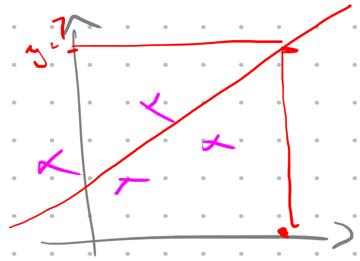
$$\underbrace{p(\theta|D)}_{\text{posterior}} = \frac{\overbrace{p(\theta)}^{\text{prior}}\ \overbrace{p(D|\theta)}^{\text{likelihood}}}{p(D)}$$

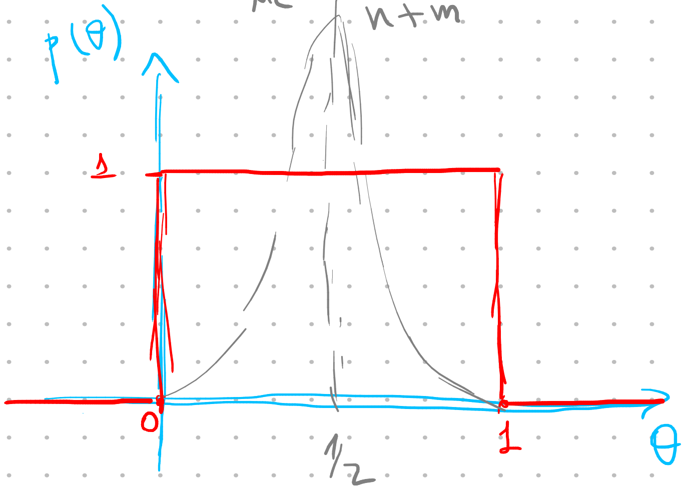
$$p(\theta|D) \propto p(\theta)\,p(D|\theta)$$



$D = hhtht$ = „n орлов, m решек"

$$p(D|\theta) = \theta^n (1-\theta)^m \xrightarrow{\theta} \max$$

$$\theta_{ML} = \frac{n}{n+m}$$



$$p(\theta) = \begin{cases} 1, & \theta \in [0,1] \\ 0, & \text{—} \end{cases}$$

$$p(\theta|D) \propto \begin{cases} \theta^n (1-\theta)^m, & \theta \in [0,1] \\ 0, & \theta \notin [0,1] \end{cases}$$

$$\theta_{MAP} = \arg\max_{\theta} p(\theta|D) = \frac{n}{n+m}$$

$$p(D) = \int p(\theta)\,p(D|\theta)\,d\theta = \int_0^1 \theta^n (1-\theta)^m\,d\theta = B(n+1, m+1)$$

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$B(n+1, m+1) = \frac{n!\,m!}{(n+m+1)!} = \frac{\Gamma(n+1)\Gamma(m+1)}{\Gamma(n+m+2)}$$

$$\underline{p(\theta|D)} = \begin{cases} \dfrac{(n+m+1)!}{n!\,m!}\,\theta^n (1-\theta)^m, & \theta \in [0,1] \\ 0, & \theta \notin [0,1] \end{cases}$$

$p(x|D)$ — predictive distribution

$$p(x|D) = \int p(x,\theta|D)\,d\theta = \int \overbrace{p(x|\theta,\cancel{D})}\, p(\theta|D)\,d\theta \approx \mathbb{E}_{q(\theta)}[p(x|\theta)]$$

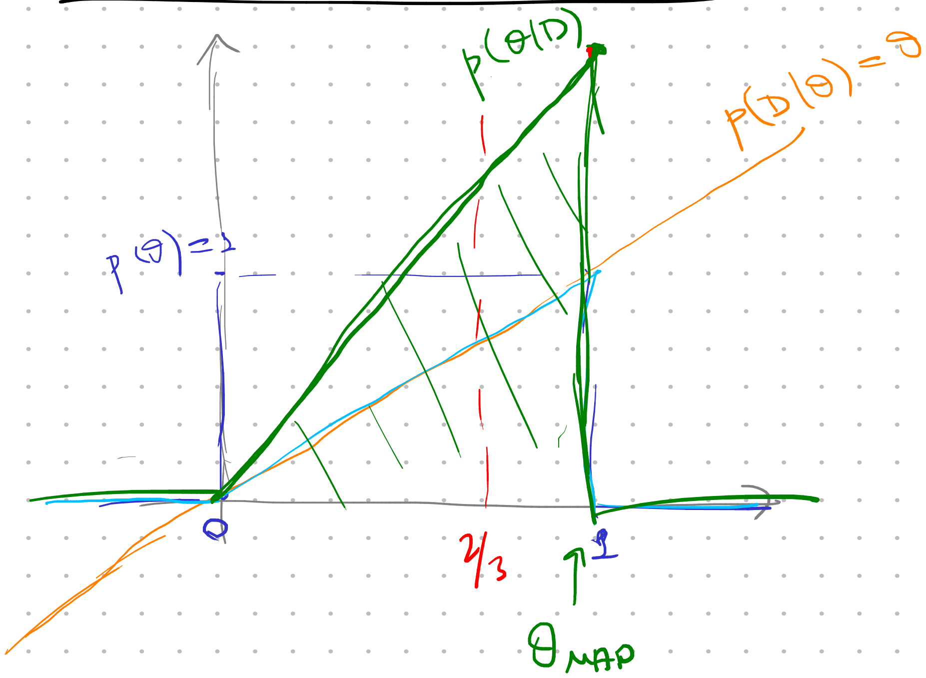
$$\boxed{p(x|D) = \int \underbrace{p(x|\theta)}_{\text{likelihood}}\ \underbrace{p(\theta|D)}_{\text{posterior}}\,d\theta} = \mathbb{E}_{p(\theta|D)}[p(x|\theta)] \approx \frac{1}{K}\sum_{k=1}^{K} p(x|\theta^{(k)})$$

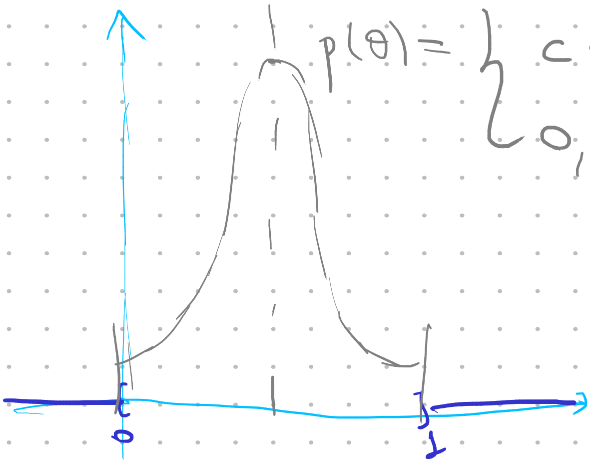
$$p(\text{open}|D) = \int p(\text{open}|\theta)\, p(\theta|D)\,d\theta =$$

$$= \int_0^1 \theta \cdot \frac{(n+m+1)!}{n!\,m!}\,\theta^n (1-\theta)^m\,d\theta =$$

$$= \frac{(n+m+1)!}{n!\,m!}\int_0^1 \theta^{n+1}(1-\theta)^m\,d\theta = \frac{\cancel{(n+m+1)!}}{n!\,\cancel{m!}} \cdot \frac{(n+1)!\,\cancel{m!}}{\cancel{(n+m+2)!}}$$

$$\boxed{p(\text{open}|D) = \frac{n+1}{n+m+2}}$$ — правило Лапласа





$$p(\theta) = \begin{cases} c \cdot e^{-\frac{1}{2\sigma^2}(\theta - 1/2)^2}, & \theta \in [0,1] \\ 0, & \text{———} \end{cases}$$

$$p(\theta|D) \propto p(\theta)\, p(D|\theta) =$$

$$= c \cdot e^{-\frac{1}{2\sigma^2}(\theta - \frac{1}{2})^2} \cdot \theta^n (1-\theta)^m$$

Conjugate priors

$$\boxed{p(\theta|\alpha)} \quad \times \quad p(D|\theta) \quad \propto \quad \boxed{p(\theta|D)}$$

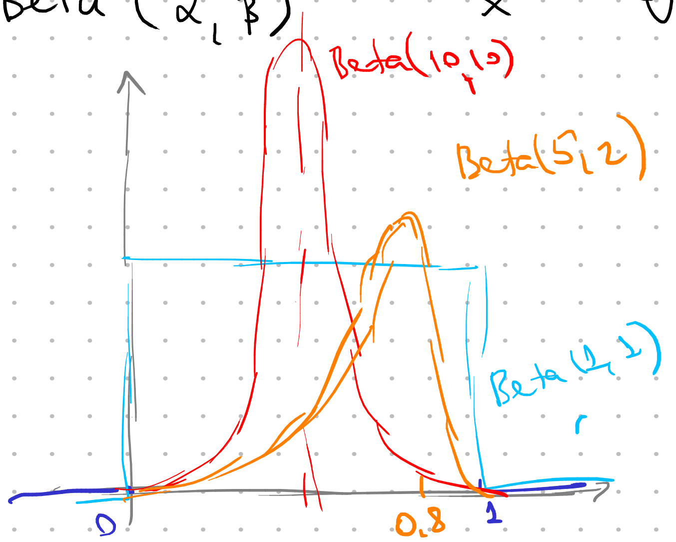
$$? \quad \times \quad \theta^n(1-\theta)^m \quad \propto \quad ?\ \ p(\theta|\alpha')$$

$$\theta^{\alpha-1}(1-\theta)^{\beta-1} \quad \times \quad \theta^n(1-\theta)^m \quad \propto \quad \theta^{(\alpha+n)-1}(1-\theta)^{(\beta+m)-1}$$

$$\mathrm{Beta}(\theta|\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad \theta \in [0,1]$$

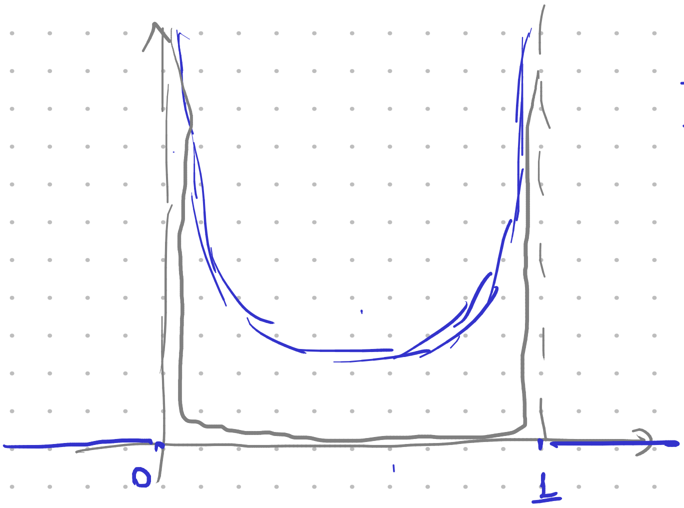
equivalent sample size

$$\mathrm{Beta}(\alpha,\beta) \quad \times \quad \theta^n(1-\theta)^m \quad \propto \quad \mathrm{Beta}(\alpha+n, \beta+m)$$



$$p(\theta)\, p(D|\theta) \propto \boxed{p(\theta|D)} \times p(D'|\theta) \propto p(\theta|D,D')$$

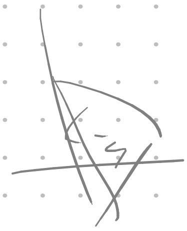
$$\frac{1}{\sqrt{\theta(1-\theta)}} \times \theta^{n}(1-\theta)^{m} \propto \theta^{n-\frac{1}{2}}(1-\theta)^{m-\frac{1}{2}}$$



$\mathrm{Beta}\left(\frac{1}{2}, \frac{1}{2}\right)$ $\mathrm{Beta}\left(\frac{1}{100}, \frac{1}{100}\right)$

$$p(x=i) = \theta_i$$

$$1-\sum_{1}^{k-1}\theta_i$$



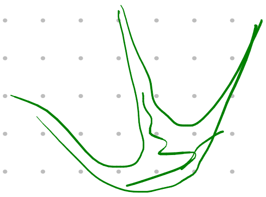
$D = 1\ 2\ 2\ 4\ 1\ 6$

$$p(D|\bar{\theta}) = \theta_1^{n_1}\, \theta_2^{n_2} \ldots \theta_k^{n_k}$$

$\times$

$$p(\bar{\theta}|\bar{\alpha}) = \mathrm{Dir}(\bar{\theta}|\bar{\alpha}) = \frac{1}{\mathrm{Dir}(\bar{\alpha})}\, \theta_1^{\alpha_1-1}\, \theta_2^{\alpha_2-1} \ldots \theta_k^{\alpha_k-1}$$

$$p(\bar{\theta}|D) = \frac{1}{\mathrm{Dir}(\bar{\alpha}+\bar{n})}\, \theta_1^{\alpha_1+n_1-1} \ldots \theta_k^{\alpha_k+n_k-1}$$



$$p(\bar{\theta}) = \mathrm{Dir}\left(\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10}\right)$$

Sparsity

topic models