$\{x_i\}_{i=1}^{N}$ $\quad P(x|p) = p(1-p)^{x}$ DATA Likelihood

PRIOR

$$P(p|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} \cdot p^{\alpha-1}(1-p)^{\beta-1}$$

MLE

1. $$\mathcal{L}^{0} = \prod_{i=1}^{N} P(x_i|p)$$
$$= \prod_{i=1}^{N} p(1-p)^{x_i} \quad = P$$
$$= p^{N} \cdot (1-p)^{\sum_{i=1}^{N} x_i}$$

2. $$\mathcal{L} = \ln(\mathcal{L}^{0}) = N.\ln(p) + \left(\sum_{i=1}^{N} x_i\right).\ln(1-p)$$

3. $$\frac{\partial \mathcal{L}}{\partial p} = 0 \iff \frac{N}{p} - \left(\sum_{i=1}^{N} x_i\right) \cdot \frac{1}{1-p} = 0$$

$$\iff \boxed{p = \frac{N}{N + \sum_{i=1} x_i}}$$

$$\prod_{i=1}^{N} p(1-p)^{x_i}$$

$$= p(1-p)^{x_1} \cdot p(1-p)^{x_2} \cdot \dots p(1-p)^{x_N}$$

$$= p^N \cdot (1-p)^{x_1 + x_2 + \dots + x_N}$$

MAP

(1) $$\boxed{\mathcal{L}^{o} = \left(\prod_{i=1}^{N} P(x_i \mid p) \right) \cdot P(p \mid \alpha, \beta)}$$

$$= p^{N} \cdot (1-p)^{\sum_{i=1}^{N} x_i} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \cdot p^{\alpha-1} (1-p)^{\beta-1}$$

$$= \boxed{p^{N+\alpha-1} \cdot (1-p)^{\sum_{i=1}^{N} x_i + \beta - 1} \cdot \underbrace{\boxed{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)}}}_{= C}}$$

(2) $$\mathcal{L} = \ln(\mathcal{L}^{o}) = (N+\alpha-1) \cdot \ln(p) + \left(\sum_{i=1}^{N} x_i + \beta - 1 \right) \cdot \ln(1-p) + \ln(C)$$

(3) $$\frac{\partial \mathcal{L}}{\partial p} = 0 \iff \dots \quad \boxed{p = \frac{N+\alpha-1}{\sum_{i=1}^{N} x_i + N + \alpha + \beta - 2}}$$

In MAP:

$$\left(\prod_{i=1}^{N} P(x_i|p)\right) \cdot P(p|\alpha,\beta) \propto \underbrace{P(p|x_i,\alpha,\beta)}_{\text{POSTERIOR}}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} \cdot \boxed{p^{\underline{N+\alpha}-1}} \cdot (1-p)^{\sum_{i=1}^{N} x_i + \beta - 1}$$

REMEMBER, PRIOR: $P(p|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} \cdot \boxed{p^{\alpha-1}} \cdot (1-p)^{\beta-1}$

BETA is a conjugate prior for the GEOMETRIC — BETA relationship.

In online update,

$$\alpha^{(t+1)} \longleftarrow \alpha^{(t)} + N$$

$$\beta^{(t+1)} \longleftarrow \beta^{(t)} + \sum_{i=1}^{N} x_i$$

Least squares: $\varepsilon = t - y \sim G(0,1)$

where $y = X.w$

$\varepsilon = t - y \sim \text{Gamma}(\alpha = 2, \beta = 2)$

$$\arg\max_{w} \prod_{i=1}^{N} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \varepsilon_i^{\alpha - 1} \cdot e^{-\beta \cdot \varepsilon_i}$$