Lecture 1

Monday, July 4, 2022 10:59 AM

STATISTICAL METHODS

Roberto Trotta

(SISSA & Imperial College)

Contact me : rtrotta@sissa.it

@R_Trotta

SISSA Theoretical & Scientific Data Science Programme :

datascience.sissa.it

Inference :

N balls inside.

W white

B black

1 random draw : what is the prob. that the ball is W ?

$$P(W) = \frac{W}{W+B}$$

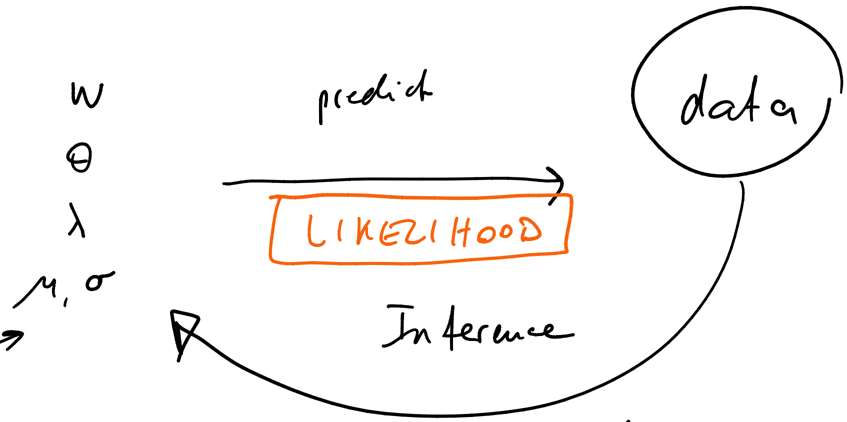
Coin : biased prob. of heads $\theta > \frac{1}{2}$

N tosses. What is the prob. distr. for H ?

$$P(H \mid N, \theta) = \binom{N}{H} \theta^{H} (1-\theta)^{N-H} \quad \text{(BINOMIAL)}$$

Forward prob : given parameter $\xrightarrow{\text{predict}}$ outcome

Inference :



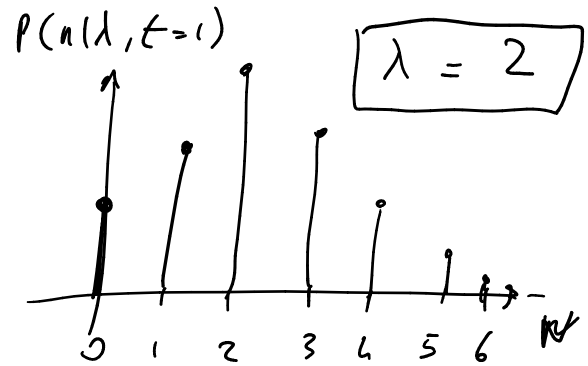
parameters
+
MODEL

Example : **Poisson** :

$$\underbrace{p(n\,|\,\lambda, t)}_{\text{prob. distribution for } n \text{ given } \lambda \ \to\ \text{likelihood}} = \frac{(\lambda t)^n}{n!} e^{-(\lambda t)}$$

prob. of n counts in time t, given rate λ

Prob. Mass function



What is λ given n ?

$\lambda \geq 0 \in \mathbb{R}$

Observed $\hat{n}$: evaluate the likelihood at $n = \hat{n}$, and look at it as a fct of λ for fixed $\hat{n}$

DEF :

$$\mathcal{L}(\lambda) \equiv p(n = \hat{n}\,|\,\lambda, t = 1)$$

likelihood function : NOT a distribution for λ !!!

λ !!!

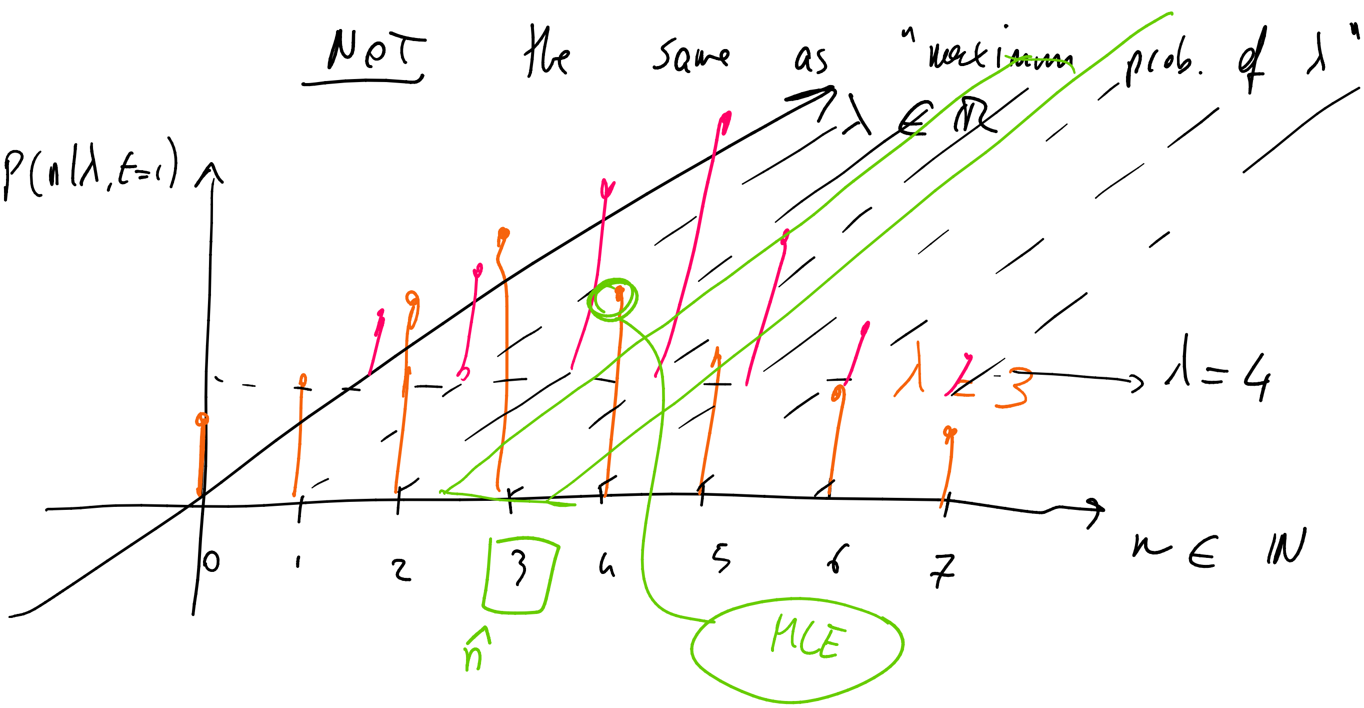
Max. likelihood principle (MLP):

$$\lambda_{MLE} = \max_{\lambda} \mathcal{L}(\lambda)$$

$$\lambda_{MLE} = \max_{\lambda} \log \mathcal{L}(\lambda)$$

λ_{MLE} maximises the prob. of obtaining the observed data.

NOT the same as "maximum prob. of λ"



Likelihood fct:

sampling distribution with parameters Θ for data d

$$\underbrace{P(d=\hat{d}}_{\text{data}} \mid \underbrace{\Theta}_{\text{parameters}}) =: \mathcal{L}(\Theta)$$

Normalization: $\displaystyle\int P(d \mid \theta)\, \underset{\text{data}}{dd} = 1$

$\boxed{\text{NOT true}}$: $\displaystyle\int \mathcal{L}(\theta)\, \underset{\text{parameter}}{d\theta} \neq 1$

Frequentist: prob. as frequency.

Probability: "The number of times the event occurs over the total number of trials, in the limit of an infinite sequence of equiprobable repetitions"

Bayesian: "Probability is a measure of the subjective degree of belief in a proposition"

More general definition

Rules of probabilities:

1) $0 \leq P(A) \leq 1$

2) $P(A) + P(\bar{A}) = 1$ (sum rule)

3) $P(A, B) = P(A \mid B)\, P(B)$ (product rule)

2+3 $\Rightarrow$ Marginalization rule:

$$P(A) = \sum_i P(A, B_i) = \sum_i P(A \mid B_i)\, P(B_i)$$

$$\Rightarrow \quad P(B \mid A)\, P(A) = P(A, B) = P(A \mid B)\, P(B)$$

$$\Rightarrow \quad P(B|A)\,P(A) = P(A,B) = P(A|B)\,P(B)$$

$$\boxed{P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}}$$

BAYES' THEOREM

$A \to \theta$ (parameter)

$B \to d$ (data)

$$\boxed{\underbrace{P(\theta|d)}_{\text{posterior distribution}} = \frac{\overbrace{P(d|\theta)}^{\text{likelihood}}\ \overbrace{P(\theta)}^{\text{prior}}}{P(d)}}$$

$$P(d) = \int P(d|\theta)\,P(\theta)\,d\theta$$

In practice :