Appendix: Software for Ordinal Data Modeling

An Introduction to MATLAB

MATLAB is an interactive high-level language for integrating computation, visualization, and programming. It was originally written to provide an easy access to matrix software developed by the LINPACK and EISPACK projects. The basic data element in MATLAB is a matrix. One performs calculations by entering calculator-type instructions in the Command Window. Alternatively, one can execute sets of commands by means of scripts called M-files and functions. MATLAB has extensive facilities for displaying vectors and matrices as graphs.

A toolbox of MATLAB version 5 functions written by the authors is available at the web site `http://www-math.bgsu.edu/~albert/ord_book/` This set of functions can be used together with the MATLAB software (available from The MathWorks, Inc., 24 Prime Park Way, Natick, MA 01760-1500) to perform all of the calculations and graphics that are illustrated in this book. Here, we outline the use of some of the MATLAB functions in this toolbox for fitting and criticizing binary, ordinal, multirater, and item response regression models.

Chapter 2 — Review of Bayesian computation

Defining the posterior density

The MATLAB functions listed below can be used to implement the different summarization methods for a posterior density with two parameters. To use these functions, a MATLAB function needs to be written which defines the logarithm of the posterior density. The following function `logpost2` computes the log of the posterior density for the log odds ratio and log odds product discussed in Section 2.2:

```
function val=logpost2(xy,data)
t1=xy(:,1);   t2=xy(:,2);
y1=data(1); n1=data(2); y2=data(3); n2=data(4);
g1=(t1+t2)/2; g2=(t2-t1)/2;
val=y1*g1-n1*log(1+exp(g1))+y2*g2-n2*log(1+exp(g2));
```

There are two inputs to this function. The matrix `xy` contains values of the parameter vector, where each row corresponds to a set of values for the two parameters, and the vector `data` contains values of data and prior parameters that are used in the definition of the posterior density. The output `val` is a vector of values of the logarithm of the density at the parameter values specified by `xy`.

The multivariate normal approximation

```
[mode,var]=laplace('logpost2',mode,numiter,data)
```

The function `laplace` implements the multivariate normal approximation described in Section 2.3. There are four inputs: `'logpost2'` is the name of the function which defines the log posterior, `mode` is a guess at the posterior mode, `numiter` is the number of iterations of the algorithm for finding the posterior mode, and `data` is a vector of prior parameters and data that is used by the function `'logpost2'`. The output is the posterior mode `mode` and the approximate variance-covariance matrix `var`.

Grid integration

```
[lint,mom]=ad_quad2('logpost2',mom,numiter,data)
```

The function `ad_quad2` implements the adaptive quadrature algorithm described in Section 2.3. One inputs the name of the function `'logpost2'` which defines the log posterior, a vector `mom` of guesses at the posterior moments of the distribution, the number of iterations `numiter` of the algorithm, and the vector `data` of data values used by `'logpost2'`. The output of the function is the logarithm of the integrand of the density `lint` and a vector `mom` of estimates at the posterior moments of the density.

Gibbs sampling

```
sim_sample=gibbs('logpost2',start,m,scale,data)
```

The function `gibbs` illustrates the Gibbs sampling algorithm, where a Metropolis-Hastings algorithm is used to sample from each full-conditional distribution. The inputs are the definition `'logpost2'` of the log posterior density, a vector `start` which contains the starting value of the algorithm, the number of iterations `m` of the algorithm, a vector `scale` of scale constants used in the MH algorithms, and the vector `data` of data values used in `'logpost2'`. The output is a matrix of simulated values from the posterior density, where each row corresponds to a different simulated value.

Chapter 3 — Regression models for binary data

Data setup

Data for a regression problem with a binary response are stored as a matrix. Each row of the matrix corresponds to a single experimental unit. The first column contains the proportion of successes, the second column contains the corresponding sample size, and the remaining columns of the matrix contain the set of regression covariates. For the statistics class dataset described in Chapter 3, each experimental unit corresponds to a binary response, and the MATLAB matrix `stat` which stores this data are given as follows:

```
stat=[
     0     1     1   525
     0     1     1   533
     1     1     1   545
     0     1     1   582
     1     1     1   581
     1     1     1   576
     1     1     1   572
     1     1     1   609
     1     1     1   559
     1     1     1   543
     1     1     1   576
     1     1     1   525
     1     1     1   574
     1     1     1   582
     1     1     1   574
     0     1     1   471
     1     1     1   595
     0     1     1   557
     0     1     1   557
     1     1     1   584
     1     1     1   599
     0     1     1   517
     1     1     1   649
     1     1     1   584
     0     1     1   463
     1     1     1   591
     0     1     1   488
     1     1     1   563
     1     1     1   553
     1     1     1   549];
```

Maximum likelihood estimation and model criticism

```
[beta,var,fitted_probs,dev_df,pearson_res,dev_res,adev_res]
     = breg_mle(stat,'l');
```

The MATLAB function `breg_mle` finds the maximum likelihood estimate of the vector of regression coefficients. The input to this function is the data matrix `stat` and the link function ('l' for logit, 'p' for probit, and 'c' for complementary log-log). As output, the vector `beta` contains the mle and the matrix `var` contains the associated variance-covariance matrix of the estimate. In addition, the function outputs the vector of fitted probabilities `fitted_probs`, a vector `dev_df` containing the deviance and associated degrees of freedom, the vector of Pearson residuals `pearson_df`, the vector of deviance residuals `dev_res`, and the vector of adjusted deviance residuals `adev_res`.

Bayesian fitting using a flat prior

```
[Mb,arate]=breg_bay(stat,m,'l');
```

The MATLAB function `breg_bay` takes a simulated sample from the posterior distribution of the regression coefficients (under a flat prior) using the Metropolis-Hastings algorithm. There are three inputs to this function: `stat` is the data matrix, `m` is the number of iterations of the simulation algorithm, and `'l'` indicates that a logistic link will be used. The output is the matrix `Mb`, where each row of the matrix corresponds to a simulated value from the posterior distribution, and the constant `arate` which is equal to the acceptance rate of the Metropolis algorithm.

Posterior fitted probabilities and residual distributions

Several functions are available for summarizing the simulated sample from the posterior distribution.

```
[fit_prob,residuals]=lfitted(Mb,stat,'l');
```

The function `lfitted` computes summaries of the posterior distributions of the fitted probabilities $\{\hat{p}_i\}$ and the residuals $\{y_i - \hat{p}_i\}$. One inputs the matrix of simulated values `Mb`, the data matrix `stat`, and the link function `'l'`. One output is the matrix `fit_prob`, where each row of the matrix corresponds to the 5th, 50th, and 95th percentiles of the posterior distribution of p_i. Similarly, a row of the output matrix `residuals` contains the 5th, 50th, and 95th percentiles of the posterior distribution of the Bayesian residual $y_i - p_i$.

```
plotfitted(stat(:,4),fit_prob,'SAT', 'Probability')
```

The function `plotfitted` is useful for graphing the summaries of the posterior distribution that are produced by the function `lfitted` as shown in Figure 3.5. The inputs to this function are the vector `stat(:,4)` that will be plotted on the horizontal axis, the summary matrix `fit_prob` of the posterior distribution of the fitted probabilities, and the strings 'SAT' and 'Probability' which are used to label the horizontal and vertical axes on the plot.

Computation of latent residuals

```
[log_scores,sz]=llatent(Mb,stat);
```

The function llatent computes the posterior means of the ordered latent residuals (assuming a logistic link) discussed in Section 3.4. This function also produces the logistic scores plot as shown in Figure 3.9. The input to this function is the matrix of simulated values from the posterior Mb and the data matrix stat. The output is the vector of logistic scores log_scores and the vector of posterior means of the ordered latent residuals sz.

Bayesian probit fitting with a flat prior using data augmentation

```
Mb=b_probg(stat,m);
```

The function b_probg implements the Gibbs sampling algorithm for fitting a Bayesian binary regression model with a probit link. There are two inputs to this function: stat, the data matrix, and m, the number of iterations of the simulation. The output matrix Mb contains the simulated sample from the posterior distribution, where the rows of the matrix correspond to the simulated variates.

Bayesian fitting using an informative prior

```
[Mb,arate]=breg_bay(stat,m,'l',prior);
```

The function breg_bay can also be used to simulate from the posterior distribution using a conditional means informative prior distribution. There are four inputs to this function: the data matrix stat, the number of iterations of the algorithm m, the string 'l' indicating a logistic link, and a matrix prior which contains the parameters of the conditional means prior. The output of this function is the matrix Mb of simulated values from the posterior and the acceptance rate arate.

In the example of Section 3.2.4, the prior information was that a student with covariate vector [1 500] would pass with probability .3, a student with covariate vector [1 600] would pass with probability .7, and each statement was worth five observations. This prior information is inputted by means of the matrix

```
prior=[.3 5 1 500
       .7 5 1 600];
```

Computation of a marginal likelihood

```
[beta,var,lmarg]=cmp(stat,prior);
```

The function cmp computes the marginal likelihood used in computing the Bayes factor discussed in Section 3.4. The input of this function is the data matrix stat and a matrix prior, which contains the parameters of a conditional means prior distribution. (The specification of the matrix prior is illustrated above.) The function has three outputs: the posterior mode beta, the associated posterior variance-covariance matrix var, and the natural logarithm of the value of the marginal likelihood lmarg.

Computation of a posterior predictive distribution

```
p_std=post_pred(Mb,conduct);
```

The function post_pred obtains a simulated sample from the posterior predictive distribution of the standard deviation of the counts $\{y_i^*\}$ as discussed in Section 3.5. The input to this function is the matrix of simulated values Mb from the posterior distribution of the regression vector and the dataset matrix conduct. The output vector p_std represents a simulated sample from the posterior predictive distribution of the standard deviation of $\{y_1^*, ..., y_n^*\}$.

Fit of a random effects model

```
[Mbeta,Ms2,p_std2]=logit_re2(conduct,m,ab);
```

The function logit_re2 fits the Bayesian random effects model discussed in Section 3.5. The inputs are the dataset conduct, the size of the simulated sample m, and the vector ab of parameter values for the inverse-gamma distribution placed on the random effects variance. The function returns the matrix Mbeta of simulated values from the marginal posterior distribution of the regression vector β, the vector Ms2 of simulated values from the marginal posterior distribution of the random effects variance σ^2, and the vector p_std of simulated values from the posterior-predictive distribution of the standard deviation of $\{y_1^*, ..., y_n^*\}$.

Chapter 4 — Regression models for ordinal data

Data setup

The data input for a ordinal response regression problem are a single matrix. Each row of the matrix corresponds to a single experimental unit. The first column contains the ordinal responses and the remaining columns contain the set of regression covariates. For the statistics class dataset described in Chapter 4, the MATLAB matrix ostat which stores this data are given as follows:

```
ostat=[
    2   1   525
    2   1   533
```

```
4   1   545
2   1   582
3   1   581
...
4   1   563
4   1   553
5   1   549]
```

Maximum likelihood estimation

```
[mle,cov,dev,devRes,fits] = ordinalMLE(ostat,C,link);
```

The function `ordinalMLE` finds the maximum likelihood estimate of the unknown category cutoffs and the regression coefficients. The inputs to this function are the data matrix `ostat`, the number of categories `C`, and a string variable `link` indicating the link function. The function outputs the maximum likelihood estimate `mle`, the associated variance-covariance matrix `cov`, the deviance statistic `dev`, the vector `devRes` containing the signed deviance contributions, and a matrix of fitted probabilities `fits`.

Bayesian fitting using flat or informative priors

```
[sampleBeta, meanZ, accept]
            = sampleOrdProb(ostat,C,mle,m,prior);
```

The function `sampleOrdProb` takes a simulated sample from the posterior distribution of the regression coefficients and category cutoffs using the MCMC algorithm described in Section 4.3.2. The inputs to this function are the data matrix `ostat`, the number of categories `C`, the vector `mle` containing the maximum likelihood estimates (found using the function `ordinalMLE`), and the number of simulated values `m`. The default is a flat prior for the unknown parameters. If an informative prior is used, `prior` is the name of a function that returns a value of the prior density at the parameter vector. The output is a matrix `sampleBeta` of simulated values where the rows correspond to the `m` samples and the columns correspond to the components of `mle`. In addition, the vector `meanZ` contains the means of the sample latent variables by observation and `accept` contains the MH acceptance rate for the category cutoffs.

Chapter 5 — Analyzing data from multiple raters

Data setup

The data input for analyzing multiple rater data consist of two matrices. The matrix `N` contains the observed ordinal responses, where the rows correspond to the "items" and the columns to the "raters." When applicable, the matrix `X` contains the design matrix for the regression model, where the rows correspond to the items and the columns to the regression parameters.

Bayesian fitting of a multiple rater model

```
[Zij, Z, S, Cats, accept]
            = sampleMulti(N,sampSize,alpha,lambda);
```

The function `sampleMulti` obtains a simulated sample from the posterior distribution of the multirater model described in Section 5.2. The inputs are the data matrix `N`, the number of iterates of the simulation `sampSize`, and the parameters `alpha` and `lambda` of the inverse-gamma distribution placed on the rater variances. The output is the matrix of simulated values `Zij` of the rater-item values $\{t_{ij}\}$, the matrix `Z` of simulated values of the latent traits Z, the matrix `S` of simulated variates from the rater variances, the matrix `Cats` of simulated values from the category cutoffs γ, and the acceptance rate `accept` of the MH algorithm used to sample the category cutoffs.

Bayesian fitting of a multiple rater model (regression case)

```
[Zij,Z,S,Cats,B,Sr,accept]
            = sampReg(N,X,bZ,sampSize,alpha,lambda);
```

The function `sampReg` obtains a simulated sample from the posterior distribution of the multirater model with regression described in Section 5.3. The inputs to this function include the data matrix `N`, the number of iterates `sampSize`, and the hyperparameters `alpha` and `lambda` as described for the previous function. In addition, this function requires the input of the design matrix `X` and the vector `bZ`, which contains the estimates of the mean latent traits. The output is `Zij`, `Z`, `S`, `Cats`, and `accept`, as described earlier; also, the function outputs `B`, a matrix containing simulated values of the posterior of the regression parameter, and `Sr`, a vector containing simulated values of the posterior of the regression variance.

Bayesian fitting of a ROC analysis

```
[Zij,Z,S,Cats,m0,m1,v0,v1]
            = roc(N,D,TmT,sampSize,alpha,lambda);
```

The function `roc` obtains a simulated sample from the posterior distribution for the ROC analysis described in Section 5.4. The input to this function is the data matrix `N`, the vector `D`, which indicates the patients' disease status (1–yes, 0–no), the vector `TmT` which indicates the treatments for all patients (0's or 1's), the number of MCMC iterates `sampSize`, and the prior parameters on the rater parameters `alpha` and `lambda`. The output is the matrix `Zij` containing the simulated rater-item values from their posterior distribution, the matrix `Z` of simulated values of the latent item traits, the matrix `S` of simulated values from the posterior on the rater variances, and the matrix `Cats` of sampled values from the category cutoffs. Also, `m0` and `m1` are vectors containing sampled values of the disease means, `v0` and `v1` are vectors of the sampled disease variances, and `accept` is the acceptance probability for the sampling algorithm on the category cutoffs.

Chapter 6 - Item response modeling

Data setup

The data are stored in the matrix form described in Section 6.3.1. The matrix has n rows and k columns, where the rows correspond to the individuals that are judged and the columns correspond to the judges or items. The entries of the matrix are 0's and 1's, where (in the exam example) 0 indicates an incorrect response and 1 indicates a correct response. In the following, the matrix `ratings` contains the ratings of the student judges on the shyness of their peers.

Bayesian fit of two-parameter probit item response model

```
[av,bv,th_m,th_s] = item_r(ratings,m_a,s_a,s_b,m);
```

The function `item_r` takes a simulated sample of the posterior distribution of the item and ability parameters of the two-parameter item response model with probit link described in Section 6.5. The input to the function is the data matrix `ratings`, the mean `m_a` and standard deviation `s_a` of the item slope parameters, the standard deviation `s_b` of the item intercept parameters, and the number of iterations `m` of the Gibbs sampling algorithm. The function returns the matrix `av` of simulated values from the marginal posterior distribution of the item slope parameters, where each row of the matrix corresponds to a single sampled vector. In addition, `bv` is the matrix of simulated values of the item intercept parameters, `th_m` is a vector containing the posterior means of the ability parameters, and `th_s` contains the respective posterior standard deviations of the ability parameters.

Posterior estimates of the item response curve

```
[pr,lo,hi] = irtpost(av,bv,theta);
```

The function `irtpost` computes summaries of the posterior densities of all of the item response curves. The inputs to the function are the matrices `av` and `bv` of simulated values from the posterior distributions of the item slope and intercept parameters, and a vector `theta` of values of the ability parameter. The function outputs a matrix `pr` of posterior medians of the probability of correct response, where the rows of the matrix correspond to the values of the ability parameter and the columns correspond to the different items. Also, the matrices `lo` and `hi` contain the respective 5th and 95th percentiles of the posterior distribution of the probability of a correct response.

Bayesian fit of one-parameter probit item response model

```
[bv,th_m,th_s] = item_r1(ratings,s_b,m);
```

The function `item_r1` takes a simulated sample of the posterior distribution of the item and ability parameters of the one-parameter item response model with probit link described in Section 6.8. The input to the function is the data matrix `ratings`, the standard deviation `s_b` of the item intercept parameters, and the number of iterations `m` of the Gibbs sampling algorithm. The function returns the matrix `bv` of simulated values of the item intercept parameters, the vector `th_m`, containing the posterior means of the ability parameters, and the vector `th_s`, containing the respective posterior standard deviations of the ability parameters.

Bayesian fit of two-parameter probit item response model with an exchangeable prior

```
[av,bv,th_m,th_s,av_m,av_s2] = item_r_h(ratings,ab,m);
```

The function `item_r_h` takes a simulated sample of the posterior distribution of the item and ability parameters of the two-parameter item response model with an exchangeable prior discussed in Section 6.10. The input to the function is the data matrix `ratings`, the vector `ab` of parameters of the inverse-gamma prior for the variance τ^2, and the number of iterations `m` of the Gibbs sampling algorithm. The function returns the matrix `av` of simulated values from the marginal posterior distribution of the item slope parameters, the matrix `bv` of simulated values of the item intercept parameters, the vector `th_m`, containing the posterior means of the ability parameters, and the vector `th_s`, containing the respective posterior standard deviations of the ability parameters.

References

Abramowitz, M. and Stegun, I.A., editors (1972), *Handbook of Mathematical Functions*, New York: Dover Publications Inc.

Albert, J.H. (1992), "Bayesian Estimation of Normal Ogive Item Response Curves Using Gibbs Sampling," *Journal of Educational Statistics*, **17**, 261-269.

Albert, J.H. (1996), *Bayesian Computation Using Minitab*, Belmont, CA: Duxbury Press.

Albert, J.H. and Chib, S. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, **88**, 669-679.

Albert, J.H. and Chib, S. (1995), "Bayesian Residual Analysis for Binary Response Regression Models," *Biometrika*, **82**, 747-759.

Antleman, G. (1997), *Elementary Bayesian Statistics* Cheltenham, Hong Kong: Edward Elgar Publishing.

Baker, F. B. (1992), *Item Response Theory: Parameter Estimation Techniques*, New York: Marcel Dekker.

Bedrick, E.J., Christensen, R., and Johnson, W. (1996), "A New Perspective on Priors for Generalized Linear Models," *Journal of the American Statistical Association*, **91**, 1450-1460.

Bernardo, J.M. and Smith, A.F.M. (1994), *Bayesian Theory*, New York: John Wiley & Sons.

Berry, D. (1995), *Statistics: A Bayesian Perspective*, Belmont, CA: Duxbury Press.

Bock, R.D. and Aitken, M. (1981), "Marginal Maximum Likelihood Estimation of Item Parameters: Application of An EM Algorithm," *Psychometrika*, **46**, 443-459.

Bock, R.D. and Lieberman, M. (1970), "Fitting a Response Model for n Dichotomously Scored Items," *Psychometrika*, **35**, 179-197.

Box, G.E.P. and Tiao, G.C (1973), *Bayesian Inference in Statistical Analysis*, Reading, MA: Addison-Wesley.

Bradlow, E.T. and Zaslavsky, A. M. (1999), "Analysis of Ordinal Survey Data with No Answer," to appear in the *Journal of the American Statistical Association*.

Bradlow, E.T., Wainer, H., and Wang, X. (1999), "A Bayesian Random Effects Model for Testlets," *Psychometrika*, to appear.

Casella, G. and George, E. (1992), "Explaining the Gibbs Sampler," *American Statistician*, **46**, 167-174.

Caulkin, J., Larkey, P., and Wei, J. (1996), "Adjusting GPA to Reflect Course Difficulty," Working paper, Heinz School of Public Policy and Management, Carnegie Mellon University.

Chaloner, K. (1991), "Bayesian Residual Analysis in the Presence of Censoring," *Biometrika,* **78**, 637-644.

Chaloner, K. and Brant, R. (1988), "A Bayesian Approach to Outlier Detection and Residual Analysis," *Biometrika,* **75**, 651-659.

Chib, S. (1995), "Marginal Likelihood from Gibbs Output," *Journal of the American Statistical Association*, **90**, 1313-1321.

Chib, S. and Greenberg, E. (1995), "Understanding the Metropolis-Hastings Algorithm," *American Statistician*, **49**, 327-335.

Cohen, J. (1960), "A Coefficient of Agreement for Nominal Tables," *Educational and Psychological Measurement*, **20**, 37-46.

Collett, D. (1991), *Modelling Binary Data,* London: Chapman and Hall.

Cowles, M.K. (1996), "Accelerating Monte Carlo Markov Chain Convergence for Cumulative-link Generalized Linear Models," *Statistics and Computing*, **6**, 101-111.

Cowles, M.K., Carlin, B.P., and Connett, J.E. (1996), "Bayesian Tobit Modeling of Longitudinal Ordinal Clinical Trial Compliance Data With Nonignorable Missingness," *Journal of the American Statistical Association*, **91**, 86-98.

Cox, D.R. (1972), "Regression Models and Life-tables" (with discussion), *Journal of the Royal Statistical Society, series B*, **34**, 187-220.

Darwin, C. (1876), *The Effect of Cross- and Self-fertilization in the Vegetable Kingdom*, 2nd edition, London: John Murray.

Dellaportas, P. and Smith, A.F.M. (1993), "Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling," *Applied Statistics,* **42,** 443-459

DiCiccio, T.J., Kass, R.E., Raftery, A.E., and Wasserman, L. (1997), "Computing Bayes Factors by Combining Simulation and Asymptotic Approximations," *Journal of the American Statistical Association*, **92**, 903-915.

Dobson, A. (1990), *An Introduction to Generalized Linear Models,* London: Chapman and Hall.

Dorn, H.F. (1954), "The Relationship of Cancer of the Lung and the Use of Tobacco," *The American Statistician*, **8**, 7-13.

Elliott, R. and Strenta A. (1988), "Effects of Improving the Reliability of the GPA on Prediction Generally and on Comparative Predictions for Gender and Race Particularly," *Journal of Educational Measurement*, **25**, 333-347.

Fahrmeir, L. and Tutz, G. (1994), *Multivariate Statistical Modeling Based on Generalized Linear Models*, New York: Springer-Verlag.

Fleiss, J.L. (1971), "Measuring Nominal Scale Agreement Among Many Raters," *Psychological Bulletin*, **88**, 322-328.

Gelfand, A.E. (1996), "Model Determination Using Sampling-based Methods," in *Markov Chain Monte Carlo in Practice*, editors Gilks, Richardson, and Spiegelhalter, London: Chapman & Hall, 145-161.

Gelfand, A.E. and Smith, A.F.M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, **85**, 398-409.

Gelfand, A.E., Hills, S.E., Racine-Poon, A., and Smith, A.F.M. (1990), "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association*, **85**, 972-985.

Gelman, A. and Meng, X.L. (1998), "Simulating Normalizing Constants: From Importance Sampling to Bridge Sampling to Path Sampling," *Statistical Science*, **13**, 163-185.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (1995), *Bayesian Data Analysis*, London: Chapman & Hall.

Gilchrist, W. (1984), *Statistical Modeling,* Chichester: John Wiley and Sons.

Gilks, W., Richardson, S., and Speigelhalter, D.J., editors, (1997), *Markov Chain Monte Carlo in Practice*, London: Chapman & Hall.

Goldman, R. and Widawski, M. (1976), "A Within-Subjects Technique for Comparing College Grading Standards: Implications in the Validity of the Evaluation of College Achievement," *Educational and Psychological Measurement*, **36**, 381-390.

Goldman, R., Schmidt, D., Hewitt, B. and Fisher, R. (1974), "Grading Practices in Different Major Fields," *American Education Research Journal*,**11**, 343-357.

Grayson, D.K. (1990), "Donner Party Deaths: A Demographic Assessment," *Journal of Anthropological Assessment*, **46**, 223-242.

Hambleton, R. and Swaminathan, H. (1985), *Item Response Theory: Principles and Applications*, Boston: Kluwer.

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J., and Ostrowski, E. (1994), *A Handbook of Small Data Sets*, London: Chapman & Hall.

Hastings, W.K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, **57**, 97-109.

Jansen, J. (1991), "Fitting Regression Models to Ordinal Data," *Biometrical Journal*, **33**, 807-815.

Johnson, V.E. (1996), "On Bayesian Analysis of Multirater Ordinal Data," *Journal of the American Statistical Association*, **91**, 42-51.

Johnson, V.E. (1997), "An Alternative to Traditional GPA for Evaluating Student Performance," *Statistical Science*, **12**, 251-278.

Johnson, V.E. (1998), "Posterior Distributions on Normalizing Constants," ISDS Discussion Paper 98-26, Duke University.

Jolayemi, E.T. (1990a), "On the Measurement of Agreement Between Two Raters," *Biometrics Journal*, **32**, 87-93.

Jolayemi, E.T. (1990b), “A Multiraters Agreement Index for Ordinal Classification,” *Biometrics Journal*, **33**, 485-492.

Kass, R.E. and Raftery, A.E. (1995), “Bayes Factors,” *Journal of the American Statistical Association*, **90**, 773-795.

Landis, J.R. and Koch, G.G. (1977a), “The Measurement of Observer Agreement for Categorical Data,” *Biometrics Journal*, 33, 159-174.

Landis, J.R. and Koch, G.G. (1977b), “An Application of Hierarchical Kappa-Type Statistics in the Assessment of Majority Agreement Among Multiple Observers,” *Biometrics Journal*, **33**, 363-374.

Larkey, P. and Caulkin J. (1992), “Incentives to Fail,” Working paper 92-51, Heinz School, Carnegie Mellon University.

Lee, P. (1989), *Bayesian Statistics: An Introduction*, New York: Oxford University Press.

Li, J., Jaszczak, R.J., Turkington, T.G., Metz, C.E., Musante, D.B., and Coleman, R.E. (1992), “ROC evaluation of Lesion Detectability of Cone Beam versus Parallel Beam Collimation in SPECT with a 3D Brain Phantom,” *Radiology*, **185**, 251.

Light, R.J. (1971), “Measures of Response Agreement for Qualitative Data: Some Generalizations and Alternatives,” *Psychological Bulletin*, **5**, 365-377.

Lord, F.M. and Novick, M.R. (1968), *Statistical Theories of Mental Test Scores*, Reading, MA: Addison-Wesley.

Manly, B.F.J. (1991), *Randomization and Monte Carlo Methods in Biology,* London: Chapman & Hall.

Masters, G.N. (1982), “A Rasch Model for Partial Credit Scoring,” *Psychometrika*, 149-174.

McCullagh, P. (1980), “Regression Models for Ordinal Data,” *Journal of the Royal Statistical Society, series B*, **42**, 109-142.

McCullagh, P. and Nelder, J.A., (1989), *Generalized Linear Models*, 2nd edition, London: Chapman & Hall.

McNeil, B.J. and Hanley, J.A. (1984), “Statistical Approaches to the Analysis of Receiver Operator Characteristic (ROC) curves,” *Medical Decision Making*, **4**, 137-150.

Mellenberg, G.J. (1994), “Generalized Linear Item Response Theory,” *Psychological Bulletin*, **115**, 300-307.

Meng, X.L. and Wong, W.H. (1996), “Simulating Ratios of Normalizing Constants Via a Simple Identity: A Theoretical Exploration,” *Statistica Sinica*, **4**, 831-860.

Metz, C.E. (1986), “ROC Methodology in Radiologic Imaging,” *Investigative Radiology*, **21**, 720-733.

Nandram, B. and Chen, M.H. (1996), “Reparameterizing the Generalized Linear Model to Accelerate Gibbs Sampler Convergence,” *Journal of Statistical Computation and Simulation*, **54**, 129-144.

Nelder, J.A. and Wedderburn, R.W.N. (1972), “Generalized Linear Models,” *Journal of the Royal Statistical Society, series A*, **135**, 370-384.

Page, E. (1994), "New Computer Grading of Student Prose, Using Modern Concepts and Software," *Journal of Experimental Education* **62(2)**, 127-142.

Patz, R.J. and Junker, B.W. (1997), " A Straightforward Approach to Markov Chain Monte Carlo Methods for Item Response Models," unpublished manuscript.

Pierce, D.A. and Schafer, D.W. (1986), "Residuals in Generalized Linear Models," *Journal of the American Statistical Association*, **81**, 977-986.

Ramsey F.L. and Schafer, D.W. (1997), *The Statistical Sleuth*, Belmont CA: Duxbury Press.

Rao, C.R. (1973), *Linear Statistical Inference and Applications*, New York: John Wiley & Sons.

Robert, C.P. (1994), *The Bayesian Choice: A Decision-Theoretic Motivation*, New York: Springer-Verlag.

Sahu, S.K. (1997), "Bayesian Estimation and Model Choice in the Item Response Models," Technical report, School of Mathematics, University of Wales, Cardiff, UK.

Samejima, F. (1969), *Estimation of Latent Ability Using a Pattern of Graded Scores,* Psychometrika, Monograph Supplement No. 17.

Samejima, F. (1972), "A General Model for Free Response Data," *Psychometrika*, **38**, 221-233.

Silvapulle, M.J. (1981), "On the Existence of Maximum Likelihood Estimators for the Binomial Response Models," *Journal of the Royal Statistical Society, series B*, **43**, 310-313.

Stout, W.F. (1990), "A New Item Response Theory Modeling Approach With Applications to Unidimensionality Assessment and Ability Estimation," *Psychometrika*, **55**, 293-325.

Strenta, A. and Elliott, R. (1987), "Differential Grading Standards Revisited," *Journal of Educational Measurement*, **24**, 281-291.

Swaminathan, H. and Gifford, J.A. (1982), "Bayesian Estimation in the Rasch Model," *Journal of Educational Statistics*, **7**, 175-192.

Swaminathan, H. and Gifford, J.A. (1985), "Bayesian Estimation in the Two-Parameter Logistic Model," *Psychometrika*, **50**, 349-364.

Swets, J.A. (1979), "ROC Analysis Applied to the Evaluation of Medical Imaging Techniques," *Investigative Radiology*, **14**, 109-121.

Tanner, M., and Wong, W.H. (1987), "The Calculation of Posterior Distributions by Data Augmentation" (with discussion), *Journal of the American Statistical Association,* **82**, 528-550.

Tanner, M. and Young, M. (1985), "Modeling Agreement Among Raters," *Journal of the American Statistical Association*, **80**, 175-180.

Tierney, L. (1994), "Markov Chains for Exploring Posterior Distributions," *Annals of Statistics*, **22**, 1701-1762.

Thissen, D. and Steinberg, L. (1984), "Taxonomy of Item Response Models," *Psychometrika*, **51**, 567-578.

Tsutakawa, R.K. and Lin, H.Y. (1986), "Bayesian Estimation of Item Response Curves," *Psychometrika*, **51**, 251-267.

Tsutakawa, R.K., Shoop, G.L., and Marienfeld, C.J. (1985), “Empirical Bayes Estimation of Cancer Mortality Rates,” *Statistics in Medicine*, **4**, 201-212.

Uebersax, J. (1992), “A Review of Modeling Approaches for the Analysis of Observer Agreement,” *Investigative Radiology*, **17**, 738-743.

Uebersax, J. (1993), “Statistical Modeling of Expert Ratings on Medical Treatment Appropriateness,” *Journal of the American Statistical Association*, **88**, 421-427.

Uebersax, J. and Grove, W.M. (1993), “A Latent Trait Finite Mixture Model for the Analysis of Rating Agreement,” *Biometrics*, **49**, 823-835.

van der Linden, W.J. and Hambleton, R.K., editors (1997), *Handbook of Modern Item Response Theory*, New York: Springer-Verlag.

Walser, P. (1969), “Untersuchung uber die Verteilung der Geburtstermine bei der mehrgebarenden Frau,” *Helvetica Paediatrica Acta*, Suppl. XX (3), **24**, 1-30.

Young, J.W. (1989), “Developing a Universal Scale for Grades: Investigating Predictive Validity in College Admissions,” PhD Thesis, Stanford University, Palo Alto, CA.

Young, J.W. (1990), “Adjusting the Cumulative GPA Using Item Response Theory,” *Journal of Educational Measurement*, **12**, 175-186.

Young, J.W. (1992), “A General Linear Model Approach to Adjusting Cumulative GPA,” *Journal of Research in Education*, **2**, 31-37.

Young, J.W. (1993), “Grade Adjustment Methods,” *Review of Educational Research*, **63**, 151-165.

Index