1 MLE for the Bernoulli/ binomial model

$$X_i \sim Ber(\theta) \tag{1}$$
$$p(D|\theta) = \theta^{N_1}(1-\theta)^{N_0} \tag{2}$$

$$\begin{aligned}
\ln\big(p(D|\theta)\big) &= \ln\Big(\theta^{N_1}(1-\theta)^{N_0}\Big) \\
&= \ln\Big(\theta^{N-1}\Big) + \ln{(1-\theta)}^{N_0} \\
&= N_1 \ln\theta + N_0 \ln(1-\theta) \\
\tfrac{\mathrm{d}}{\mathrm{d}\theta} \ln p(D|\theta) &= \frac{N_1}{\theta} - \frac{N_0}{1-\theta}
\end{aligned}$$

The log-likelihood will take a maximum when the derivative equals 0.

$$\begin{aligned}
0 &= \frac{N_1}{\theta} - \frac{N-N_1}{1-\theta} \\
0 &= N_1(1-\theta) - \theta(N-N_1) \\
0 &= N_1 - \theta N_1 - \theta N + \theta N_1 \\
0 &= N_1 - \theta(N_1 + N - N_1) \\
0 &= N_1 - \theta N \\
\hat{\theta} &= \frac{N_1}{N}
\end{aligned}$$

2 Marginal likelihood for Beta-Bernoulli model

$$p(X_{1:N}) = p(x_1)p(x_2|x_1)p(x_3|x_{1:2})...p(x_N|x_{N-1}) \tag{3}$$

$$p(X=k|D_{1:N}) = \frac{N_k + \alpha_k}{\sum_i N_i + \alpha_i} \tag{4}$$

$$(\alpha-1)! = \Gamma(\alpha) \tag{5}$$

Given $D = H, T, T, H, H \triangleq 1, 0, 0, 1, 1$

$$
\begin{aligned}
p(X=1|\alpha) &= \frac{\alpha_1}{\alpha} \\
p(X=0|\alpha, D_1) &= \frac{\alpha_0}{\alpha+1} \\
p(X=0|\alpha, D_{1:2}) &= \frac{\alpha_0+1}{\alpha+2} \\
p(X=1|\alpha, D_{1:3}) &= \frac{\alpha_0+1}{\alpha+3} \\
p(X=1|\alpha, D_{1:4}) &= \frac{\alpha_0+2}{\alpha+4}
\end{aligned}
$$

$$
\begin{aligned}
p(D) &= p(D_{1:5}) \\
&= p(D_1) \cdot p(D_2|D_1) \cdot p(D_3|D_{1:2}) \cdot p(D_4|D_{1:3}) \cdot p(D_5|D_{1:4}) \qquad \text{by (3)} \\
&= \frac{\alpha_1}{\alpha} \cdot \frac{\alpha_0}{\alpha+1} \cdot \frac{\alpha_0+1}{\alpha+2} \cdot \frac{\alpha_1+1}{\alpha+3} \cdot \frac{\alpha_1+2}{\alpha+4} \qquad \text{by (4)} \\
&= \frac{\big[\alpha_1(\alpha_1+1)(\alpha_1+2)\big]\big[\alpha_0(\alpha_0+1)\big]}{\alpha(\alpha+1)(\alpha+2)(\alpha+3)(\alpha+4)} \\
&= \frac{\big[(\alpha_1)...(\alpha_1+N_1-1)\big]\big[(\alpha_0)...(\alpha_0+N_0-1)\big]}{(\alpha)...(\alpha+N-1)} \\
&= \frac{(\alpha_1+N_1-1)!}{(\alpha_1-1)!} \cdot \frac{(\alpha_0+N_0-1)!}{(\alpha_0-1)!} \cdot \frac{(\alpha-1)!}{(\alpha+N-1)!} \\
&= \frac{\Gamma(\alpha_1+N_1)}{\Gamma(\alpha_1)} \cdot \frac{\Gamma(\alpha_0+N_0)}{\Gamma(\alpha_0)} \cdot \frac{\Gamma(\alpha)}{\Gamma(\alpha+N)} \\
&= \frac{\Gamma(\alpha_1+N_1)\,\Gamma(\alpha_0+N_0)}{\Gamma(\alpha_1)\Gamma(\alpha_0)} \frac{\Gamma(\alpha_0+\alpha_1)}{\Gamma(\alpha_0+\alpha_1+N)}
\end{aligned}
$$

3 Posterior predictive for a Beta-Binomial model

$$
\begin{aligned}
p(x|n, D) &= Bb(x|\alpha_0', \alpha_1', n) \\
&= \frac{B(x+\alpha_1', n-x+\alpha_0')}{B(\alpha_1', \alpha_0')}\binom{n}{x}
\end{aligned}
$$

Given $n = 1$

$$
\begin{aligned}
Bb(1|\alpha_0, \alpha_1, 1) &= \frac{B(1+\alpha_1, \alpha_0)}{B(\alpha_1, \alpha_0)} \binom{1}{1} \\
&= \frac{\Gamma(1+\alpha_1)\Gamma(\alpha_0)}{\Gamma(\alpha_0+\alpha_1+1)} \cdot \frac{\Gamma(\alpha_0+\alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} \\
&= \frac{\alpha_1\Gamma(\alpha_1)\Gamma(\alpha_0)}{(\alpha_0+\alpha_1)\Gamma(\alpha_0+\alpha_1)} \cdot \frac{\Gamma(\alpha_0+\alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} \\
&= \frac{\alpha_1}{\alpha_0+\alpha_1} \\
&= \frac{\alpha_1}{\alpha}
\end{aligned}
$$

3.1 Beta updating from censored likelihood

Let n represent the number of coin tosses. Let X represent the number of heads. Given $n = 5$ and $X < 3$, we need to compute the posterior $p(\theta|X < 3)$ under a $B(1, 1)$ prior up to normalization constants.

$$
\begin{aligned}
p(\theta) &= \frac{p(\theta)p(D|\theta)}{p(D)} \\
&= \frac{p(\theta) \cdot p(X < 3|\theta)}{p(X < 3)} \\
p(\theta) &\propto p(\theta) \cdot p(X < 3) \\
&\propto B(1, 1) \cdot \sum_{k=0}^{2} P(k|\theta, n) \\
&\binom{\propto \sum_{k=0}^{2} n}{k\theta^k(1-\theta)^{n-k}}
\end{aligned}
$$