Figures and Tables

from

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

This document is intended as a presentation aid for instructors who use the book for their classes. The tables and figures are from the manuscript form of the book, and might not exactly match the form presented in the book.

Table 1.1: Bayesian analogues of null hypothesis significance tests. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Traditional Analysis Name	**Bayesian Analogue**
Binomial test	Chapters 6–9, 21
t-test	Chapter 16
Simple linear regression	Chapter 17
Multiple linear regression	Chapter 18
Oneway ANOVA	Chapter 19
Multi-factor ANOVA	Chapter 20
Logistic regression	Chapter 21
Multinomial Logistic regression	Chapter 22
Ordinal regression	Chapter 23
Chi-square test (contingency table)	Chapter 24
Power analysis (sample size planning)	Chapter 13

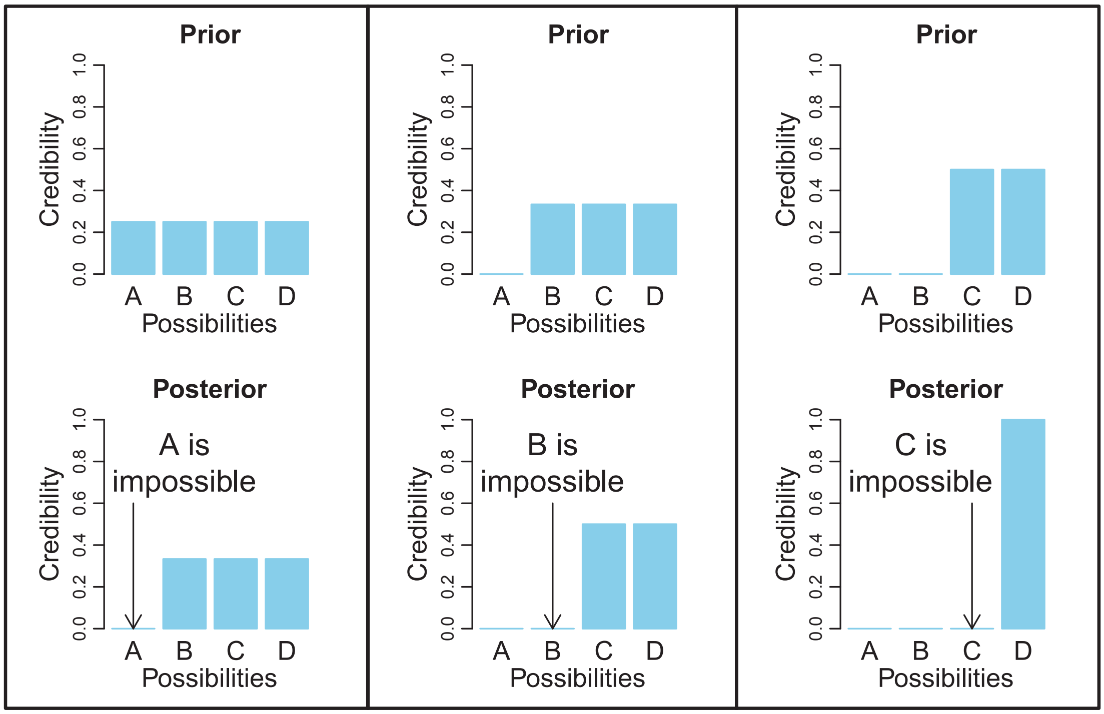

Figure 2.1: The upper-left graph shows the credibilities four possible causes for an outcome. The causes, labeled A, B, C and D, are mutually exclusive and exhaust all possibilities. The causes happen to be equally credible at the outset, hence all have prior credibility of 0.25. The lower-left graph shows the credibilities when one cause is learned to be impossible. The resulting posterior distribution is used as the prior distribution in the middle column, where another cause is learned to be impossible. The posterior distribution from the middle column is used as the prior distribution for the right column. The remaining possible cause is fully implicated by Bayesian re-allocation of credibility.

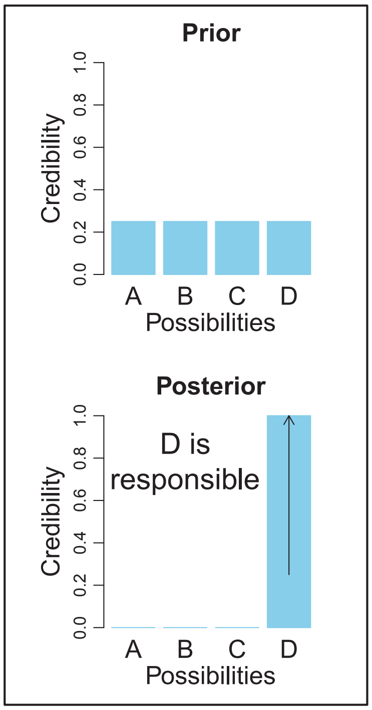

Figure 2.2: The upper graph shows the credibilities four possible causes for an outcome. The causes, labeled A, B, C and D, are mutually exclusive and exhaust all possibilities. The causes happen to be equally credible at the outset, hence all have prior credibility of 0.25. The lower graph shows the credibilities when one cause is learned to be responsible. The non-responsible causes are "exonerated" (i.e., have zero credibility as causes) by Bayesian re-allocation of credibility. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

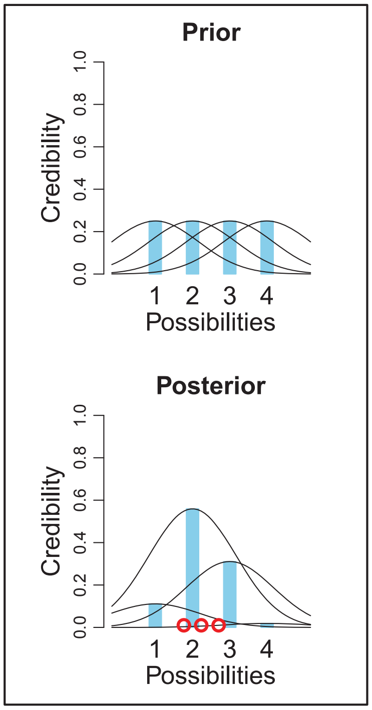

Figure 2.3: The upper graph shows the prior credibilities of four candidate means in normal distributions, located at values of 1, 2, 3, and 4. Superimposed on the means are the corresponding normal distributions. The horizontal axis is playing double duty as a scale for the means (marked by the blue bars) and for the data (suggested by the normal distributions). The three observed data values are plotted as circles on the floor of the lower panel. Bayesian reallocation of credibility across the four candidate means indicates that the mean at 2 is most credible given the data, the mean at 3 is somewhat credible, and so on. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

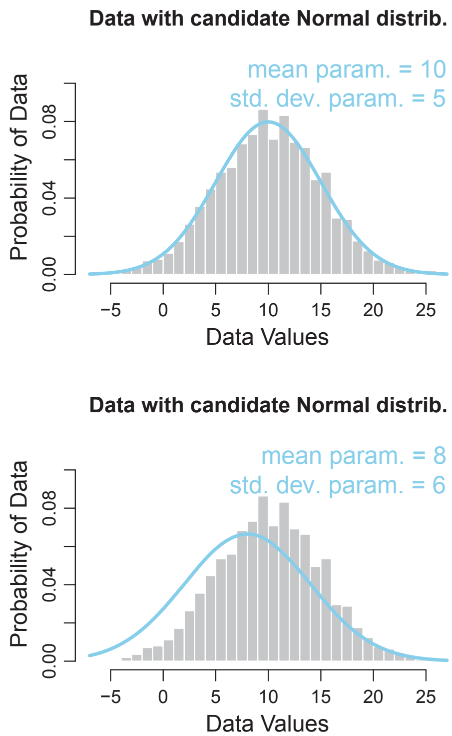

Figure 2.4: The two graphs show the same data histogram but with two different candidate descriptions by normal distributions. Bayesian analysis computes the relative credibilities of candidate parameter values. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

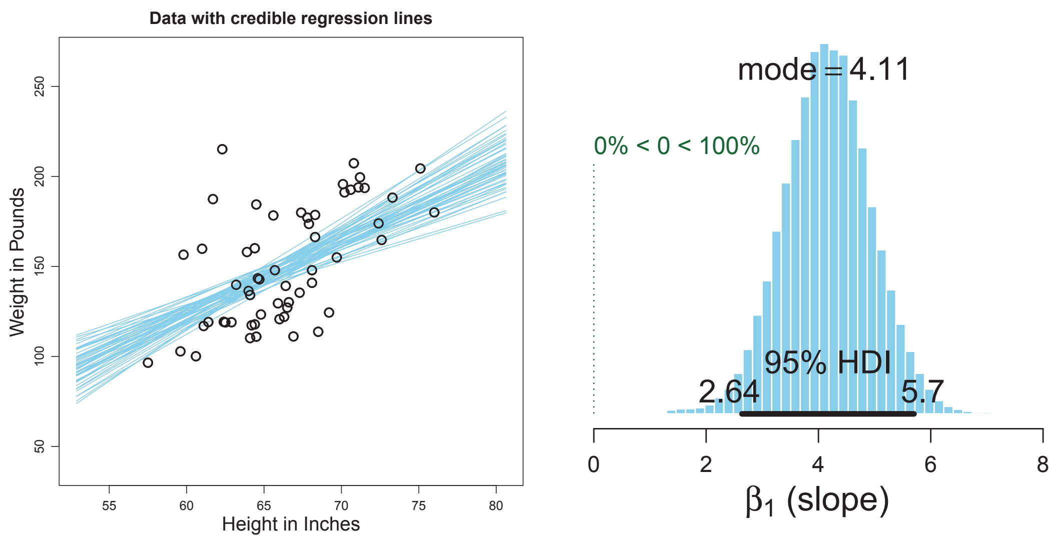

Figure 2.5: Data are plotted as circles in the scatter plot of the left panel. The left panel also shows a smattering of credible regression lines from the posterior distribution superimposed on the data. The right panel shows the posterior distribution of the slope parameter (i.e., β_1 in Eqn. 2.1). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

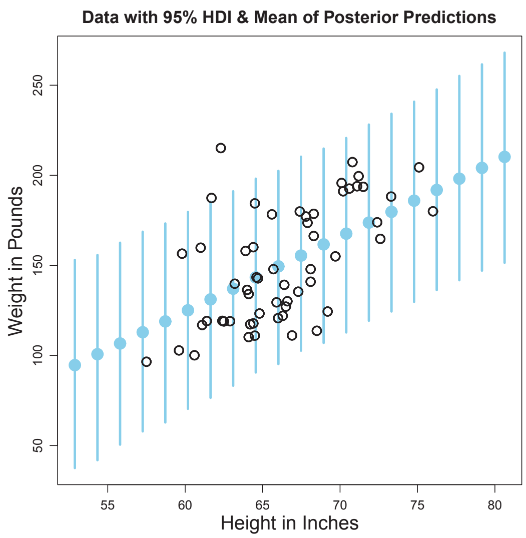

Figure 2.6: The data of Figure 2.5 are shown with posterior predicted weight values superimposed at selected height values. Each vertical bar shows the range of the 95% most credible predicted weight values, and the dot at the middle of each bar shows the mean predicted weight value.

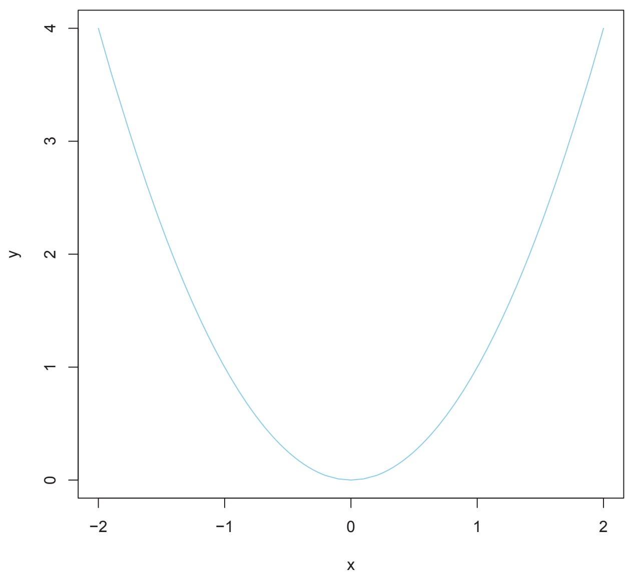

Figure 3.1: A simple graph drawn by R. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

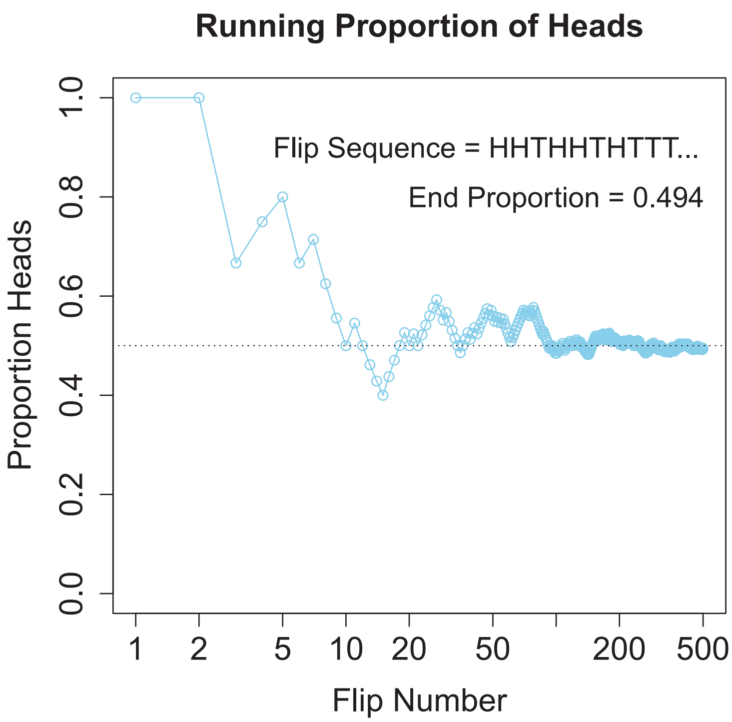

Figure 4.1: Running proportion of heads when flipping a coin. The x-axis is plotted on a logarithmic scale so that you can see the details of the first few flips but also the long-run trend after many flips. R code for producing this figure is discussed in Section 4.5.

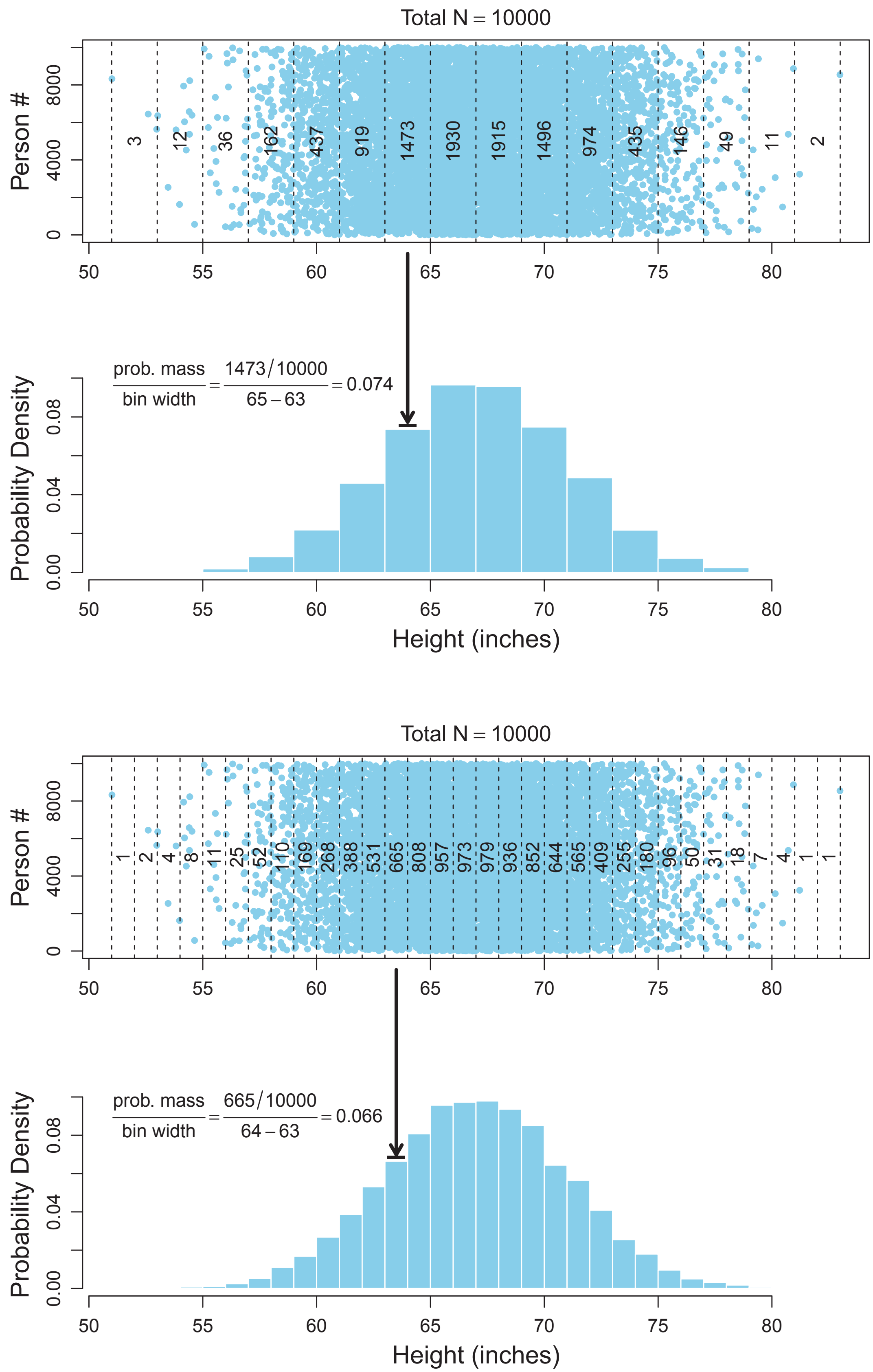

Figure 4.2: Examples of computing probability density. Within each main panel, the upper plot shows a scatter of 10,000 heights of randomly selected people, and the lower plot converts into probability density for the particular selection of bins depicted.

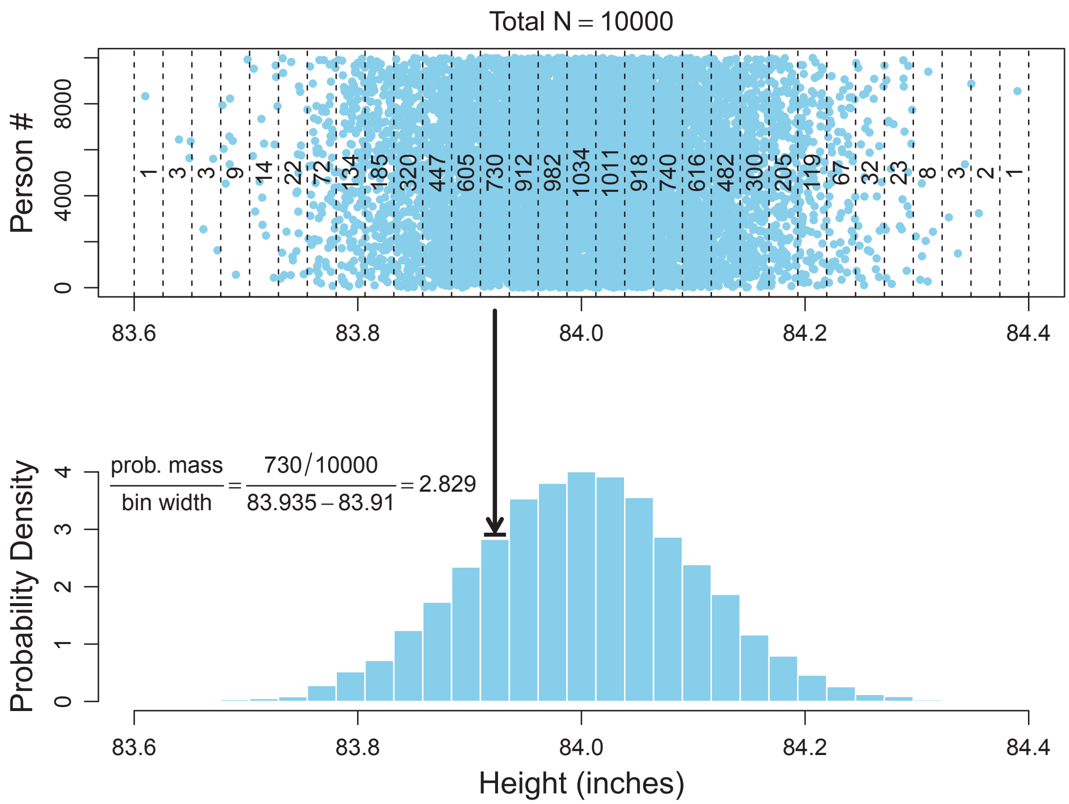

Figure 4.3: Example of probability density greater than 1.0. Here, all the probability mass is concentrated into a small region of the scale, and therefore the density can be high at some values of the scale. The annotated calculation of density uses rounded interval limits for display. (For this example, we can imagine that the points refer to manufactured doors instead of people, and therefore the y-axis of the top panel should be labelled "Door" instead of "Person.") Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

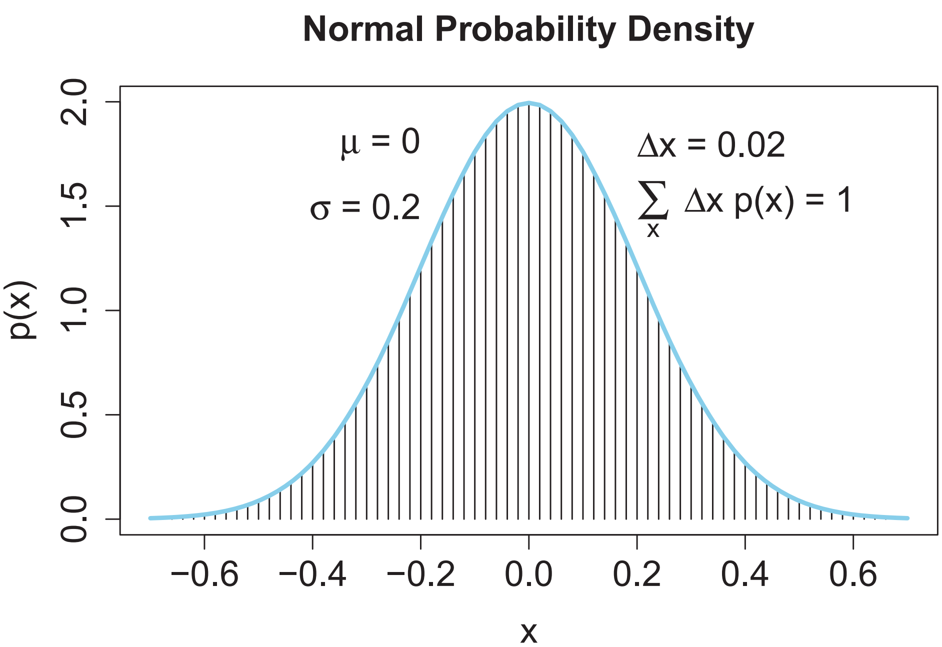

Figure 4.4: A normal probability density function, shown with a comb of narrow intervals. The integral is approximated by summing the width times height of each interval. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

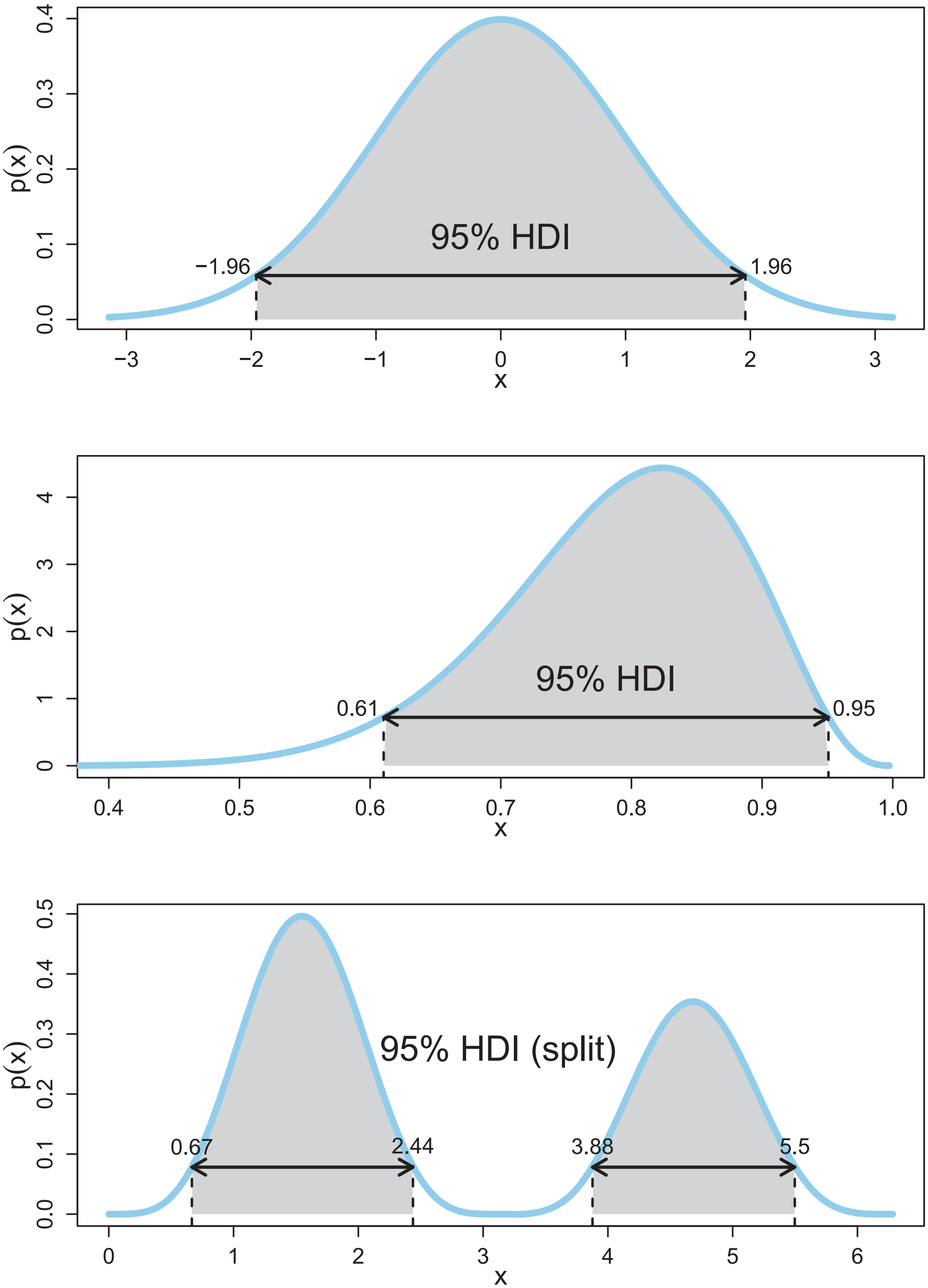

Figure 4.5: Examples of 95% highest density intervals (HDIs). For each example, all the x values inside the interval have higher density than any x value outside the interval, and the total mass of the points inside the interval is 95%. The 95% area is shaded, and it includes the zone below the horizontal arrow. The horizontal arrow indicates the width of the 95% HDI, with its ends annotated by (rounded) x values. The height of the horizontal arrow marks the minimal density exceeded by all x values inside the 95% HDI. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Table 4.1: Proportions of combinations of hair color and eye color. Some rows or columns may not sum exactly to their displayed marginals because of rounding error from the original data. Data adapted from Snee (1974). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Eye Color	Hair Color Black	Brunette	Red	Blond	Marginal (Eye Color)
Brown	.11	.20	.04	.01	.37
Blue	.03	.14	.03	.16	.36
Hazel	.03	.09	.02	.02	.16
Green	.01	.05	.02	.03	.11
Marginal (Hair Color)	.18	.48	.12	.21	1.0

Table 4.2: Example of conditional probability. Of the blue-eyed people in Table 4.1, what proportion have hair color h? Each cell shows $p(h|\text{blue}) = p(\text{blue}, h)/p(\text{blue})$ rounded to two decimal points. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

	Hair Color				Marginal
Eye Color	Black	Brunette	Red	Blond	(Eye Color)
Blue	.03/.36 = .08	.14/.36 = .39	.03/.36 = .08	.16/.36 = .45	.36/.36 = 1.0

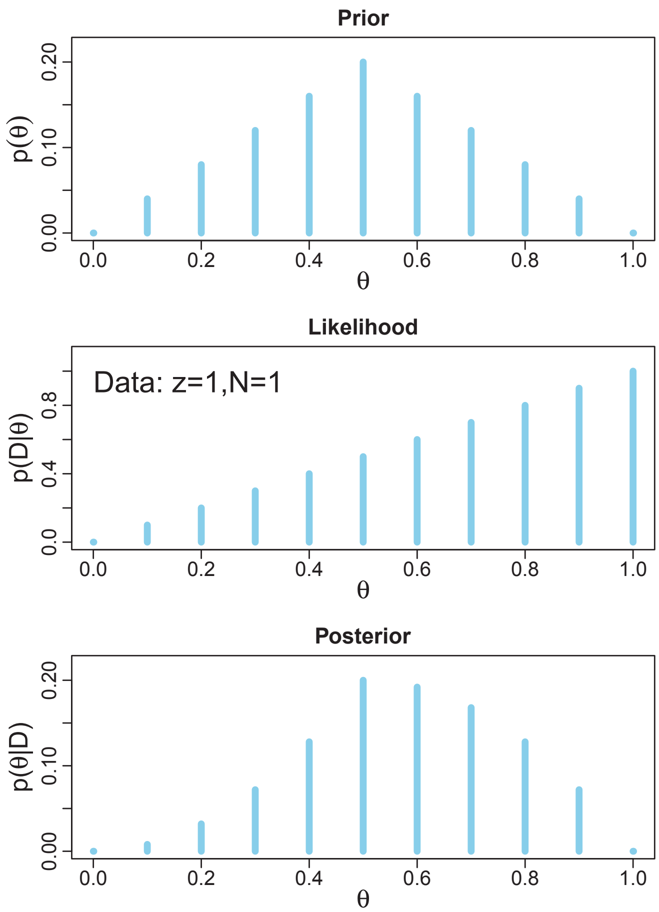

Figure 5.1: Bayes' rule applied to estimating the bias of a coin. There are discrete candidate values of θ. At each value of θ, the posterior is computed as prior times likelihood, normalized. In the data, denoted D, the number of heads is z and the number of flips is N.

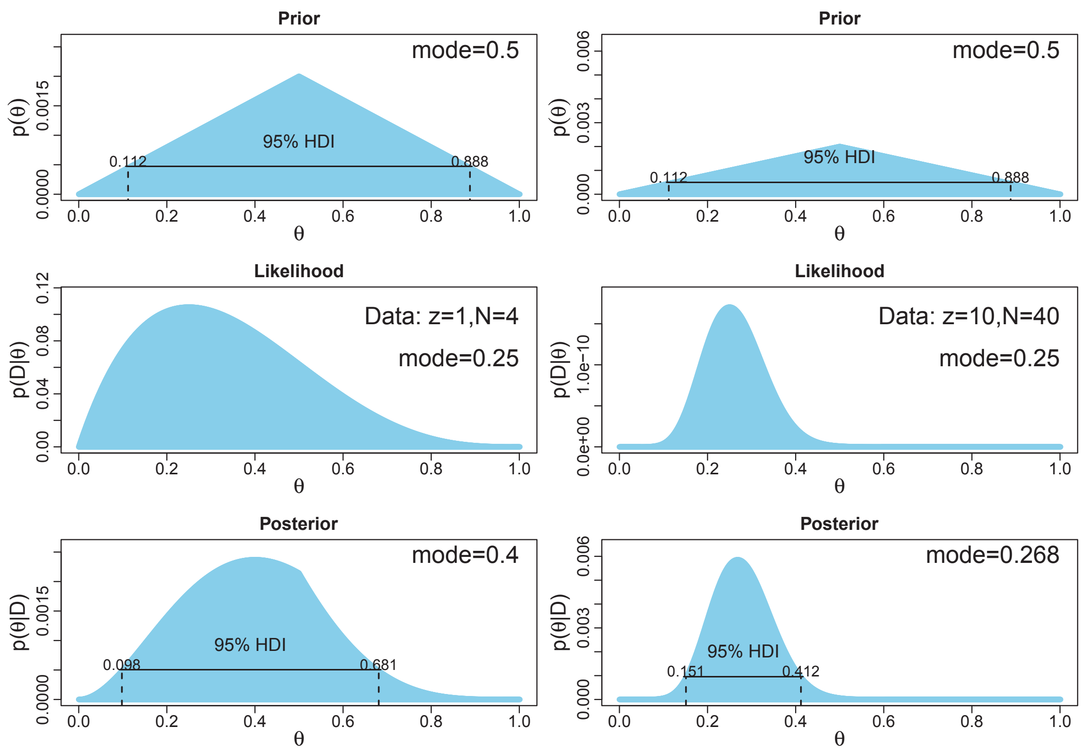

Figure 5.2: The two columns show different sample sizes with the same proportion of heads. The prior is the same in both columns but plotted on a different vertical scale. The influence of the prior is overwhelmed by larger samples, in that the peak of the posterior is closer to the peak of the likelihood function. Notice also that the posterior HDI is narrower for the larger sample. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

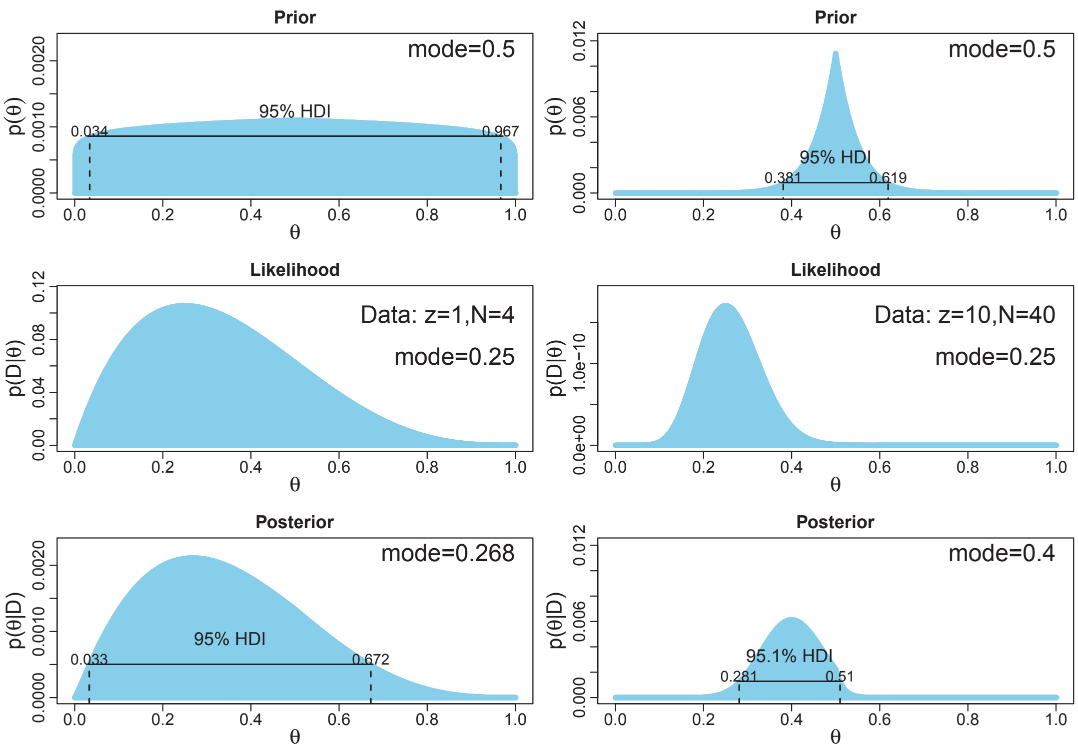

Figure 5.3: The left side is the same small sample as the left side of Figure 5.2 but with a flatter prior. The right side is the same larger sample as the right side of Figure 5.2 but with a sharper prior. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Table 5.1: A table for making Bayes' rule not merely special but spatial. When conditionalizing on row value r, the conditional probability $p(c|r)$ is simply the cell probability, $p(r, c)$, divided by the marginal probability, $p(r)$. When algebraically re-expressed as shown in the table, this is Bayes' rule. Spatially, Bayes' rule gets us from the lower marginal distribution, $p(c)$, to the conditional distribution $p(c|r)$ when focusing on row value r. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

	Column			
Row	$\ldots$	c	$\ldots$	**Marginal**
$\vdots$		$\vdots$		
r	$\ldots$	$p(r, c)$ $= p(r\|c)\, p(c)$	$\ldots$	$p(r)$ $= \sum_{c^*} p(r\|c^*)\, p(c^*)$
$\vdots$		$\vdots$		
Marginal		$p(c)$		

Table 5.2: Proportions of combinations of hair color and eye color. Some rows or columns may not sum exactly to their displayed marginals because of rounding error from the original data. Data adapted from Snee (1974). This is a Table 4.1 duplicated here for convenience. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

	Hair Color				Marginal
Eye Color	Black	Brunette	Red	Blond	(Eye Color)
Brown	.11	.20	.04	.01	.37
Blue	.03	.14	.03	.16	.36
Hazel	.03	.09	.02	.02	.16
Green	.01	.05	.02	.03	.11
Marginal (Hair Color)	.18	.48	.12	.21	1.0

Table 5.3: Example of conditional probability. Of the blue-eyed people in Table 5.2, what proportion have hair color h? Each cell shows $p(h|\text{blue}) = p(\text{blue}, h)/p(\text{blue})$ rounded to two decimal points. This is a Table 4.2 duplicated here for convenience. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

	Hair Color				Marginal
Eye Color	Black	Brunette	Red	Blond	(Eye Color)
Blue	.03/.36 = .08	.14/.36 = .39	.03/.36 = .08	.16/.36 = .45	.36/.36 = 1.0

Table 5.4: Joint and marginal probabilities of test results and disease states. For this example, the base rate of the disease is .001, as shown in the lower marginal. The test has a hit rate of .99 and a false alarm rate of .05, as shown in the row for $T = +$. For an actual test result, we restrict attention to the corresponding row of the table and compute the conditional probabilities of the disease states via Bayes' rule. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Test Result	Disease		Marginal (Test Result)
	$\theta =$ ☹ (present)	$\theta =$ ☺ (absent)	
$T = +$	$p(+\|$☹$)\,p($☹$)$ $= .99 \cdot .001$	$p(+\|$☺$)\,p($☺$)$ $= .05 \cdot (1 - .001)$	$\sum_\theta p(+\|\theta)\,p(\theta)$
$T = -$	$p(-\|$☹$)\,p($☹$)$ $= (1 - .99) \cdot .001$	$p(-\|$☺$)\,p($☺$)$ $= (1 - .05) \cdot (1 - .001)$	$\sum_\theta p(-\|\theta)\,p(\theta)$
Marginal (Disease)	$p($☹$) = .001$	$p($☺$) = 1 - .001$	1.0

Table 5.5: Applying Bayes' rule to data and parameters. When conditionalizing on row value D, the conditional probability $p(\theta|D)$ is the cell probability $p(D,\theta)$ divided by the marginal probability $p(D)$. When these probabilities are algebraically re-expressed as shown in the table, this is Bayes' rule. This table is merely Table 5.1 with its rows and columns re-named.

Data	**Model Parameter**			Marginal
	...	θ value	...	
$\vdots$		$\vdots$		$\vdots$
D value	...	$p(D,\theta)$ $= p(D\|\theta)\ p(\theta)$	...	$p(D)$ $= \sum_{\theta^*} p(D\|\theta^*)\ p(\theta^*)$
$\vdots$		$\vdots$		$\vdots$
Marginal	...	$p(\theta)$	...	

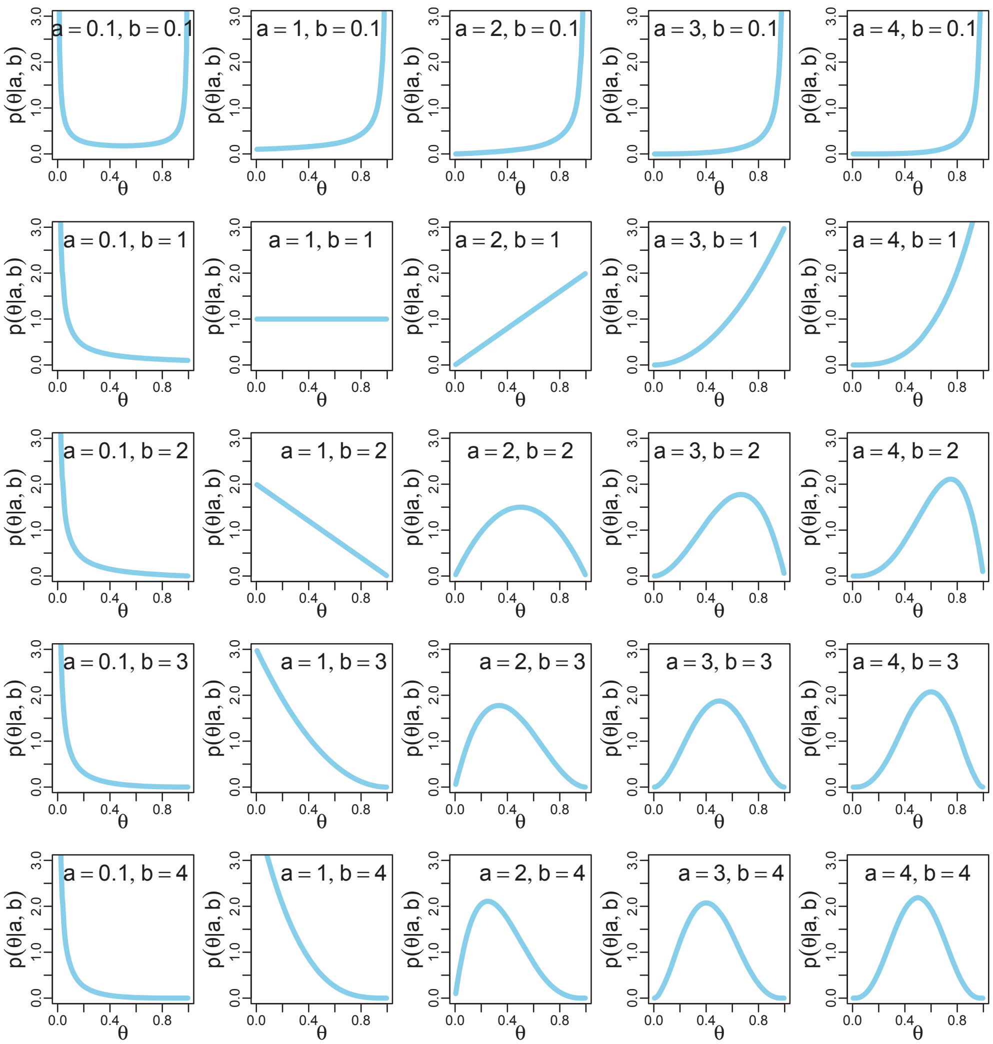

Figure 6.1: Examples of the beta distribution (Eqn. 6.1). The shape parameter a increases from left to right across the columns, while the shape parameter b increases from top to bottom across the rows. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

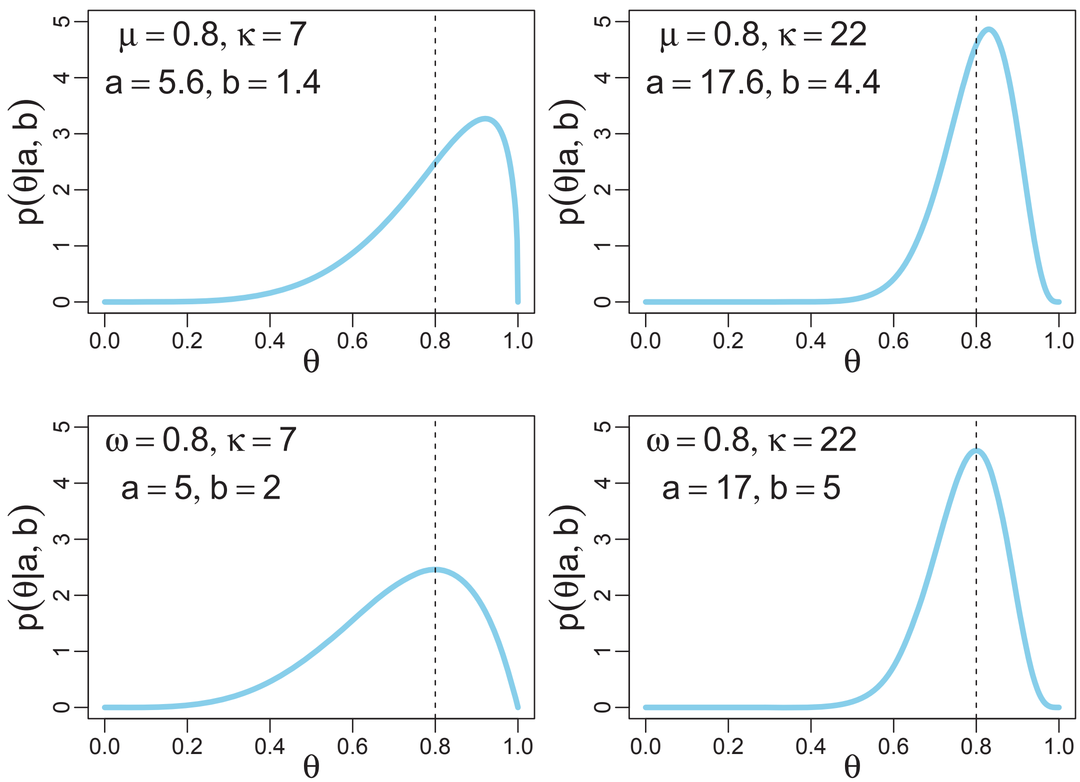

Figure 6.2: Beta distributions with a *mean* of $\mu = 0.8$ in the upper panels and a *mode* of $\omega = 0.8$ in the lower panels. Because the beta distribution is usually skewed, it can be more intuitive to think in terms of its mode instead of its mean. When κ is smaller, as in the left column, the beta distribution is wider than when κ is larger, as in the right column. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

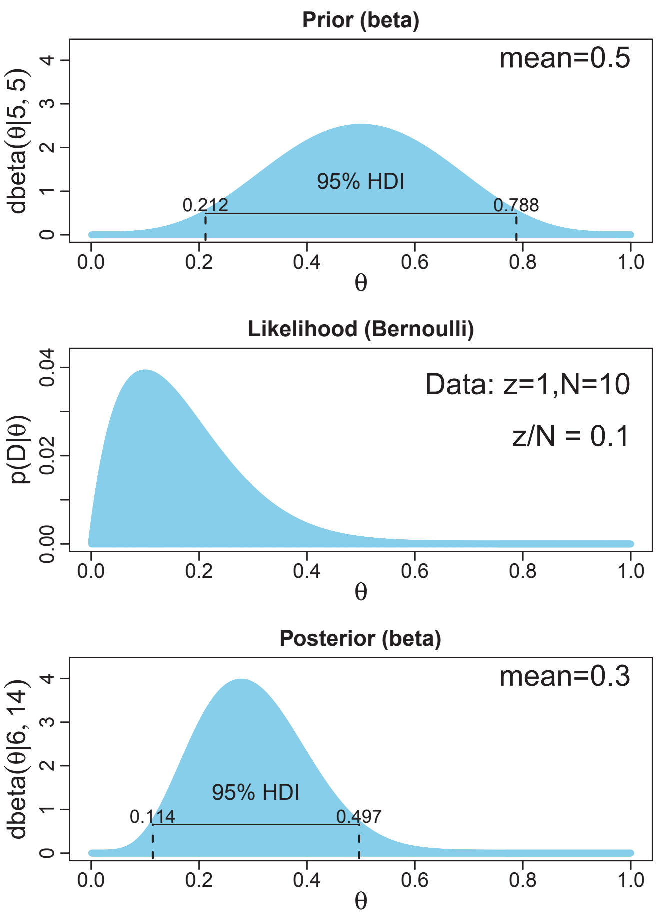

Figure 6.3: An illustration of Equation 6.9, showing that the mean of the posterior is a weighted combination of the mean of the prior and the proportion of heads in the data.

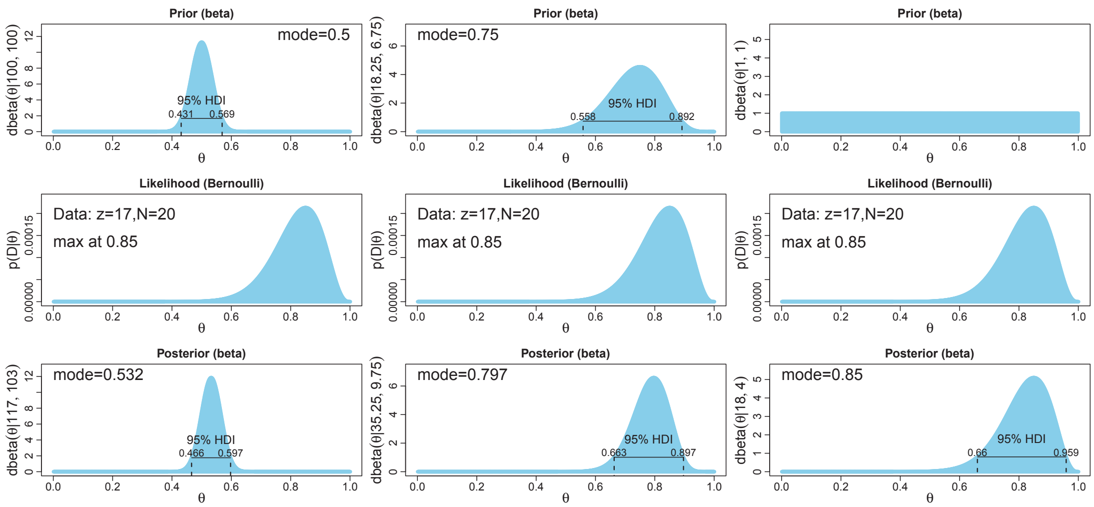

Figure 6.4: Examples of updating a beta prior distribution. The three columns show the same data with different priors. R code for this figure is described in Section 6.6. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

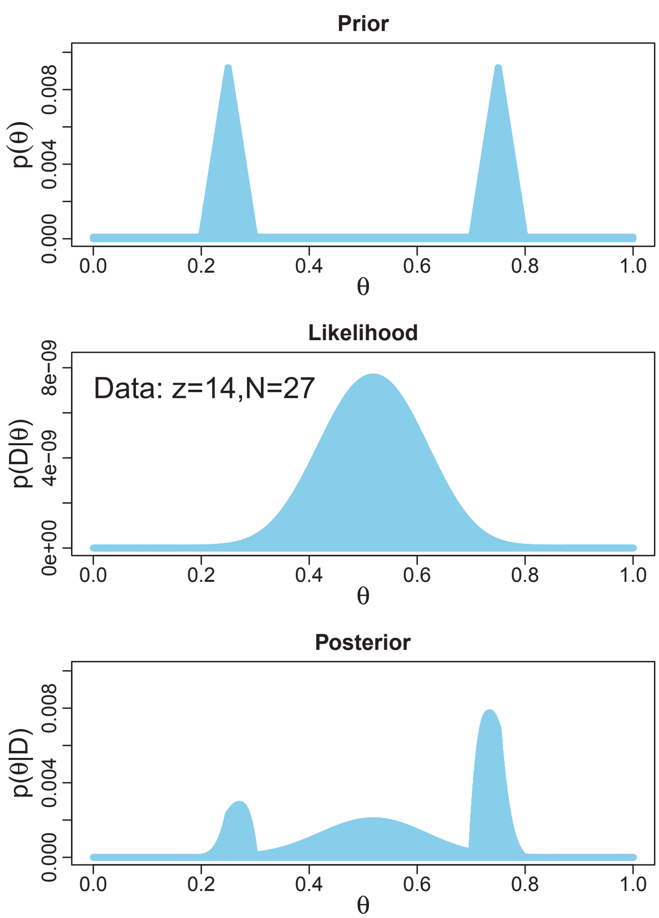

Figure 6.5: An example for which the prior distribution cannot be expressed by a beta distribution. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

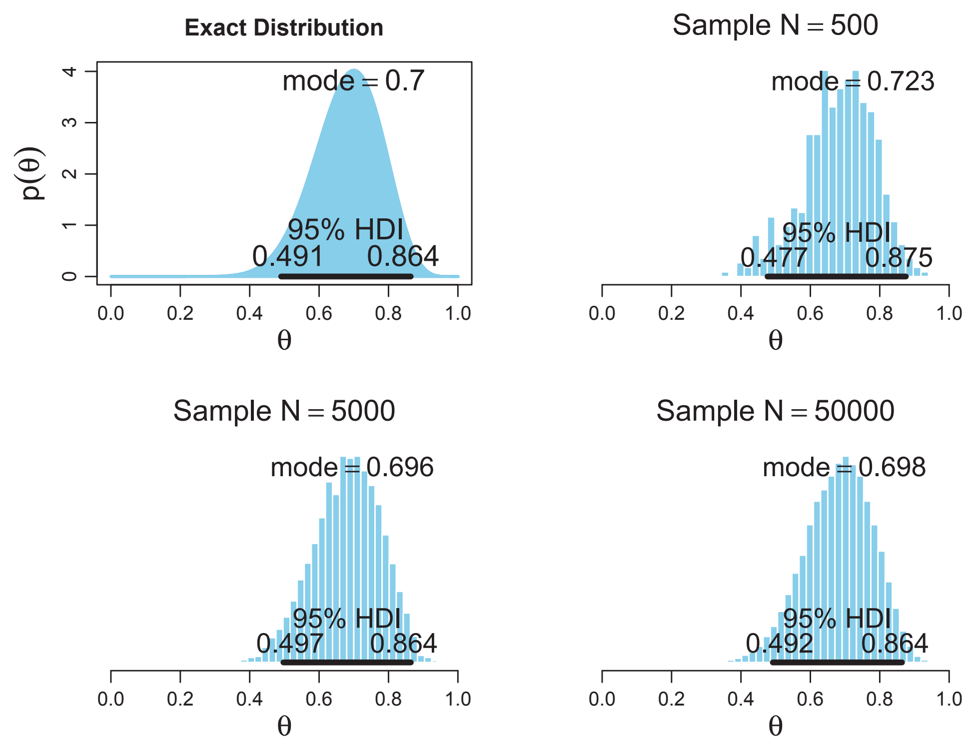

Figure 7.1: Large representative samples approximate the continuous distribution in the upper left panel. The larger the sample, the more accurate the approximation.(This happens to be a beta(θ|15, 7) distribution.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

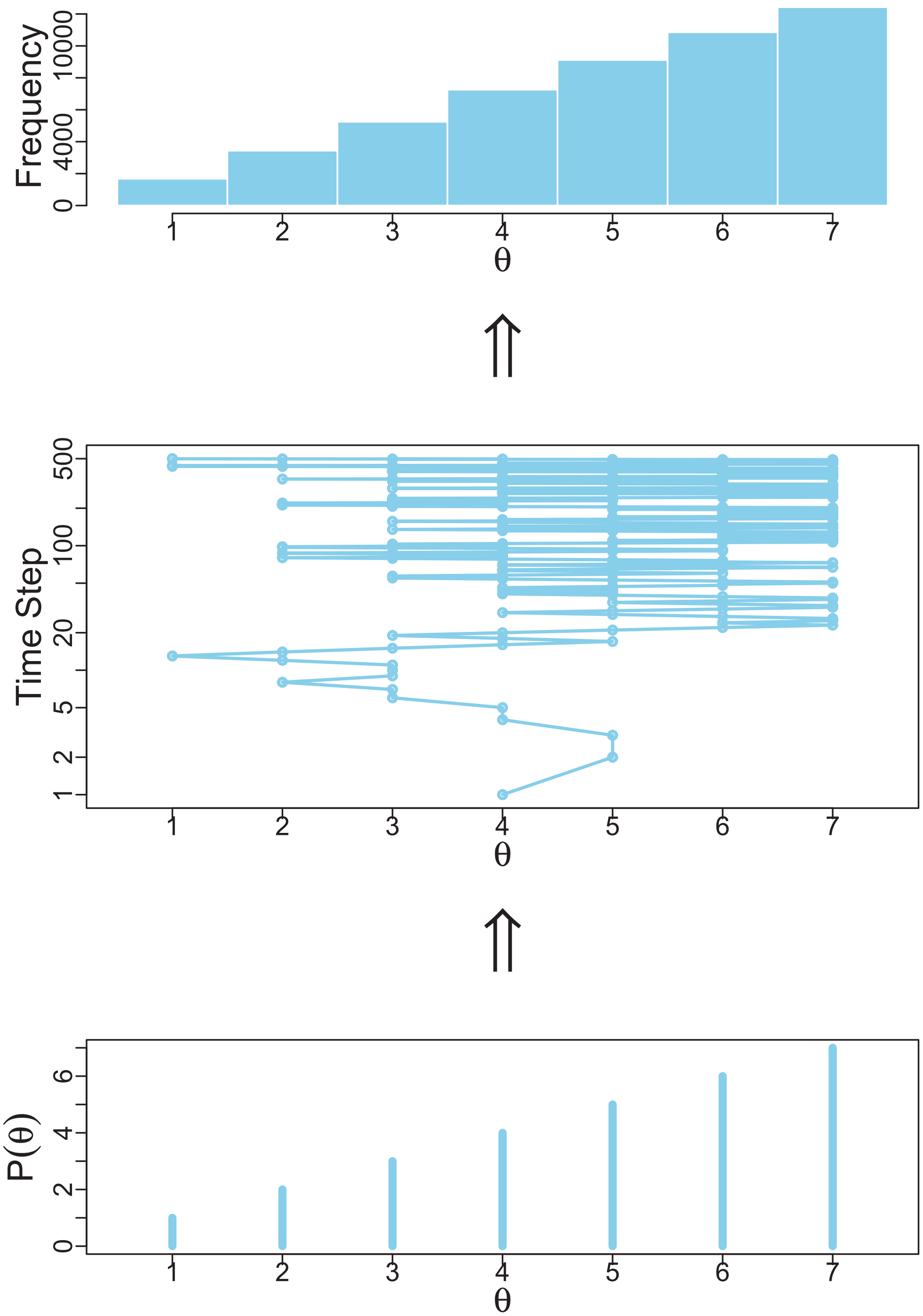

Figure 7.2: Illustration of a simple Metropolis algorithm. The bottom panel shows the values of the target distribution. The middle panel shows one random walk, at each time step proposing to move either one unit right or one unit left, and accepting the proposed move according the heuristic described in the main text. The top panel shows the frequency distribution of the positions in the walk. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

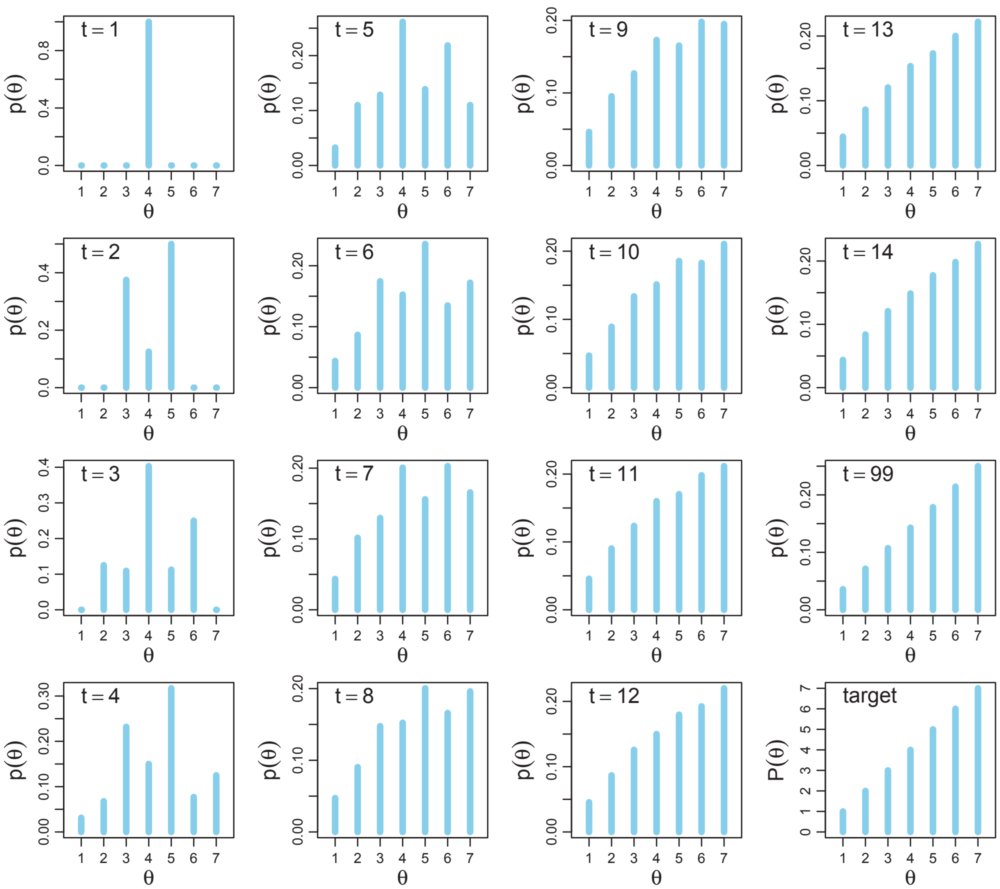

Figure 7.3: The probability of being at position θ, as a function of time t, when a simple Metropolis algorithm is applied to the target distribution in the lower right panel. The time in each panel corresponds to the step in a random walk, an example of which is shown in Figure 7.2. The target distribution is shown in the lower right panel.

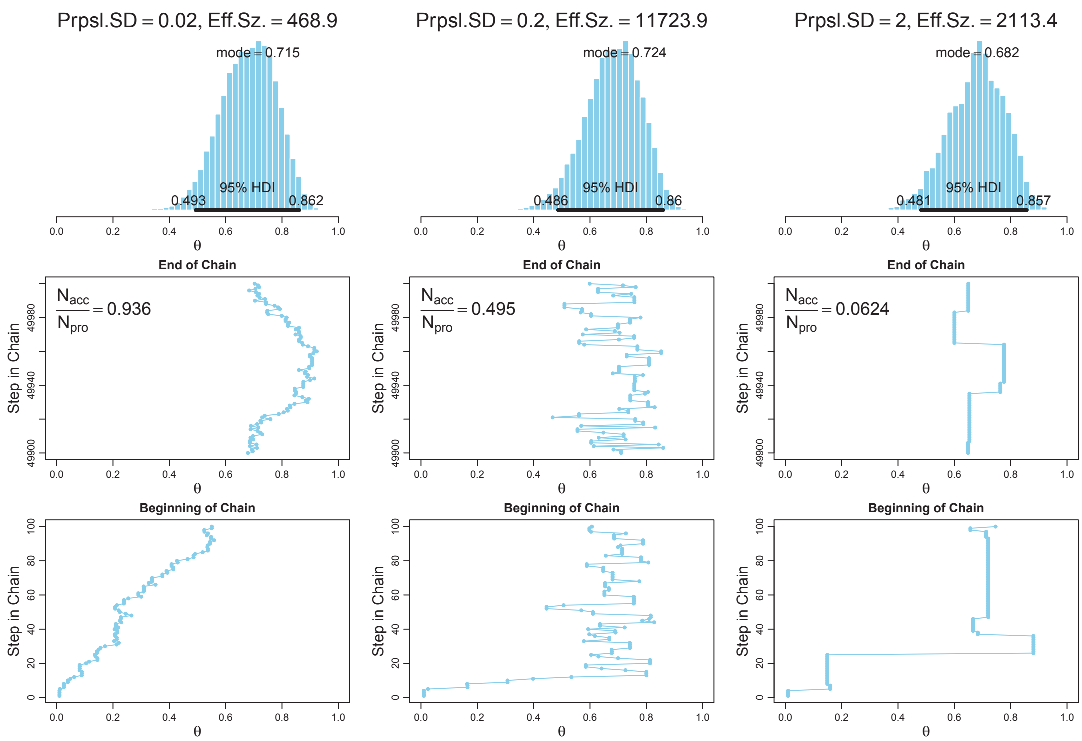

Figure 7.4: Metropolis algorithm applied to Bernoulli likelihood with beta(θ|1, 1) prior and $z = 14$ with $N = 20$. For each of the three columns, there are 50,000 steps in the chain, but, for the left column the proposal standard deviation (SD) is 0.02, for the middle column SD=0.2, and for the right column SD=2.0. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

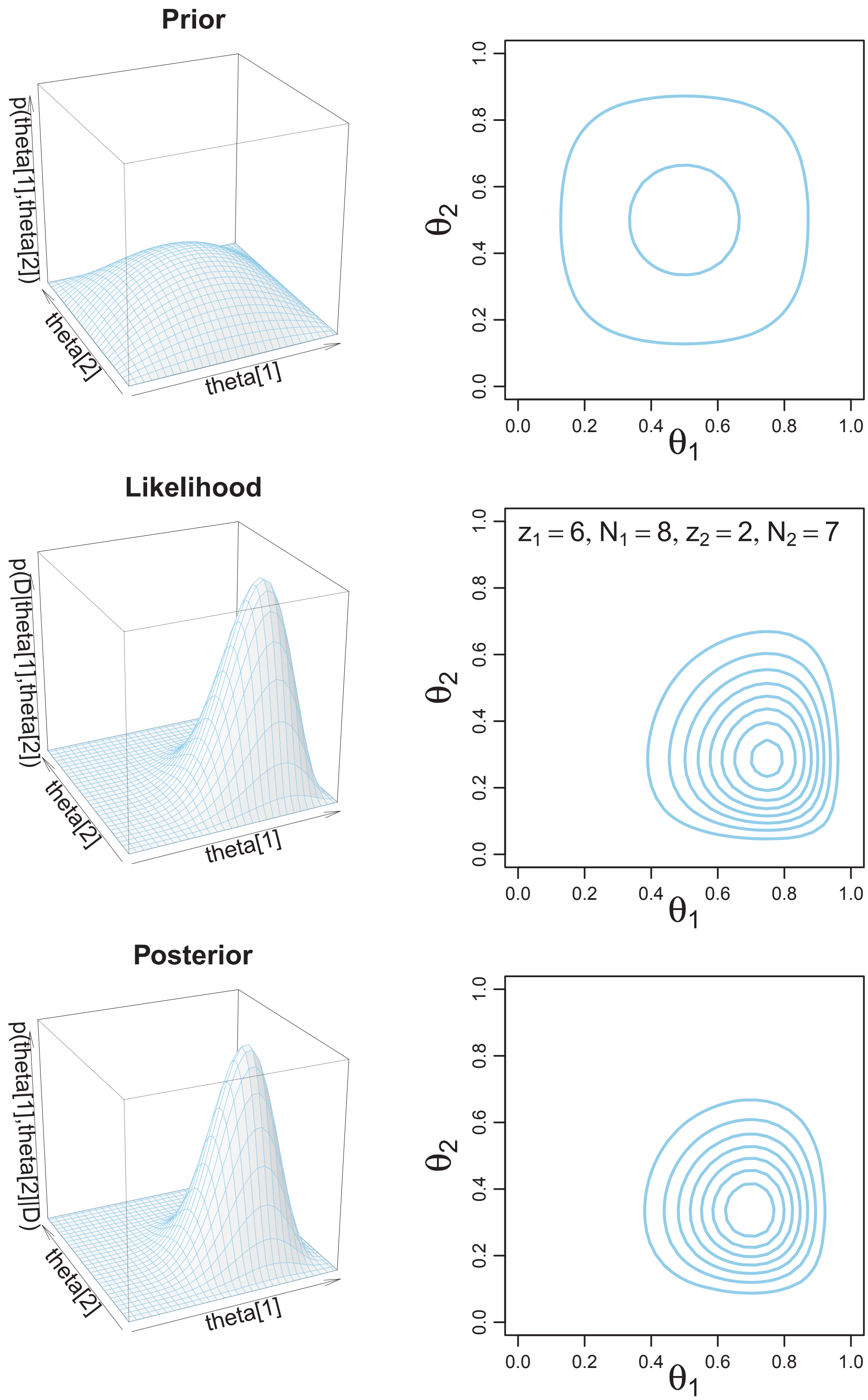

Figure 7.5: Bayesian updating of independent beta($\theta|2,2$) priors with the data annotated in the middle-right panel. Left panels show perspective surface plots; right panels show contour plots of the same distributions. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

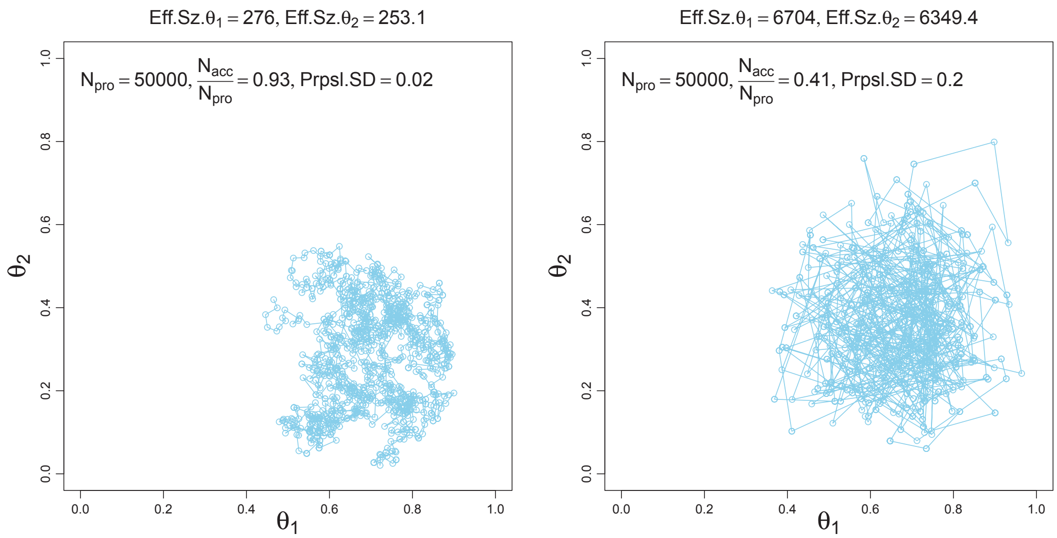

Figure 7.6: Metropolis algorithm applied to the prior and likelihood shown in Figure 7.5, p. 163. Left panel shows chain with narrow proposal distribution and right panel shows chain with moderate-width proposal distribution, as indicated by annotation "Prpsl.SD" in each panel. N_{pro} is the number of proposed jumps, and N_{acc} is the number of accepted proposals. *Many of the plotted points have multiple superimposed symbols where the chain lingered during rejected proposals.* Notice that the effective size of the chain, indicated at the top of the plot, is far less than the length of the chain (N_{pro}). Only 1,000 of the N_{pro} steps are displayed here. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

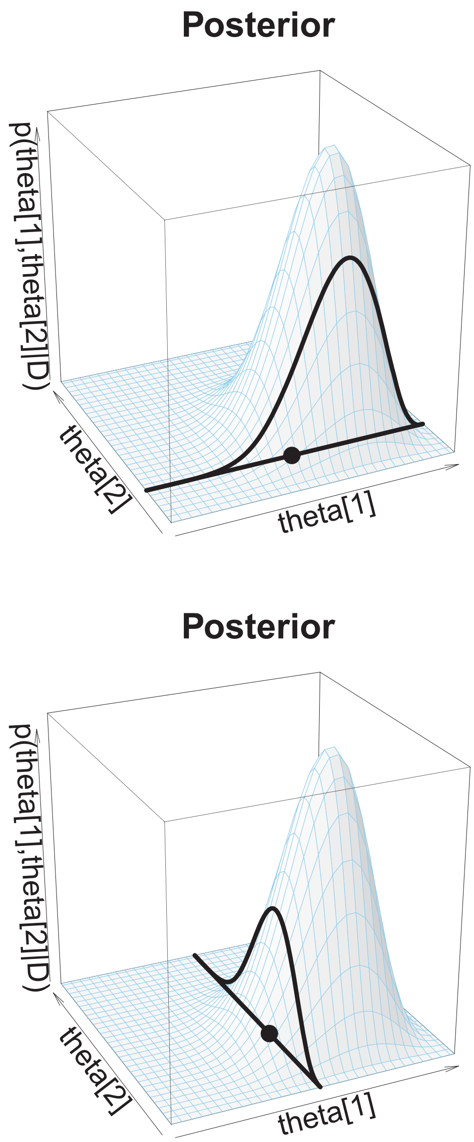

Figure 7.7: Two steps in a Gibbs sampling. In the upper panel, the heavy lines show a slice through the posterior conditionalized on a particular value of θ_2, and the large dot shows a random value of θ_1 sampled from the conditional density. The lower panel shows a random generation of a value for θ_2, conditional on the value for θ_1 determined by the previous step. The heavy lines show a slice through the posterior at the conditional value of θ_1, and the large dot shows the random value of θ_2 sampled from the conditional density. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

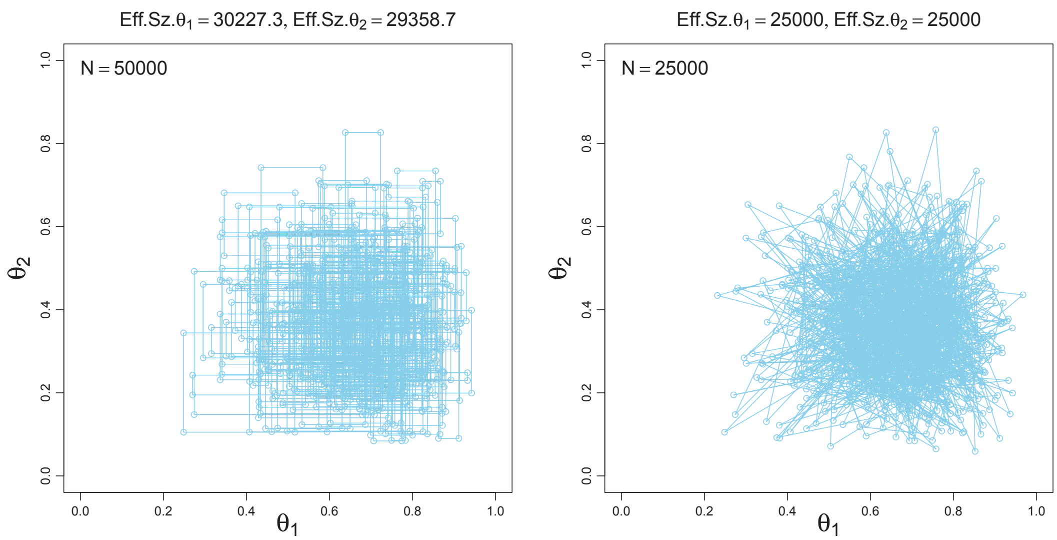

Figure 7.8: Gibbs sampling applied to the posterior shown in Figure 7.5, p. 163. The left panel shows all the intermediate steps of chain, changing one parameter at a time. The right panel shows only the points after complete sweeps through all (two) parameters. Both are valid samples from the posterior distribution. Only 1,000 of the N steps are displayed here. Compare with the results of the Metropolis algorithm in Figure 7.6. Notice that the effective size of the Gibbs sample is larger than the effective size of the Metropolis sample for the same length of chain. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

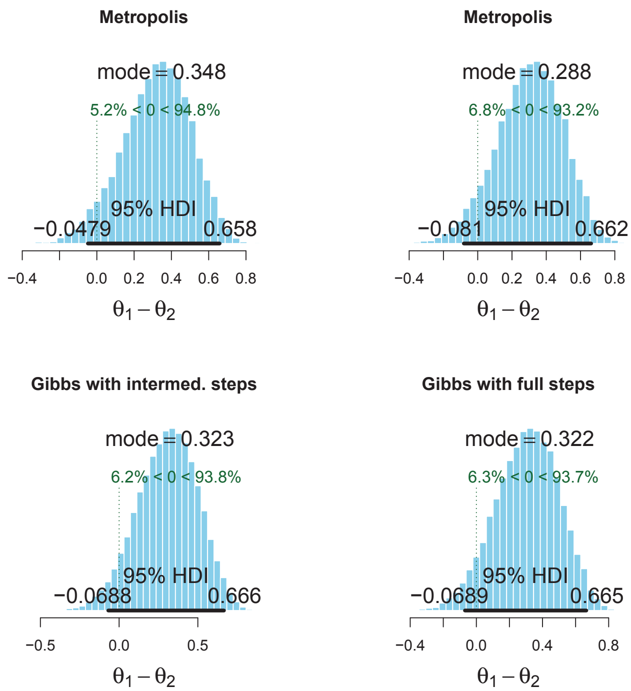

Figure 7.9: Credible posterior differences between biases. Upper panels comes from results of Metropolis algorithm in Figure 7.6; lower panels come from results of Gibbs sampling in Figure 7.8. The four distributions are nearly the same, and in the limit for infinitely long chains, should be identical. For these finite chains, the ones with longer effective size (i.e., the Gibbs sampled) are more accurate on average.

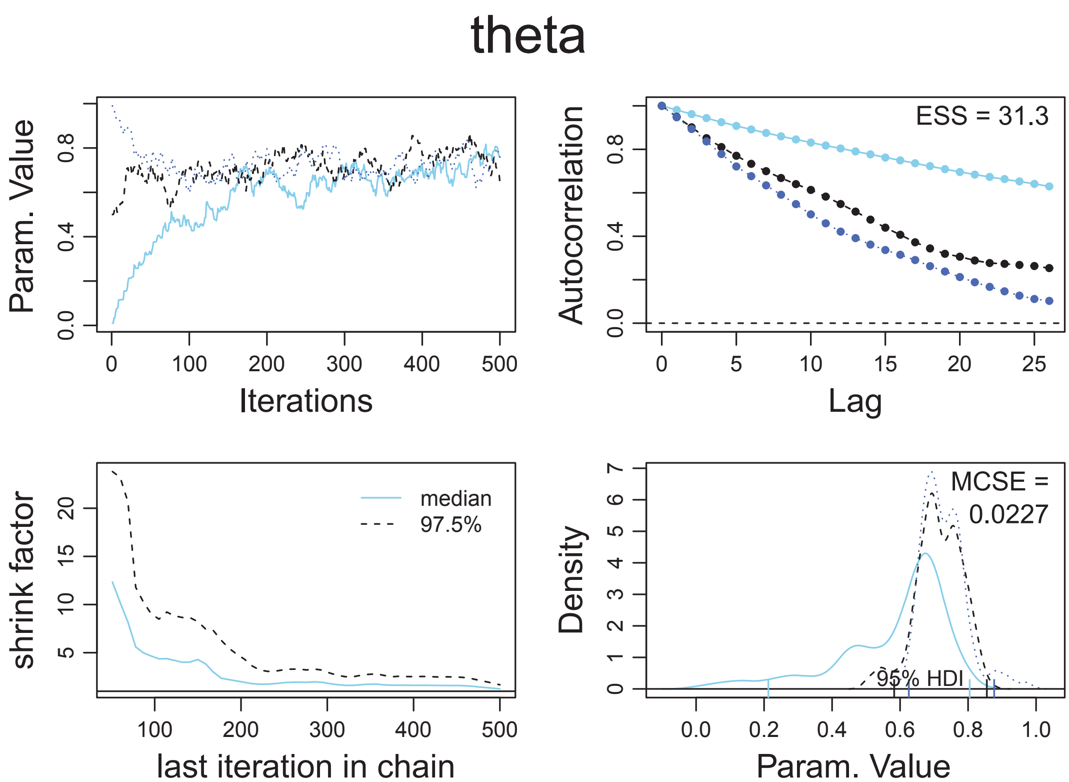

Figure 7.10: Illustration of MCMC diagnostics. Three chains were generated by starting a Metropolis algorithm at different initial values, with proposal SD=0.02 (cf. Fig. 7.4) for data $z = 35$, $N = 50$. Only steps 1–500 are shown here. See Fig. 7.11 for later steps in the chain.

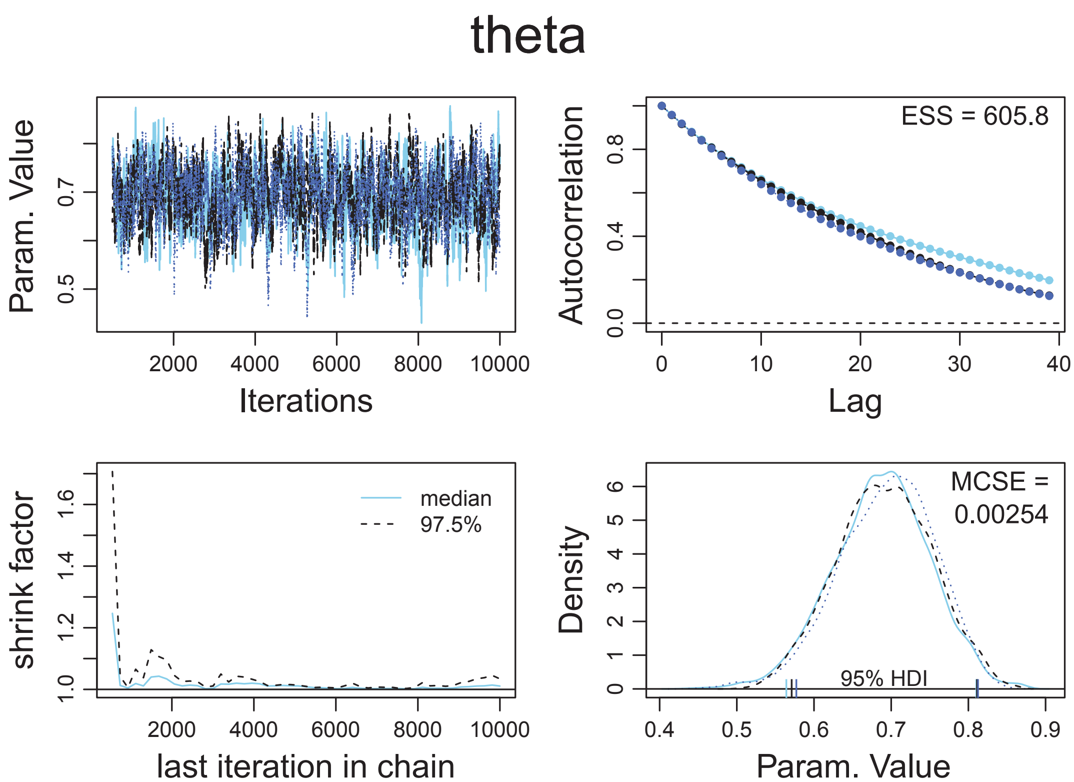

Figure 7.11: Illustration of MCMC diagnostics. Three chains were generated by starting a Metropolis algorithm at different initial values, with proposal SD=0.02 (cf. Fig. 7.4) for data $z = 35$, $N = 50$. Steps 500–10,000 are shown here. See Fig. 7.10 for earlier steps in the chain. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

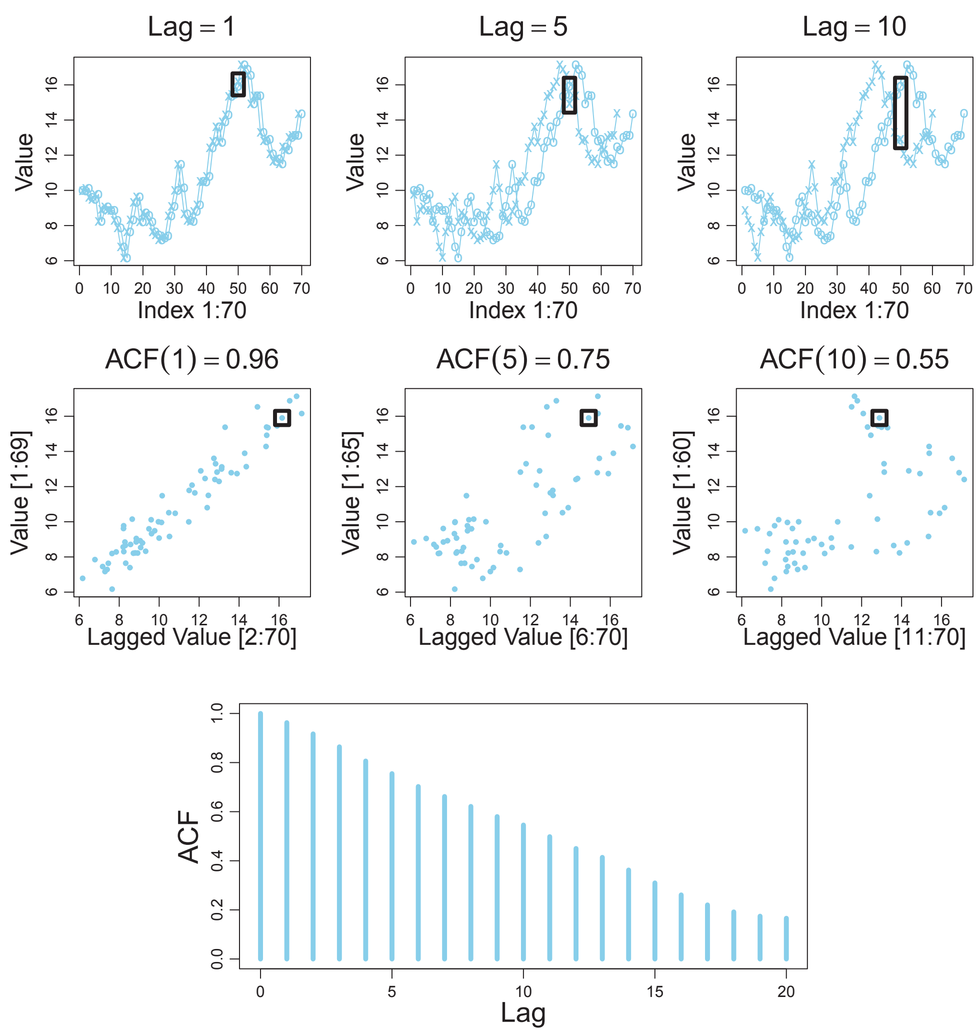

Figure 7.12: Autocorrelation of a chain. Upper panels show examples of lagged chains. Middle panels show scatter plots of chain values against lagged chain values, with their correlation annotated. Lowest panel shows the autocorrelation function (ACF). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

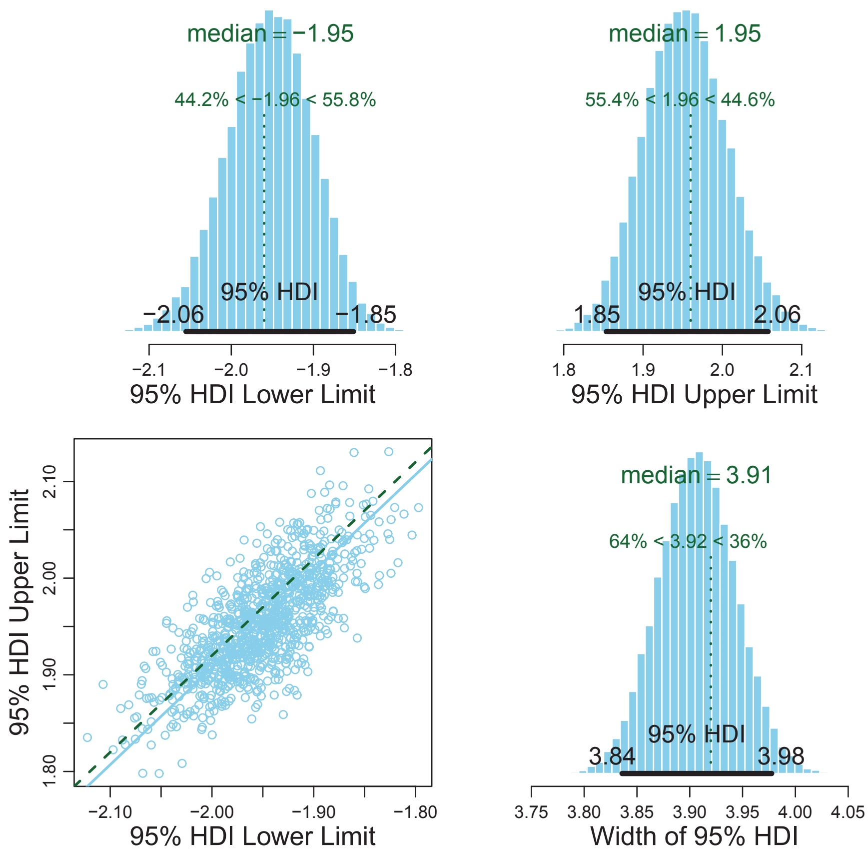

Figure 7.13: Estimated 95% HDI limits for random samples from a standardized normal distribution that have an *ESS of 10,000*. Repeated runs yield a distribution of estimates as shown here; there were 50,000 repetitions. Upper panels show estimate of HDI limits. Lower panels show estimate of HDI width. True values are indicated by dashed lines. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

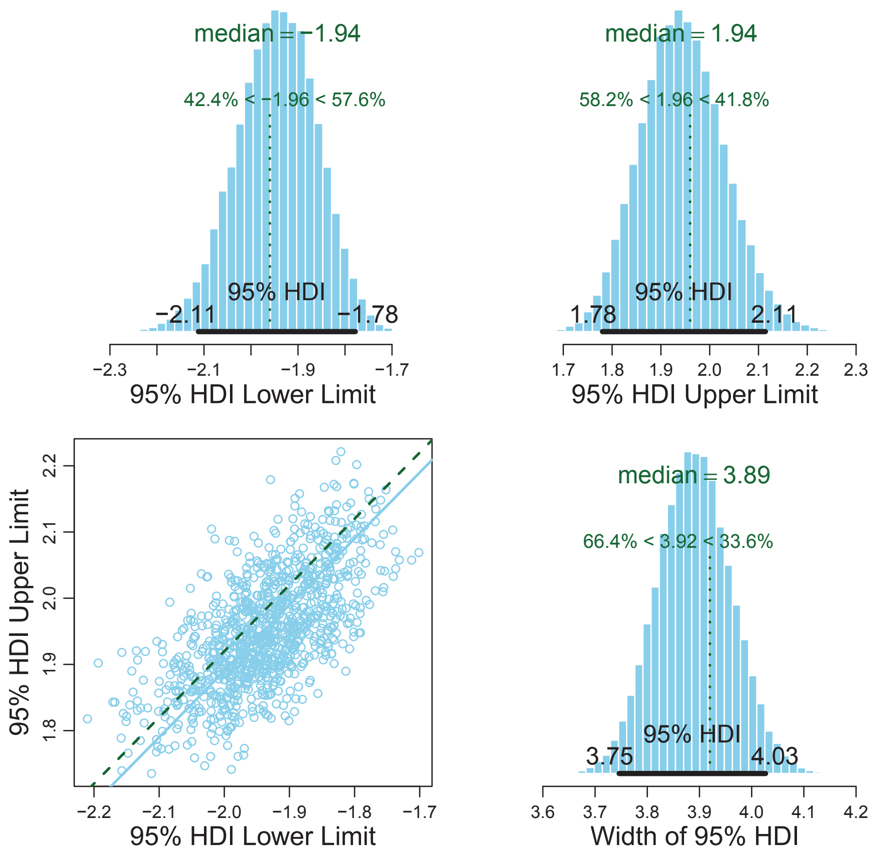

Figure 7.14: Estimated 95% HDI limits for random samples from a standardized normal distribution that have an *ESS of only 2,500*. Repeated runs yield a distribution of estimates as shown here; there were 50,000 repetitions. Upper panels show estimate of HDI limits. Lower panels show estimate of HDI width. True values are indicated by dashed lines. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

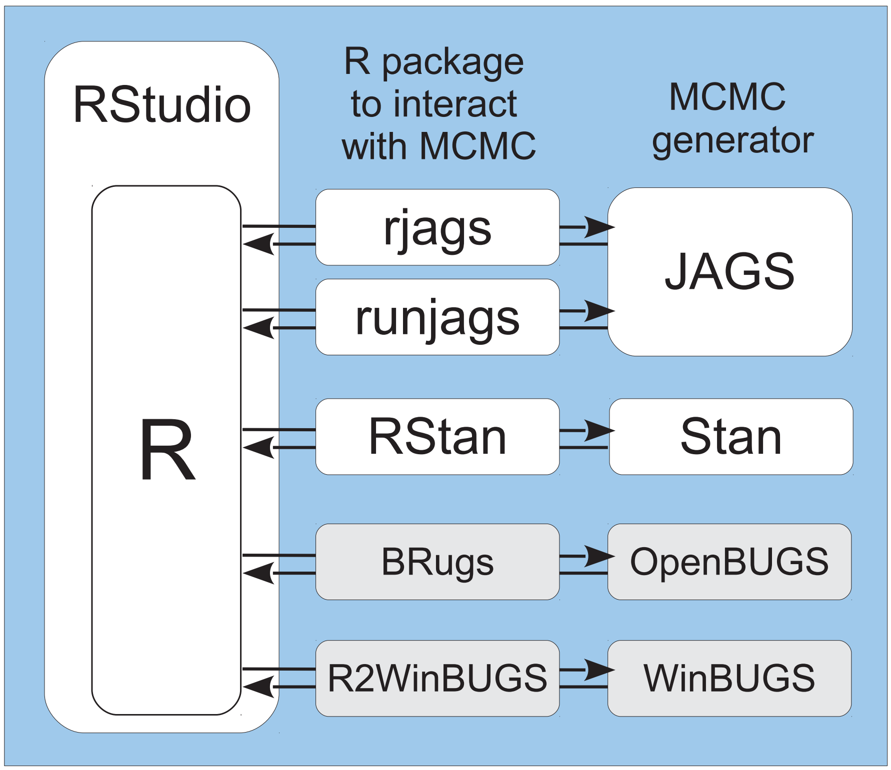

Figure 8.1: Relation of R programming language to other software tools. On the left, RStudio is an editor for interacting with R. The items on the right are various programs for generating MCMC samples of posterior distributions. The items in the middle are packages in R that interact with the MCMC generators. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

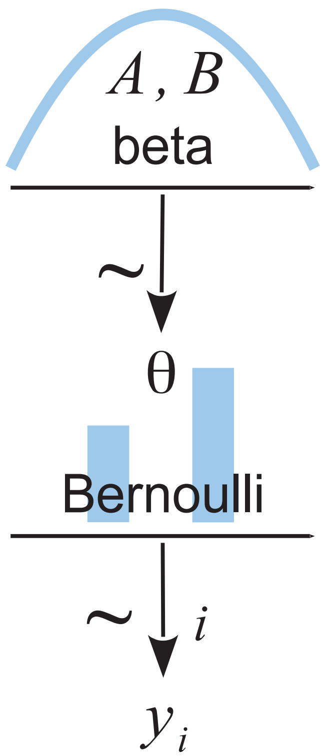

Figure 8.2: Diagram of model with Bernoulli likelihood and beta prior. The pictures of the distributions are intended as stereotypical icons, and are not meant to indicate the exact forms of the distributions. Diagrams like this should be scanned from the bottom up, starting with the data y_i and working upward through the likelihood function and prior distribution. Every arrow in the diagram has a corresponding line of code in JAGS model specification. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

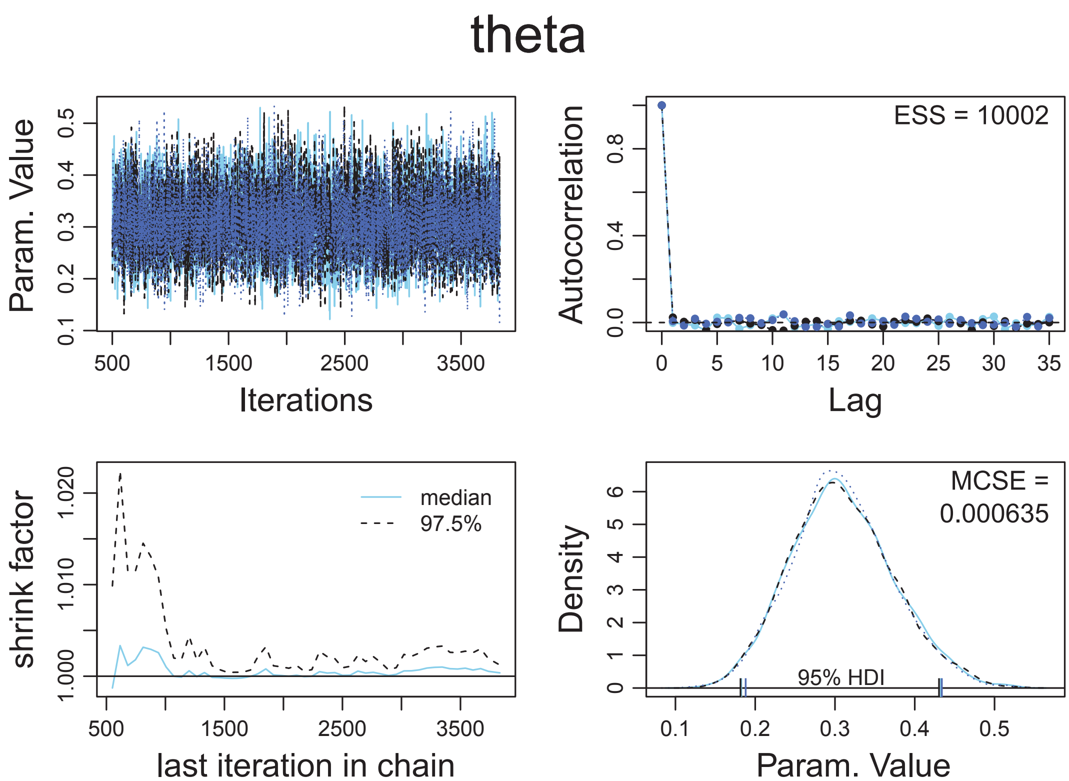

Figure 8.3: Convergence diagnostics for output of JAGS... Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

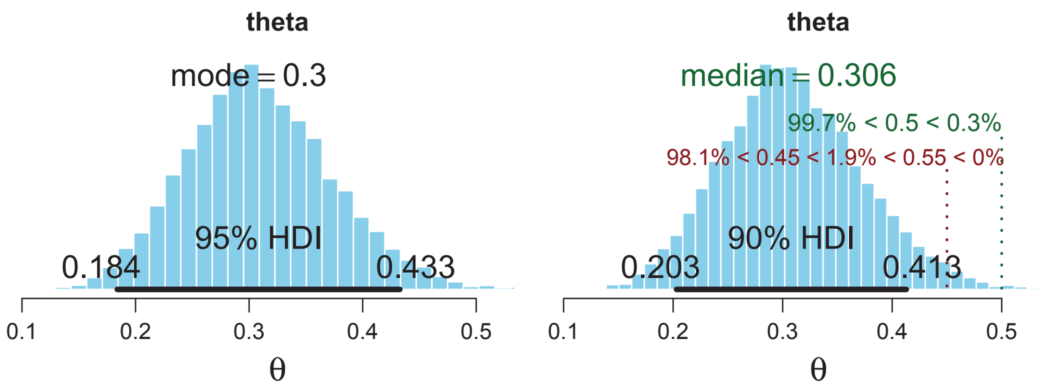

Figure 8.4: Posterior distribution based on output from JAGS, plotted twice using different options in the `plotPost` function. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

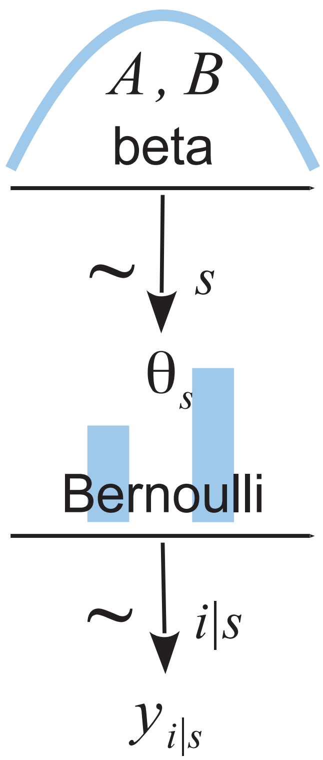

Figure 8.5: Diagram of model with Bernoulli likelihood and beta prior for multiple subjects, *s*. Notice the indices on the variables and arrows. Every arrow in the diagram has a corresponding line of code in JAGS model specification. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

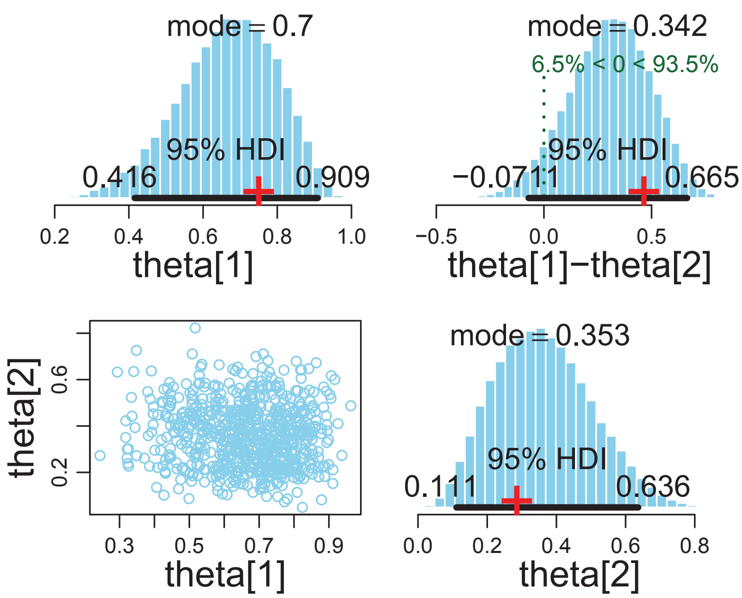

Figure 8.6: Posterior distribution created by JAGS. Compare the upper-right panel with Figure 7.9. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

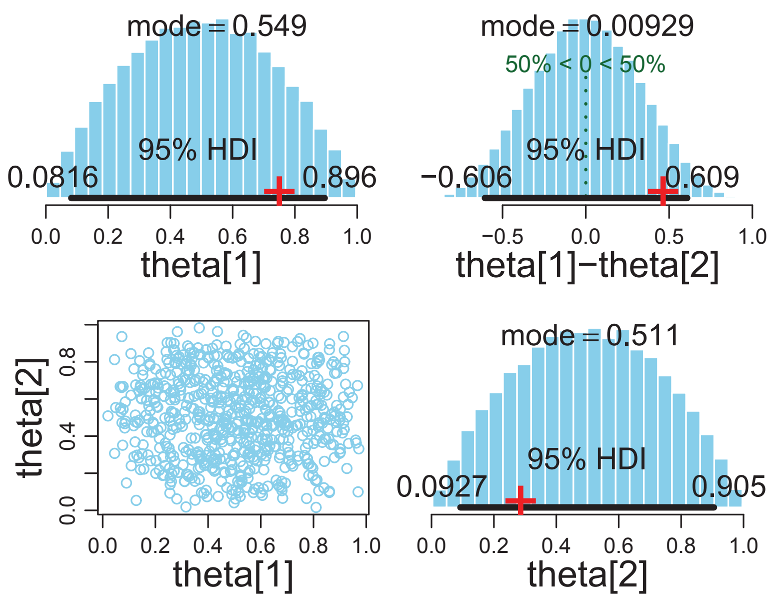

Figure 8.7: The prior distribution for difference of biases. Compare with Figure 8.6.

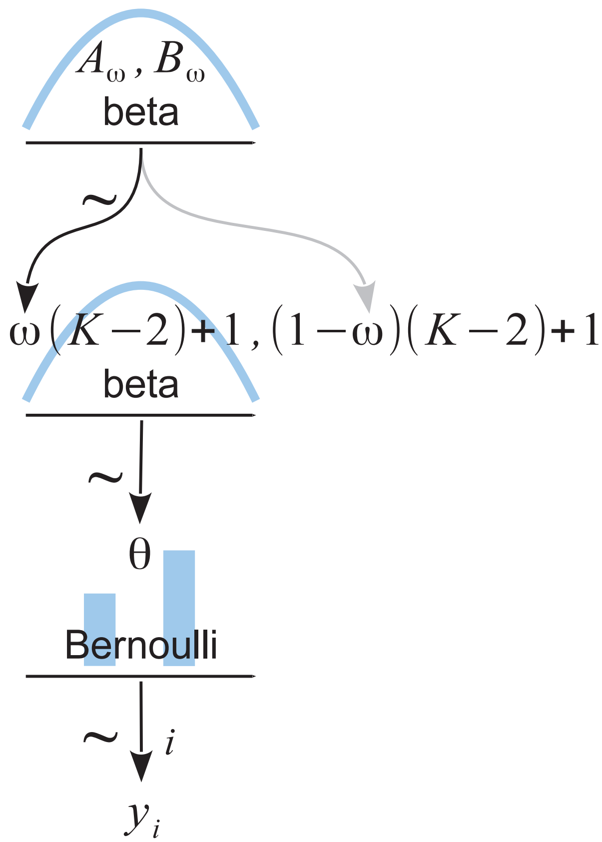

Figure 9.1: A model of hierarchical dependencies for data from a single coin. The chain of arrows illustrates the chain of dependencies in Equations 9.2, 9.4, and 9.5. (At the top of the diagram, the second instantiation of the arrow to ω is shown in grey instead of black merely to suggest that it is the same dependence already indicated by the first arrow in black.)

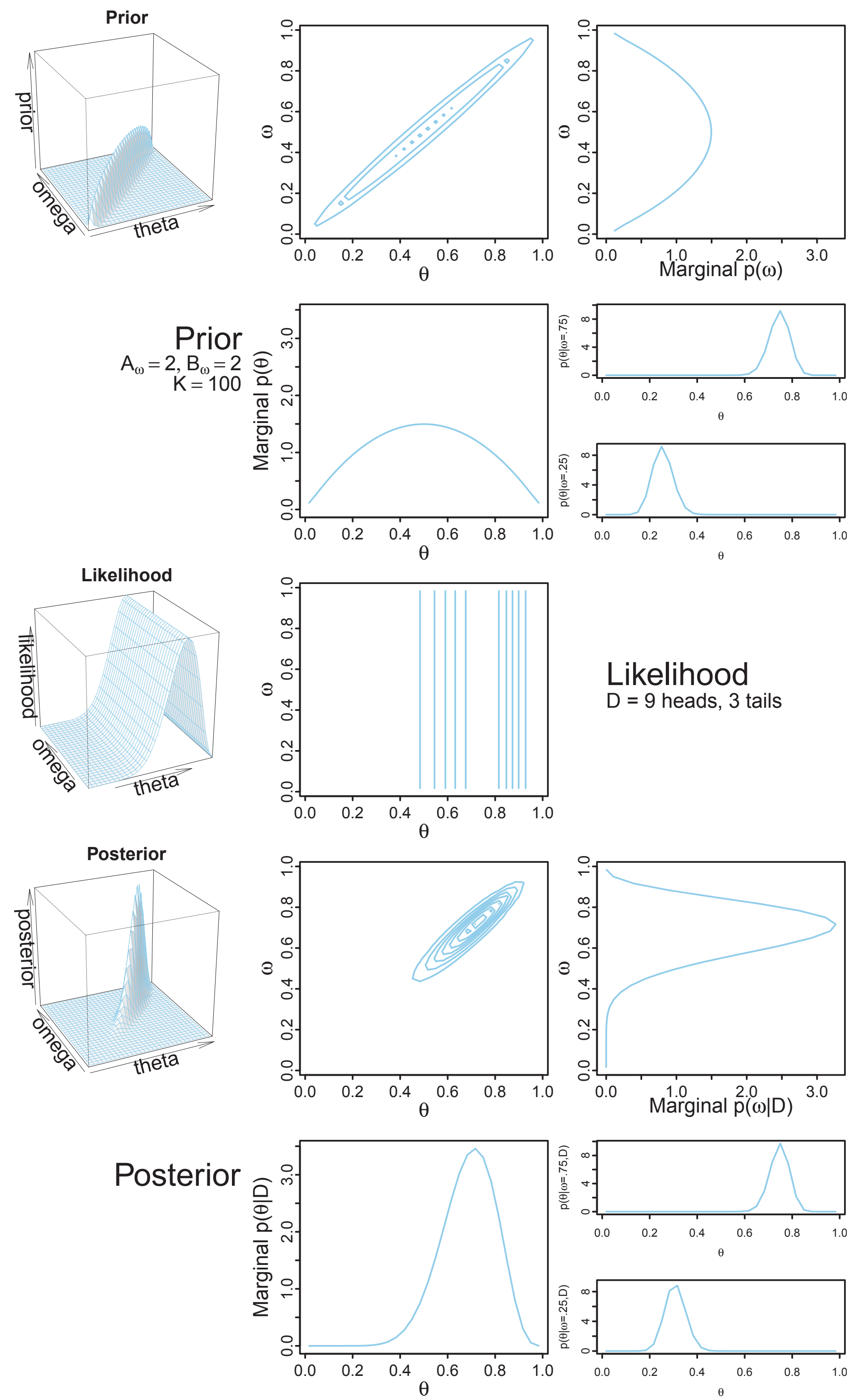

Figure 9.2: The prior has low certainty regarding ω, but high certainty regarding the dependence of θ on ω. The posterior shows that the distribution of ω has been altered noticeably by the data (see sideways plots of marginal $p(\omega)$), but the dependence of θ on ω has not been altered much (see small plots of $p(\theta|\omega)$). Compare with Figure 9.3, which uses the same data but a different prior. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

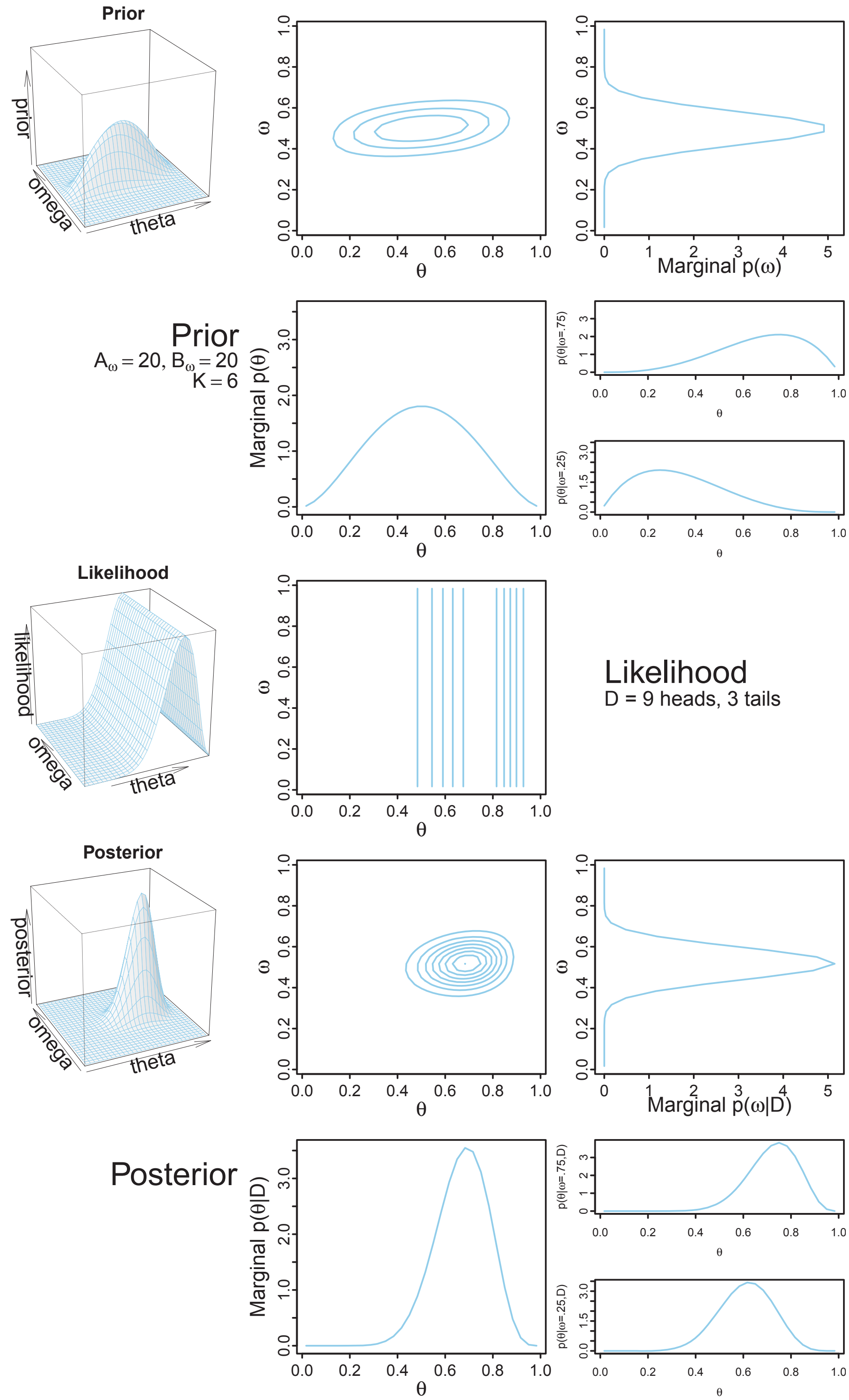

Figure 9.3: The prior has high certainty regarding ω, but low certainty regarding the dependence of θ on ω. The posterior shows that the distribution of ω has not been altered much by the data (see sideways plots of marginal $p(\omega)$), but the dependence of θ on ω has been altered noticeably (see small plots of $p(\theta|\omega)$). Compare with Figure 9.2, which uses the same data but a different prior. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

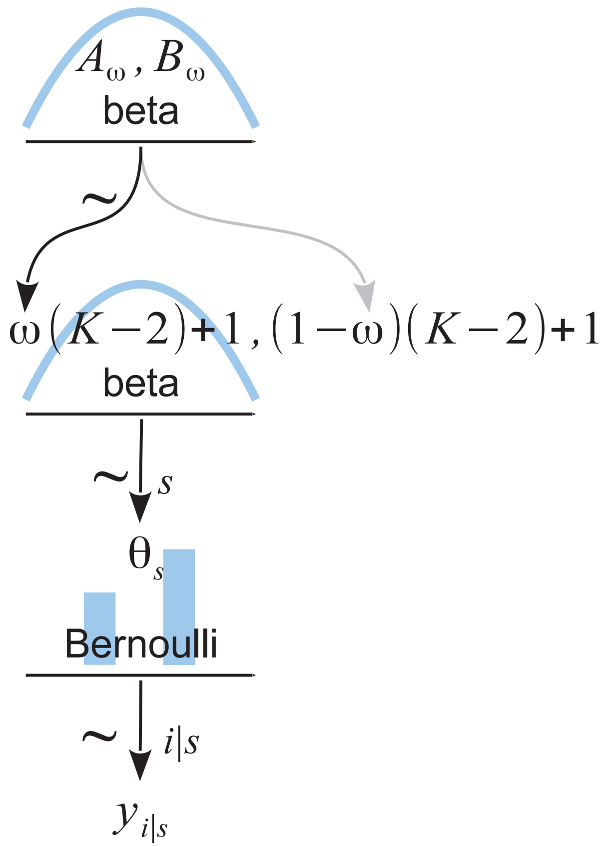

Figure 9.4: A model of hierarchical dependencies for data from several coins created independently from the same mint. A datum $y_{i|s}$, from the i^{th} flip of the s^{th} coin, depends on the value of the bias parameter θ_s for the coin. The values of θ_s depend on the value of the hyperparameter ω for the mint that created the coins. The ω parameter has a prior belief distributed as a beta distribution with shape parameters A_ω and B_ω. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

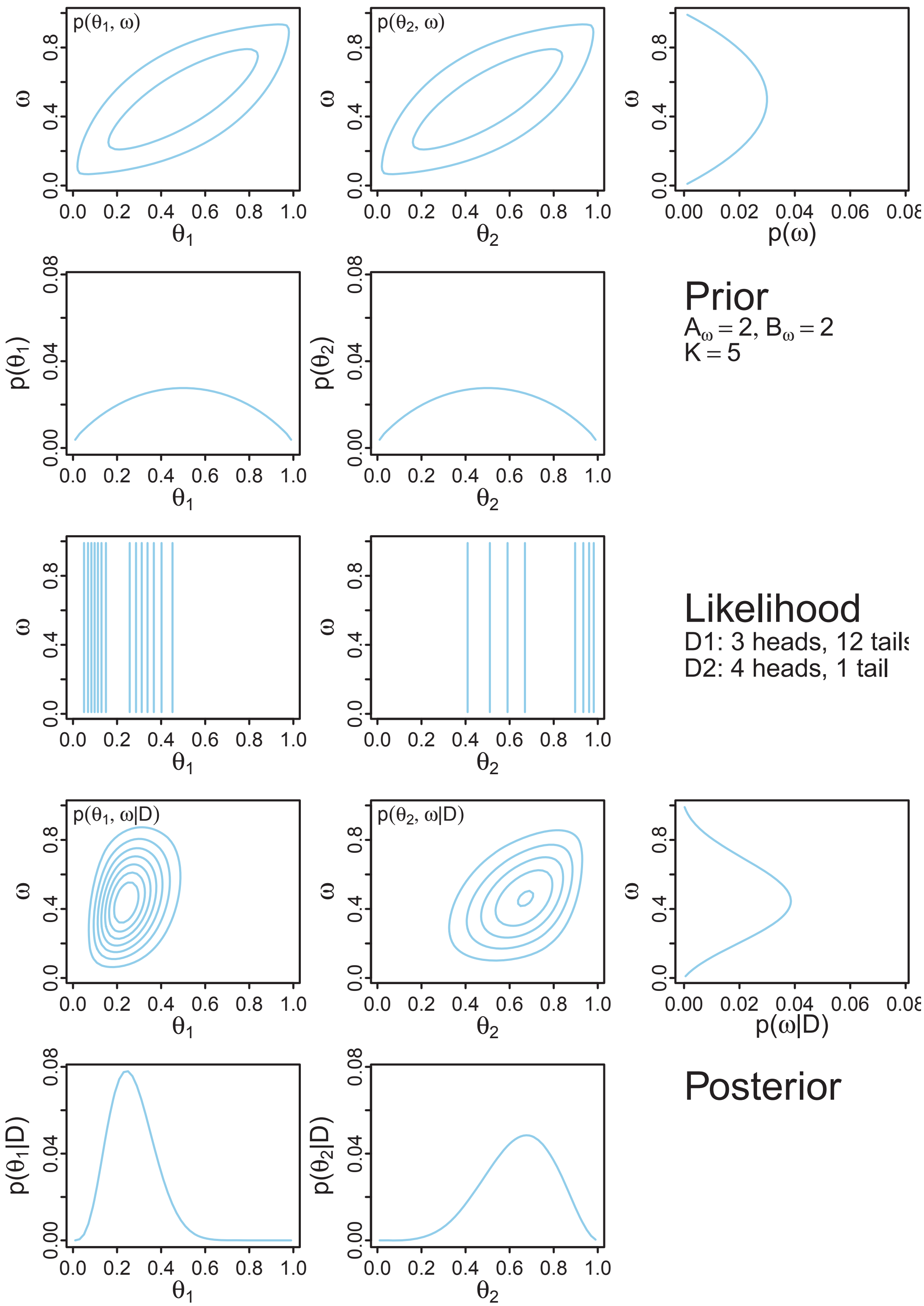

Figure 9.5: The prior imposes only a weak dependence of θ on μ (i.e., K is small), so the posteriors on θ_1 and θ_2 (bottom row) are only weakly influenced by each other's data. Compare with Figure 9.6, which uses the same data but a prior that has a strong dependence.

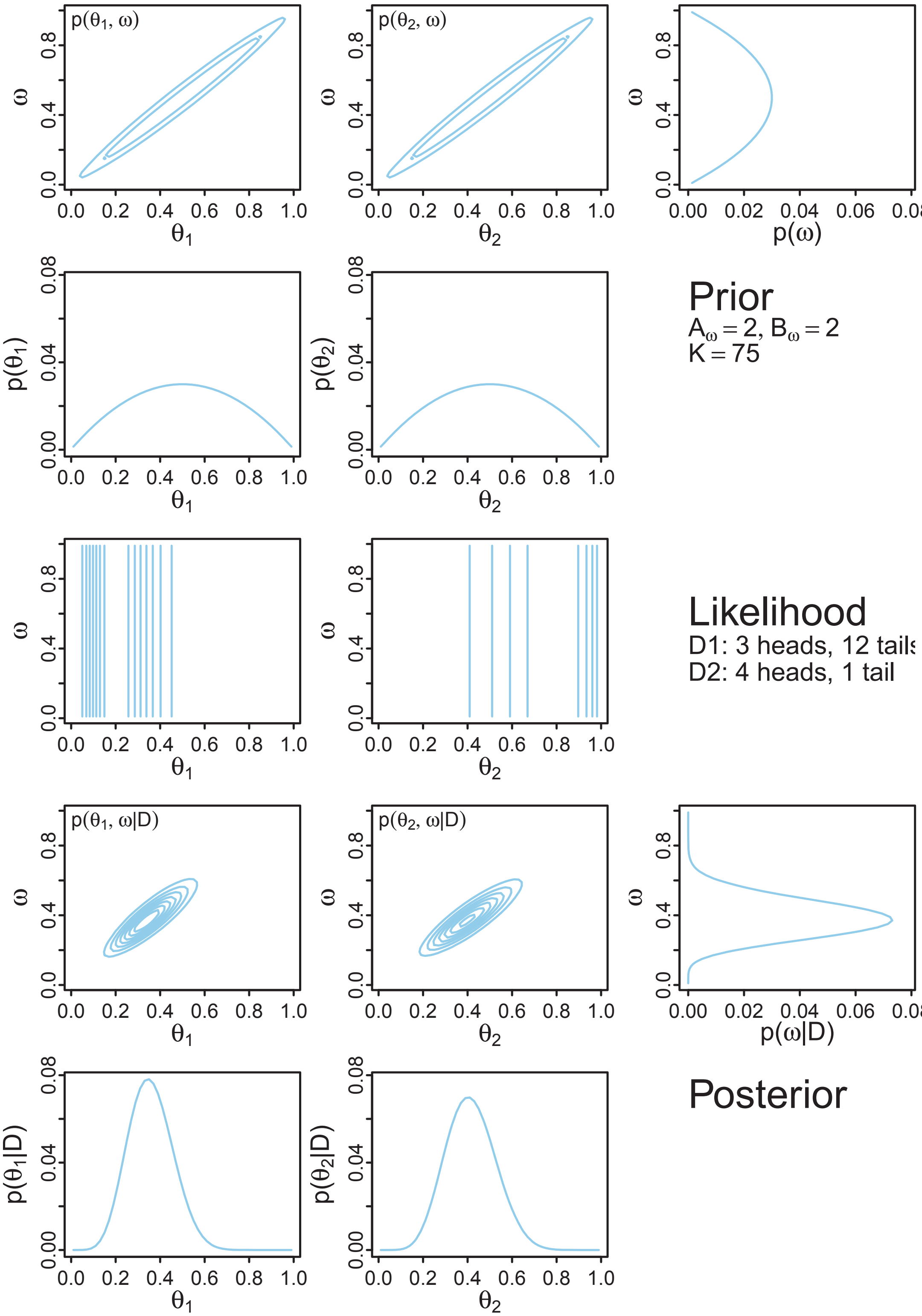

Figure 9.6: The prior imposes a strong dependency of θ on μ (i.e., K is large), so the posteriors on θ_1 and θ_2 (bottom row) are strongly influenced by each other's data, with θ_2 being pulled toward θ_1 because $N_1 > N_2$. Compare with Figure 9.5, which uses the same data but a prior that has a weak dependence.

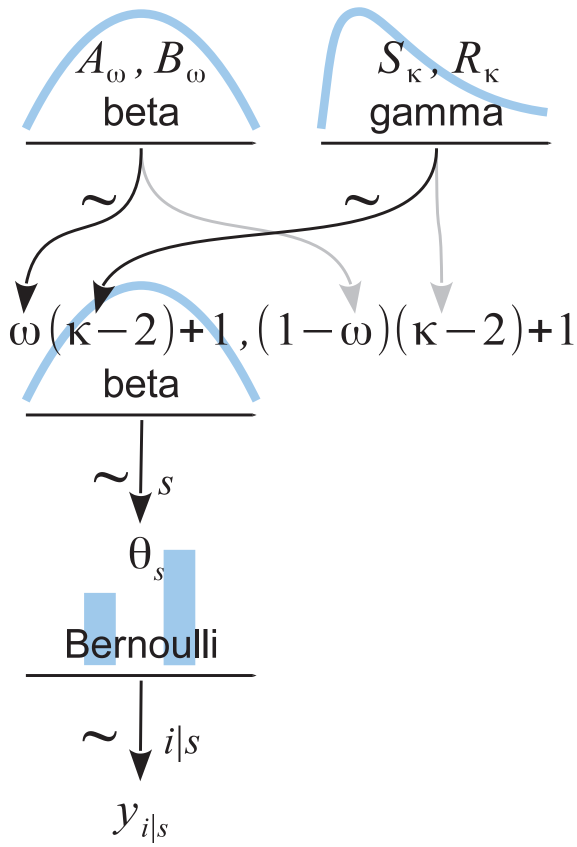

Figure 9.7: A model of hierarchical dependencies for data from several coins created independently from the same mint, with the characteristics of the mint parameterized by its mode ω and concentration κ. The value of $\kappa - 2$ has a prior distributed as a gamma density with shape and rate parameters of S_κ and R_κ. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

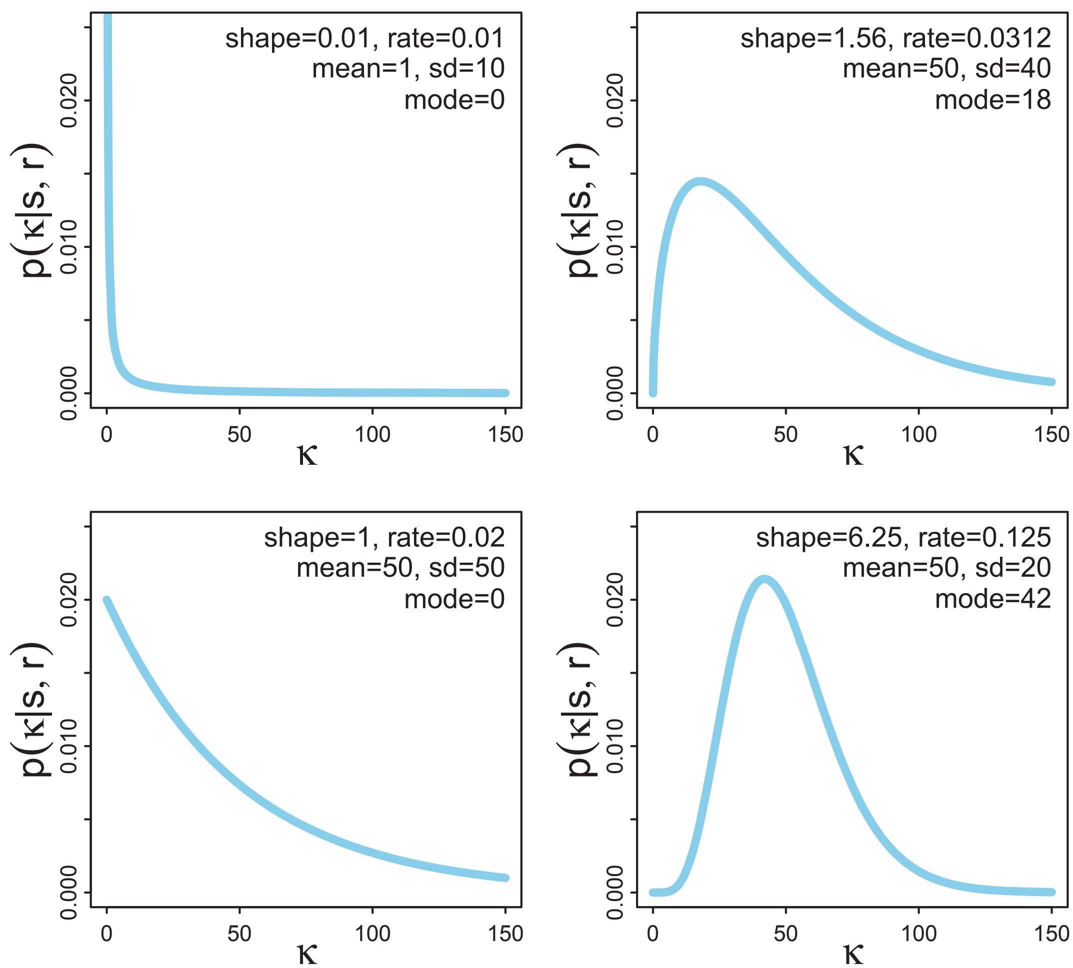

Figure 9.8: Examples of the gamma distribution. The vertical axis is $p(\kappa|s, r)$ where s is the shape and r is the rate, whose values are annotated in each panel. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

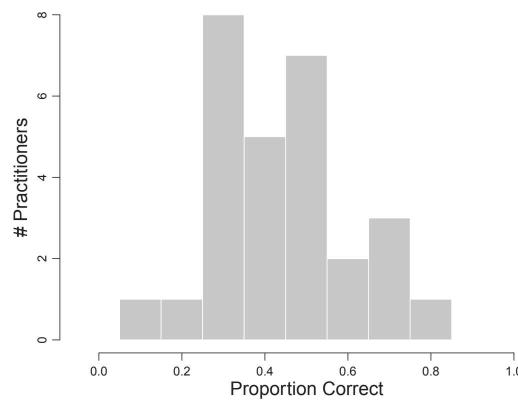

Figure 9.9: The therapeutic touch experiment of Rosa et al. (1998). Histogram of proportion correct for the 28 practitioners. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

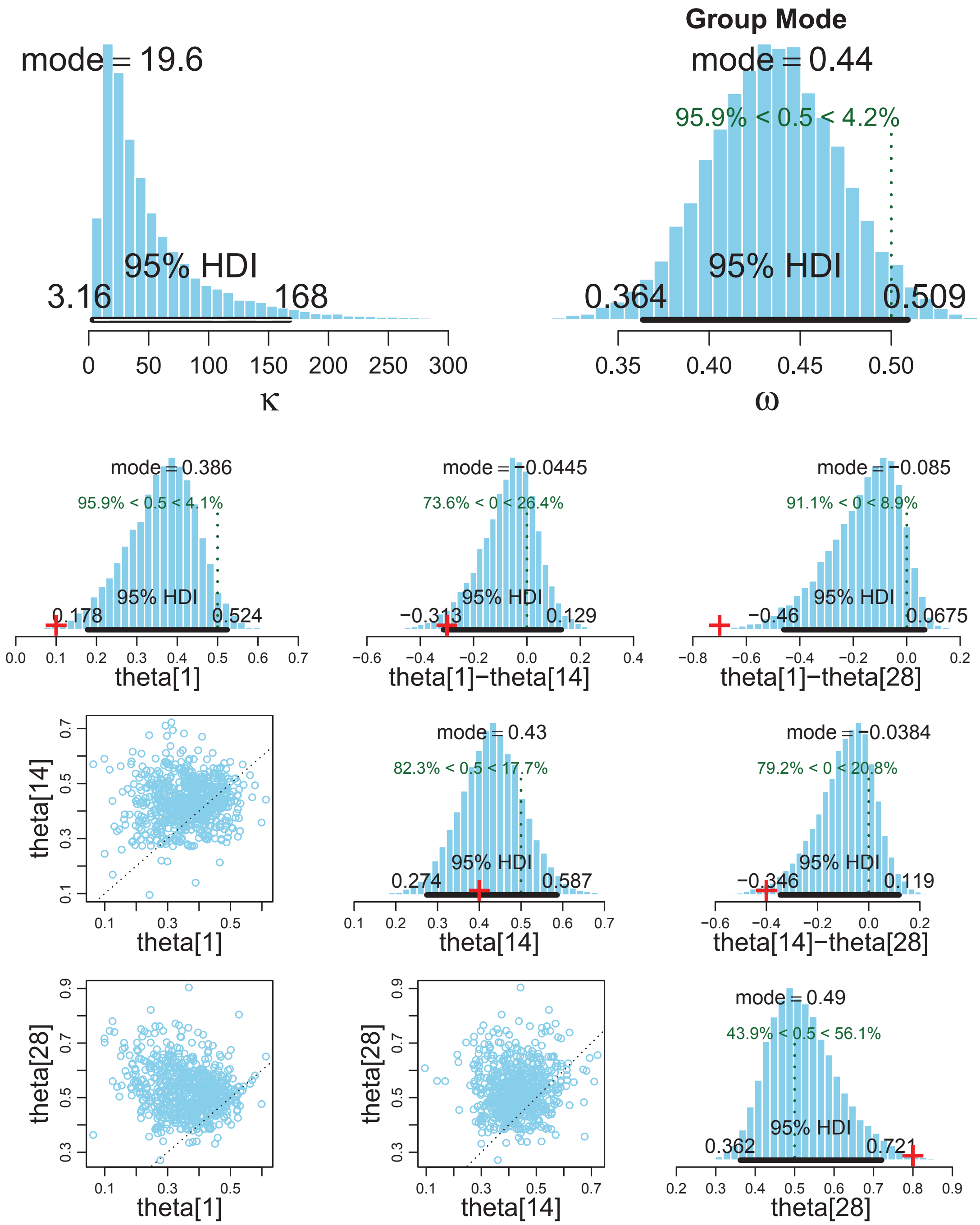

Figure 9.10: Marginal posterior distributions for the therapeutic touch data. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

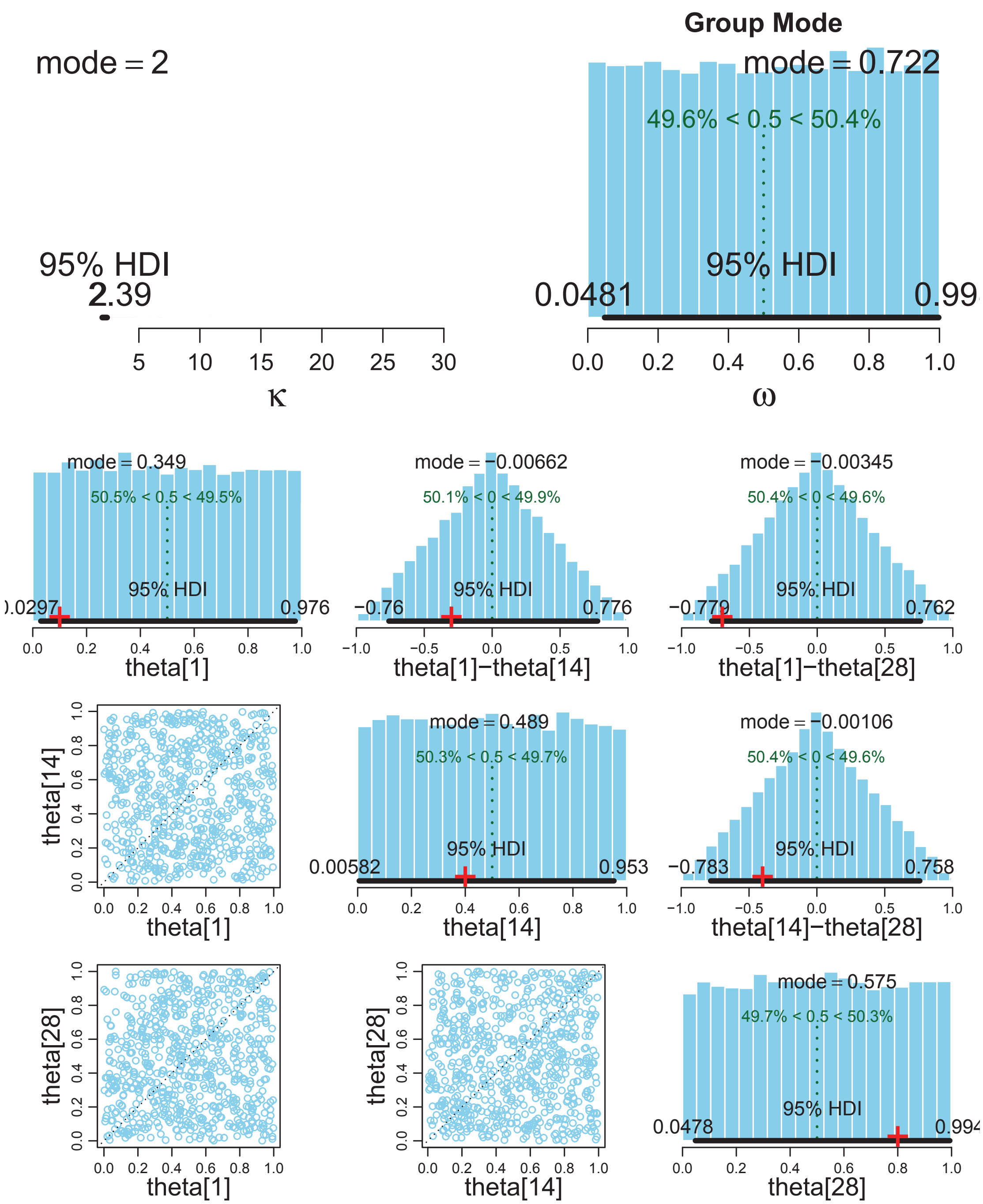

Figure 9.11: Marginal prior distributions for the therapeutic touch data. The upper-left panel, showing κ, does not plot well because it is a tall narrow peak near 2, with a long short tail extending far right. The estimated modal values of uniform distributions should be disregarded, as they are merely marking whatever random ripple happens to be a little higher than the other random ripples. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

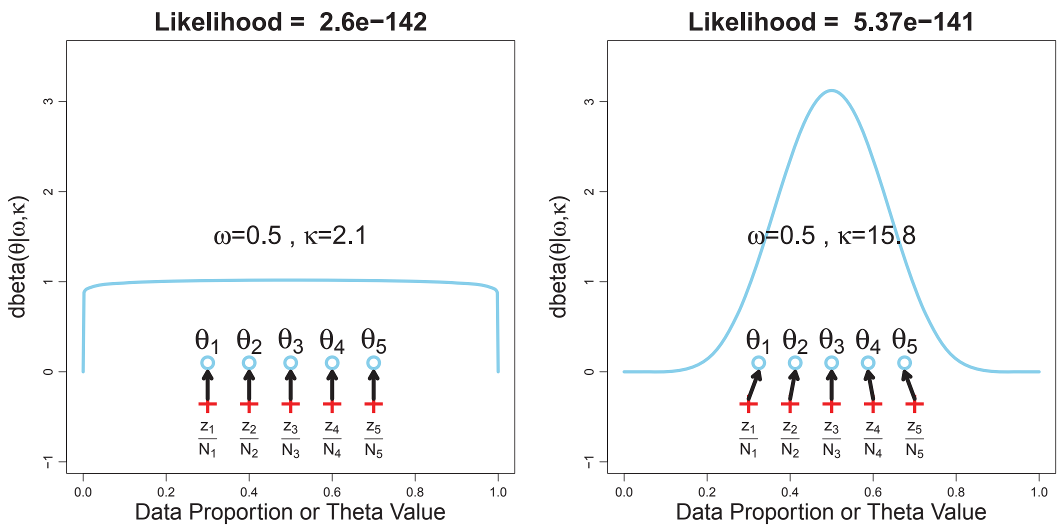

Figure 9.12: Shrinkage of individual parameter values in maximum likelihood estimation (MLE). Within each panel, the data proportion (z_s/N_s) from each individual is plotted as a "+" symbol, and the candidate value of θ_s is plotted as a circle, and the overarching beta distribution is also plotted with its candidate values of mode (ω) and concentration (κ) annotated. The left panel shows the choice of $\theta_s = z_s/N_s$ with a nearly flat beta distribution. The right panel shows the MLE, which exhibits shrinkage. Arrows highlight the shrinkage. The likelihood has improved by a factor greater than 20. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

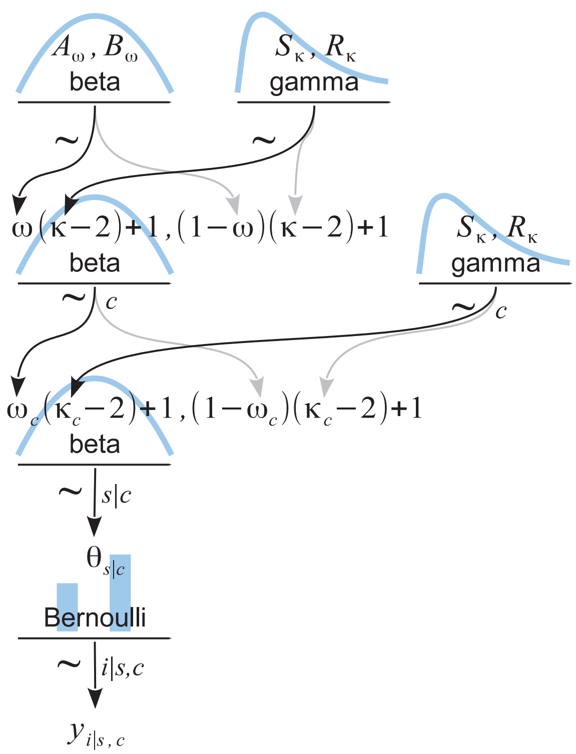

Figure 9.13: A model of hierarchical dependencies for data from several coins (indexed by subscript s) created by more than one category of mint (indexed by subscript c), with an overarching distribution across categories. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

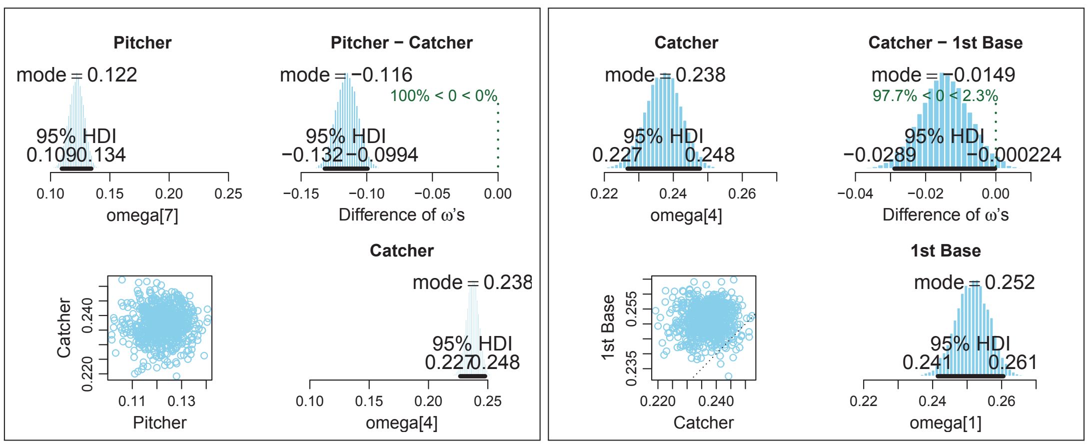

Figure 9.14: Marginal posterior distributions for baseball batting data. Left quartet shows that the pitchers have far lower batter abilities than the catchers. Right quartet shows that the catchers have marginally lower batting abilities than 1st-base men. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

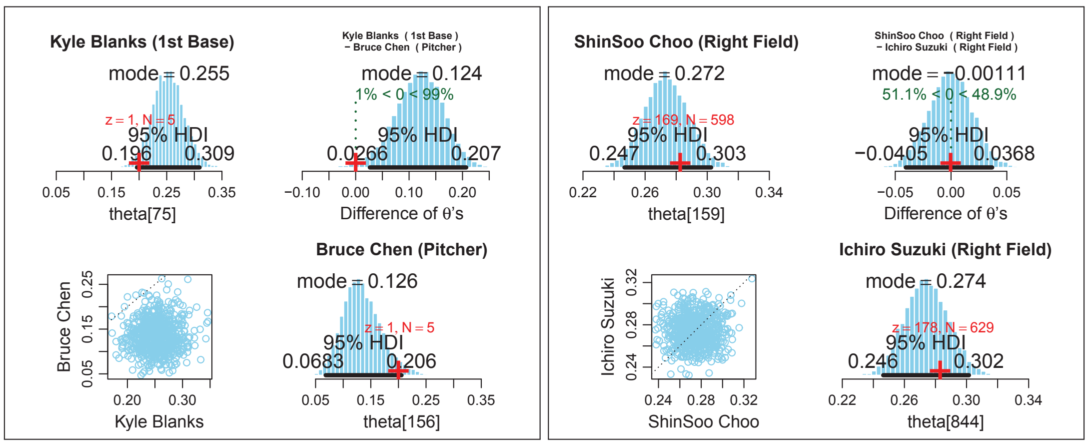

Figure 9.15: Marginal posterior distributions for baseball batting data. *Left quartet:* Individual estimates for two players with identical records of 1 hit in only 5 at-bats, but from two different positions. Although the batting records are identical, the estimated batting abilities are very different. *Right quartet:* Individual estimates for two right fielders with large numbers of at-bats. The posterior distributions of their individual performances have narrow HDI's compared with the left quartet, and are shrunken slightly toward the position-specific mode (which is about 0.247). The posterior distribution of their difference is essentially zero and the 95% HDI of the difference is very nearly contained within a ROPE from −0.04 to +0.04 (except for MCMC instability). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

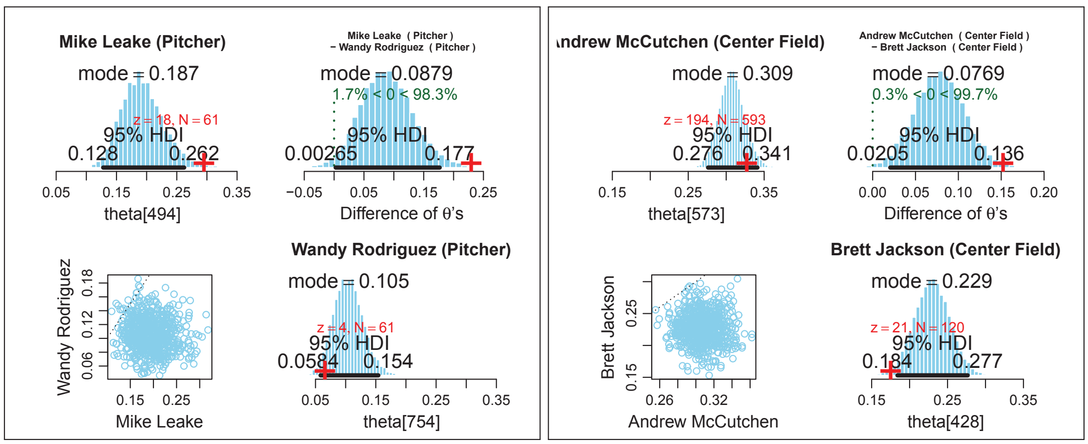

Figure 9.16: Marginal posterior distributions for baseball batting data. *Left quartet:* Two pitchers each with 61 at-bats but very different numbers of hits. Despite the difference in performance, shrinkage toward the position-specific mode leaves the posterior distribution of their difference marginally spanning zero. *Right quartet:* Two center fielders with very different batting averages, and moderately large at-bats. Despite some shrinkage toward the position-specific mode, the larger set of data makes the posterior distribution of their difference notably exclude zero. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

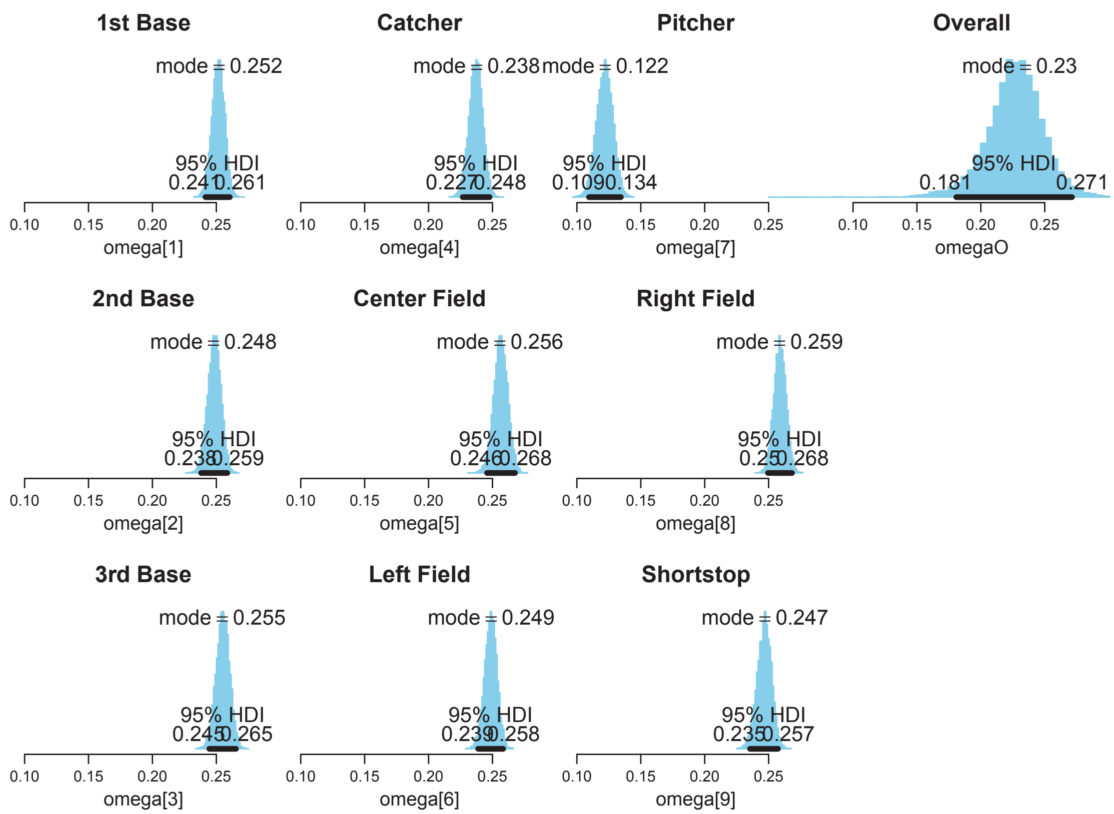

Figure 9.17: Marginal posterior distributions for baseball batting data. Notice that the estimate of the overall mode `omegaO` is less certain (wider HDI) than the estimate of position modes `omega[c]`. One reason for the different certainties is that there are dozens or hundreds of individuals contributing to each position, but only nine positions contributing to the overall level. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

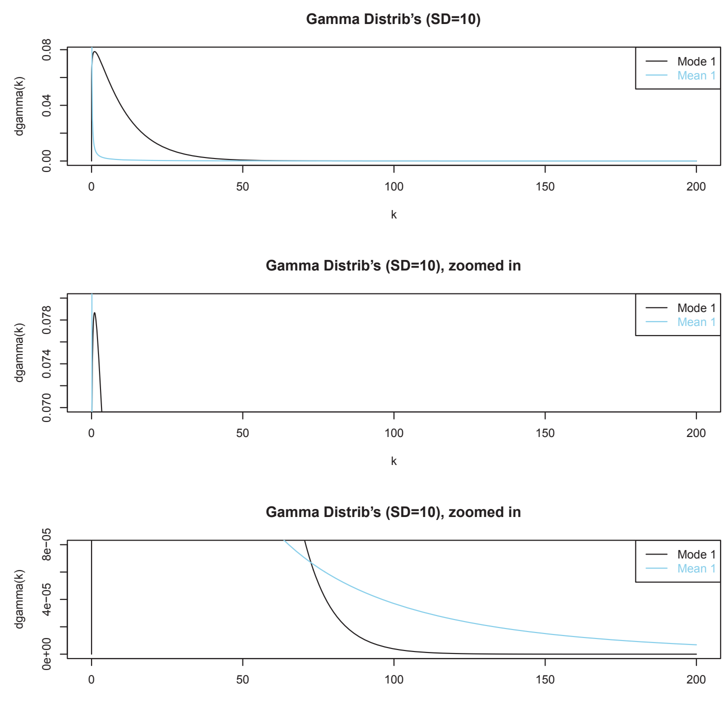

Figure 9.18: Two gamma distributions superimposed, for use with Exercise 9.1. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

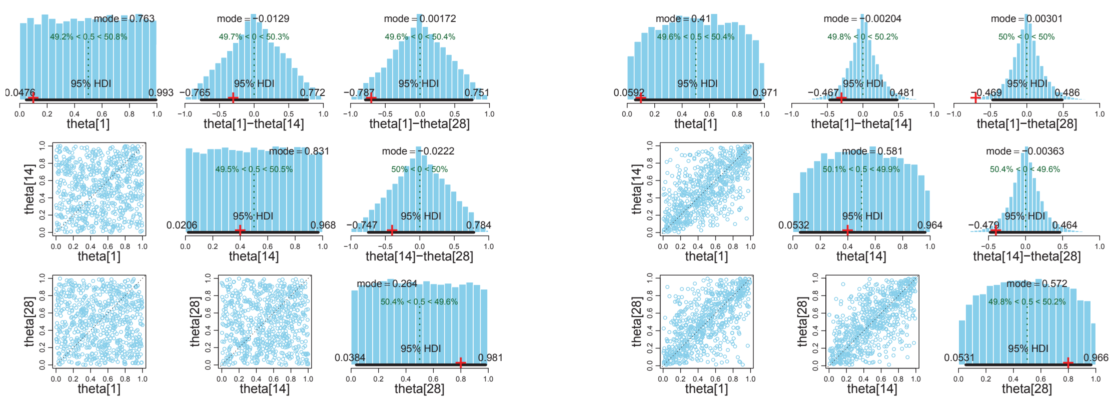

Figure 9.19: Priors on θ_s implied by different gamma distributions on κ. For use with Exercise 9.2. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

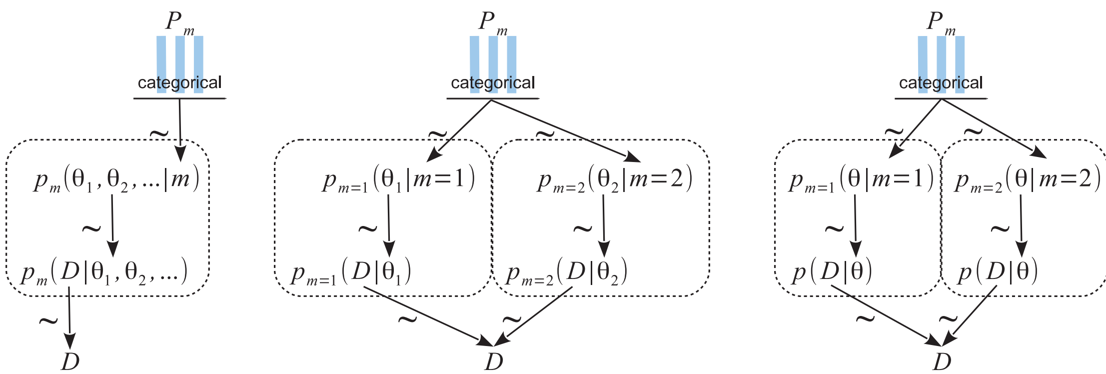

Figure 10.1: Model comparison as a single hierarchical model. Dashed boxes enclose the sub-models being compared. Left panel shows general conception, with parameters θ_m for all sub-models in a joint space. Middle panel shows the usual case in which the likelihood and prior reduce to functions of only θ_m for each m. Right panel shows the special case in which the likelihood function is the same for all m, and only the form of the prior is different for different m. (Middle and right panels depict only two sub-models, but there can be many.)

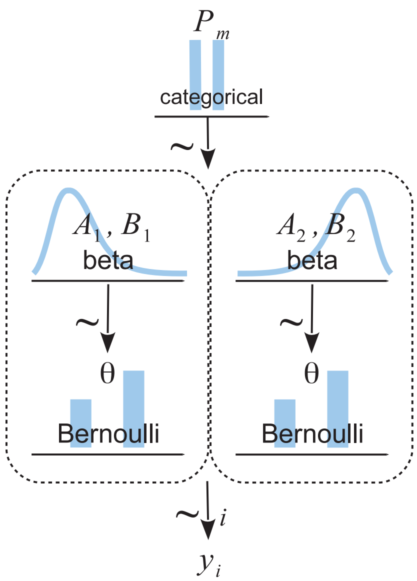

Figure 10.2: Hierarchical diagram for two models of a coin. Model 1 is a tail-biased mint; model 2 is a head-biased mint. This diagram is a specific case of the right panel of Figure 10.1, because the likelihood function is the same for both models and only the priors are different.

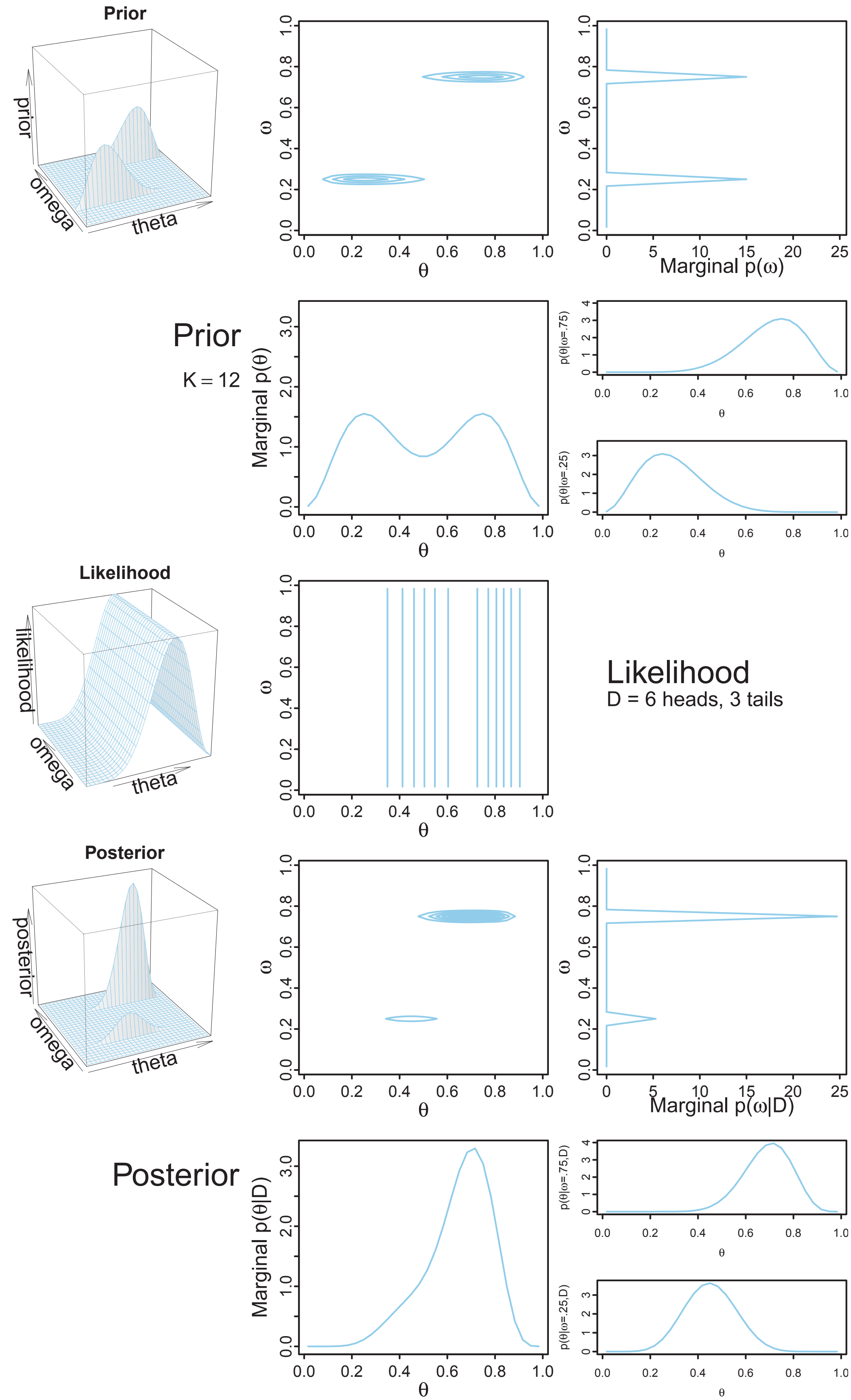

Figure 10.3: A representation of the joint ω, θ parameter space when the mode parameter, ω, is allowed only two discrete values. (For an example with a continuous distribution on ω, compare with Figure 9.2, p. 224.)

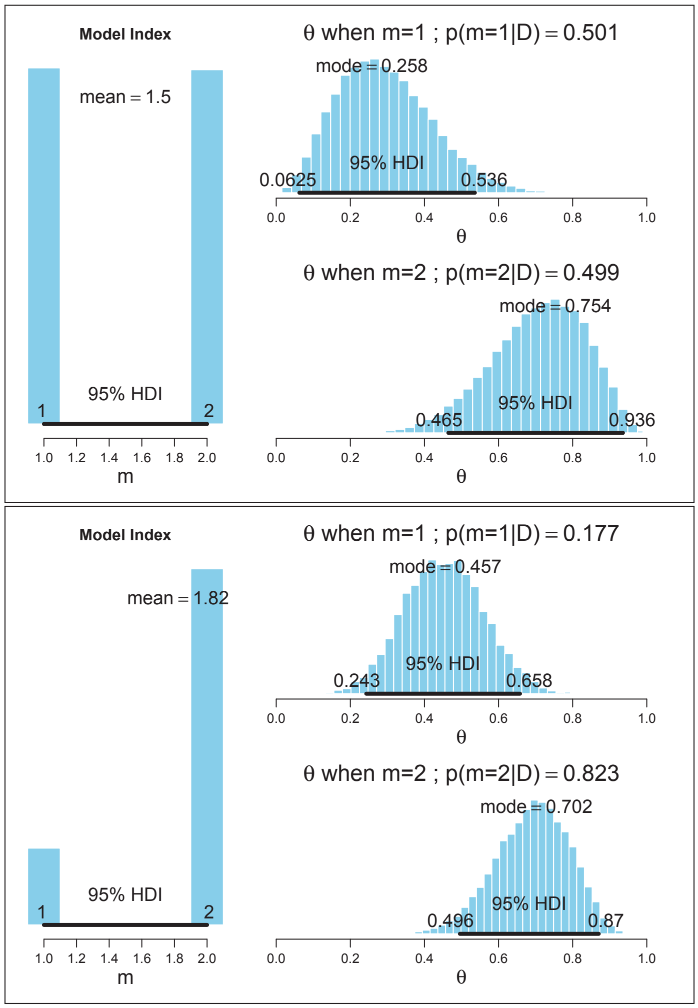

Figure 10.4: The prior and posterior distributions for script `Jags-Ydich-Xnom1subj-MbernBetaModelComp.R`. The upper frame, which shows the prior distribution, has labels that indicate $p(\theta|D)$ but the data set, D, is empty. The lower frame shows the posterior distribution. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

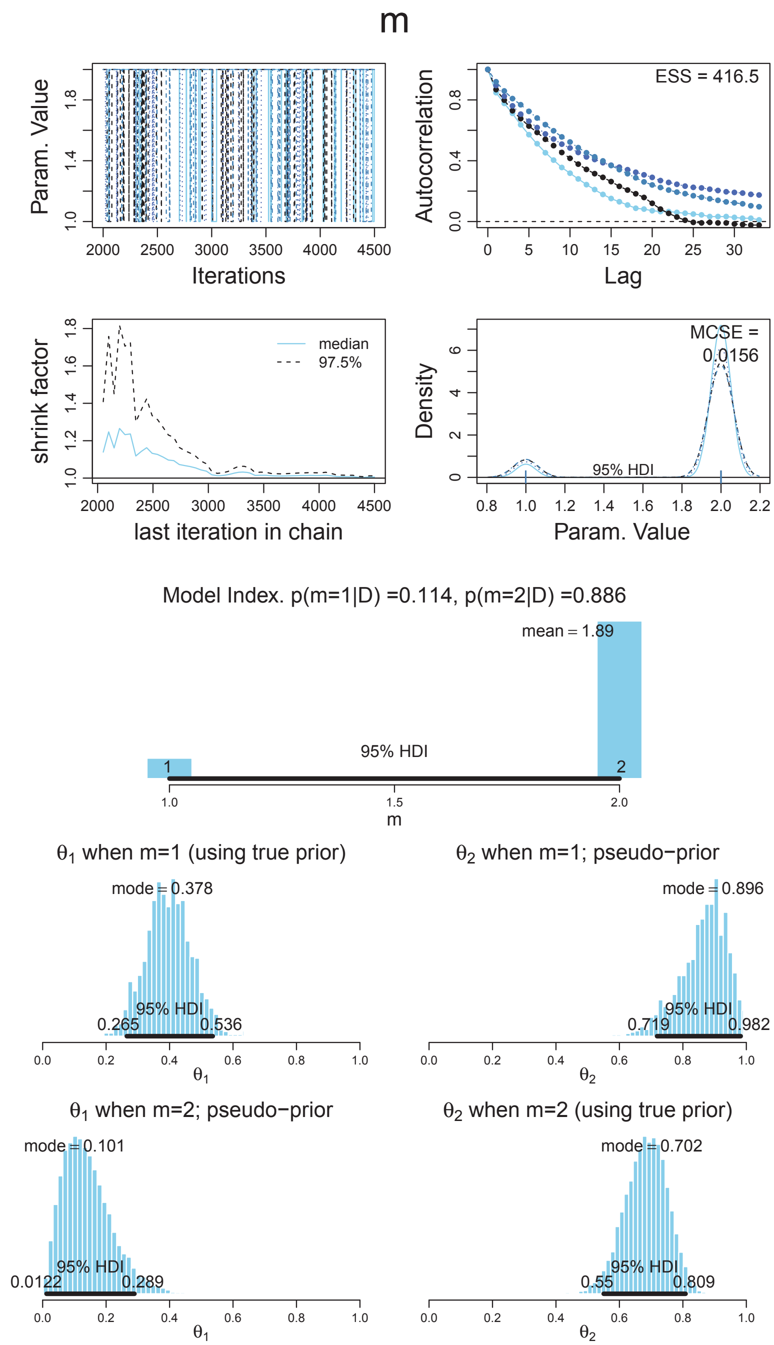

Figure 10.5: Poor jumping between models when *not* using pseudopriors; that is, when the implemented pseudoprior is the true prior. Compare with Figure 10.6. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

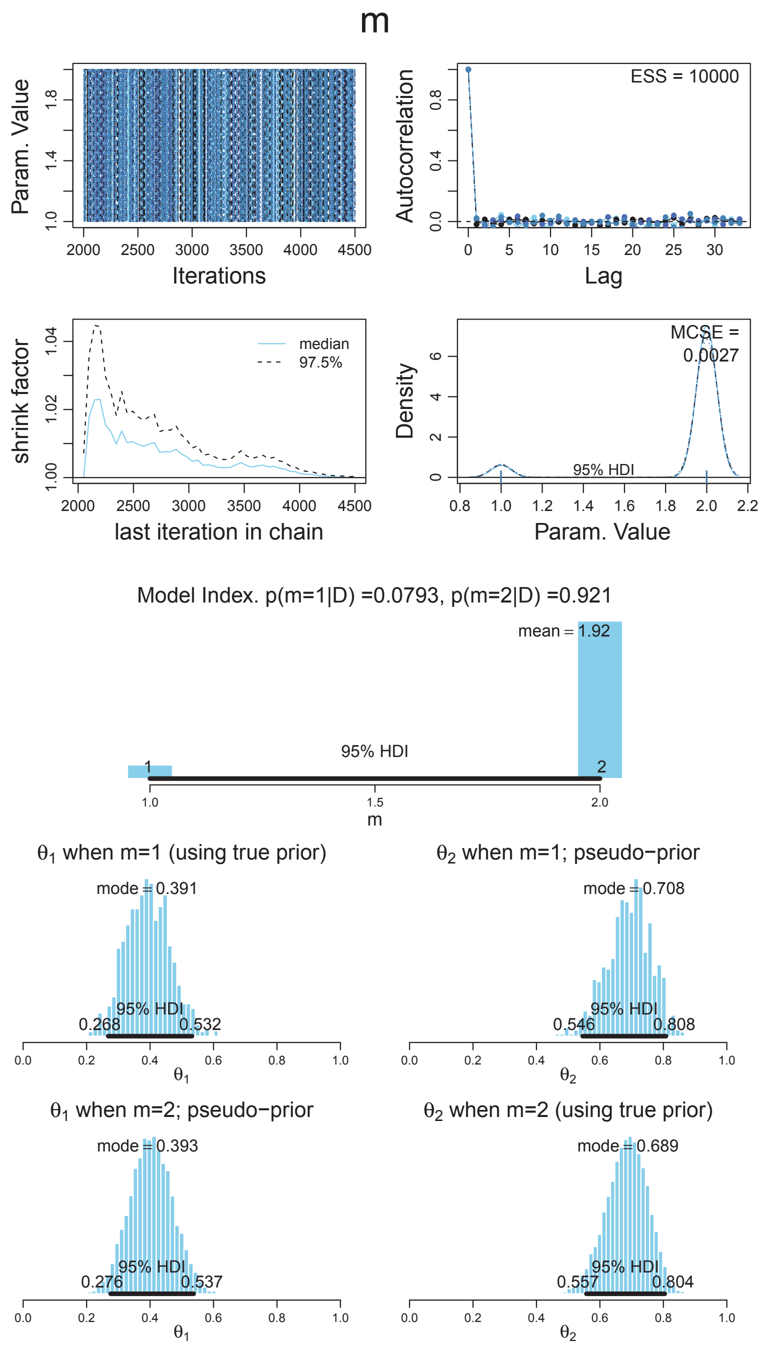

Figure 10.6: Better jumping between models when *using* pseudopriors; that is, when the implemented pseudoprior mimics the posterior. Compare with Figure 10.5. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

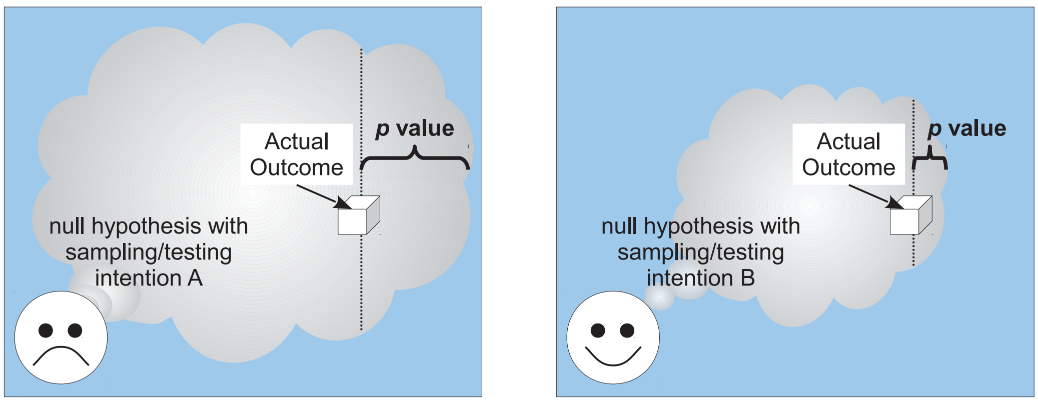

Figure 11.1: The null hypothesis generates a cloud of imaginary outcomes, most of which fall in the center of the cloud, but some of which fall beyond the actual outcome marked by block. The p value is the proportion of the cloud as or more extreme than the actual outcome. Left panel: With sampling intention A, the cloud of imaginary possibilities has a large proportion that exceeds the actual outcome, hence the p value is large. Right panel: With sampling intention B, the cloud of imaginary possibilities has a small proportion that exceeds the actual outcome, hence the p value is small. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

		N								
		1	2	3	4	5	6	7	8	...
z	0									...
	1									...
	2	-								...
	3	-	-							...
	4	-	-	-						...
	5	-	-	-	-					...
	6	-	-	-	-	-				...
	7	-	-	-	-	-	-			...
	8	-	-	-	-	-	-	-		...
	...	-	-	-	-	-	-	-	-	...

		N								
		1	2	3	4	5	6	7	8	...
z	0									...
	1									...
	2	-								...
	3	-	-							...
	4	-	-	-						...
	5	-	-	-	-					...
	6	-	-	-	-	-				...
	7	-	-	-	-	-	-			...
	8	-	-	-	-	-	-	-		...
	...	-	-	-	-	-	-	-	-	...

		N								
		1	2	3	4	5	6	7	8	...
z	0									...
	1									...
	2	-								...
	3	-	-							...
	4	-	-	-						...
	5	-	-	-	-					...
	6	-	-	-	-	-				...
	7	-	-	-	-	-	-			...
	8	-	-	-	-	-	-	-		...
	...	-	-	-	-	-	-	-	-	...

Figure 11.2: Sample space for flips of a coin, in which columns show candidate values for N and rows show candidate values for z. Left table: Space of possibilities when N is considered fixed, highlighted by shaded column (at $N = 5$). Middle table: Space of possibilities when z is considered fixed, highlighted by shaded row (at $z = 4$). Right table: Space of possibilities when duration is considered fixed, with probabilities of sample sizes suggested by differential shading of columns. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

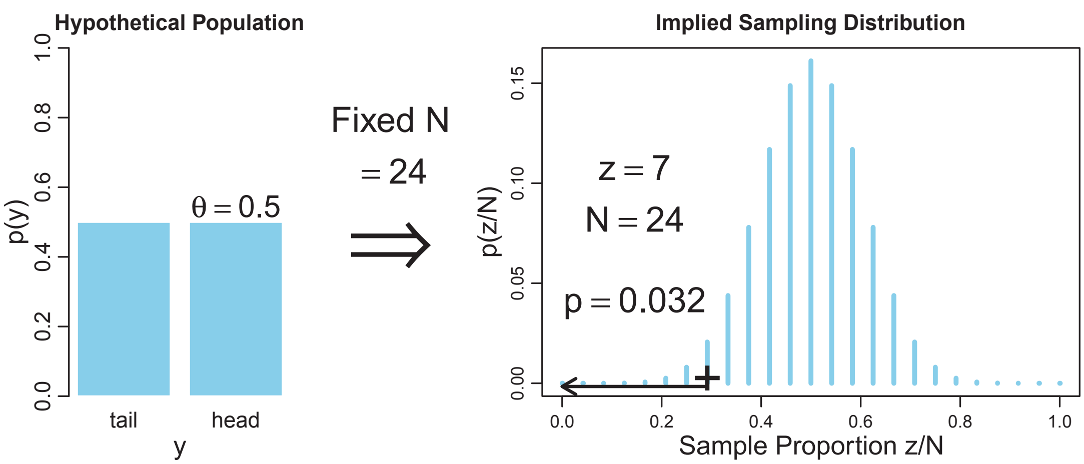

Figure 11.3: The imaginary cloud of possible outcomes when N is fixed. The null hypothesis likelihood distribution and parameter are shown on the left. The stopping intention is shown is the middle. The sampling distribution and p value are shown on the right. Compare with Figures 11.4 and 11.5.

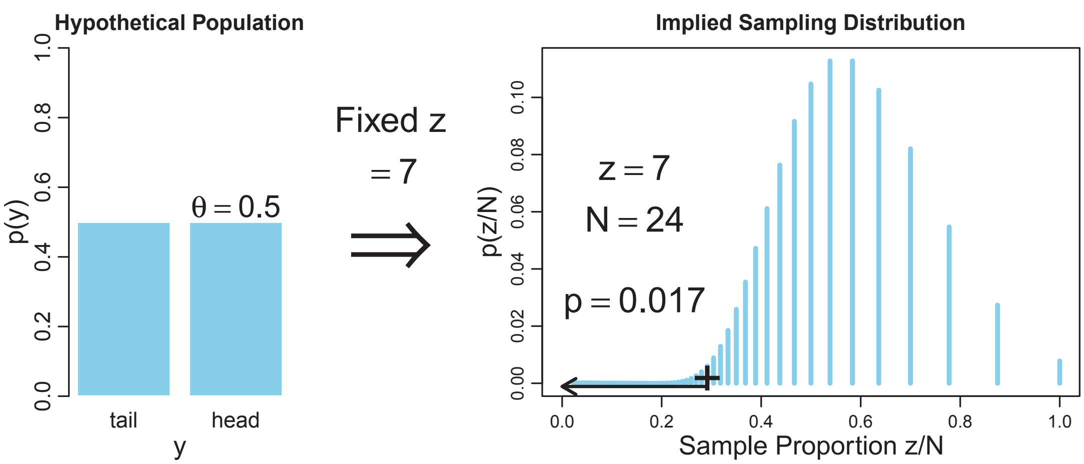

Figure 11.4: The imaginary cloud of possible outcomes when z is fixed. The null hypothesis likelihood distribution and parameter are shown on the left. The stopping intention is shown is the middle. The sampling distribution and p value are shown on the right. Compare with Figures 11.3 and 11.5. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

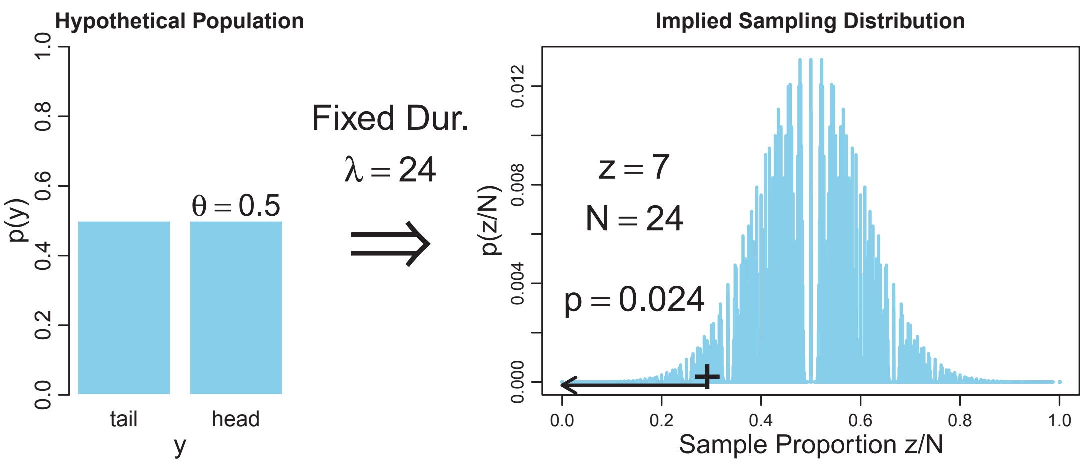

Figure 11.5: The imaginary cloud of possible outcomes when duration is fixed. The null hypothesis likelihood distribution and parameter are shown on the left. The stopping intention is shown is the middle. The sampling distribution and *p* value are shown on the right. Sample sizes are drawn randomly from a Poisson distribution with mean λ. Compare with Figures 11.3 and 11.4. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

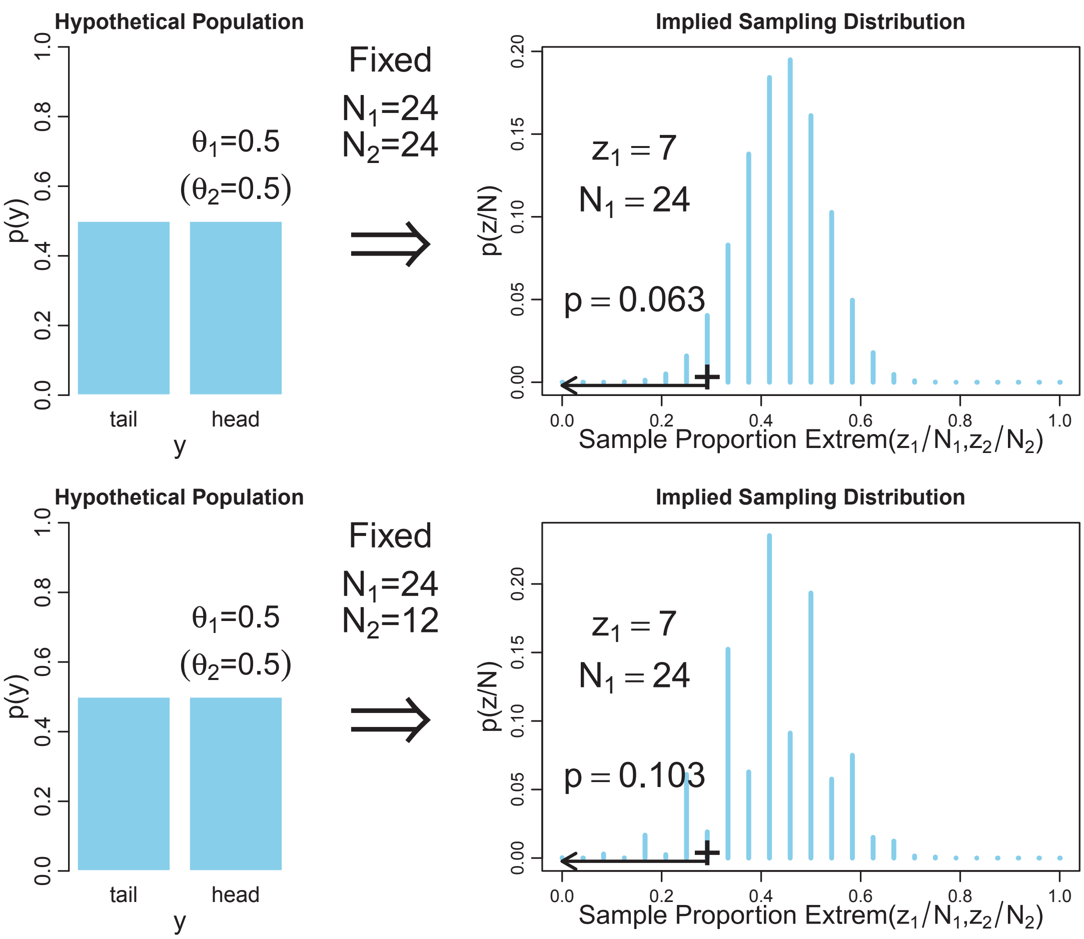

Figure 11.6: The imaginary cloud of possible outcomes when N is fixed and there are two independent tests. Upper is for $N_2 = N_1$. Lower is for $N_2 < N_1$. The null hypothesis likelihood distribution and parameter are shown on the left. The stopping and testing intentions are shown is the middle. The sampling distribution and p value are shown on the right. Compare with Figure 11.3. Extrem$\{z_1/N_1, z_2/N_2\}$ is the smaller (for left-tailed) of the hypothetically sampled proportions. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

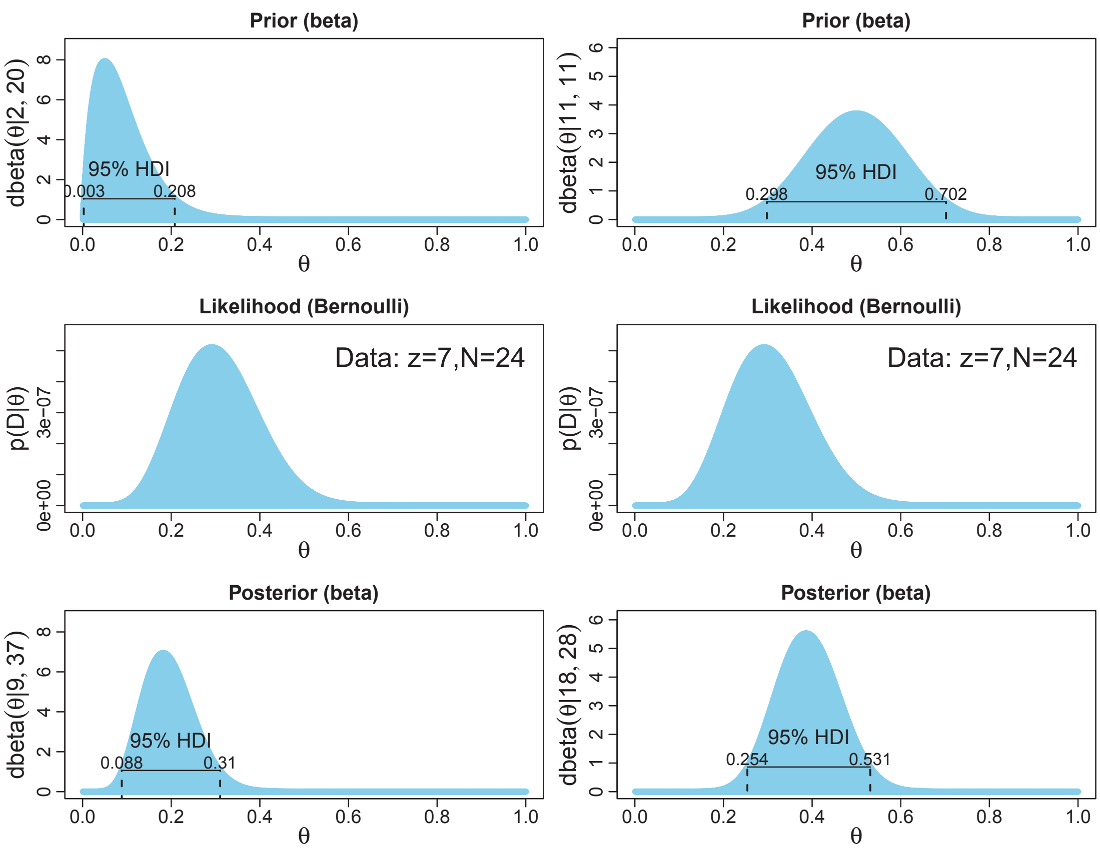

Figure 11.7: Posterior HDI for the bias of a Bernoulli process, when the prior assumes a tail-strong nail (left column) or a fair coin (right column). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

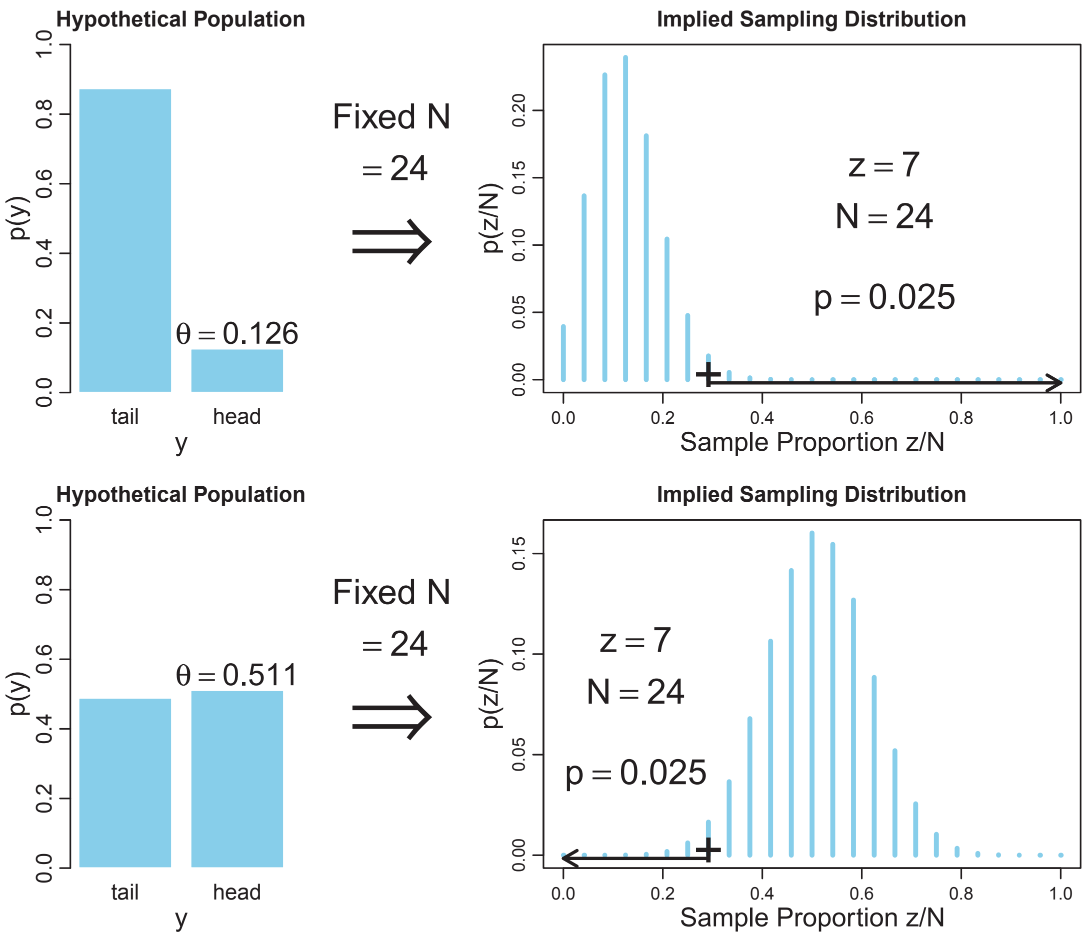

Figure 11.8: 95% confidence interval when N is fixed by the experimenter's intention extends from $\theta = 0.126$ (top row) to $\theta = 0.511$ (bottom row). Compare with Figures 11.9 and 11.10. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

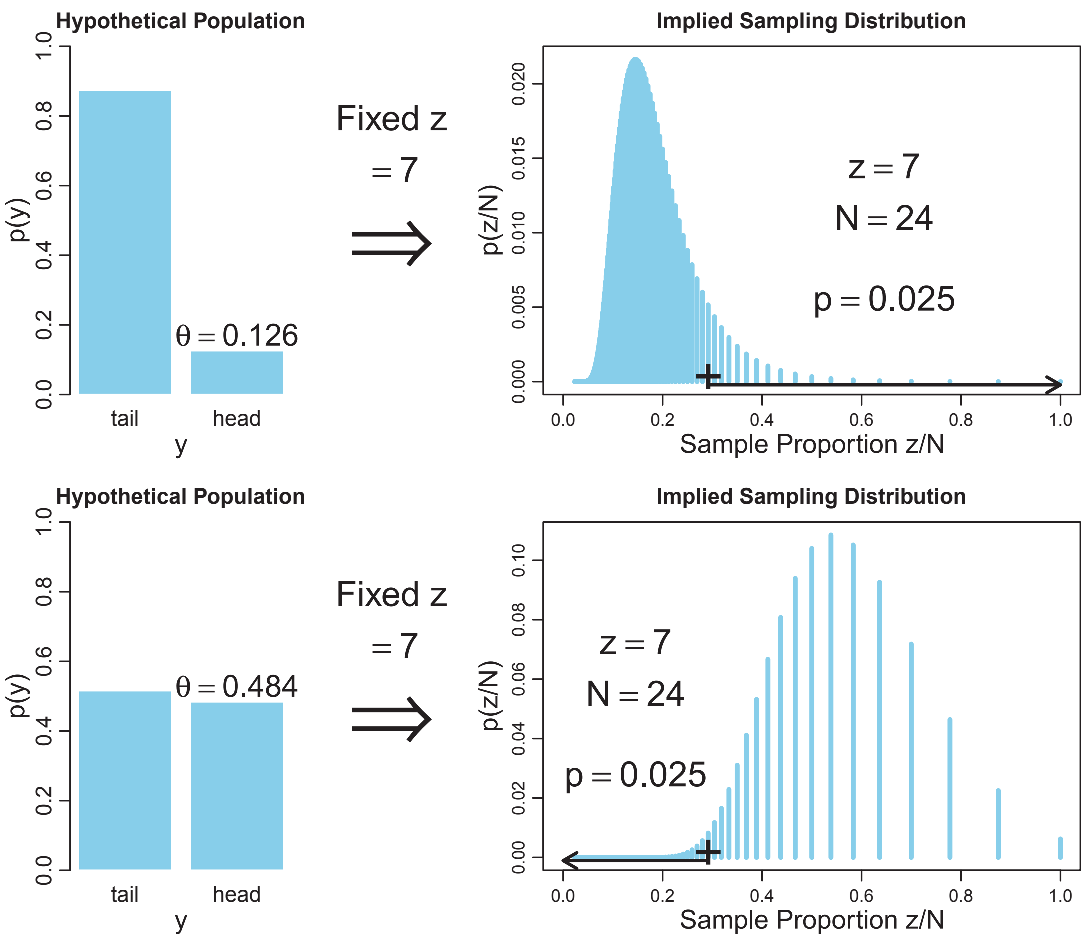

Figure 11.9: 95% confidence interval when z is fixed by the experimenter's intention extends from $\theta = 0.126$ (top row) to $\theta = 0.484$ (bottom row). Compare with Figures 11.8 and 11.10.

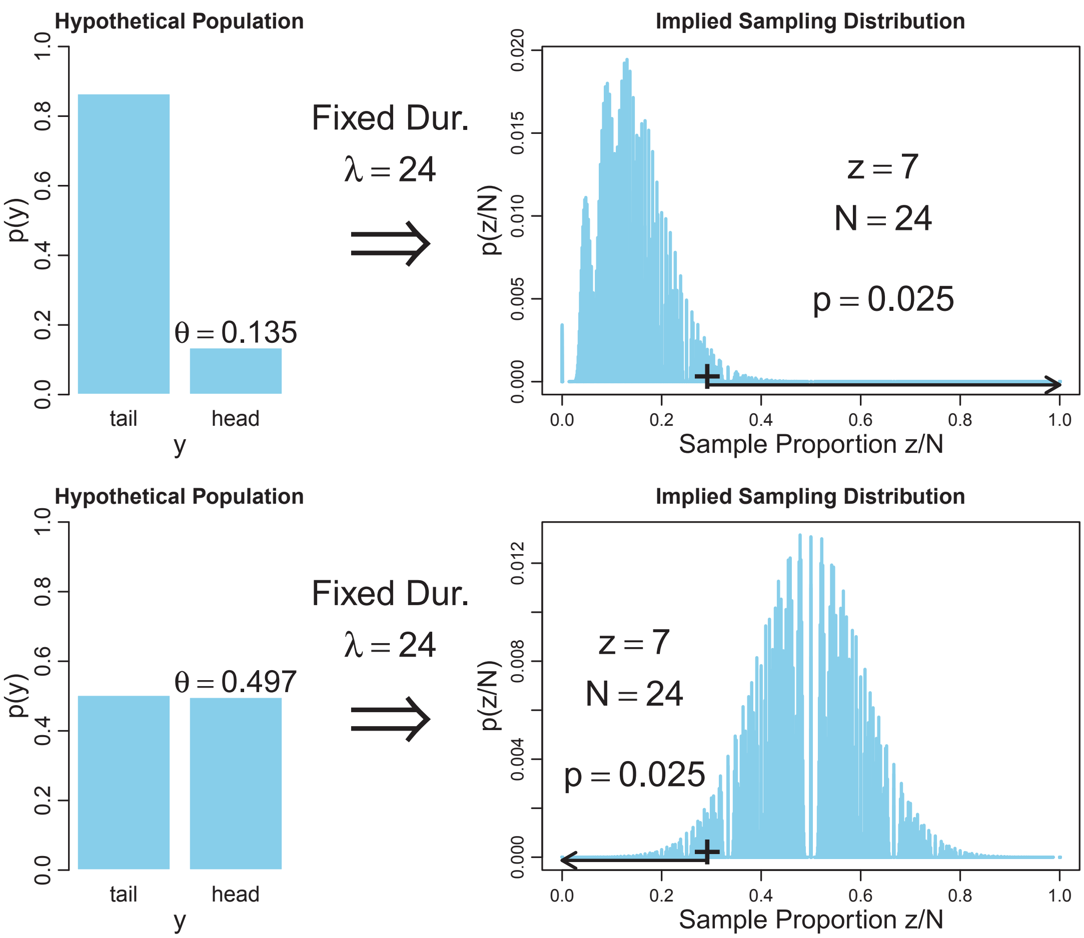

Figure 11.10: 95% confidence interval when duration is fixed by the experimenter's intention extends from $\theta = 0.135$ (top row) to $\theta = 0.497$ (bottom row). Compare with Figures 11.8 and 11.9.

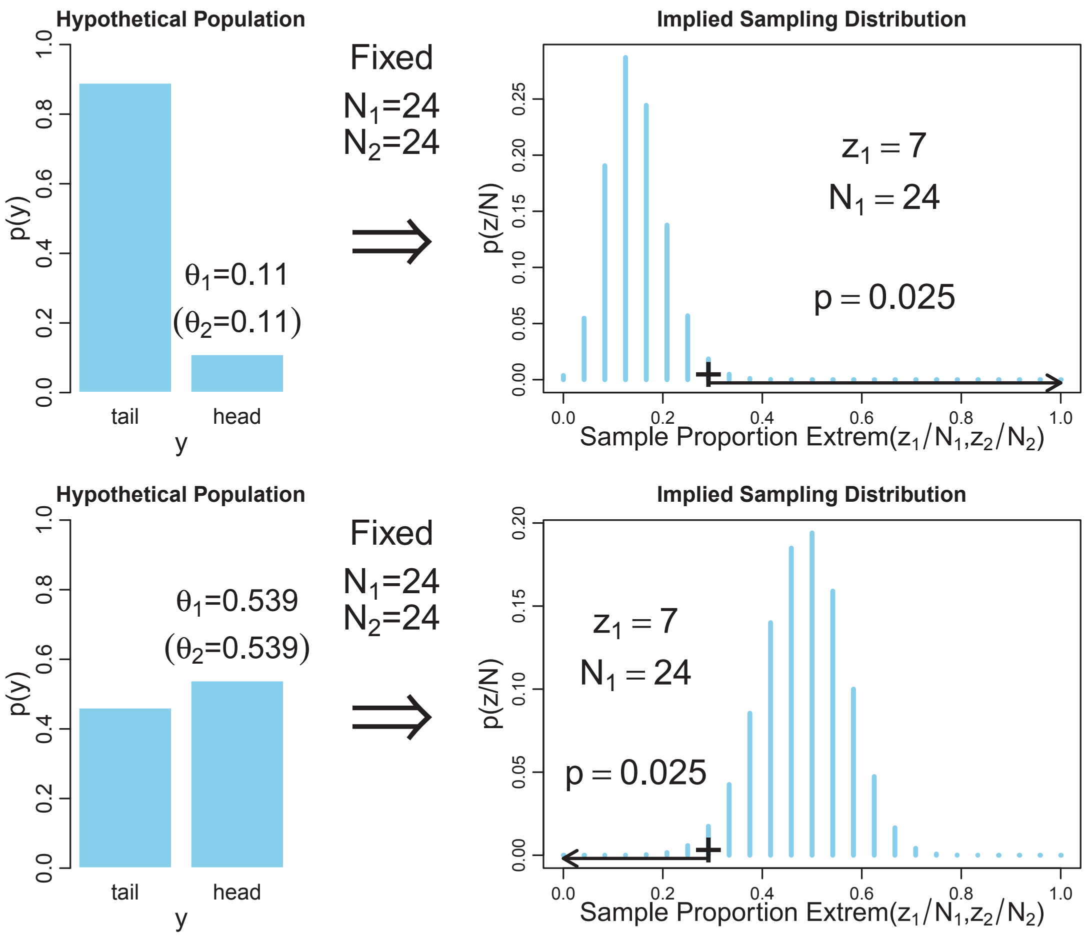

Figure 11.11: 95% confidence interval, when N is fixed by the experimenter's intention and there are two tests, extends from $\theta = 0.110$ (top row) to $\theta = 0.539$ (bottom row). Compare with Figure 11.8. Extrem$\{z_1/N_1, z_2/N_2\}$ is the hypothetical sampled proportion that is more extreme relative to θ. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

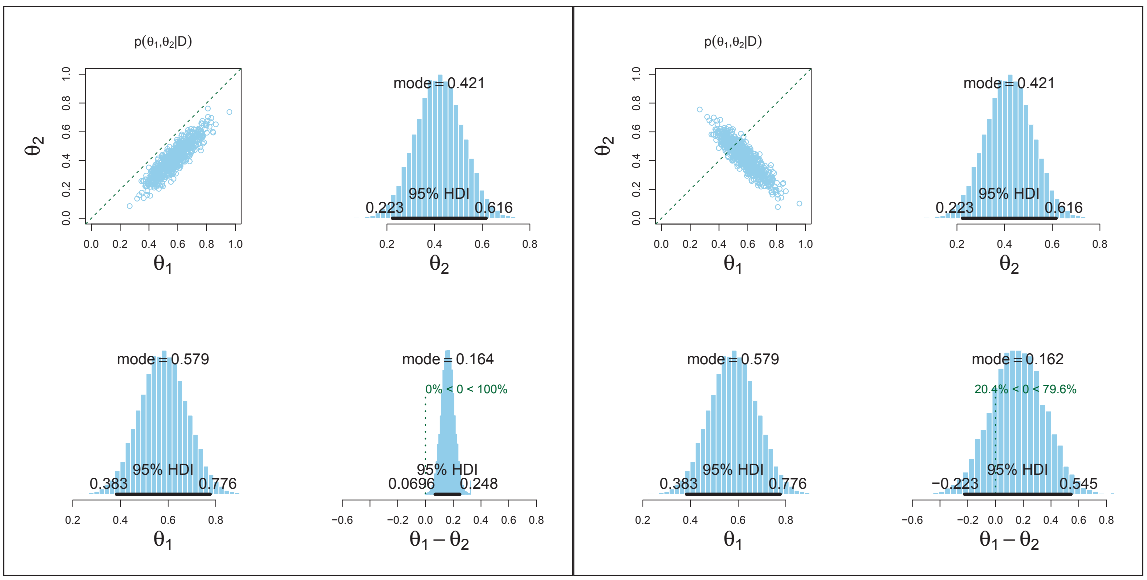

Figure 12.1: When there is a positive correlation between parameters, as shown in the left quartet, the distribution of differences is narrower than when there is a negative correlation, as shown in the right quartet. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

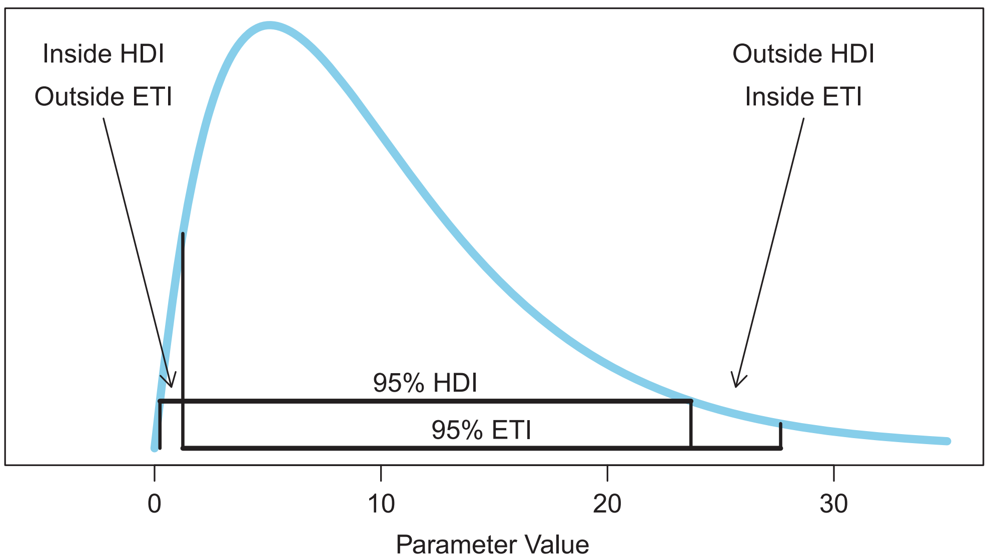

Figure 12.2: A skewed distribution has different 95% highest density interval (HDI) than 95% equal-tailed interval (ETI). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

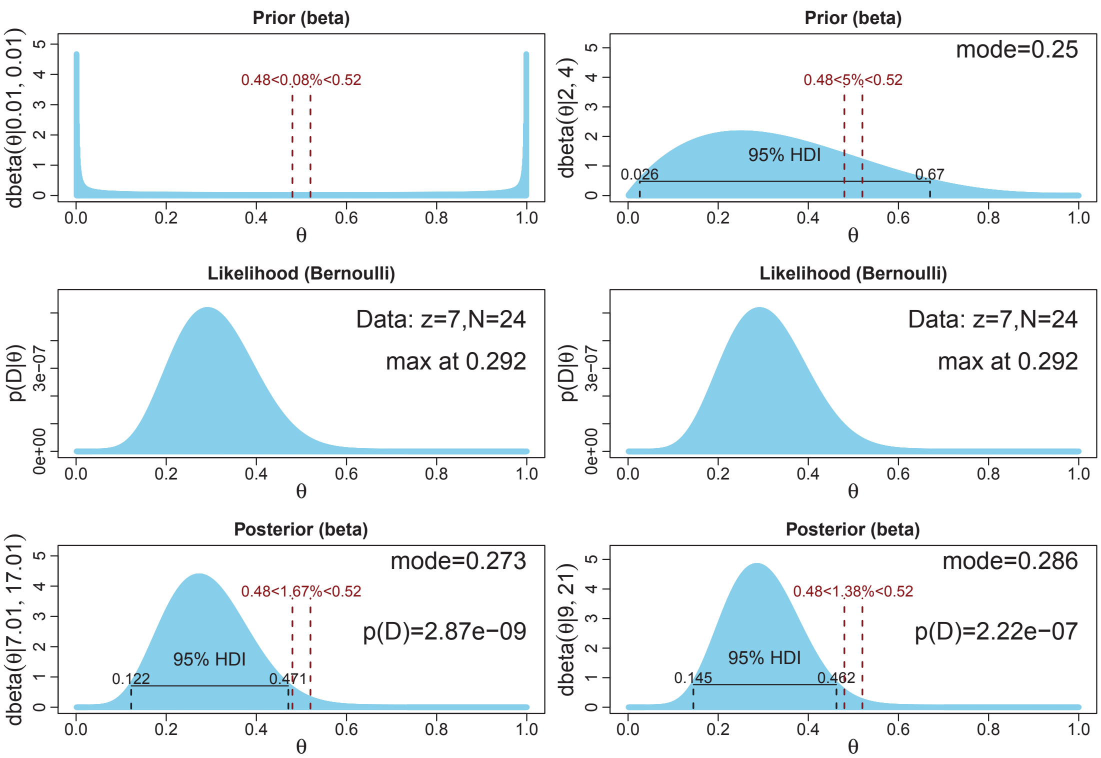

Figure 12.3: Left column: Haldane prior. Right column: Mildly informed prior. Vertical dashed lines mark a ROPE from 0.48 to 0.52. Annotation above the dashed lines indicates the percentage of the distribution within the ROPE. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

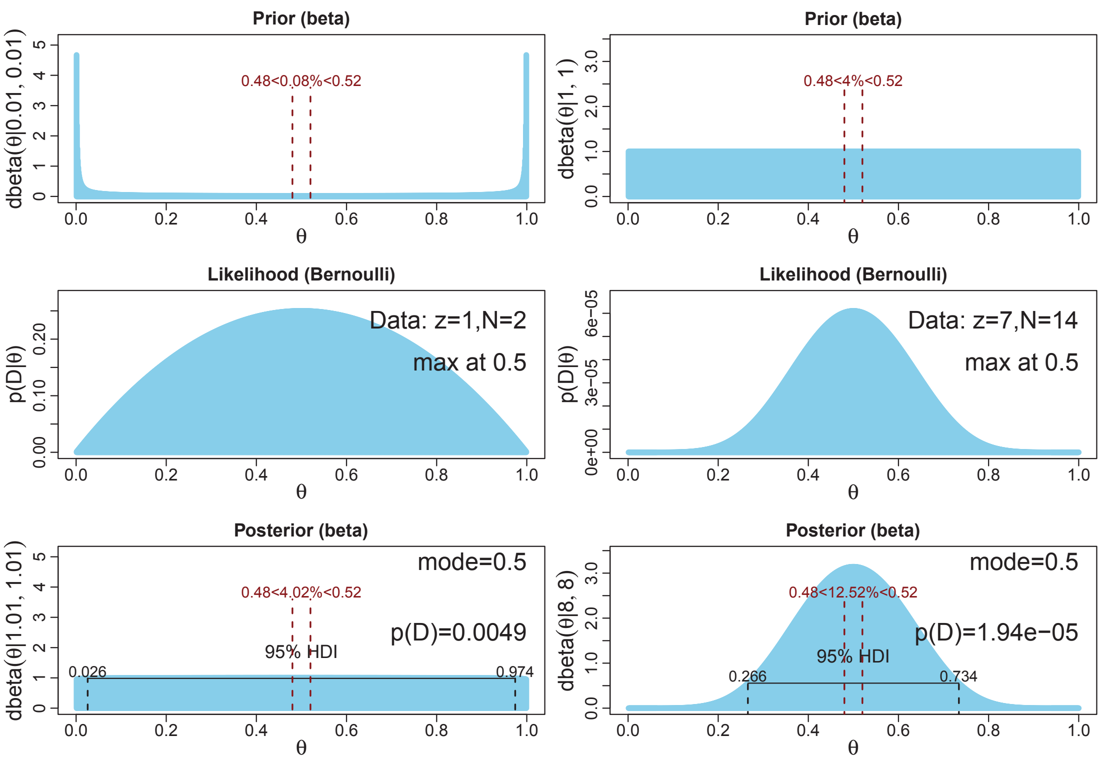

Figure 12.4: Bayes factor (model comparison) approach can accept the null even with low precision on estimate. Left column: Haldane prior. Bayes factor is 51.0 in favor of null, but 95% HDI extends from 0.026 to 0.974 (!). Right column: Uniform prior. Bayes factor is 3.14 in favor of null, but the 95% HDI extends from 0.266 to 0.734 (!).

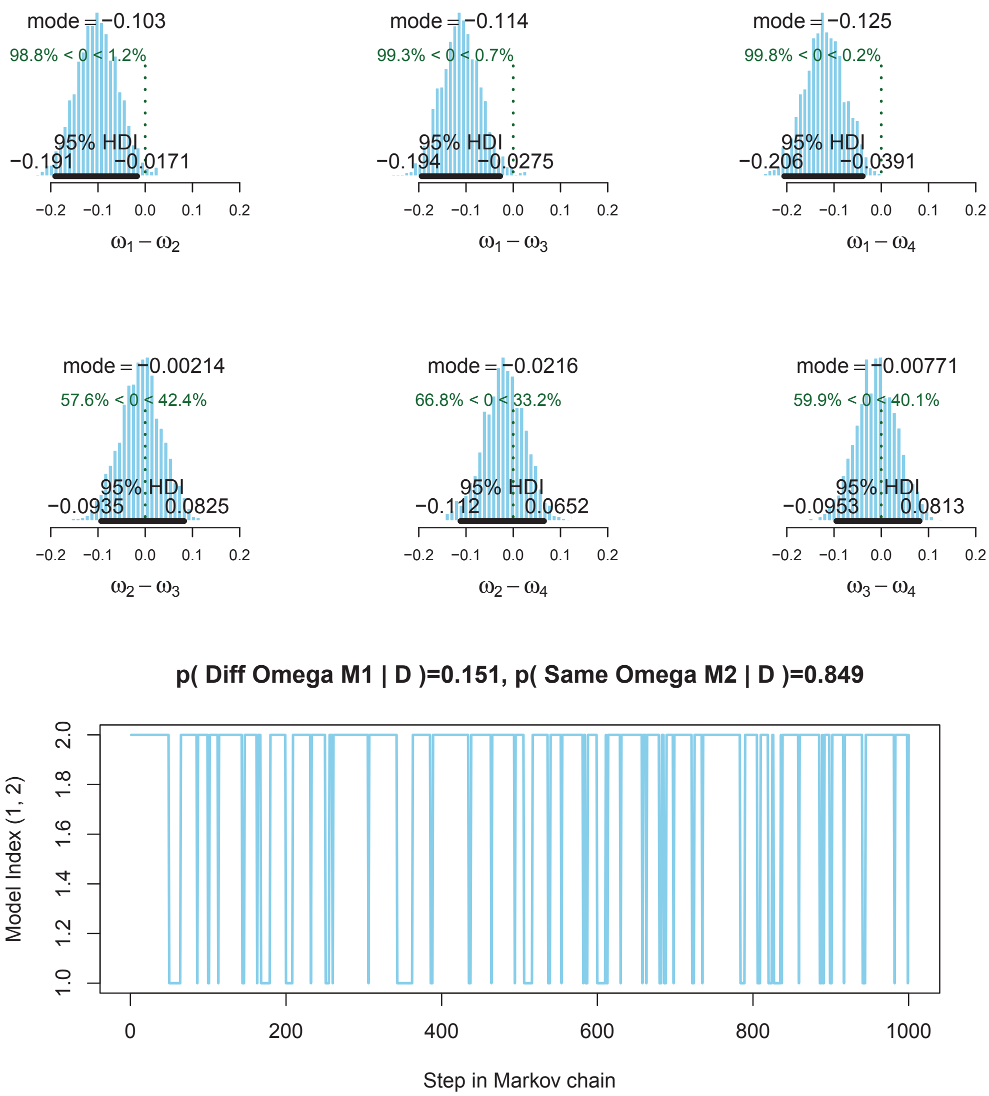

Figure 12.5: Top: Differences of posterior ω_j values for the four groups in the different-omega model. Notice that ω_1 *is credibly different from* ω_3 *and* ω_4, and possibly different from ω_2. The histograms are a bit choppy because the MCMC chain visits the different-omega model relatively rarely. Bottom: Trace plot of the model index shows that the model with a *single* omega parameter ("Same Omega M2") is *preferred* to a model with a separate omega parameter for each group ("Diff Omega M1").

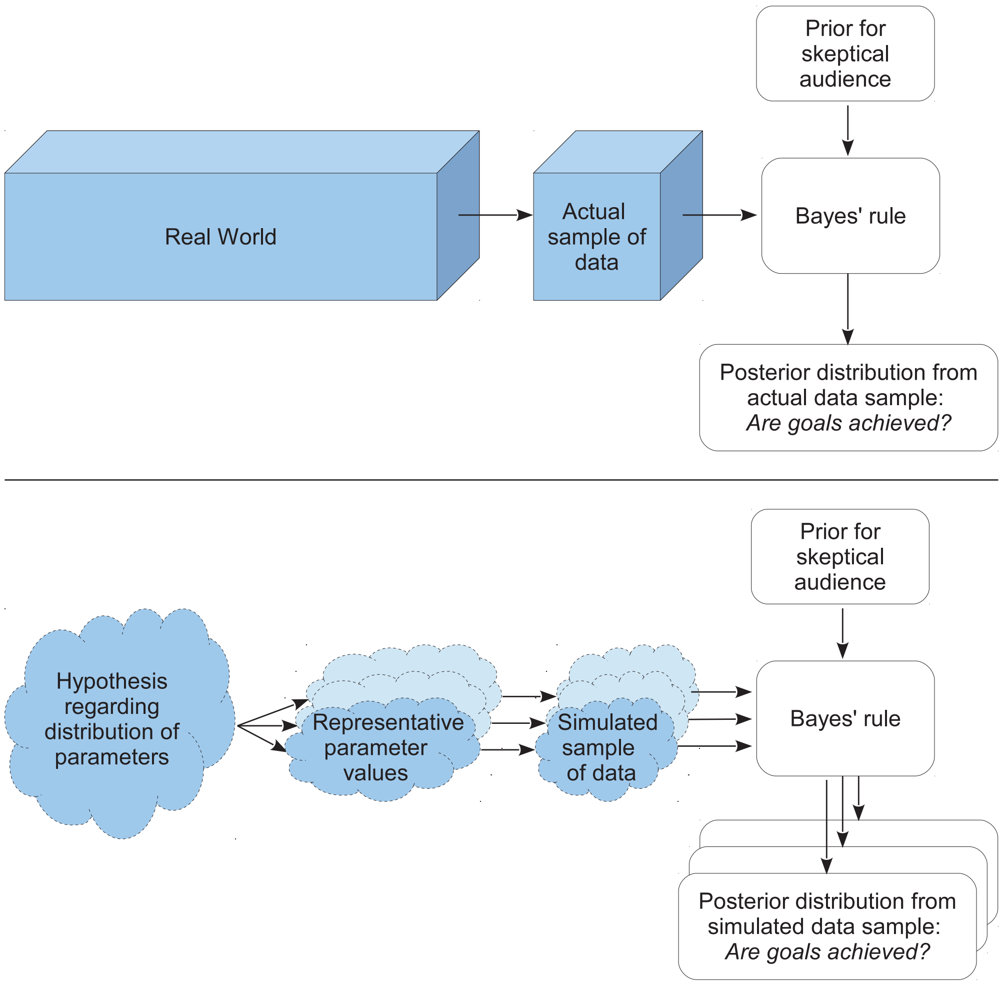

Figure 13.1: Upper diagram illustrates the flow of information in an actual Bayesian analysis, in which the data come from the real world. Lower diagram illustrates flow of information in a power analysis, in which simulated data come from random hypothetical parameters. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

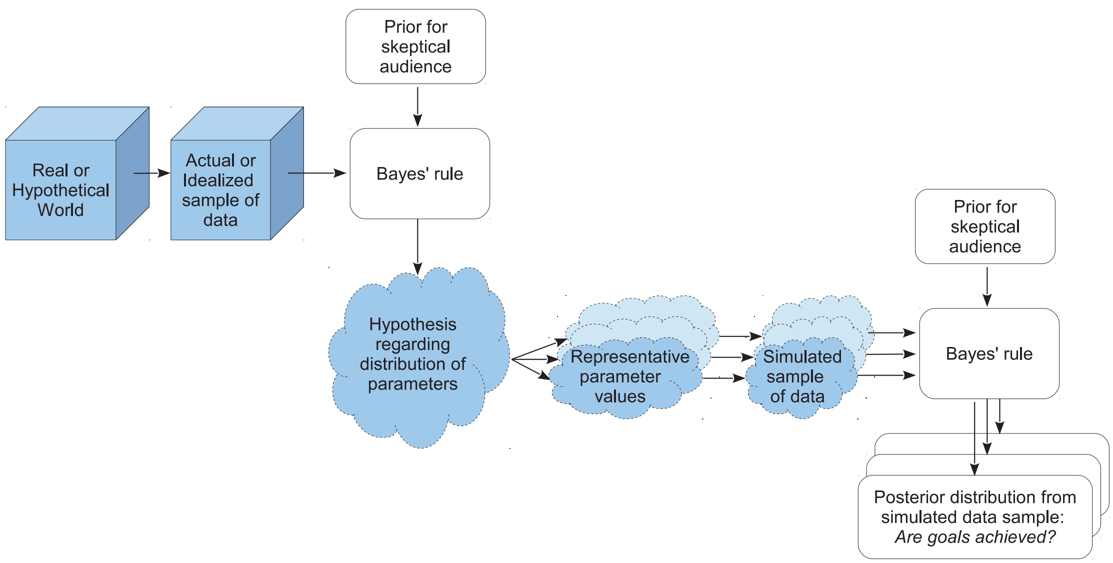

Figure 13.2: Flow of information in a power analysis when the hypothesis regarding the distribution of parameters is a posterior distribution from a Bayesian analysis on real or idealized previous data. Compare with Figure 13.1, p. 356.

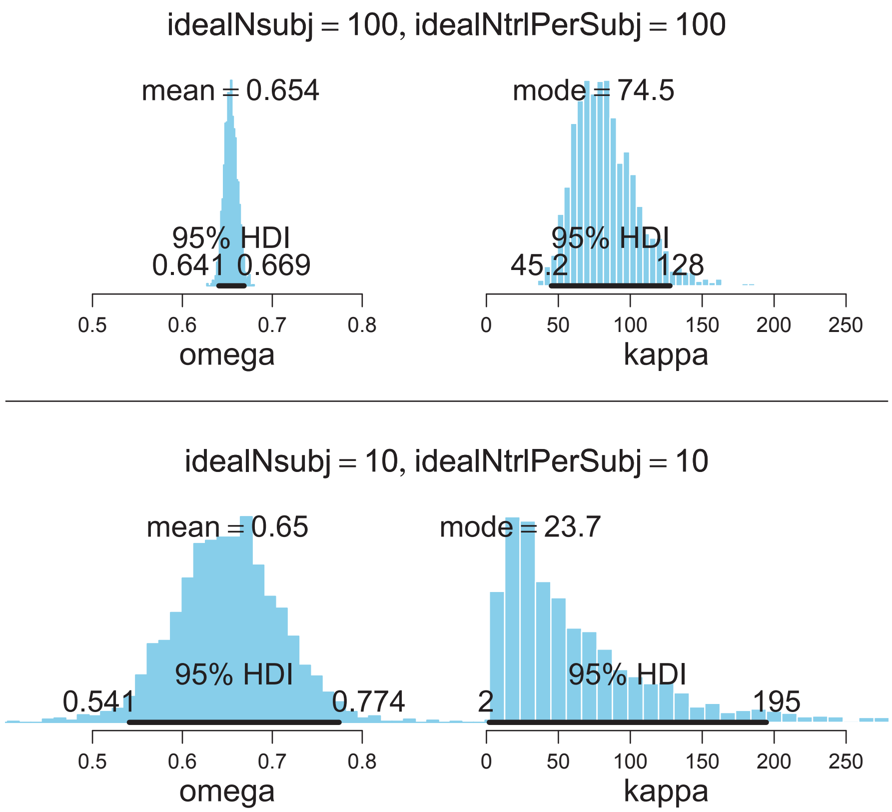

Figure 13.3: Distributions of parameters consistent with idealized data. Upper panel is for large amounts of idealized data, lower panel is for small amounts. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

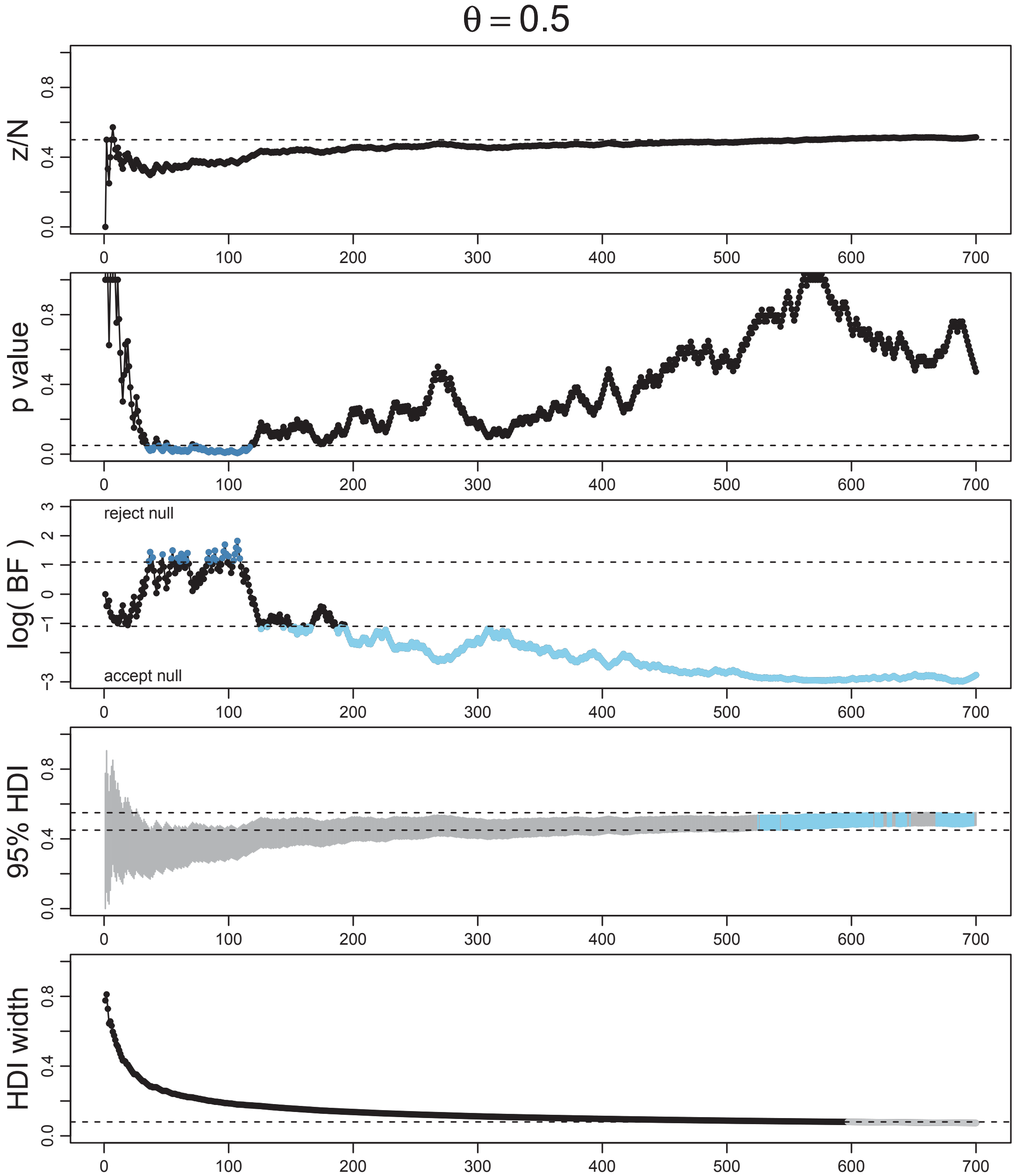

Figure 13.4: An example of a sequence of flips with testing of cumulative data at every flip. The abscissa is N. In this case the null hypothesis is true (i.e., $\theta = 0.50$). This sequence of flips (top panel) happens to show a preponderance of tails early in the sequence, hence both the p value and Bayes factor (BF) reject the null early on. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

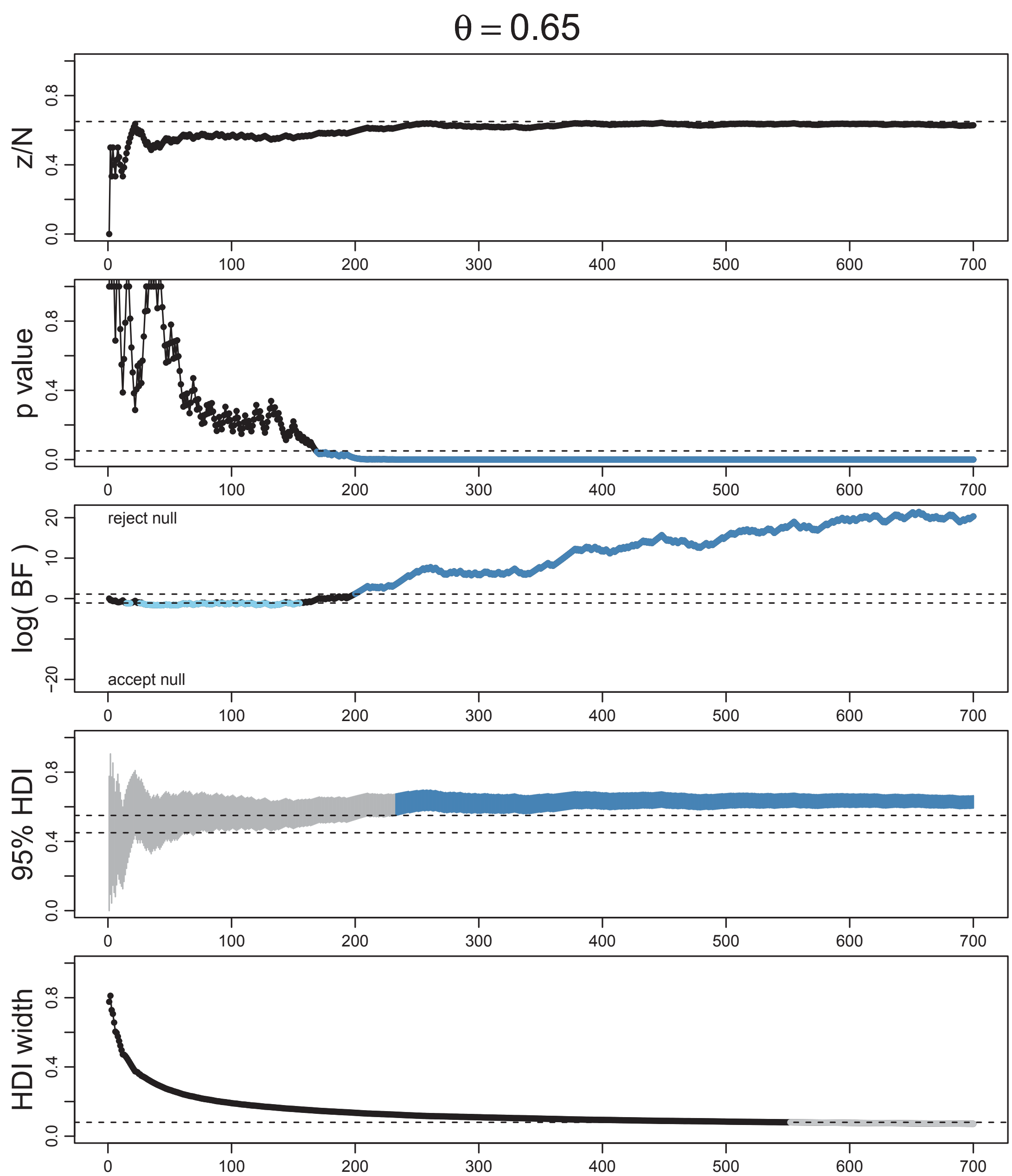

Figure 13.5: An example of a sequence of flips with testing of cumulative data at every flip. The abscissa is N. In this case the null hypothesis is *not* true (i.e., $\theta = 0.65$). This sequence of flips (top panel) happens to show proportions near 0.5 early in the sequence, hence the Bayes factor (BF) accepts the null early on.

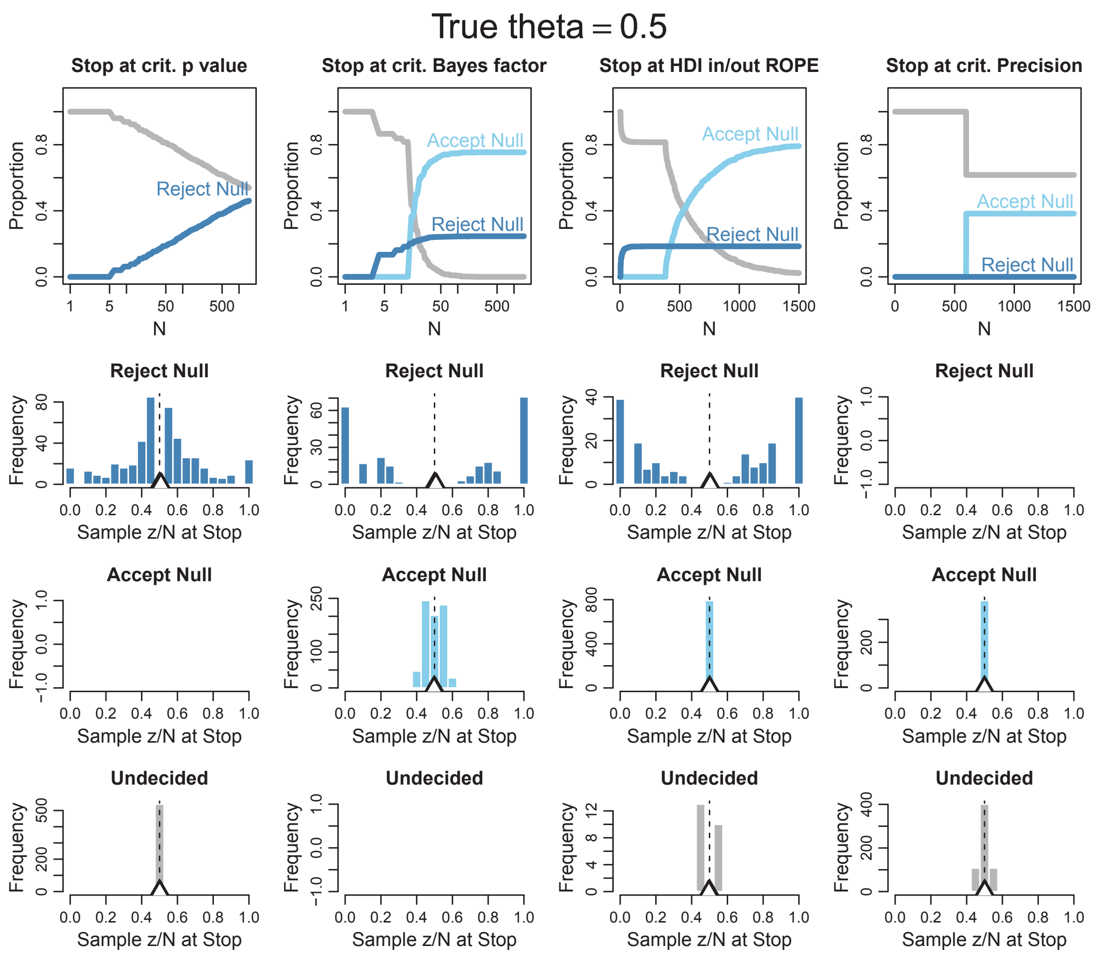

Figure 13.6: Behaviors of the four stopping rules are in the four columns, for $\theta = 0.50$. The top row shows the proportion of 1,000 sequences that make each decision (accept, reject, undecided) at each flip. The lower rows show the value of z/N when data collection stops. Black triangle marks true θ and outline triangle marks mean z/N at stopping. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

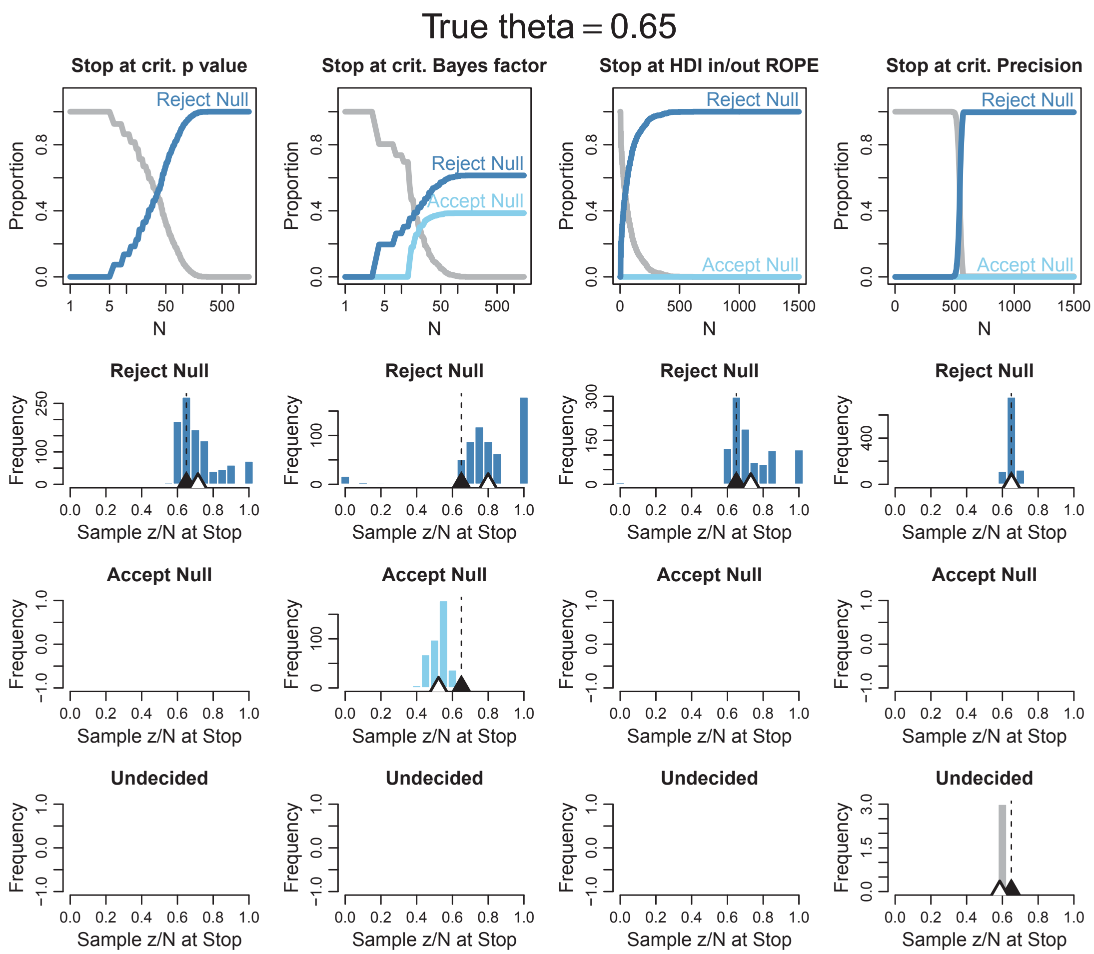

Figure 13.7: Behaviors of the four stopping rules are in the four columns, for $\theta = 0.65$. The top row shows the proportion of 1,000 sequences that make each decision (accept, reject, undecided) at each flip. The lower rows show the value of z/N when data collection stops. Black triangle marks true θ and outline triangle marks mean z/N at stopping. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Table 13.1: Minimal sample size required for 95% HDI to exclude a ROPE from 0.48 to 0.52, when flipping a single coin.

	Generating Mode ω					
Power	**.60**	**.65**	**.70**	**.75**	**.80**	**.85**
.7	238	83	40	25	16	7
.8	309	109	52	30	19	14
.9	430	150	74	43	27	16

Note. The data-generating distribution is a beta density with mode ω, as indicated by the column header, and concentration $\kappa = 2000$. The audience prior is a uniform distribution.

Table 13.2: Minimal sample size required for 95% HDI to have maximal width of 0.2, when flipping a single coin.

	Generating Mode ω					
Power	**.60**	**.65**	**.70**	**.75**	**.80**	**.85**
.7	91	90	88	86	81	75
.8	92	92	91	90	87	82
.9	93	93	93	92	91	89

Note. The data-generating distribution is a beta density with mode ω, as indicated by the column header, and with concentration $\kappa = 10$. The audience-agreeable prior is uniform.

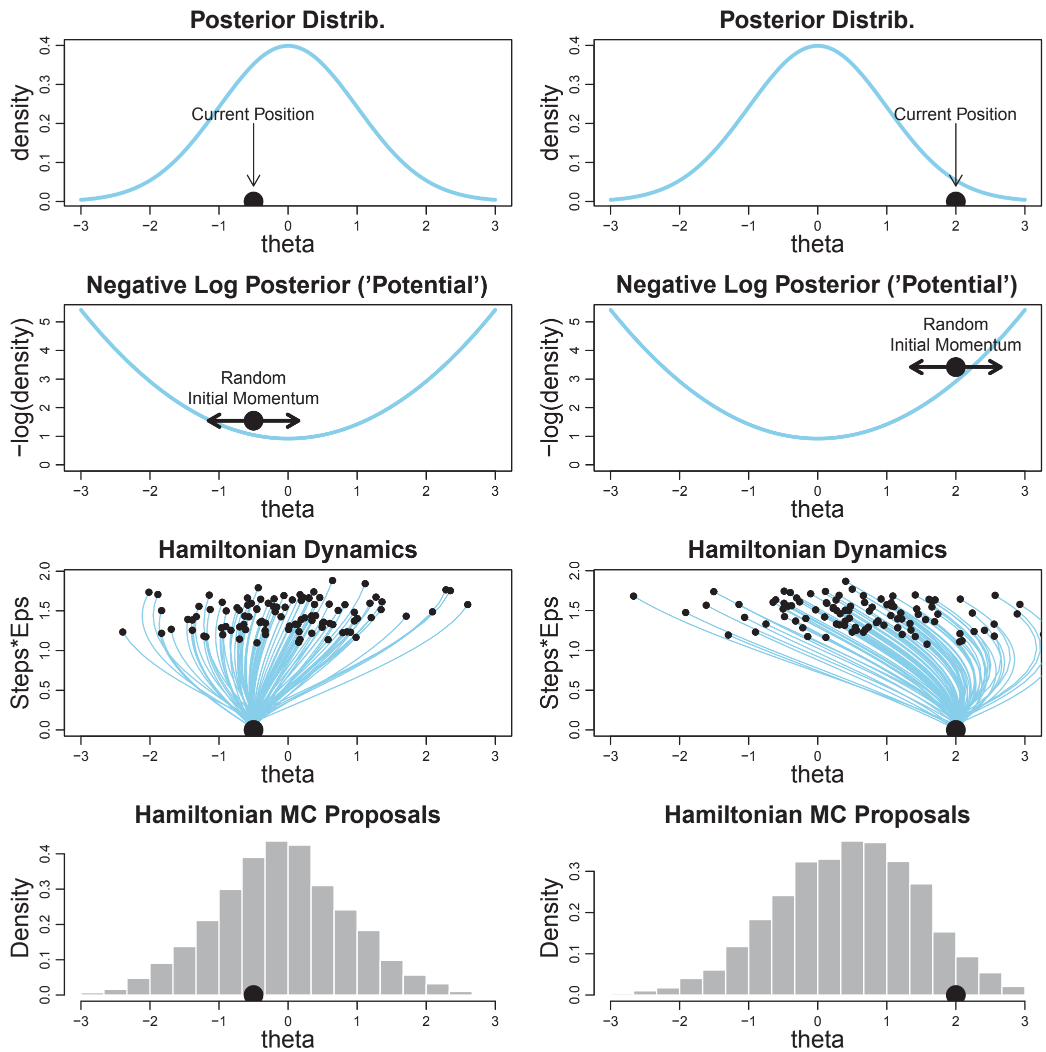

Figure 14.1: Examples of a Hamiltonian Monte Carlo proposal distributions. Two columns show two different current parameter values, marked by the large dots. First row shows posterior distribution. Second row shows the potential energy, with a random impulse given to the dot. Third row shows trajectories, which are the theta value (x-axis) as a function of time (y-axis marked Steps*Eps). Fourth row shows histograms of the proposals. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

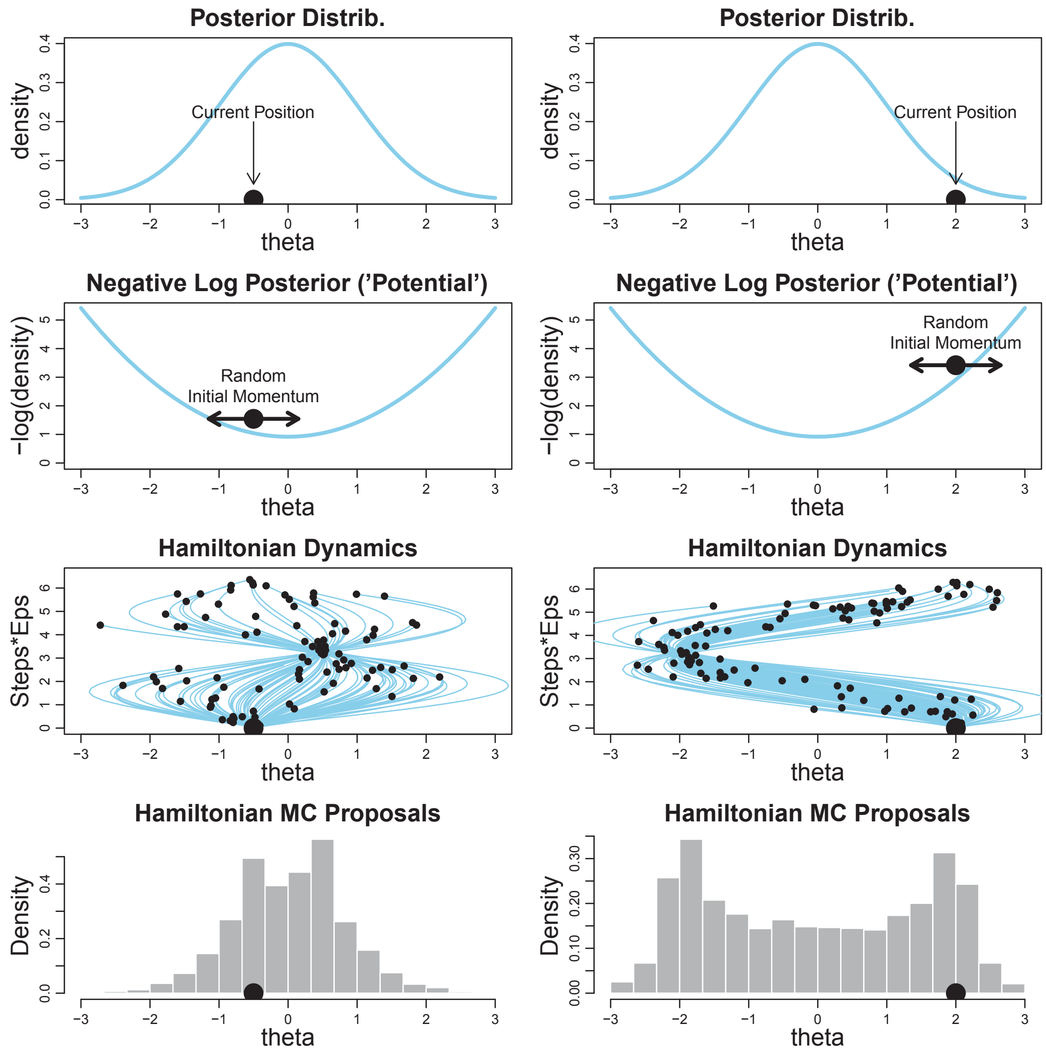

Figure 14.2: Examples of a Hamiltonian Monte Carlo proposal distributions for two different current parameter values, marked by the large dots, in the two columns. For this figure, a large range of random trajectory lengths (Steps*Eps) is sampled. Compare with Figure 14.1. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

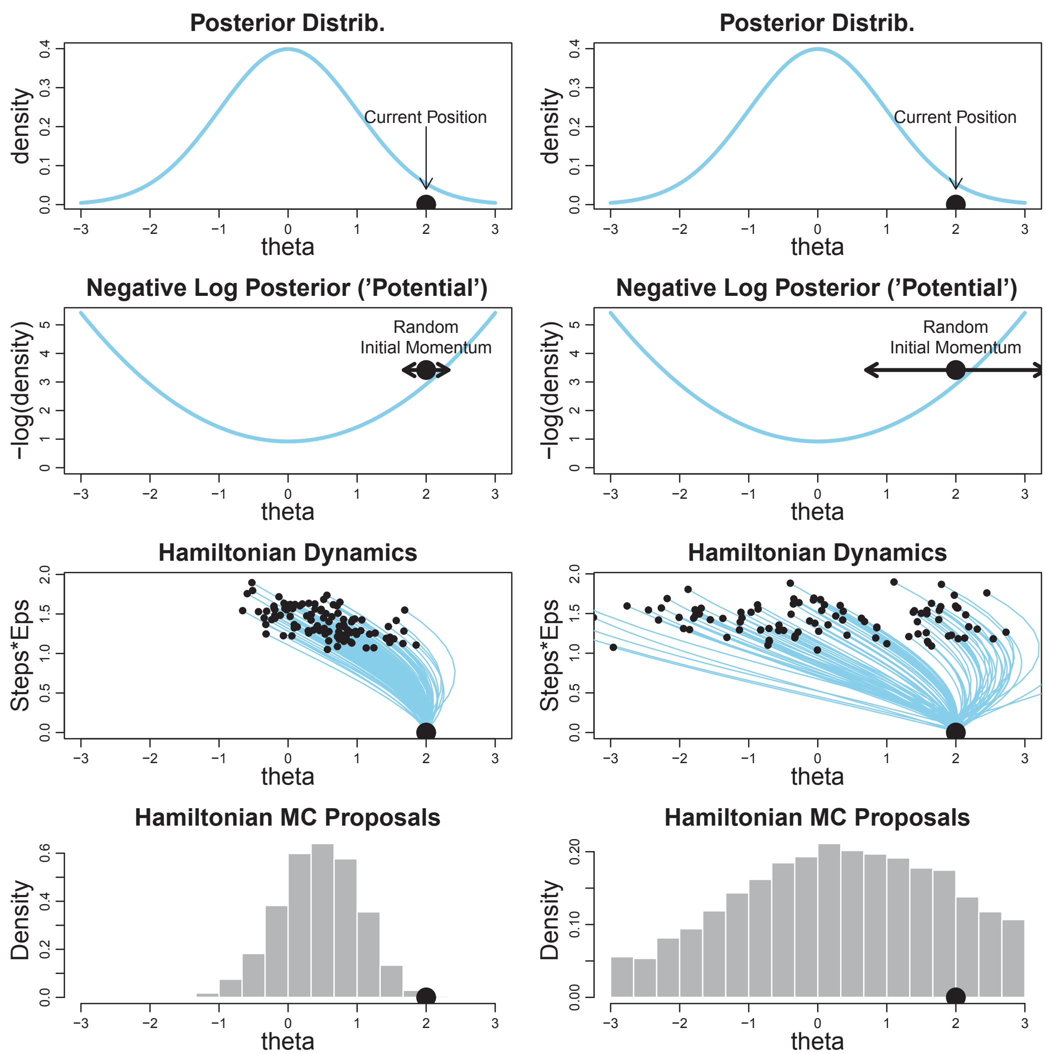

Figure 14.3: Examples of a Hamiltonian Monte Carlo proposal distributions for two different variances of the initial random momentum, indicated in the second row. Compare with Figure 14.1, which shows an intermediate variance of the initial random momentum. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

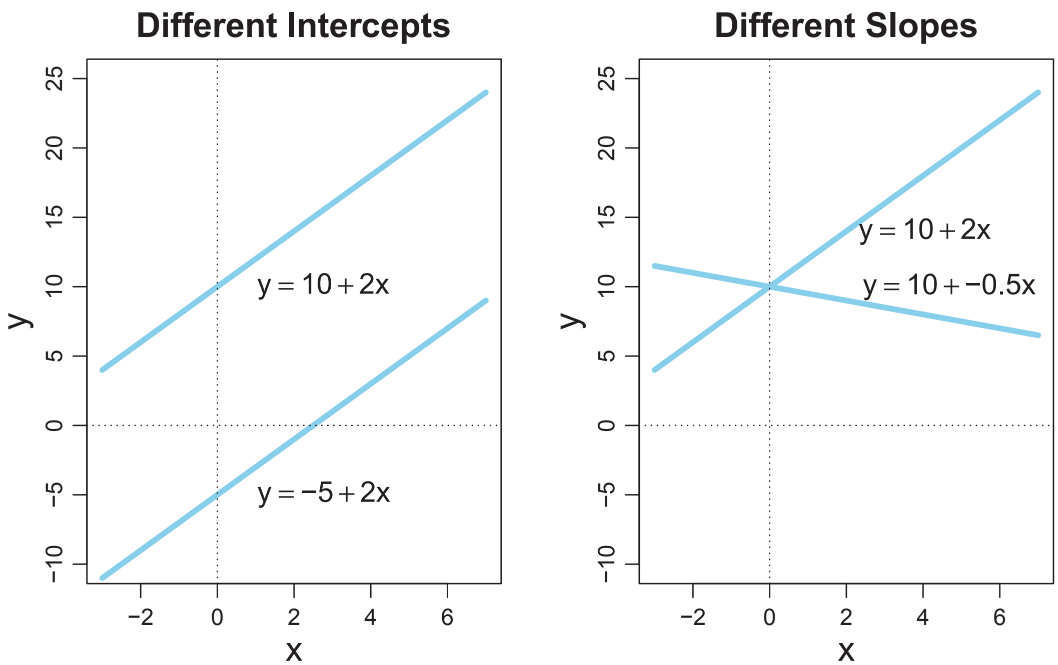

Figure 15.1: Examples of linear functions of a single *x* variable. The left panel shows examples of two lines with the same slope but different intercepts. The right panel shows examples of two lines with the same intercept but different slopes.

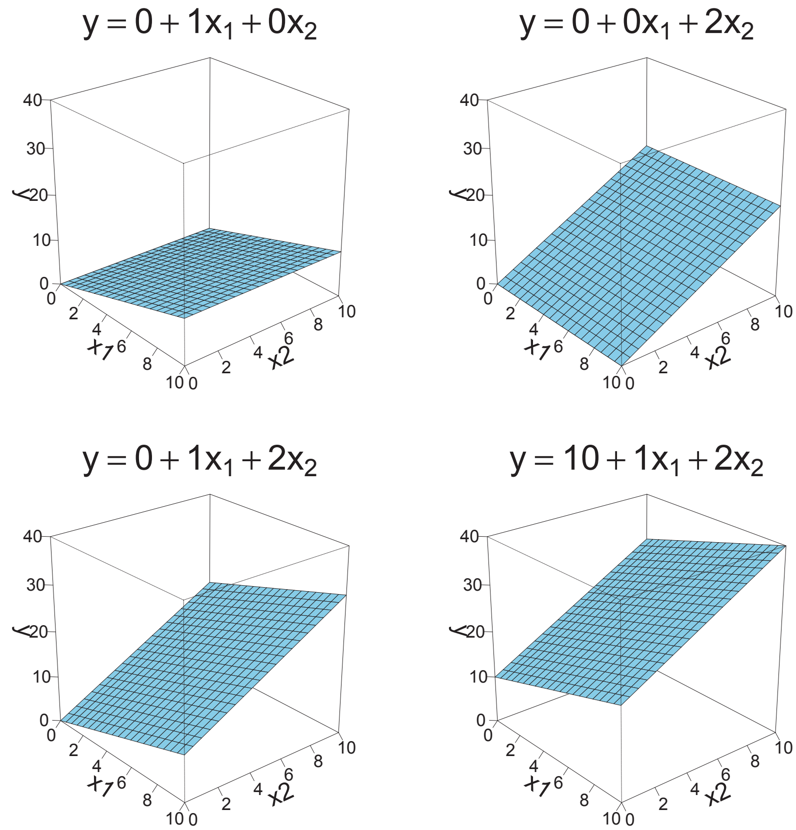

Figure 15.2: Examples of linear functions of two variables, x_1 and x_2. Upper left: Only x_1 has an influence on y. Upper right: Only x_2 has an influence on y. Lower left: x_1 and x_2 have an additive influence on y. Lower right: Non-zero intercept is added.

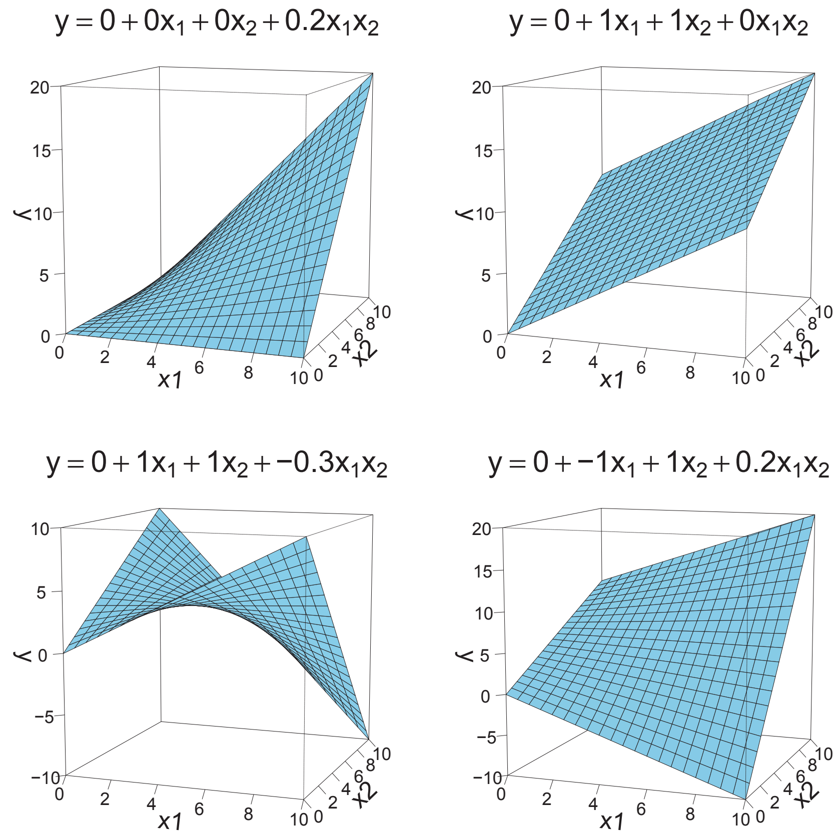

Figure 15.3: Multiplicative interaction of two variables, x_1 and x_2. Upper right panel shows *zero* interaction, for comparison. Figure 18.8, p. 502, provides additional perspective and insight.

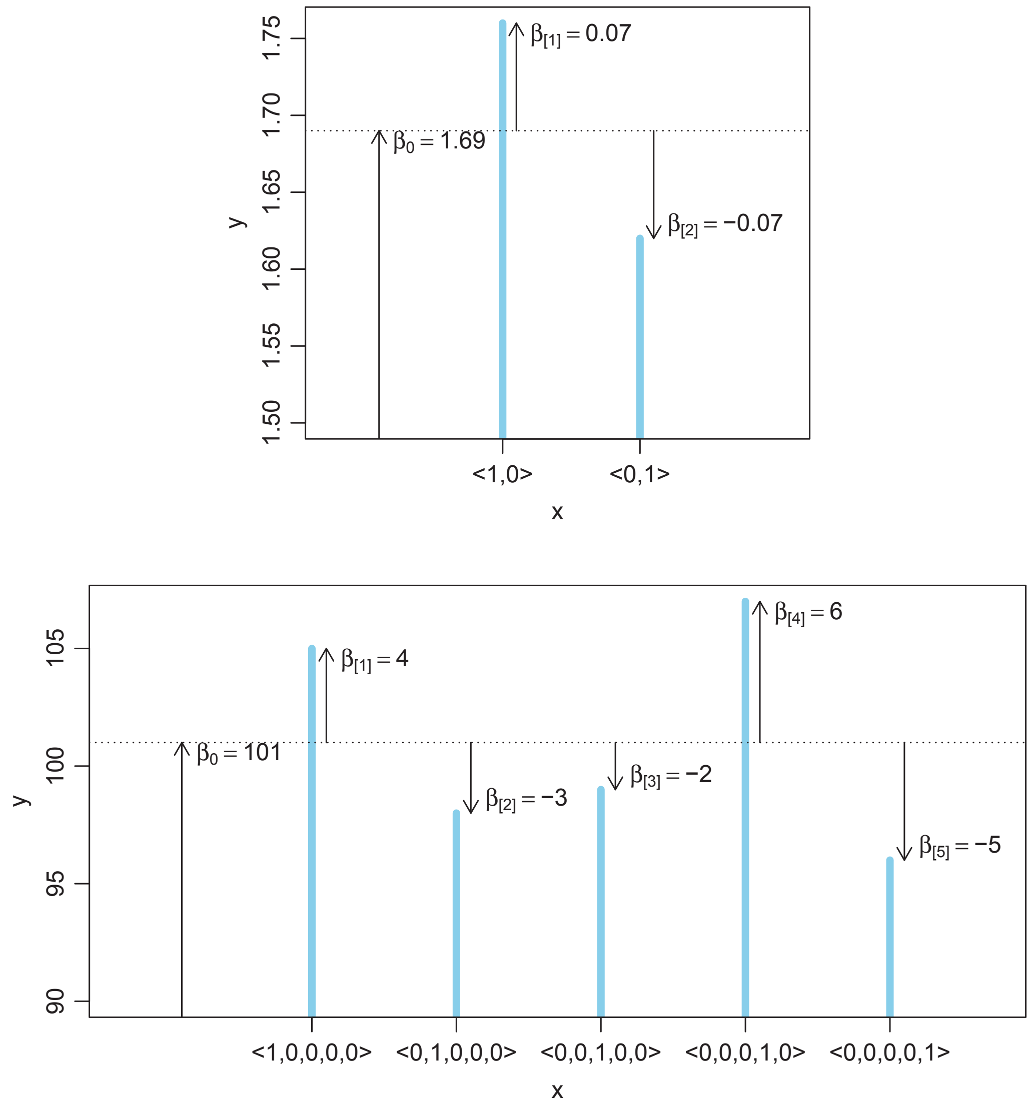

Figure 15.4: Examples of a nominal predictor (Equations 15.3 and 15.4). Upper panel shows a case with $J = 2$, lower panel shows a case with $J = 5$. In each panel, the baseline value of y is on the far left. Notice that the deflections from baseline sum to zero. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

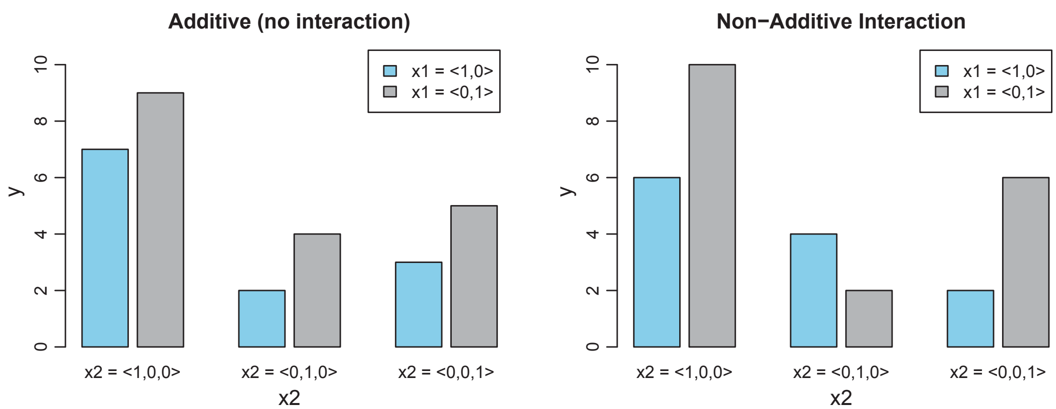

Figure 15.5: Combinations of two nominal variables. *Left:* Additive combination. Notice that the difference between levels of x_1 is the same for every level of x_2. *Right:* Non-additive interaction. Notice that the difference between levels of x_1 is *not* the same for every level of x_2. The labels elevate subscripts for readability; thus x_1 is displayed as $x1$ and x_2 is displayed as $x2$. Figure 20.1, p. 561, provides additional perspective and insight. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

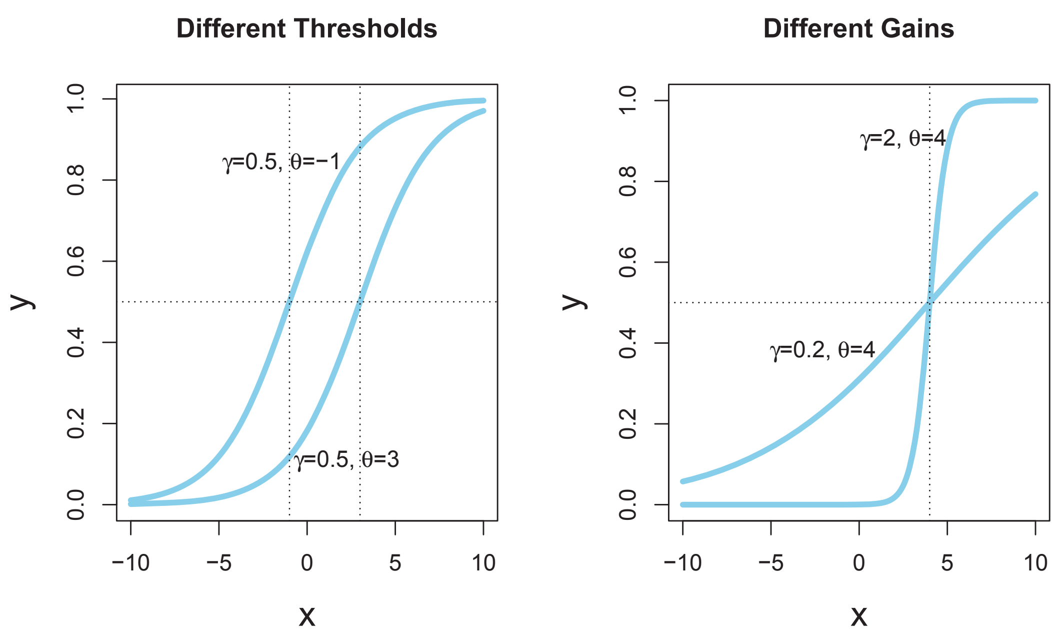

Figure 15.6: Examples of logistic functions of a single variable. The left panel shows logistics with the same gain but different thresholds. The right panel shows logistics with the same threshold but different gains. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

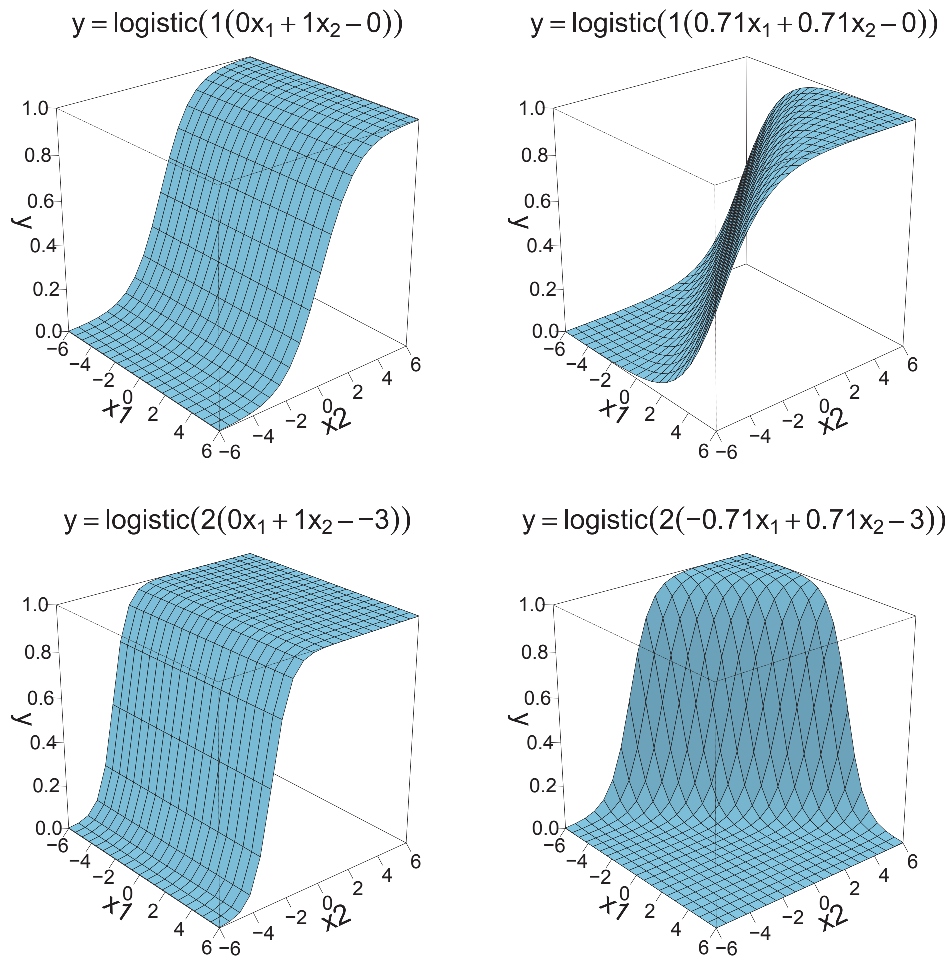

Figure 15.7: Examples of logistic functions of two variables. Top two panels show logistics with the same gain and threshold, but different coefficients on the predictors. The left two panels show logistics with the same coefficients on the predictors, but different gains and thresholds. The lower right panel shows a case with a negative coefficient on the first predictor. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

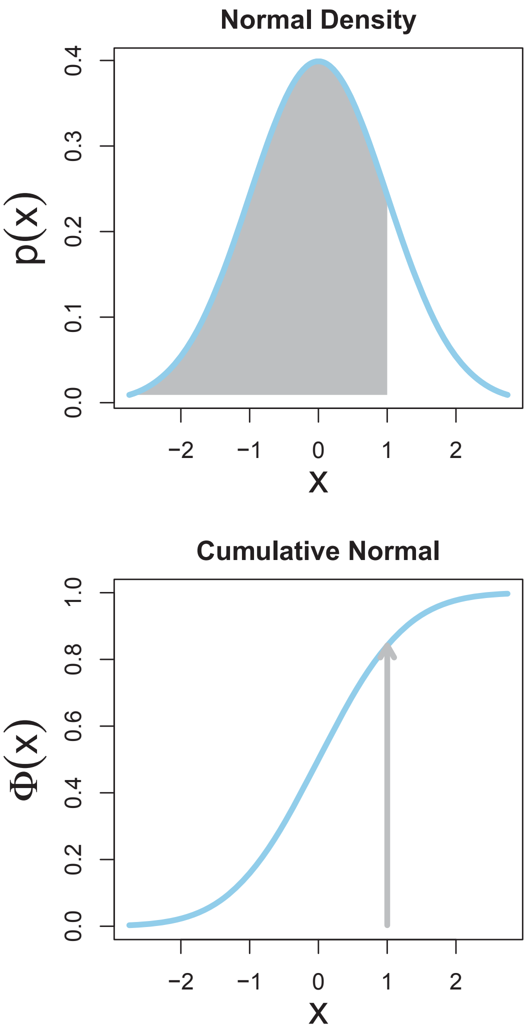

Figure 15.8: Top: A standardized normal density (i.e., with mean zero and standard deviation one). Bottom: The corresponding standardized cumulative normal function. The area under the normal distribution to the left of x (top panel) is the height of the cumulative normal at x (lower panel).

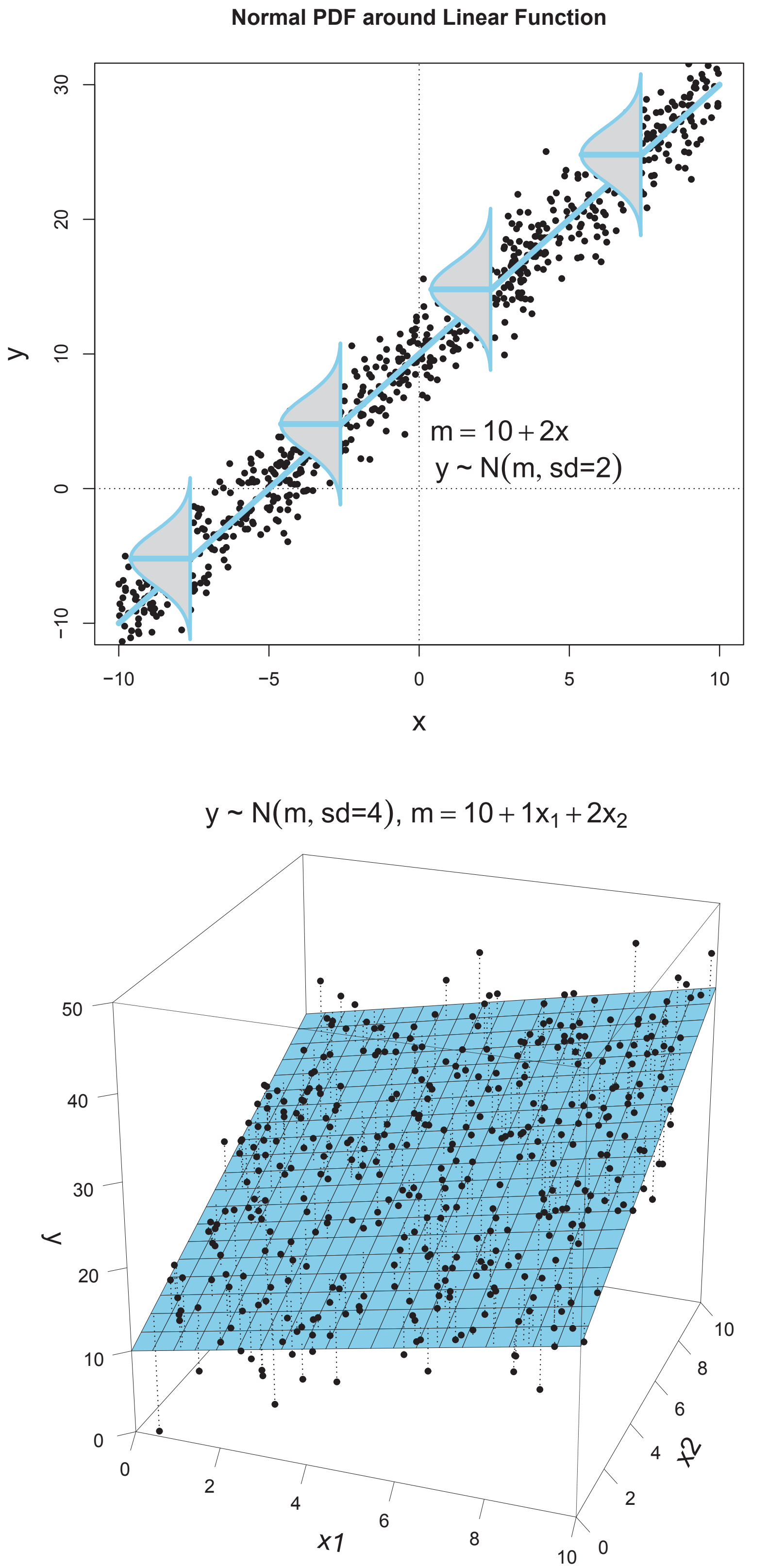

Figure 15.9: Examples of points normally distributed around a linear function. The upper panel shows normal distributions superimposed on the linear function to emphasize that the random variability is vertical, along the *y* axis, and centered on the line. The lower panel shows each datum connected to the plane by a dotted line, to again emphasize the vertical displacement from the plane. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

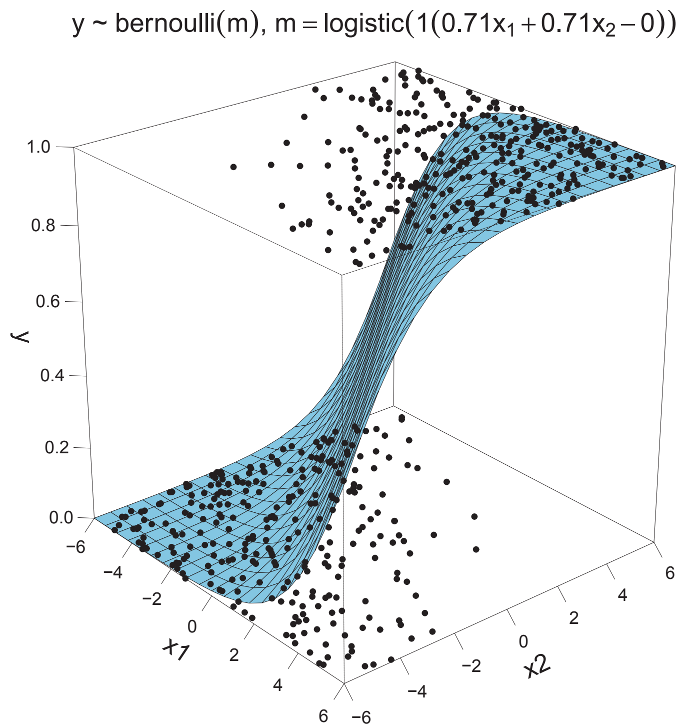

Figure 15.10: Examples of points that are Bernoulli distributed around a logistic function of two predictors. All the points are either at $y = 1$ or $y = 0$; intermediate values such as $y = .6$ cannot occur.

Table 15.1: For the generalized linear model: Typical linear functions lin(x) of the predictor variables x, for various scale types of x. The value lin(x) is mapped to the predicted data by functions shown in Table 15.2.

Scale Type of Predictor x					
Metric				Nominal	
Single Group	Two Groups	Single Predictor	Multiple Predictors	Single Factor	Multiple Factors
β_0	$\beta_{x=1}$ $\beta_{x=2}$	β_0 $+\beta_1 x$	β_0 $+ \sum_k \beta_k x_k$ $+ \sum_{j,k} \beta_{j\times k} x_j x_k$ $+ \left[\begin{array}{c} \text{higher-order} \\ \text{interactions} \end{array} \right]$	β_0 $+\vec{\beta} \cdot \vec{x}$	β_0 $+ \sum_k \vec{\beta}_k \cdot \vec{x}_k$ $+ \sum_{j,k} \vec{\beta}_{j\times k} \cdot \vec{x}_{j\times k}$ $+ \left[\begin{array}{c} \text{higher-order} \\ \text{interactions} \end{array} \right]$

Table 15.2: For the generalized linear model: typical noise distributions and inverse-link functions for describing various scale types of the predicted variable y. The value μ is a central tendency of the predicted data (not necessarily the mean). The predictor variable is x, and lin(x) is a linear function of x, such as those shown in Table 15.1.

Scale Type of Predicted y	Typical Noise Distribution $y \sim \text{pdf}\,(\mu, [\text{parameters}])$	Typical Inverse-Link Function $\mu = f\,(\text{lin}(x), [\text{parameters}])$
Metric	$y \sim \text{normal}(\mu, \sigma)$	$\mu = \text{lin}(x)$
Dichotomous	$y \sim \text{bernoulli}(\mu)$	$\mu = \text{logistic}\,(\text{lin}(x))$
Nominal	$y \sim \text{categorical}(\ldots, \mu_k, \ldots)$	$\mu_k = \dfrac{\exp(\text{lin}_k(x))}{\sum_c \exp(\text{lin}_c(x))}$
Ordinal	$y \sim \text{categorical}(\ldots, \mu_k, \ldots)$	$\mu_k = \begin{array}{l} \Phi\,((\theta_k - \text{lin}(x))\,/\sigma) \\ -\,\Phi\,((\theta_{k-1} - \text{lin}(x))\,/\sigma) \end{array}$
Count	$y \sim \text{poisson}(\mu)$	$\mu = \exp\,(\text{lin}(x))$

Table 15.3: Book chapters that discuss combinations of scale types for predicted and predictor variables of Tables 15.2 and 15.1. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

	Scale Type of Predictor x					
			Metric		Nominal	
Scale Type of Predicted y	Single Group	Two Groups	Single Predictor	Multiple Predictors	Single Factor	Multiple Factors
Metric	Ch. 16		Ch. 17	Ch. 18	Ch. 19	Ch. 20
Dichotomous	Ch's 6–9		Ch. 21			
Nominal	Ch. 22					
Ordinal	Ch. 23					
Count	Ch. 24					

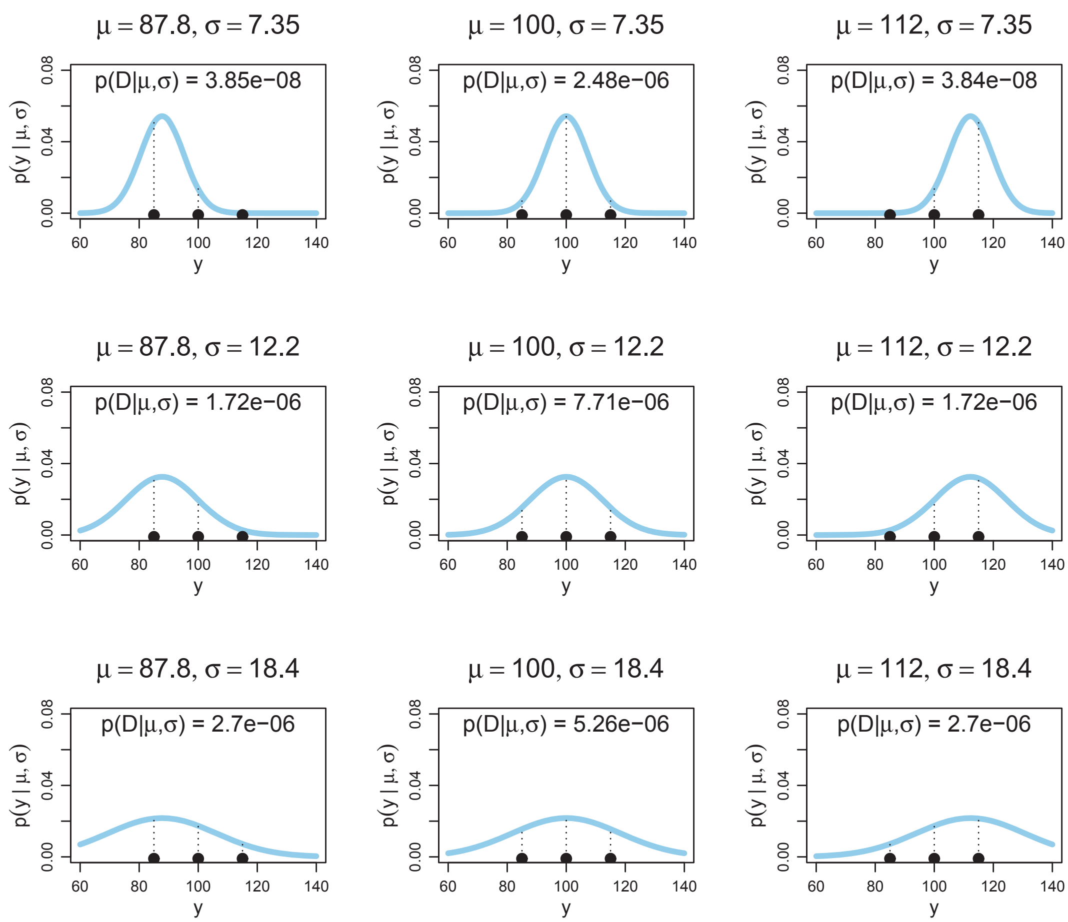

Figure 16.1: The likelihood $p(D|\mu, \sigma)$ for three data points, $D = \{85, 100, 115\}$, according to a normal likelihood function with different values of μ and σ. Columns show different values of μ, and rows show different values of σ. The probability density of an individual datum is the height of the dotted line over the point. The probability of the set of data is the product of the individual probabilities. The middle panel shows shows the μ and σ that maximize the probability of the data. (For another example, see Figure 2.4, p. 27.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

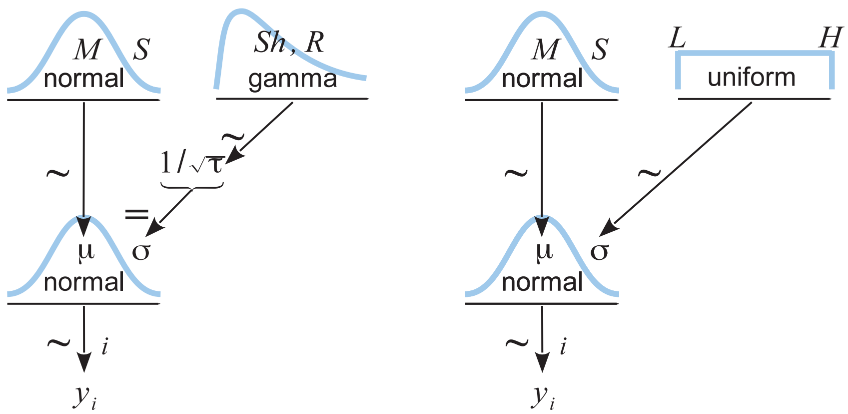

Figure 16.2: Dependencies of variables for metric data described by a normal distribution. The left panel puts a gamma prior on the precision $\tau = 1/\sigma^2$. The right panel puts a uniform prior on σ. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

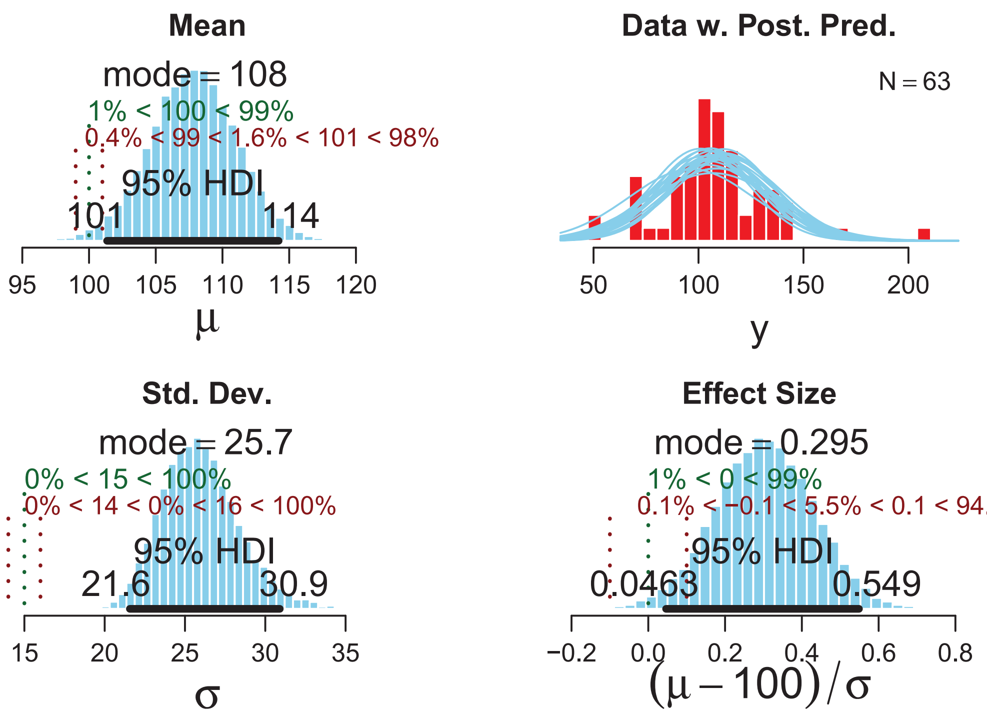

Figure 16.3: Posterior distribution of `Jags-Ymet-Xnom1grp-Mnormal` applied to fictitious IQ data from a "smart drug" group. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

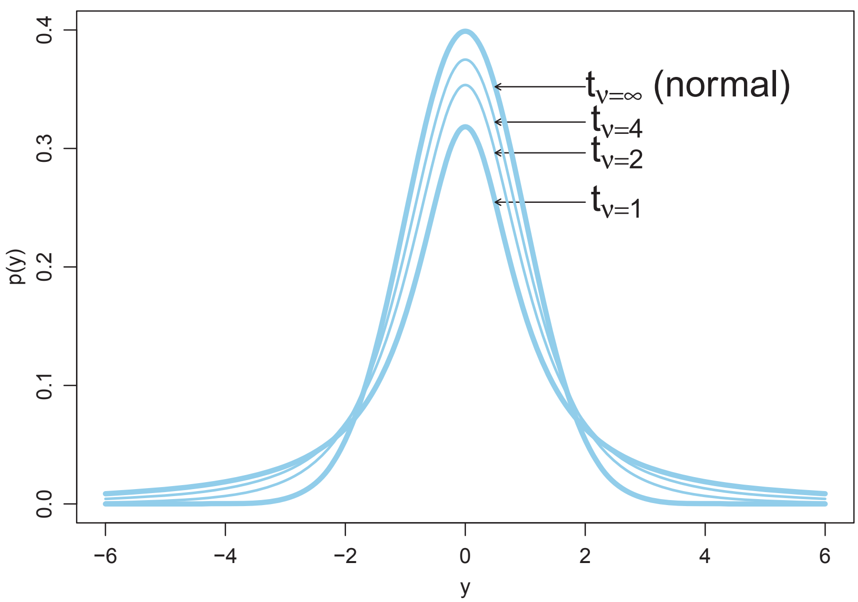

Figure 16.4: Examples of t distributions. In all cases, $\mu = 0$ and $\sigma = 1$. The normality parameter, ν, controls the heaviness of the tails. Curves for different values of ν are superimposed for easy comparison. The abscissa is labelled as y (not x) because the distribution is intended to describe predicted data.

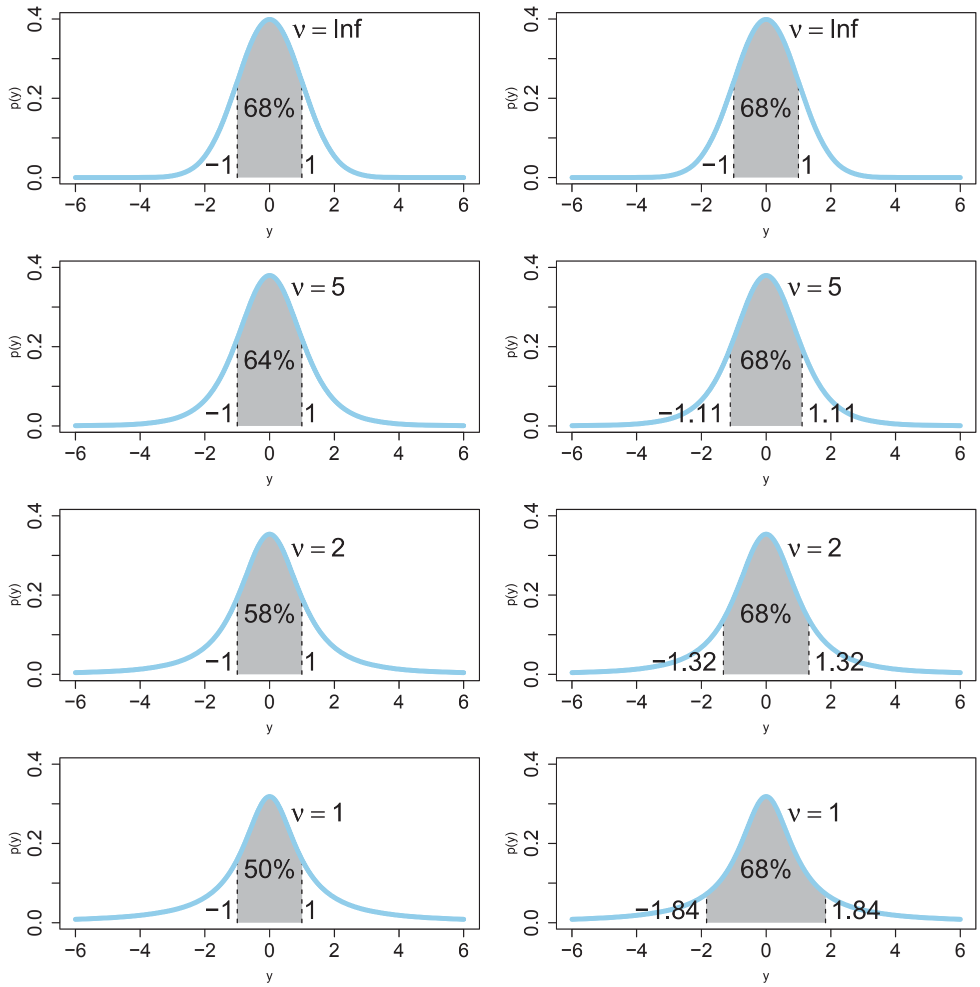

Figure 16.5: Examples of *t* distributions with areas under the curve. In all cases, $\mu = 0$ and $\sigma = 1$. Rows show different values of the normality parameter, ν. Left column shows area under the *t* distribution from $y = -1$ to $y = +1$. Right column shows values of $\pm y$ needed for an area of 68.27%, which is the area under a standardized normal curve from $y = -1$ to $y = +1$. The abscissa is labelled as y (not x) because the distribution is intended to describe predicted data.

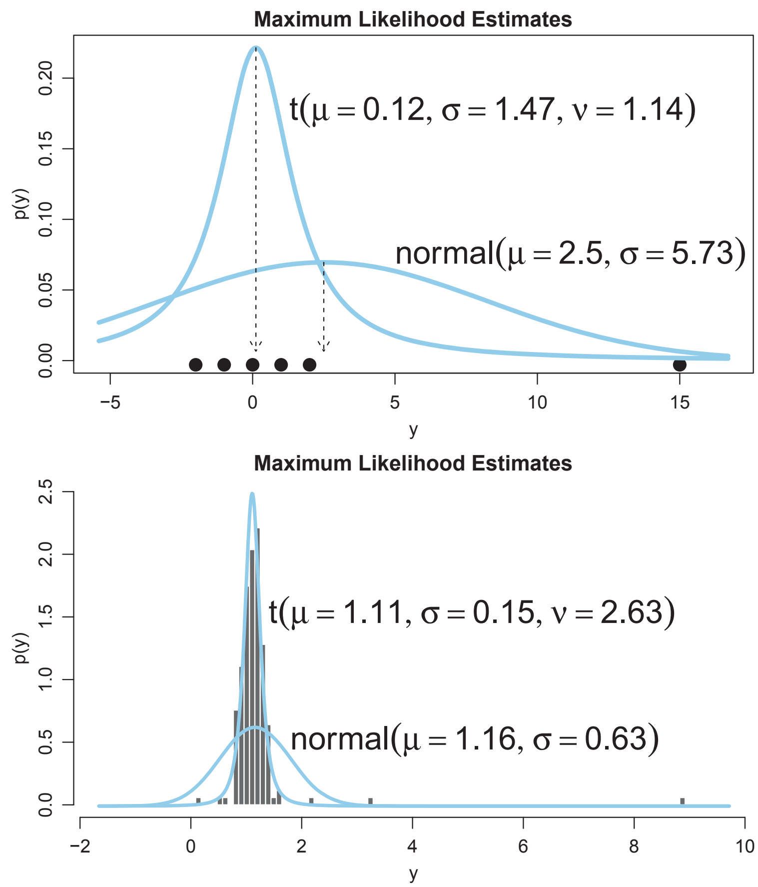

Figure 16.6: The maximum likelihood estimates of normal and *t* distributions fit to the data shown. Upper panel shows "toy" data to illustrate that the normal accommodates an outlier only by enlarging its standard deviation and, in this case, by shifting its mean. Lower panel shows actual data (Holcomb & Spalsbury, 2005) to illustrate realistic effect of outliers on estimates of the normal.

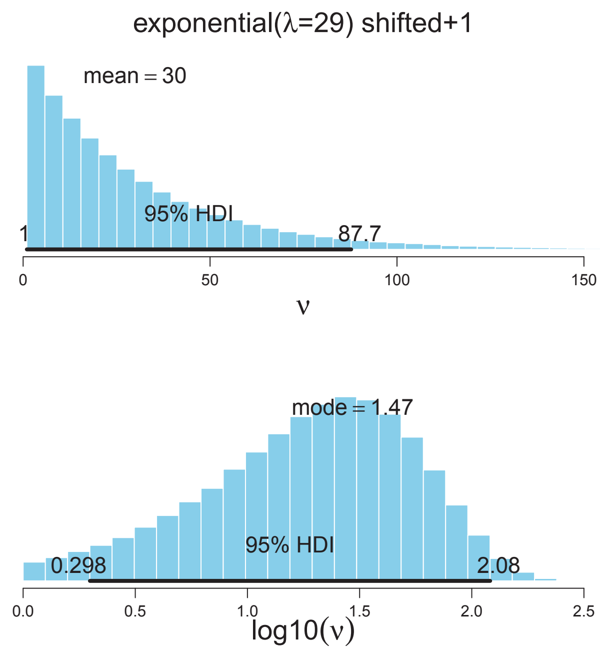

Figure 16.7: The prior on the normality parameter. Upper panel shows the shifted exponential distribution on the original scale of ν. Lower panel shows the same distribution on a logarithmic scale.

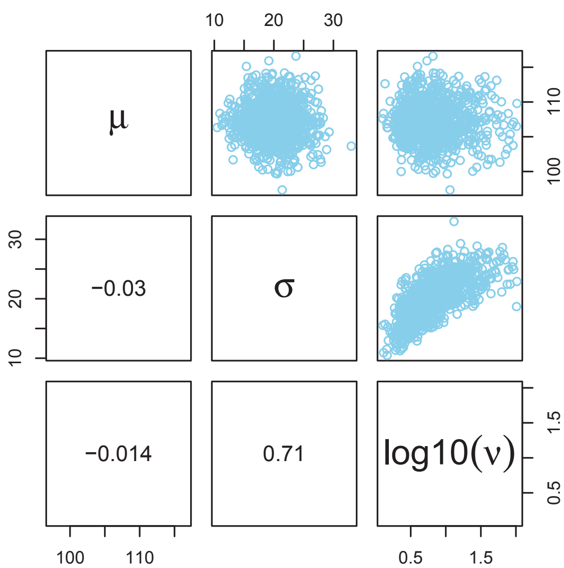

Figure 16.8: Posterior distribution of `Jags-Ymet-Xnom1grp-Mrobust` applied to fictitious IQ data from a "smart drug" group. Off-diagonal cells show scatter plots and correlations of parameters indicated in the corresponding diagonal cells. Notice the strong positive correlation of σ and log10(ν). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

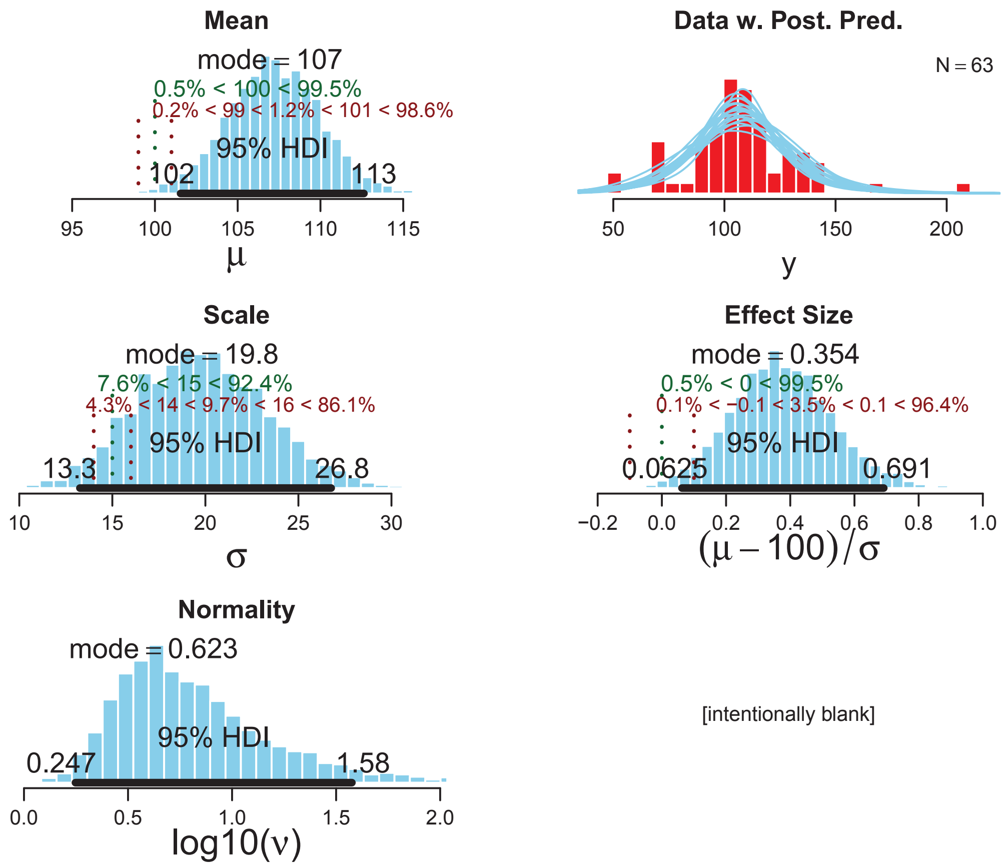

Figure 16.9: Posterior distribution of `Jags-Ymet-Xnom1grp-Mrobust` applied to fictitious IQ data from a "smart drug" group. Compare with Figure 16.3. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

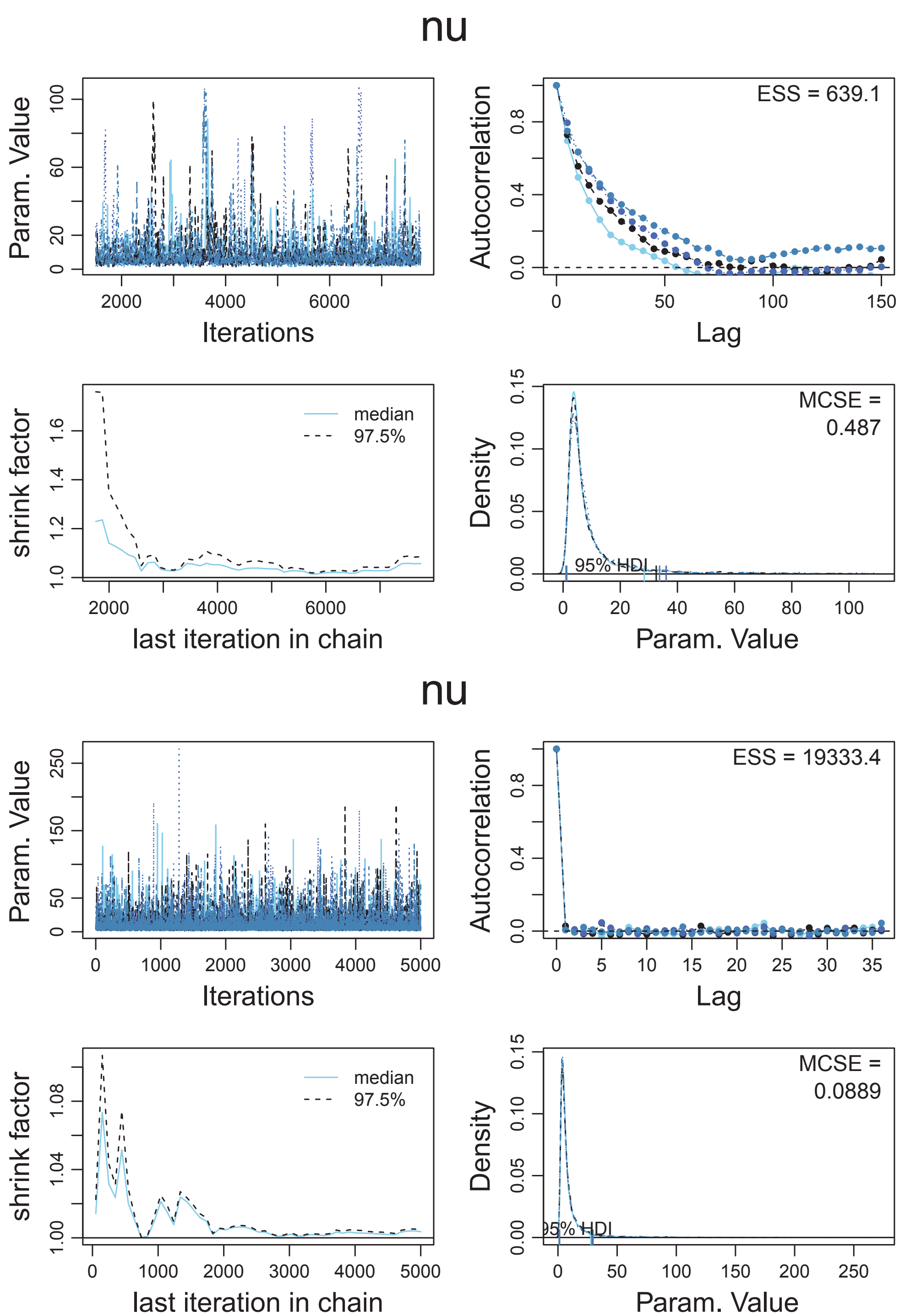

Figure 16.10: Chain diagnostics for JAGS (above) and Stan (below). Notice difference in autocorrelation in the upper right panels, and the resulting ESS. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

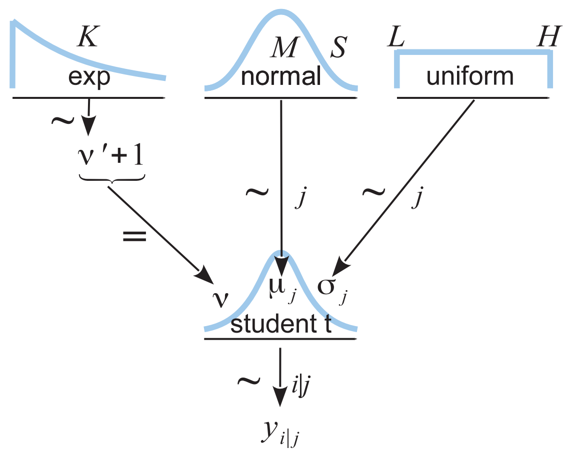

Figure 16.11: Dependency diagram for robust estimation of two groups. At the bottom of the diagram, $y_{i|j}$ is the i^{th} datum within the j^{th} group. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

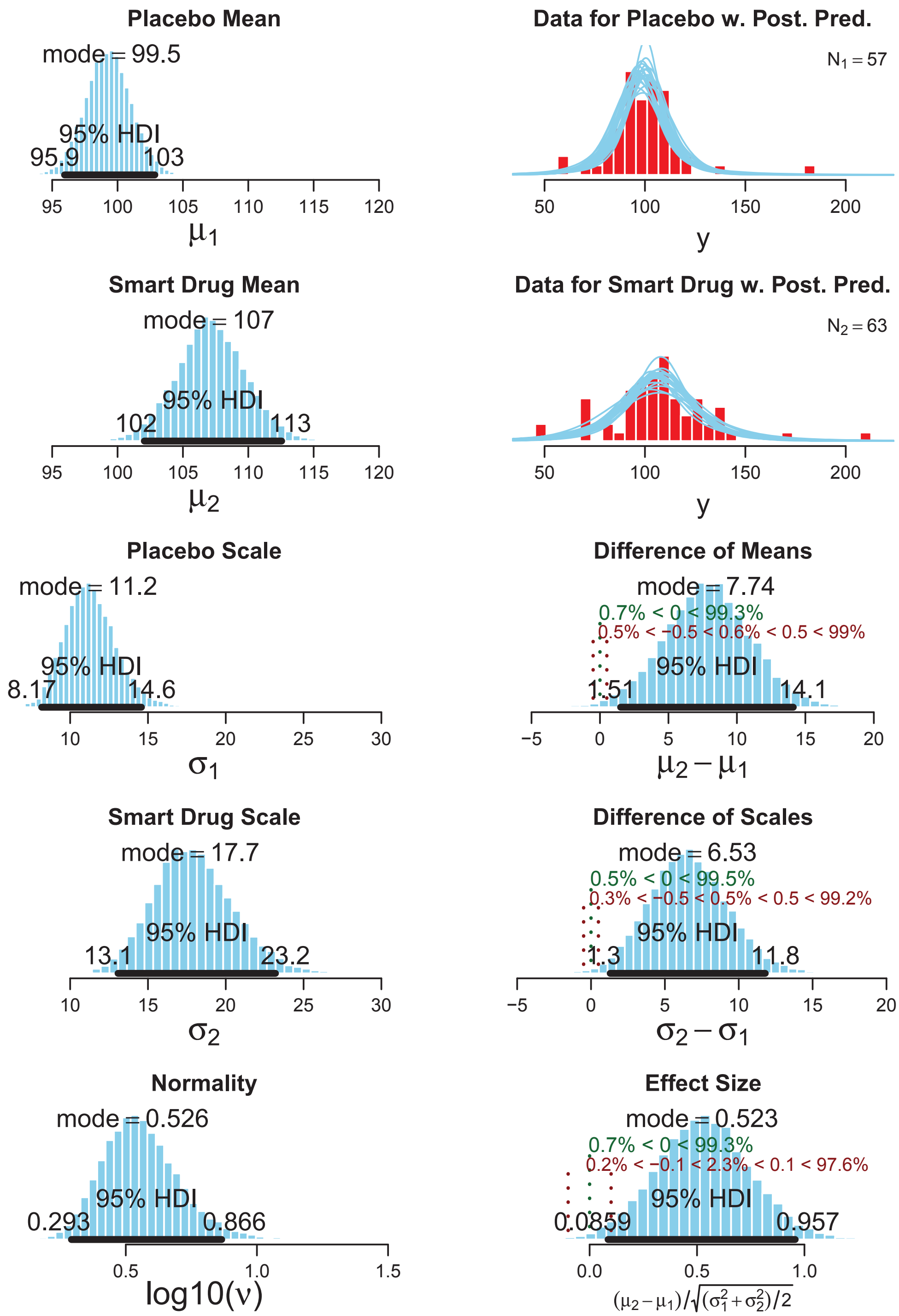

Figure 16.12: Posterior distribution for two groups. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

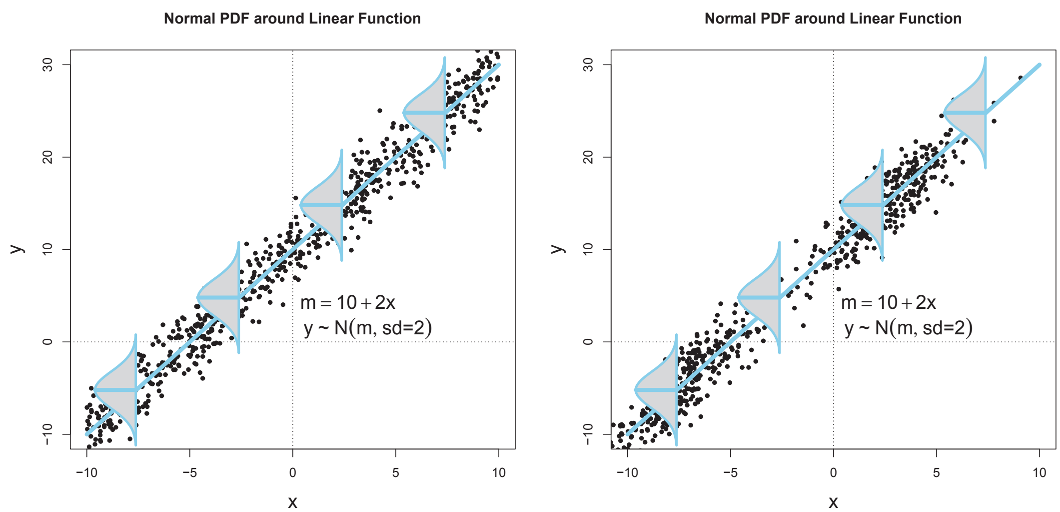

Figure 17.1: Examples of points normally distributed around a linear function. (The left panel repeats Figure 15.9, p. 406.) The model assumes that the data *y* are normally distributed vertically around the line, as shown. Moreover, the variance of *y* is the same at all values of *x*. The model puts no constraints on the distribution of *x*. The right panel shows a case in which *x* are distributed bimodally, whereas in the left panel the *x* are distributed uniformly. In both panels, there is homogeneity of variance. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

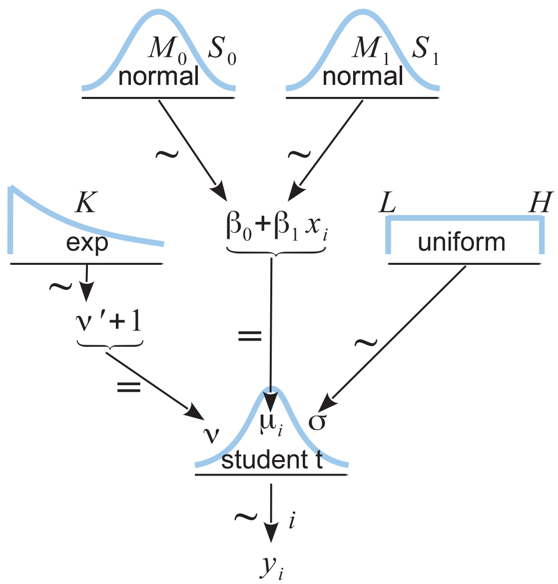

Figure 17.2: A model of dependencies for robust linear regression. The datum, y_i at the bottom of the diagram, is distributed around the central tendency μ_i, which is a linear function of x_i. Compare with Figure 16.11 on p. 437. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

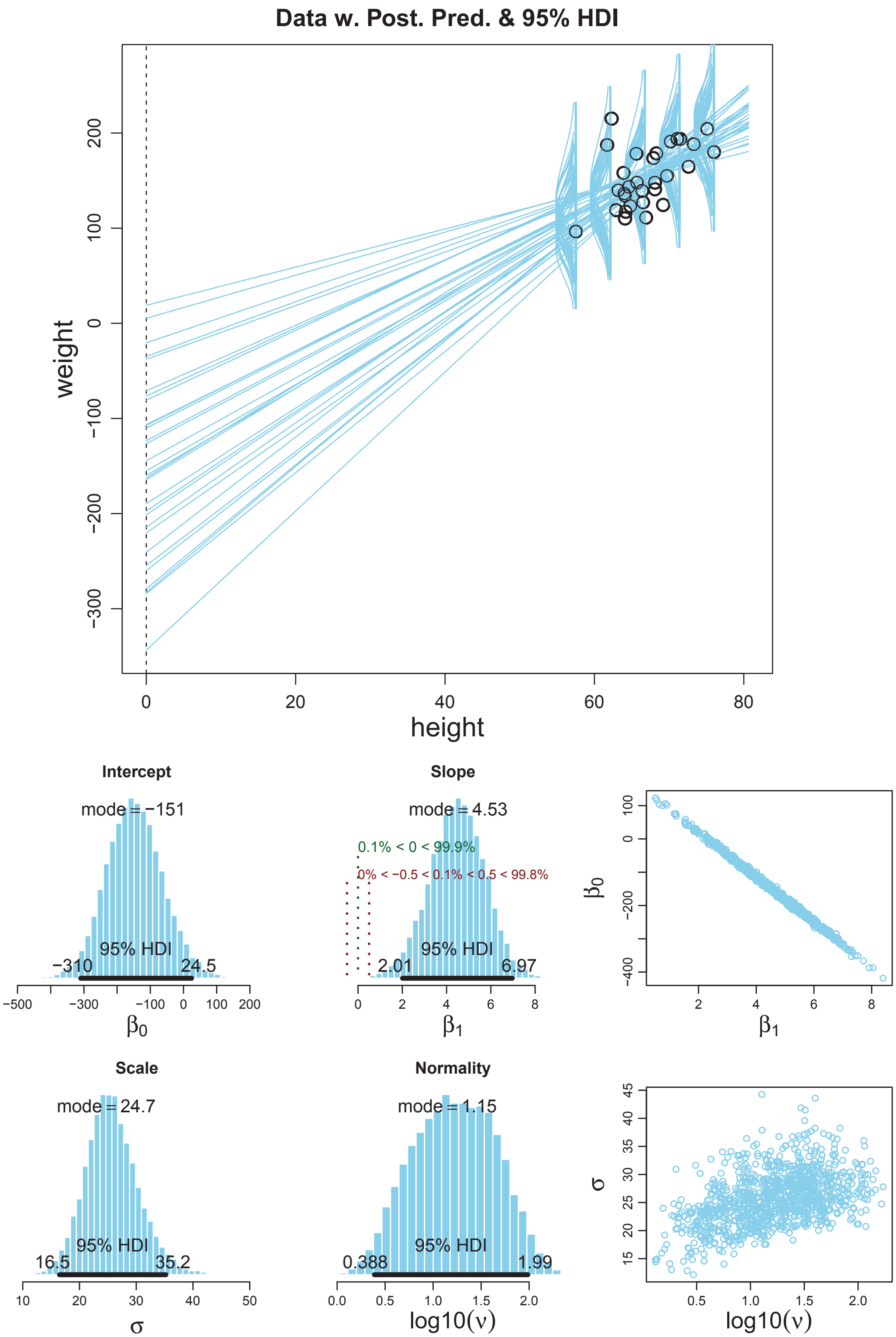

Figure 17.3: Upper panel: Data ($N = 30$) with a smattering of credible regression lines and t noise distributions superimposed. Lower panels: Marginal posterior distribution on parameters. Compare with Figure 17.4.

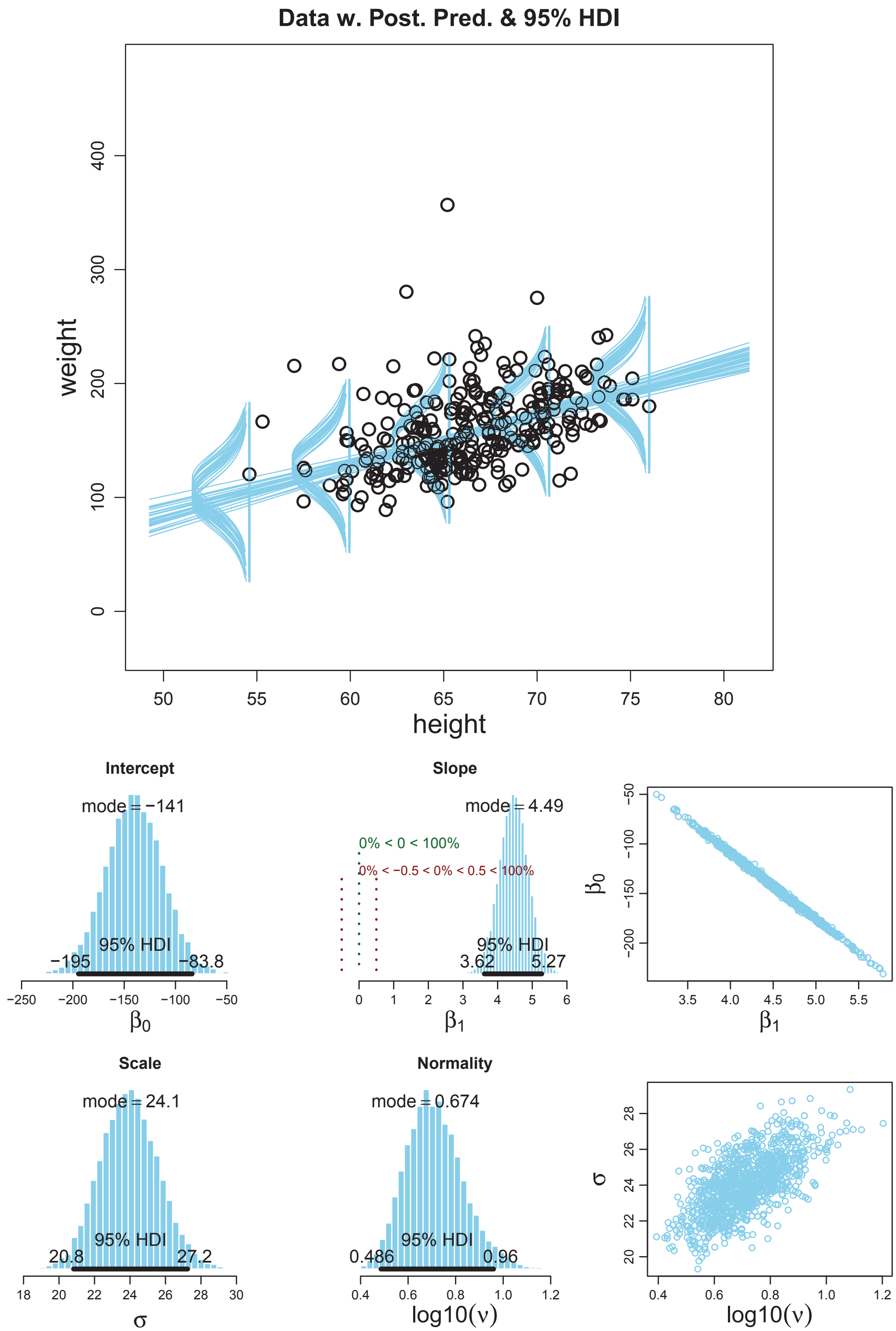

Figure 17.4: Upper panel: Data ($N = 300$) with a smattering of credible regression lines and t noise distributions superimposed. Lower panels: Marginal posterior distribution on parameters. Compare with Figure 17.3. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

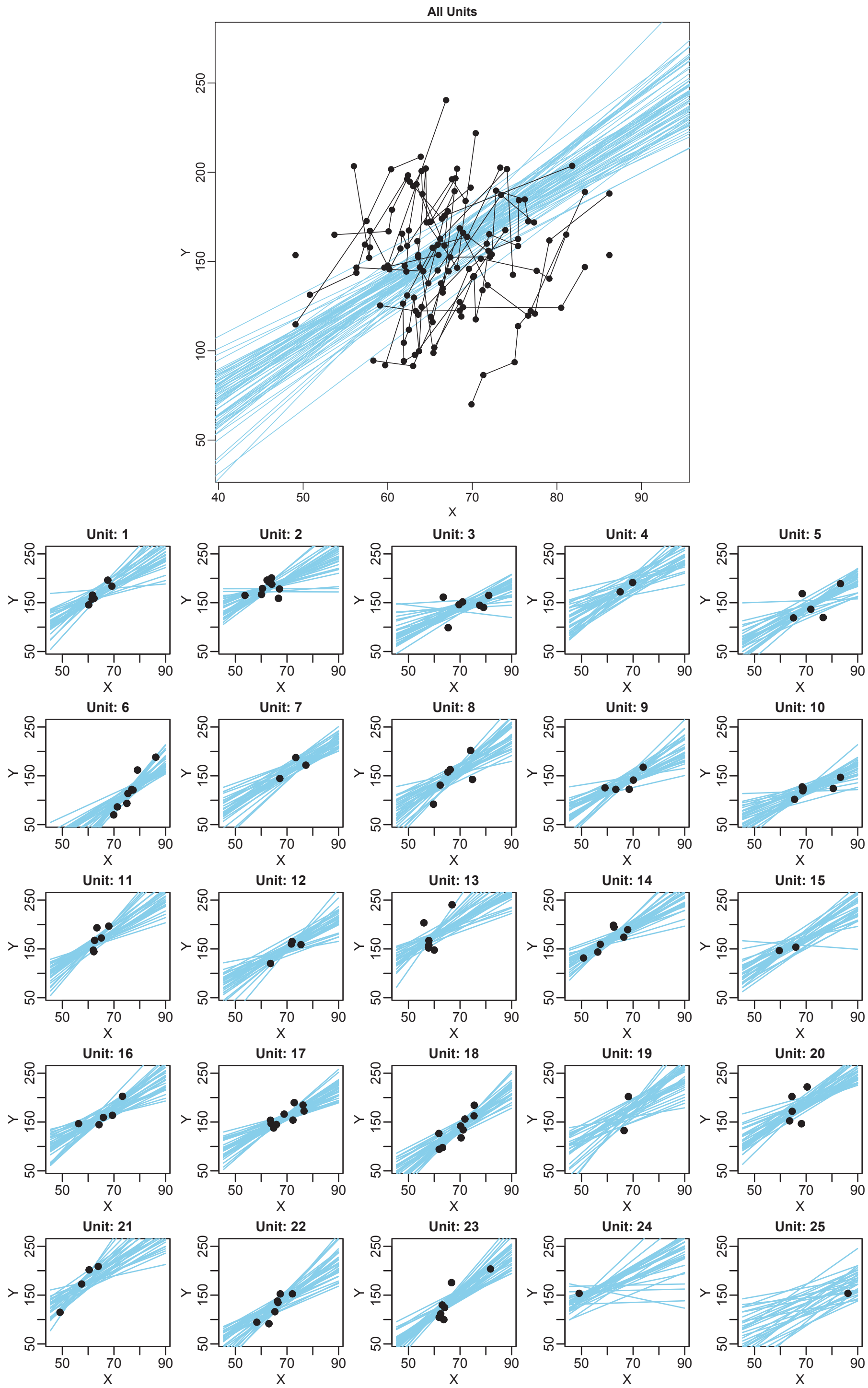

Figure 17.5: Fictitious data for demonstrating hierarchical linear regression, with posterior predicted lines superimposed. Upper panel: All data together, with individuals represented by connected segments. Lower panels: Plots of individual data. *Notice that the final two subjects have only single data points, yet the hierarchical model has fairly tight estimates of the individual slopes and intercepts.* Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

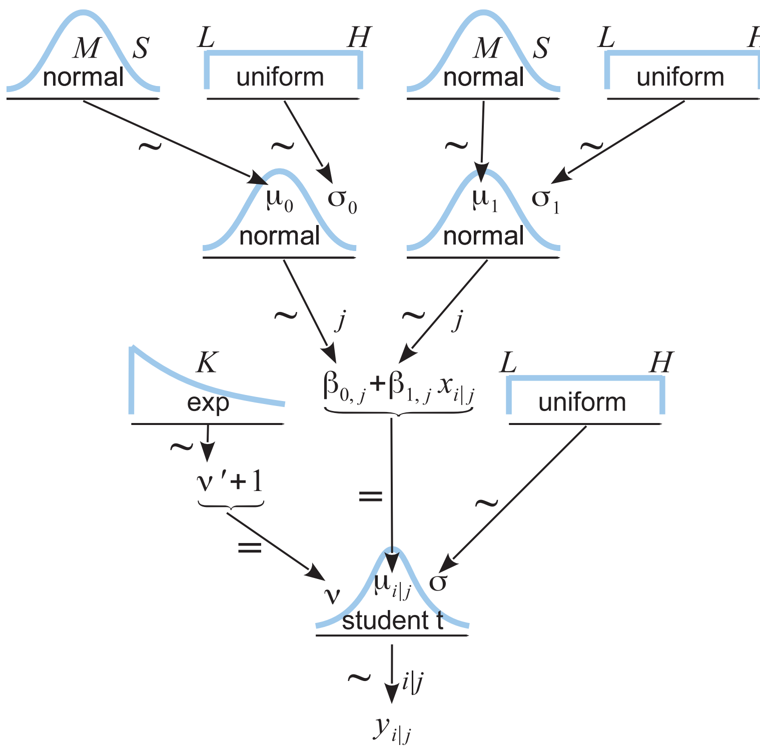

Figure 17.6: A model of dependencies for robust hierarchical linear regression. Compare with Figure 17.2 on p. 463. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

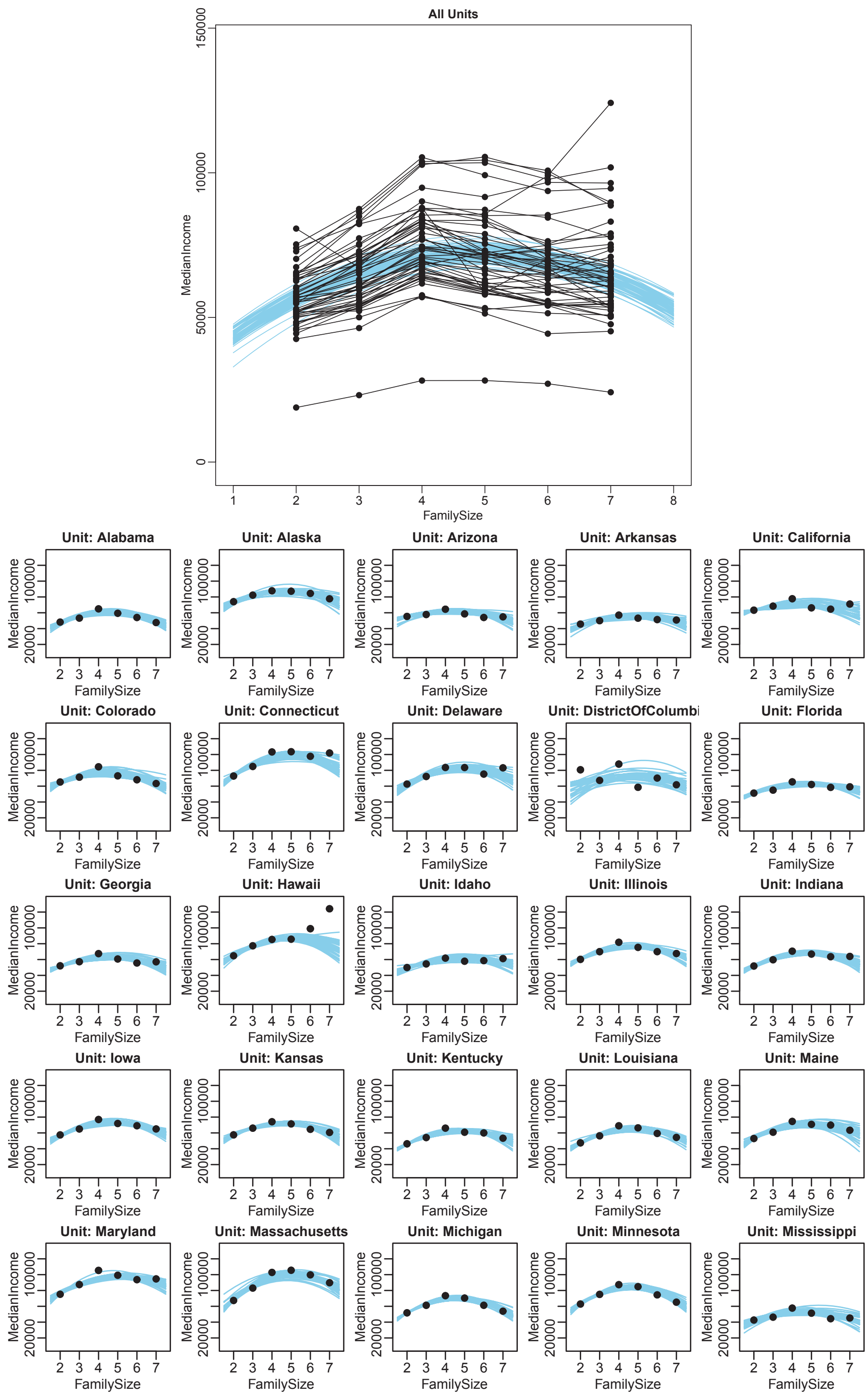

Figure 17.7: Upper panel: Median family income for various family sizes, with separate line for each of 50 states and the District of Columbia and Puerto Rico. Lower panels show data of a subset of individual states. Credible quadratic trends are superimposed. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

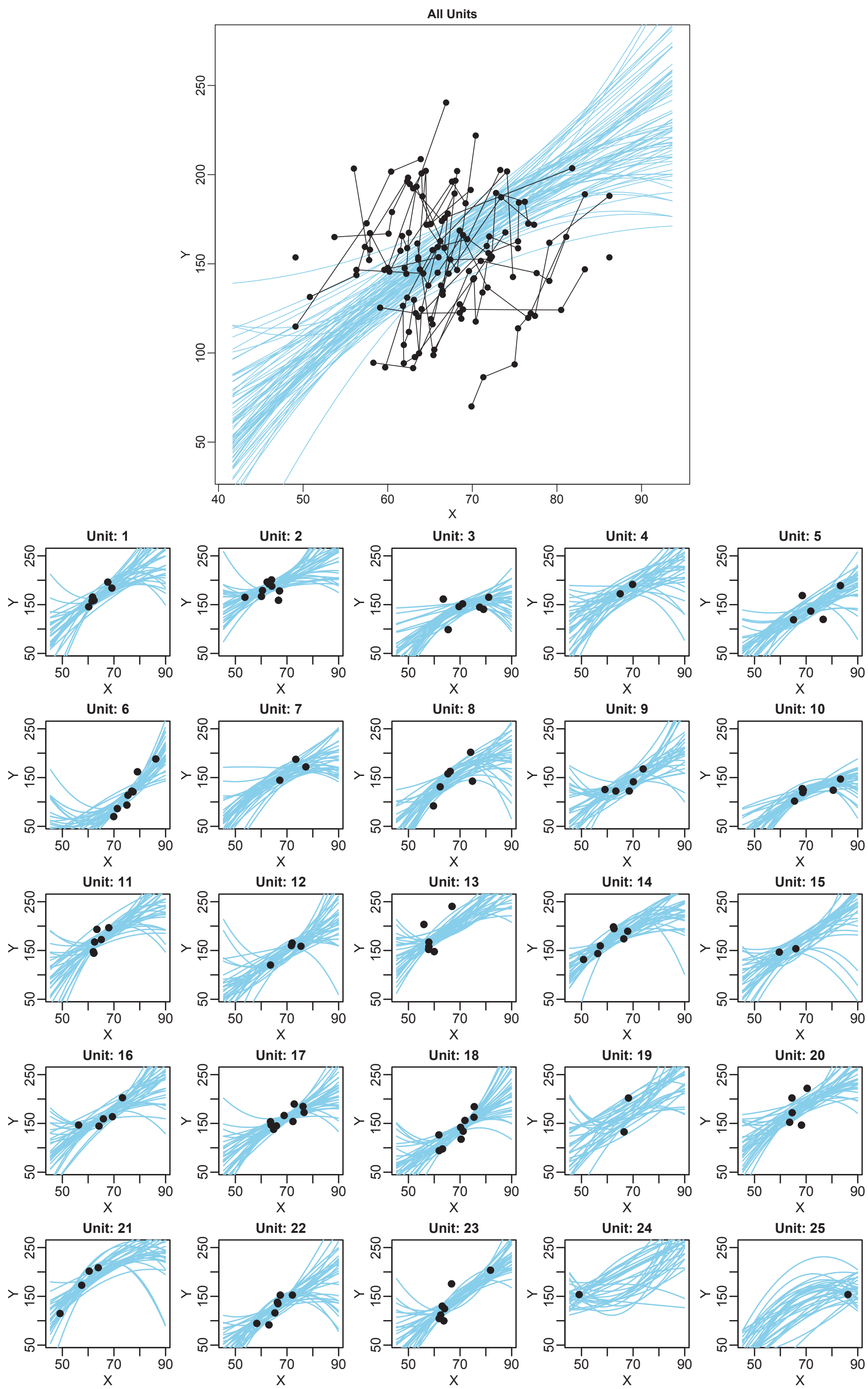

Figure 17.8: Fitting the data of Figure 17.5 with the quadratic trend model. The flexibility of the quadratic trend yields more *un*certainty in the linear trend, despite the fact that the modal estimate of the quadratic trend is nearly zero.

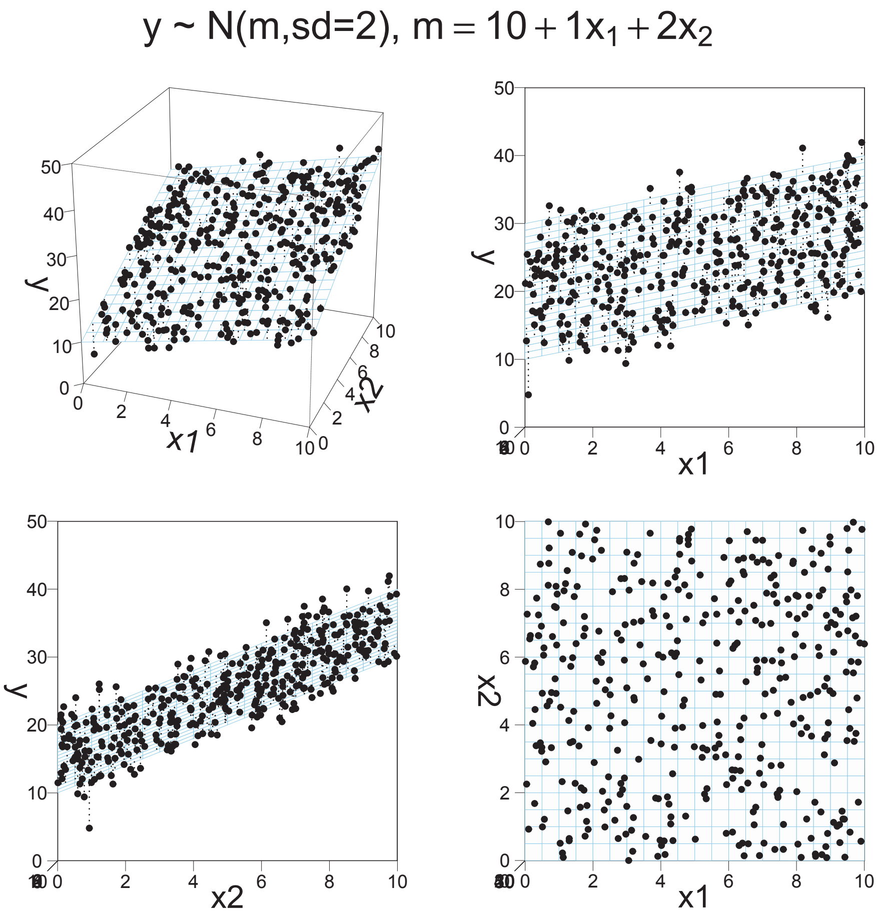

Figure 18.1: Data, y, that are normally distributed around the values in the plane. The $\langle x_1, x_2 \rangle$ values are independent of each other, as shown in the lower-right panel. The panels show different perspectives on the same plane and data. Compare with Figure 18.2. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

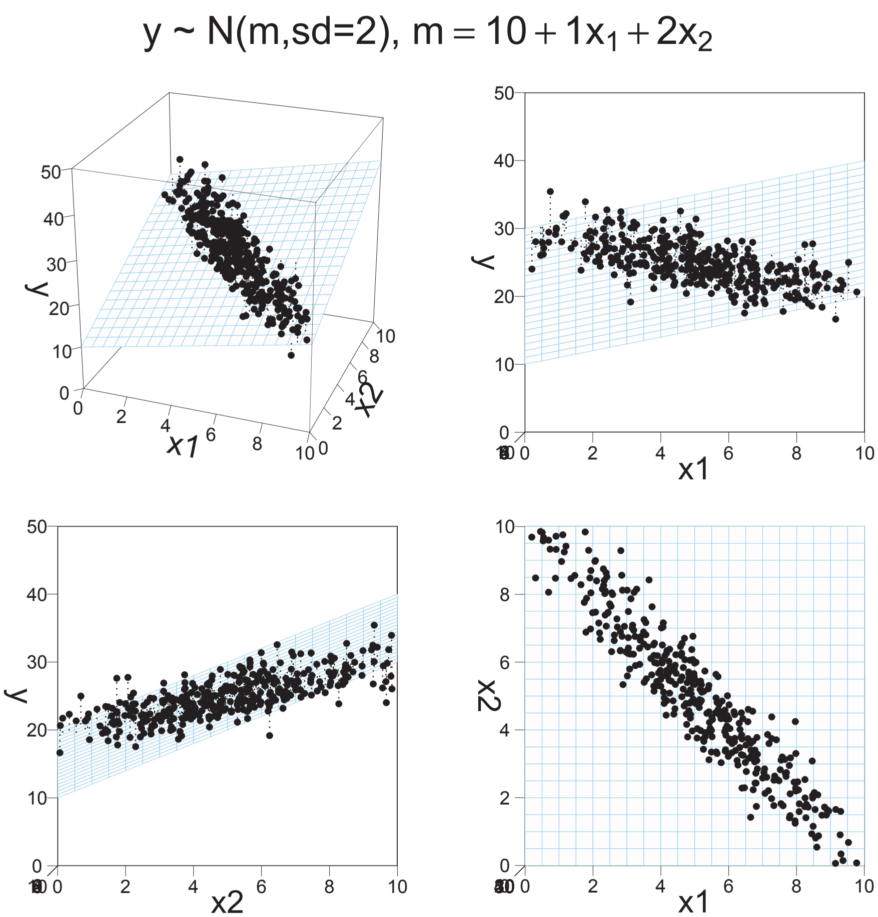

Figure 18.2: Data, y, that are normally distributed around the values in the plane. The $\langle x_1, x_2 \rangle$ values are (anti-)correlated, as shown in the lower-right panel. The panels show different perspectives on the same plane and data. Compare with Figure 18.1. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

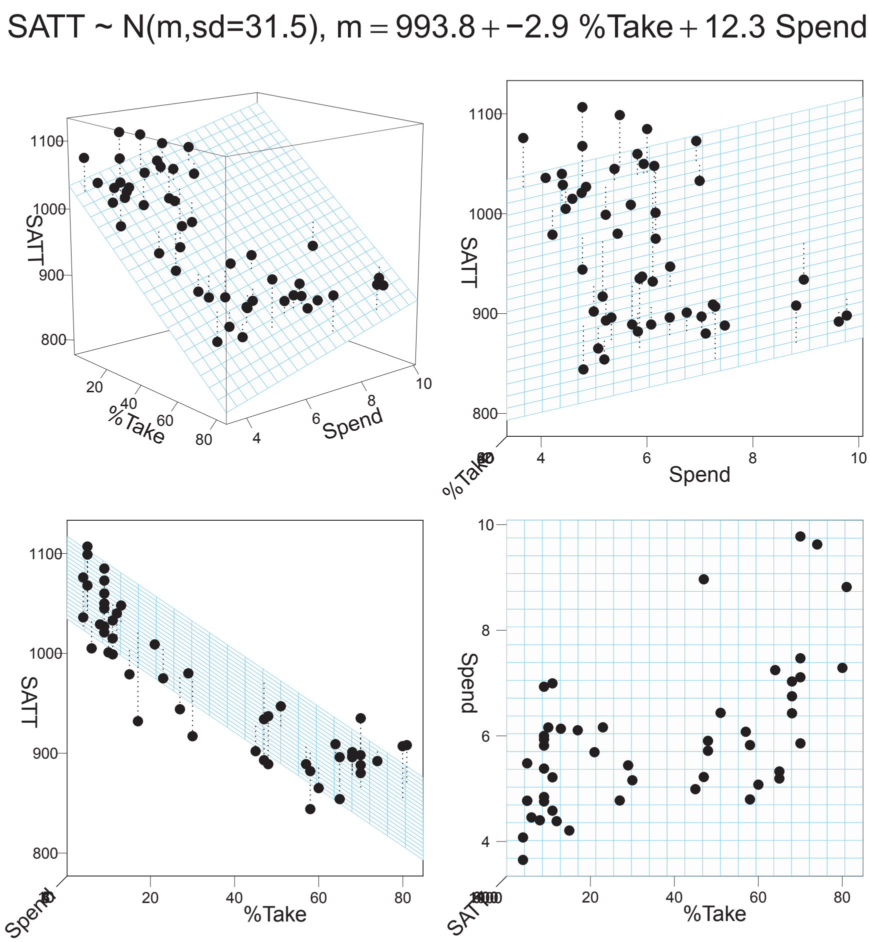

Figure 18.3: The data (Guber, 1999) are plotted as dots, and the grid shows the best fitting plane. "SATT" is the average total SAT score in a state. "%Take" is the percentage of students in the state who took the SAT. "Spend" is the spending per pupil, in thousands of dollars. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

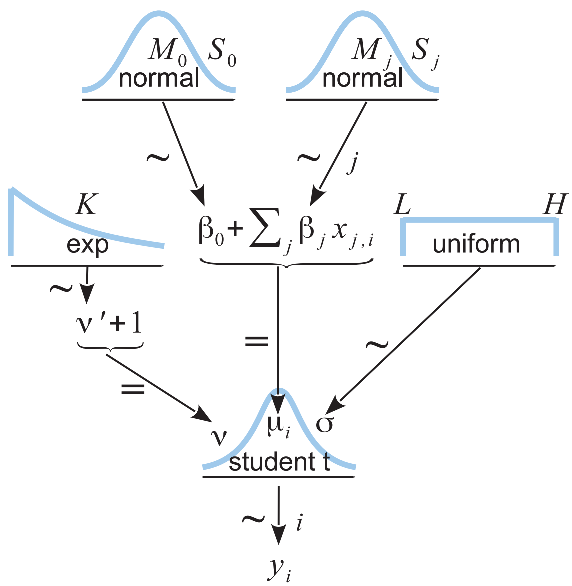

Figure 18.4: Hierarchical diagram for multiple linear regression. Compare with Figure 17.2 (p. 463). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

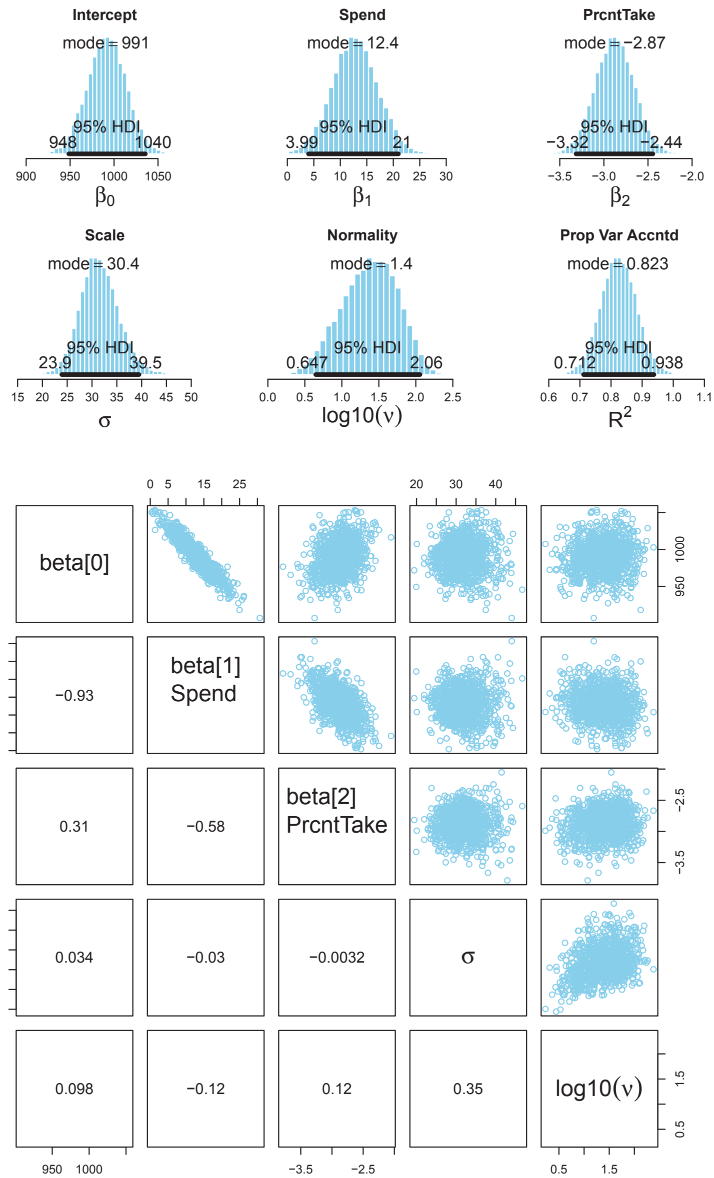

Figure 18.5: Posterior distribution for data in Figure 18.3 and model in Figure 18.4. Scatterplots reveal correlations among credible parameter values; in particular, the coefficient on Spending ("Spend") trades off with the coefficient on Percentage taking the exam ("PrcntTake"), because those predictors are correlated in the data. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

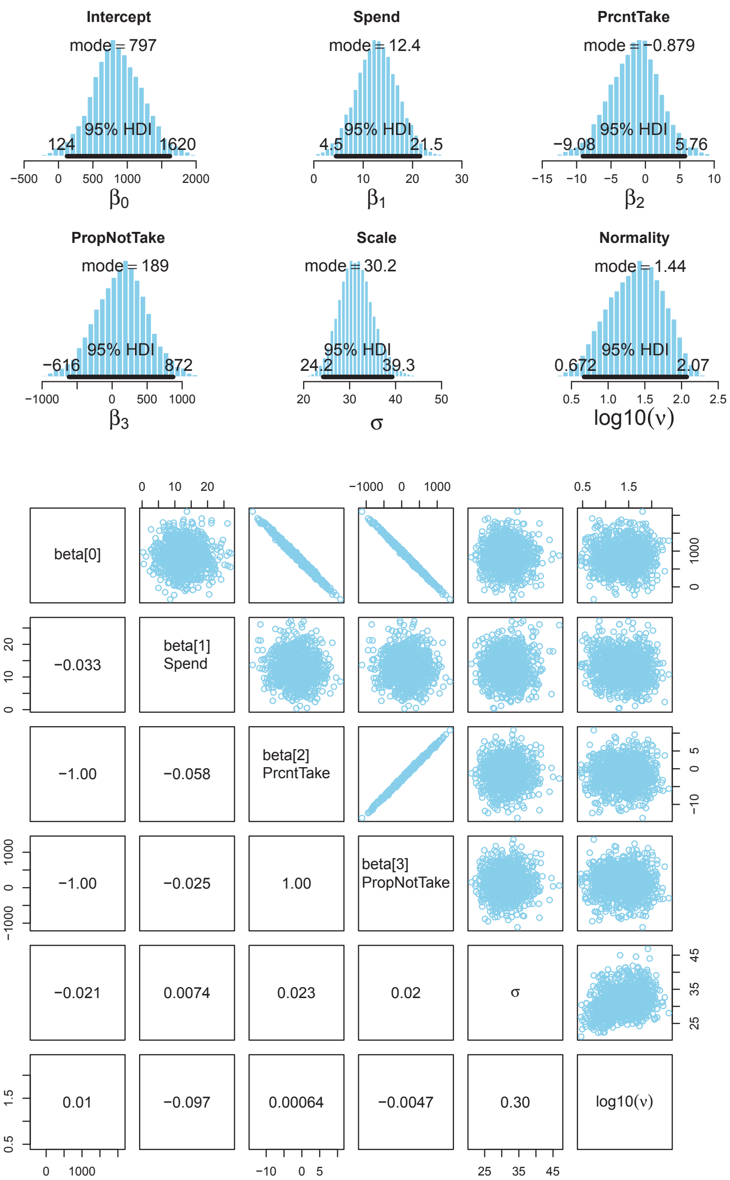

Figure 18.6: Posterior distribution for data in Figure 18.3 with a redundant predictor, the proportion of students not taking the exam. Compare with the result without a redundant predictor in Figure 18.5. Notice the perfect correlation between credible values of the regression coefficients on percentage taking the exam (PrcntTake) and proportion not taking the exam (PropNotTake). The posterior on the redundant predictors is strongly reflective of the prior distribution, which is shown in Figure 18.7. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

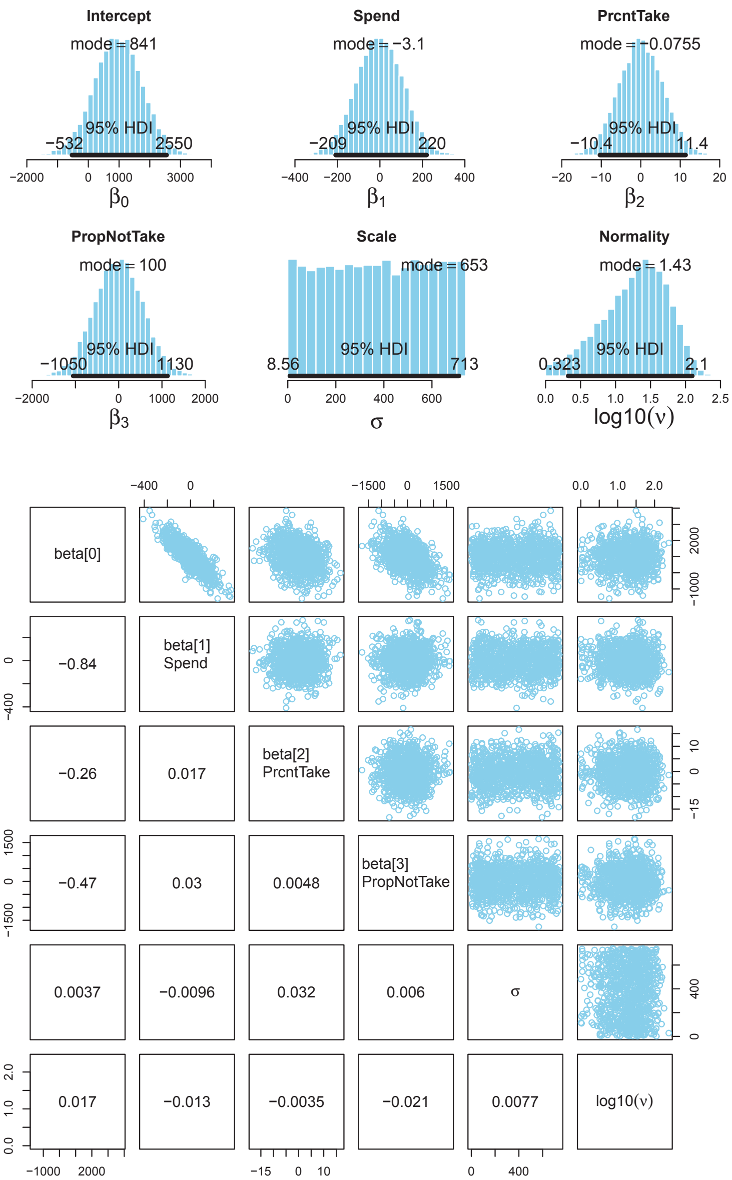

Figure 18.7: The prior distribution for the posterior distribution in Figure 18.6. Notice that the marginal posterior distributions of the redundant predictors (in Figure 18.6) is only a little narrower than the priors shown here. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

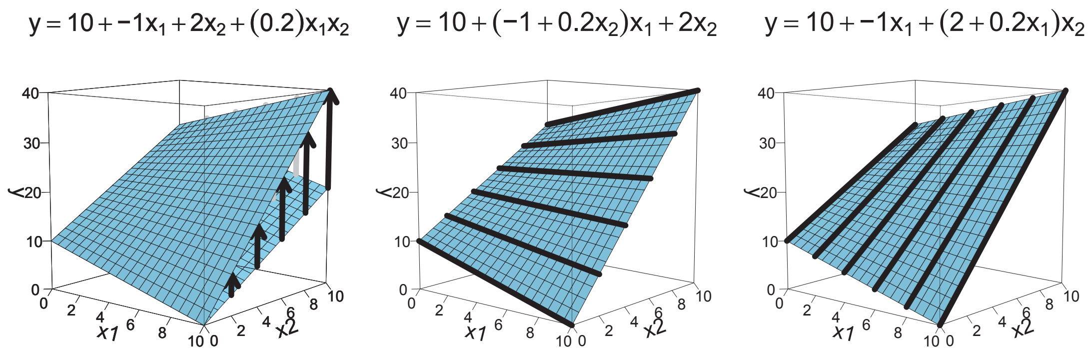

Figure 18.8: A multiplicative interaction of x_1 and x_2, parsed three ways. The left panel emphasizes that the interaction involves a multiplicative component that adds a vertical amount to the planar additive model, as indicated by the arrows that mark $\beta_{1\times2}x_1x_2$. The middle panel shows the same function, but with the terms algebraically re-grouped to emphasize that the slope in the x_1 direction depends on the value of x_2, as shown by the darkened lines that mark $\beta_1 + \beta_{1\times2}x_2$. The right panel again shows the same function, but with the terms algebraically re-grouped to emphasize that the slope in the x_2 direction depends on the value of x_1, as shown by the darkened lines that mark $\beta_2 + \beta_{1\times2}x_1$. Compare with Figure 15.3 (p. 400). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

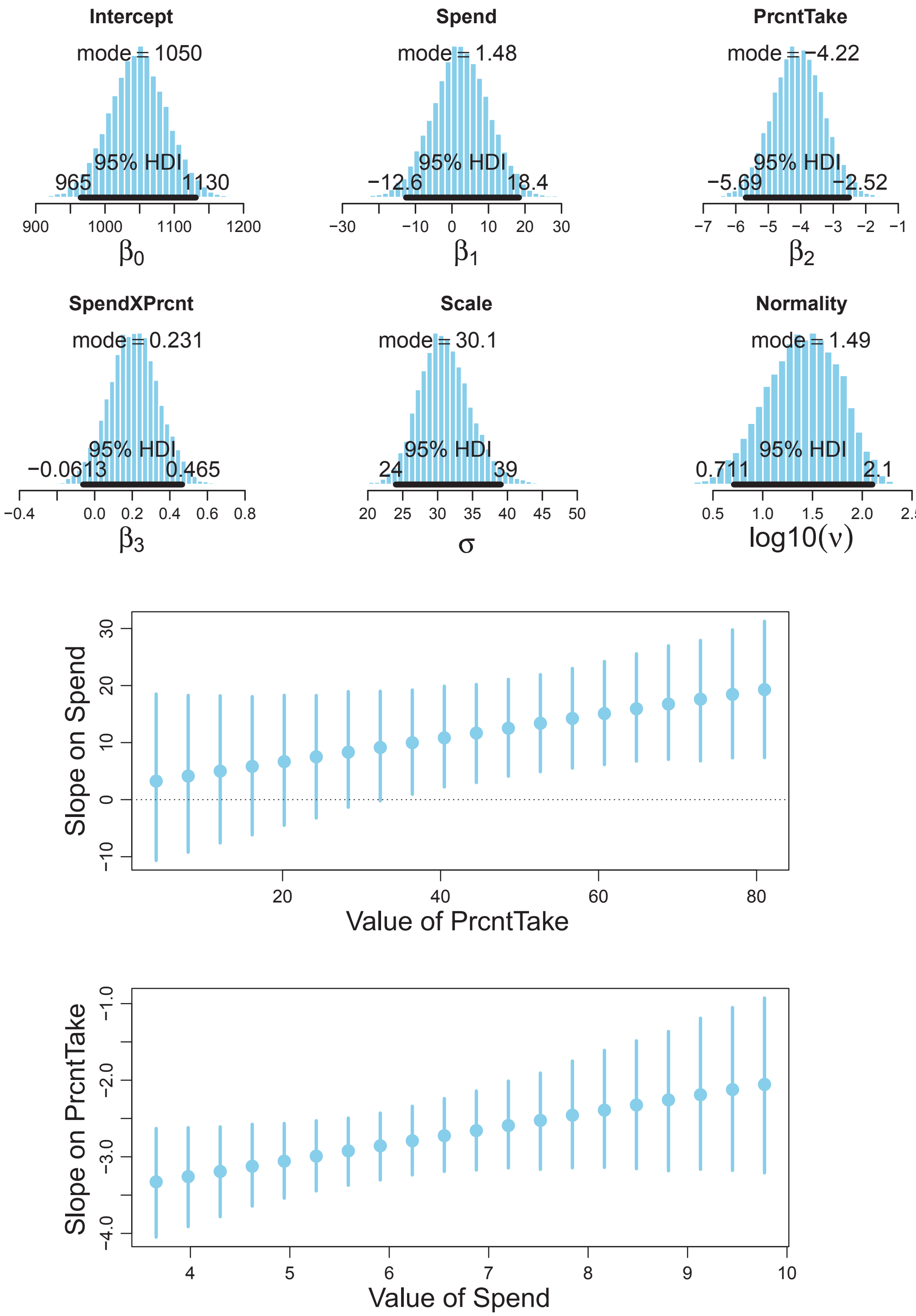

Figure 18.9: Posterior distribution when including a multiplicative interaction of Spend and PrcntTake. The marginal distribution of β_1 is the slope on Spend when PrcntTake=0, and the marginal distribution of β_2 is slope on PrcntTake when Spend=0. Lower panels show 95% HDIs and median values of slopes for other values of predictors. Slope on Spend is $\beta_1 + \beta_3 \cdot$ PrcntTake and slope on PrcntTake is $\beta_2 + \beta_3 \cdot$ Spend. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

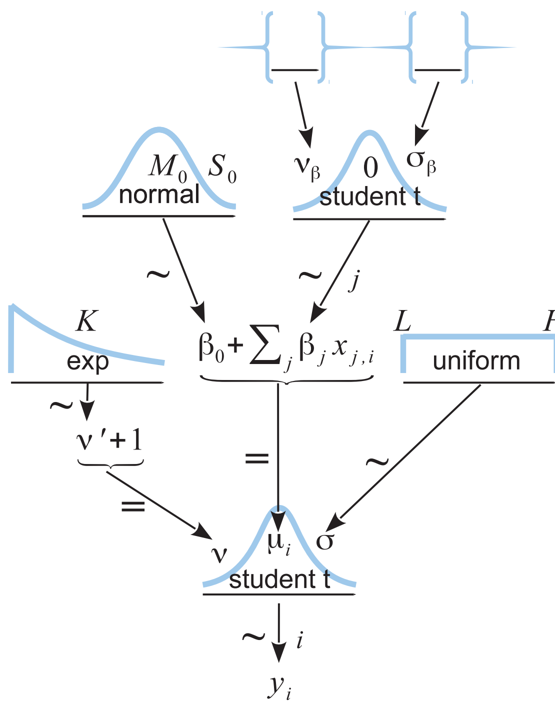

Figure 18.10: Hierarchical diagram for multiple linear regression, with a shrinkage prior across the slope coefficients. Compare with Figure 18.4 (p. 498). The empty braces at the top of the diagram indicate aspects that are optional. Typically the normality parameter ν_β is fixed at a small value, but could be estimated instead. The scale parameter σ_β could be fixed at a small value but could be estimated, in which case the standard deviation across regression coefficients is mutually informed by all the predictors.

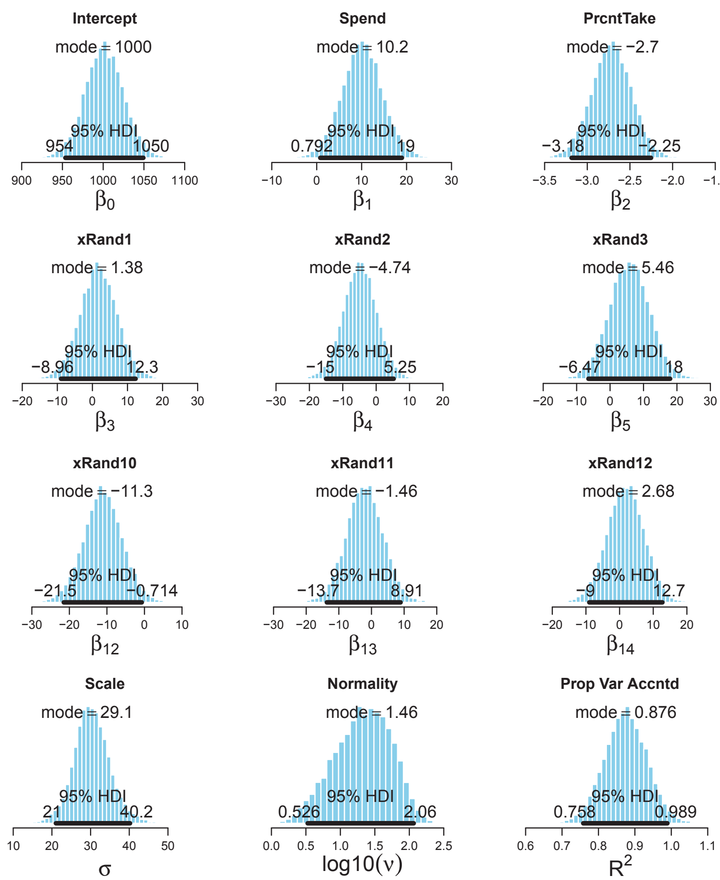

Figure 18.11: Posterior with*out* hierarchical shrinkage, using prior of Figure 18.4. Compare with results when using shrinkage prior in Figure ~~18.11~~ 18.12, especially the coefficients on Spend and xRand10. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

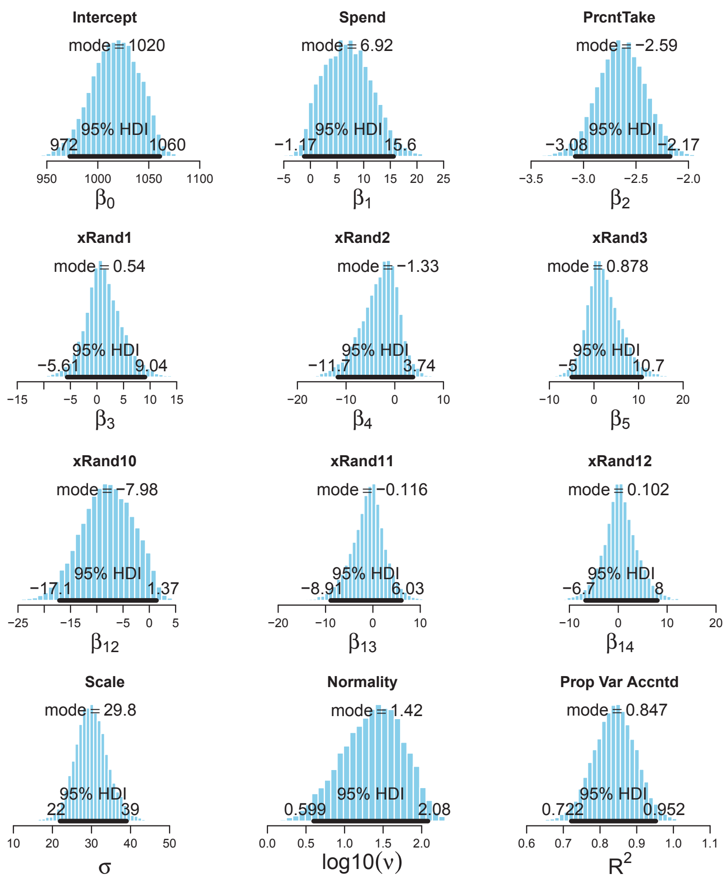

Figure 18.12: Posterior *with* hierarchical shrinkage, using hierarchical prior of Figure 18.10 with a gamma distribution (mode=1.0, sd=1.0) on standardized σ_β and $\nu_\beta = 1$. Compare with the results when not using hierarchical shrinkage in Figure 18.11, especially the coefficients on Spend and xRand10. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

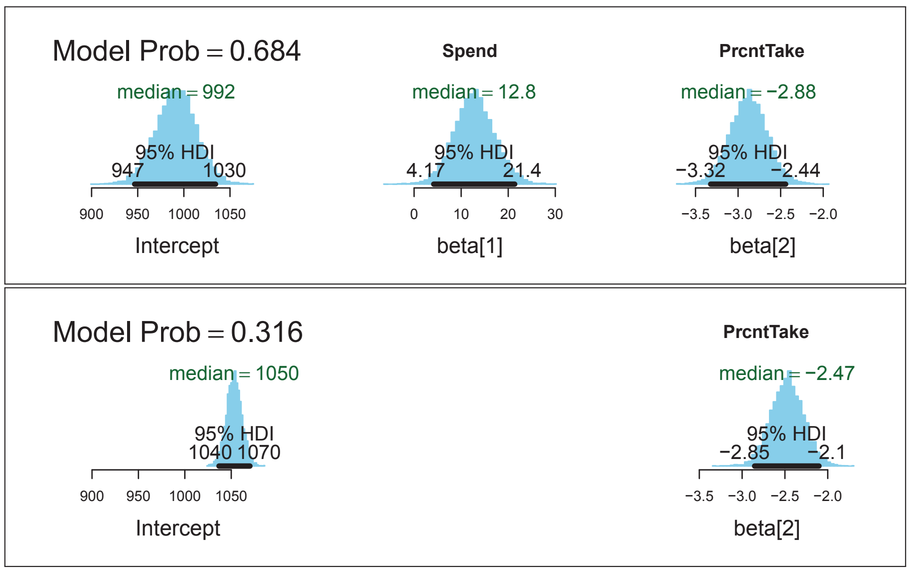

Figure 18.13: Posterior probabilities of different subsets of predictors along with the marginal posterior distributions of the included regression coefficients. The two other possible models, involving only Spend or only the intercept, had essentially zero probability. The prior probability of each model was $0.5^2 = 0.25$. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

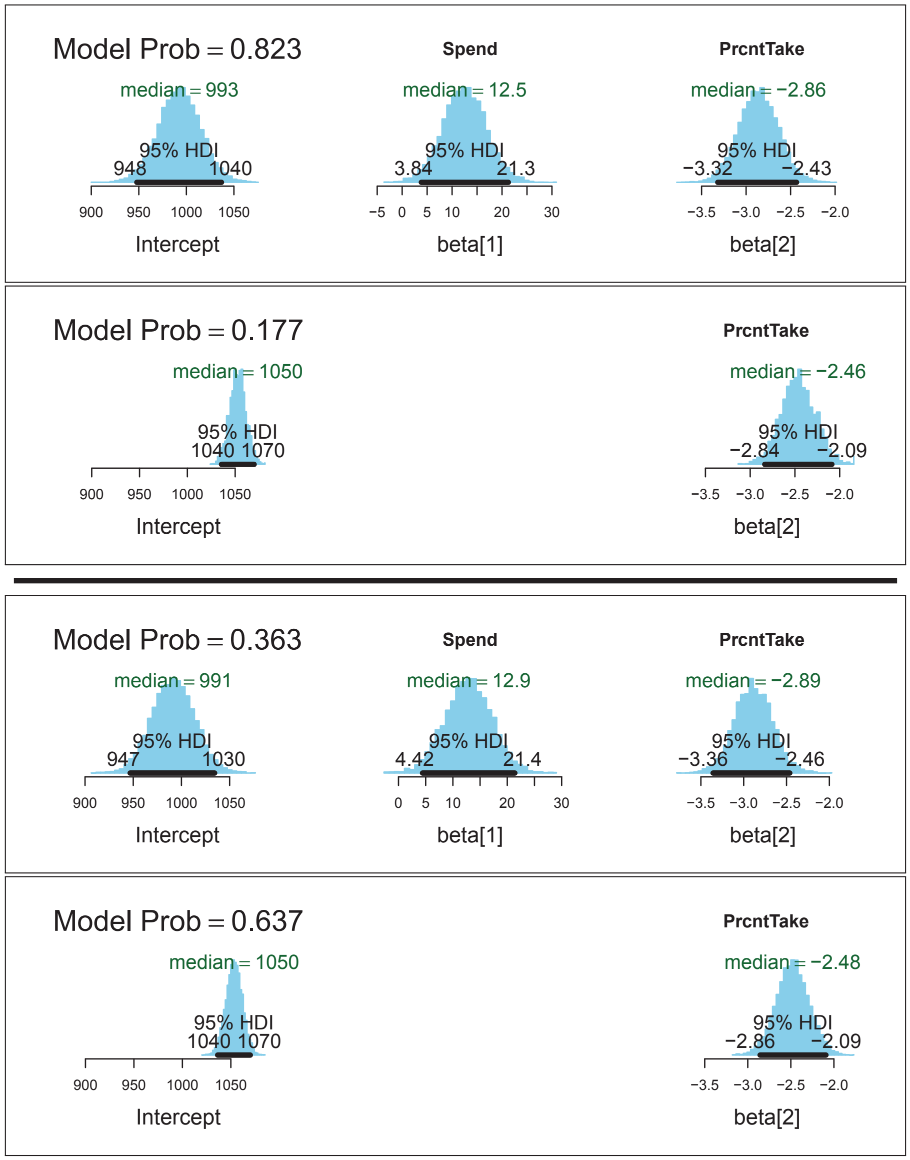

Figure 18.14: Posterior probabilities of different subsets of predictors along with the marginal posterior distributions of the included regression coefficients. Upper two panels show results when the prior on the standardized regression coefficients has SD=1; lower two panels are for SD=10. In both cases, the prior probability of each model was $0.5^2 = 0.25$.

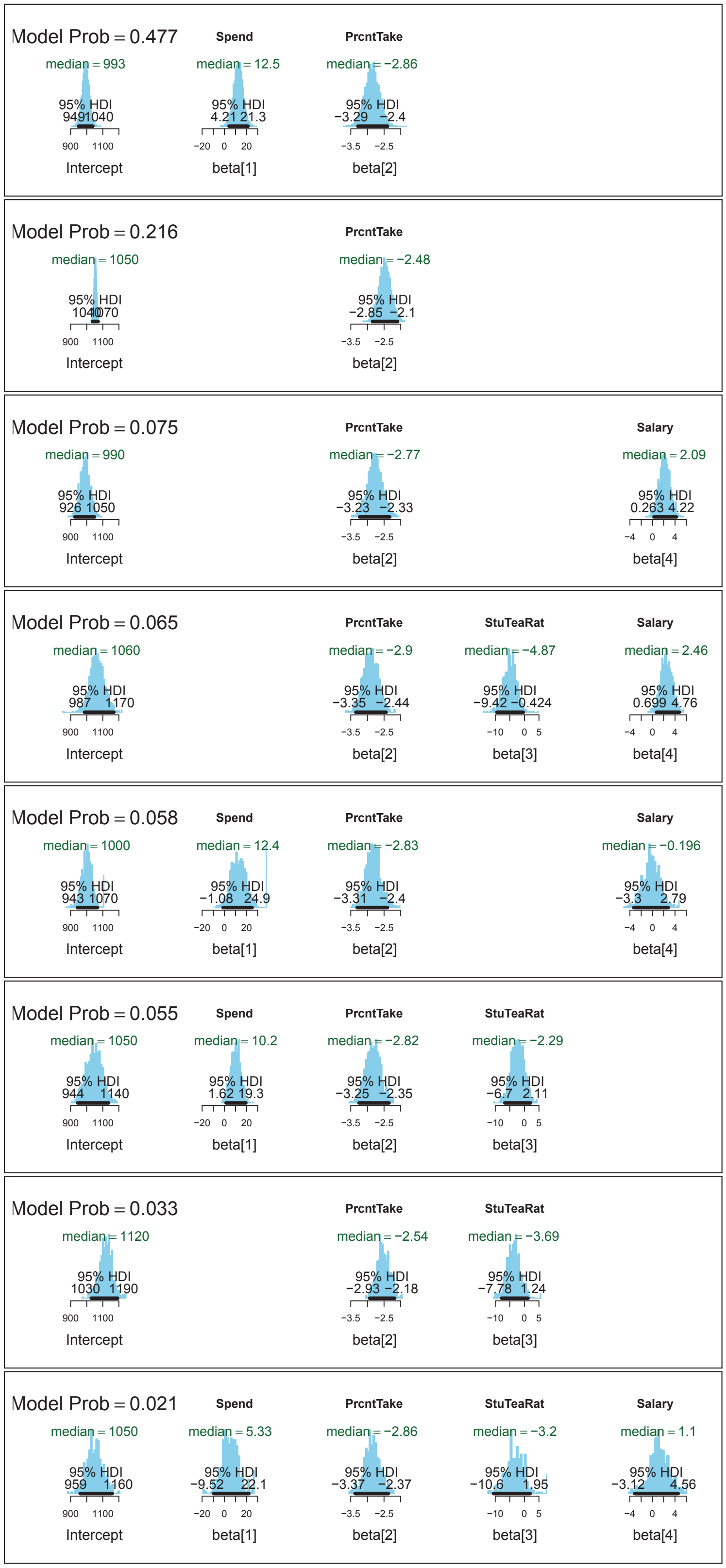

Figure 18.15: Posterior probabilities of different subsets of predictors along with the marginal posterior distributions of the included regression coefficients. The remaining eight possible models had essentially zero probability. The prior probability of each model was $0.5^4 = 0.0625$. The histograms of improbable models are jagged because the MCMC chain visited those models only rarely.

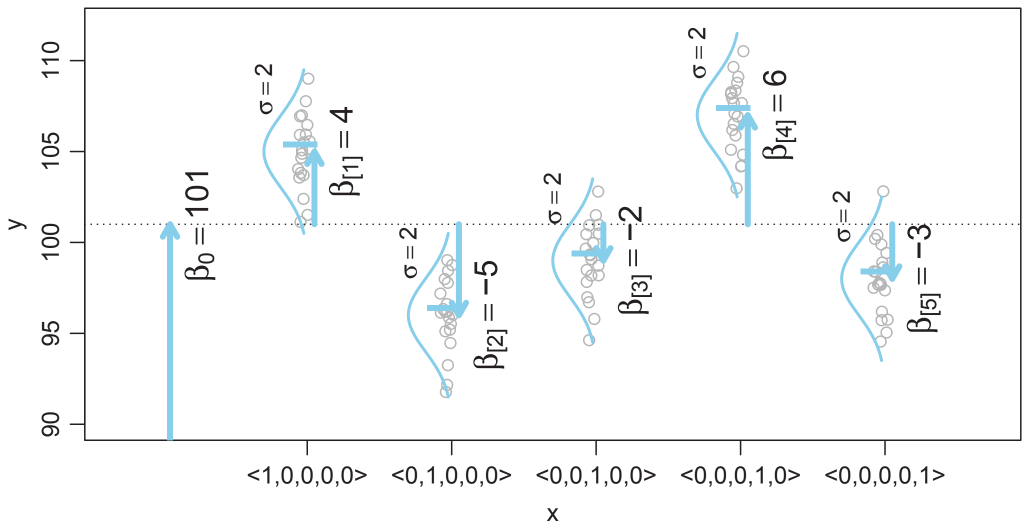

Figure 19.1: Description of data as normally distributed around group means that are conceptualized as deflections from an overall baseline. Data are indicated by circular dots (jittered left-right for visibility). The standard deviation of the data within groups is assumed to be the same for all groups and is indicated as σ. Baseline and deflections are indicated by arrows and β values. Notice that the deflections from baseline sum to zero. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

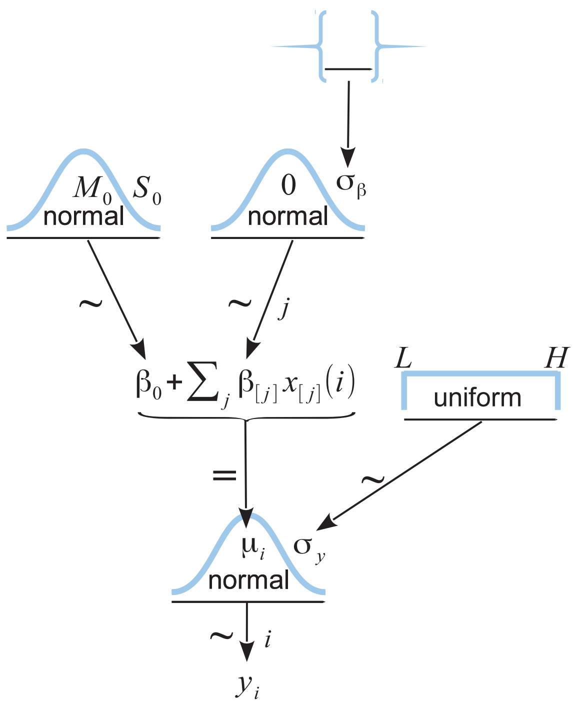

Figure 19.2: Hierarchical diagram for model that describes data from several groups of a single factor. At the top of the diagram, the empty braces indicate the prior distribution on the between-group standard deviation, σ_β, which could be a folded-t as recommended by Gelman (2006), a gamma distribution with non-zero mode, or a constant if no sharing across groups is desired. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

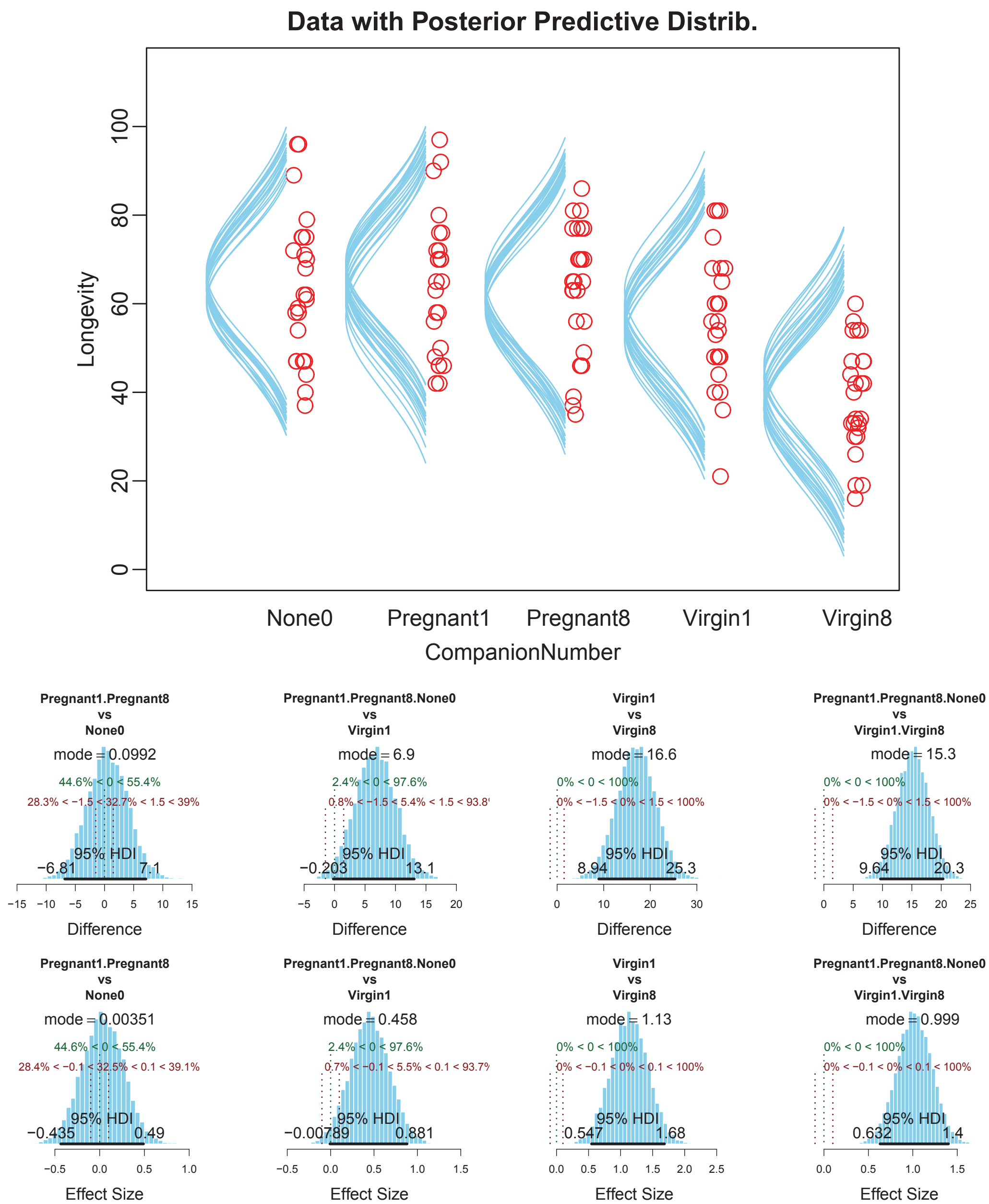

Figure 19.3: Data and posterior distribution for fruitfly longevity. Model assumes normal distributions with homogeneous variances. (Data are plotted with random left-right jitter for visibility.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

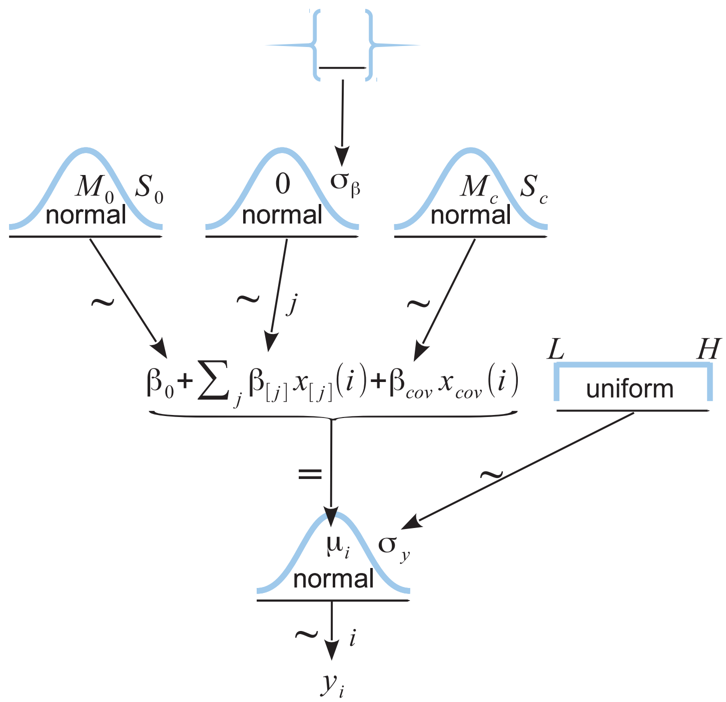

Figure 19.4: Hierarchical diagram for model that describes data from several groups of a single factor, along with a single metric covariate. Compare with Figure 19.2.

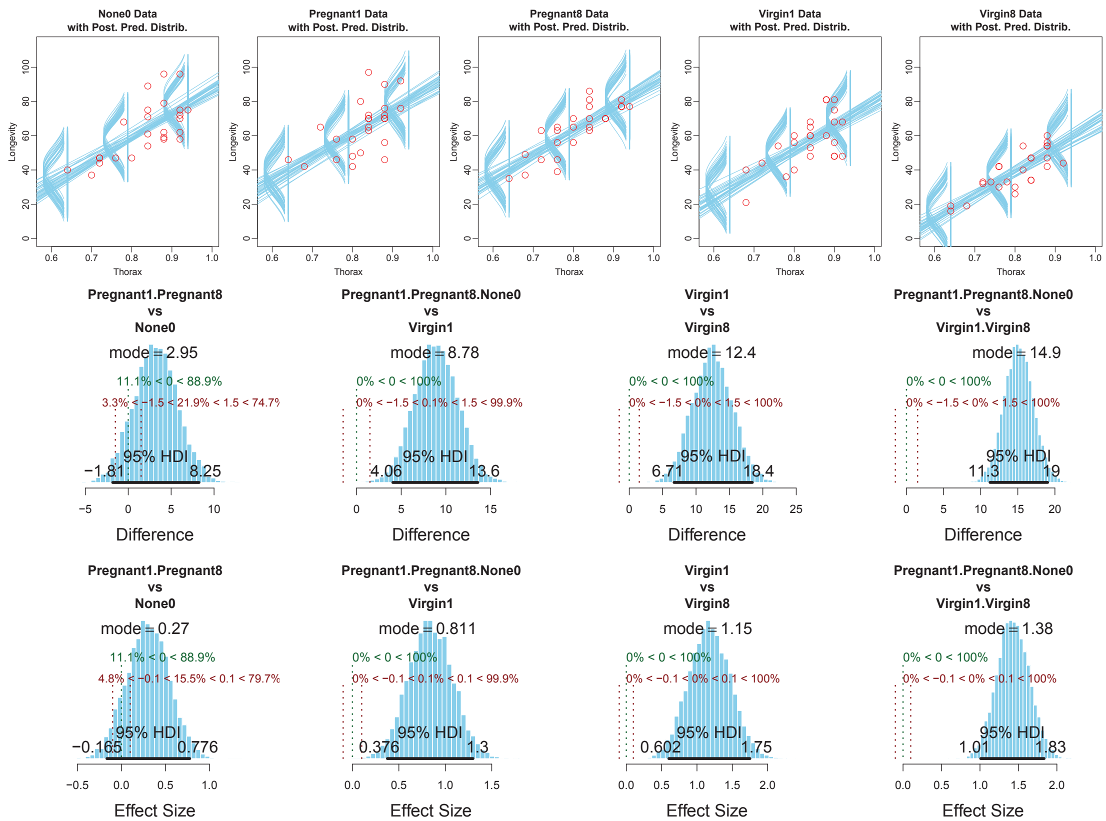

Figure 19.5: Posterior distribution for fruitfly longevity, described by normal distributions with homogeneous variances and a linear function of a covariate. Upper row shows that the within-group variance is smaller than in Figure 19.3. Lower rows show that contrasts are more precise than in Figure 19.3. In particular, here the contrast of Pregnant1 and Pregnant8 and None0 vs Virgin1 is clearly non-zero. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

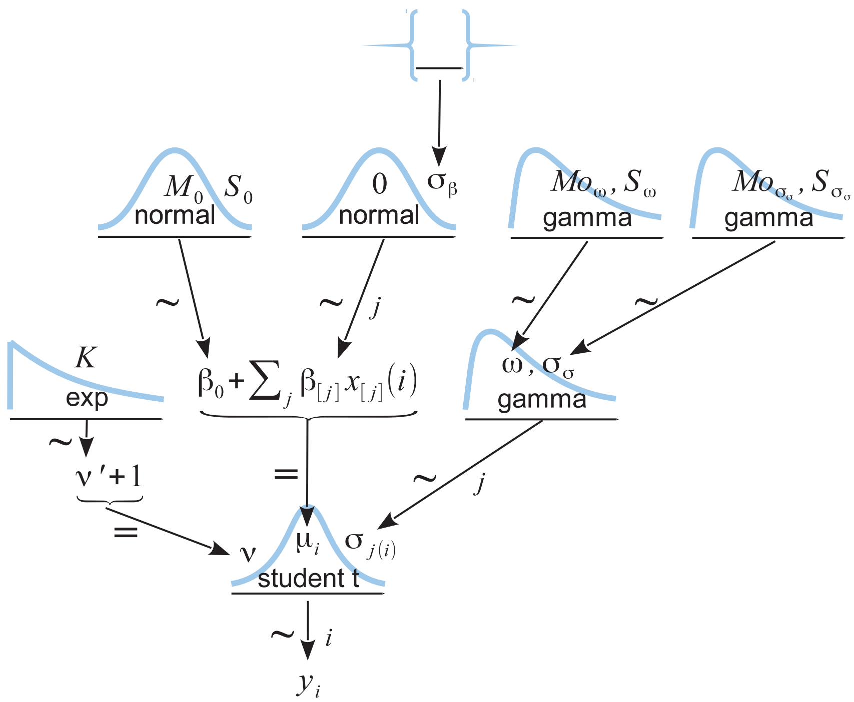

Figure 19.6: Hierarchical diagram for model that describes data from several groups of a single factor, using a heavy-tailed noise distribution and different standard deviations for each group. Compare with Figure 19.2. (The gamma distributions are parameterized by mode and standard deviation.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

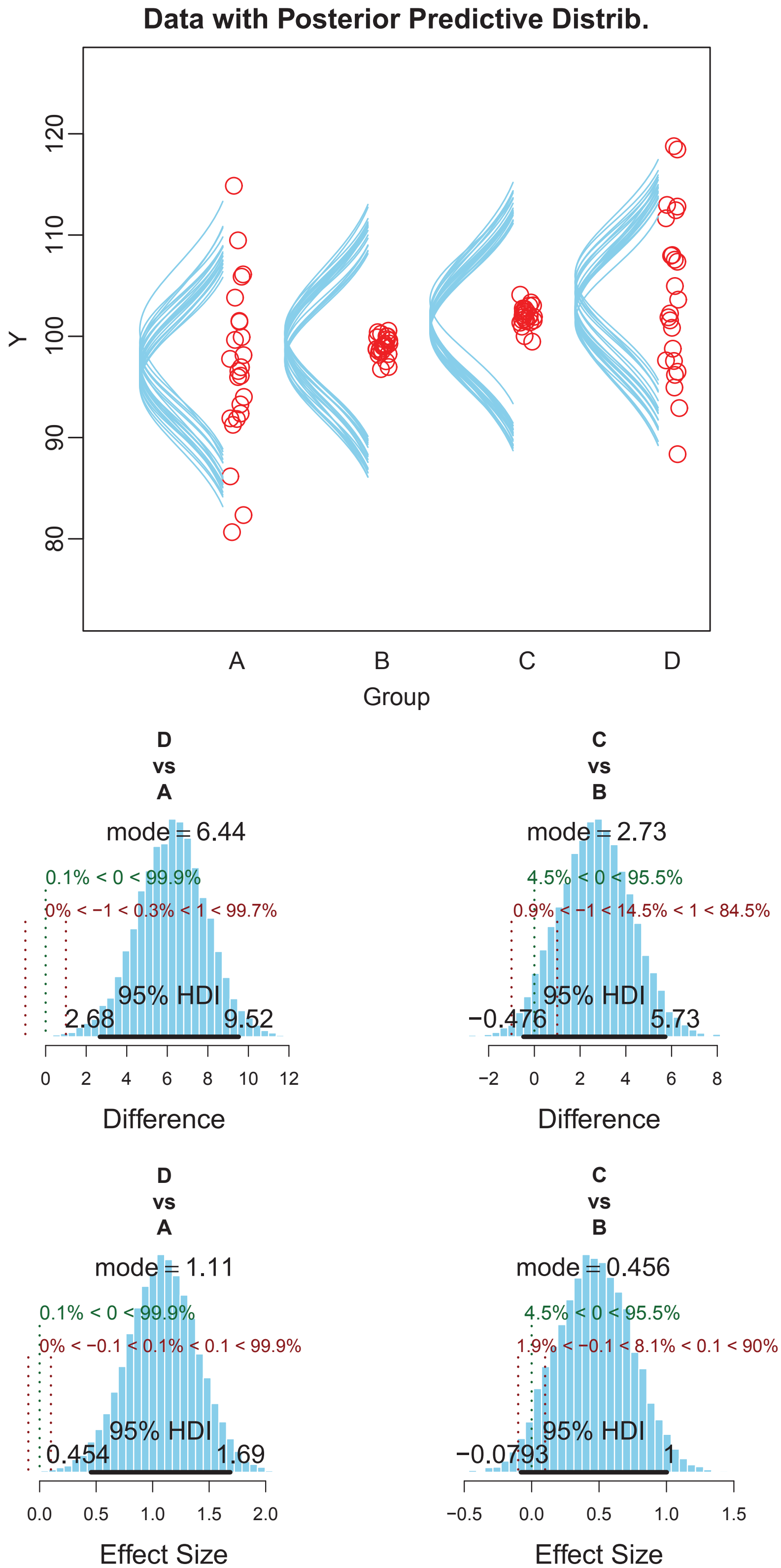

Figure 19.7: Fictitious data to illustrate groups with different variances. Here, the model assumes equal variances across groups. Compare with Figure 19.8. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

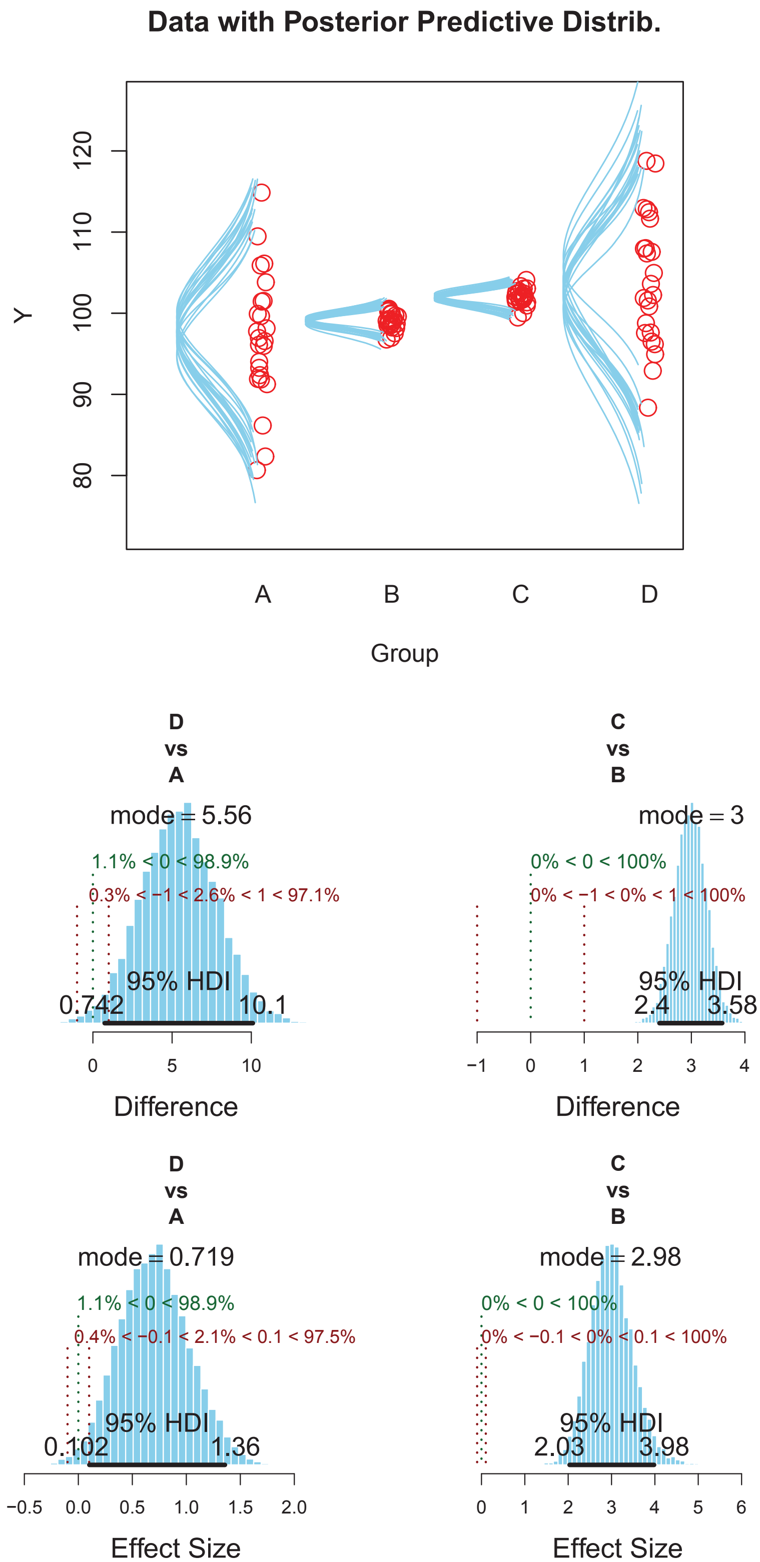

Figure 19.8: Fictitious data to illustrate groups with different variances. Here, the model assumes different variances across groups. Compare with Figure 19.7. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

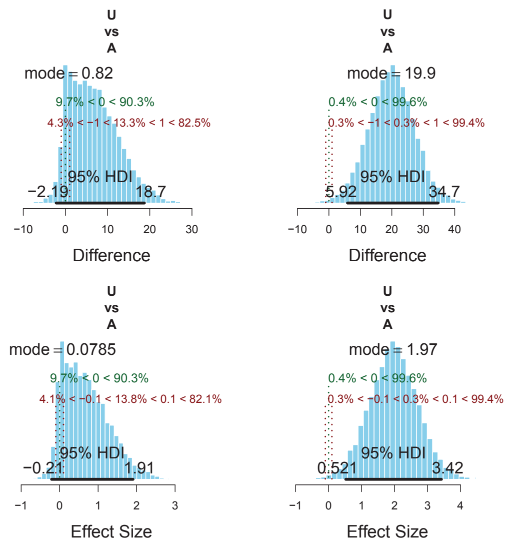

Figure 19.9: For Exercise 19.1. Left: Implosive shrinkage resulting from estimated σ_β. Right: Results for fixed σ_β. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

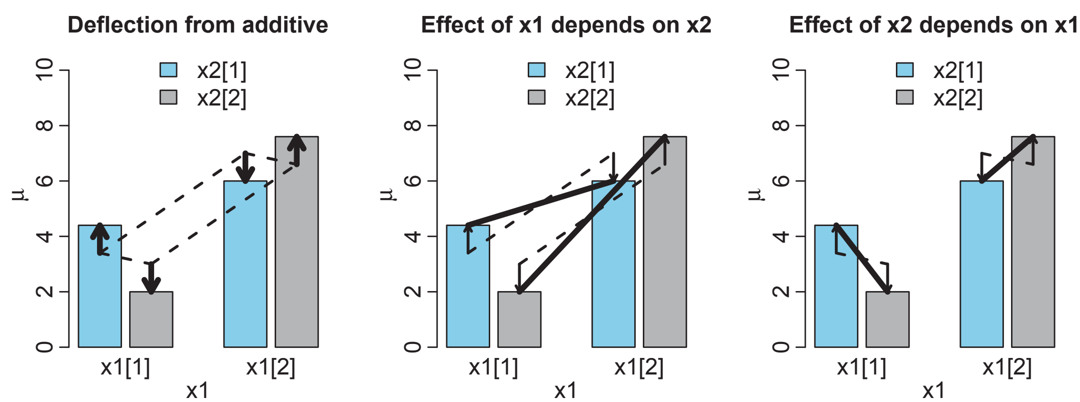

Figure 20.1: An example to illustrate the notion of interaction. Each panel plots the same four means but with different superimposed lines for different emphases expressed in the title of each panel. The dashed lines indicate the average (i.e., main) effects of the factors. Subscripts are elevated to regular size for readability; for example $x_{2[1]}$ is displayed as $x2[1]$. (Compare with Figure 15.5, p. 402.)

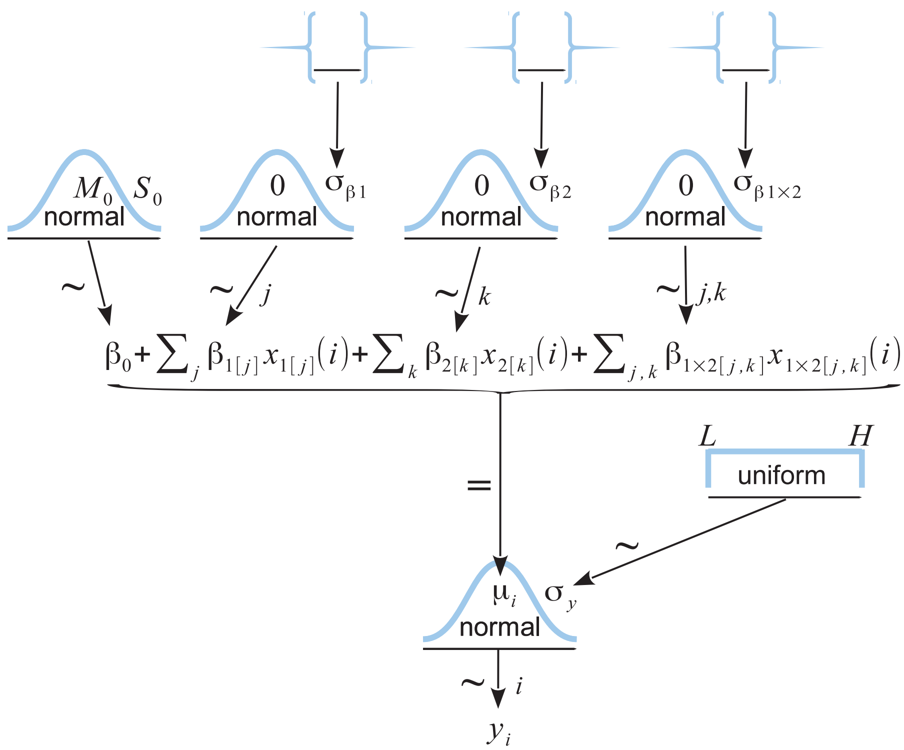

Figure 20.2: Hierarchical diagram for model that describes data from two nominal predictors. At the top of the diagram, the empty braces indicate the prior distribution on the standard deviations of the deflections, which could be a folded-*t* as recommended by Gelman (2006), a gamma distribution with non-zero mode, or a constant if no sharing across levels is desired. Compare with Figure 19.2 (p. 529). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

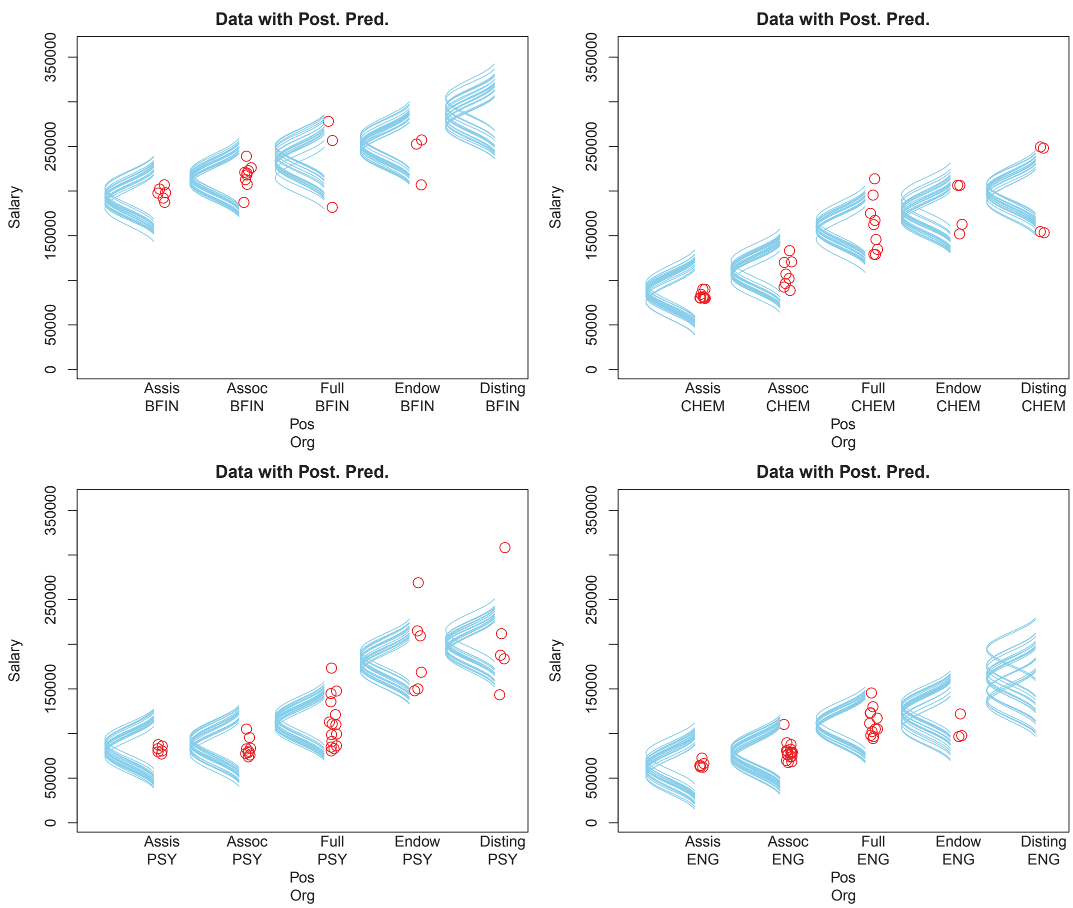

Figure 20.3: Salary data for four departments and five seniorities. The full set of data included 60 departments. Posterior predictive distributions are from a model that assumes homogeneous variances and normally distributed data within cells. (BFIN = business finance, PSY = psychology, CHEM = chemistry, ENG = english. Pos = position or rank. Org = organization or department.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

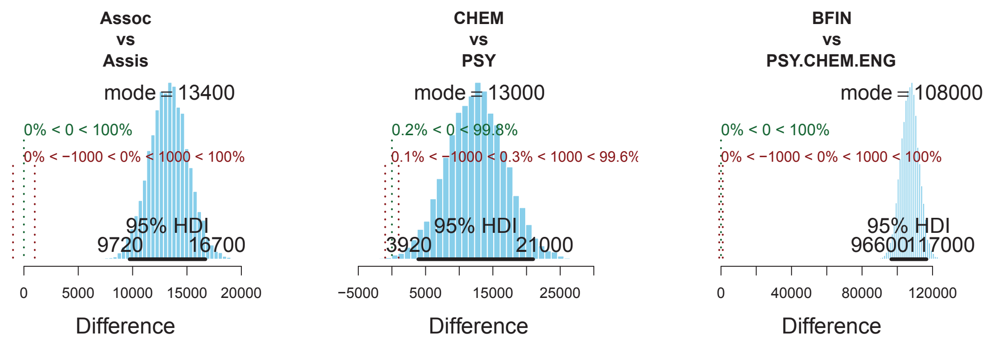

Figure 20.4: Three main effect contrasts. The left panel shows a contrast of two ranks. The right panel shows a “complex” comparison of business finance against the average of three other departments. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

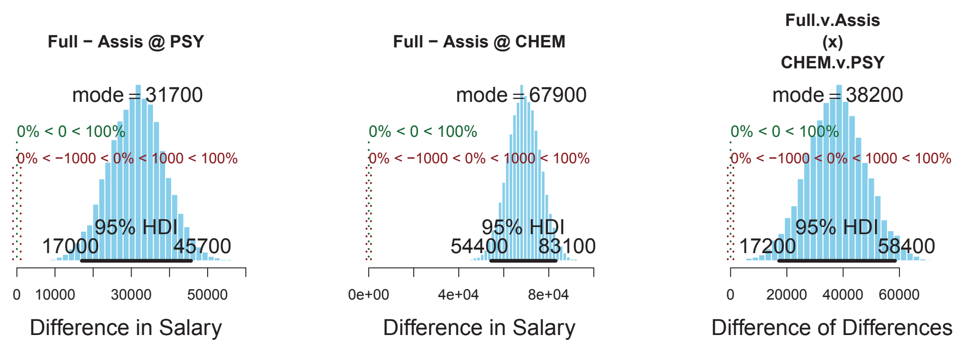

Figure 20.5: The left and middle panel show two "simple" comparisons: Each is a contrast of ranks within a level of department. The right panel shows an interaction contrast, namely, the difference of differences in the simple comparisons.

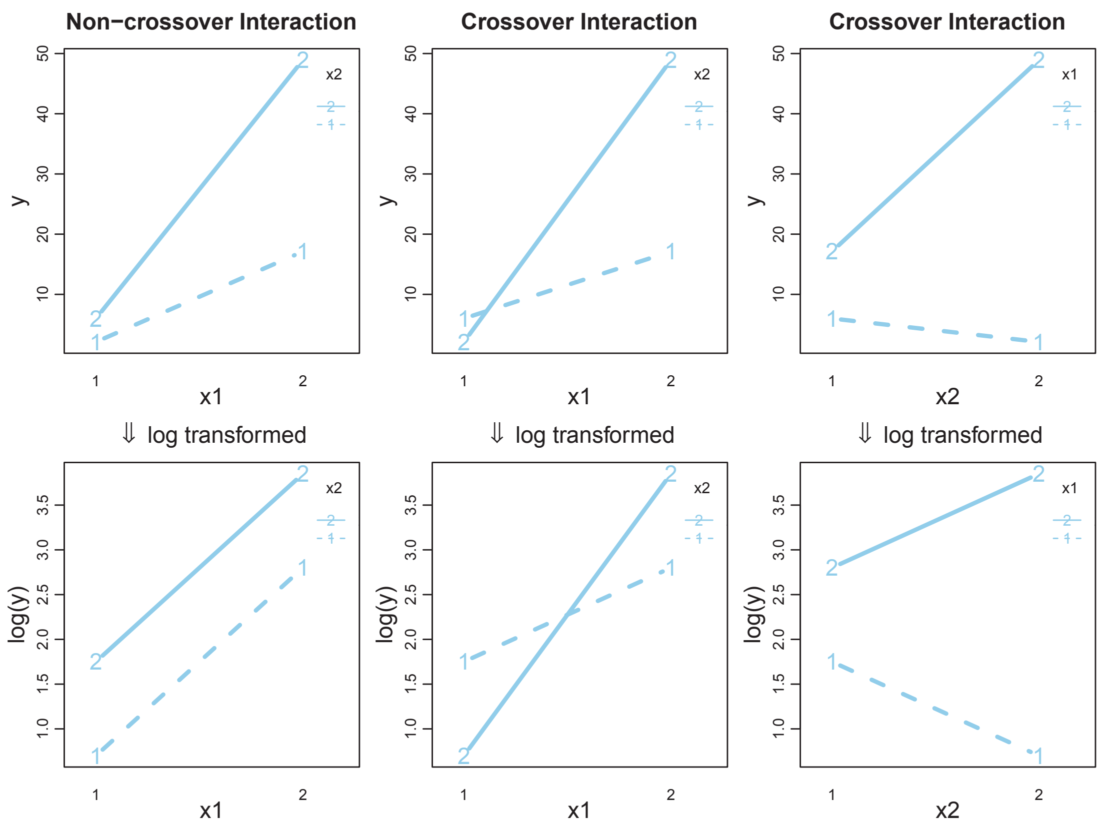

Figure 20.6: Top row shows means of original data; bottom row shows means of logarithmically transformed data. Left column shows a non-crossover interaction. Middle and right columns show a crossover interaction, the same in both columns, but plotted against x1 or x2 on the abscissa. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

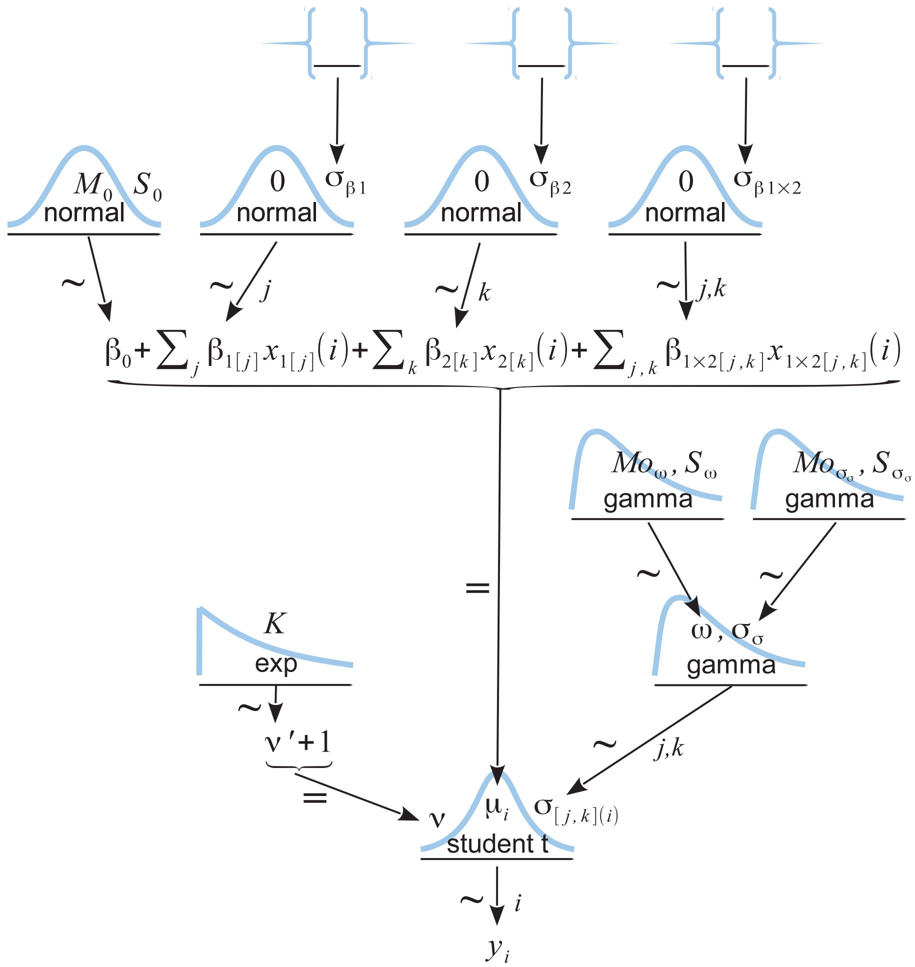

Figure 20.7: Hierarchical diagram for a model that describes data from two nominal predictors, wherein the noise distribution (at bottom of diagram) is robust to outliers and has a different standard deviation parameter, $\sigma_{[j,k]}$, for every cell. Compare with Figure 20.2 (p. 562), which assumes normally distributed noise and homogeneous variances across cells. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

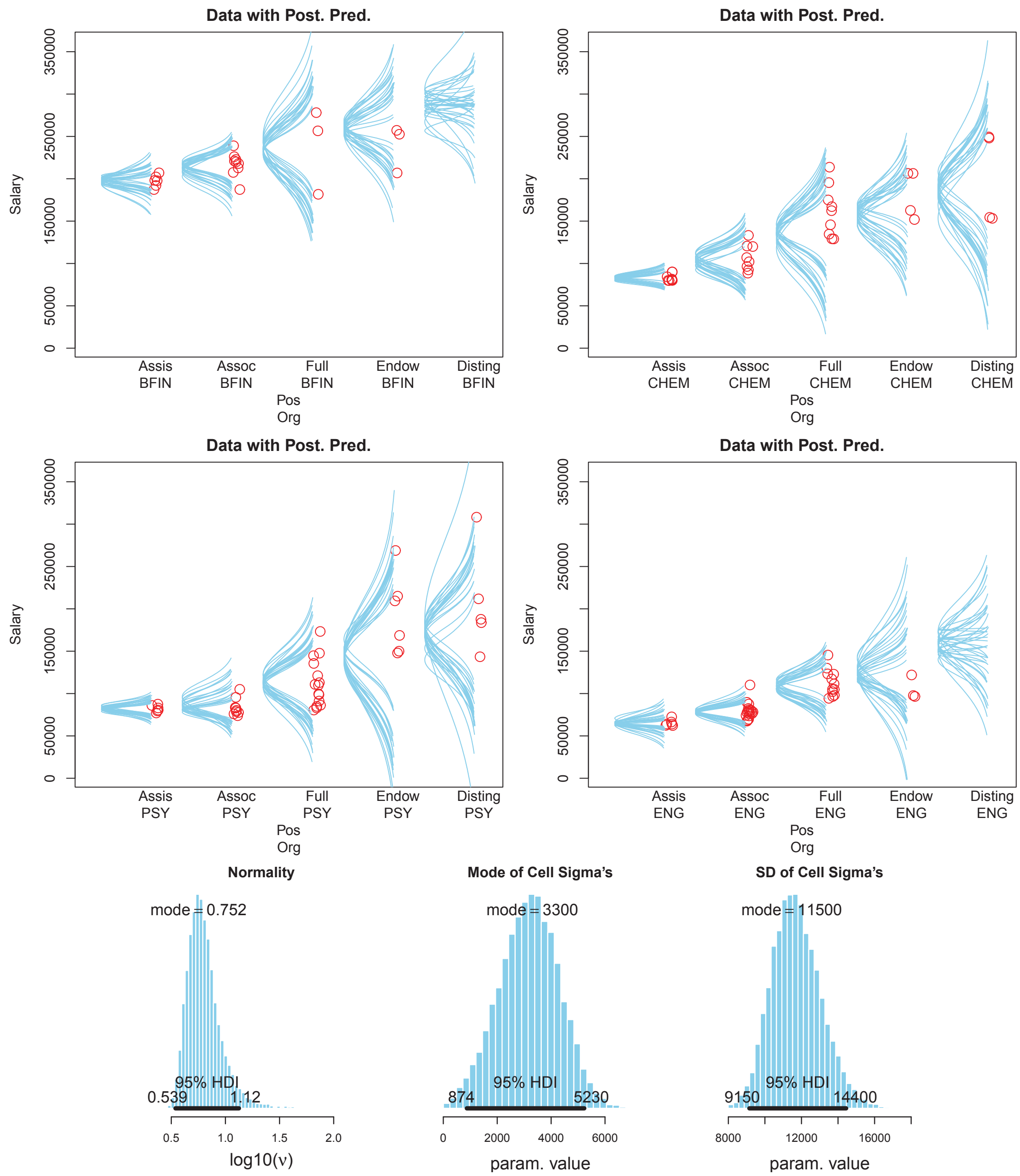

Figure 20.8: Salary data for four (of 60) departments and five seniorities. The model assumes heterogeneous variances and t-distributed data within cells. The bottom row shows marginal posterior distribution of the normality parameter, the modal cell standard deviation (ω in Figure 20.7), and the standard deviation of the estimated cell standard deviations (σ_σ in Figure 20.7). Compare with the results from a model that assumes homogeneous variances and normally distributed data within cells, shown in Figure 20.3. (BFIN = business finance, PSY = psychology, CHEM = chemistry, ENG = english. Pos = position or rank. Org = organization or department.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

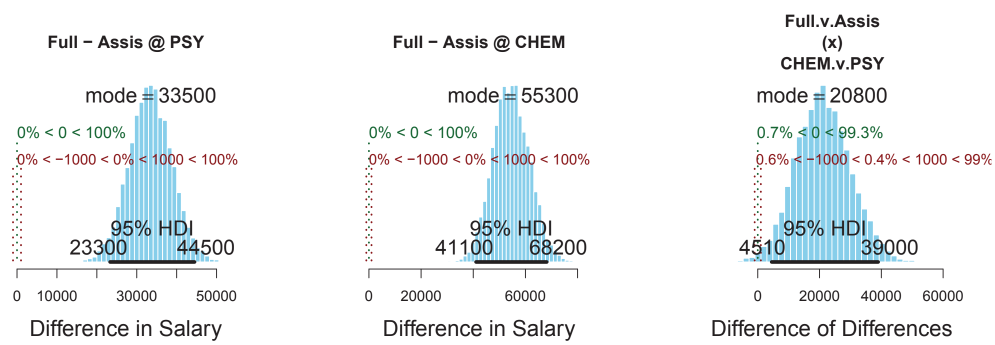

Figure 20.9: The left and middle panel show two "simple" comparisons: Each is a contrast of ranks within a level of department. The right panel shows an interaction contrast, namely, the difference of differences in the simple comparisons. The model assumes heterogeneous variances and t-distributed data within cells. Compare with Figure 20.5 (p. 565). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

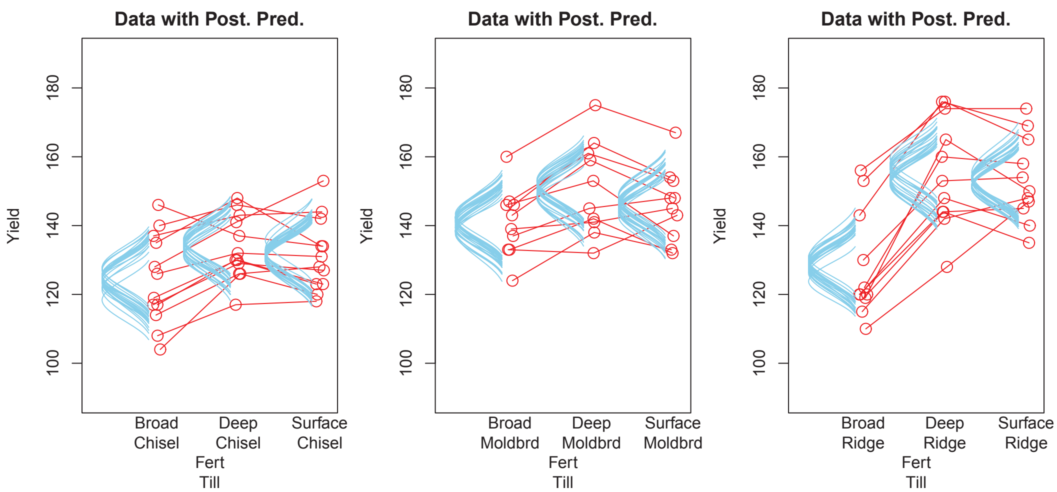

Figure 20.10: Corn production yields (bushels per acre) for different tilling (Till) and phosphorus fertilization (Fert) methods. Panels show different tilling methods (Chisel = chisel plow, Moldbrd = moldboard plow, Ridge = ridge tilling), which varied between fields. Within panels, abscissa shows different phosphorus fertilizer placements (Broad = broadcast, Deep = deep banding, Surface = surface banding), which varied within fields. Dots connected by lines indicate the same field. (Please note that these data are completely fictitious! I made them up one afternoon after skimming some information online.)

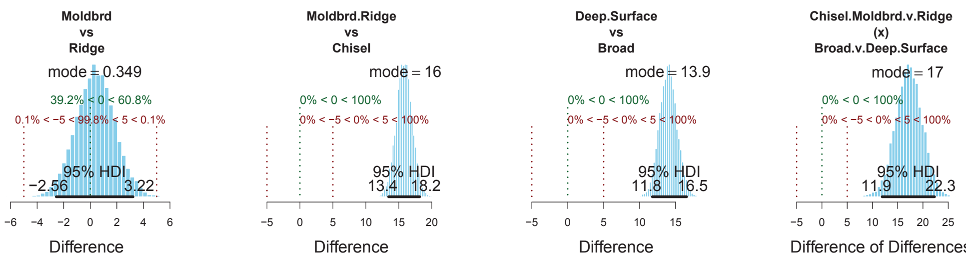

Figure 20.11: Main-effect contrasts and an interaction contrast for the corn-production data in Figure 20.10. (It is worth reiterating that these data are fictitious and might not reflect reality. The tilling and fertilization methods may also differ in effects other than current-year yield, such as cost or future-year soil quality.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

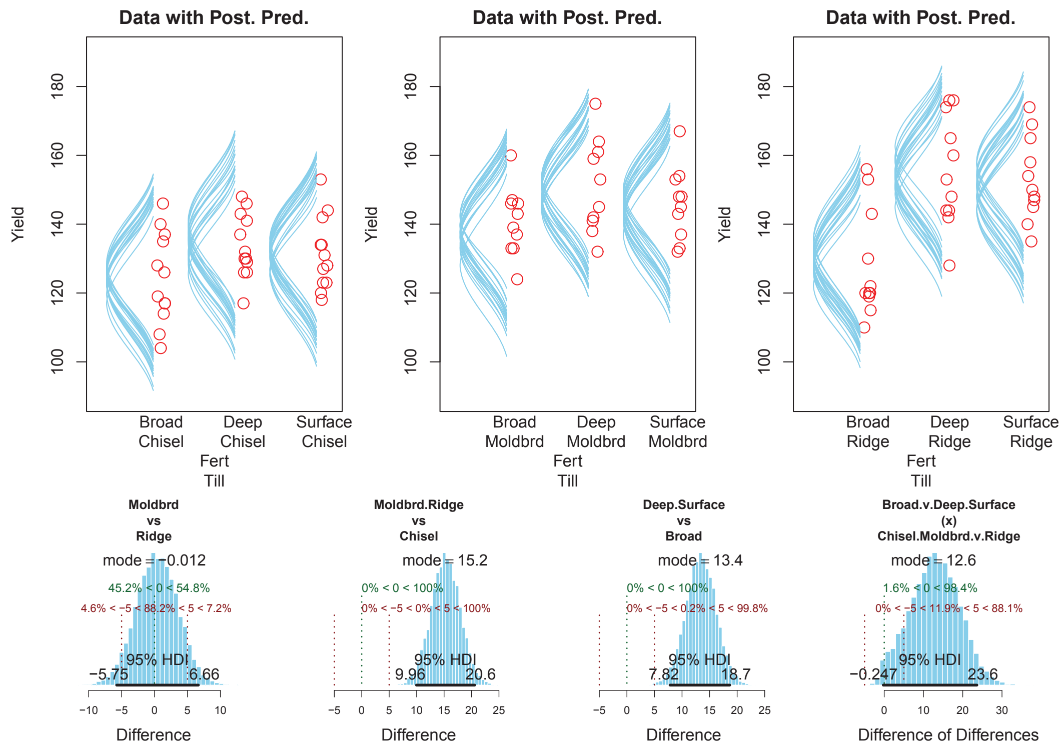

Figure 20.12: Data from Figure 20.10 (p. 570), with field/subject coding suppressed (hence no lines connecting data from the same field/subject). Because between-subject variation is modeled as noise, the posterior predictive distribution has a much larger standard deviation, and the contrasts (shown in the bottom row) are much less certain than in Figure 20.11 (p. 571). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Table 20.1: How to compute sum-to-zero deflections. Start with the cell means, $m_{1\times2[j,k]}$, where j refers to the row and k refers to the column. Then compute the marginal means, $m_{1[j]}$, $m_{2[k]}$, and m. Then compute the baseline β_0, the main effect deflections $\beta_{1[j]}$ and $\beta_{2[k]}$, and the interaction deflections $\beta_{1\times2[j,k]}$.

$m_{1\times2[1,1]}$ $\beta_{1\times2[1,1]}$ $= m_{1\times2[1,1]}$ $- (\beta_{1[1]} + \beta_{2[1]} + \beta_0)$	$m_{1\times2[1,2]}$ $\beta_{1\times2[1,2]}$ $= m_{1\times2[1,2]}$ $- (\beta_{1[1]} + \beta_{2[2]} + \beta_0)$	$m_{1[1]} = \frac{1}{K} \sum_k^K m_{1\times2[1,k]}$ $\beta_{1[1]} = m_{1[1]} - \beta_0$
$m_{1\times2[2,1]}$ $\beta_{1\times2[2,1]}$ $= m_{1\times2[2,1]}$ $- (\beta_{1[2]} + \beta_{2[1]} + \beta_0)$	$m_{1\times2[2,2]}$ $\beta_{1\times2[2,2]}$ $= m_{1\times2[2,2]}$ $- (\beta_{1[2]} + \beta_{2[2]} + \beta_0)$	$m_{1[2]} = \frac{1}{K} \sum_k^K m_{1\times2[2,k]}$ $\beta_{1[2]} = m_{1[2]} - \beta_0$
$m_{2[1]} = \frac{1}{J} \sum_j^J m_{1\times2[j,1]}$ $\beta_{2[1]} = m_{2[1]} - \beta_0$	$m_{2[2]} = \frac{1}{J} \sum_j^J m_{1\times2[j,2]}$ $\beta_{2[2]} = m_{2[2]} - \beta_0$	$m = \frac{1}{J \cdot K} \sum_{j,k}^{J,K} m_{1\times2[j,k]}$ $\beta_0 = m$

Table 20.2: Excerpt from summary table produced by function `smryMCMC` in script `Jags-Ymet-Xnom2fac-MnormalHom-Example.R`. ESS is effective sample size, defined in Equation 7.11, p. 153. In all cases, the HDI mass is 95%. All values are in units of dollars except for the ESS. Although these numbers show many digits, only the first few digits are stable because of randomness in the MCMC process.

Parameter	Mean	Median	Mode	ESS	HDI low	HDI high
b0	127,124	127,131	127,108	12,299	124,785	129,396
b1[1] Assis	-46,394	-46,415	-46,483	13,341	-49,467	-43,310
b1[2] Assoc	-33,108	-33,096	-33,052	12,987	-35,987	-30,378
b1[3] Full	-3,156	-3,159	-3,031	12,097	-6,106	-229
b1[4] Endow	26,966	26,980	27,285	13,405	22,424	31,583
b1[5] Disting	55,692	55,738	56,531	12,229	48,404	62,670
b2[21] ENG	-19,412	-19,380	-19,041	12,280	-27,416	-11,812
b2[49] PSY	6,636	6,653	6,686	12,604	353	12,494
b2[13] CHEM	19,159	19,152	19,221	14,597	12,698	25,582
b2[8] BFIN	109,184	109,200	109,156	14,287	100,185	118,579
b1b2[1,49] Assis PSY	-3,249	-3,136	-2,060	15,000	-13,588	6,682
b1b2[3,49] Full PSY	-14,993	-14,997	-15,474	11,963	-23,360	-6,463
b1b2[1,13] Assis CHEM	-12,741	-12,692	-13,110	13,224	-22,151	-3,457
b1b2[3,13] Full CHEM	12,931	12,971	13,087	12,772	3,471	22,240
ySigma	17,997	17,985	17,953	11,968	17,144	18,852

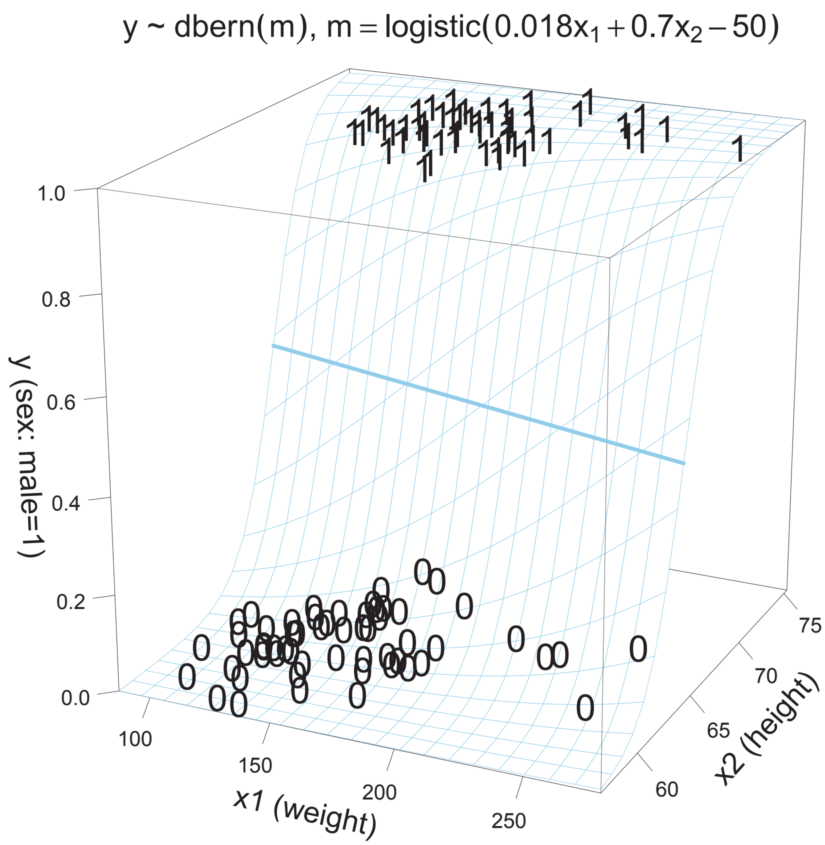

Figure 21.1: Data show gender (arbitrarily coded as male=1, female=0) as a function of weight (in pounds) and height (in inches). All 0's are located on the bottom plane of the cube, and all 1's are located on the top plane of the cube. Logistic surface shows maximum-likelihood estimate. Heavy line shows 50% threshold.

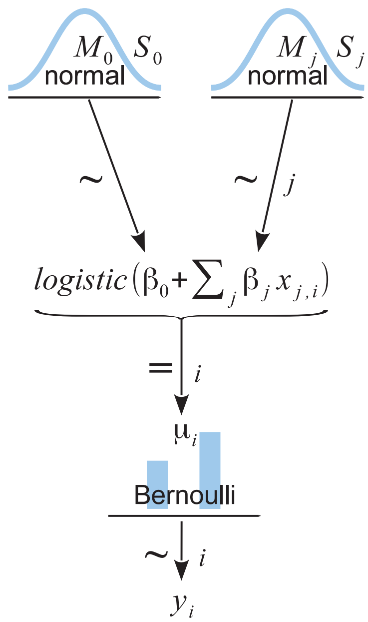

Figure 21.2: Dependency diagram for multiple logistic regression. Compare with the diagram for robust multiple linear regression in Figure 18.4 (p. 498).

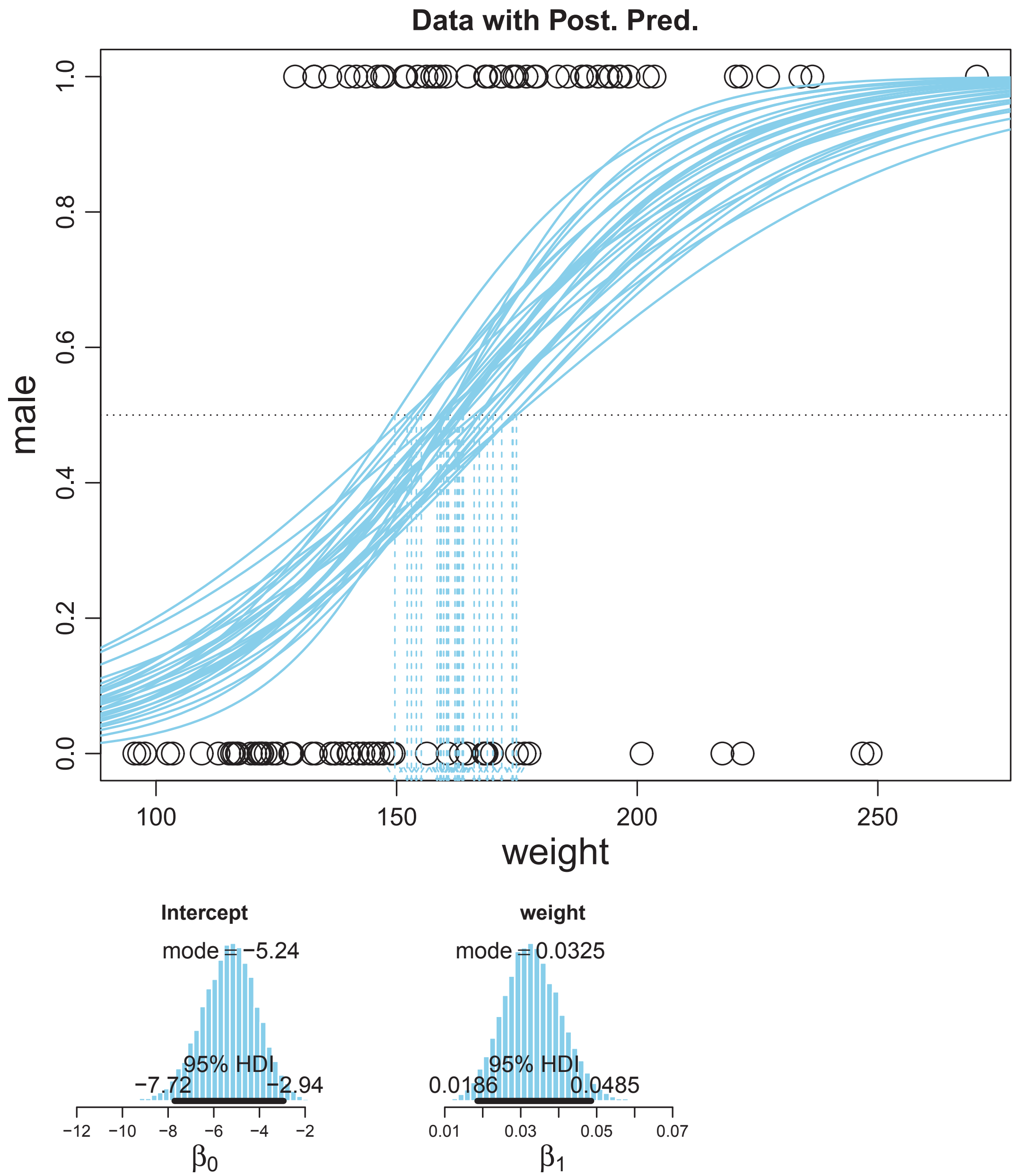

Figure 21.3: Predicting gender (arbitrarily coded as male=1, female=0) as a function of weight (in pounds), using logistic regression. Upper panel: Data are indicated by dots. Logistic curves are a random sample from the MCMC posterior. Descending arrows point to threshold weights at which the probability of male is 50%. Lower panels: Marginal posterior distribution on baseline and slope. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

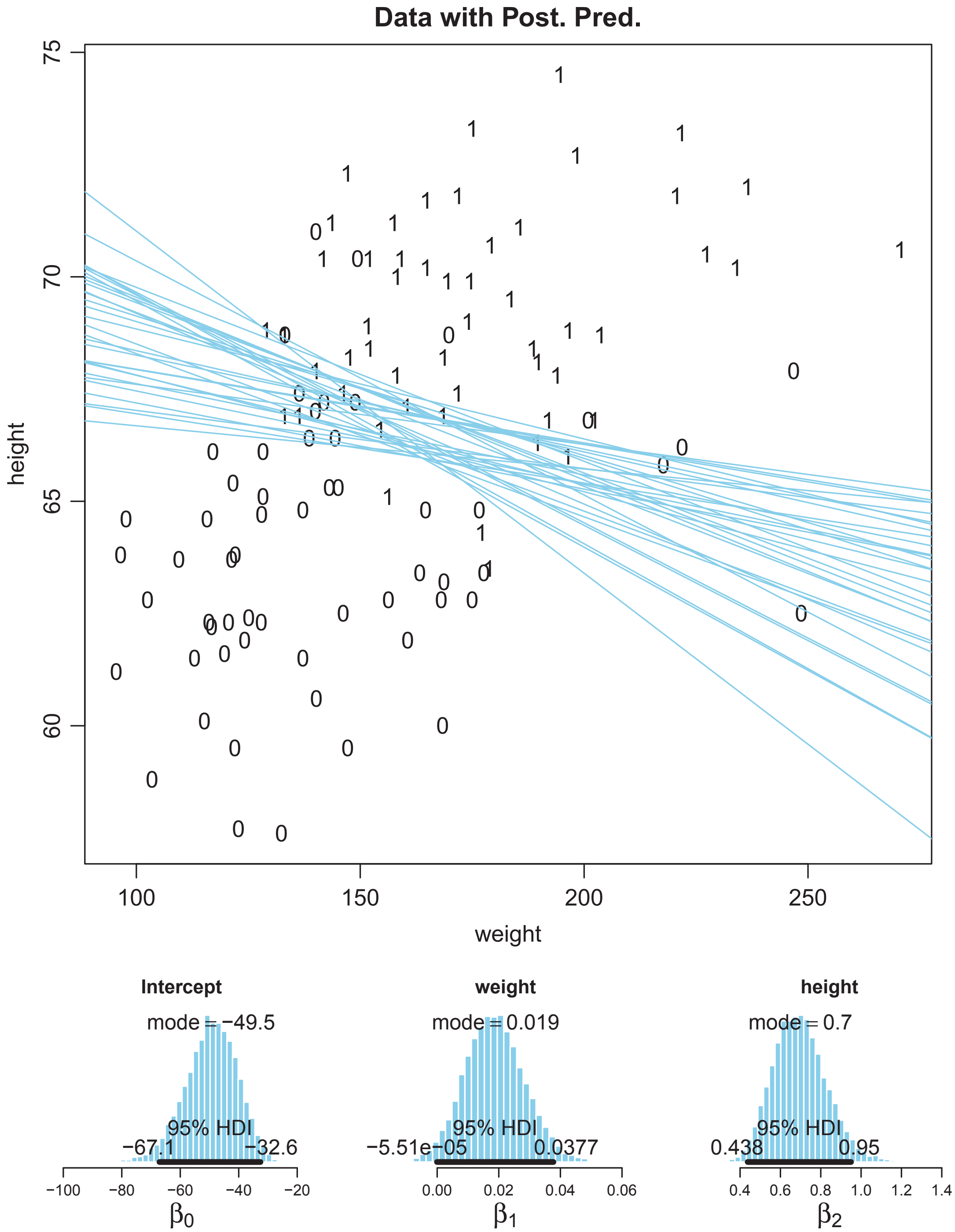

Figure 21.4: Predicting gender (arbitrarily coded as male=1, female=0) as a function of weight (in pounds) and height (in inches), using logistic regression. Upper panel: Data are indicated by 0's and 1's. Lines show a random sample from the MCMC posterior of thesholds at which the probability of male is 50%. Lower panels: Marginal posterior distribution on baseline and slopes. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

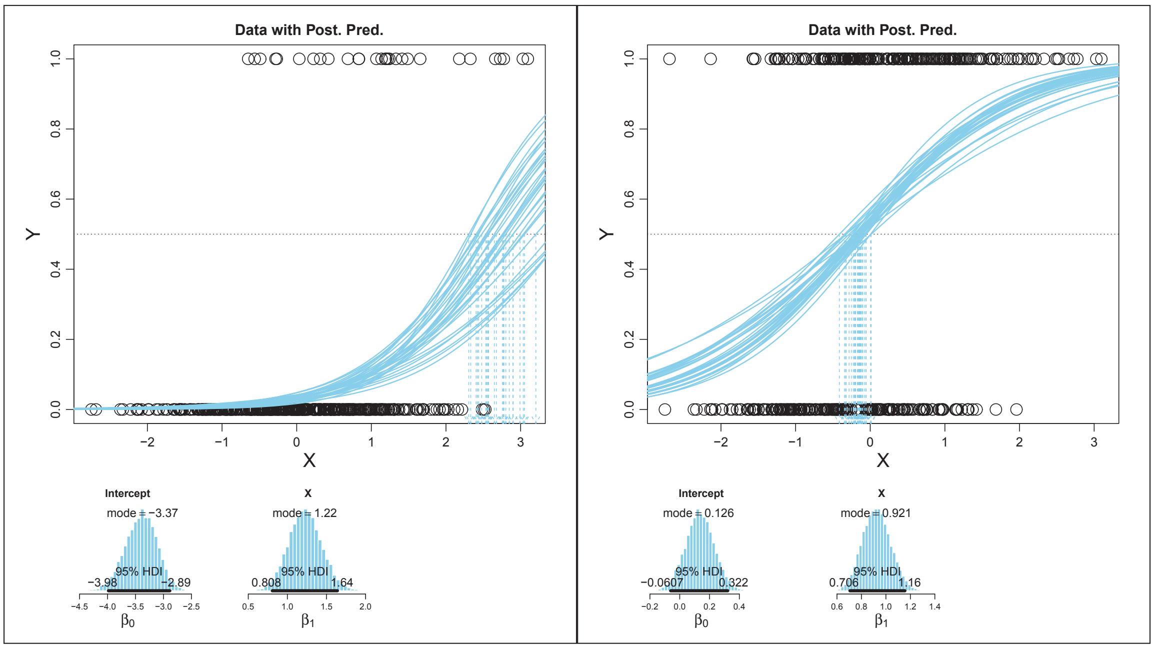

Figure 21.5: Parameter estimates are more uncertain when there are few 0's or 1's in the data. The main panels have data with the same x values, and y values randomly generated by logistic functions with the same slope ($\beta_1 = 1$) but different intercepts ($\beta_0 = -3$ and $\beta_0 = 0$).

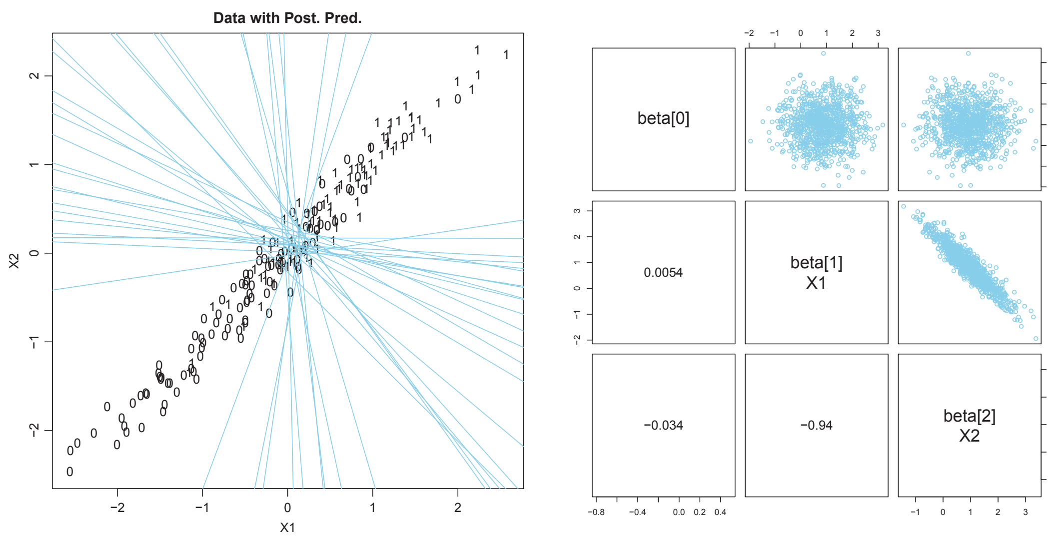

Figure 21.6: Estimates of slope parameters trade off when the predictors are correlated. Left panel shows credible 50% level contours superimposed on data. Right panel shows strong anti-correlation of credible β_1 and β_2 values. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

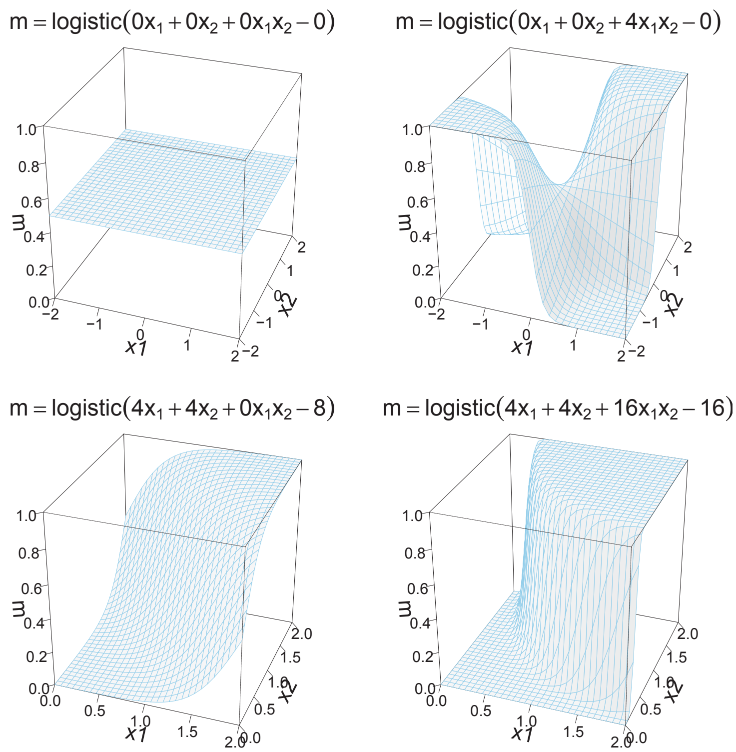

Figure 21.7: Multiplicative interaction of metric predictors in logistic regression. Left column shows examples of no interaction. Right column shows corresponding logistic surfaces with interaction. Title of each plot shows the coefficients on the predictors. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

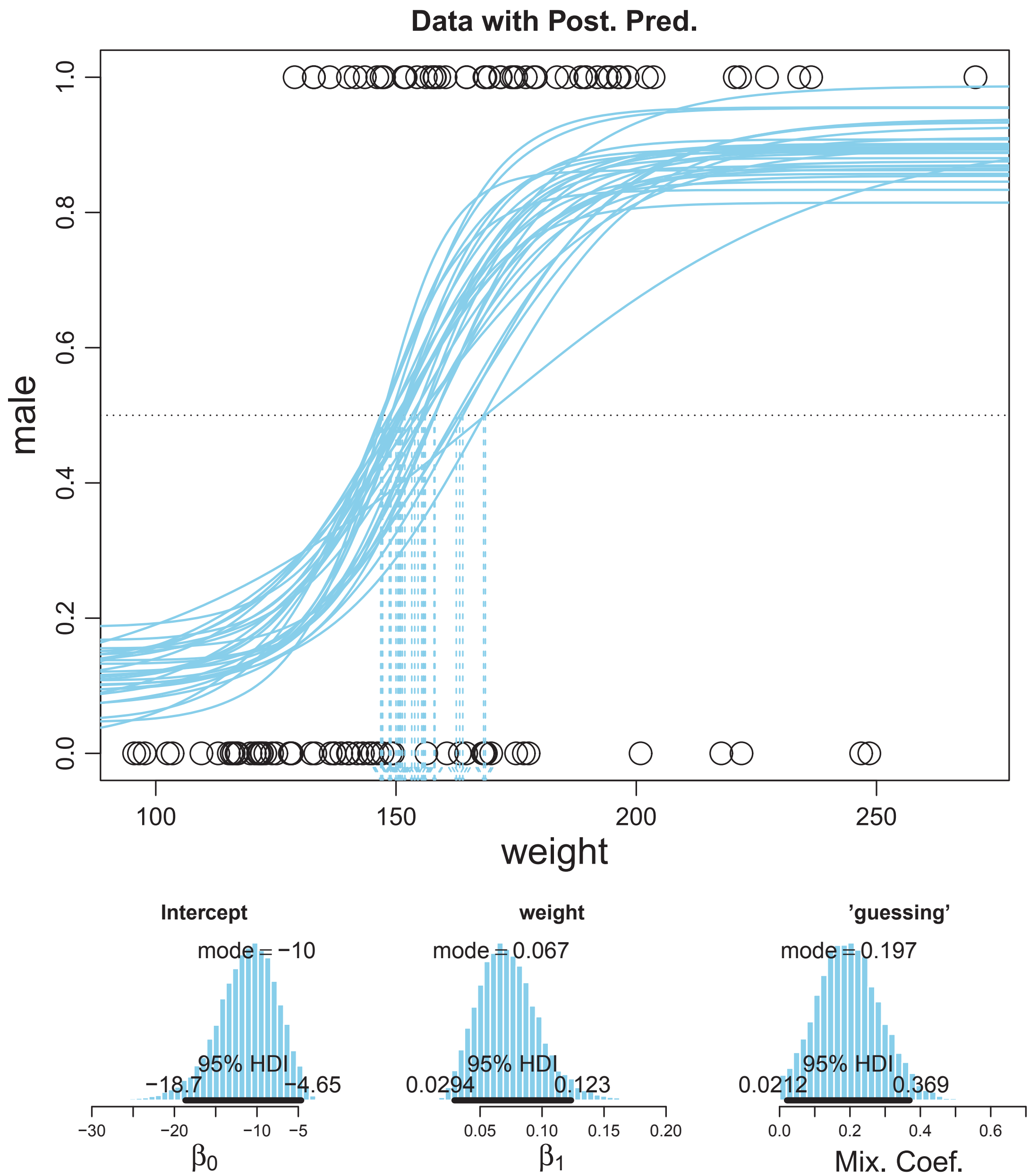

Figure 21.8: Predicting gender (arbitrarily coded as male=1, female=0) as a function of weight (in pounds), using *robust* logistic regression. Upper panel: Data are indicated by dots, the same as in Figure 21.3. Curves are a random sample from the MCMC posterior; notice asymptotes away from 0,1 limits. Descending arrows point to threshold weights at which the probability of male is 50%. Lower panels: Marginal posterior distribution on baseline, slope, and guessing coefficient. Pairwise plots are shown in Figure 21.9.

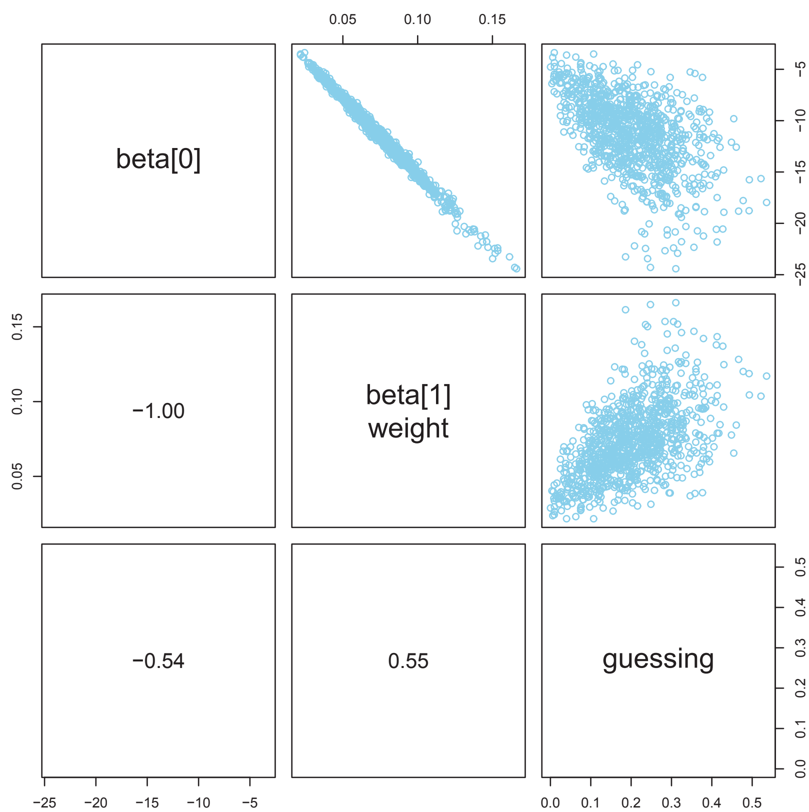

Figure 21.9: Pairwise plots for posterior distribution in Figure 21.8. Notice here that the guessing coefficient is correlated with the slope, beta[1]. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

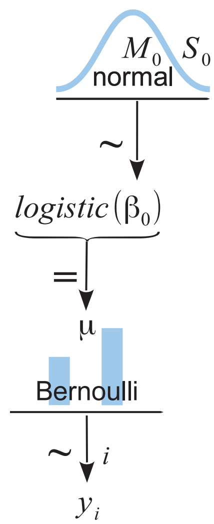

Figure 21.10: Estimating the underlying bias of a coin using a logistic function of a baseline. When $\beta_0 = 0$, then $\mu = 0.5$, and the coin is fair. Compare with the beta prior in Figure 8.2 (p. 193).

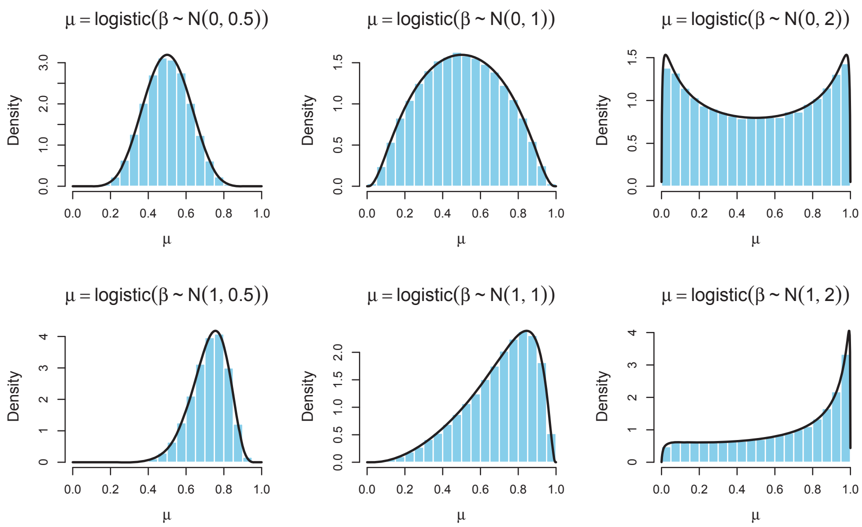

Figure 21.11: Prior distributions on μ for different choices of M_0 and S_0 in Figure 21.10. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

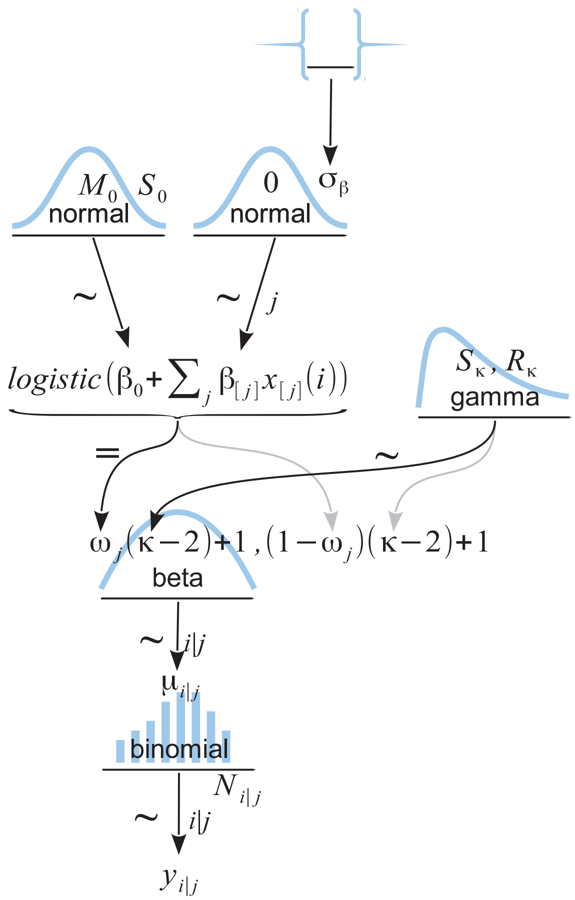

Figure 21.12: Hierarchical diagram for logistic ANOVA-like model. The top part of the structure is based on the ANOVA-like model of Figure 19.2 (p. 529). The lower part of the structure is based on the models of Figure 9.7 (p. 229) and Figure 9.13 (p. 235). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

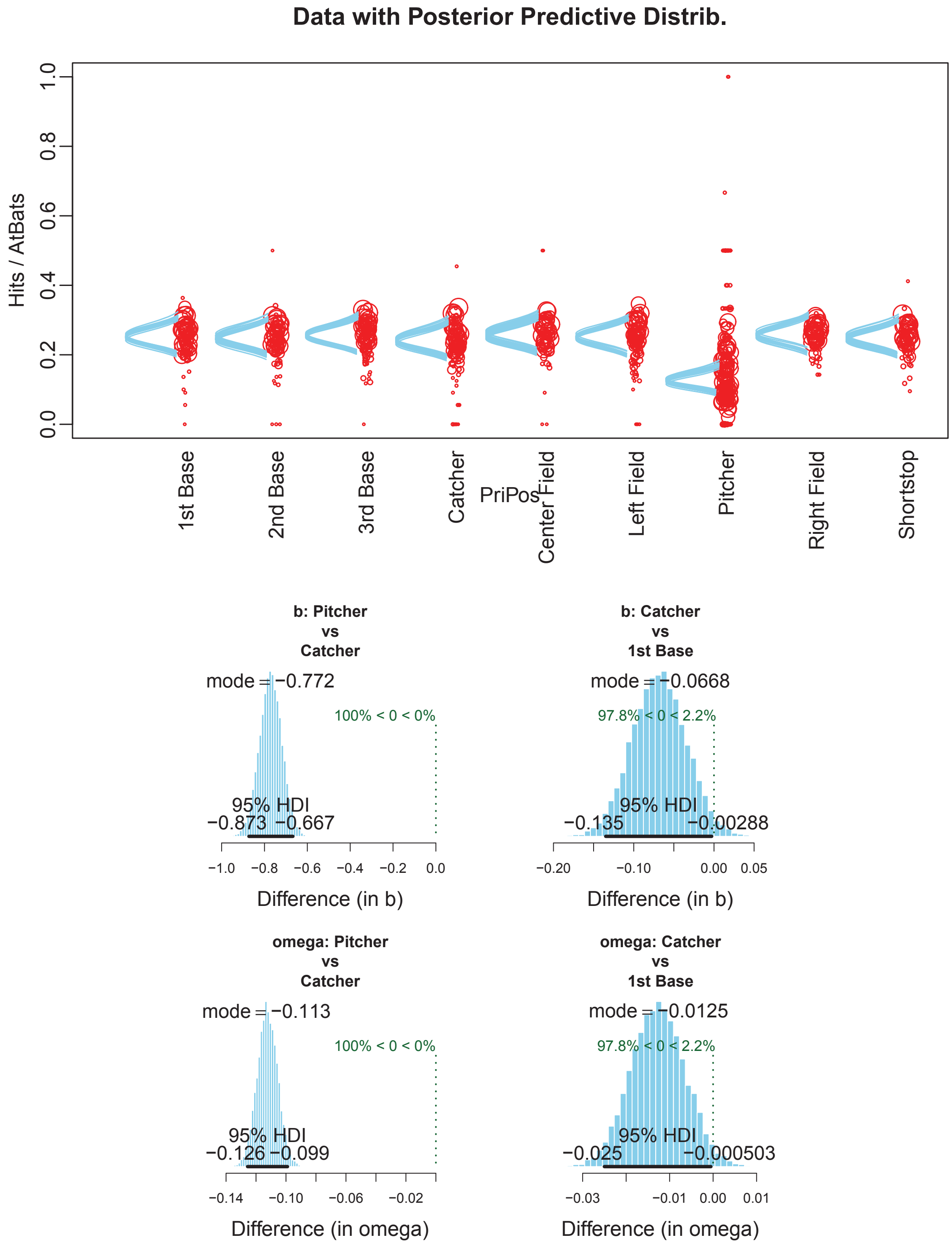

Figure 21.13: Baseball batting data are shown in the upper panel with dot size proportional to number of at-bats. The posterior predictive distributions are credible beta distributions assuming homogeneous concentration across positions. Lower panels show selected contrasts, which can be compared with the contrasts shown in Figure 9.14 (p. 236).

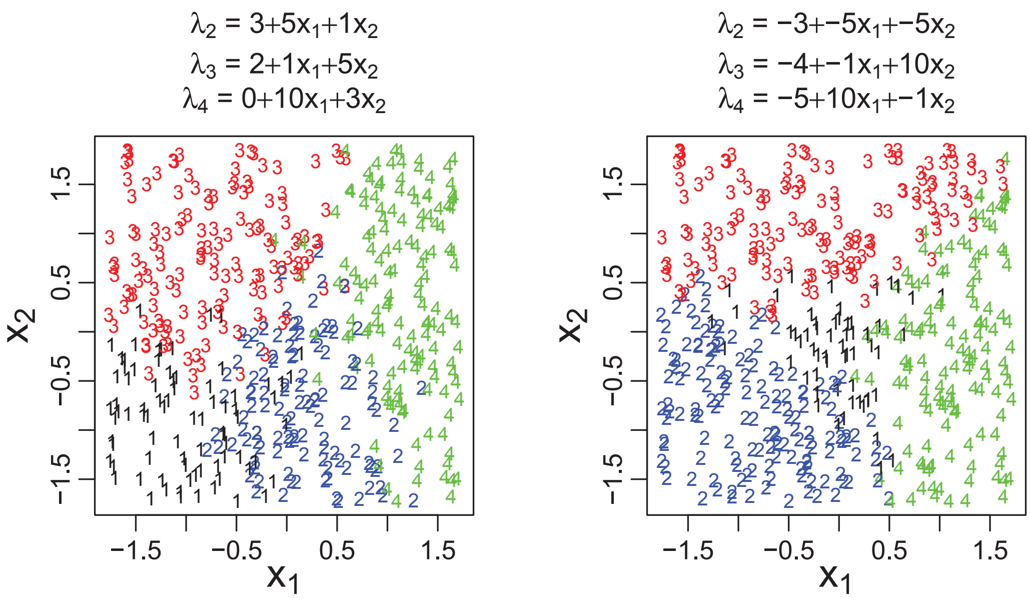

Figure 22.1: Examples of data generated from the softmax regression model. Above each panel are the specific instantiations of Equation 22.1 (for two predictors instead of only one predictor), with the reference outcome chosen to be 1, so $\lambda_1 \equiv 0$. The outcomes were sampled according to probabilities computed from the softmax function of Equation 22.2.

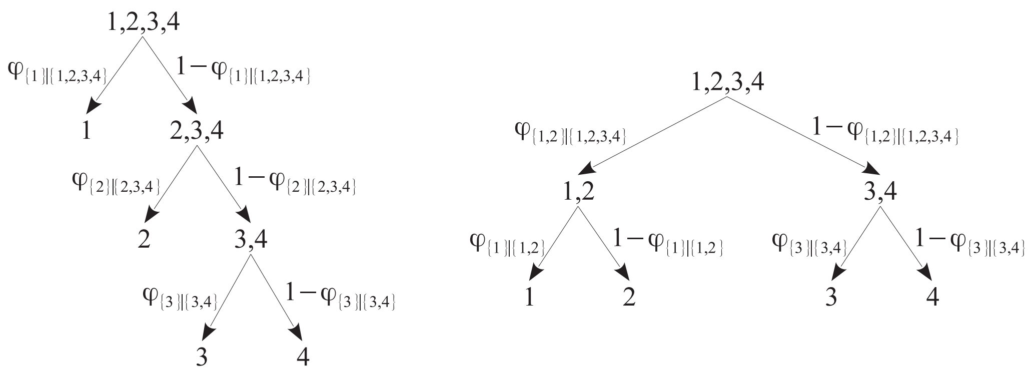

Figure 22.2: Two hierarchies of binary divisions of outcomes 1, 2, 3, and 4. Each branch is labelled with its conditional probability. In conditional logistic regression, each binary conditional probability is modeled by a logistic function. An example of data generated from the left hierarchy is shown in the left side of Figure 22.3, and an example of data generated from the right hierarchy is shown in the right side of Figure 22.3. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

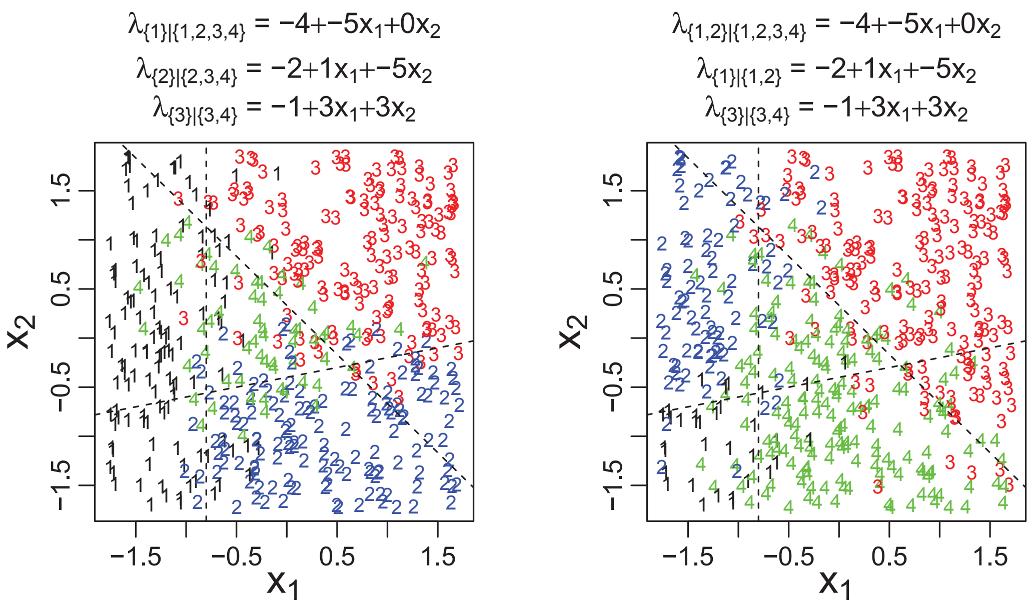

Figure 22.3: *Left panel:* Example of data generated from the conditional logistic model of Equations 22.9–22.11, which express the hierarchy in the left side of Figure 22.2. *Right panel:* Example of data generated from the conditional logistic model of Equations 22.12–22.14, which express the hierarchy in the right side of Figure 22.2. Dashed lines indicate 50% level contours of the conditional logistic functions. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

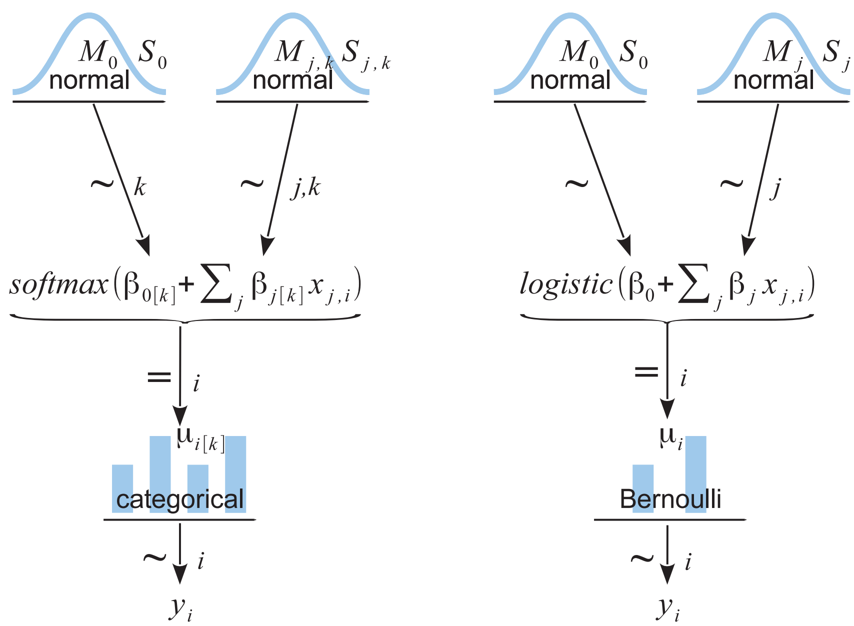

Figure 22.4: Hierarchical diagrams for softmax regression (on left) and for logistic regression (on right), both for metric predictors. Right diagram is repeated from Figure 21.2 (p. 591). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

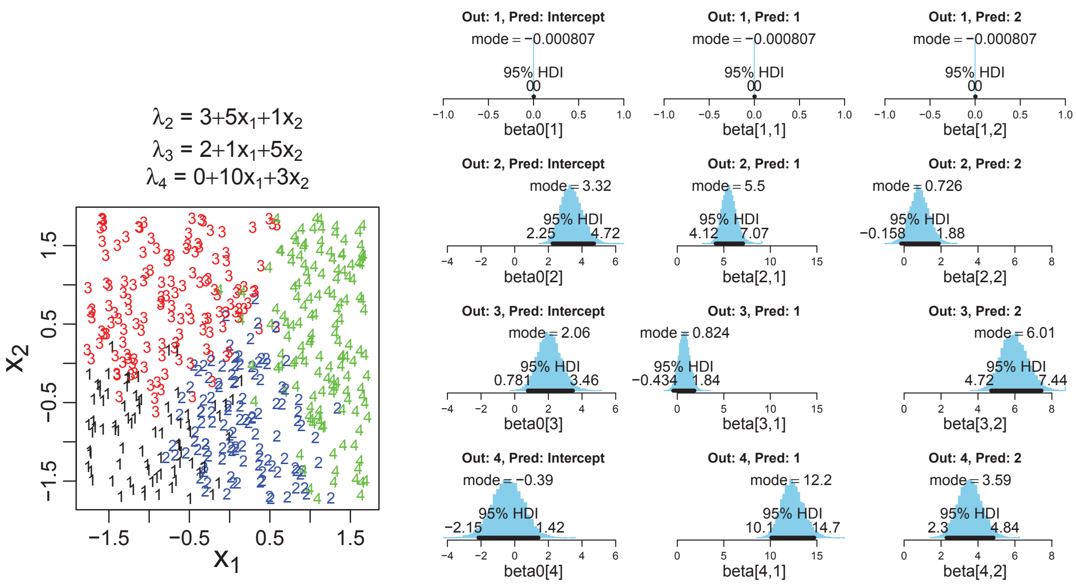

Figure 22.5: Posterior parameter estimates of the softmax model applied to data generated from a softmax model. Data are shown on the left with the true parameter values (reproduced from Figure 22.1). The four rows of marginal posterior distributions correspond to the four outcomes, and the columns of distributions correspond to β_0, β_1, and β_2. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

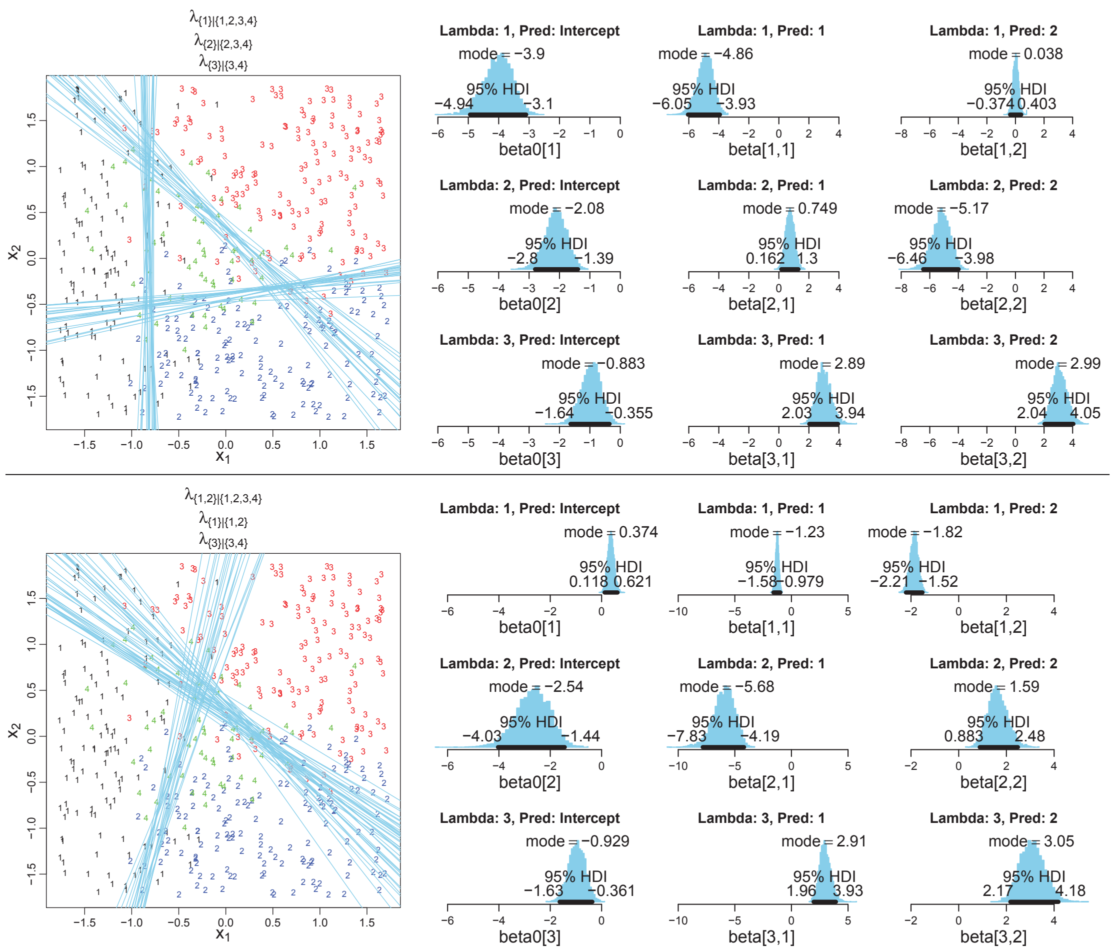

Figure 22.6: *Upper half:* Posterior parameter estimates of the conditional logistic model of Equation 22.11 for data in left panel of Figure 22.3 (i.e., the model structure matches the data generator). The three rows of distributions correspond to the three lambda functions. The columns of distributions correspond to β_0, β_1, and β_2 of the lambda function. *Lower half:* Posterior parameter estimates of the conditional logistic model of Equation 22.14 (i.e., model structure does not match data generator). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

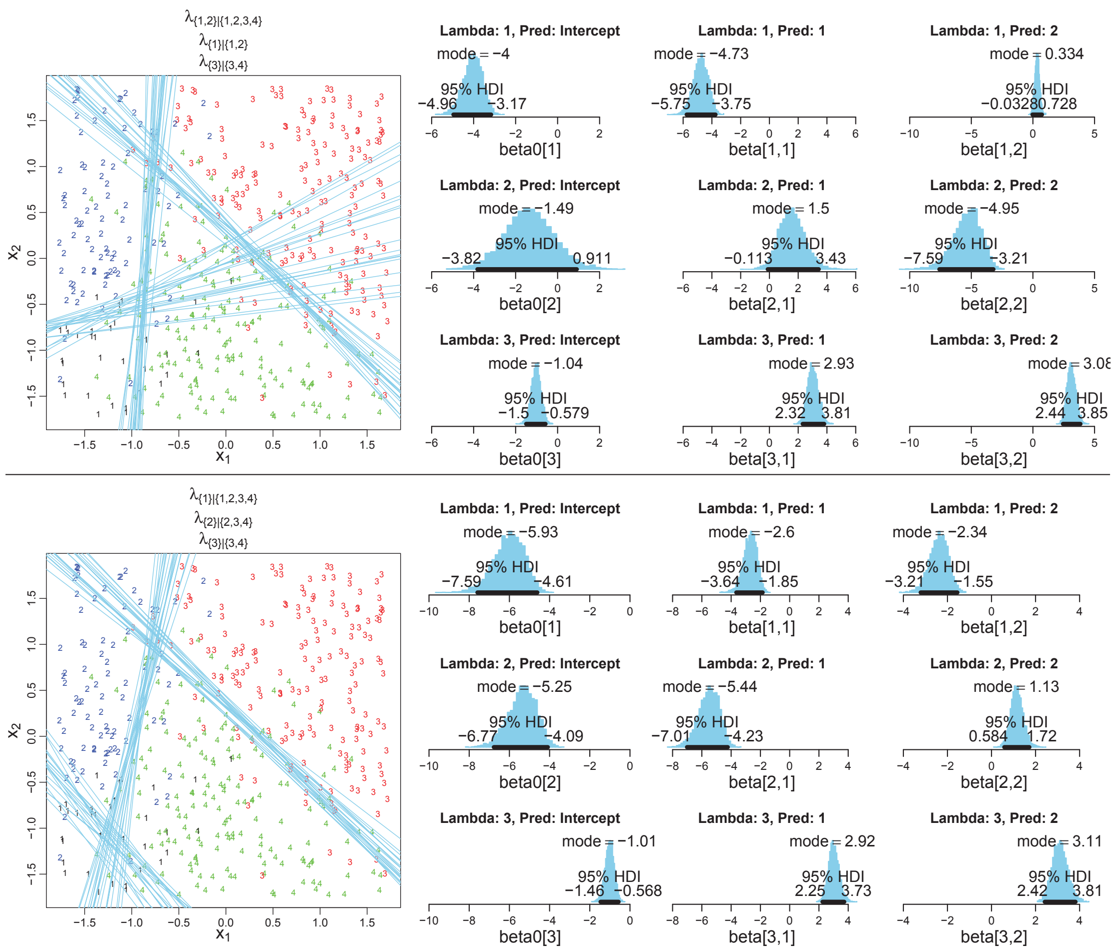

Figure 22.7: *Upper half:* Posterior parameter estimates of the conditional logistic model of Equation 22.14 for data in right panel of Figure 22.3 (i.e., the model structure matches the data generator). The three rows of distributions correspond to the three lambda functions. The columns of distributions correspond to β_0, β_1, and β_2 of the lambda function. *Lower half:* Posterior parameter estimates of the conditional logistic model of Equation 22.11 (i.e., model structure does not match data structure). Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

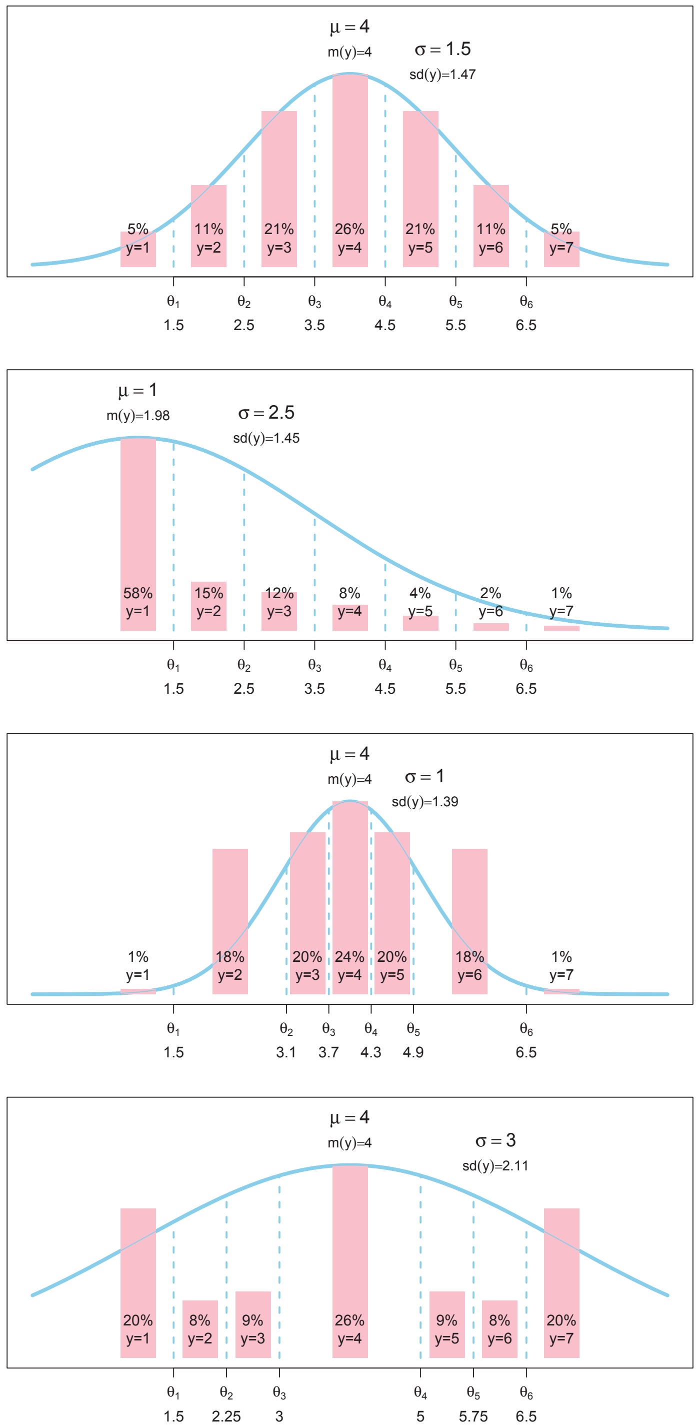

Figure 23.1: Examples of ordinal data generated by the thresholded cumulative-normal model. Ordinal outcome values (y) are indicated by bars. The horizontal axis is the underlying continuous value, which has a normal distribution with mean μ and standard deviation σ. The mean and standard deviation of the ordinal data values (treated as if metric) are annotated as m(y) and sd(y). Thresholds between ordinal outcome values on underlying continuous scale are denoted θ_j. See main text for remarks about each example.

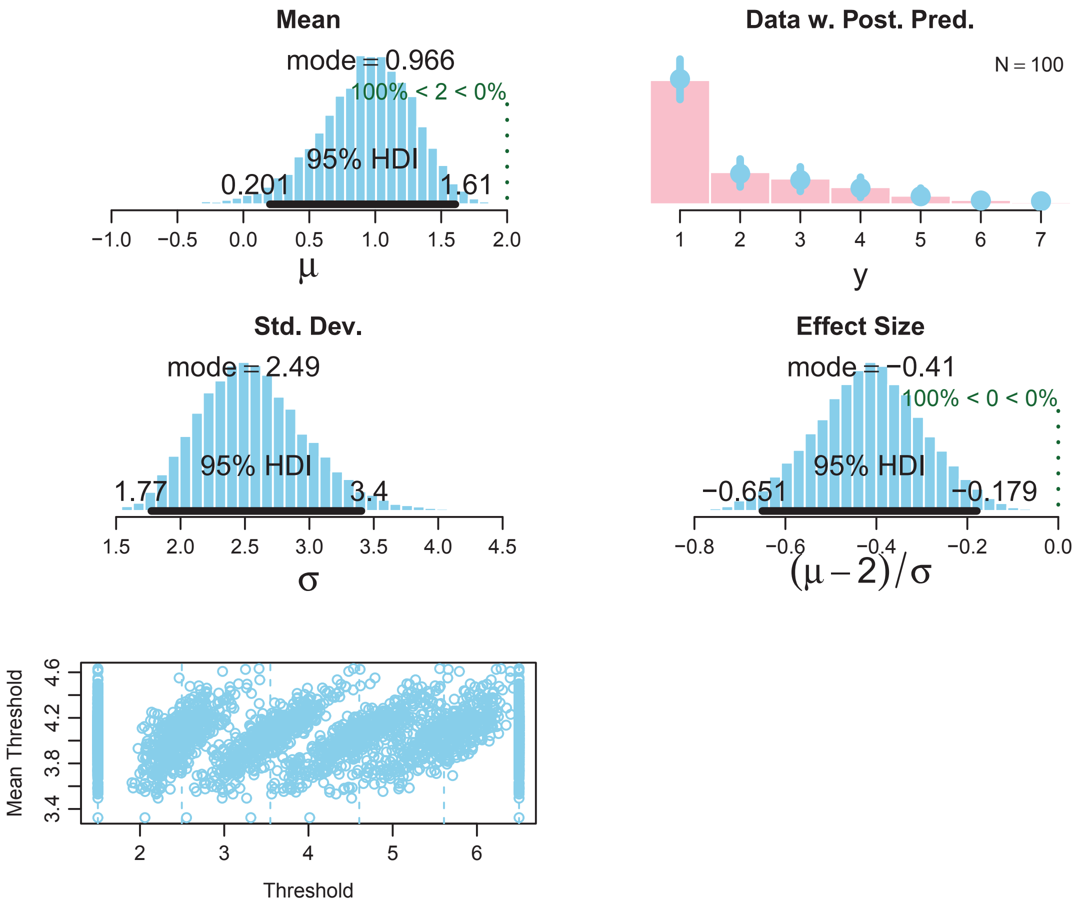

Figure 23.2: Bayesian posterior distribution for one group of ordinal data. True generating parameters are $\mu = 1.0$, $\sigma = 2.5$, $\theta_1 = 1.5$, $\theta_2 = 2.5$, $\theta_3 = 3.5$, $\theta_4 = 4.5$, $\theta_5 = 5.5$, and $\theta_6 = 6.5$. The Bayesian estimation accurately recovers the generating parameters. Posterior distribution clearly excludes a comparison value of $\mu = 2.0$. The posterior predictive distribution accurately describes the data distribution. *NHST treating data as metric:* The mean is *not* significantly different from a comparison value of $\mu = 2.0$: $M = 1.95$, $t = 0.36$, $p = 0.722$, with 95% CI of 1.67 to 2.23, and with effect size $d = 0.036$. The sample standard deviation is $S = 1.40$. The t test describes the data as normally distributed, which clearly is not the case here.

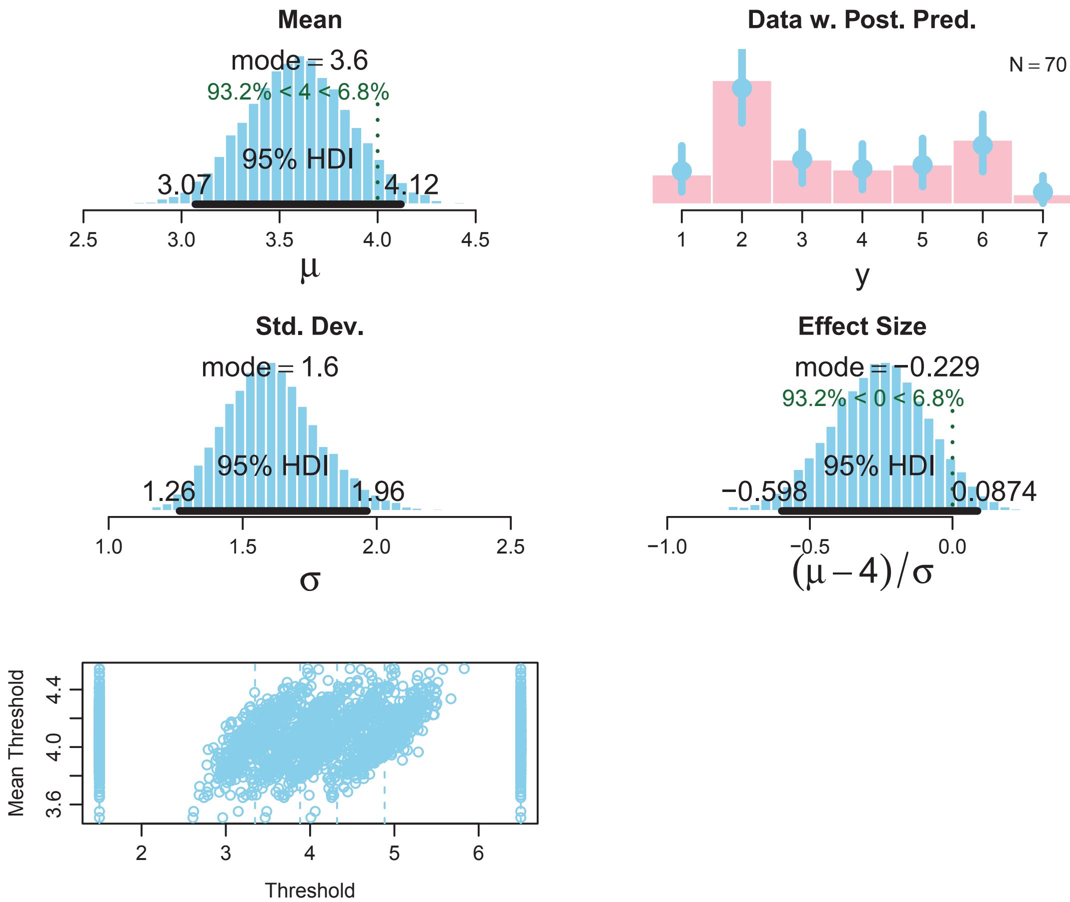

Figure 23.3: Bayesian posterior distribution for one group of ordinal data. True generating parameters are $\mu = 3.5$ and $\sigma = 1.5$, with $\theta_1 = 1.5$, $\theta_2 = 3.3$, $\theta_3 = 3.8$, $\theta_4 = 4.2$, $\theta_5 = 4.7$, and $\theta_6 = 6.5$. The Bayesian estimation accurately recovers the generating parameters. Posterior distribution includes (does not reject) a comparison value of $\mu = 4.0$. Importantly, the posterior predictive distribution nicely describes the data distribution. *NHST treating data as metric:* The mean is significantly different from $\mu = 4.0$: $M = 3.47$, $t = 2.47$, $p = 0.016$, with 95% CI of 3.04 to 3.90, and effect size $d = 0.295$. The sample standard deviation is $S = 1.79$. The t test describes the data as normally distributed which clearly is not the case here. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

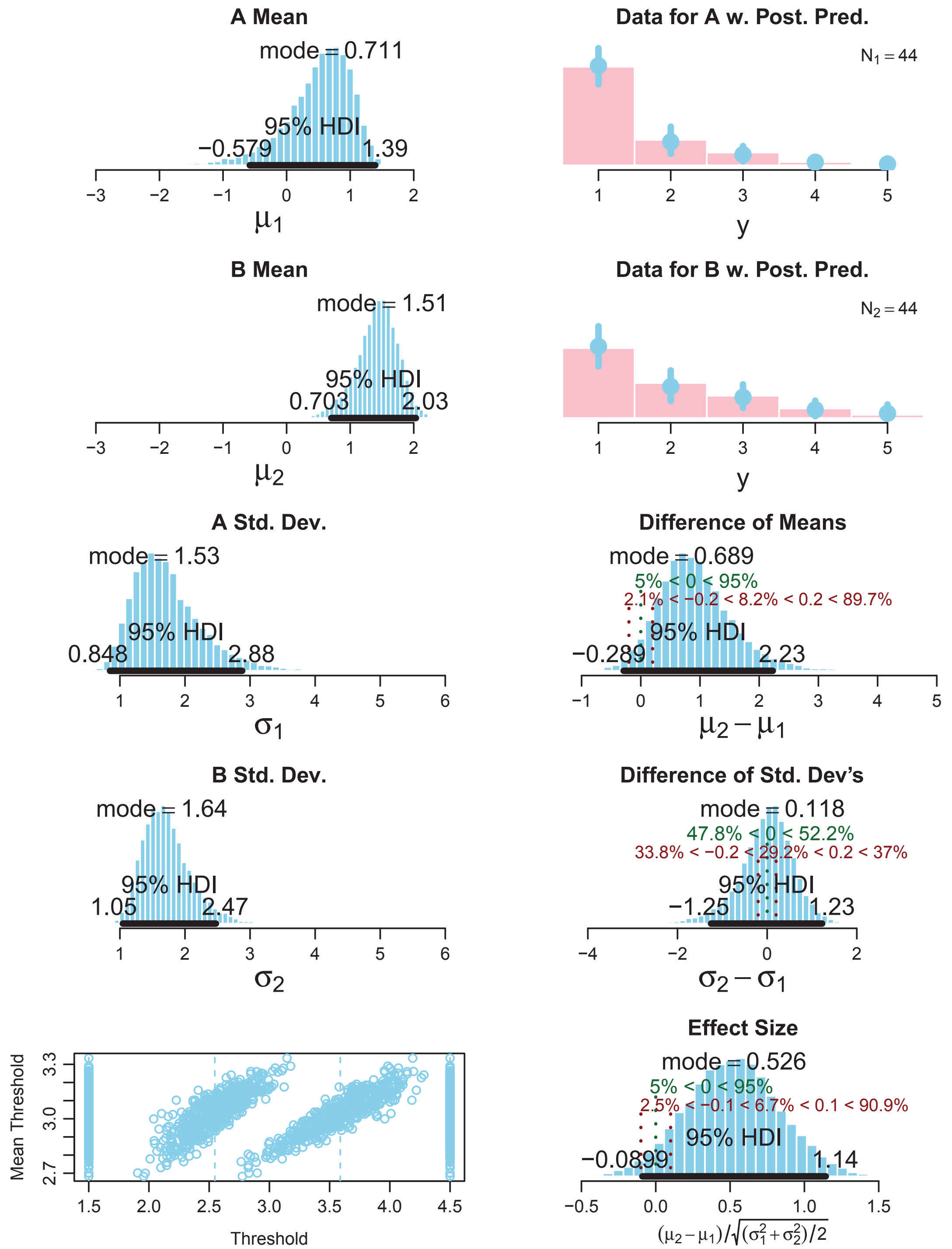

Figure 23.4: Bayesian posterior distribution for two groups of ordinal data. The true generating parameters are $\mu_1 = 0.7$, $\mu_2 = 1.5$, $\sigma_1 = 1.6$, $\sigma_2 = 1.6$ (notice equal variances), with $\theta_1 = 1.5$, $\theta_2 = 2.5$, $\theta_3 = 3.5$, and $\theta_4 = 4.5$. The Bayesian estimation accurately recovers the generating parameters. The 95% HDIs include zero effect size and zero difference of standard deviations. *NHST treating data as metric:* The means are significantly different: $M_1 = 1.43$, $M_2 = 1.86$, $t = 2.18$, $p = 0.032$, with effect size $d = 0.466$ with 95% CI of 0.036 to 0.895. An F test of the variances concludes that the standard deviations are significantly different: $S_1 = 0.76$, $S_2 = 1.07$, $p = .027$. Notice in this case that treating the values as metric greatly underestimates their variances, as well as erroneously concluding the variances are different. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

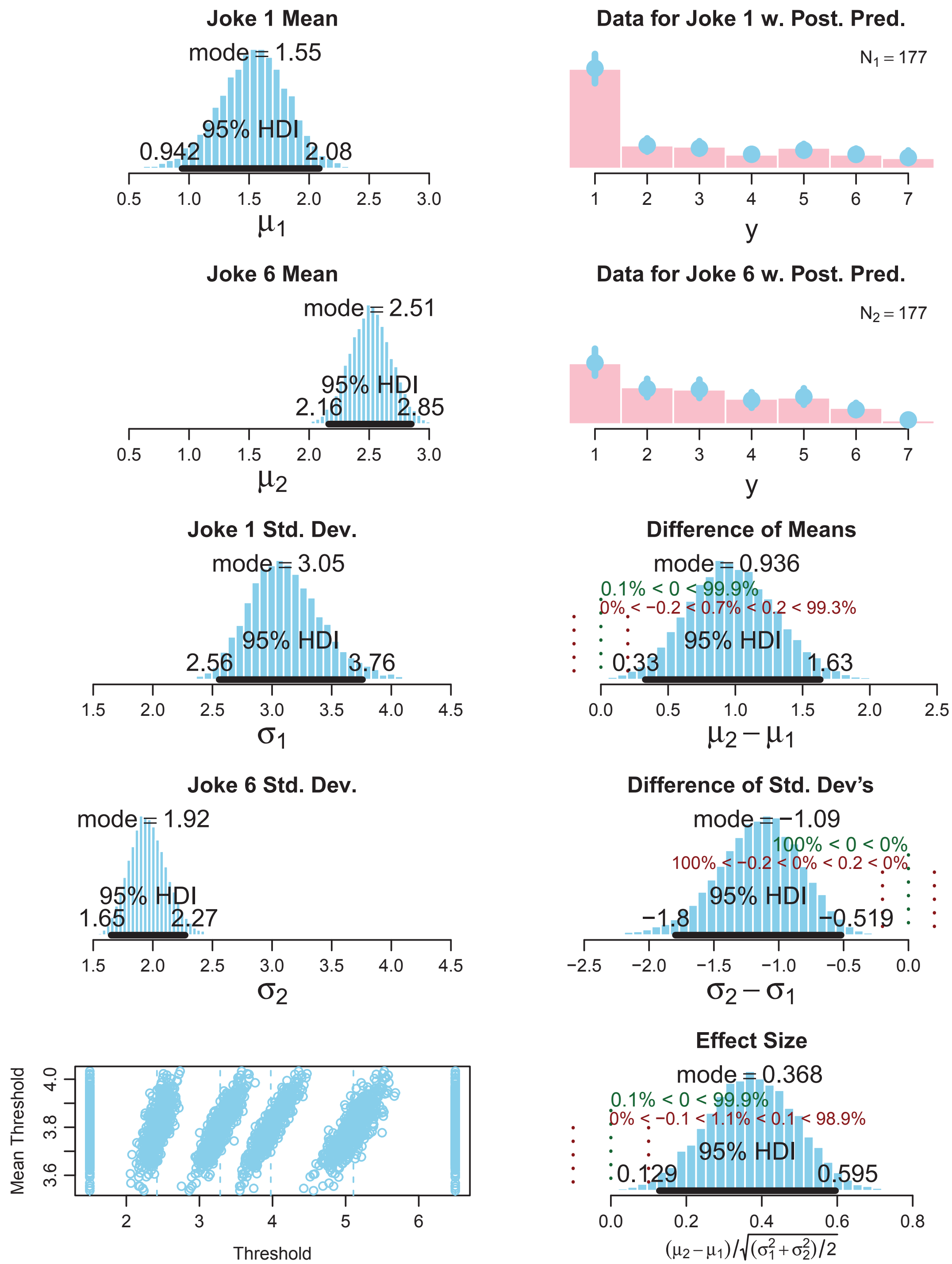

Figure 23.5: Bayesian posterior distribution for two groups of ordinal data from funniness ratings of jokes. Notice that the mean funniness ratings and standard deviations are clearly different. *NHST treating data as metric:* The means are not significantly different: $M_1 = 2.59$, $M_2 = 2.91$, $t = 1.67$, $p = 0.096$, with effect size $d = 0.178$ with 95% CI of −0.032 to 0.387. An F test of the variances concludes that the standard deviations are not significantly different: $S_1 = 1.96$, $S_2 = 1.73$, $p = .116$. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

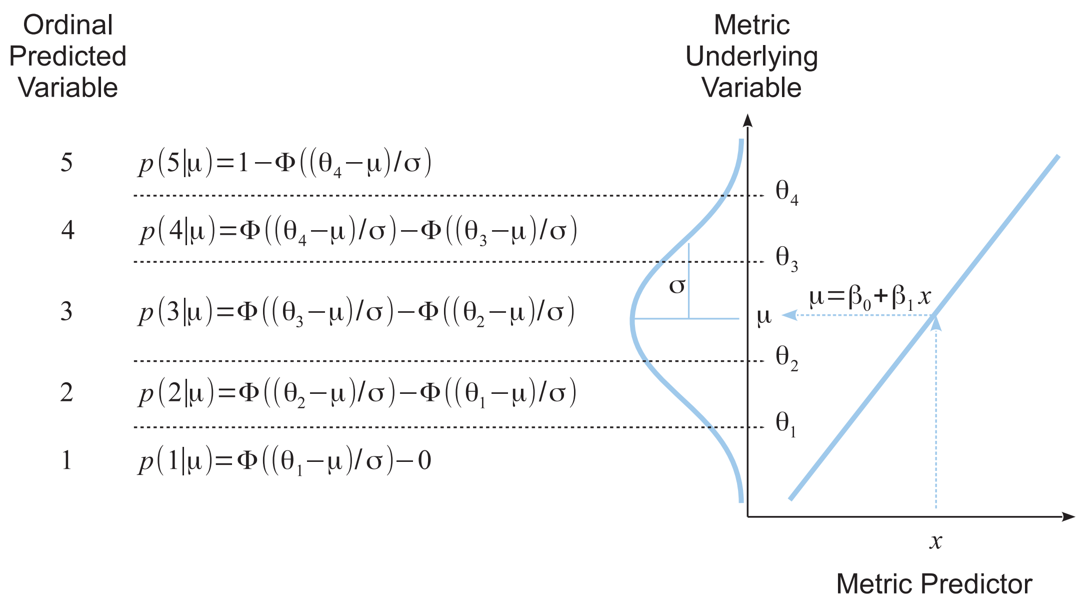

Figure 23.6: Thresholded cumulative-normal regression. Right side shows metric predictor variable mapped to metric underlying variable, as in simple linear regression of Figure 17.1 (p. 462). Left side shows mapping from metric underlying to observed ordinal variable, displaying Equations 23.1–23.3 at the corresponding intervals between thresholds. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

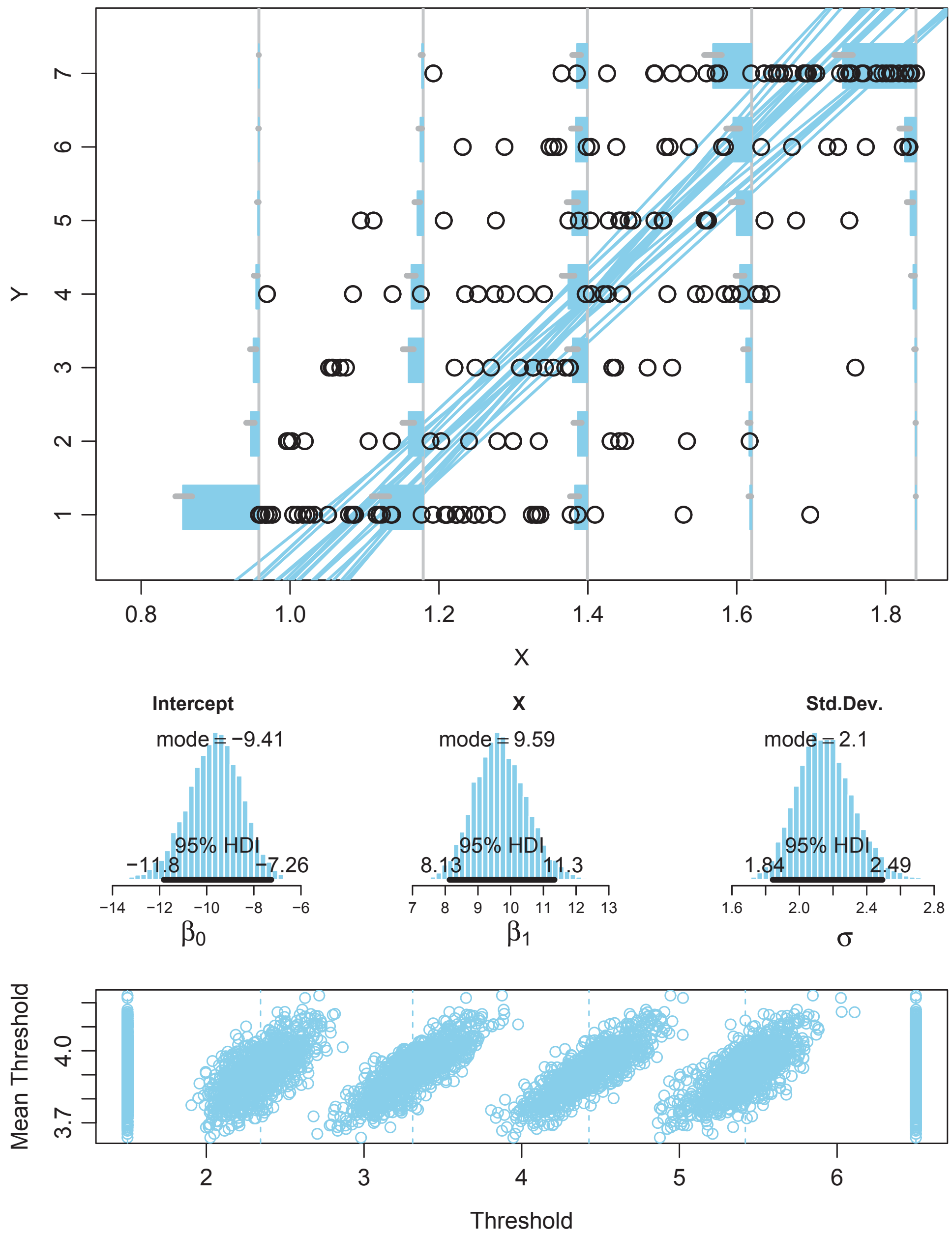

Figure 23.7: Upper panel: Data are shown as dots. Horizontal bars show mean posterior predicted probability at selected values of the predictor as marked by the vertical lines. Grey segments at tops (i.e., left end) of bars show the 95% HDI of posterior predicted probability. A smattering of credible regression lines is superimposed. The true generating parameters are $\beta_0 = -10.0$, $\beta_1 = 10.0$, $\sigma = 2.0$, $\theta_1 = 1.5$, $\theta_2 = 2.5$, $\theta_3 = 3.5$, $\theta_4 = 4.5$, $\theta_5 = 5.5$, $\theta_6 = 6.5$. Lower panels: The Bayesian estimation accurately recovers the generating parameters, as indicated by the marginal posterior distributions. *Least-squares estimate treating data as metric:* $\beta_0 = -5.42$ (SE=0.61), $\beta_1 = 6.71$ (SE=0.43), $\sigma = 1.52$. Notice that the slope is badly underestimated by least-squares estimation in this case. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

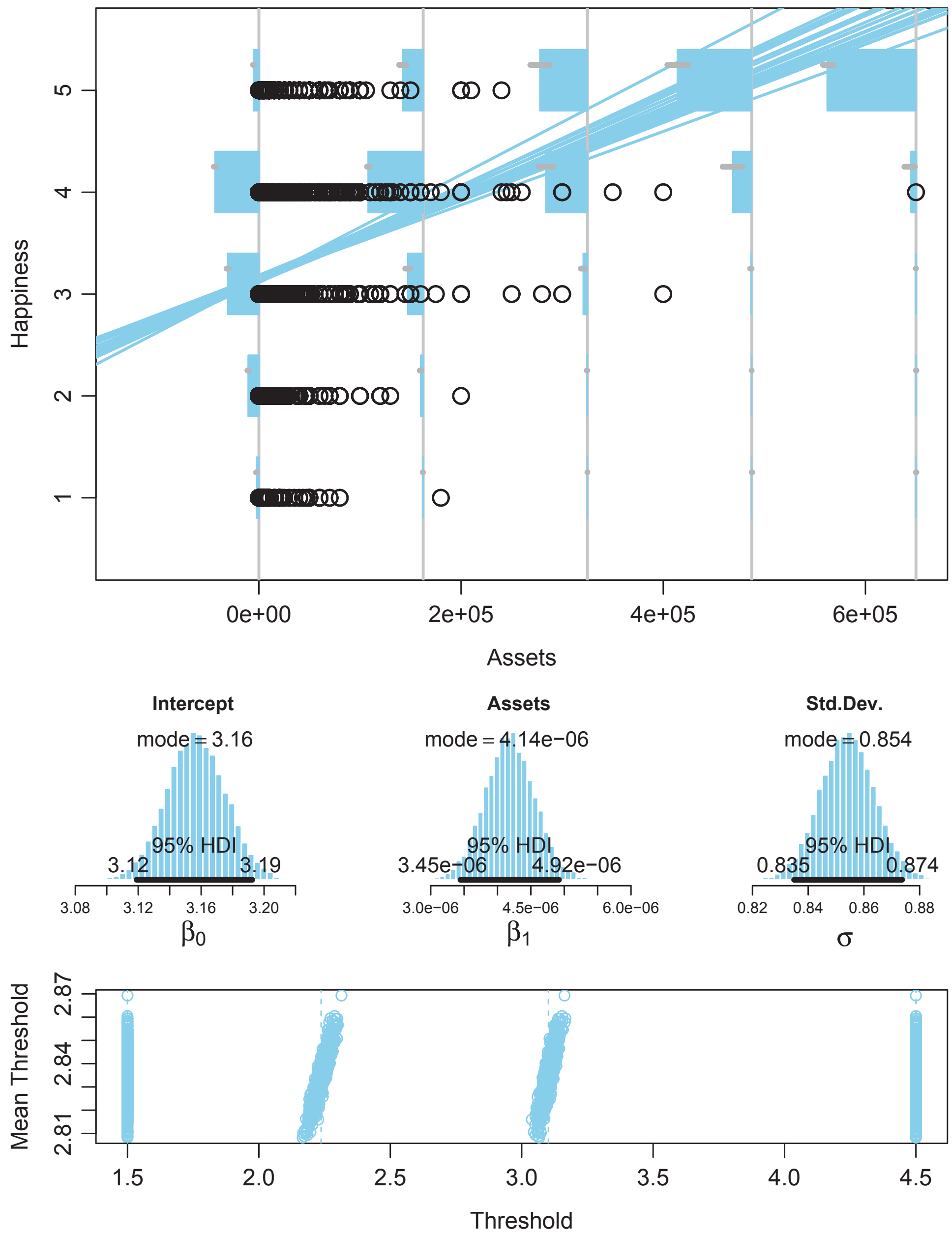

Figure 23.8: Upper panel: Happiness as a function of total household assets ($N = 6,835$, data from Shi, 2009). Assets are in 2002 Chinese yaun (2e+05 yuan was equivalent to about 24,200 US dollars in 2002). Horizontal bars show mean posterior predicted probability at selected values of the predictor as marked by the vertical lines. Grey segments at tops (i.e., left end) of bars show the 95% HDI of posterior predicted probability. A smattering of credible regression lines is superimposed. Lower panels show marginal posterior distribution on parameters. *Least-squares estimate treating data as metric:* $\beta_0 = 3.425$ (SE=0.012), $\beta_1 = 3.82e-6$ (SE=3.39e-7), $\sigma = 0.847$. Notice that the slope is estimated to be smaller by least-squares estimation in this case, and predictive probabilities are different than Bayesian.

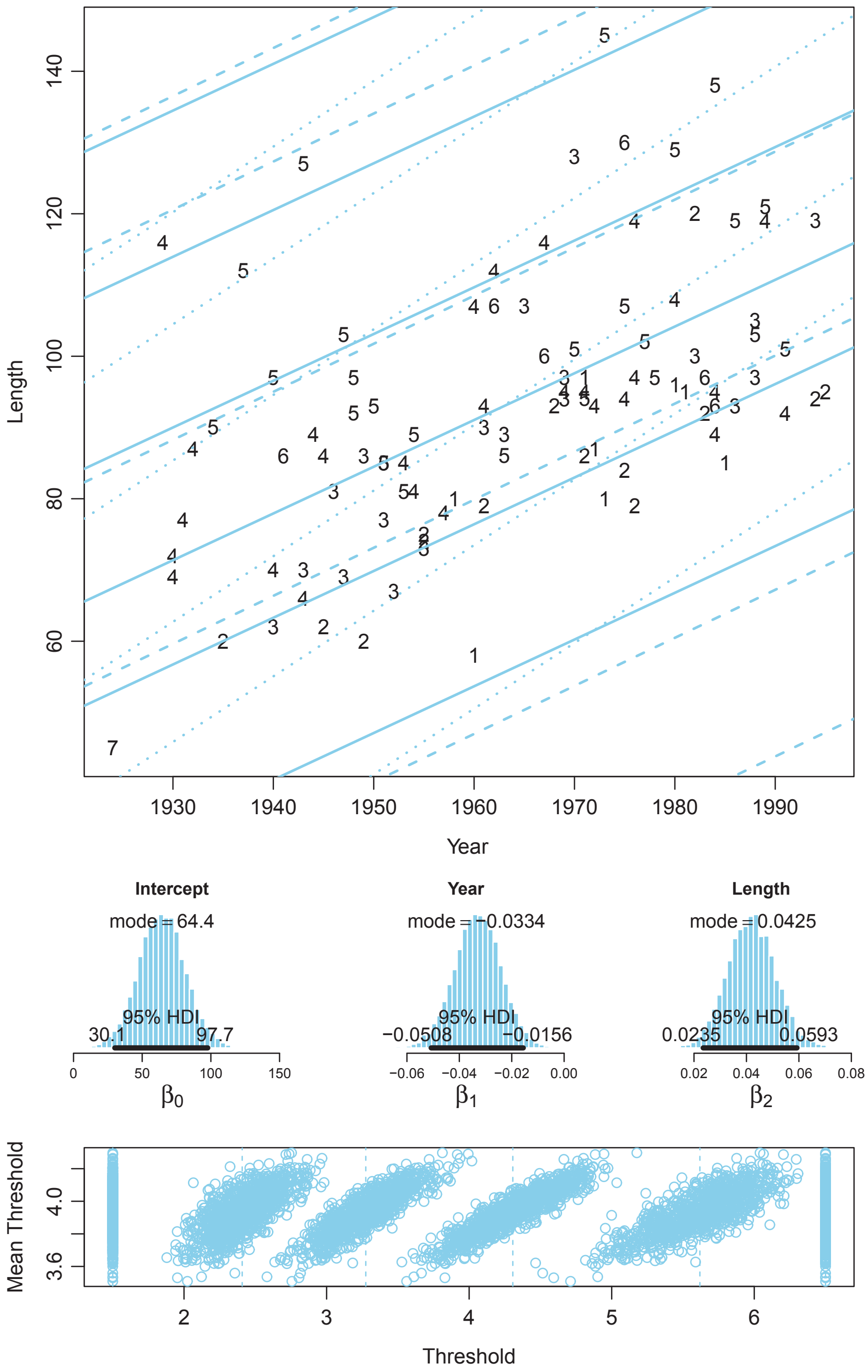

Figure 23.9: Analysis of movie rating data from Moore (2006). Not shown is the marginal posterior on σ, which has a modal value of about 1.25 and an 95% HDI from about 1.0 to 1.5. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

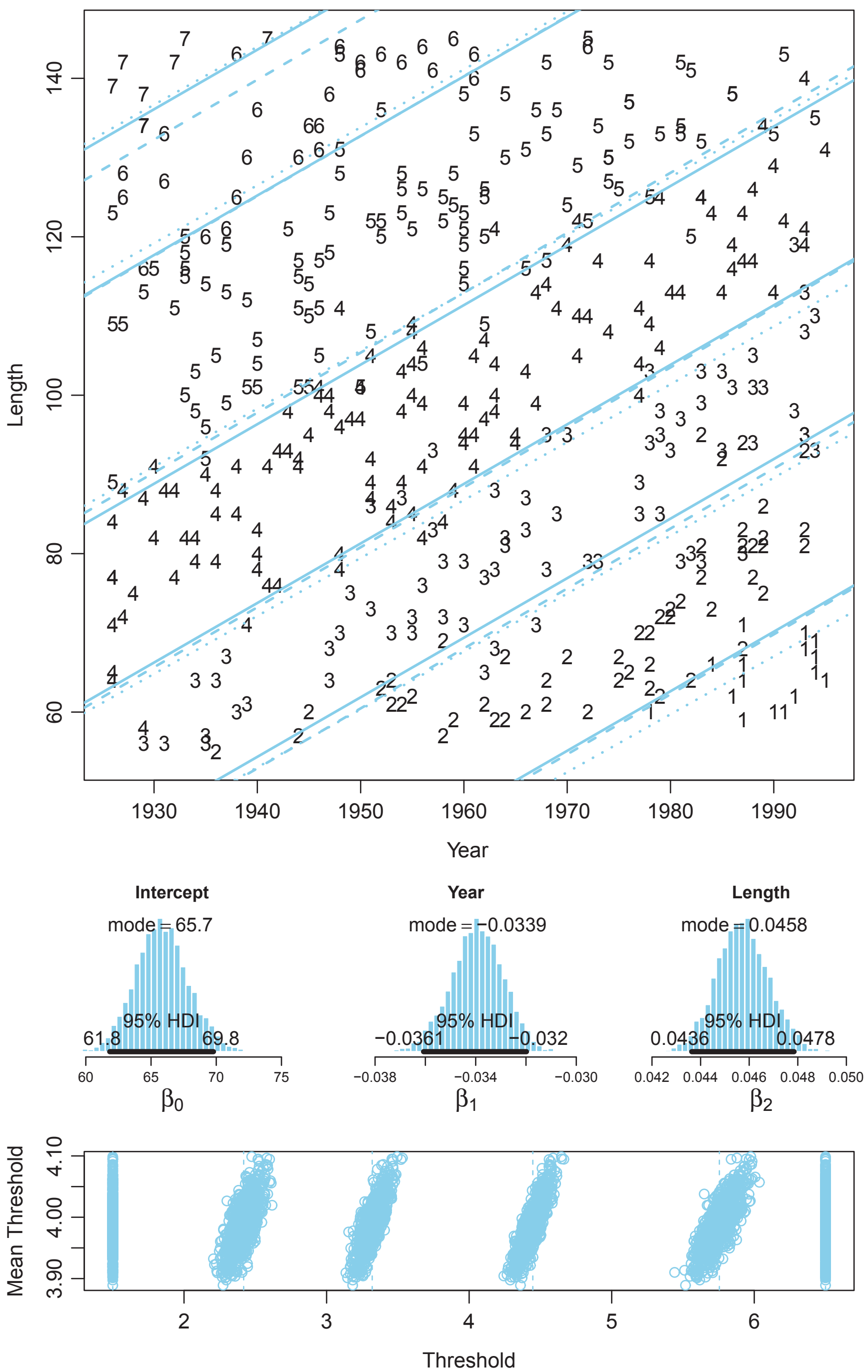

Figure 23.10: Analysis of simulated movie-rating data. These artificial data have very small noise, with generating $\sigma = 0.20$, compared with $\sigma \approx 1.25$ in the actual data of Figure 23.9. (The intercept and slope parameters are set a bit differently than the actual data so that all outcome values are present in the range of the predictors.) Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

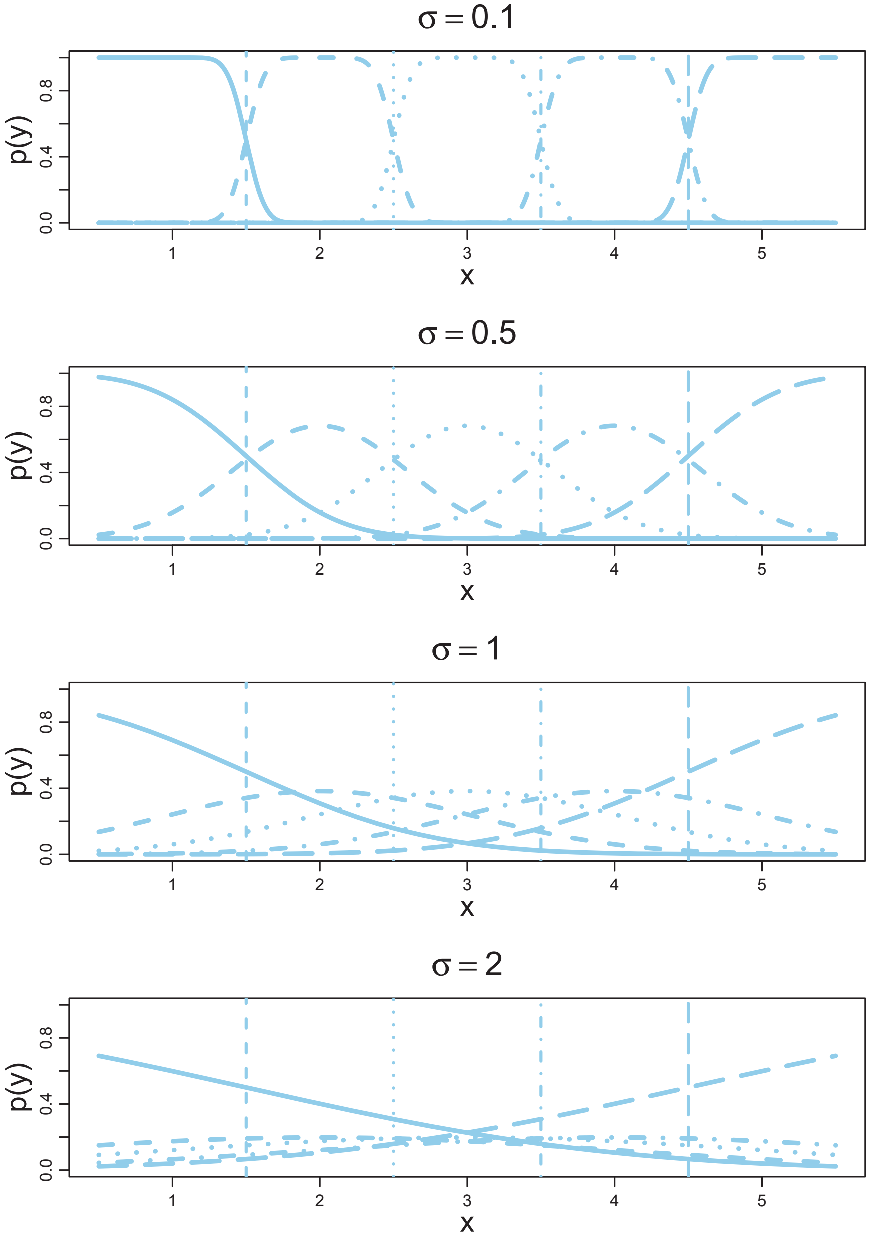

Figure 23.11: The probability of an ordinal response as a function of the predictor x. Thresholds are arbitrarily set at $\theta_1 = 1.5$, $\theta_2 = 2.5$, $\theta_3 = 3.5$, and $\theta_4 = 4.5$. Upper panel is for small noise ($\sigma = 0.1$), and lower panel is for large noise ($\sigma = 2.0$). Note that x values in the data might span only a small range on the x axis.

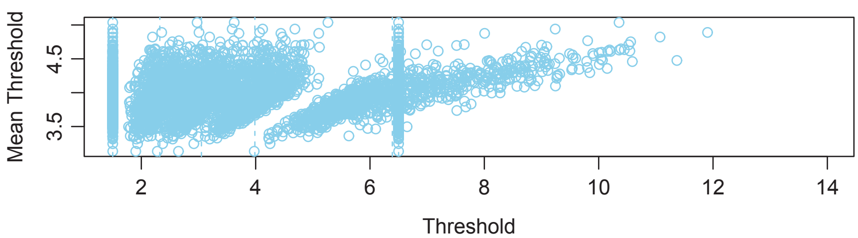

Figure 23.12: For Exercise 23.2 Posterior distribution on thresholds for movie-rating data when the model is a mixture with random guessing. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

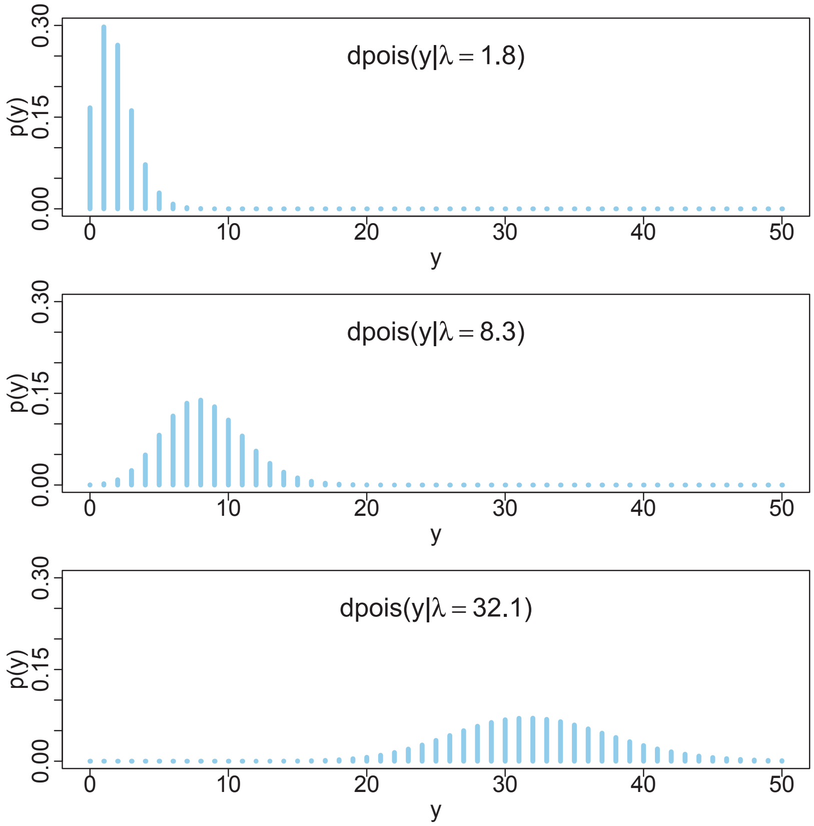

Figure 24.1: Examples of the Poisson distribution. The range of y includes integers from zero to positive infinity. The value of λ is displayed within each panel. The notation "dpois" means the Poisson probability distribution of Equation 24.3.

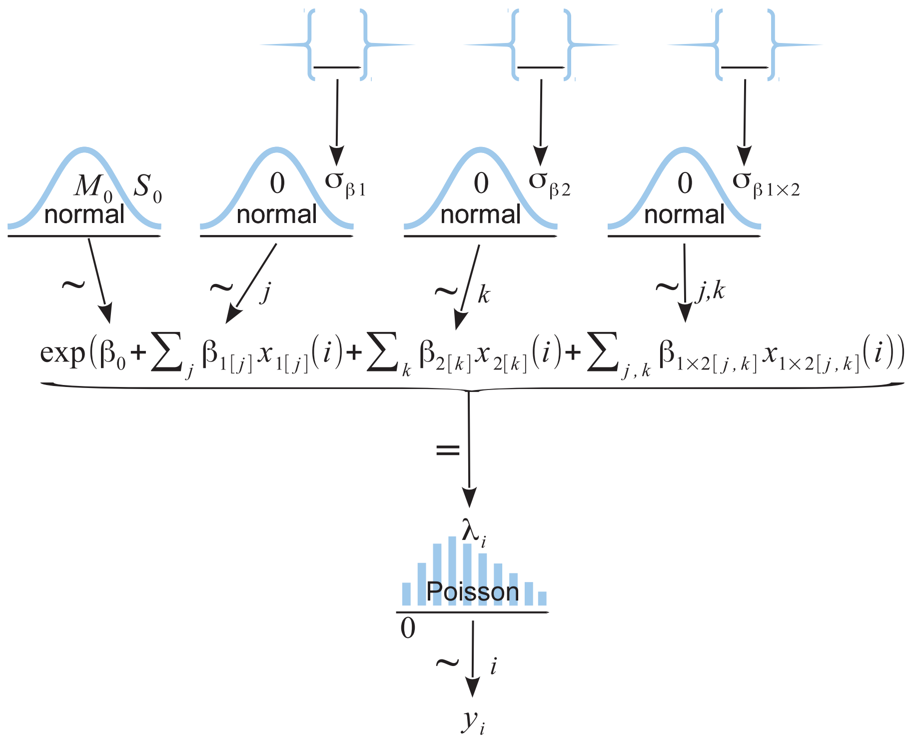

Figure 24.2: Hierarchical dependency diagram of the Poisson exponential model for two nominal predictors. Compare with the diagram for two-factor "ANOVA" in Figure 20.2 (p. 562).

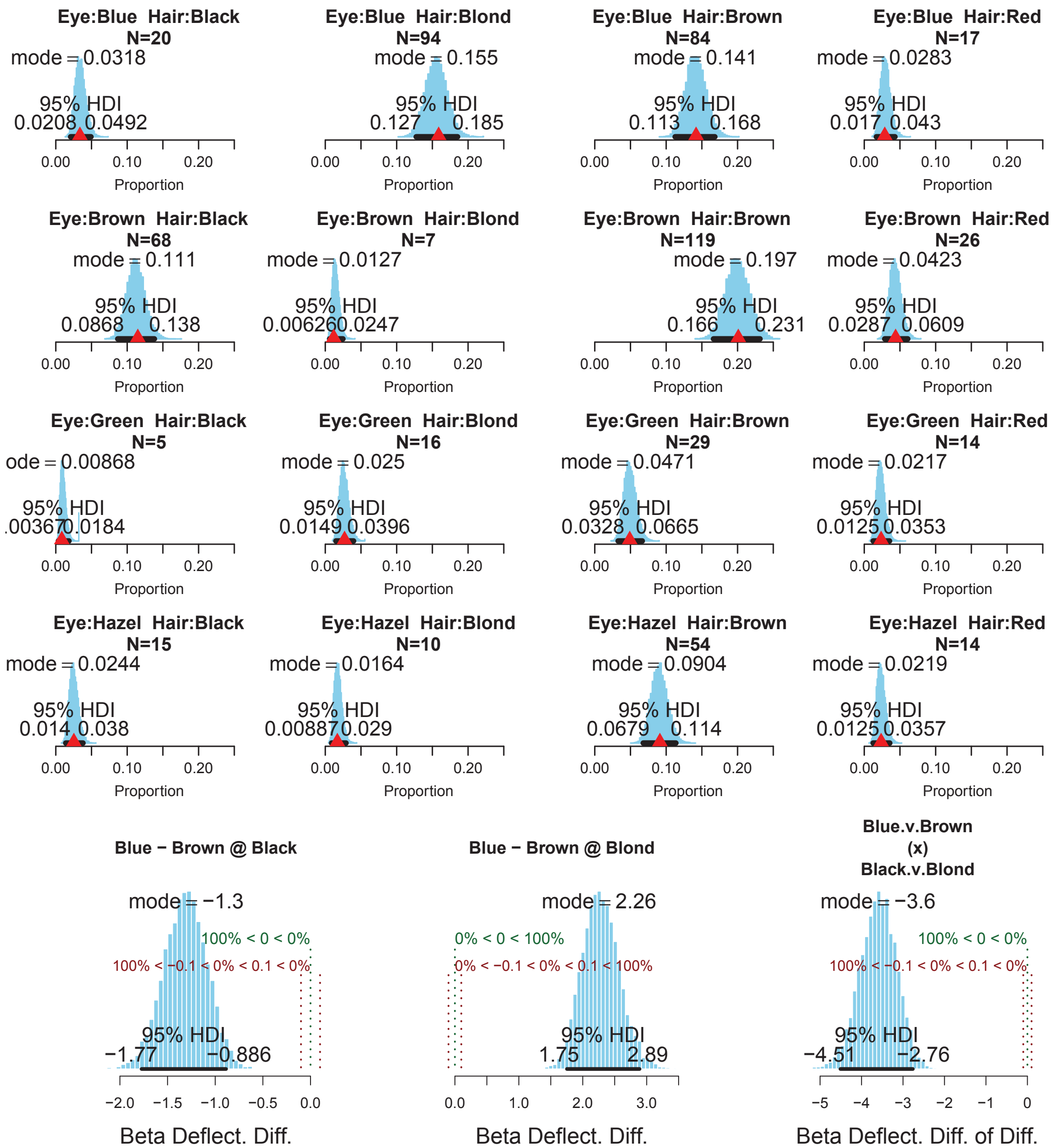

Figure 24.3: *Upper panels:* Posterior distribution of estimated cell proportions for data in Table 24.1. Triangles indicate actual data proportions. *Lower row:* Interaction contrast. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

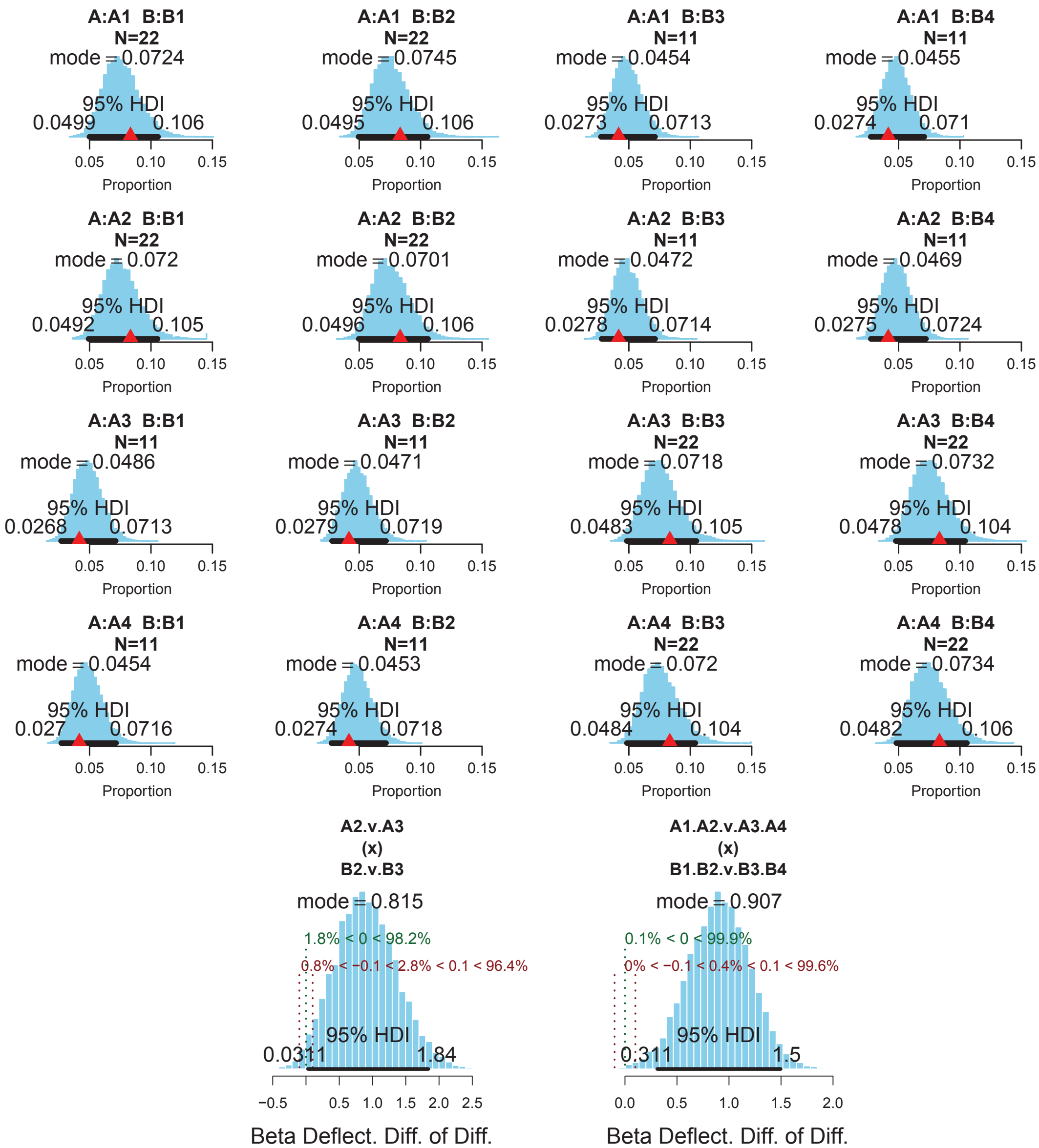

Figure 24.4: *Upper panels:* Posterior distribution of estimated cell proportions. Triangles indicate actual data proportions. *Lower row:* Interaction contrasts. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

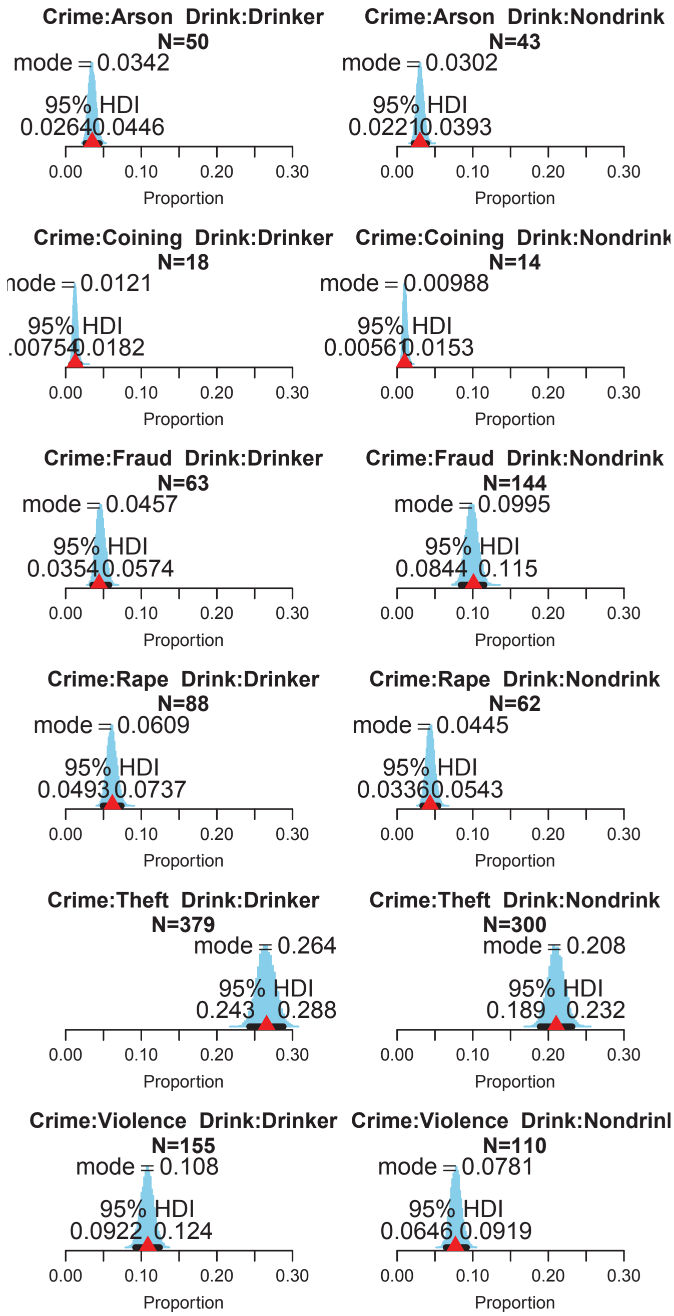

Figure 24.5: For Exercise 24.1. Posterior distribution of estimated cell proportions for crime and drinking data. Triangles indicate actual data proportions. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

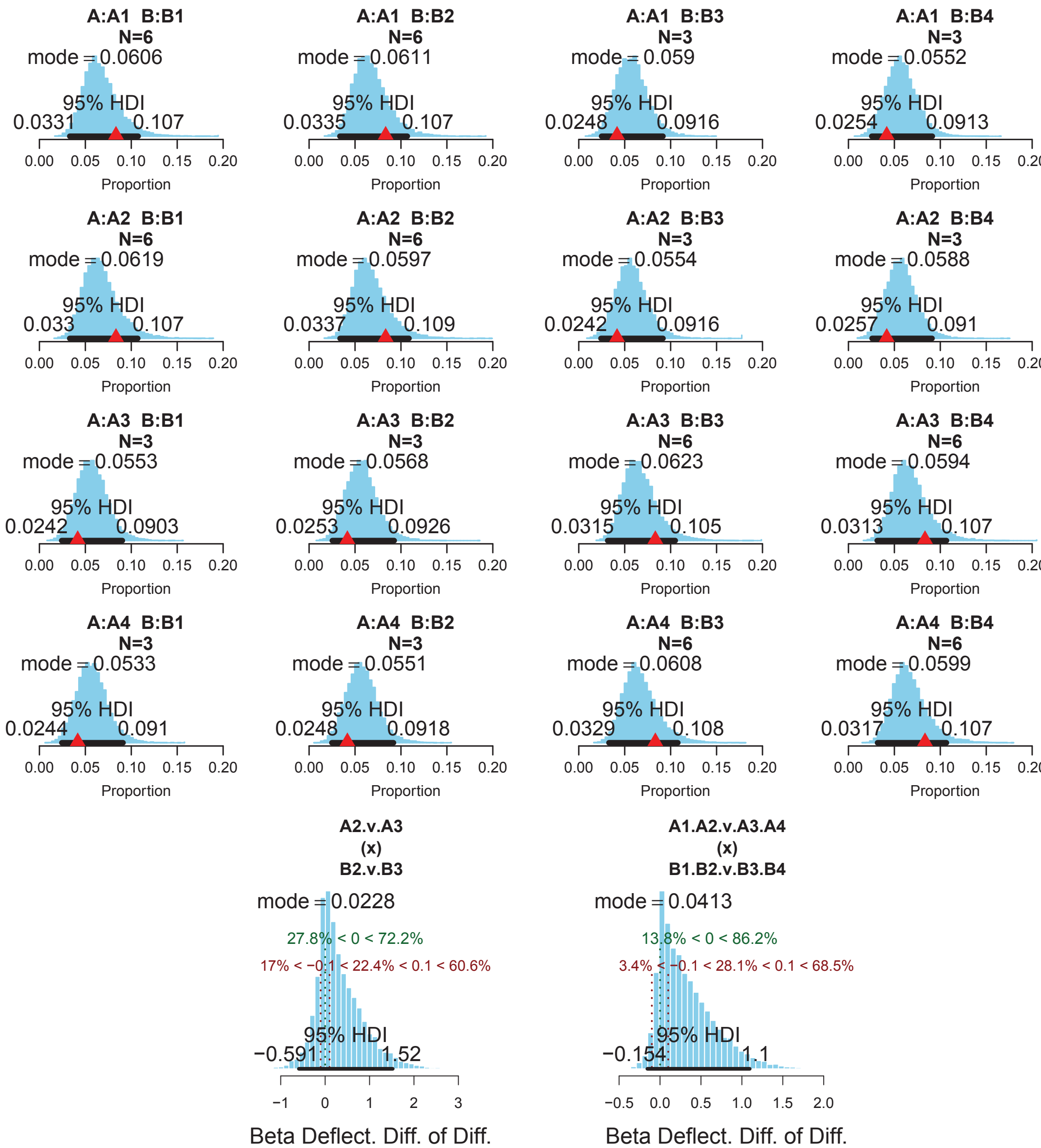

Figure 24.6: For Exercise 24.2 regarding shrinkage. *Upper panels:* Posterior distribution of estimated cell proportions. Triangles indicate actual data proportions. *Lower row:* Interaction contrast. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

Table 24.1: Counts of combinations of hair color and eye color. Data adapted from Snee (1974).

Eye Color	Hair Color: Black	Blond	Brown	Red	Marginal (Eye Color)
Blue	20	94	84	17	215
Brown	68	7	119	26	220
Green	5	16	29	14	64
Hazel	15	10	54	14	93
Marginal (Hair Color)	108	127	286	71	592

Table 24.2: Example of exponentiated linear model with zero interaction. Margins show the values of the β's, and cells show $\lambda_{r,c} = \exp(\beta_0 + \beta_r + \beta_c)$ as in Equation 24.1. Notice that every row has the same relative probabilities, namely, 10, 100, 1. In other words, the row and column attributes are independent. Notice also that the row and column deflections sum to zero, as required by Equation 20.2 (p. 538).

$\beta_0 = 4.60517$	$\beta_{c=1} = 0$	$\beta_{c=2} = 2.30259$	$\beta_{c=3} = -2.30259$
$\beta_{r=1} = 0$	100	1000	10
$\beta_{r=2} = 2.30259$	1000	10000	100
$\beta_{r=3} = -2.30259$	10	100	1

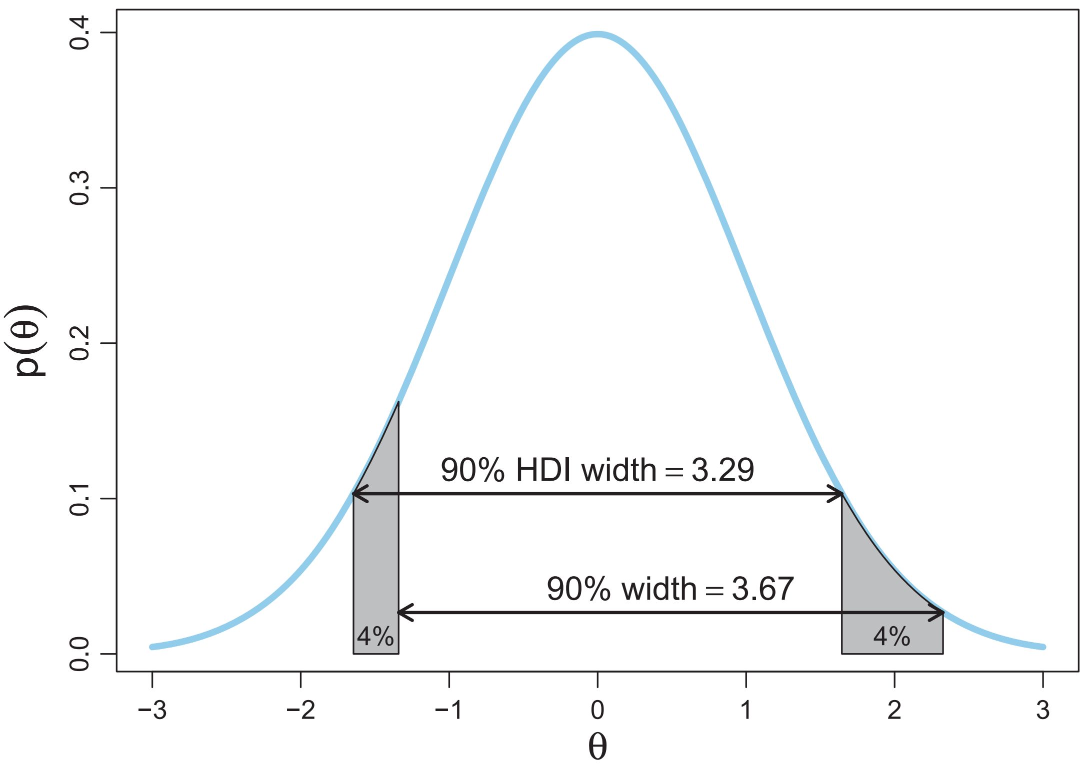

Figure 25.1: For a unimodal distribution, the HDI is the narrowest interval of that mass. This figure shows the 90% HDI and another interval that has 90% mass. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

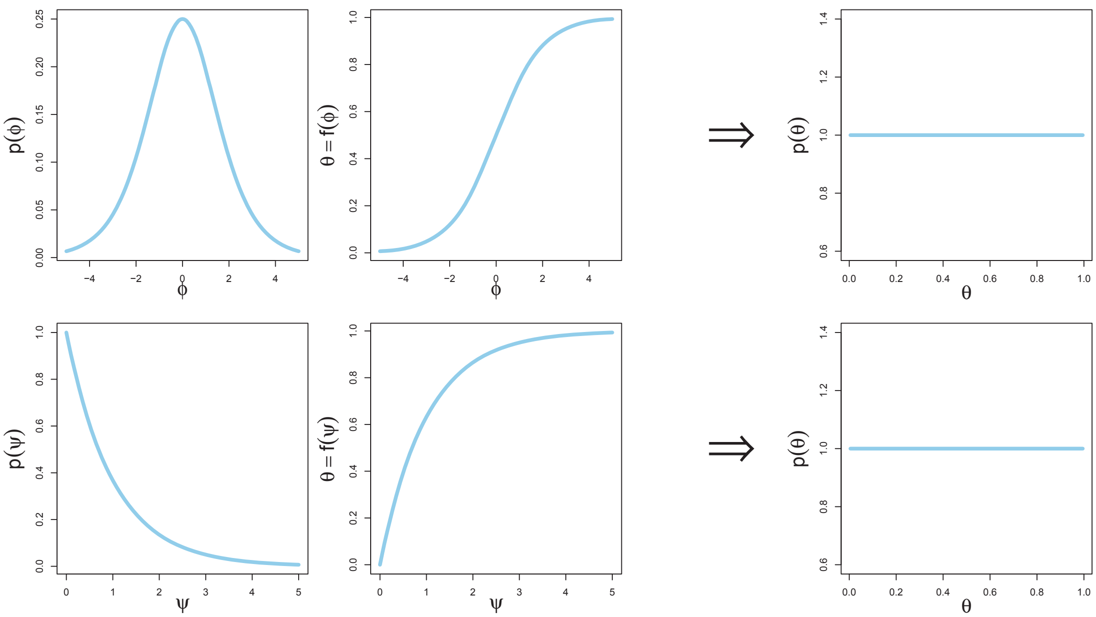

Figure 25.2: *Top row:* A reparameterization that maps a peaked distribution over $\phi \in [-\infty, +\infty]$ to a uniform distribution over $\theta \in [0, 1]$. *Bottom row:* A reparameterization that maps a descending exponential distribution over $\psi \in [0, +\infty]$ to a uniform distribution over $\theta \in [0, 1]$.

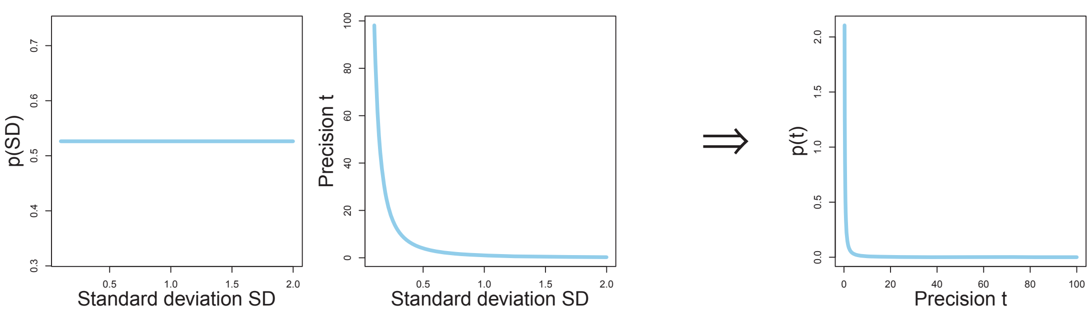

Figure 25.3: A uniform distribution on standard deviation, mapped to the corresponding distribution on precision. Copyright © Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. 2nd Edition.* Academic Press / Elsevier.

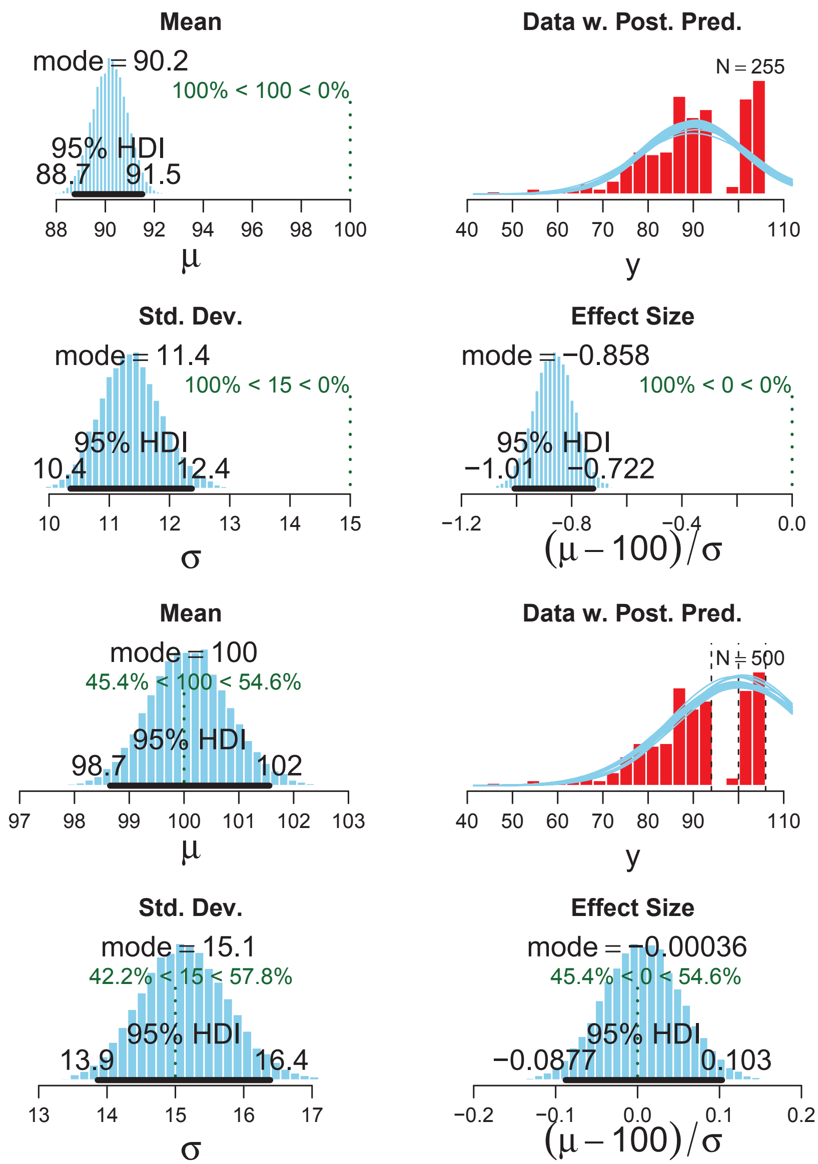

Figure 25.4: Data randomly generated from a normal distribution with $\mu = 100$ and $\sigma = 15$, then censored between 94 and 100 and above 106. *Upper quartet:* Censored data omitted from analysis; parameter estimates are too small. *Lower quartet:* Censored data imputed in known bins; parameter estimates are accurate.