posterior prior likelihood

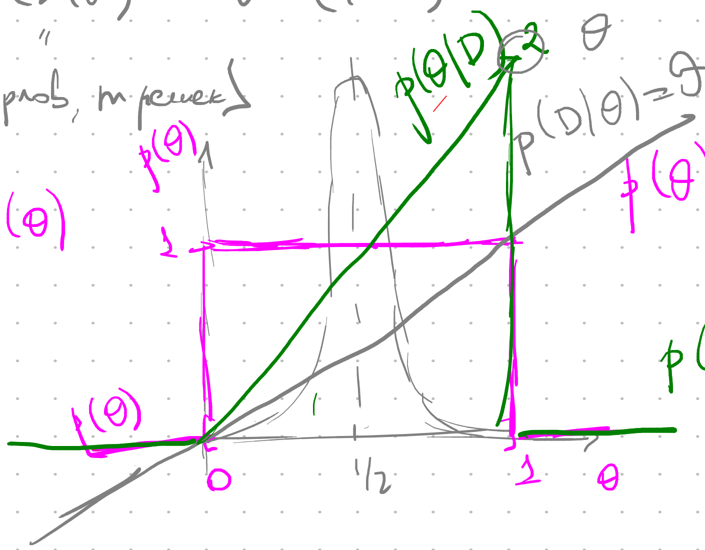
$$\underbrace{p(\bar\theta|D)}_{\text{posterior}} = \frac{\overbrace{p(\bar\theta)}^{\text{prior}}\overbrace{p(D|\bar\theta)}^{\text{likelihood}}}{\underbrace{p(D)}_{\text{evidence}}} \qquad p(D|\bar\theta) \xrightarrow[\bar\theta]{} \max, \quad \bar\theta_{ML}$$

$\bar\theta$ max $\bar\theta_{MAP}$

$$p(\bar x|D) = \int p(\bar x|\bar\theta)\, p(\bar\theta|D)\, d\bar\theta = \mathbb{E}_{p(\bar\theta|D)}\left[p(\bar x|\bar\theta)\right]$$

$$p(D|\theta) = \theta^n (1-\theta)^m \xrightarrow[\theta]{} \max \qquad \theta_{ML} = \frac{n}{n+m}$$

{n орлов, m решек}



$p(\theta)$

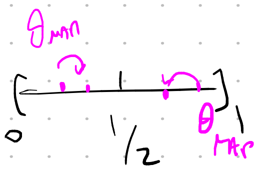
$$p(\theta) = \begin{cases} 1, & \theta \in [0,1] \\ 0, & \text{—''—} \end{cases}$$

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

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

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

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



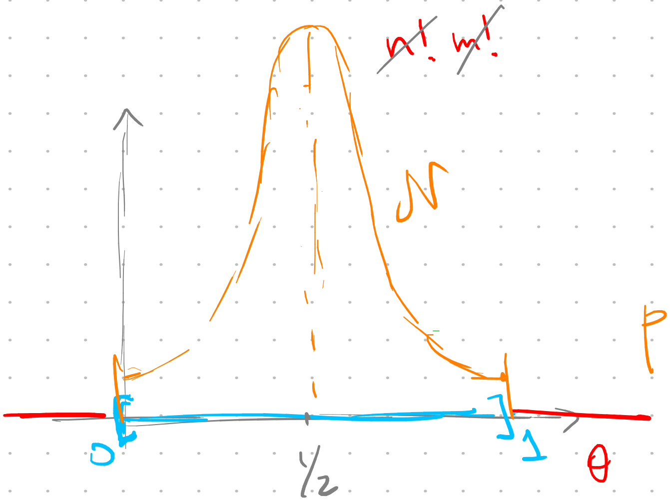
$$p(\text{орёл}|D) = \int p(\text{орёл}|\theta)\, p(\theta|D)\, d\theta =$$

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

$$= \frac{\cancel{(n+m+1)!}}{\cancel{n!}\,\cancel{m!}} \cdot \frac{(n+1)\cancel{!}\, \cancel{m!}}{(n+m+2)\cancel{!}} = \frac{n+1}{n+m+2}$$

Bayesian smoothing

Laplace's rule of succession



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

$$p(\theta) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(\theta - \frac{1}{2})^2} \\ 0, & \text{—''—} \end{cases} \qquad \int_0^1 \dots d\theta$$

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

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

Beta-распр.

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

hyperparameters

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

Conjugate priors: $p(\bar\theta|\bar\alpha)$ is a conjugate prior family for $p(D|\bar\theta)$ if $\forall D \ \forall \bar\alpha$

$$p(\bar\theta|D,\bar\alpha) = \frac{p(\bar\theta|\bar\alpha)\, p(D|\bar\theta)}{\int p(\bar\theta|\bar\alpha)\, p(D|\bar\theta)\, d\bar\theta} = p(\bar\theta|\bar\alpha')$$

$$\bar\theta = (\theta_1, \theta_2, \ldots, \theta_k = 1-(\theta_1+\ldots+\theta_{k-1}))$$

Topic modeling

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

$$\{n_1, n_2, \ldots, n_k\}$$

$$\times$$

$$\mathrm{Dir}(\bar\theta|\bar\alpha) = p(\bar\theta|\bar\alpha) = \frac{1}{\mathrm{Dir}(\bar\alpha)} \cdot \theta_1^{\alpha_1-1}\theta_2^{\alpha_2-1}\ldots\theta_k^{\alpha_k-1}, \quad \theta_i \in [0,1], \ \sum \theta_i = 1$$

$$p(\bar\theta|D) \propto \mathrm{Dir}(\bar\theta|\bar\alpha+\bar n)$$

$$p(i|D) = \frac{n_i+1}{\sum n_j + k}$$

t - test , d - disease

$$p(d \mid t) = \frac{\overset{0.95 \times 0.01 \,\sim\, 0.01}{p(t \mid d)\, p(d)}}{\underbrace{p(t \mid d)\, p(d)}_{0.95 \times 0.01 \,\sim\, 0.01} + \underbrace{p(t \mid \bar{d})\, p(\bar{d})}_{0.05 \times 0.99 \,\sim\, 0.05}} \approx 1/6$$