$$p(\theta|\tau) = \frac{\Gamma(\tau_1+\tau_2)}{\Gamma(\tau_1)\Gamma(\tau_2)} \underbrace{\theta^{\tau_1-1}(1-\theta)^{\tau_2-1}}, \quad \tau>0 \; ; \quad p(x|\theta) = \theta^x (1-\theta)^{1-x} \quad x \in \{0,1\} \quad \theta \in 0,1$$

$$\underbrace{\prod_{n=1}^{N} p(x_n|\theta)}$$

$$p(\theta|X) \propto$$

$$\theta^{\tau_1-1}(1-\theta)^{\tau_2-1}\,\theta^{\sum_{n=1}^{N} x_n}(1-\theta)^{N-\sum_{n=1}^{N} x_n} =$$

Beta

$$\underbrace{\frac{\Gamma(\tau_1+\tau_2+N)}{\Gamma(\tau_1+\sum_{n=1}^{N} x_n)\Gamma(\tau_2+N-\sum_{n=1}^{N} x_n)}\,\theta^{\tau_1+\sum_{n=1}^{N} x_n - 1}(1-\theta)^{\tau_2+(N-\sum_{n=1}^{N} x_n)-1}}$$

$$\log p(x|\lambda) = \langle \lambda, \phi(x) \rangle - A(\lambda)$$

$$\log p(\lambda; \tau, n_0) = \langle \lambda, \tau \rangle - n_0 A(\lambda) + \log H(\tau, n_0)$$

$$\log p(\lambda|X) \overset{+}{\propto} \langle \lambda, \tau \rangle - n_0 A(\lambda) + \langle \lambda, \sum_{n=1}^{N} \phi(x_n) \rangle - N A(\lambda) \Rightarrow p(\lambda|X) = p\left(\lambda; \tau + \sum_{n=1}^{N} \phi(x_n), n_0 + N\right)$$

$$p(x^*|X) = \int p(x^*|\lambda)\, p(\lambda|X)\, d\lambda = H(\tau', n_0') \underbrace{\int \exp\left[\langle \lambda, \tau' \rangle - n_0' A(\lambda) + \langle \lambda, \phi(x^*) \rangle - A(\lambda)\right] d\lambda}_{\left[H(\tau' + \phi(x^*), n_0' + 1)\right]^{-1}}$$

$$= \frac{H(\ldots)}{H(\ldots)}.$$

$$p(\lambda|\tau, n_0) = H(\tau, n_0)\exp(\langle\lambda, \tau\rangle - n_0 A(\lambda))$$

Hint: 1. $\nabla p(\lambda|\tau, n_0) = \ldots ?$

2. $\int \nabla p(\lambda|\tau, n_0)\, d\lambda =$

$= \nabla \int p(\lambda|\tau, n_0)\, d\lambda = \nabla 1 = 0.$

1. $\nabla p(\lambda|\tau, n_0) = p(\lambda|\tau, n_0)\left[\tau - n_0 \nabla A(\lambda)\right]$

1.+2. $\Rightarrow \int p(\lambda|\tau, n_0)\left[\tau - n_0 \nabla A(\lambda)\right] = 0$

$$\frac{\tau}{n_0} = \langle \nabla A(\lambda)\rangle_{p(\lambda|\tau, n_0)} \rightarrow \text{mean parametrization}$$

$\tau' \rightarrow \tau + \sum \phi$

$n_0' \rightarrow n_0 + N$

$\Rightarrow \dfrac{\tau'}{n_0'} = \dfrac{\tau + \sum\phi}{n_0 + N}$

$$\frac{\tau + \sum_{n=1}^{N} \phi(x_n)}{n_0 + N} = \langle \nabla A(\lambda) \rangle_{p(\lambda|X)}$$

$$\nabla A(\lambda) = \nabla \log \int \exp(\langle \lambda, \phi(x) \rangle)\, dx =$$

$$= \frac{1}{\int \exp(\langle \lambda, \phi(x) \rangle)\, dx} \int \exp(\langle \lambda, \phi(x) \rangle)\, \phi(x)\, dx = \langle \phi(x) \rangle_{p(x|\lambda)}$$

MLE $\frac{N}{n_0+N} \frac{1}{N} \sum_{n=1}^{N} \phi(x_n)$

prior + $\frac{\tau}{n_0} \cdot \frac{n_0}{n_0+N}$

$$\Rightarrow \frac{\tau + \sum_{n=1}^{N} \phi(x_n)}{n_0 + N} = \mathbb{E}_{p(x|\lambda)p(\lambda|X)} \phi(x)$$