Supplement to "Detecting chance: A solution to the null sensitivity problem in subliminal priming"

Richard D. Morey, Paul L. Speckman,
Jeffery N. Rouder, and Michael S. Pratte

April 5, 2007

This document is a supplement to Rouder, Morey, Speckman, and Pratte's "Detecting chance: A solution to the null sensitivity problem in subliminal priming." We present model development and computer code for the mass-at-chance (MAC) model. The model is specified and analyzed in the Bayesian framework; Rouder and Lu (2005) provide a tutorial on Bayesian hierarchical analysis for psychologists.

1 Data Representation

In the paradigm, each of I participants observes N_i trials ($i = 1, \ldots, I$). Let y_{ij} denote the outcome of the jth trial ($j = 1, \ldots, N_i$) for the ith participant. If the outcome is a correct response, then $y_{ij} = 1$; otherwise, $y_{ij} = 0$.

2 Model Specification

Outcomes are modeled as Bernoulli events:

$$y_{ij}|p_i \overset{indep}{\sim} \text{Bernoulli}(p_i). \tag{1}$$

Parameter p_i is a function of latent ability x_i:

$$p_i = \Phi(x_i \vee 0),$$

where $a \vee b = \max(a, b)$ and Φ is the cumulative distribution function of the standard normal distribution.

Individuals' latent abilities are assumed to be normally distributed:

$$x_i|\mu, \sigma^2 \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2).$$

3 Priors

Priors are needed for parameters μ and σ^2. The choice of a normal and a truncated inverse gamma for mu and σ^2 is convenient and flexible:

$$\begin{aligned} \mu &\sim \text{Normal}(\mu_0, \sigma_0^2), \\ \sigma^2 &\sim \text{TIG}_{t_{\sigma^2}}(a_0, b_0), \end{aligned}$$

where $\text{TIG}_{t_{\sigma^2}}$ denotes the truncated inverse gamma[1] distribution with pdf proportional to

$$f(x|a_0, b_0) \quad \propto \quad x^{-(a_0+1)} \exp\{-\frac{b_0}{x}\} I_{(0<x<t_{\sigma^2})}. \tag{2}$$

Before analysis, values of μ_0, σ_0^2, a_0, b_0, and t_{σ^2} must be specified. Unfortunately, the choice of values for these parameters of the prior affects selection of participants. We have experimented with a number of choices and recommend $(\mu_0 = 0, \sigma_0^2 = 1, a_0 = -.5, b_0 = 0, t_{\sigma^2} = 1)$ as reasonable. This recommendation introduces a subtle bias against selection. This bias is appropriate as it is important to control the rate of selecting above-chance-pperforming individuals. With these prior parameters the prior on σ is flat over the interval (0,1).

- *Proof.* If $a_0 = -.5$, $b_0 = 0$, and $t_{\sigma^2} = 1$, then

  $$\sigma^2 \quad \propto \quad (\sigma^2)^{-.5} I_{(0<\sigma^2<1)}$$

  by (2). Transforming variables,

  $$\begin{aligned} \sigma &\propto (\sigma^2)^{-.5} \left(\frac{d}{d\sigma}\sigma^2\right) I_{(\sqrt{0}<\sigma<\sqrt{1})} \\ &\propto \sigma^{-1}\sigma I_{(0<\sigma<1)} \\ &\propto I_{(0<\sigma<1)} \end{aligned}$$

  This is the density function for a uniform random variable over (0,1).

When selecting prior parameters, it is important to choose parameters which yield proper posterior distributions. One way to do this is to ensure that that the priors are proper; i.e., for σ^2, a_0, b_0, and t_{σ^2} are chosen in such a way that

$$0 < \int_0^{t_{\sigma^2}} f(\sigma^2|a_0, b_0) d\sigma^2 < +\infty.$$

This will guarantee the propriety of the posterior distributions. In this case, choosing $a_0 > 0$ and $b_0 > 0$ is sufficient to ensure that the prior is proper. For other values of a_0 and b_0, the prior is not guaranteed to be proper.

[1] When $t_{\sigma^2} = +\infty$, $a_0 > 0$, and $b_0 > 0$ the truncated inverse gamma distribution reduces to the inverse gamma distribution.

4 Overview of Analysis

The basic steps in analysis are to specify the joint posterior, then integrate the joint posterior via Markov chain Monte Carlo (MCMC) sampling to obtain marginal posterior distributions. We describe these steps in turn. Preceding this, however, we define a set of additional parameters that are useful in deriving posterior distributions.

5 Additional Latent Parameters

We adapt the formulation of Albert & Chib (1995) and define the following latent parameters to aid analysis of probit models:

$$w_{ij} \stackrel{indep}{\sim} \text{Normal}(x_i \vee 0, 1).$$

If y_{ij} is defined as

$$y_{ij} = \begin{cases} 0 & \text{if and only if } w_{ij} < 0, \\ 1 & \text{if and only if } w_{ij} \geq 0, \end{cases}$$

then $Pr(w_{ij} \geq 0) = p_i$, and (1) holds.

6 Joint Posterior

Let $\mathbf{Y}$ be the collection of data, $\mathbf{Y} = \left\langle \langle y_{ij} \rangle_{j=1}^{j=N_i} \right\rangle_{i=1}^{i=I}$, and $\mathbf{W}$ be the collection of latent parameters, $\mathbf{W} = \left\langle \langle w_{ij} \rangle_{j=1}^{j=N_i} \right\rangle_{i=1}^{i=I}$. The joint posterior of all parameters is proportional to

$$\begin{aligned}
[\mathbf{W}, \mathbf{x}, \mu, \sigma^2 | \mathbf{Y}] \quad \propto \quad & \prod_i \left\{ \prod_{\substack{j=1 \\ (y_{ij}=0)}}^{N_i} \exp\left\{ -\frac{(w_{ij} - (0 \vee x_i))^2}{2} \right\} I_{(w_{ij}<0)} \right. \\
& \left. \times \prod_{\substack{j=1 \\ (y_{ij}=1)}}^{N_i} \exp\left\{ -\frac{(w_{ij} - (0 \vee x_i))^2}{2} \right\} I_{(w_{ij}>0)} \right\} \\
& \times \prod_i \sigma^{-1} \exp\left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\
& \times \exp\left\{ -\frac{(\mu - \mu_0)^2}{2\sigma_0^2} \right\} (\sigma^2)^{-(a_0+1)} \exp\left\{ -\frac{b_0}{\sigma^2} \right\}. \quad (3)
\end{aligned}$$

7 Conditional Posteriors

In order to obtain the marginal posterior distributions, we use Gibbs sampling (Gelfand & Smith, 1990). As input, Gibbs sampling requires conditional posterior distributions. These conditional posterior distributions are provided by Facts A through Fact D, below. Proofs of Facts A, B, and C are standard and may be found in Gelman, Carlin, Stern, and Rubin (2004). The proof of Fact D is provided here.

- *Fact A. Conditional posterior of* $w_{ij}|x_i, y_{ij}$: Let $\text{TN}_{t+}(\mu, \sigma^2)$ denote the $\text{Normal}(\mu, \sigma^2)$ distribution truncated below at t and $\text{TN}_{t-}(\mu, \sigma^2)$ denote the $\text{Normal}(\mu, \sigma^2)$ truncated above at t, respectively. The pdfs for these distributions are

$$f_{t+}(x|\mu, \sigma^2) = \frac{I_{(x>t)}}{\Phi(\frac{\mu-t}{\sigma})\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad (4)$$

$$f_{t-}(x|\mu, \sigma^2) = \frac{I_{(x<t)}}{\Phi(\frac{t-\mu}{\sigma})\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}. \quad (5)$$

  The conditional posterior for $w_{ij}|x_i, y_{ij}$ is

$$w_{ij}|x_i; (y_{ij} = 1) \sim \text{TN}_{0+}(x_i \vee 0, 1)$$
$$w_{ij}|x_i; (y_{ij} = 0) \sim \text{TN}_{0-}(x_i \vee 0, 1).$$

- *Fact B. Conditional posterior of* $\mu|\mathbf{x}, \sigma^2$: Let

$$\tau^2 = \left(\frac{I}{\sigma^2} + \frac{1}{\sigma_0^2}\right)^{-1}$$
$$\eta = \tau^2 \left(\frac{\sum_i x_i}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right).$$

  Conditional posterior for $\mu|\mathbf{x}, \sigma^2$ is

$$\mu|\mathbf{x}, \sigma^2 \sim \text{Normal}(\eta, \tau^2).$$

- *Fact C. Conditional posterior of* $\sigma^2|\mathbf{x}, \mu$:

$$\sigma^2|\mathbf{x}, \mu \sim \text{TIG}_{t_{\sigma^2}}\left(a_0 + \frac{I}{2}, b_0 + \frac{1}{2}\left(\sum_i (x_i - \mu)^2\right)\right).$$

- *Fact D: Conditional posterior of* $x_i|\mathbf{W}, \mu, \sigma^2$: Let

$$\lambda_i^2 = \left(N_i + \frac{1}{\sigma^2}\right)^{-1}$$
$$\nu_i = \lambda_i^2 \left(\sum_j^{N_i} w_{ij} + \frac{\mu}{\sigma^2}\right)$$
$$\rho_i = \frac{\sigma\Phi(-\frac{\mu}{\sigma})}{\sigma\Phi(-\frac{\mu}{\sigma}) + \lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\{-\frac{1}{2}(\frac{\mu^2}{\sigma^2} - \frac{\nu_i^2}{\lambda_i^2})\}}.$$

Conditional posterior pdf of $x_i|\mathbf{W},\mu,\sigma^2$ is

$$[x_i|\mathbf{W},\mu,\sigma^2] \quad = \quad \rho_i f_{0-}(x_i|\mu,\sigma^2) + (1-\rho_i) f_{0+}(x_i|\nu_i,\lambda_i^2), \qquad (6)$$

where f_{0+} and f_{0-} are the pdfs of normal distributions truncated below and above at 0, respectively, and are defined in (4) and (5).

Proof of Fact D: Inspection of (3) reveals that the conditional posterior density of $x_i|\mathbf{W},\mu,\sigma^2$ satisfies

$$\begin{aligned}[x_i|\mathbf{W},\mu,\sigma^2] \quad \propto \quad & \exp\left\{-\frac{1}{2}\sum_{j=1}^{N_i}(w_{ij}-(x_i\vee 0))^2\right\} \\ & \times \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}.\end{aligned}$$

This right-hand side may be simplified as

$$\begin{aligned}[x_i|\mathbf{W},\mu,\sigma^2] \quad \propto \quad & \Bigg[I_{(x_i<0)}\exp\left\{-\frac{1}{2}\sum_{j=1}^{N_i}w_{ij}^2\right\} \\ & + I_{(x_i>0)}\exp\left\{-\frac{1}{2}\sum_{j=1}^{N_i}(w_{ij}-x_i)^2\right\}\Bigg] \\ & \times \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}.\end{aligned}$$

Distributing and completing the square yields

$$\begin{aligned}[x_i|\mathbf{W},\mu,\sigma^2] \quad \propto \quad & I_{(x_i<0)}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\} \\ & + I_{(x_i>0)}\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\} \\ & \times \exp\left\{-\frac{(x_i-\nu_i)^2}{2\lambda_i^2}\right\}.\end{aligned}$$

Multiplying by constants equal to 1 yields

$$\begin{aligned}[x_i|\mathbf{W},\mu,\sigma^2] \quad \propto \quad & \left[\left(\frac{\sqrt{2\pi\sigma^2}\Phi(-\frac{\mu}{\sigma})}{\sqrt{2\pi\sigma^2}\Phi(-\frac{\mu}{\sigma}}\right) I_{(x_i<0)}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}\right] \\ & + \left[\left(\frac{\sqrt{2\pi\lambda_i^2}\Phi(-\frac{\nu_i}{\lambda_i})}{\sqrt{2\pi\lambda_i^2}\Phi(-\frac{\nu_i}{\lambda_i})}\right) I_{(x_i>0)}\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}\right. \\ & \left.\times \exp\left\{-\frac{(x_i-\nu_i)^2}{2\lambda_i^2}\right\}\right].\end{aligned}$$

Rearranging terms and dividing the expression by $\sqrt{2\pi}$ gives

$$
\begin{aligned}
[x_i|\mathbf{W},\mu,\sigma^2] \quad &\propto \quad \sigma\Phi(-\frac{\mu}{\sigma})\frac{I_{(x_i<0)}}{\Phi(-\frac{\mu}{\sigma})\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\} \\
&\quad +\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\} \\
&\quad \times\frac{I_{(x_i>0)}}{\Phi(\frac{\nu_i}{\lambda_i})\sqrt{2\pi\lambda_i^2}}\exp\left\{-\frac{(x_i-\nu_i)^2}{2\lambda_i^2}\right\}.
\end{aligned}
$$

Substituting (4) and (5) yields

$$
\begin{aligned}
[x_i|\mathbf{W},\mu,\sigma^2] \quad &\propto \quad \sigma\Phi(-\frac{\mu}{\sigma})f_{0-}(x_i|\mu,\sigma^2) \\
&\quad +\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}f_{0+}(x_i|\nu_i,\lambda_i^2).
\end{aligned} \tag{7}
$$

The normalization constant is

$$
C \quad = \quad \frac{1}{\sigma\Phi(-\frac{\mu}{\sigma})+\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}}.
$$

Multiplying (7) by C yields a proper density

$$
\begin{aligned}
[x_i|\mathbf{W},\mu,\sigma^2] \quad &= \quad \frac{\sigma\Phi(-\frac{\mu}{\sigma})f_{0-}(x_i|\mu,\sigma^2)}{\sigma\Phi(-\frac{\mu}{\sigma})+\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}} \\
&\quad +\frac{\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}f_{0+}(x_i|\nu_i,\lambda_i^2)}{\Phi(-\frac{\mu}{\sigma})\sigma+\lambda_i\Phi(\frac{\nu_i}{\lambda_i})\exp\left\{-\frac{1}{2}(\frac{\mu}{\sigma^2}-\frac{\nu_i^2}{\lambda_i^2})\right\}} \\
&= \quad \rho_i f_{0-}(x_i|\mu,\sigma^2)+(1-\rho_i)f_{0+}(x_i|\nu_i,\lambda_i^2).
\end{aligned}
$$

8 MCMC sampling

Sampling from the conditional posterior distributions derived in Facts A-D requires the ability to sample from normal, truncated inverse gamma, truncated normal, and Bernoulli distributions. Routines for sampling the normal distributions are found in most statistical applications. Samples from a truncated inverse gamma distribution are obtained by taking the reciprocal of samples from a truncated gamma distribution. We sample from the truncated gamma distribution using inverse cdf sampling (Devroye, 1986). Ahrens and Dieter (1974, 1982) provide algorithms for sampling the gamma distribution. Sampling $w_{ij}|x_i,y_{ij}$ and $x_i|\mathbf{W},\mu,\sigma^2$ requires sampling from truncated normal distributions. We describe our method.

8.1 Sampling from the Truncated Normal Distribution

We use inverse cdf sampling (Devroye, 1986). Unfortunately, we encountered occasional underflow errors using the standard approximation for Φ. The following variant using the standard function for $\Psi(x) = \ln \Phi(x)$ worked well. Let U be a sample from a Uniform$(0, 1)$. Then it is easy to see that

$$\sigma \Psi^{-1}(\ln U + \Psi(\frac{t-\mu}{\sigma})) + \mu \quad \sim \quad \mathrm{TN}_{t-}(\mu, \sigma^2).$$

Further, by the symmetry of the normal distribution,

$$-\sigma \Psi^{-1}(\ln U + \Psi(\frac{\mu - t}{\sigma})) + \mu \quad \sim \quad \mathrm{TN}_{t+}(\mu, \sigma^2).$$

The standard approximation to Ψ is provided in the R statistical package.

8.2 Sampling from the conditional $x_i|\mathbf{W}, \mu, \sigma^2$

Sampling from (6) requires sampling from a Bernoulli(ρ_i) and a truncated normal distribution. Again, we found occasional underflow problems with the straightforward calculation of ρ_i. However, calculation of the log odds-ratio of ρ_i, denoted here by γ_i, proved satisfactory. Let

$$\begin{aligned}
\gamma_i &= \ln \left\{ \frac{\sigma \Phi(-\frac{\mu}{\sigma})}{\lambda_i \Phi(\frac{\nu_i}{\lambda_i}) \exp\left\{-\frac{1}{2}(\frac{\mu^2}{\sigma^2} - \frac{\nu_i^2}{\lambda_i^2})\right\}} \right\} \\
&= \ln \sigma + \ln \Phi\left(-\frac{\mu}{\sigma}\right) - \ln \lambda_i - \ln \Phi\left(\frac{\nu_i}{\lambda_i}\right) \\
&\quad + \frac{1}{2}\left(\frac{\mu^2}{\sigma^2} - \frac{\nu_i^2}{\lambda_i^2}\right).
\end{aligned}$$

Probability ρ_i is then expressed as

$$\rho_i = \frac{1}{1 + e^{-\gamma_i}}.$$

To sample $x_i|\mathbf{W}, \mu, \sigma^2$, samples of the following are drawn:

$$\begin{aligned}
r_i &\sim \mathrm{Bernoulli}(\rho_i), \\
t_i &\sim \mathrm{TN}_{0-}(\mu, \sigma^2), \\
q_i &\sim \mathrm{TN}_{0+}(\nu_i, \lambda_i^2).
\end{aligned}$$

With these samples, a sample of $x_i|\mathbf{W}, \mu, \sigma^2$ is given as

$$x_i|\mathbf{W}, \mu, \sigma^2 = \begin{cases} t_i & \text{if } r_i = 1 \\ q_i & \text{if } r_i = 0. \end{cases}$$

To save time in analysis, only one of t_i or q_i may be sampled, depending on the value of r_i.

9 Selecting Participants

The marginal posterior distributions obtained via Gibbs sampling can be used to compute the quantity $\Pr(x_i < 0|\mathbf{Y})$. It is natural to select those participants as at chance if this probability exceeds a criterion (denoted by c_0).

10 Statistical foundation of $\mathbf{Pr}(x_i < 0|\mathbf{Y})$

It is desirable to show that if a participant is selected with a criterion c_0, then the probability that this participant is truly at chance exceeds c_0; i.e. $\Pr(x_i < 0|\mathbf{Y}) \geq c_0 \Rightarrow Pr(x_i < 0) \geq c_0$. Unfortunately, this statement does not hold in the finite sample case. We show here that if μ and σ^2 are known, the implication holds. Let $\omega_i = \Pr(x_i < 0|y_i, \mu, \sigma^2)$.

Lemma 1 *$Pr(x_i < 0|w_i \geq c_0) \geq c_0$.*

Proof: For notational clarity, let C_i be the event that participant i is truly at chance, i.e.

$$C_i = \{x_i \leq 0\}.$$

Let R be the set of all outcomes y such that $\omega_i \geq c_0$.

$$R = \{y : w_i(y) > c_0\}.$$

Substituting the definition of ω_i yields

$$R = \{y : \Pr(C_i|y_i = y, \mu, \sigma) \geq c_0\}. \quad (8)$$

With this notation, $\Pr(x_i < 0|w_i \geq c_0) = \Pr(C_i|y_i \in R)$. To complete the proof, we show $\Pr(C_i|y_i \in R) \geq c_0$.

By the definition of conditional probability,

$$\begin{aligned} \Pr(C_i|y_i \in R) &= \frac{\Pr(C_i \cap \{y_i \in R\})}{\Pr(y_i \in R)} \\ &= \frac{\sum_{y \in R} \Pr(C_i|y_i = y)\Pr(y_i = y)}{\sum_{y \in R} \Pr(y_i = y)} \end{aligned}$$

Substituting the inequality $\Pr(C_i|y_i \in R) \geq c_0$ into the above equation yields

$$\begin{aligned} \Pr(C_i|y_i \in R) &\geq \frac{\sum_{y \in R} c_0 \Pr(y_i = y)}{\sum_{y \in R} \Pr(y_i = y)} \\ &\geq c_0, \end{aligned}$$

which completes the proof.

The quantity $\Pr(x_i < 0|\mathbf{Y})$ serves as an approximation to ω_i . Given the consistency of Bayesian estimators, in the limit $I \to \infty$, the quantity $\Pr(x_i < 0|\mathbf{Y})$ converges to ω_i. Lemma 1 is, therefore, useful in characterizing the asymptotic properties of selection.

11 Implementation

The following implementation analyzes the data in the companion manuscript. The code is written for **R**, which is freely available from the Comprehensive R Archive Network (http://cran.r-project.org/).

```
# This code takes a vector of data y and estimates
# MAC model parameters. Vector y contains the number of
# correct responses for each participant; vector N contains
# the total number of observations for each participant.

rm(list=ls()) # Clear workspace

#-----Helper functions------------------

# Function to sample from a TN(mu,sigma^2) truncated above
# at value trunc.at
rtnorm.lower = function(n=1,mu=0,sigma=1,trunc.at=0)
{
  sigma * qnorm(log(runif(n)) +
    pnorm((trunc.at-mu)/sigma,log.p=T),log.p=T) + mu
}

# Function to sample from a TN(mu,sigma^2) truncated below
# at value trunc.at
rtnorm.upper = function(n=1,mu=0,sigma=1,trunc.at=0)
{
  -sigma * qnorm(log(runif(n)) +
    pnorm((-trunc.at+mu)/sigma,log.p=T),log.p=T) + mu
}

# Create a general rtnorm function; this function will
# generate n normal deviates with mean mu and standard
# deviation sigma, truncated at trunc.at. n.lower values
# will be from normal truncated above 0, and n-n.lower
# from the normal truncated below 0.
rtnorm = function(n=1,mu=0,sigma=1,trunc.at=0,n.lower=1)
{
  c(rtnorm.lower(n.lower,mu,sigma,trunc.at),
  rtnorm.upper(n-n.lower,mu,sigma,trunc.at))
}

# Function to sample from an TIG(t,a,b)
rtruncinvgamma = function(n,t,a,b)
{
  u = runif(n)
```

```
  q = 1-pgamma(1/t,a,b)
  1/qgamma(1-u*q,a,b)
}


#-------End Helper Functions-------------

# Data; 27 subjects from the experiment described in Rouder,
# Morey, Speckman, and Pratte's 'Differentiating Chance
# Performance from above-chance performance: A solution to
# the null sensitivity problem in evaluating subliminal
# priming.'
# y = Number or correct responses
y = c(150,142,154,155,136,138,211,140,148,159,164,150,158,138,
      148,146,163,145,180,155,148,147,134,134,167,149,147)
# N = Number of observations
N = c(284,288,287,288,288,288,288,288,285,287,288,288,288,288,
      288,288,288,288,288,288,287,287,288,286,288,288,288)

# For housekeeping
obs.fewer=max(N)-N # How many observations LESS does each have?

# Initialize variables
M      = 50000      # Length of MCMC chain
I      = length(y)  # Number of participants
burnin = 5000       # Throw away this many at start of chain


w       = matrix(nrow=max(N),ncol=I)
x       = matrix(nrow=M,ncol=I)
mu      = matrix(nrow=M)
sigma.2 = matrix(nrow=M)

# Initialize prior parameters
mu_0      = 0
sigma_0.2 = 1
a_0       = -.5
b_0       = 0
t_sig2    = 1

# Starting values
p.hat = (y+1)/(N+2) # Estimate p from data for starting values
x[1,]      = qnorm(p.hat)
mu[1]      = 0
sigma.2[1] = 1
for(i in 1:I) # Find reasonable starting values for w_ij
```

```
{
  w[,i] = c(rtnorm(n=N[i],mu=x[1,i],sigma=1,
            trunc.at=0,n.lower=N[i]-y[i]),
            rep(NA,obs.fewer[i]))
}

# Main MCMC loop
for(m in 2:M) # Loop through iterations
{
  for(i in 1:I)  # Loop through participants
  {
    # Sampling x_i
    lambda_i.2 = (N[i] + (1/sigma.2[m-1]))^-1
    nu_i     = lambda_i.2*(sum(w[,i],na.rm=T) +
                mu[m-1]/sigma.2[m-1])
    gamma_i  = log(sqrt(sigma.2[m-1])) +
                pnorm(-mu[m-1]/sqrt(sigma.2[m-1]),log.p=T) -
                log(sqrt(lambda_i.2)) -
                pnorm(nu_i/sqrt(lambda_i.2),log.p=T) +
                (1/2)*(mu[m-1]^2/sigma.2[m-1] - nu_i^2/lambda_i.2)
    rho_i    = 1/(1+exp(-gamma_i))
    r_i      = rbinom(1,1,rho_i) # Sample Bernoulli
    if(r_i==0){
      x[m,i] = rtnorm.upper(n=1,mu=nu_i,
                 sigma=sqrt(lambda_i.2),trunc.at=0)
    }else{
      x[m,i] = rtnorm.lower(n=1,mu=mu[m-1],
                 sigma=sqrt(sigma.2[m-1]),trunc.at=0)
    }
    # Sampling w_ij
    w[,i] = c(rtnorm(n=N[i],mu=max(0,x[m,i]),
               sigma=1,trunc.at=0,n.lower=N[i]-y[i]),
               rep(NA,obs.fewer[i]))
  } # End loop through participants

  # Sampling mu
  tau.2 = (I/sigma.2[m-1] + 1/sigma_0.2)^-1
  eta   = tau.2*(sum(x[m,])/sigma.2[m-1] + mu_0/sigma_0.2)
  mu[m] = rnorm(1,eta,sqrt(tau.2))

  # Sampling sigma^2
  sigma.2[m] = rtruncinvgamma(1, t_sig2, a_0 + I/2,
                 b_0 + sum((x[m,]-mu[m])^2)/2)

  # Report progress
  if(!(m%%500)) { cat('Iteration number:',m,'\n') }
```

```
} # End main MCMC loop

# Estimate parameters (be sure to check for convergence)
mu.estimate      = mean(mu[burnin:M])
sigma.2.estimate = mean(sigma.2[burnin:M])
pr.x.lt.0        = colMeans(x[burnin:M,]<0)

# Create a plot of the output
plot(y/N,pr.x.lt.0)
abline(h=.95)
```

References

[1] J.H. Ahrens and U. Dieter. Computer methods for sampling from gamma, beta, poisson and binomial distributions. *Computing*, 12:223–246, 1974.

[2] J.H. Ahrens and U. Dieter. Generating gamma variates by a modified rejection technique. *Association for Computing Machinery. Communications of the ACM*, 25:47–54, 1982.

[3] J. Albert and S. Chib. Bayesian residual analysis for binary response regression models. *Biometrika*, 82:747–759, 1995.

[4] L. Devroye. *Non-uniform random variate generation.* Springer-Verlag Inc, 1986.

[5] A. Gelfand and A. F. M. Smith. Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409, 1990.

[6] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis (2nd edition).* Chapman and Hall, London, 2004.

[7] J. N. Rouder and J. Lu. An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin and Review*, 12:573–604, 2005.