BDA HM6

Piero Birello

February 2025

10.1

The problem can be expressed as:

$$p(m|D) = \frac{p(D|m)p(m)}{\sum_m p(D|m)p(m)} \tag{1}$$

where:

$$p(D|m) = \frac{B(z+a, N-z+b)}{B(a,b)} \tag{2}$$

When $w1 = 0.25$, $w2 = 0.75$, $k = 6$, $p(m = 1) = p(m = 2) = 0.5$, $z = 7$, $N = 10$ we obtain $p(m = 1|D) = 0.25$ and $p(m = 2|D) = 0.75$. If we change $k = 202$ we obtain $p(m = 1|D) = 0.016$ and $p(m = 2|D) = 0.984$. The results are completely different based on our choice for the prior distribution for the coins' biases. Actually, in the second case the concentration parameter is high, meaning we have high confidence that the coins' biases are near the values $w1$ and $w2$ respectively. Since the number of heads is high (7/10) we will prefer the model with $w2 = 0.75$. On the other hand, when $k = 6$ there's high uncertainty on model parameters, and we cannot be that sure when selecting the best model.

10.2

Figure 1 and Figure 2 show the prior and posterior probabilities for the two proposed models. The prior was generated inserting empty data and sampling via MCMC. The collapsed posterior across models is shown in Figure 3. This is very similar to the posterior according to model 2, since $p(m = 2|D) = 0.822$. However, it has a heavier left tail taking into account the posterior according to model 1.

If we attempt to solve numerically the previous exrcise using this code, we get the same exact values for $k = 6$, i.e., $p(m = 2|D) = 0.75$. When $k = 202$, we can actually never sample from model 1, so that we get $p(m = 2|D) = 1$.

10.3

In this exercise we run `Jags-Ydich-Xnom1subj-MbernBetaModelCompPseudoPrior.R` first setting pseudo-priors to the true values of priors (Figure 4, Figure 5), then to values proposed in the book (Figure 6, Figure 7). From the diagnostics, we see that the effective sample size is way larger in the second case, and the autocorrelation function decays faster. This demonstrates that sampling is being correctly performed. As a result, posterior distributions for θ_1 and θ_2 are consistent under $m = 1$ and $m = 2$, which is not the case when pseduo-priors are set to true values.

In Figure 8, Figure 9 we show results when we set the pseudo-prior distributions to be broad. In this case parameter values sampled according to the other model (values of θ_2 when $m = 1$ and viceversa) follow an

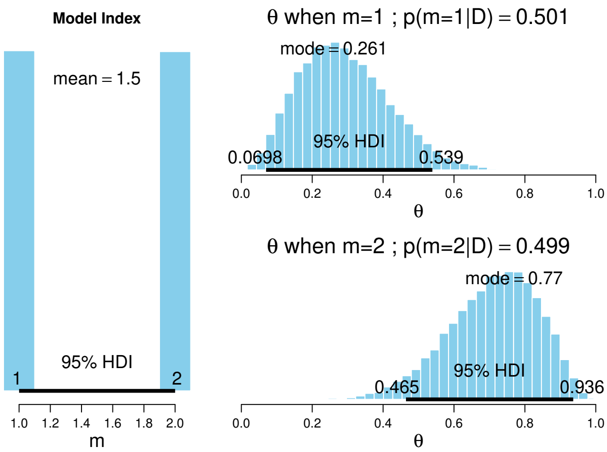

Figure 1: Exercise 10.2 a, prior of the two models.

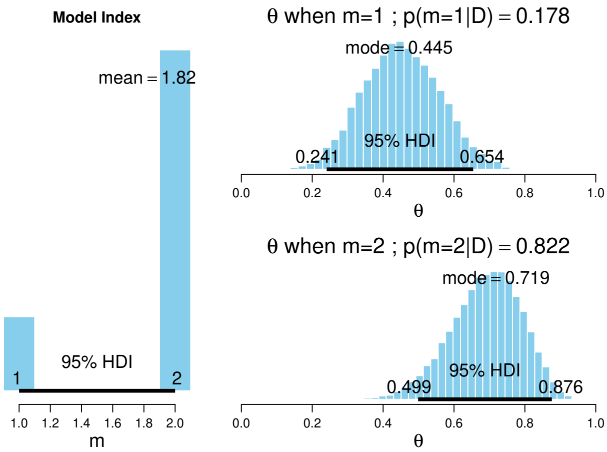

Figure 2: Exercise 10.2 a, posterior of the two models.

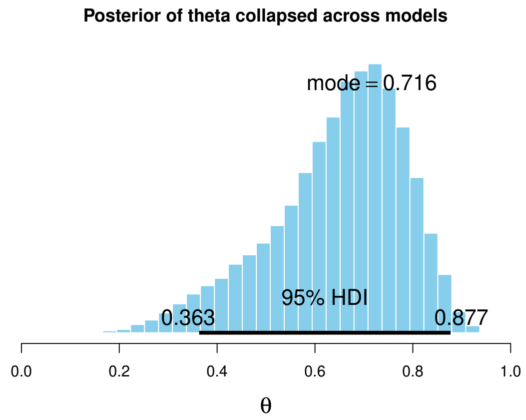

Figure 3: Exercise 10.2, posterior collapsed across models.

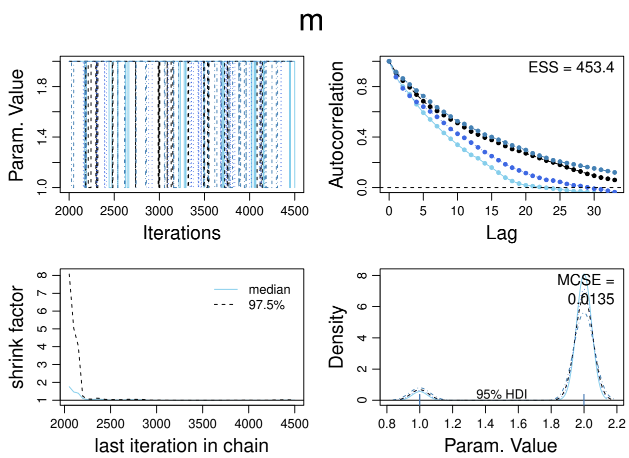

Figure 4: Exercise 10.3, pseudo-priors set to true priors, diagnostic.

almost flat distribution over the entire support. This is suboptimal with respect to choosing pseudo-priors mirroring the posterior distribution of parameters. Still, the sampling performance is better than the one when using no pseudo-priors at all, as we can see from the ESS and faster decaying autocorrelation of the Monte Carlo chain.

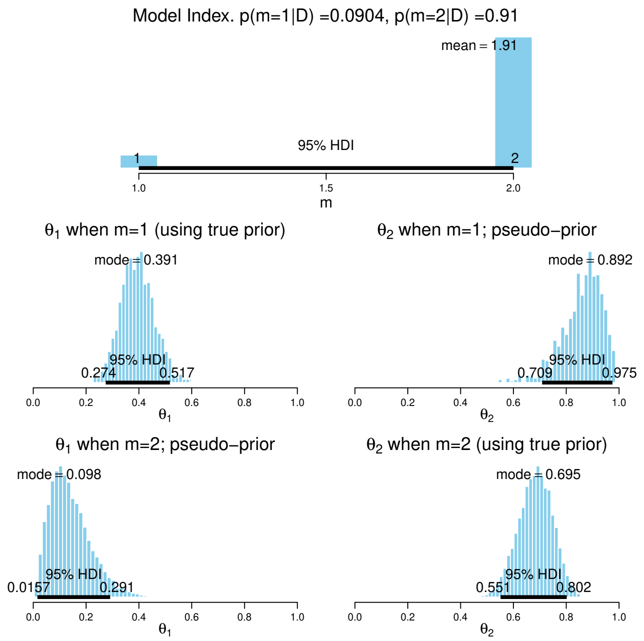

Figure 5: Exercise 10.3, pseudo-priors set to true priors, posterior.

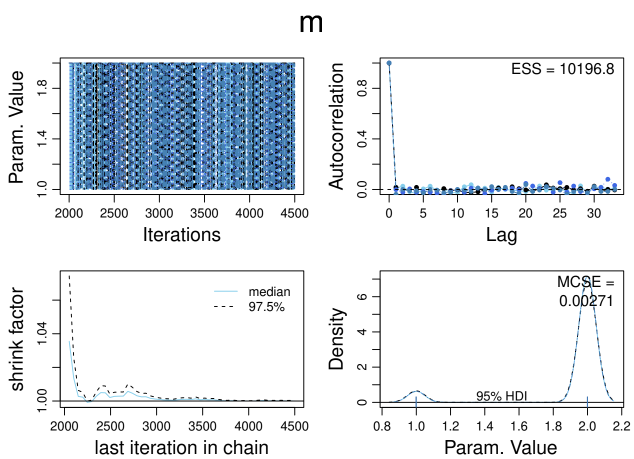

Figure 6: Exercise 10.3, pseudo-priors set different from true priors, diagnostic.

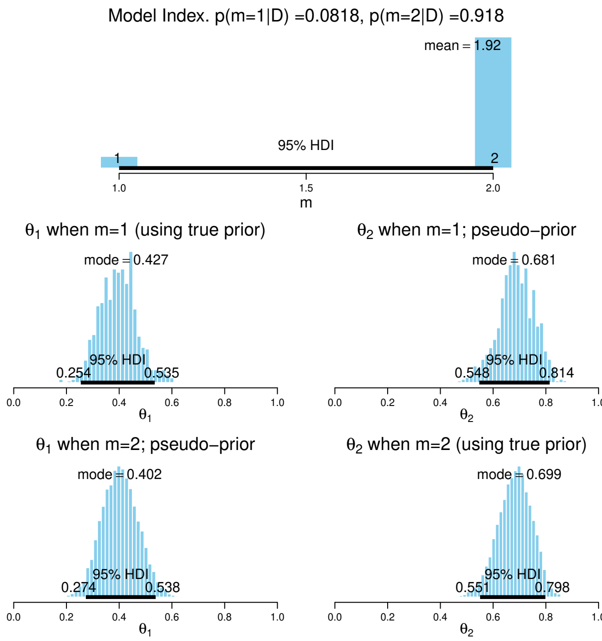

Figure 7: Exercise 10.3, pseudo-priors set different from true priors, posterior.

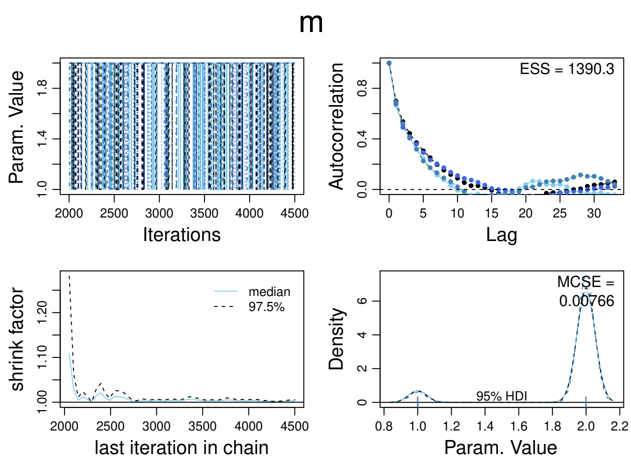

Figure 8: Exercise 10.3, pseudo-prior distribution is set broad, diagnostic.

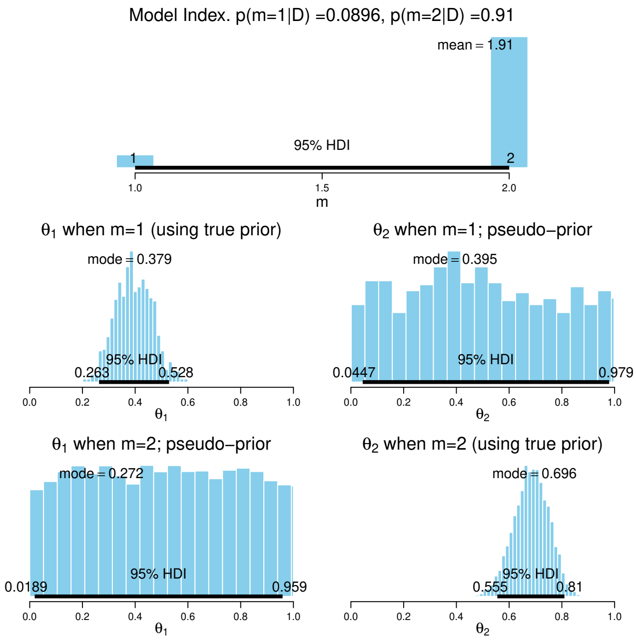

Figure 9: Exercise 10.3, pseudo-prior distribution is set broad, posterior.