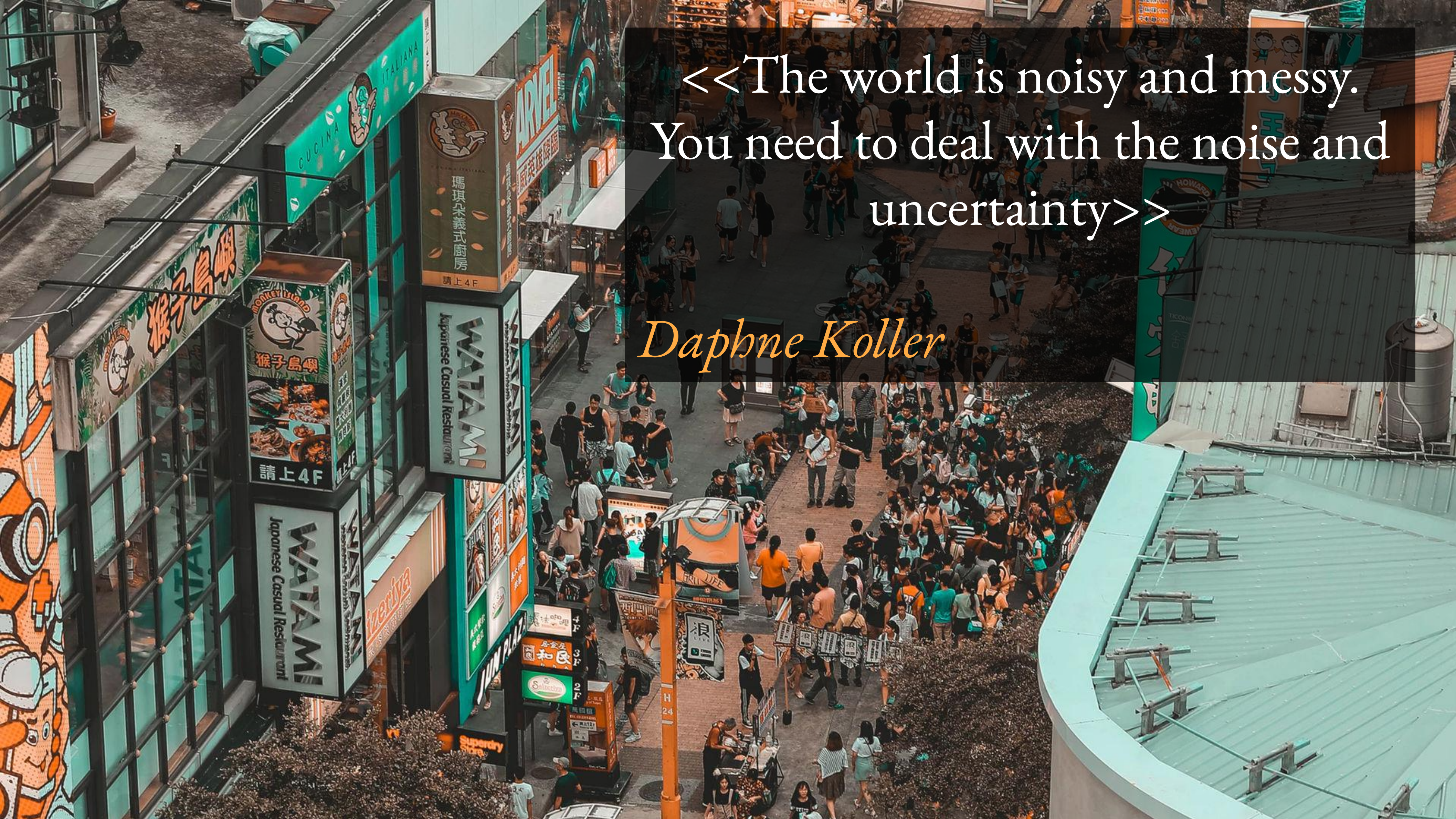
Probabilistic Modeling and Bayesian Networks

Introduction

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

INFORMATIKA FAKULTATEA
FACULTAD DE INFORMÁTICA

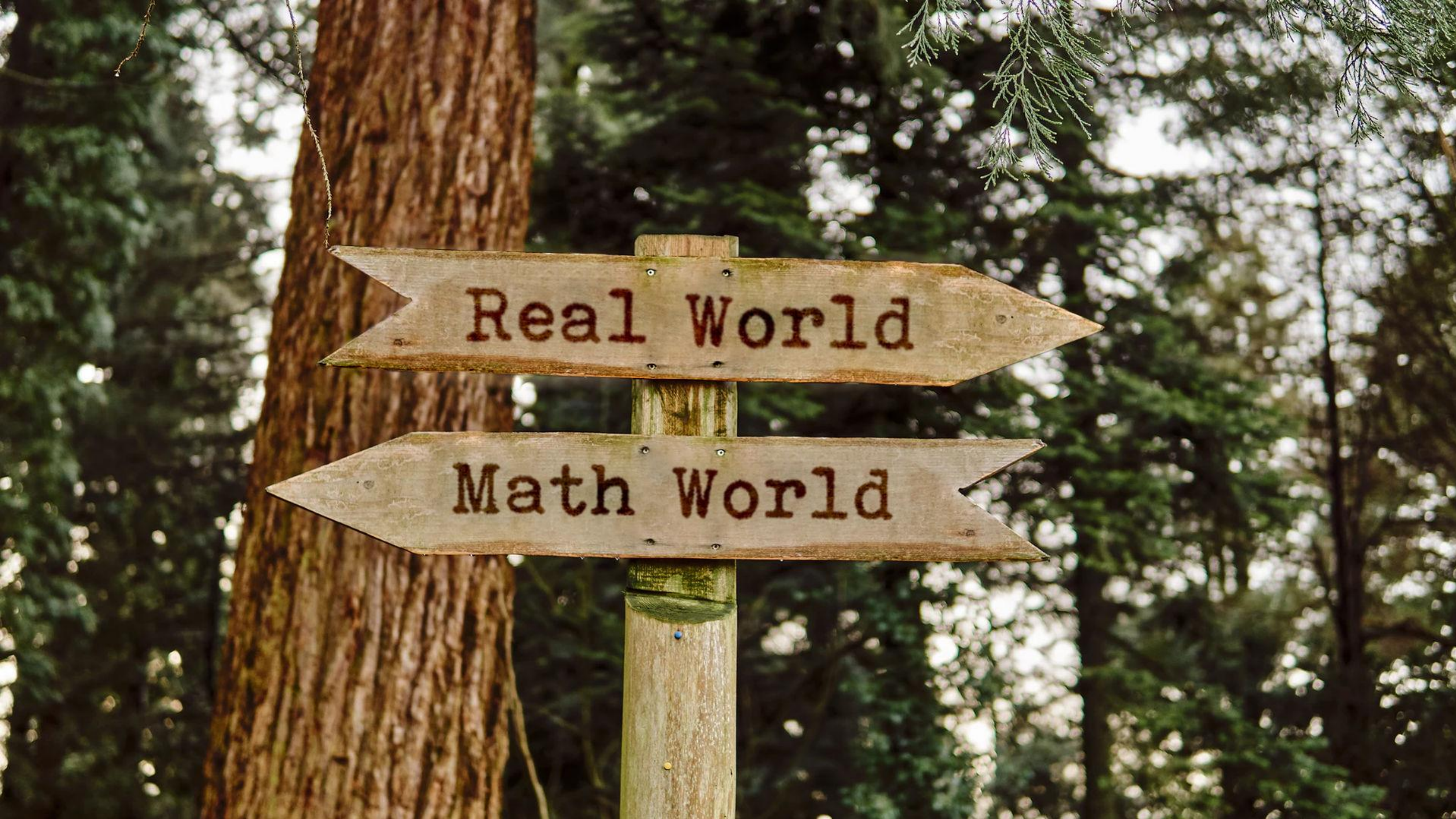
Konputazio Ingeniaritza eta Sistema Adimentsuak Unibertsitate Masterra
Máster Universitario en Ingeniería Computacional y Sistemas Inteligentes



<<The world is noisy and messy. You need to deal with the noise and uncertainty>>

Daphne Koller


Parmesan Pepper

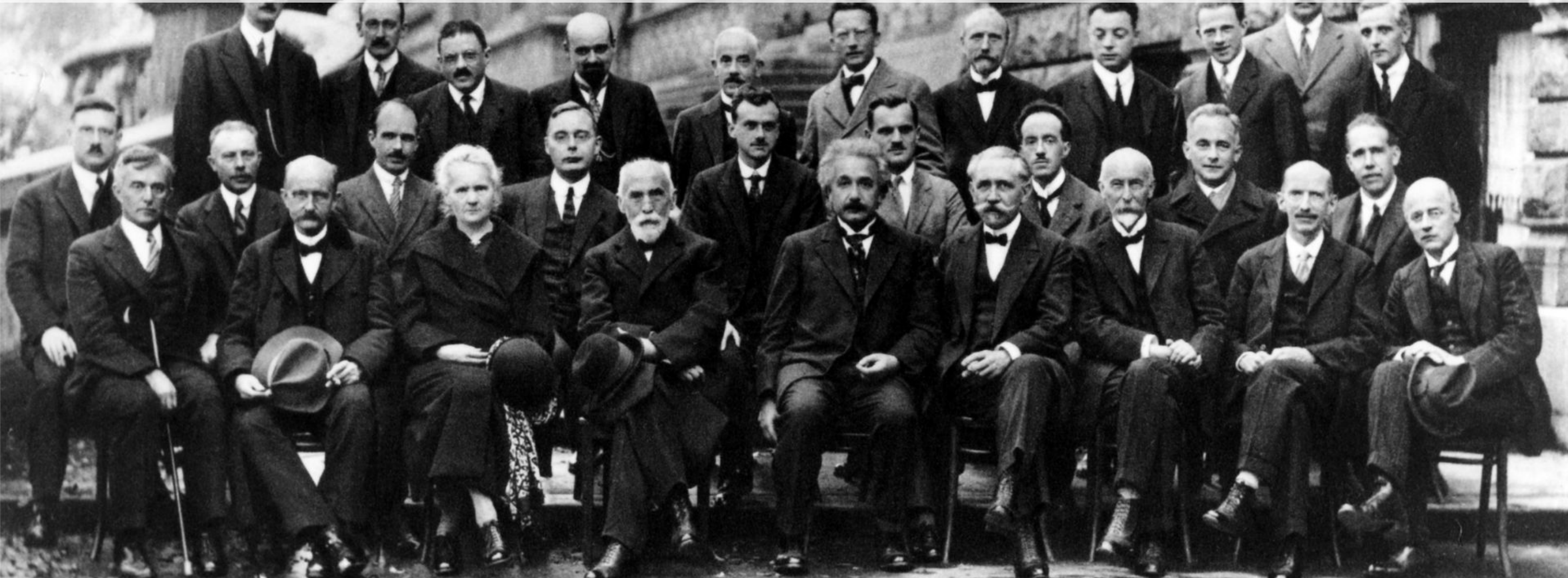

Real World
Math World

What is Probability?

and uncertainty?

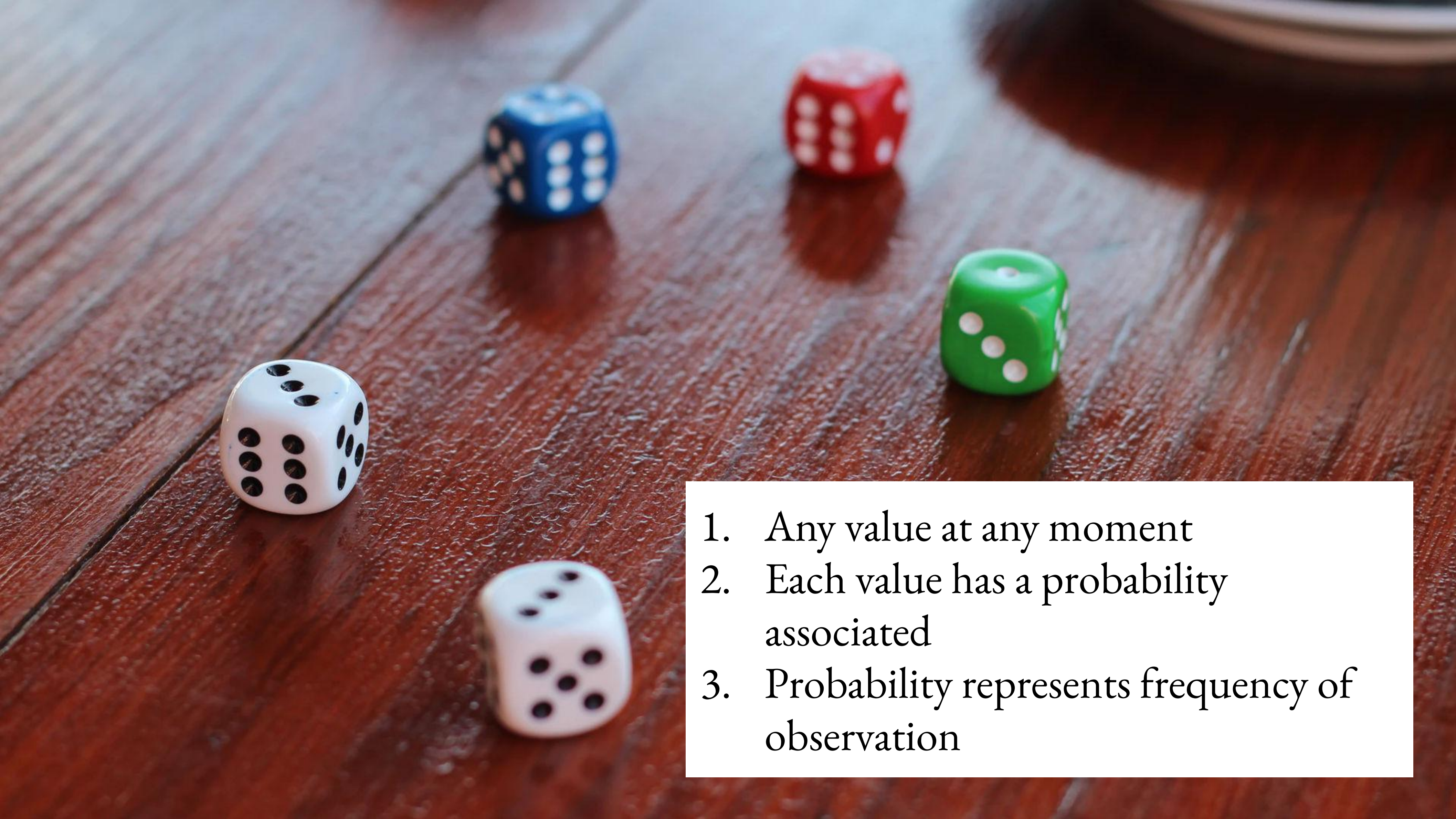
≪As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality≫

Albert Einstein



Random Variables

Basic Concepts



1. Any value at any moment
2. Each value has a probability associated
3. Probability represents frequency of observation

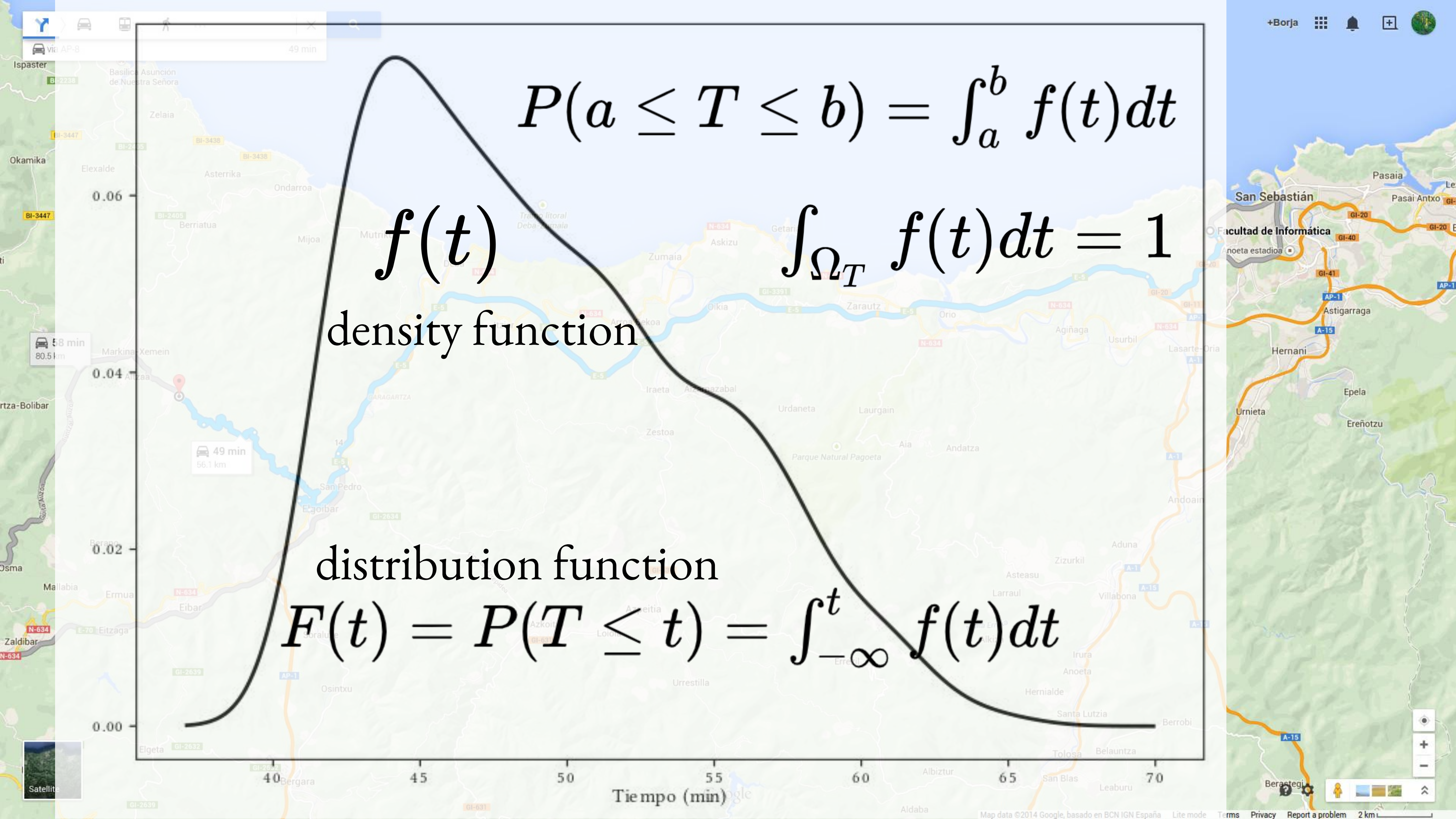
What's the probability of getting ⚅ (P(X=6))?

What's the probability of getting an even number? And ANY number?

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

$$P(X = x_i or X = x_j) = P(X = x_i) + P(X = x_j)$$

$$P(X \in \Omega_X) = \sum_{x \in \Omega_X} P(X = x) = 1$$



$P(a \leq T \leq b) = \int_a^b f(t)dt$
$f(t)$
$\int_{\Omega_T} f(t)dt = 1$
density function
distribution function
$F(t) = P(T \leq t) = \int_{-\infty}^t f(t)dt$
0.00
0.02
0.04
0.06
40
45
50
55
60
65
70
Tiempo (min)

SUMMARY

- A random variable can take any value at any moment
- Each value has an associated probability
- Random variables can be ...

DISCRETE VARIABLES

Probability mass function, P(X)

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

$$P(X \in A) = \sum_{x \in A} P(X = x)$$

$$\sum_{x \in \Omega_X} P(X = x) = 1$$

CONTINUOUS VARIABLES

Probability density function, f(X)

$$f(X) \geq 0$$

$$P(a \leq X \leq b) = \int_a^b f(X) dX$$

$$\int_{\Omega_X} f(X) dX = 1$$

Random Variables

Vectors of Variables

JOINT PROBABILITY

$$P(X, Y)$$

MARGINAL PROBABILITY

$$P(X) \qquad P(Y)$$

$X \in \{H, T\}$

$Y \in \{1, 2, 3, 4, 5, 6\}$

Provided that I have the joint probability, what's the probability of Y=6?

MARGINALIZATION

$$P(Y = y) = \sum_{x \in \Omega_X} P(X = x, Y = y)$$

$$P(Y = 6) = P(Y = 6, X = H) + P(Y = 6, X = T)$$

1. Toss the coin
2. If the result is R, roll the red dice; otherwise, roll the green dice

* The green dice is a regular dice, but the red dice only has even numbers

What's the probability of Red and 1? and Red and 2?

CONDITIONAL PROBABILITIES

$$P(Y = 1 | X = R) = 0$$
$$P(Y = 2 | X = R) = 0.33$$

$$P(X, Y) = P(X) P(Y|X)$$
$$P(X, Y) = P(Y) P(X|Y)$$

GENERALIZATION: CHAIN RULE

$$P(X_1, \ldots, X_n) = P(X_1) \cdot P(X_2|X_1) \cdot \ldots \cdot P(X_n|X_1, \ldots, X_{n-1})$$

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1})$$

BAYES' THEOREM

$$P(X|Y) = \frac{P(X,Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)}$$

INDEPENDENCE

$$P(Y = 3|X = H)?$$

$$P(Y = 3|X = T)?$$

X and Y are independent if:

$$P(X, Y) = P(X)P(Y)$$

$$P(X|Y) = P(X); P(Y|X) = P(Y)$$

CONDITIONAL INDEPENDENCE

X and Y are conditionally independent given Z if:

$$P(X|Y,Z) = P(X|Z); P(Y|X,Z) = P(Y|Z)$$

$$P(X,Y|Z) = P(X|Z)P(Y|Z)$$

SUMMARY

Types of probability

$$P(X, Y) \qquad P(X) \qquad P(X|Y)$$

Chain Rule

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | X_1, \ldots, X_{i-1})$$

Bayes' Theorem

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

Independency

$$P(X, Y) = P(X)P(Y) \qquad P(X|Y) = P(X)$$

EXERCISE

$P(Y\|X)$	1	2	3	4	5
R	0	0.4	0	0.3	0
G	0.1	0.2	0.1	0.2	0.1

	$P(X)$
R	0.6
G	0.4

$$P(Y = 6|X = R)$$
$$P(Y = 6)$$
$$P(Y = 2, X = G)$$
$$P(X = G|Y = 5)$$

Are X and Y independent?

MONTY HALL PARADOX

Probabilistic Modeling and Bayesian Networks

Introduction

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

INFORMATIKA FAKULTATEA
FACULTAD DE INFORMÁTICA

Konputazio Ingeniaritza eta Sistema Adimentsuak Unibertsitate Masterra
Máster Universitario en Ingeniería Computacional y Sistemas Inteligentes