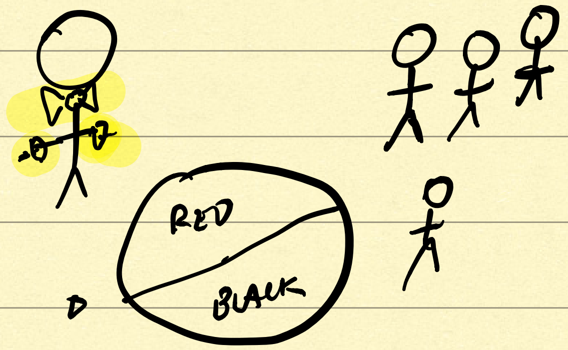
CASINO

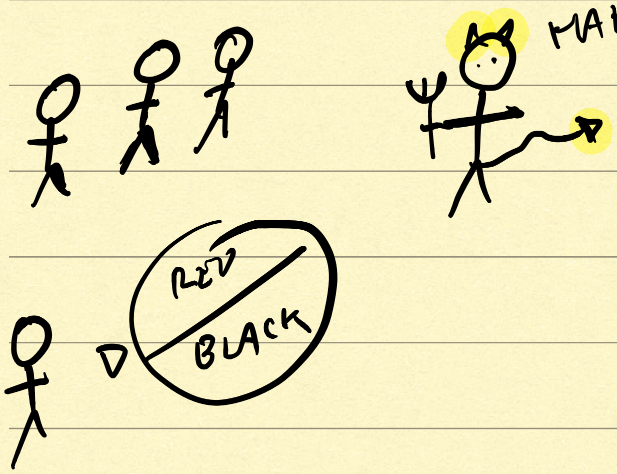


SPIN #1 SPIN #2 SPIN #3

L L L

$\mathbb{P}(\text{SPIN } \#4 = L) = ?$

PROFESSOR MAHAMAHA!



SPIN #1 SPIN #2 SPIN #3

L L L

$\mathbb{P}(\text{SPIN } \#4 = L) = ?$

PROBABILITY

MODEL, PARAMETERS → OBSERVATION

EX: i.i.d.

L, L, L, ?

LEARNING
INFERENCE
REGRESSION
CLASSIFICATION

PREDICTION

STATISTICAL LEARNING

PROBABILITY DENSITY $p(t) = \lambda e^{-\lambda t} = p(t;\lambda)$

PROBABILITY MASS $p(k) = \frac{\lambda^k e^{-\lambda}}{k!} = p(k;\lambda)$

THE LIKELIHOOD FUNCTION IS THE PROBABILITY MASS / DENSITY VIEWED AS A FUNCTION OF THE PARAMETERS

EX. $L(\lambda) = \lambda e^{-\lambda t}$

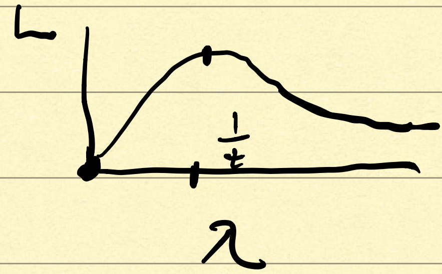
NOTE $\int L(\lambda)\, d\lambda \neq 1$

$L(\lambda) \geq 0$

TO FIND A PARAMETER θ FROM AN OBSERVATION X, TAKE THE $\theta = \hat{\theta}$ THAT MAXIMIZES $L(\theta)$.

→ MAXIMUM LIKELIHOOD ESTIMATION

EX T - TIME



$$L(\lambda) = \lambda e^{-\lambda t}$$

$$\frac{\partial L}{\partial \lambda} = 0 \quad \ldots \quad \hat{\lambda} = \frac{1}{t}$$

EX SUPPOSE YOU HAVE N OBSERVATIONS THAT INDEPENDENT & IDENTICALLY DISTRIBUTED (i.i.d.) THEN

$$L(\theta) = \prod_{i=1}^{N} p(x_i ; \theta)$$

EX N - NORMAL RANDOM VARIABLES

$$L(\mu, \sigma) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^{N} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{N} (x_i - \mu)^2\right)$$

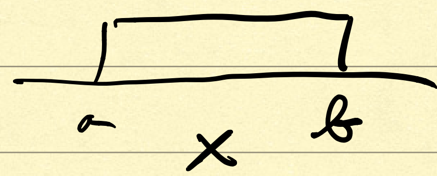
$$\frac{\partial L}{\partial \mu} = 0, \quad \frac{\partial L}{\partial \sigma} = 0 \quad \ldots$$

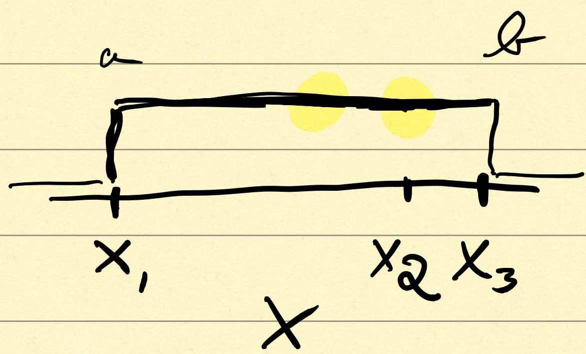
$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

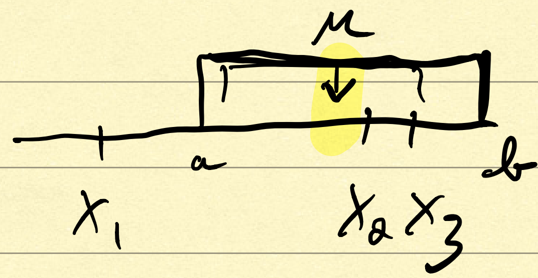
$$\hat{\sigma} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{\mu})^2}$$

EX UNIFORM RANDOM VARIABLES

$$X_i \sim \text{UNIF}(a, b)$$



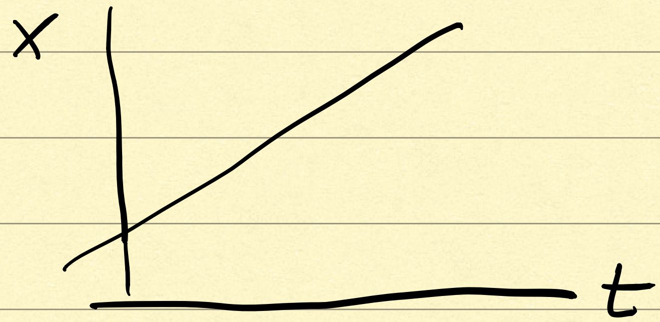




MAXIMUM LIKELIHOOD

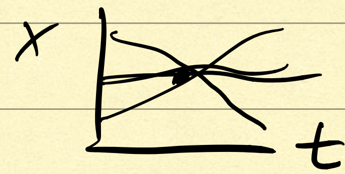
MOMENT MATCHING

EX $\frac{dX}{dt} = \beta_1 \qquad X(0) = \beta_0$

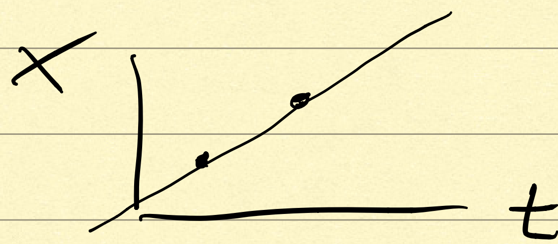


$X(t) = \beta_0 + \beta_1 t$

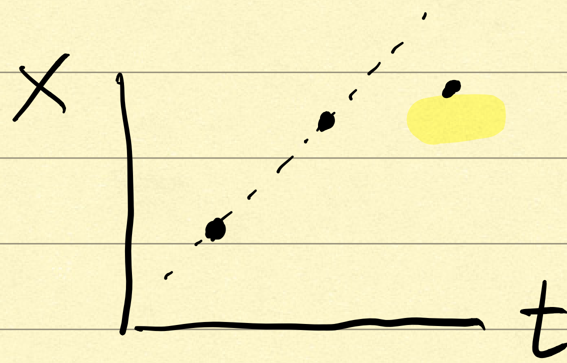
1 DATA POINT:



2 DATA POINTS:



3 DATA POINTS:



$L(\beta_0, \beta_1) = 0$

EX $\frac{dX}{dt} = \beta_1 \quad , \quad X(0) = \beta_0$

$$Y(t) = X(t) + E_i$$

E_i - RANDOM VARIABLE