Application to the historical example

1. **the likelihood**

$$f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^S(1-\theta)^{n-S} \qquad \text{where } S = \sum_{i=1}^{n} y_i$$

2. **the prior**

   . . .

3. **the posterior**

   . . .

Application to the historical example

1. **the likelihood**

$$f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^S(1-\theta)^{n-S} \qquad \text{where } S = \sum_{i=1}^{n} y_i$$

2. **the prior**

   Uniform: $\pi(\theta) = 1$

3. **the posterior**

   ...

Application to the historical example

1. **the likelihood**

$$f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^S(1-\theta)^{n-S} \qquad \text{where } S = \sum_{i=1}^{n} y_i$$

2. **the prior**

Uniform: $\pi(\theta) = 1$

3. **the posterior**

$$p(\theta|\boldsymbol{y}) = \frac{\theta^S(1-\theta)^{n-S}}{f(\boldsymbol{y})} = p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\theta^S(1-\theta)^{n-S}$$

Application to the historical example

1. **the likelihood**

$$f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^S(1-\theta)^{n-S} \qquad \text{where } S = \sum_{i=1}^{n} y_i$$

2. **the prior**

   Uniform: $\pi(\theta) = 1$

3. **the posterior**

$$p(\theta|\boldsymbol{y}) = \frac{\theta^S(1-\theta)^{n-S}}{f(\boldsymbol{y})} = p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\theta^S(1-\theta)^{n-S}$$

To answer the question of interest, we can then compute: ...

Application to the historical example

1. **the likelihood**

$$f(\boldsymbol{y}|\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{(1-y_i)} = \theta^S(1-\theta)^{n-S} \qquad \text{where } S = \sum_{i=1}^{n} y_i$$

2. **the prior**

   Uniform: $\pi(\theta) = 1$

3. **the posterior**

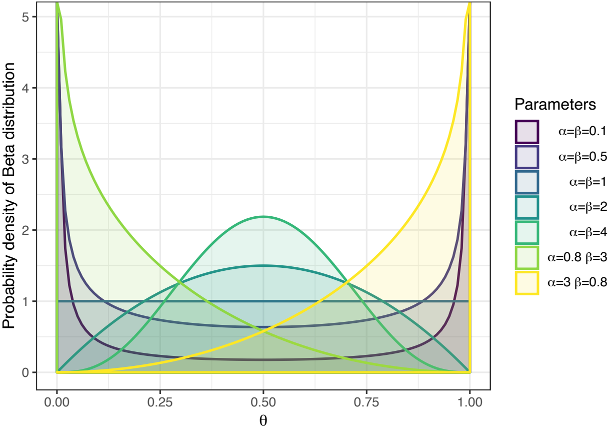
$$p(\theta|\boldsymbol{y}) = \frac{\theta^S(1-\theta)^{n-S}}{f(\boldsymbol{y})} = p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\theta^S(1-\theta)^{n-S}$$

To answer the question of interest, we can then compute:

$$P(\theta \geq 0.5|\boldsymbol{y}) = \int_{0.5}^{1} p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\int_{0.5}^{1}\theta^S(1-\theta)^{n-S}\mathrm{d}\theta \approx 1.15\ 10^{-42}$$

The Beta distribution

$$f(\theta) = \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\theta^{\alpha-1}(1-\theta)^{\beta-1} \text{ for } \alpha > 0 \text{ and } \beta > 0$$



Examples of various parametrizations for the Beta distribution

Construction of a Bayesian model

Conjugacy of the Beta distribution

Beta *prior*: $\pi = \text{Beta}(\alpha, \beta)$

Conjugacy of the Beta distribution

Beta *prior*: $\pi = \text{Beta}(\alpha, \beta)$

Corresponding *posterior*: $p(\theta|\mathbf{y}) \propto \theta^{\alpha+S-1}(1-\theta)^{\beta+(n-S)-1}$

...

The $\propto$ symbol means: "proportional to"

Conjugacy of the Beta distribution

Beta *prior*: $\pi = \text{Beta}(\alpha, \beta)$

Corresponding *posterior*: $p(\theta|\mathbf{y}) \propto \theta^{\alpha+S-1}(1-\theta)^{\beta+(n-S)-1}$

$\Rightarrow \theta|\mathbf{y} \sim \text{Beta}(\alpha + S, \beta + (n - S))$

The $\propto$ symbol means: "proportional to"

Conjugacy of the Beta distribution

Beta *prior*: $\pi = \text{Beta}(\alpha, \beta)$

Corresponding *posterior*: $p(\theta|\mathbf{y}) \propto \theta^{\alpha+S-1}(1-\theta)^{\beta+(n-S)-1}$

$\Rightarrow \theta|\mathbf{y} \sim \text{Beta}(\alpha + S, \beta + (n-S))$

This is called a **conjugated distribution** because the ***posterior*** and the ***prior*** belong to the **same parametric family**

The $\propto$ symbol means: "proportional to"

Impact of the *prior* choice

Interpretation of the *prior*	Parameters of the Beta distribution	$P(\theta \geq 0.5 \mid \boldsymbol{y})$
#boys > #girls	$\alpha = 0.1, \beta = 3$	$1.08\ 10^{-42}$
#boys < #girls	$\alpha = 3, \quad \beta = 0.1$	$1.19\ 10^{-42}$
#boys = #girls	$\alpha = 4, \quad \beta = 4$	$1.15\ 10^{-42}$
#boys ≠ #girls	$\alpha = 0.1, \beta = 0.1$	$1.15\ 10^{-42}$
non-informative	$\alpha = 1, \quad \beta = 1$	$1.15\ 10^{-42}$

For 493,472 newborns including 241,945 girls

Impact of the *prior* choice

Interpretation of the *prior*	Parameters of the Beta distribution	$P(\theta \geq 0.5 \mid \mathbf{y})$
#boys > #girls	$\alpha = 0.1, \beta = 3$	$1.08\ 10^{-42}$
#boys < #girls	$\alpha = 3, \quad \beta = 0.1$	$1.19\ 10^{-42}$
#boys = #girls	$\alpha = 4, \quad \beta = 4$	$1.15\ 10^{-42}$
#boys ≠ #girls	$\alpha = 0.1, \beta = 0.1$	$1.15\ 10^{-42}$
non-informative	$\alpha = 1, \quad \beta = 1$	$1.15\ 10^{-42}$

For 493,472 newborns including 241,945 girls

Interpretation of the *prior*	Parameters of the Beta distribution	$P(\theta \geq 0.5 \mid \mathbf{y})$
#boys > #girls	$\alpha = 0.1, \beta = 3$	0.39
#boys < #girls	$\alpha = 3, \quad \beta = 0.1$	0.52
#boys = #girls	$\alpha = 4, \quad \beta = 4$	0.46
#boys ≠ #girls	$\alpha = 0.1, \beta = 0.1$	0.45
non-informative	$\alpha = 1, \quad \beta = 1$	0.45

For 20 newborns including 9 girls

Impact of the *prior* choice for 20 observed births