Gibbs           Metropolis
## -----
## Fixed Part
##
##      FP_Intercept           -1.04 *          -1.05 *
##                               [-1.24; -0.85]   [-1.25; -0.87]
##      FP_age                  -0.02 *          -0.02 *
##                               [-0.03; -0.01]   [-0.03; -0.01]
##      FP_lcOne_child           0.68 *          0.69 *
##                               [ 0.49; 0.87]   [ 0.50; 0.89]
##      FP_lcTwo_children         0.82 *          0.83 *
##                               [ 0.61; 1.03]   [ 0.61; 1.04]
##      FP_lcThree_plus          0.82 *          0.83 *
##                               [ 0.60; 1.04]   [ 0.63; 1.04]
##      FP_urbanUrban            0.50 *          0.51 *
##                               [ 0.29; 0.72]   [ 0.30; 0.74]
##
## Level-2
##
##      RP2_var_Intercept        0.16 *          0.16 *
##                               [ 0.08; 0.27]   [ 0.08; 0.27]
##      RP2_cov_Intercept_urbanUrban -0.16 *      -0.17 *
##                               [-0.32; -0.06]  [-0.32; -0.07]
##      RP2_var_urbanUrban        0.27 *          0.28 *
##                               [ 0.09; 0.55]   [ 0.12; 0.54]
##
## Level-1
##
##      RP1_var_bcons_1          1.00 *          1.00 *
##                               [ 1.00; 1.00]   [ 1.00; 1.00]
## =====
## * 0 outside the confidence interval
cat("The effective sample sizes\n")
## The effective sample sizes
ESS.aa <- effectiveSize(mymodel17@chains[, 2:11])
ESS.bb <- effectiveSize(mymodel18@chains[, 2:11])
ctable <- cbind(round(ESS.aa), round(ESS.bb))
colnames(ctable) <- c("ESS(Gibbs)", "ESS(Metropolis)")

```

```
print(ctable)
```

##	ESS(Gibbs)	ESS(Metropolis)
## FP_Intercept	564	63
## FP_age	1731	241
## FP_lcOne_child	1770	281
## FP_lcTwo_children	1761	205
## FP_lcThree_plus	1347	109
## FP_urbanUrban	601	92
## RP2_var_Intercept	740	136
## RP2_cov_Intercept_urbanUrban	521	123
## RP2_var_urbanUrban	399	129
## RP1_var_bcons_1	5000	5000

Hox, Joop, and Leoniek Wijngaards-de Meij. 2014. "The Multilevel Regression Model." In *The Sage Handbook of Regression Analysis and Causal Inference*, edited by Henning Best and Christof Wolf, 133–52. SAGE Publications Ltd. doi:10.4135/9781446288146.n7.

Zhang, Zhengzheng, Richard M. A. Parker, Christopher M. J. Charlton, George Leckie, and William J. Browne. 2016. "R2MLwiN: A Package to Run Mlwin from Within R." *Journal of Statistical Software* 72 (10). doi:10.18637/jss.v072.i10.