

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \text{Pr}(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\text{Pr}(\theta < 0.8)$ is very small.

We choose $\theta \sim \text{Be}(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1-\theta)^4}{B(77, 5)}, & 0 < \theta < 1 \\ &= 128107980\theta^{76}(1-\theta)^4, & 0 < \theta < 1.\end{aligned}\tag{2.7}$$

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \Pr(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\Pr(\theta < 0.8)$ is very small.

We choose $\theta \sim Be(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1-\theta)^4}{B(77, 5)}, & 0 < \theta < 1 \\ &= 128107980\theta^{76}(1-\theta)^4, & 0 < \theta < 1.\end{aligned}\tag{2.7}$$

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \Pr(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\Pr(\theta < 0.8)$ is very small.

We choose $\theta \sim Be(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1-\theta)^4}{B(77, 5)}, & 0 < \theta < 1 \\ &= 128107980\theta^{76}(1-\theta)^4, & 0 < \theta < 1.\end{aligned}\tag{2.7}$$

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \Pr(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\Pr(\theta < 0.8)$ is very small.

We choose $\theta \sim Be(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1-\theta)^4}{B(77, 5)}, \quad 0 < \theta < 1 \\ &= 128107980\theta^{76}(1-\theta)^4, \quad 0 < \theta < 1.\end{aligned}\tag{2.7}$$

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \Pr(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\Pr(\theta < 0.8)$ is very small.

We choose $\theta \sim Be(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1-\theta)^4}{B(77, 5)}, & 0 < \theta < 1 \\ &= 128107980\theta^{76}(1-\theta)^4, & 0 < \theta < 1.\end{aligned}\tag{2.7}$$

Example 2.2

Consider an experiment to determine how good a music expert is at distinguishing between pages from Haydn and Mozart scores. Let $\theta = \Pr(\text{correct choice})$.

Suppose that, before conducting the experiment, we have been told that the expert is very competent.

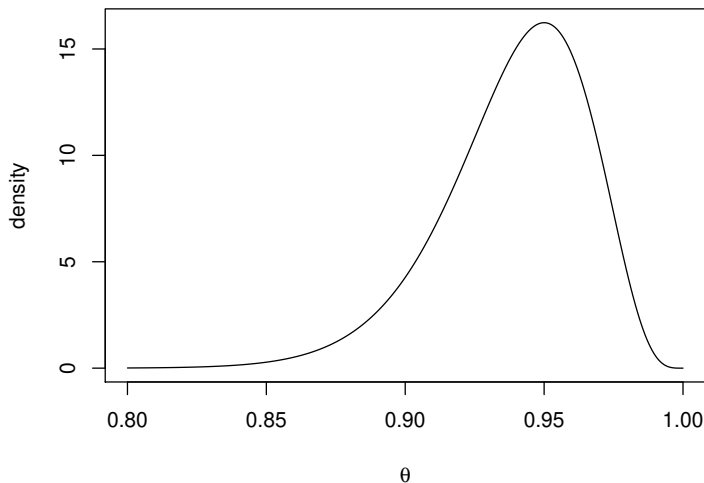
In fact, it is suggested that we should have a prior distribution which has a mode around $\theta = 0.95$ and for which $\Pr(\theta < 0.8)$ is very small.

We choose $\theta \sim Be(77, 5)$ (see Example 3.2, Chapter 3), with probability density function

$$\begin{aligned}\pi(\theta) &= \frac{\theta^{76}(1 - \theta)^4}{B(77, 5)}, & 0 < \theta < 1 \\ &= 128107980\theta^{76}(1 - \theta)^4, & 0 < \theta < 1.\end{aligned}\tag{2.7}$$

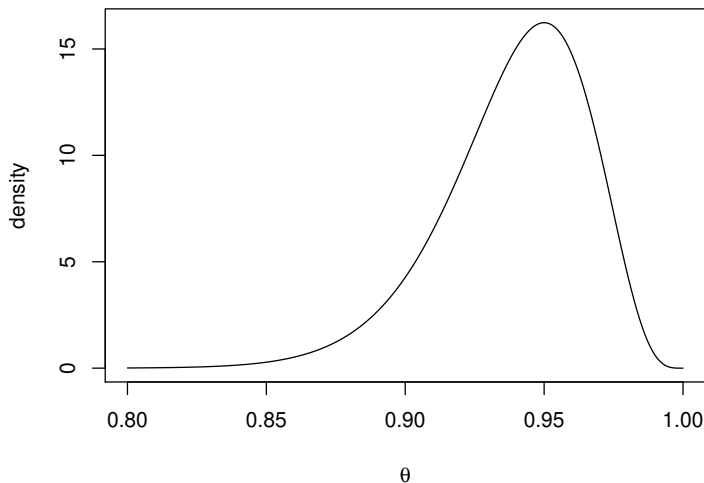
Example 2.2

A graph of this prior density is given in Figure 2.4.



Example 2.2

A graph of this prior density is given in Figure 2.4.



Example 2.2

In the experiment, the music expert makes the correct choice 9 out of 10 times.

Determine the posterior distribution for θ given this information.

Example 2.2

In the experiment, the music expert makes the correct choice 9 out of 10 times.

Determine the posterior distribution for θ given this information.

Solution to Example 2.2 (1/2)

We have an observation on the random variable $X|\theta \sim \text{Bin}(10, \theta)$.

This gives a likelihood function of

$$\begin{aligned} f(x = 9|\theta) &= {}^{10}C_9 \theta^9 (1 - \theta)^1 \\ &= 10\theta^9(1 - \theta) \end{aligned} \tag{2.8}$$

which favours values of θ near its maximum $\theta = 0.9$.

We combine these two sources of information using Bayes' Theorem.

Solution to Example 2.2 (1/2)

We have an observation on the random variable $X|\theta \sim \text{Bin}(10, \theta)$.

This gives a likelihood function of

$$\begin{aligned} f(x = 9|\theta) &= {}^{10}\text{C}_9 \theta^9 (1 - \theta)^1 \\ &= 10\theta^9 (1 - \theta) \end{aligned} \tag{2.8}$$

which favours values of θ near its maximum $\theta = 0.9$.

We combine these two sources of information using Bayes' Theorem.

Solution to Example 2.2 (1/2)

We have an observation on the random variable $X|\theta \sim \text{Bin}(10, \theta)$.

This gives a likelihood function of

$$\begin{aligned} f(x = 9|\theta) &= {}^{10}C_9 \theta^9 (1 - \theta)^1 \\ &= 10\theta^9 (1 - \theta) \end{aligned} \tag{2.8}$$

which favours values of θ near its maximum $\theta = 0.9$.

We combine these two sources of information using Bayes' Theorem.

Solution to Example 2.2 (1/2)

We have an observation on the random variable $X|\theta \sim \text{Bin}(10, \theta)$.

This gives a likelihood function of

$$\begin{aligned} f(x = 9|\theta) &= {}^{10}\text{C}_9 \theta^9 (1 - \theta)^1 \\ &= 10\theta^9 (1 - \theta) \end{aligned} \tag{2.8}$$

which favours values of θ near its maximum $\theta = 0.9$.

We combine these two sources of information using Bayes' Theorem.

Solution to Example 2.2 (2/2)

The posterior density function is

$$\pi(\theta|x = 9) \propto \pi(\theta)f(x = 9|\theta)$$

$$\propto \frac{\theta^{76}(1 - \theta)^4}{B(77, 5)} \times 10 \theta^9(1 - \theta), \quad 0 < \theta < 1$$

$$= 128107980 \theta^{76}(1 - \theta)^4 \times 10 \theta^9(1 - \theta), \quad 0 < \theta < 1$$

$$= k \theta^{85}(1 - \theta)^5, \quad 0 < \theta < 1. \quad (2.9)$$

We can recognise this density function as one from the Beta family. In fact, the posterior distribution is $\theta|x = 9 \sim Be(86, 6)$.

Solution to Example 2.2 (2/2)

The posterior density function is

$$\pi(\theta|x = 9) \propto \pi(\theta)f(x = 9|\theta)$$

$$\propto \frac{\theta^{76}(1 - \theta)^4}{B(77, 5)} \times 10 \theta^9(1 - \theta), \quad 0 < \theta < 1$$

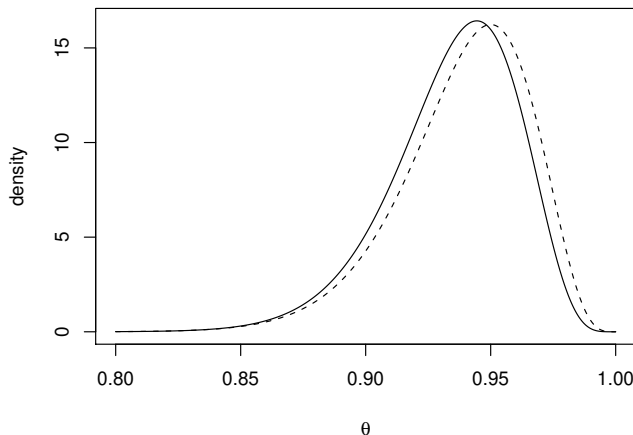
$$= 128107980 \theta^{76}(1 - \theta)^4 \times 10 \theta^9(1 - \theta), \quad 0 < \theta < 1$$

$$= k \theta^{85}(1 - \theta)^5, \quad 0 < \theta < 1. \quad (2.9)$$

We can recognise this density function as one from the Beta family. In fact, the posterior distribution is $\theta|x = 9 \sim Be(86, 6)$.

Example 2.2: Summary

The changes in our beliefs about θ are described by the prior and posterior distributions shown in Figure 2.5 and summarised in Table 2.1.



Example 2.2: Summary

	Prior (2.7)	Likelihood (2.8)	Posterior (2.9)
$Mode(\theta)$	0.950	0.900	0.944
$E(\theta)$	0.939	—	0.935
$SD(\theta)$	0.0263	—	0.0256

Notice that, having observed only a 90% success rate in the experiment, the posterior mode and mean are smaller than their prior values.

Also, the experiment has largely confirmed our ideas about θ , with the uncertainty about θ being only very slightly reduced.

Example 2.2: Summary

	Prior (2.7)	Likelihood (2.8)	Posterior (2.9)
$Mode(\theta)$	0.950	0.900	0.944
$E(\theta)$	0.939	—	0.935
$SD(\theta)$	0.0263	—	0.0256

Notice that, having observed only a 90% success rate in the experiment, the posterior mode and mean are smaller than their prior values.

Also, the experiment has largely confirmed our ideas about θ , with the uncertainty about θ being only very slightly reduced.

Example 2.2: Summary

	Prior (2.7)	Likelihood (2.8)	Posterior (2.9)
$Mode(\theta)$	0.950	0.900	0.944
$E(\theta)$	0.939	—	0.935
$SD(\theta)$	0.0263	—	0.0256

Notice that, having observed only a 90% success rate in the experiment, the posterior mode and mean are smaller than their prior values.

Also, the experiment has largely confirmed our ideas about θ , with the uncertainty about θ being only very slightly reduced.