Дмитрий Петрович Ветров

БММО лекция

02.09.16

$$p(x|y) \overset{def}{=} \frac{p(x,y)}{p(y)}$$

Conditional {условное} = $\dfrac{\text{Joint}}{\text{Marginal}}$ {совместное}

~~(Prior)~~ ~~{априорное}~~

$$p(x,y) = p(x)\cdot p(y)$$

$$p(y|x)\cdot p(x) = p(x,y) = p(x|y)\cdot p(y)$$

Product rule:

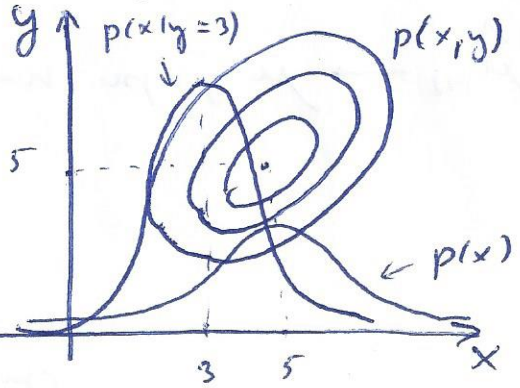
$$p(x_1, x_2, \dots, x_n) = p(x_n | x_1 \dots x_{n-1}) \cdot \dots \cdot p(x_2|x_1)\cdot p(x_1)$$

$$p(y|x) = \frac{p(x|y)\cdot p(y)}{p(x)}$$ {формула Байеса}

$$\int p(y)\,dy = 1, \quad \int p(y|x)\,dy = 1, \quad \int p(y|x)\,dx \geq 0$$

$$p(y|x) = \frac{p(x|y)\cdot p(y)}{p(x)} \;\Big|\int dy$$

$$p(x) = \int \overbrace{p(x|y)p(y)}^{p(x,y)}\,dy = \mathbb{E}_y\, p(x|y) = \int p(x,y)\,dy$$



p(x,y)
p(x|y=3)
p(x)

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{p(y|x)\cdot p(x)}{\int p(y|x)p(x)dx}$$

Th Bayes {y известен}

{правдоподобие}

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

{апостериорное распределение} {обоснованность}

	Frequentist частотний подход	Bayesian байесовский подход
Интерпретация случайности	Объективная неопределённость	Субъективное незнание
Величины	Случайные и детерминированные x – сл. в. μ – параметр	Все случайные $x_1, \ldots, x_n \sim N(x\|\mu, 1)$ x, μ – сл. в.
Метод вывода	ML $\frac{k}{n}$	Th Bayes $\frac{k+1}{n+2}$ Коррекция на априорное представление

$$p(X|\mu) = \prod_{i=1}^{n} p(x_i|\mu) \to \max_{\mu}$$

$$\ln p(X|\mu) = \sum_{i=1}^{n} \ln p(x_i|\mu) \to \max_{\mu}$$

$$\hat{\mu}_{ML} = \arg\max \ln p(X|\mu)$$

$\mathbb{E}\,\hat{\mu}_{ML} = \mu \quad \forall\, n \gg 1$, $\hat{\mu}_{ML} \to \mu$, при $n \to \infty$

$\mathbb{D}\,\hat{\mu}_{ML}$ – наименьшая

$$p(X \mid \lambda) = \prod_{i=1}^{n} p(x_i \mid \lambda) \to \max_{\lambda}$$

$$\mathcal{L}(\lambda) = \sum_{i=1}^{n} \ln \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \sum_{i=1}^{n} (\ln e^{-\lambda} + x_i \ln \lambda -$$

$$- \ln x_i) = \sum_{i=1}^{n} (x_i \ln \lambda - \lambda - \ln x_i)$$

$$\mathcal{L}'(\lambda) = \sum_{i=1}^{n} \left(\frac{x_i}{\lambda} - 1 \right) = 0$$

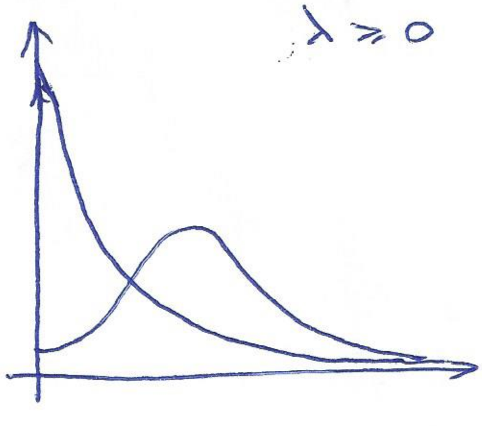
$$\frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0 \quad , \quad \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\# \quad p(\lambda \mid X) = \frac{p(X \mid \lambda)\, p(\lambda)}{p(X)} \ominus$$

$$G(\lambda \mid a, b) = \frac{b^a \cdot \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \qquad a, b > 0, \quad \lambda \geq 0$$

$$\Gamma(a) = \int_0^{\infty} \lambda^{a-1} e^{-\lambda} d\lambda$$

$$\mathbb{E}\xi = \frac{a}{b}, \quad \mathbb{D}\xi = \frac{a}{b^2}$$



$$\ominus \frac{1}{Z} \left(\prod_{i=1}^{n} \left[\frac{e^{-\lambda} \lambda^{x_i}}{(x_i)!} \right] \cdot \lambda^{a-1} e^{-b\lambda} \right) =$$

$$= \frac{1}{Z} \lambda^{a-1} e^{-b\lambda - \lambda n} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{(x_i)!} = \frac{1}{Z} e^{-\lambda(b+n)} \lambda^{a-1} \cdot \lambda^{\sum_{i=1}^{n} x_i} =$$

$$= \frac{1}{Z} e^{-b\lambda - n\lambda} \cdot \lambda^{a-1+\sum_{i=1}^{n} x_i}$$

$$\mathbb{E}\eta = \frac{a + \sum_{i=1}^{n} x_i}{b + n}$$

02.09.16 БММО семинар

x, t, θ sum: $\int p(a, b)\,db = p(a)$

$p(x, t, \theta)$ prod: $p(a, b) = p(a|b)\,p(b)$

Больница

$d \quad t$

$p(d=1) = 10^{-4}$

$p(t=1|d=1) = 99\%$

$p(t=0|d=0) = 99\%$

$p(d=1|t=1)$ - ?

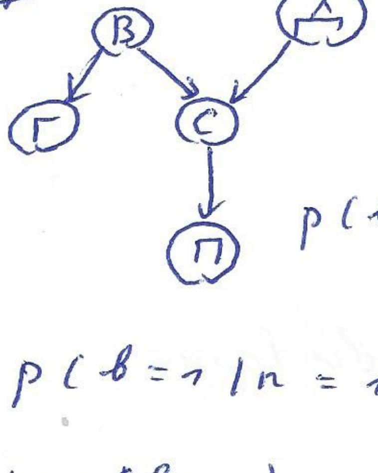
$$p(d=1|t=1) = \frac{p(t=1, d=1)}{p(t=1)} = \frac{p(t=1|d=1)\,p(d=1)}{p(t=1)} \ominus$$

$$p(t=1) = p(t=1|d=0)\cdot p(d=0) + p(t=1|d=1)\,p(d=1)$$

$$p(t=1) = 0{,}99\cdot 10^{-4} + (1-0{,}99)\cdot 10^{-4} = 10^{-4}$$

$$\ominus \frac{(1-0{,}99)\cdot 10^{-4}}{10^{-4}} = 10^{-1} \; ?$$

#



$$p(г, б, с, д, п) = p(б)\,p(д)\,p(г|б)\cdot p(с|б, д)\cdot p(п|с)$$

$$p(б=1|г=1) = \frac{p(г=1|б=1)\,p(б=1)}{\sum_б p(г=1|б)\,p(б)}$$

$$p(б=1|п=1) = \frac{p(п=1|б=1)\,p(б=1)}{\sum_б p(п=1|б)\,p(б)} \ominus$$

$p(п=1, б=1)$

$$= \sum_{г,с,д} p(п=1, г, с, д, б=1) \cancel{p(б=1)} =$$

$$= \sum_{г,с,д} p(б=1)\,p(д)\,p(г|б=1)\,p(с|б=1, д)\cdot p(п=1|с)$$

стр 1

$$p(b=1 \mid r=1, h=1) = \frac{p(b=1, r=1, h=1)}{p(r=1, h=1)} =$$

$$= \frac{\sum_{c,g} p(b=1, r=1, h=1, c, g)}{\sum_{b,c,g} p(r=1, h=1, b, c, g)}$$

$X_1 \sim Poiss(\lambda_1)$

$X_2 \sim Poiss(\lambda_2)$ $\quad p(x_i) = e^{-\lambda_i} \frac{\lambda_i^k}{k!}$

$$p(X_1 + X_2 = k) = \sum_{s=0}^{k} p(X_1 = s) p(X_2 = k-s) =$$

$$= \sum_{s=0}^{k} \frac{e^{-\lambda_1}}{s!} \lambda_1^s \frac{e^{-\lambda_2} \lambda_2^{k-s}}{(k-s)!} = e^{-\lambda_1-\lambda_2} \sum_{s=0}^{k} \frac{k! \cancel{e^{-\lambda_1} e^{-\lambda_2}} \lambda_1^{k} \lambda_2^{s-k}}{k!\, s!(k-s)!} =$$

$$= \frac{e^{-\lambda_1-\lambda_2}}{k!} \sum_{s=0}^{k} C_k^s \cancel{e^{-\lambda_1} e^{-\lambda_2}} \lambda_1^{k} \lambda_2^{s-k} = \frac{e^{-\lambda_1-\lambda_2}}{k!} (\lambda_1 + \lambda_2)^k$$

$$\varphi(t) = \sum_{k=0}^{\infty} t^k \cancel{\varphi} P(X=k) = \sum_{k=0}^{\infty} t^k e^{-\lambda} \frac{\lambda^k}{k!} =$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(t\lambda)^k}{k!} = e^{t\lambda} e^{-\lambda} = e^{\lambda(t-1)}$$

$$\begin{cases} e^{\lambda_1 (t-1)} \\ e^{\lambda_2 (t-1)} \end{cases} \longrightarrow e^{(\lambda_1 + \lambda_2)(t-1)}$$

Правдоподобие — вероятность выборки при условии параметра.

$x_1, \ldots, x_n \sim Poiss(\lambda)$

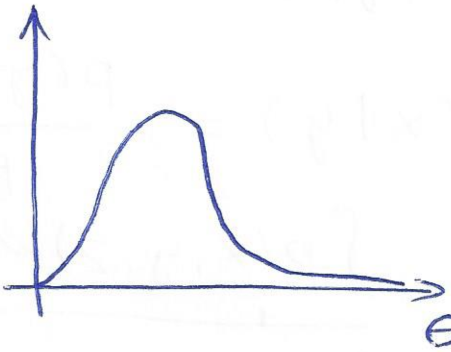
$$p(x_i = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Оценка	Точечные (реже интервальные)	Posteriors (распределения)
Применимость	$n >> 1$ $\{ \frac{n}{k} >> 1 \}$ k – число настраиваемых параметров	$\forall n \in \mathbb{Z}_+$

1) Регуляризация ML

2) Композицируемость

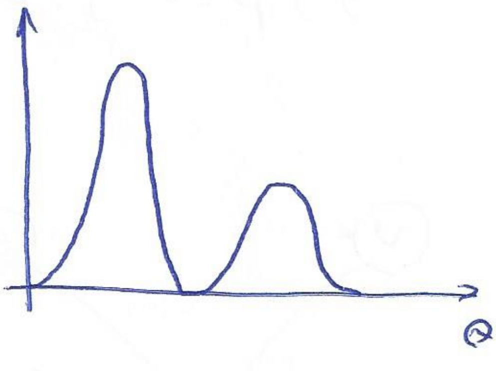
$p(\theta)$ Prior



1) x_1

$$p(\theta | x_1) = \frac{p_1(x_1|\theta) p(\theta)}{\int p_1(x_1|\theta) p(\theta) d\theta}$$

хвост



2) x_2

$$p(\theta | x_1, x_2) = \frac{p_2(x_2|\theta) \cdot p(\theta|x_1)}{\int p_2(x_2|\theta) p(\theta|x_1) d\theta}$$

$$p(z|x,y) \propto \frac{p(z|x) \cdot p(z|y)}{p(z)}$$

3) Стримминг

$p(\theta | x_1, \dots, x_n) \to x_{n+1}$

$$p(\theta | x_1, \dots, x_{n+1}) = \frac{p(x_{n+1}|\theta) \cdot p(\theta | x_1 \dots x_n)}{\int p(x_{n+1}|\theta) \cdot p(\theta|x_1 \dots x_n) d\theta}$$

④ Моделирование с латентами

⑤ Масштабируемость

$$\frac{A,\ A \to B}{B} \qquad \frac{A \to B,\ B}{A - ?} \qquad \frac{p(A),\ p(B|A),\ \cancel{p(B)}}{p(A|B)}$$

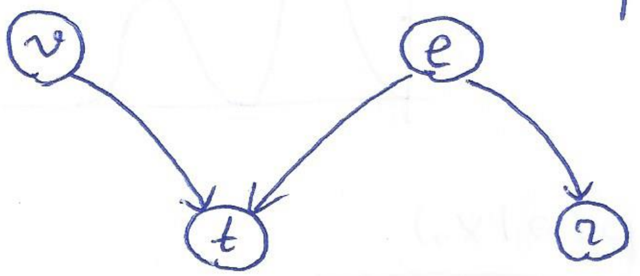
Приближённые квазилогические утверждения

$$p(A|B) = \frac{p(B|A)p(A)}{\int p(B|A)\cdot p(A)\,dA}$$

$p(x, y, z)$

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(x,y)}{p(y)} =$$

$$= \frac{\int p(x,y,z)\,dz}{\iint p(x,y,z)\,dx\,dz}$$



$$p(v,t,e,r) = p(t|v,e)\cdot p(r|e)\cdot p(e)\cdot p(v)$$

$p(v=1) = 2\cdot 10^{-4}$

$p(e=1) = 10^{-2}$

$p(r=1|e=1) = 0.5$

$p(r=1|e=0) = 0$

$p(t=1\|v,e)$	$e=1$	$e=0$
$v=1$	1	1
$v=0$	0.1	0

$$p(v=1 \mid t=1) = \frac{p(t=1 \mid v=1)\, p(v=1)}{\sum_v p(t=1 \mid v)\, p(v)} \; (=)$$

$$p(t \mid v) = \sum_e p(t \mid v, e)\, p(e)$$

$$p(x) = \int p(x \mid y)\, p(y)\, dy$$

$$p(x \mid z) = \int p(x \mid y, z)\, p(y)\, dy$$

$$p(t=1 \mid v=1) = \overbrace{p(t=1 \mid v=1, e=0)}^{1}\, p(e=0) +$$

$$+ \underbrace{p(t=1 \mid v=1, e=1)}_{1}\, p(e=1) = 1$$

$$\overset{1}{p(t=1 \mid v=0)} = p(t=1 \mid v=0, e=0)\, p(e=0) +$$

$$+ p(t=1 \mid v=0, e=1)\, p(e=1) = 10^{-3}$$

$$(=) \; \frac{1 \cdot 2 \cdot 10^{-4}}{2 \cdot 10^{-4} + 10^{-3}(1 - 2 \cdot 10^{-4})} \approx \frac{2}{2+10} = \frac{1}{6} \approx 16\,\%$$

$$p(v) = 2 \cdot 10^{-3} \longrightarrow \frac{2 \cdot 10^{-3}}{2 \cdot 10^{-3} + 10^{-3}(1 - 2 \cdot 10^{-3})} \approx \frac{2 \cdot 10^{-3}}{3 \cdot 10^{-3}} \approx 67\,\%$$

$$p(v=1 \mid r=1, t=1) = \frac{p(v=1, r=1, t=1)}{p(t=1, r=1)} =$$

$$= \frac{\sum_e p(v=1, r=1, t=1, e)}{\sum_{v,e} p(v, t=1, r=1, e)} = \frac{\sum_e p(t=1 \mid v=1, e)\, p(r=1 \mid e)\, p(e)\, p(v=1)}{\sum_{v,e} p(t=1 \mid v, e)\, p(r=1 \mid e)\, p(e) \cdot p(v)} \approx 2\,\%$$

$$p(v=1 \mid e=1) = p(v=1 \mid e=0) \; ?$$