$\mu_{MLE} = m/N$

assuming m/N head/tail exp Beta distributed

$\beta(\mu|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}$

$$P(\mu|D) = \frac{P(D|\mu)P(\mu)}{P(D)}$$

$(D = x_1, x_2 \ldots x_n)$

Prior

$P(D) =$ Constant independent of μ

(introduce another var & integrate $\forall\mu$)

sum rule

$p(D) = \int_{all\ \mu} p(D,\mu)\,d\mu$

after integration

P(D) is a constant

$$\propto p(x_1, x_2, \ldots, x_n|\mu)\,p(\mu)$$

$$= \prod_{n=1}^{N} p(x_n|\mu)\,p(\mu)$$

$$= \prod_{n=1}^{N} \left(\mu^{x_n}(1-\mu)^{1-x_n}\right)p(\mu)$$

$$= \mu^{m}(1-\mu)^{n-m} \cancel{\times} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\mu^{a-1}(1-\mu)^{b-1}$$

$$\propto \mu^{m}(1-\mu)^{n-m}\mu^{a-1}(1-\mu)^{b-1}$$

$$P(\mu|D) \propto \mu^{m+a-1}(1-\mu)^{\underbrace{(n-m)}_{l}+b-1} \Big\}\ \text{posterior distribution}$$

$$P(\mu|D) \propto \mu^{m+a-1}(1-\mu)^{l-b+1}$$

normalizing

$$\cancel{\propto} P(\mu|D) = \frac{\mu^{m+a-1}(1-\mu)^{l-b+1}}{\int \cdots d\mu}$$

$$\frac{\Gamma(m+a+n-m+b)}{\Gamma(m+a)\Gamma(n-m+b)}\mu^{m+a-1}(1-\mu)^{n-m+b-1}$$

$$p(\mu|D) = \beta(m+a, n-m+b)$$

$$P(x=1|D) = \int p(x=1|\mu)\, P(\mu|D)\, d\mu$$

$$P(x) = \sum_y P(x, y)$$

↳ Prob. of getting a head given μ

$$\int \mu\, p(\mu|D)\, d\mu = E(\mu|D)$$

$$\beta(m+a, l+b) = P(\mu|D)$$

$$E(\mu|D) = \frac{m+a}{m+a+l+b}$$

$$\mu_{ML} = \frac{m}{m+l} \qquad l = m-N$$

$$p(x=1|\mu) = \frac{m+a}{m+a+l+b}$$

when $n \to \infty$

$$p(\mu) = Beta(a,b) = Beta(2,3) \qquad \frac{m+a}{m+a+l+b} \approx \left(\frac{m}{m+l} = \frac{m}{N}\right)$$

There can be a relation between 2 things which we dunno