5.1. $p(☹) = 0.001$

$p(T=+|\theta=☹) = 0.99$

$p(T=+|\theta=☺) = 0.05$

The subject is tested once and diagnosed positive, then re-tested and diagnosed negative

$$p(\theta=☹|T=-) = \frac{p(T=-|\theta=☹)\,p(\theta=☹|T=+)}{\sum_\theta p(T=-|\theta)\,p(\theta)}$$

$$= \frac{p(T=-|\theta=☹)}{\sum_\theta p(T=-|\theta)p(\theta)}\left[\frac{p(T=+|\theta=☹)\,p(\theta=☹)}{\sum_\theta p(T=+|\theta)\,p(\theta)}\right]$$

$$= \frac{0.01 \times 0.99 \times 0.001}{(0.01 \times 0.001 + 0.95 \times 0.999) \times (0.99 \times 0.001 + 0.05 \times 0.999)}$$

$$= 2.0 \times 10^{-4}$$

5.2. A)

	$\theta = ☹$	$\theta = ☺$	
D = +	99	~~54945~~ 4995	~~55044~~ 5094
D = -	1	~~44955~~ 94905	~~44956~~ 94906
	100	99 900	N = 100 000

B) proportion of people who have the disease, given that test is positive:

$$\frac{99}{5094\ \cancel{55044}} = \cancel{1.8 \times 10^{-3}}\ 1.9 \times 10^{-2}$$

Using Bayes rule: $p(\theta=☹|T=+) = \frac{p(T=+|\theta=☹)\,p(\theta=☹)}{p(T=+)} = 1.9\times10^{-2}$

C) N = 10 000 000

99.9% are expected not to have the disease

0.1% are expected to have the disease

N = 10 000 000

10 000 → 9999 → 100

9 990 000 → 499500 → ~~24975~~ 474525

D) Proportion of people who test positive first and negative then and who actually have the disease?

$$\frac{100}{474525+100} = 2.1 \times 10^{-4}$$

Same result as 5.1

5.3 A) $p(\theta = \ddot{\frown} \mid T = +) = \dfrac{p(T=+ \mid \theta = \ddot{\frown})\, p(\theta = \ddot{\frown})}{p(T=+)} = \dfrac{0.01 \times 0.001}{0.01 \times 0.001 + 0.05 \times 0.999} = \cancel{0.019}$ 1.05×10^{-5}

B) $p(\theta = \ddot{\frown} \mid T = +) = \dfrac{p(T=+ \mid \theta = \ddot{\frown})\, p(\theta = \ddot{\frown} \mid T = -)}{\sum_\theta P(T=+ \mid \theta)\, p(\theta)}$

$$= \frac{p(T=+ \mid \theta = \ddot{\frown})}{\sum_\theta p(T=+ \mid \theta)\, p(\theta)} \quad \frac{P(T=- \mid \theta = \ddot{\frown})}{\sum_\theta P(T=- \mid \theta)\, p(\theta)} \quad p(\theta = \ddot{\frown}) = 2.0 \times 10^{-4}$$

We obtain the same expression

5.4. See Latex

6.1 See Latex

6.2 100 randomly sampled people, 58 preferred candidate A, 42 preferred candidate B.

A) Before poll, prior belief was uniform, i.e., Beta (1,1).
After poll, HDI spans from 0.483 to 0.673 (see latex for figure)

B) 100 more people selected, 57 A, 43 B. Now 95 % HDI spans from 0.506 to 0.642.

6.4 We know a coin to be biased to usually come up either heads or tails ~~either~~
we can choose Beta(0.1, 0.1). See latex for posterior after 4 heads out of 5 flips.

6.5 A) You have a strong prior belief that the coin is fair, since it was minted by the government. We can translate this into a prior mode $t = 0.5$ and an effective prior sample size $n = 1000$. Then $a = b = 500$. My predicted probability for head at the 11th flip is the mode of the posterior, $\hat{\theta} = 0.504$. I.e., I am just slightly suspicious that the coin might be unfair. See Latex for figure.

B) Now we have a suspicious coin, that may be biased towards one of the two outcomes, so I will use Beta (0.1, 0.1). 9 heads out of 10. The mode of the posterior is now $\hat{\theta} = 0.988$

BDA HM2

Piero Birello

January 2025

5.4

All examples 0-10 (figs. 1 to 11) show how the interaction between prior and likelihood shapes the posterior distribution for θ, the parameter describing the probability for heads, given a certain number of head outcomes z over a total of N coin flips. In examples 0 and 1 discrete priors are adopted, while examples 2-10 use continuous ones. Examples 0 and 1 are almost equivalent, if not for the number of points the prior distribution is defined on. Example 2 uses a flat prior, so that the posterior is proportional to the likelihood. In example 3 the prior is composed by two delta functions. The effect of the likelihood consists in selecting the correct value among the two possible for the parameter. In examples 4-6 we see how different priors determine the posterior given the same outcome $z = 1, N = 4$. In 7-9 we see the same result in frequency, but with higher number of trials, i.e., $z = 10, N = 40$: the posterior now resembles the likelihood more. At last, example 10 shows what happens using a prior composed by two triangular peaks (similar to a Gaussian mixture distribution).

6.1

figs. 12 to 14 show progressive updates of the posterior distribution when extracting H, H, T starting from a $Beta(4, 4)$, using the code from $BernBeta.R$. In fig. 15 we show the final posterior distribution when extracting T, H, H. Due to data order invariance, the final posterior distribution is identical.

6.2

We use $BernBeta.R$. Results are shown in figs. 16 and 17.

6.4

We use $BernBeta.R$. Results are shown in fig. 18

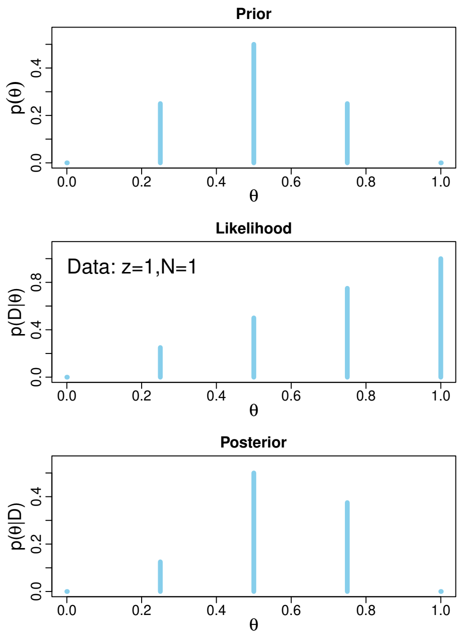

Figure 1: Example 0

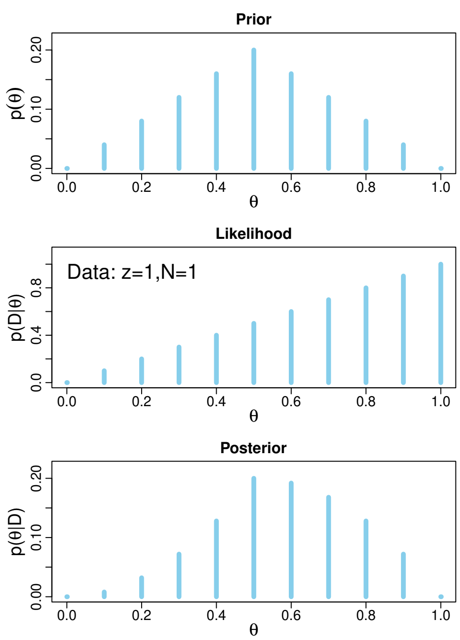

Figure 2: Example 1

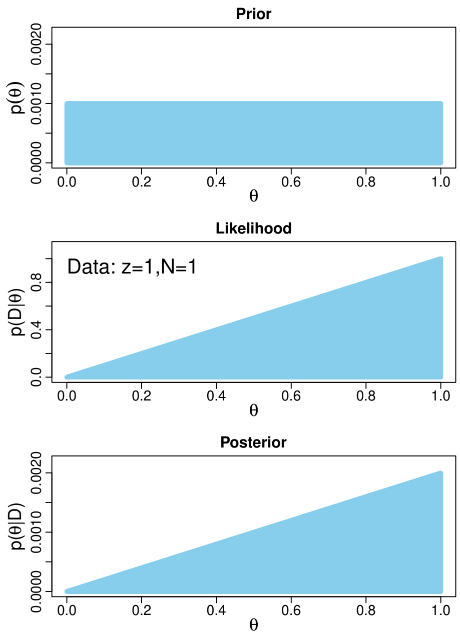

Figure 3: Example 2

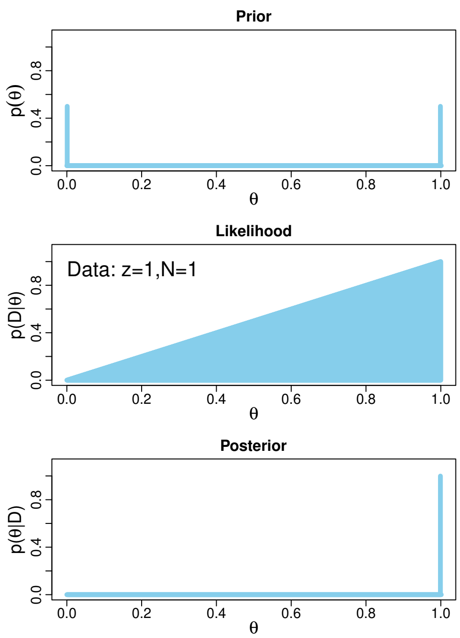

Figure 4: Example 3

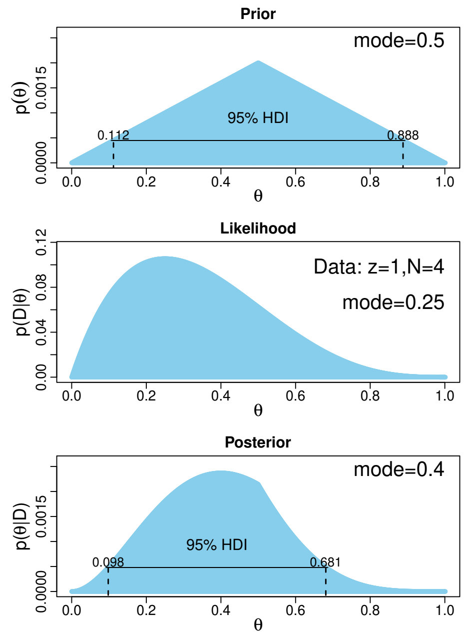

Figure 5: Example 4

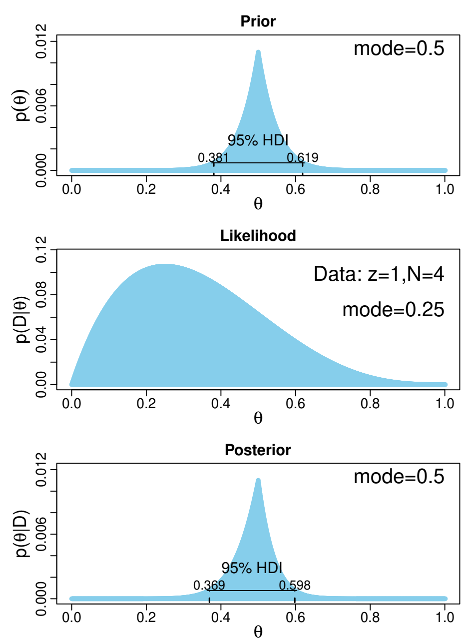

Figure 6: Example 5

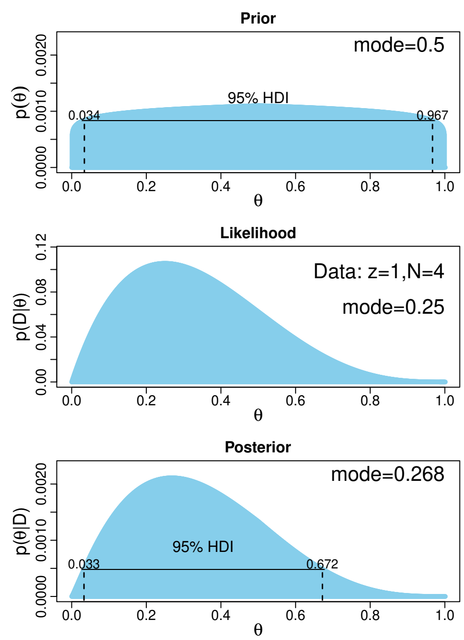

Figure 7: Example 6

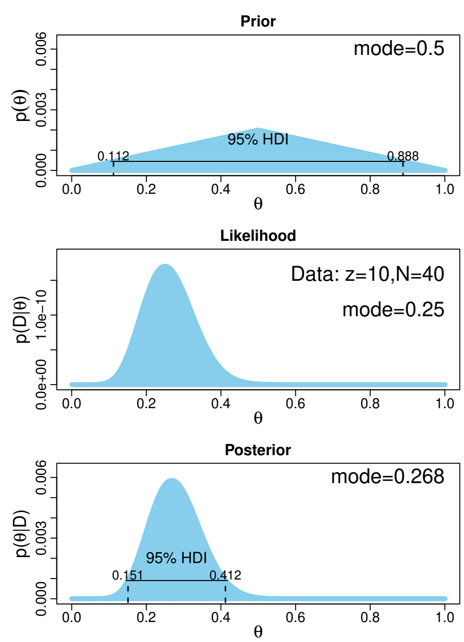

Figure 8: Example 7

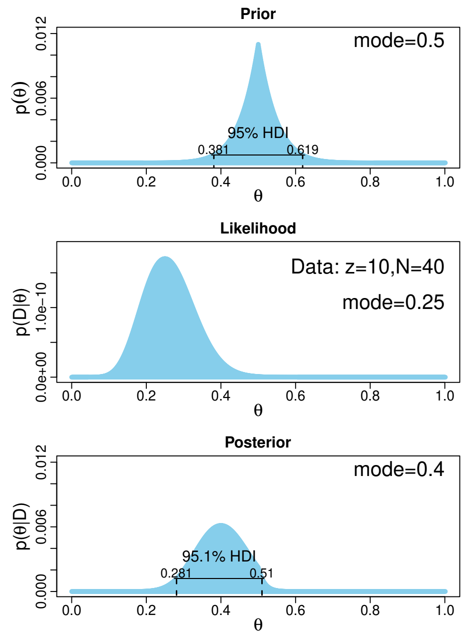

Figure 9: Example 8

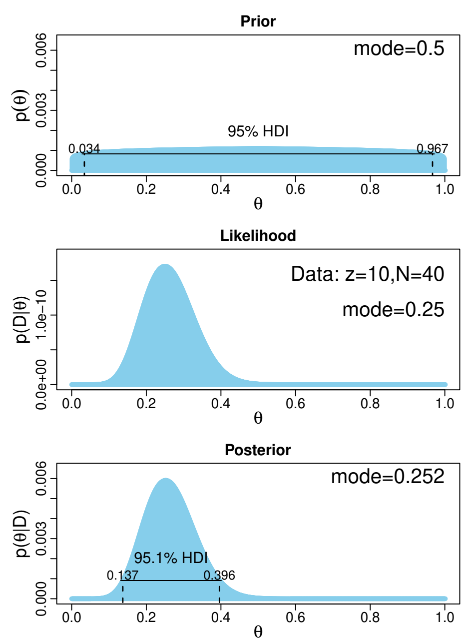

Figure 10: Example 9

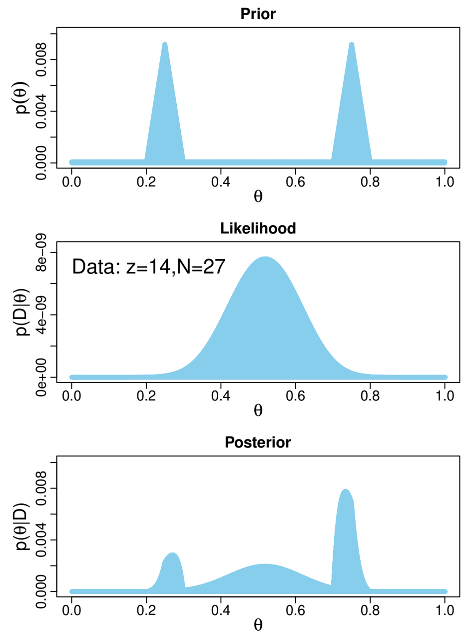

Figure 11: Example 10

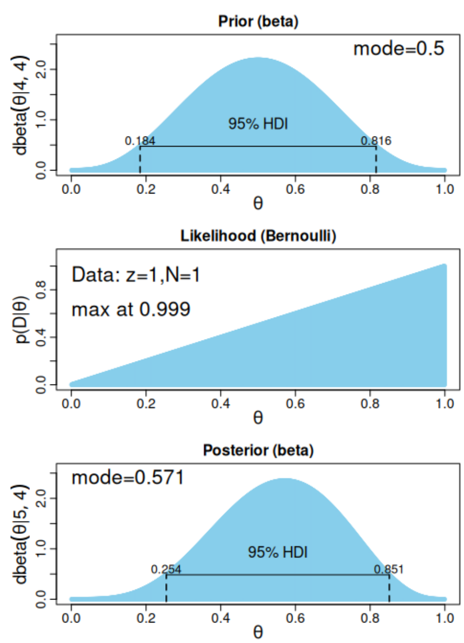

Figure 12: Caption

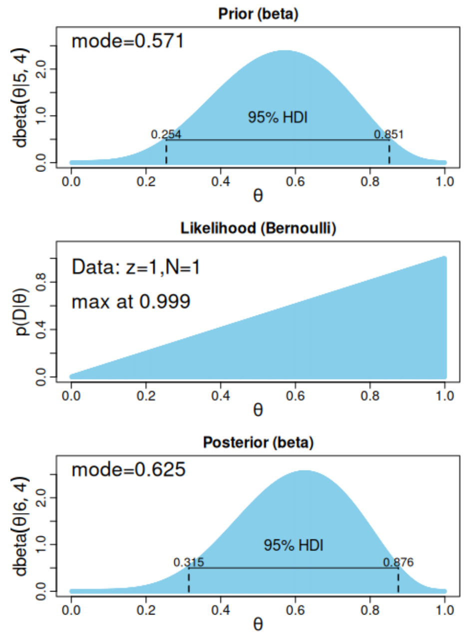

Figure 13: Caption

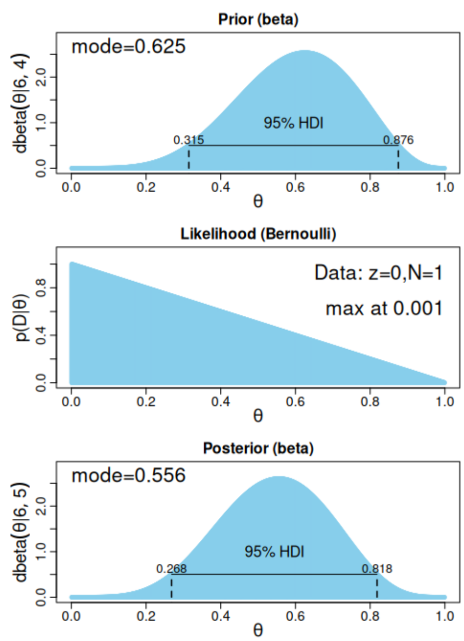

Figure 14: Caption

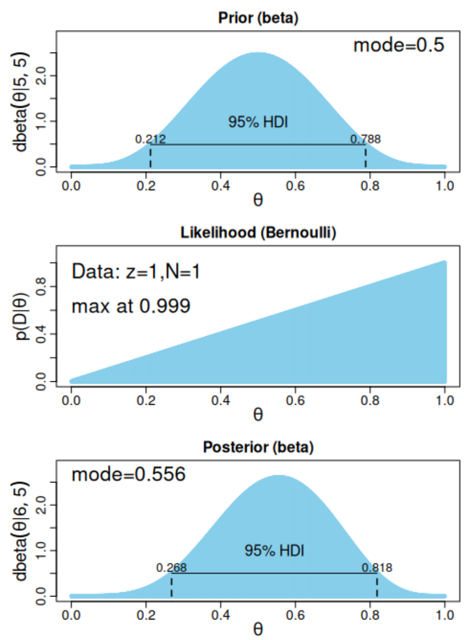

Figure 15: Caption

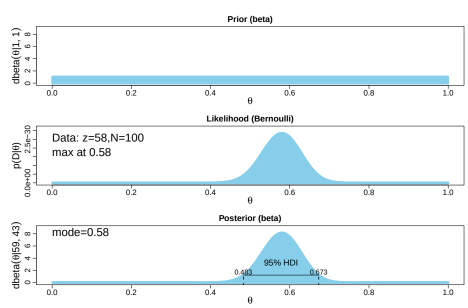

Figure 16: Exercise 6.2 A

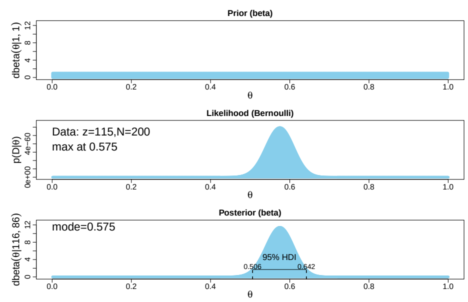

Figure 17: Exercise 6.2 B

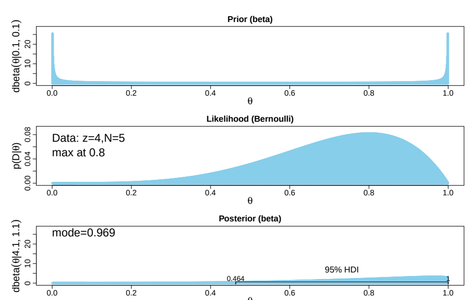

Figure 18: Exercise 6.4

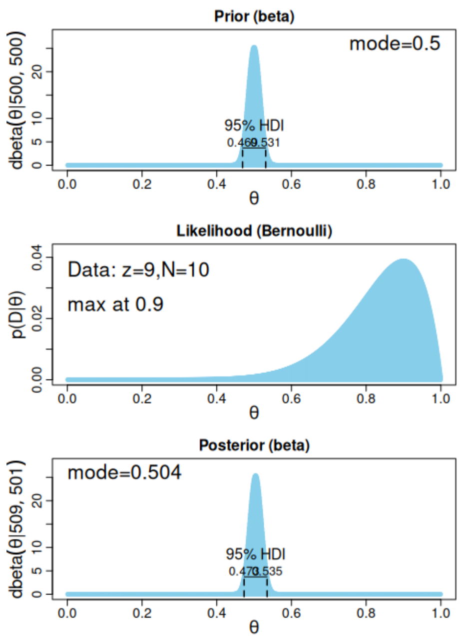

Figure 19: 6.5 A

6.5

We use *BernBeta.R*. Results are shown in figs. 19 and 20.

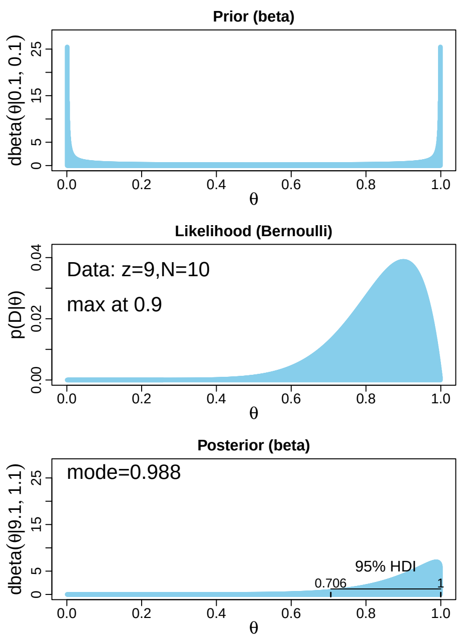

Figure 20: 6.5 B