Workshop 1: Conjugate Bayesian inference in **R** Solutions

1. (i) The posterior distribution of the success rate is

$$\begin{aligned} p(\theta \mid y) &\propto f(y \mid \theta)p(\theta) \\ &= \binom{n}{y}\theta^y(1-\theta)^{n-y}\frac{1}{B(a,b)}\theta^{a-1}(1-\theta)^{b-1} \\ &\propto \theta^{a+y-1}(1-\theta)^{b+n-y-1}, \end{aligned}$$

which we recognise as the kernel of a beta distribution with parameters $a+y$ and $b+n-y$. Therefore,

$$\theta \mid y \sim \text{Beta}(a+y, b+n-y).$$

Taking $a = 9.2$, $b = 13.8$, $n = 20$, and $y = 15$, results in a Beta$(24.2, 18.8)$ distribution.

(ii) The posterior mean is $24.2/(24.2 + 18.8) = 0.563$. Using the function `hpd` from the package `TeachingDemos` (see `R` script), we obtain the HPD interval $(0.416, 0.708)$.

(iii) By computing the 2.5% and 97.5% percentiles of the posterior distribution, we obtain the symmetric credible interval $(0.414, 0.706)$. The two intervals (HPD and credible) are basically the same because in this case the posterior distribution is unimodal (and also practically symmetric around the mean).

(iv) The probability that the true success rate is greater than 0.6 is 0.316.

(v) Under a uniform prior, i.e., with a Beta$(1, 1)$ prior distribution, the above probability changes to 0.904. With a Jeffreys' prior, it is 0.918.

(vi) Let z denotes the number of positive responses in further $m = 40$ patients. We must first calculate the posterior predictive distribution

$$\begin{aligned} f(z \mid y) &= \int_{\Theta} f(z \mid \theta)p(\theta \mid y)\mathrm{d}\theta \\ &= \int_0^1 \binom{m}{z}\theta^z(1-\theta)^{m-z}\frac{1}{B(a+y,b+n-y)}\theta^{a+y-1}(1-\theta)^{b+n-y-1}\mathrm{d}\theta \\ &= \binom{m}{z}\frac{1}{B(a+y,b+n-y)}\int_0^1 \theta^{a+y+z-1}(1-\theta)^{b+n-y+m-z-1}\mathrm{d}\theta \\ &= \binom{m}{z}\frac{B(a+y+z,b+n-y+m-z)}{B(a+y,b+n-y)}\int_0^1 \frac{1}{B(a+y+z,b+n-y+m-z)}\theta^{a+y+z-1}(1-\theta)^{b+n-y+m-z-1}\mathrm{d}\theta \\ &= \binom{m}{z}\frac{B(a+y+z,b+n-y+m-z)}{B(a+y,b+n-y)} \end{aligned}$$

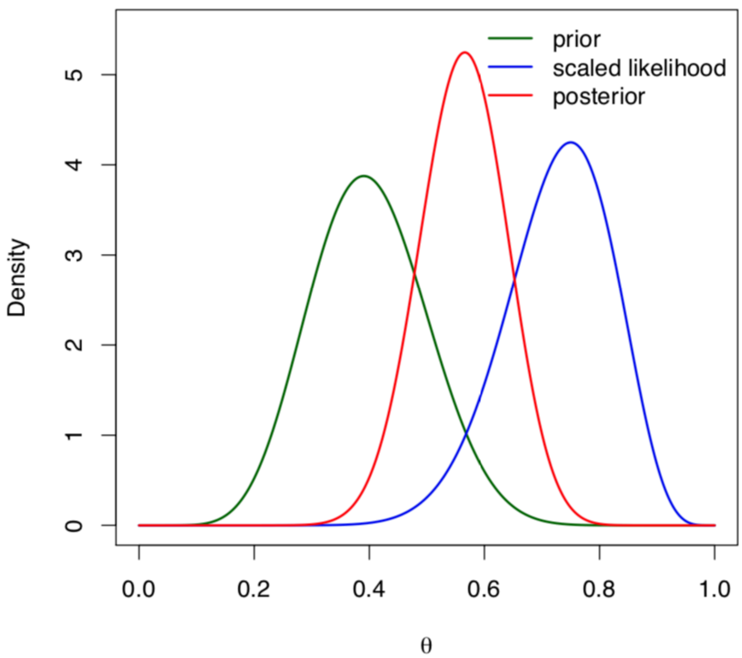
This is the Beta-Binomial Distribution. It is now straightforward to find that Pr(z ≥ 25) = 0.329 (see R script).

(vii) We start by calculating the prior predictive distribution

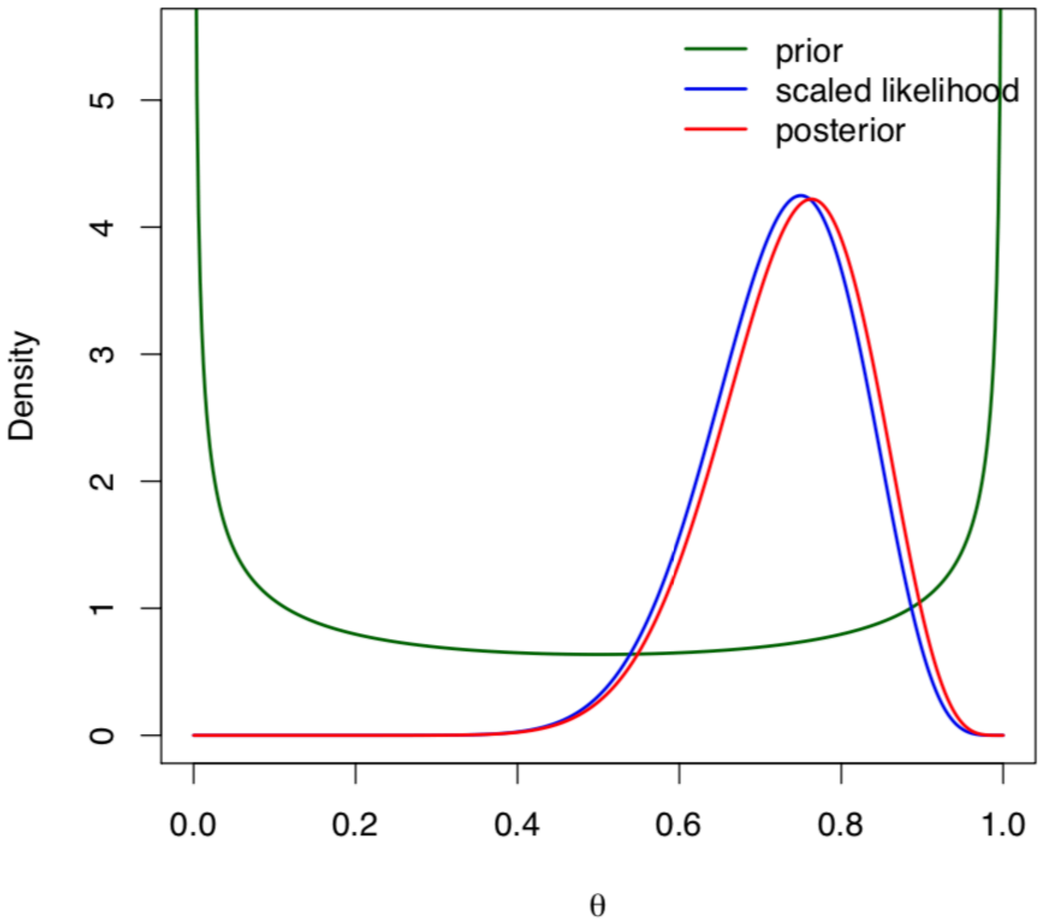
$$
\begin{aligned}
f(y) &= \int_\Theta f(y \mid \theta) p(\theta) \mathrm{d}\theta \\
&= \int_0^1 \binom{n}{y} \theta^y (1-\theta)^{n-y} \frac{1}{B(a,b)} \theta^{a-1} (1-\theta)^{b-1} \mathrm{d}\theta \\
&= \binom{n}{y} \frac{1}{B(a,b)} \int_0^1 \theta^{a+y-1} (1-\theta)^{b+n-y-1} \mathrm{d}\theta \\
&= \binom{n}{y} \frac{B(a+y, b+n-y)}{B(a,b)}
\end{aligned}
$$

The prior predictive probability of observing at least 15 positive responses can then be computed from the last expression and it is 0.01526 (see R script for further details). This suggests some evidence that the data and the prior are incompatible.

(viii) The prior/likelihood/posterior plot is shown in the figure below

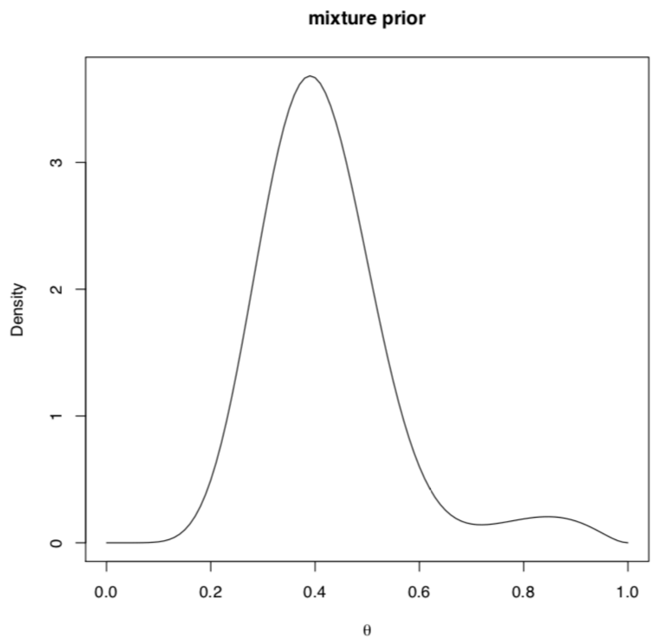


There is not much overlap between the support of the prior and the likelihood and the prior has considerable effect on the posterior. Re-doing the same plot but now with Jeffreys' prior we can appreciate that now the prior has basically no effect on the posterior, with most information coming from the likelihood.



2. (i) Solving for a and b gives a Beta$(12, 3)$ prior.

(ii) The mixture prior $\theta \sim \pi\text{Beta}(a_1, b_1) + (1-\pi)\text{Beta}(a_2, b_2)$ is plotted in the figure below.



(iii) We will start by finding the posterior distribution of θ.

$$
\begin{aligned}
p(\theta \mid y) &\propto \binom{n}{y}\theta^{y}(1-\theta)^{n-y}\left\{\pi\frac{1}{B(a_1,b_1)}\theta^{a_1-1}(1-\theta)^{b_1-1}+(1-\pi)\frac{1}{B(a_2,b_2)}\theta^{a_2-1}(1-\theta)^{b_2-1}\right\}\\
&\propto \pi\frac{1}{B(a_1,b_1)}\theta^{a_1+y-1}(1-\theta)^{b_1+n-y-1}+(1-\pi)\frac{1}{B(a_2,b_2)}\theta^{a_2+y-1}(1-\theta)^{b_2+n-y-1}\\
&= \pi\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}\frac{1}{B(a_1+y,b_1+n-y)}\theta^{a_1+y-1}(1-\theta)^{b_1+n-y-1}\\
&\quad+(1-\pi)\frac{B(a_2+y,b_2+n-y)}{B(a_2,b_2)}\frac{1}{B(a_2+y,b_2+n-y)}\theta^{a_2+y-1}(1-\theta)^{b_2+n-y-1}\\
&= \pi\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}\text{Beta}(\theta \mid a_1+y,b_1+n-y)\\
&\quad+(1-\pi)\frac{B(a_2+y,b_2+n-y)}{B(a_2,b_2)}\text{Beta}(\theta \mid a_2+y,b_2+n-y).
\end{aligned}
$$

We are almost there, but note that the 'weights' $\pi\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}$ and $(1-\pi)\frac{B(a_2+y,b_2+n-y)}{B(a_2,b_2)}$ do not sum up to one. Renormalising, we finally obtain that

$$
\theta \mid y \sim \omega_1\text{Beta}(\theta \mid a_1+y,b_1+n-y)+(1-\omega_1)\text{Beta}(\theta \mid a_2+y,b_2+n-y)
$$

with

$$
\omega_1=\pi\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}\left(\pi\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}+(1-\pi)\frac{B(a_2+y,b_2+n-y)}{B(a_2,b_2)}\right)^{-1}
$$

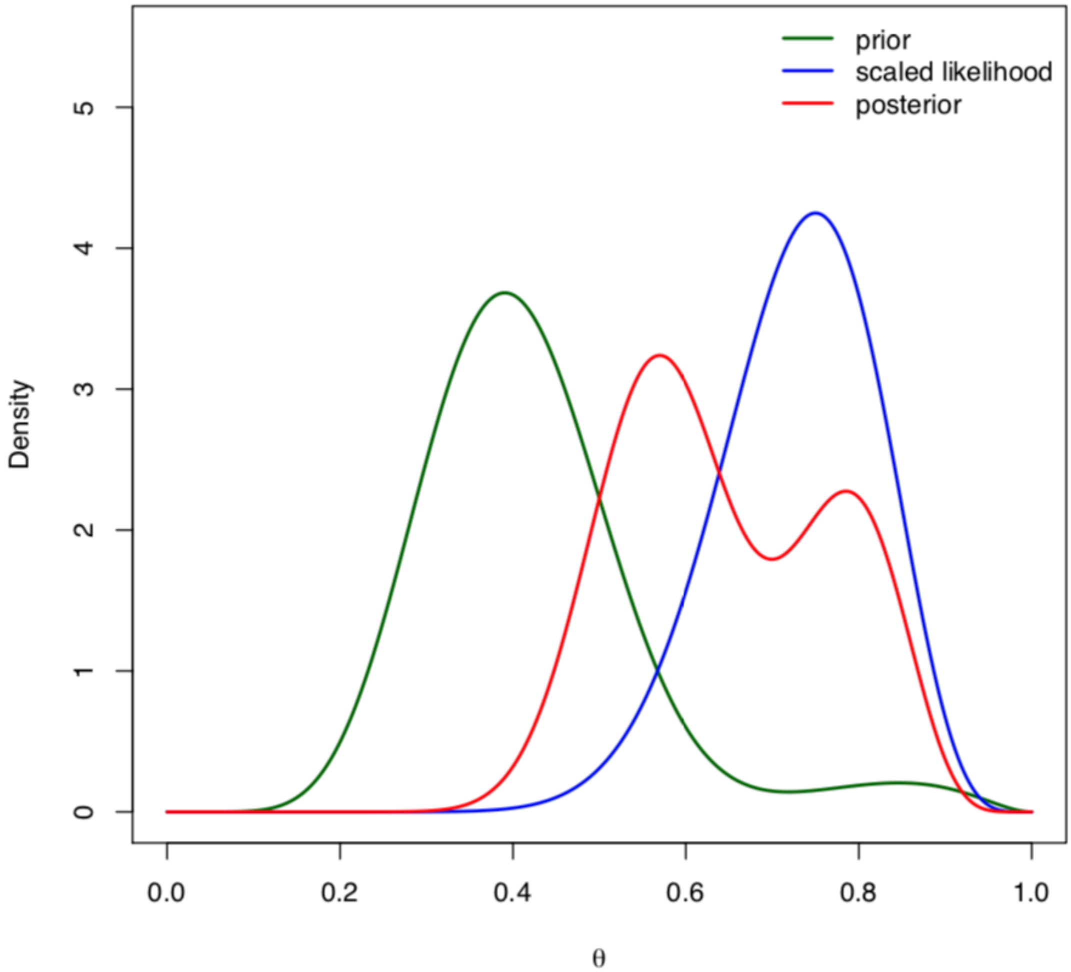
We are now ready to compute the required probability (see R script), which turns out to be 0.58062.

(iv) The procedure is similar to the one in 1. (vii), the only difference is the computation of the prior predictive distribution. In this case,

$$
\begin{aligned}
f(y) &= \int_{\Theta} f(y \mid \theta)p(\theta)\mathrm{d}\theta\\
&= \int_0^1 \binom{n}{y}\theta^{y}(1-\theta)^{n-y}\left\{\pi\frac{1}{B(a_1,b_1)}\theta^{a_1-1}(1-\theta)^{b_1-1}+(1-\pi)\frac{1}{B(a_2,b_2)}\theta^{a_2-1}(1-\theta)^{b_2-1}\right\}\mathrm{d}\theta\\
&= \pi\binom{n}{y}\frac{1}{B(a_1,b_1)}\int_0^1 \theta^{a_1+y-1}(1-\theta)^{b_1+n-y-1}\mathrm{d}\theta+(1-\pi)\binom{n}{y}\frac{1}{B(a_2,b_2)}\int_0^1 \theta^{a_2+y-1}(1-\theta)^{b_2+n-y-1}\mathrm{d}\theta\\
&= \pi\binom{n}{y}\frac{B(a_1+y,b_1+n-y)}{B(a_1,b_1)}+(1-\pi)\binom{n}{y}\frac{B(a_2+y,b_2+n-y)}{B(a_2,b_2)}
\end{aligned}
$$

The prior predictive probability of observing at least 15 positive responses is now 0.0514 (see R script for further details), which does not provide strong evidence of incompatibility.

(v) The prior/likelihood/posterior plot is shown in the figure below



3. (i) The survey information is equivalent to a normal prior for the mean SBP with mean 120 and variance 100.

(ii) We have seen in class that

$$\theta \mid \mathbf{y}, \sigma^2 \sim \mathrm{N}\left(\frac{\frac{\mu_0}{\sigma_0^2} + n\frac{\bar{\mathbf{y}}}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}}\right).$$

In this case we have $\bar{\mathbf{y}} = 130$, $n = 2$, $\sigma^2 = 25$, $\mu_0 = 120$, $\sigma_0^2 = 100$, leading to a posterior mean of 128.89 and a 95% credible interval of $(122.356, 135.422)$. The mle is 130 and a 95% confidence interval is $\bar{y} \pm 1.96\frac{\sigma}{\sqrt{n}} = (123.070, 136.930)$.

(iii) With the 2 new observations the sample mean is unchanged, $\bar{\mathbf{y}} = 130$, but now $n = 4$. The 95% credible interval is now $(124.658, 134.165)$.