Bayes' Rule, Discrete Random Variables and probability distribution

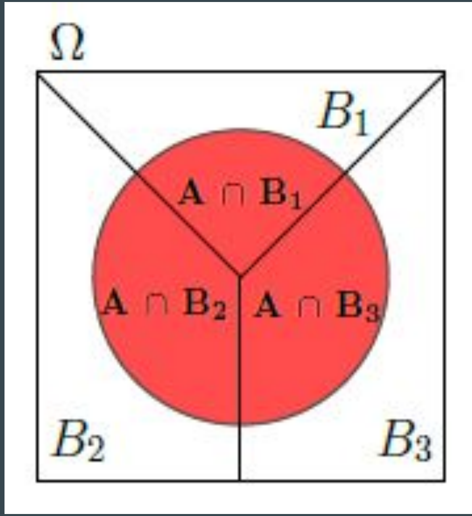
•••

Lesson 3

Contents

Law of Total Probability

- Suppose the sample space Ω is divided into 3 disjoint events B1, B2 and B3.
- Then for any event A:
- $P(A) = P(A \cap B1) + P(A \cap B2) + P(A \cap B3)$



$$P(A) = P(A|B1).P(B1) + P(A|B2).P(B2) + P(A|B3).P(B3)$$

Example

An urn contains 5 red balls and 2 green balls. Two balls are drawn one after the other. What is the probability that the second ball is red?

Example

An urn contains 5 red balls and 2 green balls. Two balls are drawn one after the other. What is the probability that the second ball is red?

Solution:

The sample space is $\Omega = \{rr, rg, gr, gg\}$. Let R1 be the event 'the first ball is red', G1 = 'first ball is green', R2 = 'second ball is red', G2 = 'second ball is green'. We are asked to find P(R2).

$$P(R2) = P(R2|R1)P(R1) + P(R2|G1)P(G1)$$

Bayes' Rule

- For two events A and B Bayes' theorem (also called Bayes' rule and Bayes' formula) says:
- $P(B \mid A) = P(A \mid B) \cdot P(B) \ / \ P(A)$

- Comments:
  - Bayes' rule tells us how to 'invert' conditional probabilities, i.e. to find $P(B|A)$ from $P(A|B)$.
  - In practice, $P(A)$ is often computed using the law of total probability.

Example

Toss a coin 5 times. Let H_1 = 'first toss is head' and let H_A = 'all 5 tosses are heads'.

Solution:

Then $P(H_1|H_A) =$ __ .

$P(H_A|H_1) =$ __.

Example

Consider a routine screening test for a disease. Suppose the frequency of the disease in the population (base rate) is 0.5%. The test is highly accurate with a 5% false positive rate and a 10% false negative rate.

You take the test and it comes back positive. What is the probability that you have the disease?

Example

Events:

D+ = ‘you have the disease’

D– = ‘you do not have the disease

T+ = ‘you tested positive’

T– = ‘you tested negative’.

$$P(D+ \mid T+)$$

Example

Using :

- $P(D+) = 0.005$
- $P(D-) =$ ______
- $P(T- \mid D+) = 0.1$ (false negative)
- $P(T+ \mid D+) =$ ___
- $P(T+ \mid D-) =$ ____ (false positive)

Example

$$P(D+ \mid T+) = \frac{P(T+ \mid D+) \cdot P(D+)}{P(T+)}$$

$$P(D+ \mid T+) = \frac{P(T+ \mid D+) \cdot P(D+)}{P(T+ \mid D+) \cdot P(D+) + P(T+ \mid D-) \cdot P(D-)}$$

Discrete Random Variables

Roll a die twice and record the outcomes as (i, j), where i is the result of the first roll and j the result of the second. We can take the sample space to be

$$\Omega = \{(1, 1), (1, 2), (1, 3), \ldots, (6, 6)\} = \{(i, j) \mid i, j = 1, \ldots, 6\}$$

The probability function is $P(i, j) = 1/36$.

In this game, you win $500 if the sum is 7 and lose $100 otherwise.

Discrete Random Variables

We give this function the name X and describe it formally by:

$$X(i, j) = \begin{cases} 500 & \text{if } i + j = 7 \\ -100 & \text{if } i + j \mathrel{!}= 7 \end{cases}$$

A random variable assigns a number to each outcome in a sample space.

Discrete Random Variables

Definition: Let Ω be a discrete sample space. A discrete random variable is a function

$$X : \Omega \longrightarrow \mathrm{IR}$$

that takes a discrete set of values $\Omega = \{ \omega 1 , \omega 2 , \ldots , \omega n \}$

The event X = 500 is the set {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}, i.e. the set of all outcomes that sum to 7. So P(X = 500) = 1/6.

Probability mass function - pmf

Definition: The probability mass function (pmf) of a discrete random variable X is the function $p(a) = P(X = a)$ (or $p_X(a)$)

1. We always have $0 \leq p(a) \leq 1$.
2. We allow *a* to be any number. If *a* is a value that X never takes, then $p(a) = 0$.

Probability mass function - pmf

Inequalities with random variables describe events. For example $X \leq a$ is the set of all outcomes ω such that $X(\omega) \leq a$.

Example: If our sample space is the set of all pairs of (i, j) coming from rolling two dice and $Z(i, j) = i + j$ is the sum of the dice then

- $Z \leq 4 = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}$ -> $P(Z(\omega) \leq 4) = ??$
- $Z == 13 = \{\}$ -> $P(Z(\omega) = 13) = ??$

Cumulative distribution function - cdf

Definition: The cumulative distribution function (cdf) of a discrete random variable X is the function F given by $F(a) = P(X \leq a)$. We will often shorten this to distribution function.

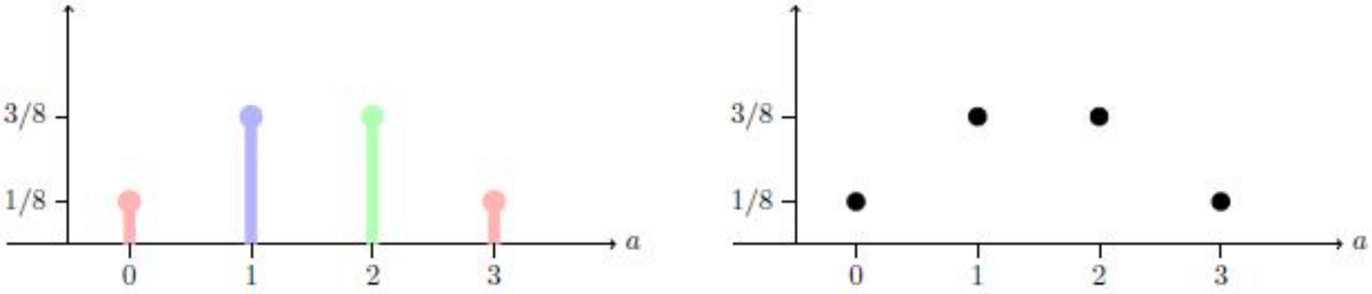
Properties:

- $0 \leq F(a) \leq 1$
- F is non-decreasing. Symbolically if $a \leq b$ then $F(a) \leq F(b)$.
- $\lim_{a \to \infty} F(a) = 1, \quad \lim_{a \to -\infty} F(a) = 0.$

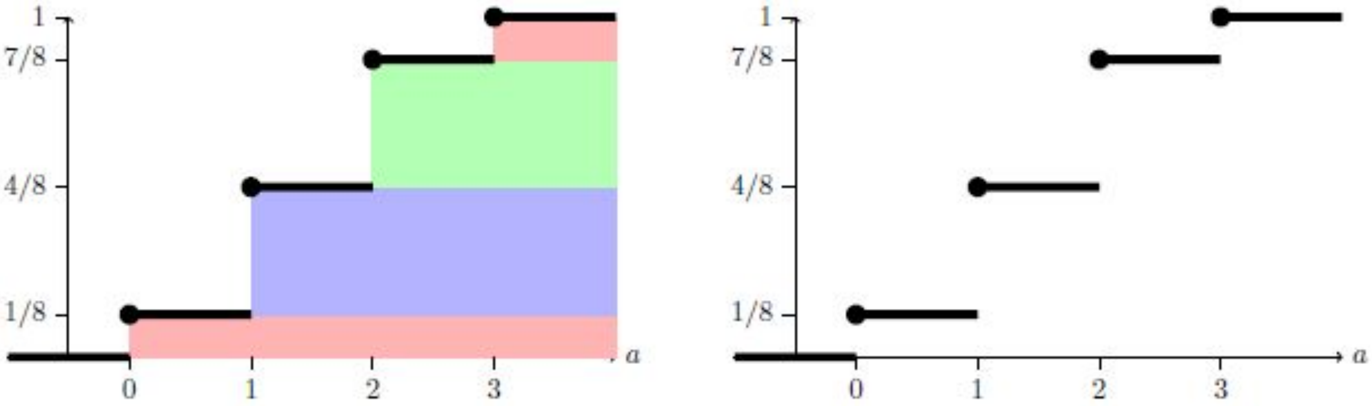
Cumulative distribution function VS Probability mass function

Example 1: let X be the number of heads in 3 tosses of a fair coin:

value a:	0	1	2	3
pmf p(a):	__	3/8	__	1/8
cdf F(a):	1/8	__	__	1



Probability mass function for X



Cumulative distribution function for X

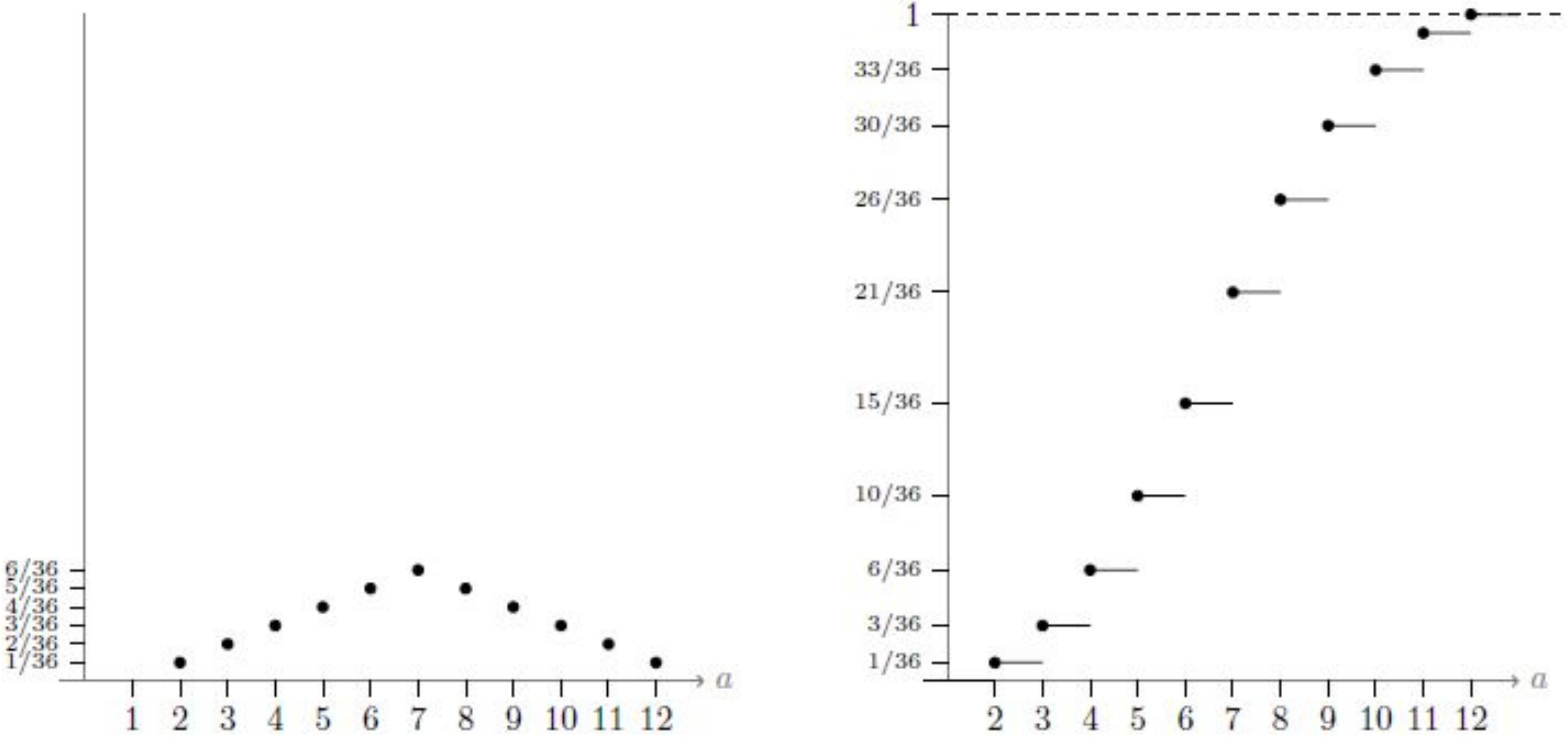
Cumulative distribution function VS Probability mass function

Example 2: Let Ω be our earlier sample space for rolling 2 dice. Define the random variable M to be the maximum value of the two dice:

$$M(i, j) = \max(i, j)$$

For example, the roll (3,5) has maximum 5, i.e. $M(3, 5) = 5$.

- $F(8) = P(M \leq 8) =$ __
- $F(-2) = P(M \leq -2) =$ __
- $F(2.5) = P(M \leq 2.5) =$ __
- $F(\pi) = P(M \leq \pi) = 9/36$



pmf and cdf for the maximum of two dice

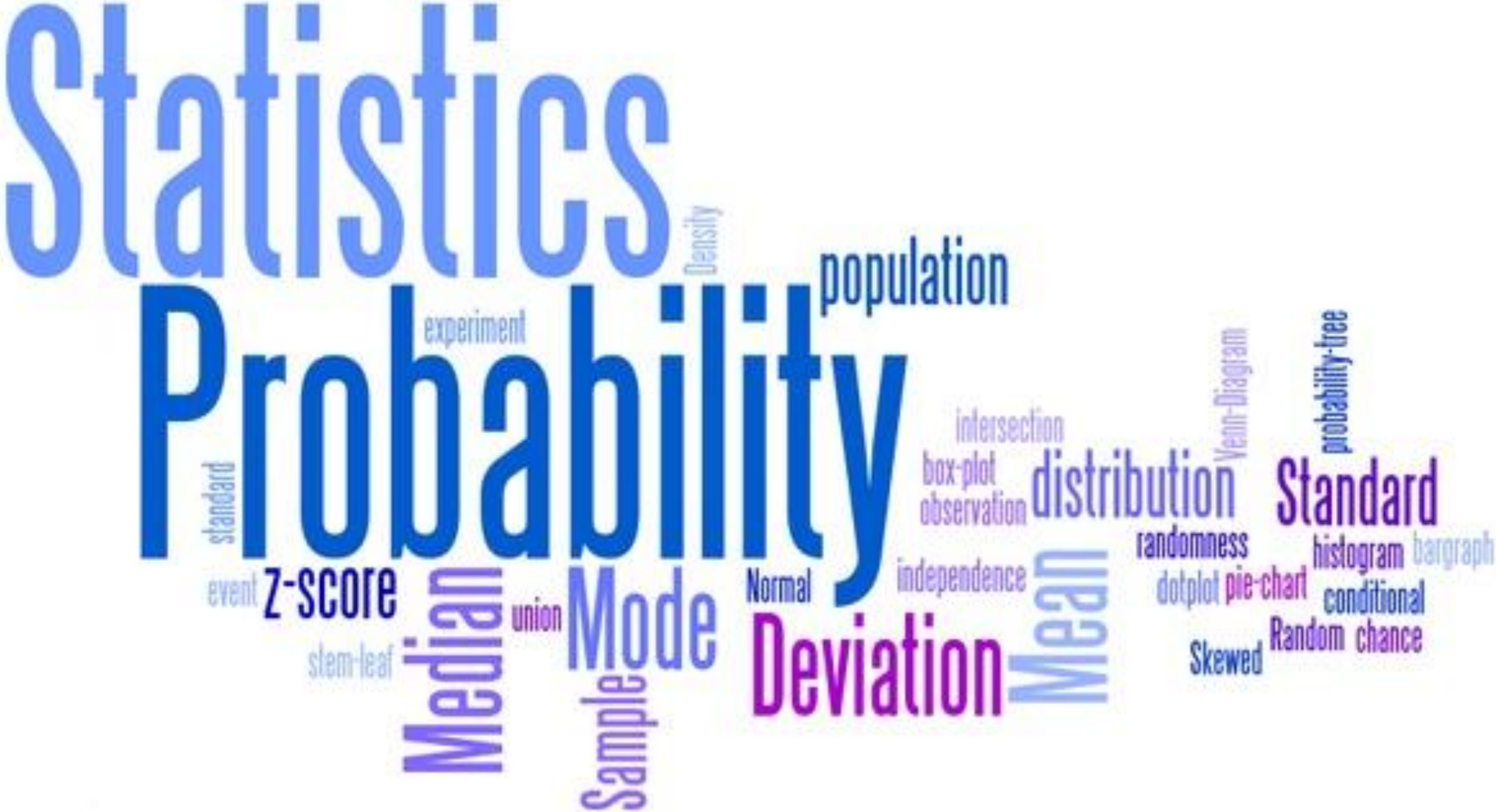
Math Team

Christian Córdova

Bachelor in Mathematics

José Castro

Bachelor in Computer Science


Statistics
Density
population
Probability
experiment
probability-tree
Venn-Diagram
intersection
box-plot
distribution
Standard
observation
standard
randomness
histogram
bargraph
event
z-score
Median
Normal
independence
dotplot
pie-chart
conditional
union
Mode
Mean
Random
chance
stem-leaf
Deviation
Skewed
Sample