x – наблюдаеми променливи, t – скрити

вероятностни модели

$p(T|X,\Theta)$, $p(X,T|\Theta)$

discriminative generative

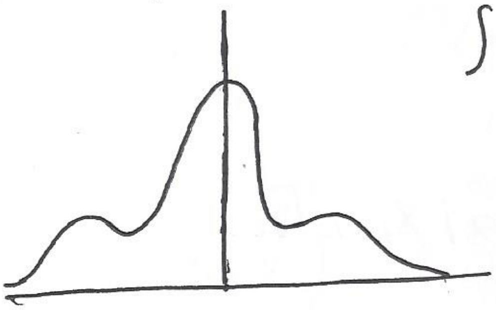
(X_{tr}, T_{tr})

$p(T,\Theta|X)$ $p(X,T,\Theta)$

$$p(T,\Theta|X) = p(T|X,\Theta)\cdot p(\Theta|X) = p(T|X,\Theta)\cdot p(\Theta)$$

Stage	Given	Unknown	Wanted (Bayes)	wanted (Freq.)
Training	X_{tr}, T_{tr}	Θ	$p(\Theta \mid X_{tr}, T_{tr})$	$\Theta_{ML} = \arg\max_{\Theta}(p(T_{tr} \mid X_{tr},\Theta))$
Testing	X_{test}	T_{test}	$p(T_{test} \mid X_{test}, X_{tr}, T_{tr}) = \int_{\Theta} p(T_{test} \mid X_{test},\Theta)\cdot p(\Theta \mid X_{tr}, T_{tr})d\Theta$	$p(T_{test} \mid X_{test},\Theta_{ML})$

$$p(\Theta|X_{tr},T_{tr}) = \frac{p(T_{tr},\Theta|X_{tr})}{\int p(T_{tr},\Theta|X_{tr})d\Theta} = \frac{p(T_{tr}|X_{tr},\Theta)p(\Theta)}{\int p(T_{tr}|X_{tr},\Theta)p(\Theta)d\Theta}$$



Θ_{MP}

$$\Theta_{MP} = \arg\max_{\Theta} p(\Theta|X_{tr},T_{tr})$$

$$p(\Theta|X_{tr},T_{tr}) \simeq \delta(\Theta-\Theta_{MP})$$

$$p(T_{test}|X_{test},X_{tr},T_{tr}) = \int p(T_{test}|X_{test},\Theta)\,\delta(\Theta-\Theta_{MP})d\Theta =$$

$$= p(T_{test}|X_{test},\Theta_{MP})$$

1) $p_1(T,\theta|X) = p(T|X,\theta)\,p_1(\theta)$
$p_2(T,\theta|X) = p(T|X,\theta)\,p_2(\theta)$

задача выбора модели

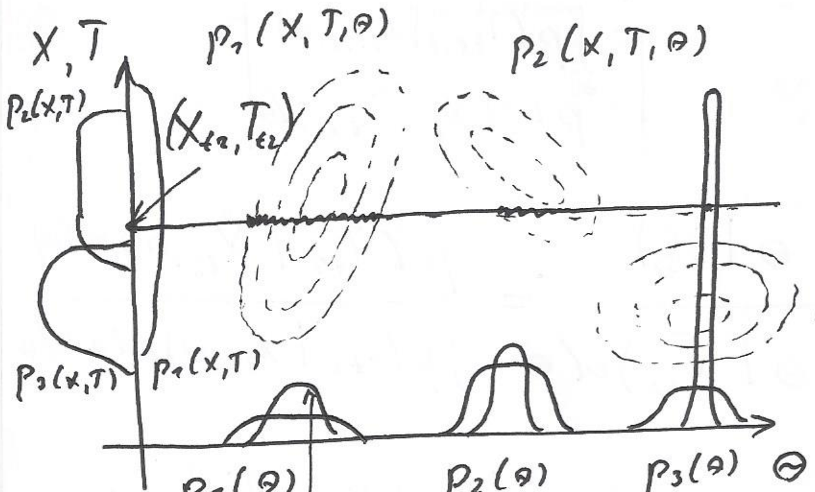
2) $p(X,T|\theta)\,p_1(\theta)$
$p(X,T|\theta)\,p_2(\theta)$

1) $\hat{j} = \arg\max_j p_j(T_{t2}|X_{t2}) = \arg\max_j \int p(T_{t2}|X_{t2},\theta)\,p_j(\theta)\,d\theta$

2) $\hat{j} = \arg\max_j p_j(X_{t2},T_{t2}) = \arg\max_j \int p(X_{t2},T_{t2}|\theta)\,p_j(\theta)\,d\theta$

$$p(\theta|X_{t2},T_{t2}) = \frac{p(X_{t2},T_{t2}|\theta)\,p(\theta)}{\underbrace{\int p(X_{t2},T_{t2}|\theta)\,p(\theta)\,d\theta}_{\text{evidence}}}$$

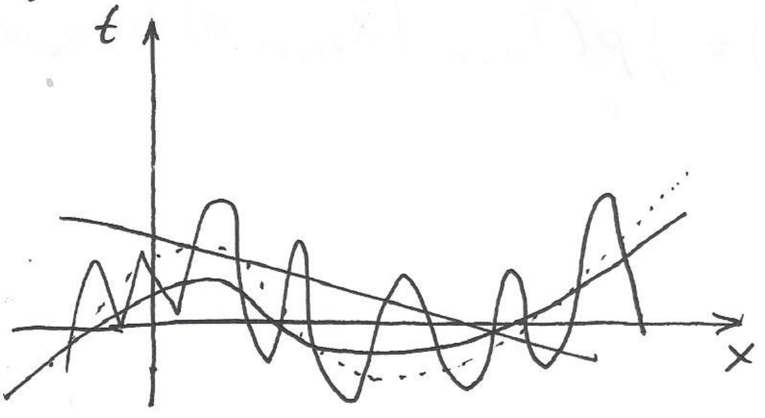
метод наибольшей обоснованности.



$p_1(X,T,\theta)$
$p_2(X,T,\theta)$
$p_3(X,T,\theta)$

$p_1(\theta|X_{t2},T_{t2}),\ p_2(\theta|X_{t2},T_{t2}),\ p_3(\theta|X_{t2},T_{t2})$

одномерная задача регрессии



убийцы, жертвы

$$m = \begin{cases} 0 - \text{белый} \\ 1 - \text{чёрный} \end{cases}, \quad v = \begin{cases} 0 - \text{белый} \\ 1 - \text{чёрный} \end{cases}$$

$$d = \begin{cases} 1, \text{ казнь} \\ 0, \text{ тюрьма} \end{cases}$$

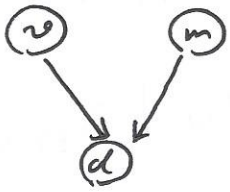
$m=0$

	$d=1$	$d=0$
$v=0$	19	132
$v=1$	0	9

$m=1$

	$d=1$	$d=0$
$v=0$	11	52
$v=1$	6	97

1) $p(d=1) = \theta = p(d=1 \mid v, m) \;\; \forall v, m$



2) $p(d=1 \mid v=0) = \alpha$

$p(d=1 \mid v=1) = \beta$

$p(d=1 \mid v=0) \approx 14\%$

$p(d=1 \mid v=1) \approx 5\%$

3) $p(d=1 \mid m=0) = \gamma$

$p(d=1 \mid m=1) = \delta$

$p(d=1 \mid m=0) \approx 12\%$

$p(d=1 \mid m=1) \approx 10\%$

4)

$p(d=1 \mid m, v)$	$m=0$	$m=1$
$v=0$	τ	χ
$v=1$	γ	ζ

$p_j(k \mid w, \theta)\, p_j(\theta), \quad p_j \sim U[0,1]$

$$\text{Evidence}_1 = \int \theta^{19}(1-\theta)^{132} C_{151}^{19} \cdot \theta^{0} \cdot (1-\theta)^{9} C_{9}^{0} \cdot \theta^{11}(1-\theta)^{52} C_{63}^{11} \cdot \theta^{6}(1-\theta)^{97} C_{103}^{6}\, d\theta = C \cdot C \cdot C \cdot C \cdot B(37, 291) = C \cdot C \cdot C \cdot C \cdot 2{,}8 \cdot 10^{-51}$$

$$\text{Evidence}_2 = \iint \alpha^{19}(1-\alpha)^{132} C_{151}^{19} \cdot \beta^{0}(1-\beta)^{9} C_{9}^{0} \cdot \alpha^{11}(1-\alpha)^{52} C_{63}^{11} \cdot \beta^{6}(1-\beta)^{97} C_{103}^{6}\, d\alpha\, d\beta = CCCC \cdot B(\ldots) \cdot B(\ldots) = C \cdot C \cdot C \cdot C \cdot 4{,}7 \cdot 10^{-51}$$

$$\text{Evidence}_3 = CCCC \iint \gamma^{19}(1-\gamma)^{132} \cdot \gamma^{0}(1-\gamma)^{9} \cdot \delta^{11}(1-\delta)^{52} \cdot \delta^{6}(1-\delta)^{97}\, d\gamma\, d\delta = CCCC \cdot 0{,}27 \cdot 10^{-51}$$

$$\text{Evidence}_4 = CCCC \iiiint \tau^{19}(1-\tau)^{132} \gamma^{0}(1-\gamma)^{9} \chi^{11}(1-\chi)^{52} \zeta^{6}(1-\zeta)^{97}\, d\tau\, d\gamma\, d\chi\, d\zeta = CCCC \cdot 0{,}18 \cdot 10^{-51}$$

16.09.16 Бимо семинар

выборка (T, X)

$$p(\theta, T \mid X, \alpha) = p(T \mid X, \theta)\, p(\theta \mid \alpha)$$

α – семейство моделей

$$-\sum_{i=1}^{n} (\theta^T x_i - t_i)^2 - \frac{1}{2\alpha}\|\theta\|^2 + const(\alpha)$$

$$p(T_{t2} \mid X_{t2}, \alpha) = \int p(T_{t2}, \theta \mid X_{t2}, \alpha)\, d\theta \to \max_{\alpha}$$

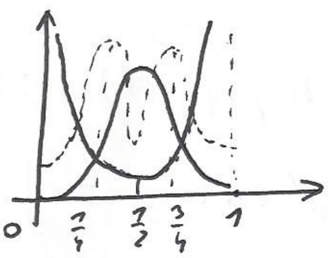
$$p(\theta \mid X_{t2}, T_{t2}, \alpha) = \frac{p(T_{t2}, \theta \mid X_{t2}, \alpha)}{p(T_{t2} \mid X_{t2}, \cancel{X_{ts}}, \cancel{T_{ts}}, \alpha)}$$

$X_{ts},\ T_{ts}$

$$p(T_{ts} \mid X_{ts}, X_{t2}, T_{t2}, \alpha) = \int p(T_{ts} \mid X_{ts}, \theta) \cdot p(\theta \mid X_{t2}, T_{t2}, \alpha)\, d\theta$$

$$p(\theta, T, \alpha \mid X) = p(\theta, T \mid X, \alpha)\, p(\alpha) = p(T \mid X, \theta) \cdot p(\theta \mid \alpha) \cdot p(\alpha)$$

$$p(T_{ts} \mid X_{ts}, X_{t2}, T_{t2}) = \int p(T_{ts} \mid X_{ts}, X_{t2}, T_{t2}, \alpha) \cdot$$
$$\cdot\, p(\alpha \mid X_{t2}, T_{t2})\, d\alpha$$



\# $p(k \mid N, q) = C_N^k q^k (1-q)^{N-k}$

$$p(q \mid \alpha) = \sum_i w_i^\alpha B(q \mid a_i^\alpha, b_i^\alpha)$$

~~$\alpha \in \{1, 2, 3\}$~~ $\alpha = \overline{1,3}$

$$p(k, q \mid N, \alpha) = p(k \mid N, q) \cdot p(q \mid \alpha)$$

$$p(k \mid N, \alpha) = \int p(k, q \mid N, \alpha)\, dq = \int p(k \mid N, q)\, p(q \mid \alpha)\, dq =$$

$$= \int C_N^k q^k (1-q)^{N-k} \Big[\sum_i w_i^\alpha \frac{1}{B(a_i^\alpha, b_i^\alpha)} q^{a_i^\alpha - 1} (1-q)^{b_i^\alpha - 1}\Big] dq =$$

$$= C_N^k \sum_i w_i^\alpha \frac{1}{B(a_i^\alpha, b_i^\alpha)} \int q^{k + a_i^\alpha - 1} (1-q)^{N-k+b_i^\alpha - 1}\, dq =$$

стр 4

$$= C_N^k \sum_i w_i^\alpha B(k + a_i^\alpha, N - k + b_i^\alpha) / B(a_i^\alpha, b_i^\alpha)$$

$$p(q|k,N,\alpha) \overset{p(q|\alpha)\cdot}{=} \frac{p(k,q|N,\alpha)}{p(k|N,\alpha)} \propto C_N^k q^k (1-q)^{N-k} \cdot$$

$$\cdot \sum_i w_i^\alpha B(q|a_i^\alpha, b_i^\alpha) = C_N^k \sum_i w_i \frac{1}{B(a_i^\alpha, b_i^\alpha)} q^{k+a_i^\alpha-1} (1-q)^{N+b_i^\alpha-1-k} =$$

$$= C_N^k \sum_i \frac{w_i}{B(a_i^\alpha, b_i^\alpha)} \frac{q^{k+a_i^\alpha-1}(1-q)^{N+b_i^\alpha-k-1}}{B(a_i^\alpha+k, N-k+b_i^\alpha)} B(a_i^\alpha+k, N-k+b_i^\alpha) \propto$$

$$\propto C_N^k \sum_i \frac{w_i^\alpha \cancel{B(\alpha a_i} \, Beta(q|a_i^\alpha+k, N-k+b_i^\alpha)}{B(a_i^\alpha, b_i^\alpha) \cdot B(a_i^\alpha+k, N-k+b_i^\alpha)}$$

$$Beta(q|a,b) = \frac{1}{B(a,b)} q^{a-1}(1-q)^{b-1}$$

N_1, k_1

$$p(k_1|N_1,k,N,\alpha) = \int p(k_1|N_1,q)\, p(q|N,k,\alpha)\, dq =$$

$$= \int C_{N_1}^{k_1} q^{k_1}(1-q)^{N_1-k_1} \cdot \sum_i w_i^\alpha Beta(q|\tilde{a}_i^\alpha, \tilde{b}_i^\alpha)\, dq =$$

$$= \int C_{N_1}^{K_1} \sum_i \frac{w_i^\alpha}{B(\tilde{a}_i^\alpha, \tilde{b}_i^\alpha)} \frac{q^{k_1+\tilde{a}_i^\alpha-1}(1-q)^{N_1-k_1+\tilde{b}_i^\alpha-1}}{B(k_1+\tilde{a}_i^\alpha, N_1-k_1+\tilde{b}_i^\alpha)} B(..)\, dq =$$

$$= C_{N_1}^{K_1} \cancel{\int} \sum_i w_i \int \frac{Beta(q|\tilde{a}_i^\alpha+k, \tilde{b}_i^\alpha+N_1-k_1)}{B(\tilde{a}_i^\alpha, \tilde{b}_i^\alpha)} \cdot B(k_1+\tilde{a}_i^\alpha, N_1-k_1+\tilde{b}_i^\alpha)\, dq =$$

$$= C_{N_1}^{K_1} \sum_i w_i \frac{Beta(q|\tilde{a}_i^\alpha+k, \tilde{b}_i^\alpha+N_1-k)}{B(\tilde{a}_i^\alpha, \tilde{b}_i^\alpha)}$$

$$p(q|\alpha) = \sum_i w_i B(q|a^\alpha, b^\alpha), \quad p(\alpha) = \frac{1}{3}$$

$$p(\alpha|N,k) = \frac{p(\alpha,k|N)}{p(k|N)} \propto, \quad p(k,q,\alpha|N) = p(k|N,q)\cdot p(q|\alpha) p(\alpha)$$

$$\propto \frac{\int p(k|N,q)\, p(q|\alpha)\, p(\alpha)\, dq}{\int\int p(k|N,q)\, p(q|\alpha)\, p(\alpha)\, dq\, d\alpha}$$

cmp 5

$$p(k_1 \mid N_1, k, N) = \int p(k_1 \mid N_1, N, k, \alpha)\, p(\alpha \mid k, N)\, d\alpha$$

$$p(n_1, \dots, n_6 \mid \Theta) = ?$$

$$C^N_{x_1 \dots x_\ell} = \frac{N!}{x_1! \cdot \ldots \cdot x_\ell!}$$

$$p(n_1 \dots n_\ell \mid \Theta) = \frac{1}{Z} \prod_{i=1}^{n} q_i^{\ell}$$

$$p(x \mid q, N) = C^N_{x_1 \dots x_\ell} \prod_{i=1}^{\ell} q_i^{x_i}$$

$$\mathcal{D}(q \mid \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{\ell} q^{\alpha_i - 1}$$

$$B(\alpha) = \frac{\prod_{i=1}^{\ell} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{\ell} \alpha_i\right)}$$

$$p(\bar{q} \mid \bar{x}, N, \alpha) \propto p(\bar{x} \mid \bar{q}, N) \cancel{\propto} p(\bar{q} \mid \alpha) \propto Dir(\alpha_1 + x_1, \dots, \alpha_\ell + x_\ell)$$

③ N городов, в городе i население N_i, x_i умерло из N_i, Θ_i - ур-нь смертности в i-м городе

$$p(x, \Theta \mid N, \alpha) = \prod_{i=1}^{N} p(x_i \mid N_i, \Theta_i)\, p(\Theta_i \mid \alpha)$$

$$p(x_i \mid N_i, \Theta_i) = C^{x}_{N_1}$$