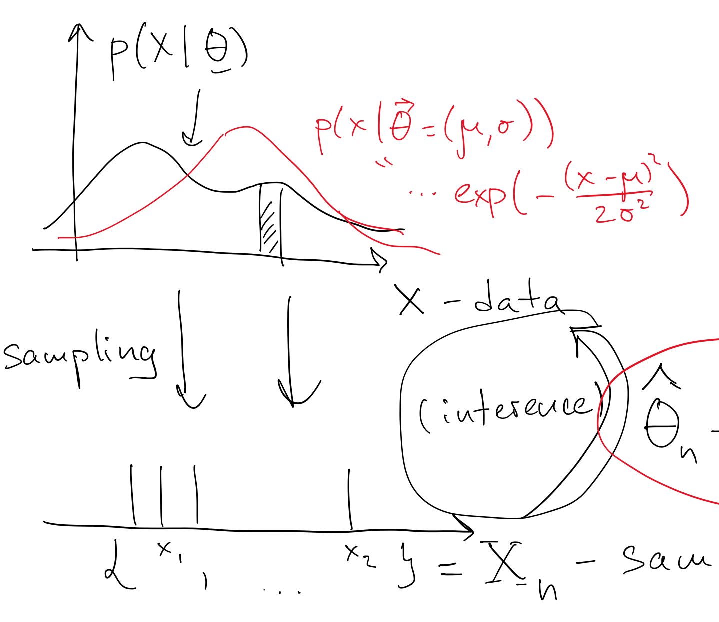
Откус Байес 1

пятница, 26 января 2024 г.   20:03



$p(x|\vec{\theta}=(\mu,\sigma))$ " $\dots \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

X - data

Classical Stat. Inference

(θ - some unknown constant)

(inference) $\hat{\theta}_n$ - estimate of θ

$\{x_1, \dots x_2\} = X_n$ - sample

Bayesian inference

θ - random variable

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

$p(A,B) = $ ($A \cap B$)
$= p(A|B)\cdot p(B)$

posterior

$$p(\theta|X) = \frac{p(X|\theta)\cdot p(\theta)}{p(X)}$$

likelihood: $p(X|\theta)$ prior: $p(\theta)$ evidence: $p(X)$

$p(\theta|X) = p(\theta=\theta|X=x)$

Holy Grail of Bayesian Stats

$X = \{x_1, \dots, x_n\}$ - data

assumption: i.i.d

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

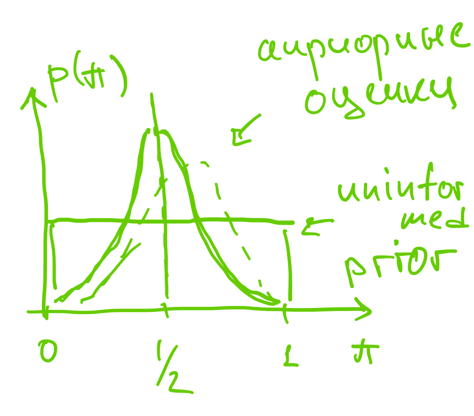
$\hat{\theta}_n$ - ? Maximum Likelihood Estimation (MLE)

$$\hat{\theta}_{MLE} = \arg\max_{\theta} p(X|\theta)$$

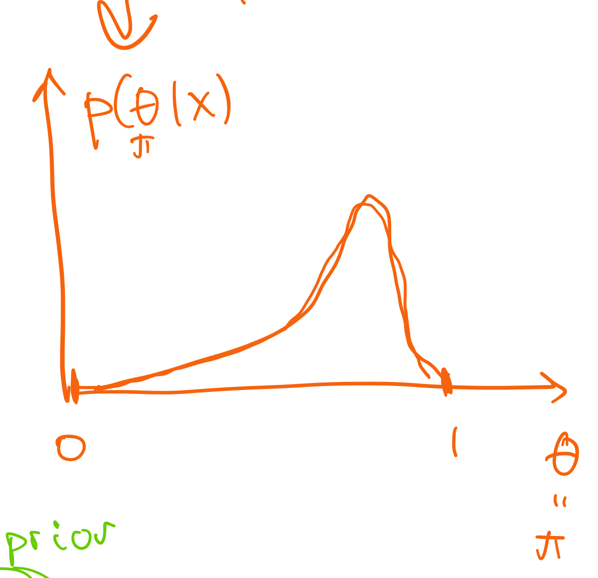
"Bayesian coin"

$\to 1, \ \underset{\pi}{\overset{\theta}{\shortparallel}} \in [0,1]$

$\to 0, \ 1-\pi$



"+X"



$$\underset{\text{posterior}}{p(\theta|X)} = \frac{\overset{\text{likelihood}}{p(X|\theta)}\cdot \overset{\text{prior}}{p(\theta)}}{p(X)}$$

evidence

$p(X) = \int d\theta \ p(X,\theta) = \int p(X|\theta)\cdot p(\theta)$ / Σ

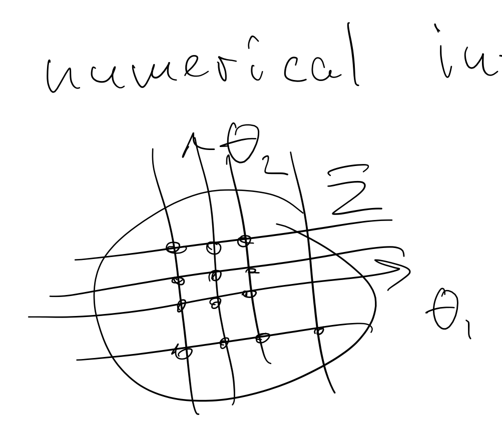
marginalizing θ out

if θ is high-dimensional (e.g. NN)

$\int / \Sigma$ is hard

but we have computers!

1) numerical integration



2) Monte-Carlo

$$p(\theta|X) = \frac{p(X|\theta)\, p(\theta)}{p(X)}$$

does not depend on θ!

$$p(\theta|X) \propto p(X|\theta)\cdot p(\theta)$$

we can sample θ!

$\theta_{initial} \to \theta_{new}$ → accept → $\Theta = \{\theta_{init}, \theta_{new}\}$

→ reject → ...

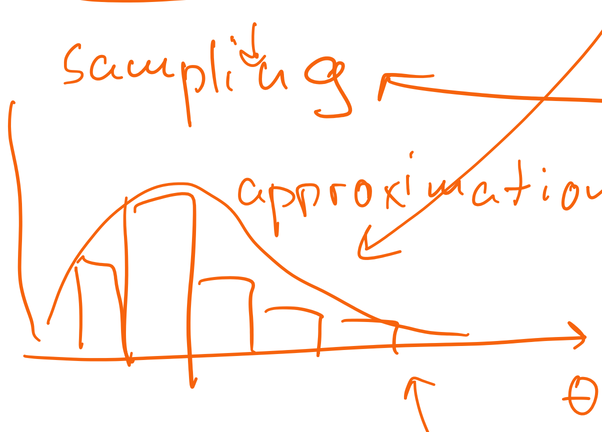
$\Theta = \{\theta_{initaal}\}$

N

$\Theta = \{\theta_1, \theta_2, \dots\} \sim p(\theta|X)$ posterior

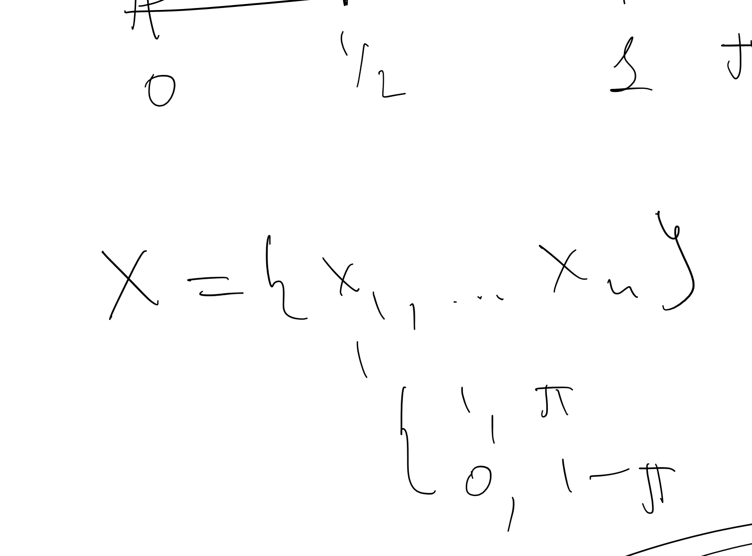
sampling

probabilistic programming (PyMC, Stan)



$\hat{\theta}_{prior} = \cancel{\alpha} \ \mathbb{E}_{p(\theta)} \theta$ - априорная оценка

$\hat{\theta}_{Bayes} = \mathbb{E}_{p(\theta|X)} \theta = \Sigma_{\Theta} \theta_i$ - постериорная оценка



$$p(\pi) \propto \pi^{\alpha-1}\cdot(1-\pi)^{\beta-1}$$

Beta-distr

$\int_0^1$

$X = \{x_1, \dots x_n\}$

$\begin{cases} 1, & \pi \\ 0, & 1-\pi \end{cases}$

$X_n = x_1 + x_2 + \dots + x_n$

$$p(X|\pi) = \binom{n}{k} \pi^k (1-\pi)^{n-k}$$ likelihood

k

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$$p(\pi|X) \propto p(X|\pi)\cdot p(\pi)$$

$$\propto \pi^{(\alpha-1)+k}\cdot(1-\pi)^{(\beta-1)+(n-k)}$$

when (likelihood, prior are such that) posterior $\in$ the same family as the prior

— such likelihood (Binomial) and prior (Beta) are called conjugate

(сопряженные)