BayesSDT: Software for Bayesian inference with signal detection theory

MICHAEL D. LEE
University of California, Irvine, California

This article describes and demonstrates the BayesSDT MATLAB-based software package for performing Bayesian analysis with equal-variance Gaussian signal detection theory (SDT). The software uses WinBUGS to draw samples from the posterior distribution of six SDT parameters: discriminability, hit rate, false alarm rate, criterion, and two alternative measures of bias. The software either provides a simple MATLAB graphical user interface or allows a more general MATLAB function call to produce graphs of the posterior distribution for each parameter of interest for each data set, as well as to return the full set of posterior samples.

This article describes and demonstrates the simple BayesSDT software package—developed in MATLAB—for performing Bayesian analysis using signal detection theory (SDT). As background, we first briefly describe SDT and the Bayesian approach to its analysis. We then present illustrative examples of the BayesSDT software and its outputs.

Signal Detection Theory

SDT—developed by Peterson, Birdsall, and Fox (1954) and applied to psychophysics by Tanner and Swets (1954)—is a very general, useful, and widely employed method for drawing inferences from data in psychology (see Green & Swets, 1966; Macmillan & Creelman, 2004, for detailed introductions). SDT analysis is often applied to data that take a 2 × 2 form—as shown in Table 1—in which there are "signal" trials and "noise" trials, and "yes" responses and "no" responses. When a "yes" response is given for a signal trial, it is called a *hit*. When a "yes" response is given for a noise trial, it is called a *false alarm*. When a "no" response is given for a signal trial, it is called a *miss*. When a "no" response is given for a noise trial, it is called a *correct rejection*. The data needed for an SDT analysis are just the counts of hits, false alarms, misses, and correct rejections.

The key assumptions of SDT are shown in Figure 1 and involve representation and decision-making assumptions. Representationally, the idea is that signal and noise trials can be represented as values along a unidimensional "strength" construct. Both types of trials are assumed to produce strengths that vary according to a Gaussian distribution along this dimension. The signal strengths are assumed to be greater—on average—than the noise strengths; thus, the signal strength distribution has a greater mean. We deal with only the most common equal-variance form of SDT, in which both the distributions are assumed to have the same variance. The decision-making assumption of SDT is that "yes" and "no" responses are produced by comparing the strength of the current trial with a fixed criterion. If the strength exceeds the criterion, then a "yes" response is made; otherwise, a "no" response is made.

Figure 1 provides a formal version of the equal-variance SDT model. Since the underlying strength scale has arbitrary units, the variances are fixed to 1, and the mean of the noise distribution is fixed to 0. The mean of the signal distribution is d' and measures the discriminability of the signal and noise trials as being the distance between their two distributions.

The strength value $\frac{1}{2}d'$ is special because it is the criterion value at which both signal and noise distributions are equally likely. In this sense, using a criterion of $\frac{1}{2}d'$ corresponds to balanced or unbiased responding. The actual criterion used for responding is denoted k, and distance between this criterion and the unbiased criterion is denoted c. This makes c a measure of bias, because it corresponds to how different the actual criterion is from the unbiased one. Positive values of c correspond to a bias toward saying "no" and thus to an increase in correct rejections at the expense of an increase in misses. Negative values of c correspond to a bias toward saying "yes" and thus to an increase in hits at the expense of an increase in false alarms.

Figure 1 also shows an alternative measure of bias, β, which is the likelihood ratio of the signal-to-noise distribution at the criterion, expressed on a log odds scale. Again, positive values of β correspond to a bias toward saying "no," whereas negative values correspond to a bias toward saying "yes." It has been argued that one advantage of the c measure over β is that c is independent of d', whereas β is not (Snodgrass & Corwin, 1988).

The SDT model—with its representation and decision-making assumptions—naturally makes predictions about

M. D. Lee, mdlee@uci.edu

Table 1
Signal Detection Theory Terminology

Response	Signal Trial	Noise Trial
"Yes"	Hit	False alarm
"No"	Miss	Correct rejection

hit and false alarm rates and thus maps onto the counts in Table 1. In Figure 1, the hit rate, h, is shown as the proportion of the signal distribution above the criterion k. Similarly, the false alarm rate, f, is the proportion of the noise distribution above the criterion k. The usefulness of SDT is that, through this relationship, it is possible to take the sort of data in Table 1 and convert the counts of hits and false alarms into psychologically meaningful measures of discriminability and bias. Discriminability is a measure of how easily signal and noise trials can be distinguished. Bias is a measure of how the decision-making criterion being used relates to the balanced criterion.

A Graphical Model for Bayesian Inference

Bayesian inference. The Bayesian approach to statistical inference (see, e.g., Gelman, Carlin, Stern, & Rubin, 2004; Jaynes, 2003; Lee & Wagenmakers, 2005) has a number of desirable properties in inferring the parameters of SDT models from data. These advantages stem from the basic Bayesian philosophy of always representing what is known and unknown about parameters of interest using probability distributions. The distributions provide complete representations of uncertainty and can be updated in a coherent and principled way using probability theory. Bayesian methods have recently found widespread appli-

cation in psychology—for instance, as models of human cognition (e.g., Anderson, 1991; Griffiths & Tenenbaum, 2005; Kemp, Bernstein, & Tenenbaum, 2005; Sanborn, Griffiths, & Navarro, 2006; Tenenbaum & Griffiths, 2001) and as methods for relating psychological models to data (e.g., Kuss, Jäkel, & Wichmann, 2005; Lee, 2008; Lee & Webb, 2005; Navarro, Griffiths, Steyvers, & Lee, 2006; Pitt, Myung, & Zhang, 2002; Rouder, Lu, Speckman, Sun, & Jiang, 2005).

In the context of SDT, Bayesian advantages include avoiding the need for edge corrections that can be a problem in standard analyses (see Snodgrass & Corwin, 1988),[1] and posterior distributions are not constrained to take particular parametric forms; thus, they better represent parameter values—like very high hit or false alarm rates—that are near the boundary of the theoretically possible range. Bayesian methods are also automatically exact for any sample size and are sensitive to how many data are available in making inferences. Additional potential advantages of Bayesian methods for SDT, which go beyond the functionality of the software presented in the present article, include hierarchical extensions to model individual differences in participants and in items (e.g., Lee & Paradowski, 2007; Rouder & Lu, 2005) and the straightforward consideration of alternative unequal variance and other extended SDT models without requiring analytic results for inference.

Graphical model. Graphical models provide a convenient formalism for expressing many Bayesian models (see Griffiths, Kemp, & Tenenbaum, in press; Jordan, 2004; Lee, 2008, for two psychological introductions and one statistical introduction). In graphical modeling, a graph is created in which nodes represent parameters and data. The structure of the graph then indicates dependencies between the parameters and data, with children depending on their parents. The practical advantage of graphical models is that sophisticated and relatively general-purpose Markov chain Monte Carlo (MCMC) algorithms exist that can sample from the full joint posterior distribution of the parameters conditional on the observed data. Our software relies on WinBUGS (Spiegelhalter, Thomas, & Best, 2004), which uses a range of MCMC computational methods, including adaptive rejection sampling, slice sampling, and Metropolis–Hastings (see, e.g., Chen, Shao, & Ibrahim, 2000; Gilks, Richardson, & Spiegelhalter, 1996; MacKay, 2003) to perform posterior sampling.

Figure 2 shows the graphical model we use for making inferences about the SDT parameters in Figure 1. The convention is used that observed variables (i.e., data) have shaded nodes, whereas unobserved variables (i.e., the parameters to be inferred from data) are not shaded. In addition, continuous variables have circular nodes, whereas discrete variables have square nodes. Finally, stochastic variables in our model are shown with single-bordered nodes, whereas deterministic variables (i.e., those that are defined simply as functions of other variables) are shown as double-bordered nodes.

The best way to understand the graphical model is to begin with the discriminability d' and bias c parameters.
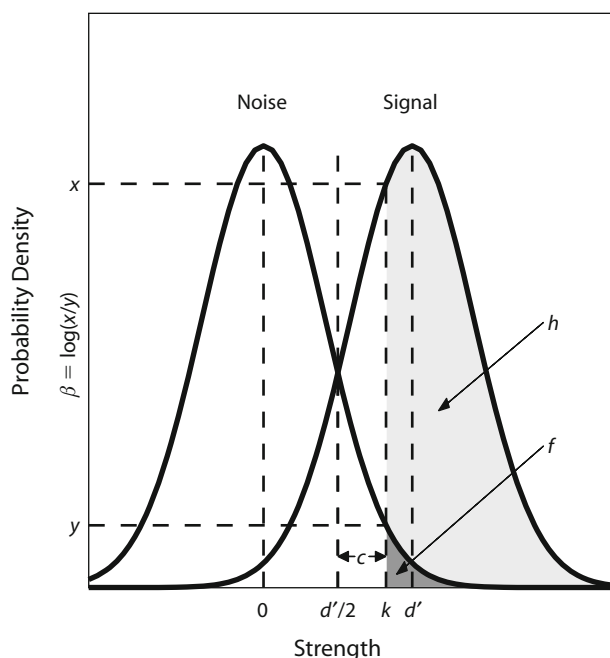

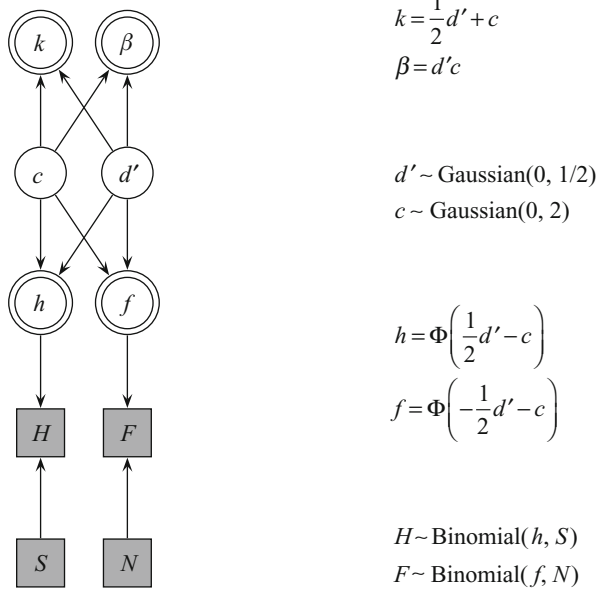**Figure 1. Equal-variance Gaussian signal detection theory framework.**

$$k = \frac{1}{2}d' + c$$

$$\beta = d'c$$

$$d' \sim \text{Gaussian}(0, 1/2)$$

$$c \sim \text{Gaussian}(0, 2)$$

$$h = \Phi\left(\frac{1}{2}d' - c\right)$$

$$f = \Phi\left(-\frac{1}{2}d' - c\right)$$

$$H \sim \text{Binomial}(h, S)$$

$$F \sim \text{Binomial}(f, N)$$

Figure 2. Graphical model for inferring discriminability (d'), criterion (k), two measures of bias (c and β), and hit (h) and false alarm (f) rates from H observed hits out of S signal trials and F observed false alarms out of N noise trials.

In order to relate these parameters to the observed data, d' and c can be reparameterized (under the model given by SDT) into hit and false alarm rates h and f. This means that there is a deterministic relationship between h and f on the one hand, and between d' and c on the other, given by

$$h = \Phi\left(\frac{1}{2}d' - c\right) \tag{1}$$

$$f = \Phi\left(-\frac{1}{2}d' - c\right) \tag{2}$$

where $\Phi(\cdot)$ is the standard Gaussian cumulative density function.

The number of signal trials, S, and the number of noise trials, N, are observed data, as are the hit count H and the false alarm count F. The hit and false alarm counts follow a binomial distribution depending on both the hit and false alarm rates, and the number of signal and noise trials, so that

$$H \sim \text{Binomial}(S, h) \tag{3}$$

and

$$F \sim \text{Binomial}(N, f). \tag{4}$$

In this way, SDT relates discriminability d' and bias c parameters to the observed data so that posterior inferences can be drawn. From the posterior distributions of d' and c, the posterior distributions for the criterion k and alternative bias measure β can be calculated according to the standard results

$$k = \frac{d'}{2} + c \tag{5}$$

$$\beta = cd'. \tag{6}$$

The final component of the Bayesian graphical model involves prior assumptions. We choose priors on discriminability and bias that correspond to the assumption of uniform priors for the hit and false alarm rates, consistent with standard practice. It can be proven[2] that the required distributions are

$$d' \sim \text{Gaussian}(0,2) \tag{7}$$

and

$$c \sim \text{Gaussian}(0,\tfrac{1}{2}). \tag{8}$$

WinBUGS implementation. The graphical model shown in Figure 2 is straightforward to implement in WinBUGS, using the script in Table 2. The required inputs are the hit, false alarm, miss, and correct-rejection counts, and the outputs are samples from the posterior distribution of the d', h, f, k, c, and β parameters of the SDT model.[3]

The BayesSDT Software Package

The BayesSDT software package consists of a number of MATLAB scripts and functions, as well as WinBUGS scripts. It can be downloaded from www.socsci.uci.edu/~mdlee/. BayesSDT requires that MATLAB and WinBUGS 1.4.2 or later be installed. WinBUGS is available from www.mrc-bsu.cam.ac.uk/bugs/. BayesSDT also requires the MatBugs MATLAB function available at www.cs.ubc.ca/~murphyk/Software/MATBUGS/matbugs.html, which allows MATLAB and WinBUGS to pass information to each other.

The BayesSDT software provides two MATLAB interfaces with the WinBUGS implementation of the SDT graphical model. The easiest to use is a graphical user interface (GUI), but a MATLAB function call is also provided to allow more than four data sets to be analyzed simultaneously.

Graphical user interface. The BayesSDT GUI is started by running the BayesSDT_GUI MATLAB script. A demonstration of the interface is displayed in Figure 3, showing the entry of three illustrative data sets for analysis. The first is the "many" data set, involving 70 hit counts and 50 false alarm counts from 100 signal and 100 noise trials. The second is the "few" data set, involving 7 hit counts and 5 false alarm counts from 10 signal and 10 noise trials. The idea is that the many data set has the same rate of hits and false alarms as does the few data set, to illustrate how Bayesian analysis represents the decrease in uncertainty with additional data. The third, "perfect" data set involves 10 hits and 0 false alarms from 10 signal and 10 noise trials. The purpose of this data set is to illustrate how Bayesian analysis deals with edge effects.

As Figure 3 shows, the GUI allows a label to be entered for each data set. The color, width, and style properties of the lines used to draw the posterior distribution for each data set can also be specified using standard MATLAB options. The `help plot` command in MATLAB provides a full list. Checkboxes allow the selection of the SDT parameters to be analyzed. Finally, the number of posterior samples to be generated—as well as the number of bins used to display the final posterior densities—can be

Table 2
WinBUGS Script for Bayesian SDT Graphical Model

```
# BAYESIAN SIGNAL DETECTION THEORY
#
# INPUT VARIABLES
# H, F, M, C are the hit, false alarm,
# miss and correct rejection signal detection counts
#
# OUTPUT VARIABLES
# d, k are discriminability and criterion
# h, f are the hit and false-alarm rates
# c, b are measures of bias

model {

    # Relating observed counts to underlying Hit and False Alarm rates
    # Number of Signal Trials is sum of Hit and Miss counts
    S <- H+M
    # Number of Noise Trials is sum of False Alarm and Correct Rejection counts
    N <- F+C
    # Hit counts are Binomial
    H ~ dbin(h,S)
    # False alarm counts are Binomial
    F ~ dbin(f,N)

    # Reparameterization, converting Hit and False Alarm rates
    #  to Discriminability and Bias indices
    h <- phi(d/2-c)
    f <- phi(-d/2-c)
    k <- d/2+c
    b <- d*c

    # These priors over Discriminability and Bias correspond
    #  to uniform priors over the Hit and False Alarm rates
    MEAND <- 0
    MEANC <- 0
    LAMBDAD <- 1/2
    LAMBDAC <- 2
    c ~ dnorm(MEANC,LAMBDAC)
    d ~ FSDT)

}
```

specified. Additional samples and bins lead to smoother approximations to the posterior density, but come at the expense of additional computing time. As a guide to speed, in the present article, collecting and analyzing 10^6 samples and 100 bins for the three data sets—many, few, and perfect—took 273 sec on a PC with a processor speed of 2.4 GHz and 2 Gb of RAM.

Figure 4 shows the graphical output of the BayesSDT software, which is the main result of the analysis, although BayesSDT GUI also automatically generates a text file—BayesSDToutput.txt—listing the means and standard deviations for each of the parameters. In Figure 4, the posterior distributions for the many data set show that the parameters are estimated with relatively little uncertainty. For the few data set, however, the posterior distributions now represent the much greater degree of uncertainty. For each parameter, the mode is the same, but the range of possible values the parameter could take is greater. This is a natural consequence of the parameter estimates being based on fewer data, and it is well handled by the Bayesian approach. For the perfect data set, the modal hit and false alarm rates are 1.0 and 0.0, but other possibilities have some density; thus, the remaining discriminability, bias, and criterion pa-

rameters are all well defined without the need for edge corrections. It is also worth noting that these hit and false alarm posteriors do not take the Gaussian form assumed by standard analyses, but instead are sensitive to the theoretical bounds on the values of these rate parameters.

MATLAB function. Rather than using the BayesSDT GUI, it is possible to call the BayesSDT function directly. This is especially useful for analyzing more than four data sets simultaneously. The BayesSDT function takes a single argument, D, which is a structured variable with 14 fields.

`D.ndatasets` is the number of data sets to be analyzed.

`D.sdt`. This is a matrix with the rows representing data sets and the columns representing signal detection data counts. The four columns correspond to hits, false alarms, misses, and correct rejections, respectively.

`D.nsamples` is the number of posterior samples to generate.

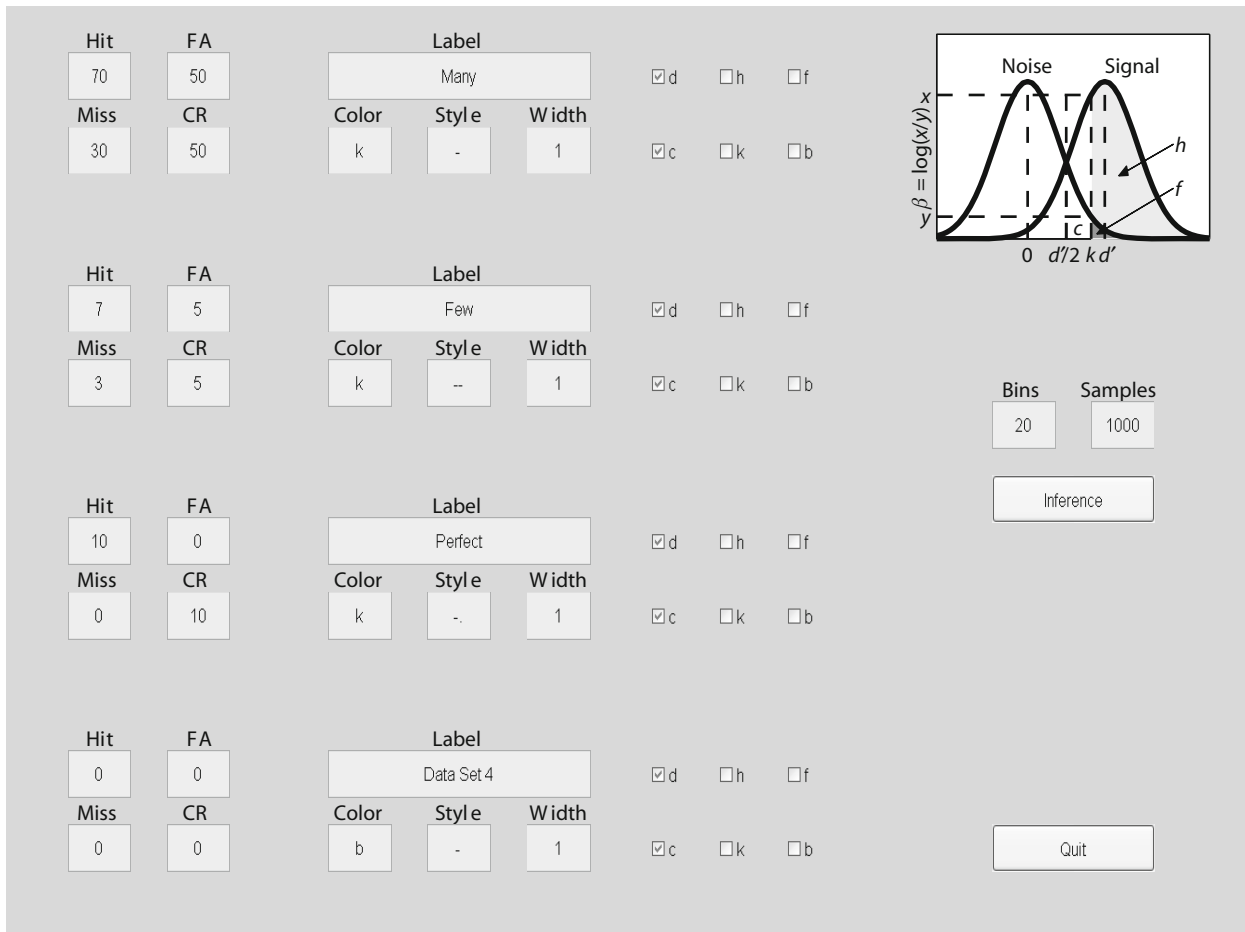`D.nbins` is the number of bins to use in drawing histograms of the posterior densities.

Figure 3. The BayesSDT graphical user interface.

D.labels is a character array with a string label for each data set.

D.linecolor is a character array with a MATLAB color for each data set, using the standard MATLAB plot colors.

D.linestyle is a character array with a MATLAB color for each data set, using the standard MATLAB line styles.

D.linewidth is an array of positive numbers containing line widths for each data set.

D.dcheck, D.hcheck, D.fcheck, D.kcheck, D.ccheck, and D.bcheck are vectors of 0 and 1 entries for each data set, indicating whether or not the d', h, f, k, c, and β parameters, respectively, should be analyzed.

The BayesSDT function generates graphs of posterior distributions for the analyzed parameters. It also returns two structured variables—samples and stats—returned by WinBUGS. The fields in samples give the full list of posterior samples of each parameter for each data set. These samples allow more advanced Bayesian analyses based on the full-joint posterior distribution

(e.g., Lee, 2008). The fields in stats provide means and standard deviations for the posterior samples of each parameter for each data set. The syntax for the Bayes-SDT function call is simply [samples stats] = BayesSDT(D).

Conclusion

SDT is a highly useful and widely used method for analyzing psychological data. Bayesian inference for SDT has a number of attractive properties, using the full posterior distributions of parameters to represent what is known about discriminability, bias, criterion, and hit and false alarm rates. The BayesSDT software package provides a simple WinBUGS implementation of a suitable Bayesian graphical model and provides a MATLAB GUI and function call for performing the Bayesian analysis.
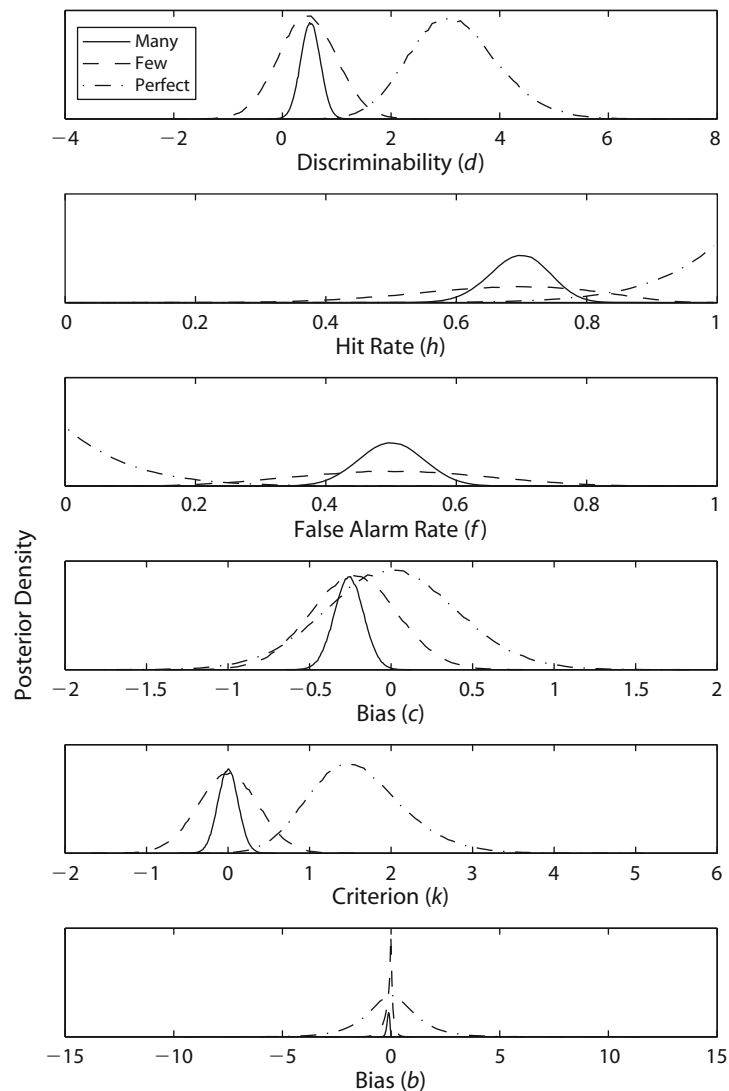
Figure 4. Posterior distributions for discriminability (d'), hit (h) and false alarm (f) rates, criterion (k), and two measures of bias (c and β) for the illustrative "many," "few," and "perfect" data sets.

REFERENCES

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, **98**, 409-429.

Angus, J. E. (1994). The probability integral transform and related results. *SIAM Review*, **36**, 652-654.

Chen, M.-H., Shao, Q.-M., & Ibrahim, J. G. (2000). *Monte Carlo methods in Bayesian computation*. New York: Springer.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). London: Chapman & Hall/CRC.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (Eds.) (1996). *Markov chain Monte Carlo in practice*. Boca Raton, FL: Chapman & Hall/CRC.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (in press). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational psychology*. Cambridge: Cambridge University Press.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, **51**, 334-384.

Jaynes, E. T. (2003). *Probability theory: The logic of science* (G. L. Bretthorst, Ed.). New York: Cambridge University Press.

Jordan, M. I. (2004). Graphical models. *Statistical Science*, **19**, 140-155.

Kemp, C., Bernstein, A., & Tenenbaum, J. B. (2005). A generative theory of similarity. In B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* [CD-ROM]. Mahwah, NJ: Erlbaum.

Kuss, M., Jäkel, F., & Wichmann, F. A. (2005). Bayesian inference for psychometric functions. *Journal of Vision*, **5**, 478-492.

Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, **15**, 1-15.

Lee, M. D., & Paradowski, M. J. (2007). Group decision-making on an optimal stopping problem. *Journal of Problem Solving*, **1**, 53-73.

Lee, M. D., & Wagenmakers, E.-J. (2005). Bayesian statistical inference in psychology: Comment on Trafimow (2003). *Psychological Review*, **112**, 662-668.

Lee, M. D., & Webb, M. R. (2005). Modeling individual differences in cognition. *Psychonomic Bulletin & Review*, **12**, 605-621.

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge: Cambridge University Press.

MACMILLAN, N. A., & CREELMAN, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Hillsdale, NJ: Erlbaum.

NAVARRO, D. J., GRIFFITHS, T. L., STEYVERS, M., & LEE, M. D. (2006). Modeling individual differences using Dirichlet processes. *Journal of Mathematical Psychology*, **50**, 101-122.

PETERSON, W. W., BIRDSALL, T. G., & FOX, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Group on Information Theory*, **PGIT-4**, 171-212.

PITT, M. A., MYUNG, I. J., & ZHANG, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, **109**, 472-491.

ROUDER, J. N., & LU, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, **12**, 573-604.

ROUDER, J. N., LU, J., SPECKMAN, P. [L.], SUN, D., & JIANG, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, **12**, 195-223.

SANBORN, A. N., GRIFFITHS, T. L., & NAVARRO, D. J. (2006). A more rational model of categorization. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 726-731). Mahwah, NJ: Erlbaum.

SNODGRASS, J. G., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

SPIEGELHALTER, D. J., THOMAS, A., & BEST, N. G. (2004). *WinBUGS version 1.4 user manual*. Cambridge: Medical Research Council Biostatistics Unit.

TANNER, W. P., JR., & SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, **61**, 401-409.

TENENBAUM, J. B., & GRIFFITHS, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral & Brain Sciences*, **24**, 629-640.

NOTES

1. As Rouder and Lu (2005) pointed out, the common edge correction of adding a ½ count to both hits and false alarms can be interpreted as implicitly assuming a Bayesian prior of Beta(½, ½). In this sense, frequentist edge corrections are already Bayesian, although the choice of Beta(½, ½) as a prior seems a strange one, since it gives, for example, more prior probability to a hit rate of 0.1 than to a hit rate of 0.5.

2. The proof relies on the probability integral transform theorem (see, e.g., Angus, 1994), which says that if X is a continuous real-valued random variable with strictly increasing cumulative distribution function F_X, then $F_X(X)$ is uniformly distributed on (0, 1). This means, for example, that if $X \sim$ Gaussian(0, 1), then $\Phi(X) \sim$ Uniform(0, 1). I thank Geoff Iverson for showing me the result.

3. Note that in defining the priors on d' and c, WinBUGS uses a mean and precision (i.e., the reciprocal of the variance) parameterization for the Gaussian distribution.