Bayesian statistics with R
3. Analyses by hand

Olivier Gimenez

March 2021

Back to Bayes

A simple example

- Let us take a simple example to fix ideas.
- 120 deer were radio-tracked over winter.
- 61 close to a plant, 59 far from any human activity.
- Question: is there a treatment effect on survival?

	Released	Alive	Dead	Other
treatment	61	19	38	4
control	59	21	38	0

- So, $n = 57$ deer were assigned to the treatment group of which $k = 19$ survived the winter.

- So, $n = 57$ deer were assigned to the treatment group of which $k = 19$ survived the winter.
- Of interest is the probability of over-winter survival, call it θ, for the general population within the treatment area.

- So, $n = 57$ deer were assigned to the treatment group of which $k = 19$ survived the winter.
- Of interest is the probability of over-winter survival, call it θ, for the general population within the treatment area.
- The obvious estimate is simply to take the ratio $k/n = 19/57$.

- So, $n = 57$ deer were assigned to the treatment group of which $k = 19$ survived the winter.
- Of interest is the probability of over-winter survival, call it θ, for the general population within the treatment area.
- The obvious estimate is simply to take the ratio $k/n = 19/57$.
- How would the classical statistician justify this estimate?

- Our model is that we have a Binomial experiment (assuming independent and identically distributed draws from the population).

- Our model is that we have a Binomial experiment (assuming independent and identically distributed draws from the population).
- K the number of alive individuals at the end of the winter, so that $P(K = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$.

- Our model is that we have a Binomial experiment (assuming independent and identically distributed draws from the population).
- K the number of alive individuals at the end of the winter, so that $P(K = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$.
- The classical approach is to maximise the corresponding likelihood with respect to θ to obtain the entirely plausible MLE:

$$\hat{\theta} = k/n = 19/57$$

.

The Bayesian approach

- The Bayesian starts off with a prior.

The Bayesian approach

- The Bayesian starts off with a prior.
- Now, the one thing we know about θ is that is a continuous random variable and that it lies between zero and one.

The Bayesian approach

- The Bayesian starts off with a prior.
- Now, the one thing we know about θ is that is a continuous random variable and that it lies between zero and one.
- Thus, a suitable prior distribution might be the Beta defined on $[0, 1]$.

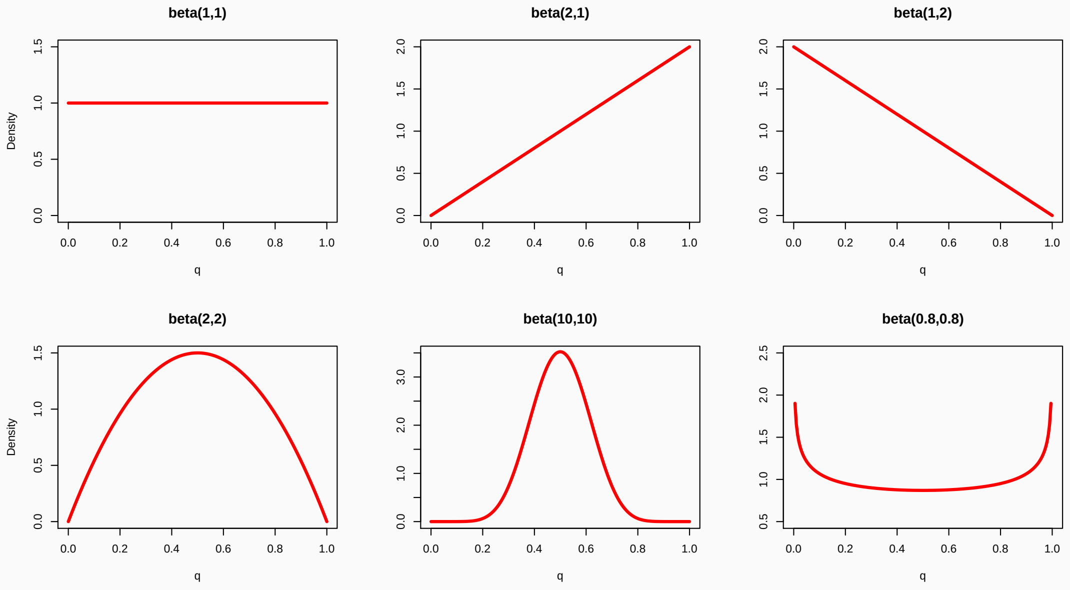
The Bayesian approach

- The Bayesian starts off with a prior.
- Now, the one thing we know about θ is that is a continuous random variable and that it lies between zero and one.
- Thus, a suitable prior distribution might be the Beta defined on $[0, 1]$.
- What is the Beta distribution?

What is the Beta distribution?

$$q(\theta \mid \alpha, \beta) = \frac{1}{\text{Beta}(\alpha, \beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

with $\text{Beta}(\alpha, \beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and $\Gamma(n) = (n-1)!$


beta(1,1)
beta(2,1)
beta(1,2)
beta(2,2)
beta(10,10)
beta(0.8,0.8)
Density
q
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.5
2.0
2.5
3.0

The Bayesian approach

- We assume a priori that $\theta \sim Beta(a, b)$ so that $\Pr(\theta) = \theta^{a-1}(1-\theta)^{b-1}$

The Bayesian approach

- We assume a priori that $\theta \sim Beta(a, b)$ so that $\Pr(\theta) = \theta^{a-1}(1-\theta)^{b-1}$
- Then we have:

$$
\begin{aligned}
Pr(\theta \mid k) &\propto \binom{n}{k}\theta^{k}(1-\theta)^{n-k}\;\theta^{a-1}(1-\theta)^{b-1} \\
&\propto \theta^{(a+k)-1}(1-\theta)^{(b+n-k)-1}
\end{aligned}
$$

The Bayesian approach

- We assume a priori that $\theta \sim Beta(a, b)$ so that $\Pr(\theta) = \theta^{a-1}(1-\theta)^{b-1}$
- Then we have:

$$
\begin{aligned}
Pr(\theta \mid k) &\propto \binom{n}{k}\theta^{k}(1-\theta)^{n-k}\ \theta^{a-1}(1-\theta)^{b-1} \\
&\propto \theta^{(a+k)-1}(1-\theta)^{(b+n-k)-1}
\end{aligned}
$$

- That is:

$$
\theta \mid k \sim Beta(a+k, b+n-k)
$$

The Bayesian approach

- We assume a priori that $\theta \sim Beta(a, b)$ so that $\Pr(\theta) = \theta^{a-1}(1-\theta)^{b-1}$
- Then we have:

$$\begin{aligned} Pr(\theta \mid k) &\propto \binom{n}{k}\theta^{k}(1-\theta)^{n-k}\ \theta^{a-1}(1-\theta)^{b-1} \\ &\propto \theta^{(a+k)-1}(1-\theta)^{(b+n-k)-1} \end{aligned}$$

- That is:

$$\theta \mid k \sim Beta(a+k, b+n-k)$$

- Take a Beta prior with a Binomial likelihood, you get a Beta posterior (conjugacy)

Application to the deer example

- Posterior distribution of survival is $\theta \sim Beta(a+k, b+n-k)$.

Application to the deer example

- Posterior distribution of survival is $\theta \sim Beta(a+k, b+n-k)$.
- If we take a Uniform prior, i.e. $Beta(1,1)$, then we have:

Application to the deer example

- Posterior distribution of survival is $\theta \sim Beta(a + k, b + n - k)$.
- If we take a Uniform prior, i.e. $Beta(1, 1)$, then we have:
- $\theta_{treatment} \sim Beta(1 + 19, 1 + 57 - 19) = Beta(20, 39)$

Application to the deer example

- Posterior distribution of survival is $\theta \sim Beta(a+k, b+n-k)$.
- If we take a Uniform prior, i.e. $Beta(1,1)$, then we have:
- $\theta_{treatment} \sim Beta(1+19, 1+57-19) = Beta(20,39)$
- Note that in this specific situation, the posterior has an explicit expression, easy to manipulate.

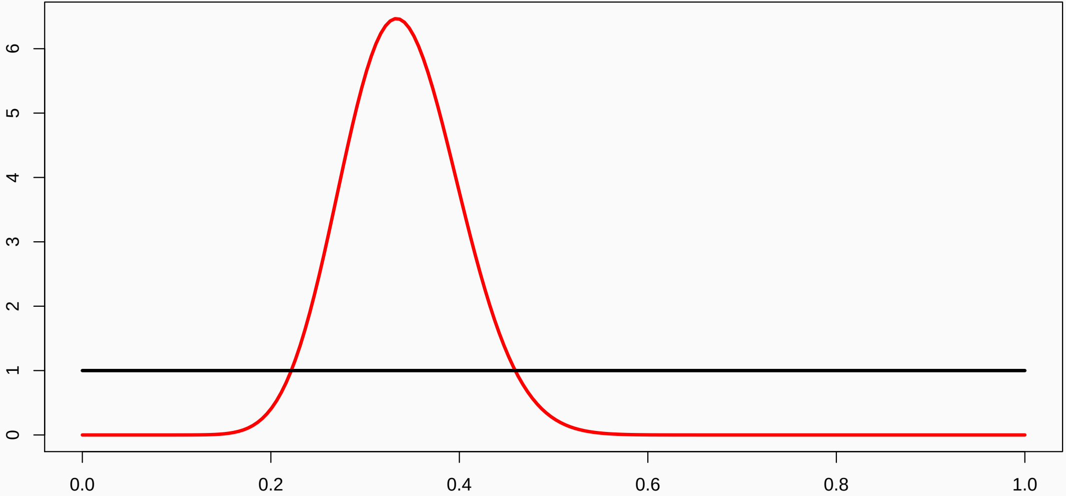
Application to the deer example

- Posterior distribution of survival is $\theta \sim Beta(a+k, b+n-k)$.
- If we take a Uniform prior, i.e. $Beta(1,1)$, then we have:
- $\theta_{treatment} \sim Beta(1+19, 1+57-19) = Beta(20, 39)$
- Note that in this specific situation, the posterior has an explicit expression, easy to manipulate.
- In particular, $E(Beta(a,b)) = \dfrac{a}{a+b} = 20/59$ to be compared with the MLE $19/57$.

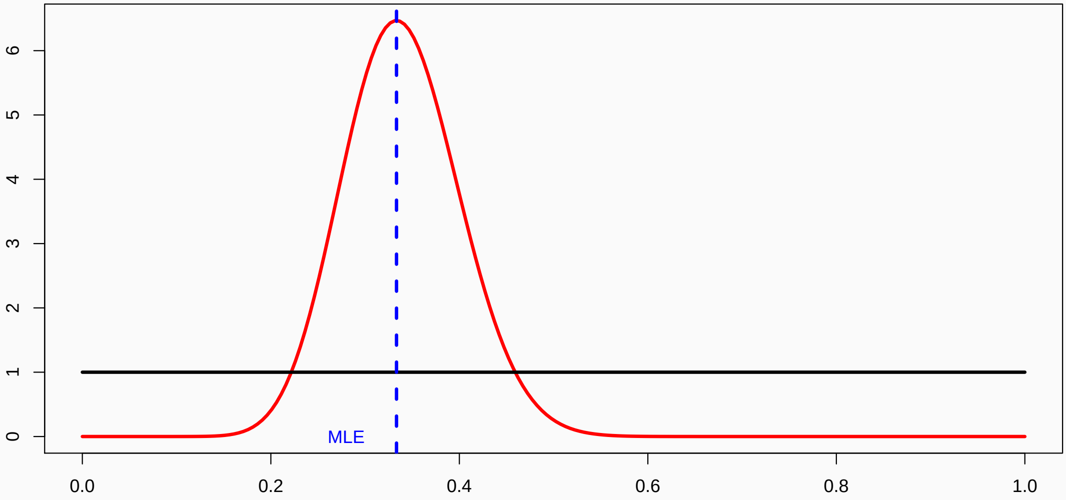
A general result

This is a general result, the Bayesian and frequentist estimates will always agree if there is sufficient data, so long as the likelihood is not explicitly ruled out by the prior.

Prior $Beta(1,1)$ and posterior survival $Beta(20,39)$



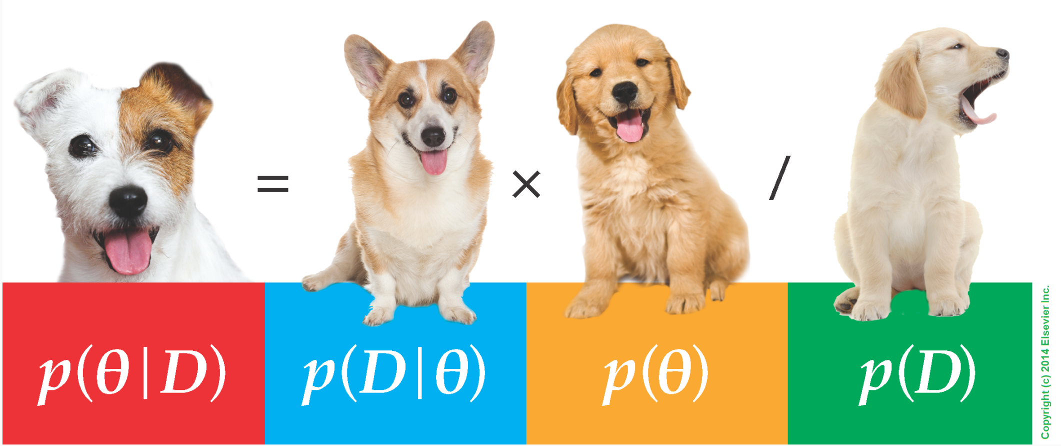
Prior $Beta(1, 1)$ and posterior survival $Beta(20, 39)$



Notation

Our model so far

$$
\begin{aligned}
y &\sim \text{Binomial}(N, \theta) && \text{[likelihood]} \\
\theta &\sim \text{Beta}(1, 1) && \text{[prior for } \theta\text{]}
\end{aligned}
$$


$p(\theta|D) = p(D|\theta) \times p(\theta) / p(D)$
Copyright (c) 2014 Elsevier Inc.