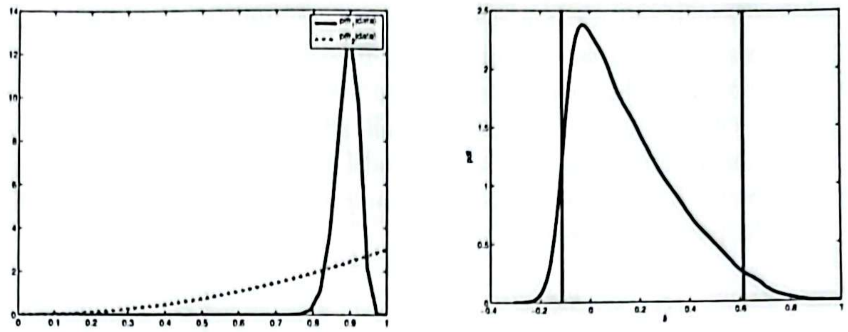

(a) (b)

Figure 5.5 (a) Exact posteriors $p(\theta_i|\mathcal{D}_i)$. (b) Monte Carlo approximation to $p(\delta|\mathcal{D})$. We use kernel density estimation to get a smooth plot. The vertical lines enclose the 95% central interval. Figure generated by `amazonSellerDemo`,

On the face of it, you should pick seller 2, but we cannot be very confident that seller 2 is better since it has had so few reviews. In this section, we sketch a Bayesian analysis of this problem. Similar methodology can be used to compare rates or proportions across groups for a variety of other settings.

Let θ_1 and θ_2 be the unknown reliabilities of the two sellers. Since we don't know much about them, we'll endow them both with uniform priors, $\theta_i \sim \text{Beta}(1, 1)$. The posteriors are $p(\theta_1|\mathcal{D}_1) = \text{Beta}(91, 11)$ and $p(\theta_2|\mathcal{D}_2) = \text{Beta}(3, 1)$.

We want to compute $p(\theta_1 > \theta_2|\mathcal{D})$. For convenience, let us define $\delta = \theta_1 - \theta_2$ as the difference in the rates. (Alternatively we might want to work in terms of the log-odds ratio.) We can compute the desired quantity using numerical integration:

$$\begin{aligned} p(\delta > 0|\mathcal{D}) &= \int_0^1 \int_0^1 \mathbb{I}(\theta_1 > \theta_2)\text{Beta}(\theta_1|y_1 + 1, N_1 - y_1 + 1) \\ & \quad \text{Beta}(\theta_2|y_2 + 1, N_2 - y_2 + 1)d\theta_1 d\theta_2 \end{aligned} \quad (5.11)$$

We find $p(\delta > 0|\mathcal{D}) = 0.710$, which means you are better off buying from seller 1! See `amazonSellerDemo` for the code. (It is also possible to solve the integral analytically (Cook 2005).)

A simpler way to solve the problem is to approximate the posterior $p(\delta|\mathcal{D})$ by Monte Carlo sampling. This is easy, since θ_1 and θ_2 are independent in the posterior, and both have beta distributions, which can be sampled from using standard methods. The distributions $p(\theta_i|\mathcal{D}_i)$ are shown in Figure 5.5(a), and a MC approximation to $p(\delta|\mathcal{D})$, together with a 95% HPD, is shown Figure 5.5(b). An MC approximation to $p(\delta > 0|\mathcal{D})$ is obtained by counting the fraction of samples where $\theta_1 > \theta_2$; this turns out to be 0.718, which is very close to the exact value. (See `amazonSellerDemo` for the code.)

5.3 Bayesian model selection

In Figure 1.18, we saw that using too high a degree polynomial results in overfitting, and using too low a degree results in underfitting. Similarly, in Figure 7.8(a), we saw that using too small