



Group-wise adaptive testing

This example illustrates the use of BILOG-MG with multiple groups and multiple subtests. It is designed to illustrate some of the more complicated features of the program, including user-specified priors on the latent distributions and priors on the item parameters.

Based on previous test performance, examinees are assigned to two groups for adaptive testing. Out of a set of 45 items, group 1 is assigned items 1 through 25, and group 2 is assigned items 21 through 45. Thus, there are 5 items linking the test forms administered to the groups.

Twenty of the 25 items presented to group 1 belong to subtest 1 (items 1-15 and 21-25); twenty items also belong to subtest 2 (items 6-25). Of the 25 items presented to group 2, 20 belong to subtest 1 (items 21-40) and 20 to subtest 2 (items 21-25 and 31-45).

In all, there are 35 items from the set of 45 assigned to each subtest. (This extent of item overlap between subtests is not realistic, but it illustrates that more than one subtest can be scored adaptively provided they each contain link items between the test forms.)

This example also illustrates how user-supplied priors for the latent distributions are specified with IDIST=1 on the CALIB command. The points and weights for these distributions are supplied in the QUAD commands. Note that with IDIST=1, there are separate QUAD commands for each group for each subtest. Within each subtest the points are the same for each group. This is a requirement of the program. But as the example shows, the points for the groups may differ by subtest. If IDIST has been set to 2, sets of weights have to be supplied by group. The set of points then applies to all subtests.

The PRIOR command for each subtest is placed after the QUAD commands for that subtest. The presence of the PRIOR command is indicated using the READPRIOR option on the CALIB command. In this example, only the prior for the standard deviation of the thresholds is supplied on the PRIOR command. Default values are used for the other prior distributions. The means of the distributions are kept fixed at their specified values by using the NOFLOAT option on the CALIB command.

The score distribution in the respondent population is estimated in the form of a discrete distribution on NQPT=16 points by adding the EMPIRICAL option to the CALIB command. This discrete distribution will be used in the place of the prior in MML estimate of the item parameters. When NGROUP>1, separate score distributions will be estimated for the groups. The first group serves as the reference group (REFERENCE=1). If the REFERENCE keyword is omitted, the first group will by default be used as the reference group. When NGROUP>1, the FLOAT option is the default. By using NOFLOAT here, the means of the prior distributions on item parameters are kept fixed at the specified values during estimation.

In the scoring phase, the empirical prior from phase 2 is used as prior distribution for the scale scores (IDIST=3). Rescaling of scores to the scale and location in the sample of scale score estimates is requested by setting RSCTYPE to 3. The presence of the INFO keyword indicates that information output is required. In this case INFO=1 and test information curves will be printed to the phase 3 output file. In combination with the YCOMMON and POP options, the test information curves will be expressed in comparable units and an estimate of the classical reliability coefficient, amongst other information, will be calculated for each subtest.

```

EXAMPL08.BLM -
                GROUP-WISE ADAPTIVE TESTING WITH TWO SUBTESTS
>GLOBAL DFNAME='EXAMPL08.DAT', NPARAM=2, NTEST=2, SAVE;
>SAVE SCORE='EXAMPL08.SCO';
>LENGTH NITEMS=(35,35);
>INPUT  NTOT=45, SAMPLE=2000, NGROUP=2, KFNAME='EXAMPL08.DAT', NALT=5,
        NFORMS=2, NIDCH=5;
>ITEMS  INUM=(1(1)45), INAME=(C01(1)C45);
>TEST1  TNAME=SUBTEST1, INAME=(C01(1)C15,C21(1)C40);
>TEST2  TNAME=SUBTEST2, INAME=(C06(1)C25,C31(1)C45);
>FORM1  LENGTH=25, INUM=(1(1)25);
>FORM2  LENGTH=25, INUM=(21(1)45);
>GROUP1 GNAME=POP1, LENGTH=25, INUM=(1(1)25);
>GROUP2 GNAME=POP2, LENGTH=25, INUM=(21(1)45);
(5A1,T1,I1,T1,I1,T7,25A1)
>CALIB  IDIST=1, READPRIOR, EMPIRICAL, NQPT=31, CYCLE=25, TPRIOR, NEWTON=5,
        CRITERION=0.01, REFERENCE=1, NOFLOAT;
>QUAD1  POINTS=(-0.4598E+01,-0.3560E+01,-0.2522E+01,-0.1484E+01,-0.4453E+00,
        0.5930E+00, 0.1631E+01, 0.2670E+01, 0.3708E+01, 0.4746E+01),
        WEIGHTS=(0.2464E-05, 0.4435E-03, 0.1724E-01, 0.1682E+00, 0.3229E+00,
        0.3679E+00, 0.1059E+00, 0.1685E-01, 0.6475E-03, 0.8673E-05);
>QUAD2  POINTS=(-0.4598E+01,-0.3560E+01,-0.2522E+01,-0.1484E+01,-0.4453E+00,
        0.5930E+00, 0.1631E+01, 0.2670E+01, 0.3708E+01, 0.4746E+01),
        WEIGHTS=(0.2996E-04, 0.1300E-02, 0.1474E-01, 0.1127E+00, 0.3251E+00,
        0.3417E+00, 0.1816E+00, 0.2149E-01, 0.1307E-02, 0.3154E-04);
>PRIOR  TSIGMA=(1.5(0)35);
>QUAD1  POINTS=(-0.4000E+01,-0.3111E+01,-0.2222E+01,-0.1333E+01,-0.4444E+00,
        0.4444E+00, 0.1333E+01, 0.2222E+01, 0.3111E+01, 0.4000E+01),
        WEIGHTS=(0.1190E-03, 0.2805E-02, 0.3002E-01, 0.1458E+00, 0.3213E+00,
        0.3213E+00, 0.1458E+00, 0.3002E-01, 0.2805E-02, 0.1190E-03);
>QUAD2  POINTS=(-0.4000E+01,-0.3111E+01,-0.2222E+01,-0.1333E+01,-0.4444E+00,
        0.4444E+00, 0.1333E+01, 0.2222E+01, 0.3111E+01, 0.4000E+01),
        WEIGHTS=(0.1190E-03, 0.2805E-02, 0.3002E-01, 0.1458E+00, 0.3213E+00,
        0.3213E+00, 0.1458E+00, 0.3002E-01, 0.2805E-02, 0.1190E-03);
>PRIOR  TSIGMA=(1.5(0)35);
>SCORE  IDIST=3, RSCTYPE=3, INFO=1, YCOMMON, POP, NOPRINT;

```

Phase 1 output

Phase 1 echoes the assignment of items to subtests, forms, and groups. Classical item statistics are computed for each subtest in each group. Output for subtest 1 and group 1 (POP1) is given below.

```

SUBTEST  1      SUBTEST1
  GROUP   1      POP1          200 OBSERVATIONS
  GROUP   2      POP2          200 OBSERVATIONS

SUBTEST  2      SUBTEST2
  GROUP   1      POP1          200 OBSERVATIONS
  GROUP   2      POP2          200 OBSERVATIONS

```

```

SUBTEST  1      SUBTEST1
ITEM STATISTICS FOR GROUP:  1      POP1
                                ITEM*TEST CORRELATION
ITEM  NAME      #TRIED  #RIGHT  PCT      LOGIT/1.7  PEARSON  BISERIAL
-----
  1    C01        200.0   170.0   0.850    -1.02      0.408    0.625
  2    C02        200.0   164.0   0.820    -0.89      0.396    0.580
  3    C03        200.0   154.0   0.770    -0.71      0.451    0.626

```

| | | | | | | | |
|-----|-----|-------|-------|-------|-------|-------|-------|
| 4 | C04 | 200.0 | 143.0 | 0.715 | -0.54 | 0.400 | 0.532 |
| 5 | C05 | 200.0 | 140.0 | 0.700 | -0.50 | 0.586 | 0.772 |
| 6 | C06 | 200.0 | 135.0 | 0.675 | -0.43 | 0.441 | 0.574 |
| ... | | | | | | | |
| 19 | C24 | 200.0 | 83.0 | 0.415 | 0.20 | 0.590 | 0.746 |
| 20 | C25 | 200.0 | 76.0 | 0.380 | 0.29 | 0.558 | 0.711 |

Phase 2 output

Phase 2 estimates empirical latent distributions for each group and item parameters for each subtest. The arbitrary mean and standard deviation of reference group 1 determine the origin and unit of the ability scales.

```
>CALIB IDIST=1, READPRIOR, EMPIRICAL, NQPT=31, CYCLE=25, TPRIOR, NEWTON=5,
CRITERION=0.01, REFERENCE=1, NOFLOAT;
```

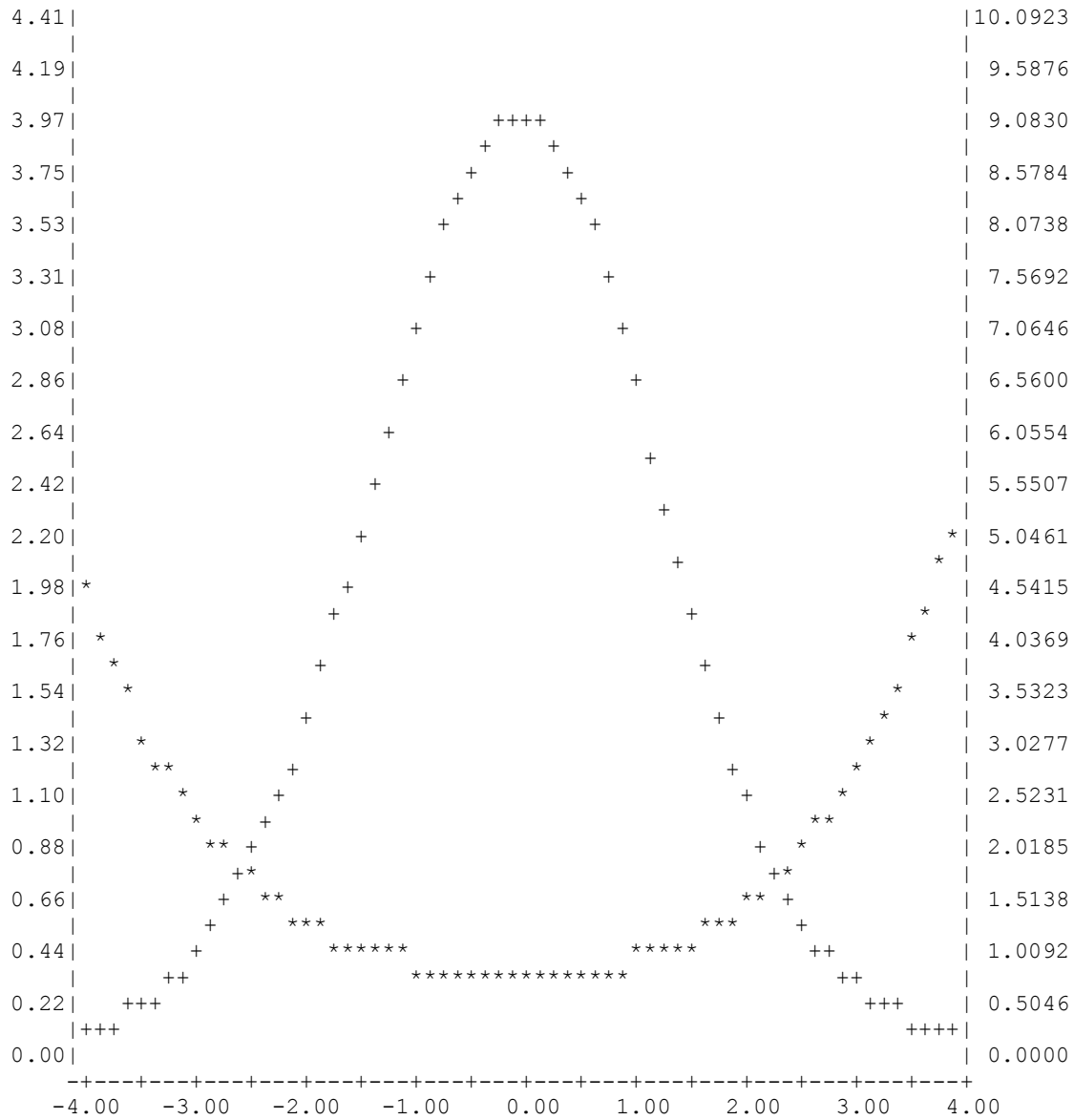
| ITEM | INTERCEPT S.E. | SLOPE S.E. | THRESHOLD S.E. | LOADING S.E. | ASYMPTOTE S.E. | CHISQ (PROB) | DF |
|------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-----|
| C01 | 1.435 0.196* | 0.930 0.187* | -1.542 0.211* | 0.681 0.137* | 0.000 0.000* | 8.5 (0.2037) | 6.0 |
| C02 | 1.196 0.162* | 0.823 0.163* | -1.453 0.215* | 0.635 0.126* | 0.000 0.000* | 7.7 (0.2580) | 6.0 |
| C03 | 1.028 0.160* | 0.922 0.169* | -1.115 0.153* | 0.678 0.124* | 0.000 0.000* | 5.8 (0.4441) | 6.0 |
| ... | | | | | | | |
| C38 | -0.962 0.164* | 1.098 0.182* | 0.876 0.115* | 0.739 0.122* | 0.000 0.000* | 6.7 (0.3520) | 6.0 |
| C39 | -1.144 0.173* | 0.879 0.170* | 1.302 0.169* | 0.660 0.128* | 0.000 0.000* | 1.7 (0.8927) | 5.0 |
| C40 | -1.044 0.133* | 0.632 0.123* | 1.652 0.268* | 0.534 0.104* | 0.000 0.000* | 3.0 (0.8143) | 6.0 |

```
LARGEST CHANGE = 0.008756
* STANDARD ERROR
171.9 233.0
(0.9990)
```

| PARAMETER | MEAN | STN DEV |
|------------|--------|---------|
| SLOPE | 0.862 | 0.154 |
| LOG(SLOPE) | -0.165 | 0.180 |
| THRESHOLD | -0.164 | 0.908 |

Phase 3 output

The only new feature in Phase 3 is the use of the YCOMMON option to place the information plots for the subtests on the same scale. This permits visual comparison of the relative precision of the subtests according to the heights of the information curves. To illustrate, the ICC for subtest 1, form is given below. The POP option also provides IRT estimated reliability for each subtest.



MAXIMUM INFORMATION APPROXIMATELY 0.9046D+01 AT 0.0000
 FOR A NORMAL POPULATION WITH MEAN 0.000 AND S.D. 1.000
 AVERAGE INFORMATION= 0.7232D+01 AND RELIABILITY INDEX= 0.879