x, y – с.в.

Плотность x:

(i) $p(x) \geq 0$

(ii) $\int_{\mathbb{R}^d} p(x) = 1$

Совместная плотность

$p(x, y)$ через

$p(x \mid y)$ ← пронаблюдали одну из координат

$= \dfrac{p(x,y)}{p(y)}$

$$p(\theta \mid D) = \frac{p(\theta)\, p(D \mid \theta)}{p(D)}$$

model parameter — data

posterior (то, что мы думаем)

prior (то, что мы думали о модели до получения данных)

likelihood (поведение модели при параметре)

это нормировка скорее

Если x, y – дискр., то

$$\frac{p(x,y)}{p(y)} = \frac{\mathbb{P}\{X=x, Y=y\}}{\mathbb{P}\{Y=y\}}$$

Если $p(x,y) = p(x) \cdot p(y)$, то x, y – н.з.

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

отсюда следует, что

$$p(x \mid y) = \frac{p(y \mid x) p(x)}{p(y)} \qquad (*)$$

ОБОБЩЕНИЕ – ПРАВИЛО ПРОИЗВЕДЕНИЯ

$$p(x_1, \ldots, x_n) = p(x_1) \cdot p(x_2 \mid x_1) \cdot p(x_3 \mid x_2, x_1) \cdot \ldots \cdot p(x_n \mid x_1, \ldots, x_{n-1})$$

В ЧАСТНОСТИ, ЕСЛИ

$x_1, \ldots, x_n$ – попарно независимы, то

$$p(x_1, \ldots, x_n) = p(x_1) \cdot p(x_2) \cdot \ldots \cdot p(x_n)$$

"МНОГОМЕРНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ = ПРОИЗВ. ОДНОМЕРНЫХ"

ПРИМЕР – ГЕНЕРАЦИЯ СЛОВ

x_1 x_2 ...

ПРОИНТЕГРИРУЕМ (*):

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

усредняем по всем y

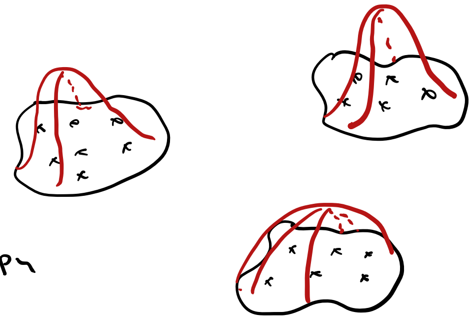
Если вдруг y – дискретна, то

$$\sum_y \mathbb{P}\{Y=y \mid X=x\} = \sum \frac{\mathbb{P}\{Y=y\} \cdot \mathbb{P}\{X=x, Y=y\}}{\mathbb{P}\{X=x\}}$$

k - номер кластера

$X \in \mathbb{R}^2$ - расположение

$P(X, k)$ - непрерывен внутри кластеров



КАК ИСПОЛЬЗУЕТСЯ В МАШ. ОБУЧЕНИИ?

Пусть $P = \{P_\theta\}$ - параметризованное семейство распределений

$X_1 \sim X_2 \sim \dots \sim X_n \sim P_\theta(x)$

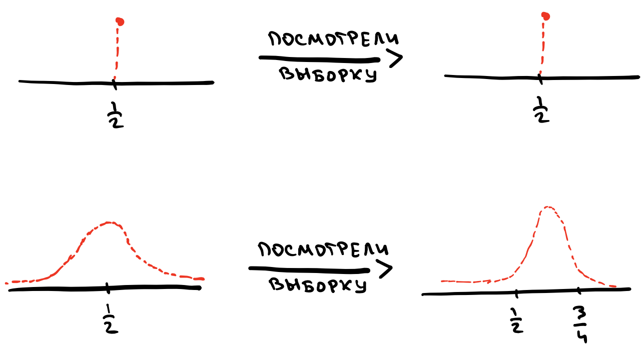
Можем искать максимум правдоподобия

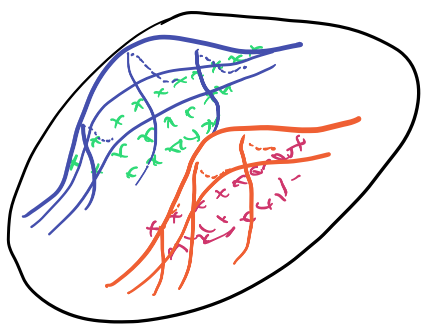
"0111011" $\quad X_i \sim Bin(\theta) \quad$ $\theta = \frac{3}{4}$ - максимум правдоподобия

$$p(\theta \mid D) = \frac{p(\theta) \cdot p(D \mid \theta)}{p(D)}$$

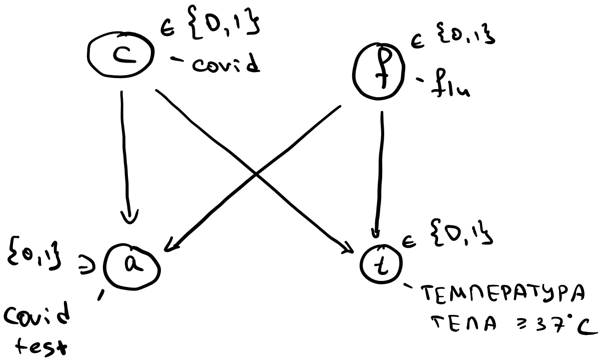
Если $\theta = \delta(x - \frac{1}{2})$ (мы прям уверены, что монетка честная)

то после обзора выборки ничего не изменится





Тест, диагностирует ковид



t, но он не влияет

$$p(a, c, f, t) = p(a \mid c, f) \cdot p(t \mid c, f)\, p(c)\, p(f)$$

$P(a \mid c, f)$	$f = 0$	$f = 1$
$c = 0$	0,05	0,15
$c = 1$	0,95	0,95

$P(t=1 \mid c, f)$	$f = 0$	$f = 1$
$c = 0$	10^{-3}	1
$c = 1$	1	1

$P(c=1) = 80 \cdot 10^{-3} \qquad P(f=1) = 90 \cdot 10^{-4}$

Как найти $\mathbb{P}(c=1 \mid a=1) = \dfrac{\mathbb{P}(c=1, a=1)}{\mathbb{P}(a=1)} =$

$$= \frac{\sum_{t,f} P(c=1, a=1, t, f)}{\sum_{t,f,c} P(a=1, t, f, c)}$$

УЖЕ ЗНАЕМ

$$\mathbb{P}(c=1 \mid a=1) = \frac{P(a=1 \mid c=1) \cdot P(c=1)}{P(a=1)}$$

$$P(a=1 \mid c=1) = \sum_{f=0}^{1} P(a=1, f \mid c=1) = \sum_{f=0}^{1} P(a=1 \mid c=1, f) \cdot P(f)$$