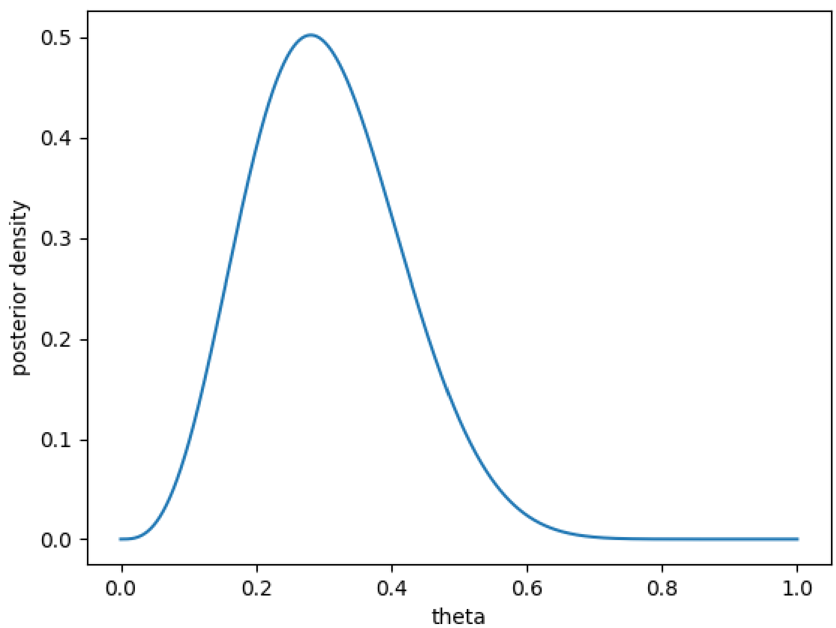
Xinyu Tan

– 1. Posterior inference

$$
\begin{aligned}
p(\theta|y<3) &\propto p(\theta)p(y<3|\theta) && (1)\\
&= \text{Beta}(4,4)\sum_{i=0}^{2} p(y=i|\theta) && (2)\\
&= \text{Beta}(4,4)\left((1-\theta)^2 + 10\theta(1-\theta) + 45\theta^2\right)(1-\theta)^8 && (3)
\end{aligned}
$$

Density plot:



– 2. Predictive distributions

Denote the result of i-th coin spin x_i, $x_i \in \{T, H\}$. We are given that $x_1 = T$ and $x_2 = T$; for any coin with $p(\text{head}) = p_h$, the probability of the event that until n-th spin a head shows up is:

$$
\begin{aligned}
p(E) &= p(x_n = H, x_{n-1} = T, \cdots, x_3 = T | x_1 = T, x_2 = T)\\
&= \frac{p(x_n = H, x_{n-1} = T, \cdots)}{p(x_1 = T, x_2 = T)}\\
&= \frac{p_h(1-p_h)^{n-1}}{(1-p_h)^2}\\
&= p_h(1-p_h)^{n-3}
\end{aligned}
$$

Let's translate the n-th spin to additional spin. Suppose additional spin is m, then total spin is $m+2$, $m \geq 1$. Then $p(E) = p_h(1-p_h)^{m-1}$.

Now, we have two coins with $p(\text{head}|C_1) = 0.6$ and $p(\text{head}|C_2) = 0.4$.

$$\begin{aligned} p(E) &= p(E|C_1)p(C_1) + p(E|C_2)p(C_2) \\ &= 0.6 \times 0.4^{m-1} \times 0.5 + 0.4 \times 0.6^{m-1} \times 0.5 \end{aligned}$$

The expectation of additional spins until a head shows up is:

$$\mathbb{E} = \sum_{m=1}^{\infty} mp(E) = \sum_{m=1}^{\infty} m\left(0.3 \times 0.4^{m-1} + 0.2 \times 0.6^{m-1}\right) = \frac{0.3}{0.6^2} + \frac{0.2}{0.4^2} = 2.08$$

- **5. Posterior distribution as a compromise between prior information and data**

  (a) Posterior predictive distribution:

$$\Pr(y = k) = \int_0^1 \Pr(y = k|\theta)d\theta = \int_0^1 \binom{n}{k}\theta^k(1-\theta)^{n-k}d\theta = \frac{1}{n+1} \tag{4}$$

  (b) Posterior mean: $\frac{\alpha+y}{\alpha+\beta+n}$

  Let's consider one case. If $\frac{\alpha}{\alpha+\beta} > \frac{y}{n} \to \alpha n > (\alpha+\beta)y$, we then have

$$\frac{\alpha+y}{\alpha+\beta+n} - \frac{y}{n} = \frac{\alpha n - (\alpha+\beta)y}{(\alpha+\beta)(\alpha+\beta+n)} > 0$$

  and

$$\frac{\alpha+y}{\alpha+\beta+n} - \frac{\alpha}{\alpha+\beta} = \frac{(\alpha+\beta)y - \alpha n}{(\alpha+\beta)(\alpha+\beta+n)} < 0$$

  .

  Hence

$$\frac{y}{n} < \frac{\alpha+y}{\alpha+\beta+n} < \frac{\alpha}{\alpha+\beta}$$

  .

  In a similar fashion, we can show that when $\frac{\alpha}{\alpha+\beta} < \frac{y}{n}$, $\frac{\alpha}{\alpha+\beta} < \frac{\alpha+y}{\alpha+\beta+n} < \frac{y}{n}$, and when $\frac{\alpha}{\alpha+\beta} = \frac{y}{n}$, $\frac{\alpha+y}{\alpha+\beta+n} = \frac{\alpha}{\alpha+\beta} = \frac{y}{n}$

- **18. Poisson model**

  Gamma distribution:

$$p(\theta) = Gamma(\alpha, \beta) \propto \theta^{\alpha-1}e^{-\beta\theta}$$

  If

$$p(y|\theta) \propto \theta^{\sum_{i=1}^n y_i}e^{-(\sum_{i=1}^n x_i)\theta}$$

  , then the posterior distribution

$$p(\theta|y) \propto p(\theta)p(y|\theta) = \theta^{\sum_{i=1}^n y_i+\alpha-1}e^{-(\sum_{i=1}^n x_i+\beta)\theta}$$

  Hence,

$$p(\theta|y) = Gamma(\sum_{i=1}^n y_i + \alpha, \sum_{i=1}^n x_i + \beta)$$

- **19. Exponential model with conjugate prior distribution**

  (a) If $y|\theta$ is exponentially distributed with rate θ and given gamma prior distribution, then

$$p(\theta|y) \propto \theta e^{-y\theta}\theta^{\alpha-1}e^{-\beta\theta} = \theta^{\alpha}e^{-(y+\beta)\theta} \sim Gamma(\alpha+1, y+\beta)$$

  Therefore, gamma prior is conjugate for inferences about θ.

(b) Inverse Gamma distribution:

$$p(\phi) \propto \phi^{-(\alpha+1)} e^{-\beta/\phi}$$

Posterior distribution of ϕ:

$$p(\phi|y) \propto p(\phi)p(y|\phi) = \phi^{-(\alpha+2)} e^{-(\beta+y)/\phi} \sim \text{Inv-}Gamma(\alpha+1, \beta+y)$$