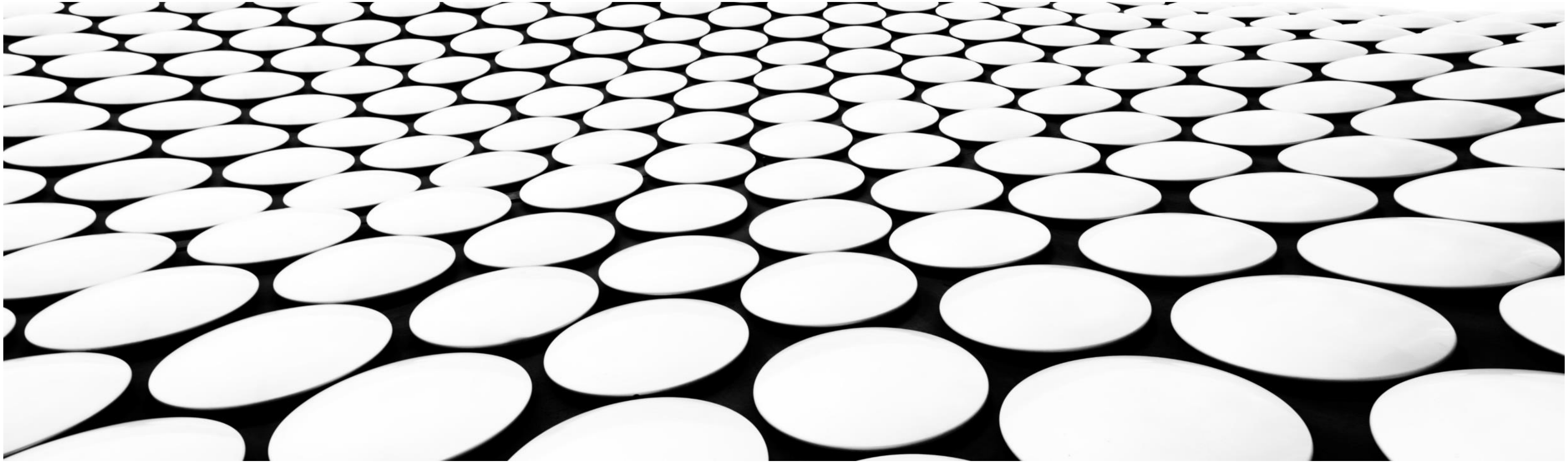
NAIVE BAYES CLASSIFIER

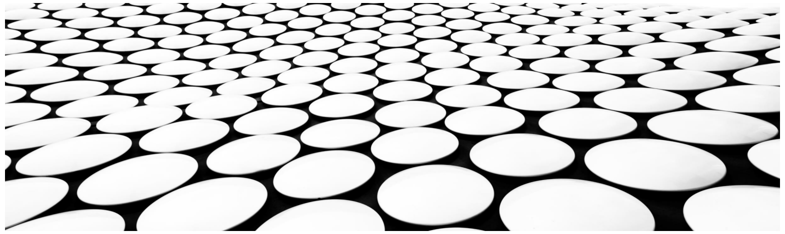
PART 2 OF 3 – HYPOTHESIS TESTING AND CLASSIFICATION



GOALS

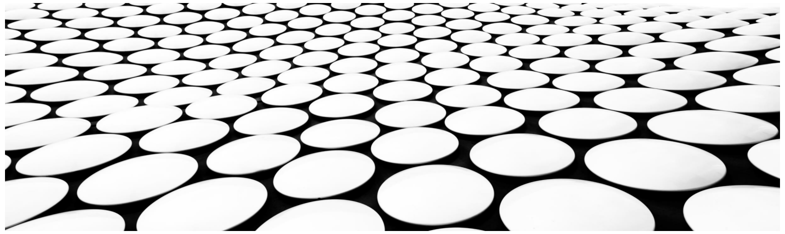
- Bayes Theorem
- Hypothesis testing
- Classification example

BAYES RULE



p(A|B) = p(B|A)*p(A) /p(B)

HYPOTHESIS TESTING



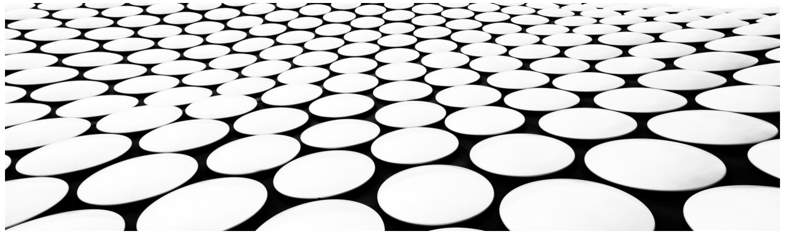
One of the most important results of probability theory is the general form of Bayes' theorem.

Assume there are individual hypotheses, h_i, from a set of hypotheses, H. Assume a set of evidences, E.

$p(h_i|E) = (p(E|\ h_i) * p(h_i))\ /\ p(E)$

This equation may be read, "***The probability of an hypothesis h_i given a set of evidences E is . . .***"

GENERAL FORM OF BAYES THEOREM

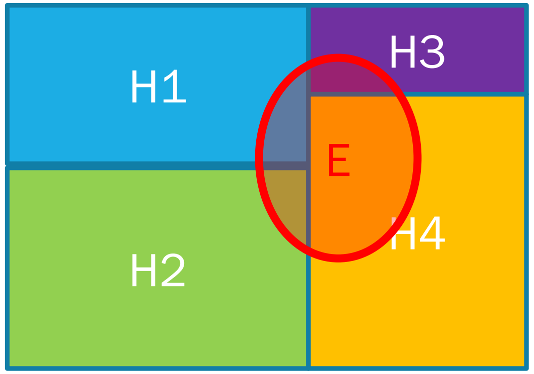


The general form of Bayes' theorem assumes that the set of hypotheses, h_i **partitions** the evidence set E:

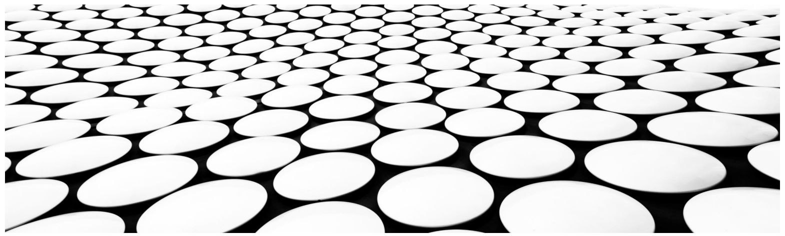
$p(hi|E)$ is the probability that h_i is true given evidence E
$p(h_i)$ is the probability that h_i is true overall
$p(E|h_i)$ is the probability of observing evidence E when h_i is true
n is the number of possible hypotheses

$$p(h_i|E) = \frac{p(E|h_i) \times p(h_i)}{\sum_{k=1}^{n} p(E|h_k) \times p(h_k)}$$

P(E)



PRIOR / POSTERIOR



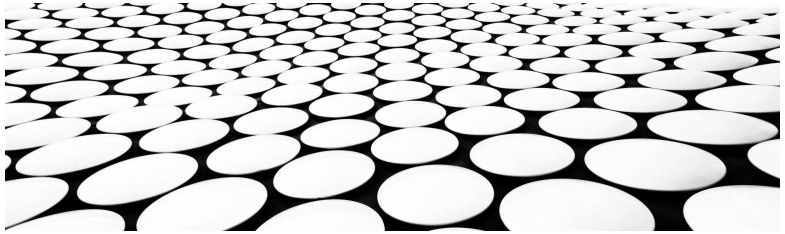
Prior probability
The prior probability, is the probability, generally an unconditioned probability, of an hypothesis or an event.

The prior probability of an event is symbolized: p(E), and the prior probability of an hypothesis is symbolized $p(h_i)$.

Posterior probability
The posterior (after the fact) probability, generally a conditional probability, of an hypothesis is the probability of the hypothesis given some evidence.

The posterior probability of an hypothesis given some evidence is symbolized: $p(h_i \mid E)$.

EXAMPLE OF BAYES THEOREM



Suppose that you go out to purchase a bicycle.

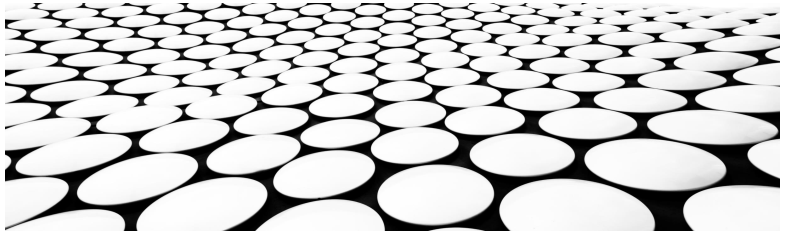
Only 3 sellers: D1, D2, D3
p(D1) = 0.2, P(D2) = 0.4, P(D3) = 0.4

You're interested in bicycle A.
If you go to D_1, the probability of buying A is 0.2. At seller D_2, the probability of buying is 0.4, and at D_3, the probability of buying is 0.3.

Assuming you bought the bicycle, what is the probability that you purchased it from D2?

Rephrase : What is the posterior probability of the hypothesis D2 given the evidence that you bought A?

EXAMPLE OF BAYES THEOREM



Only 3 sellers: D1, D2, D3
p(D1) = 0.2, P(D2) = 0.4, P(D3) = 0.4

You're interested in bicycle A.
If you go to D_1, the probability of buying A is 0.2. At seller D_2, the probability of buying is 0.4, and at D_3, the probability of buying is 0.3.

Assuming you bought the bicycle, what is the probability that you purchased it from D2?

<u>Space partitioning</u>

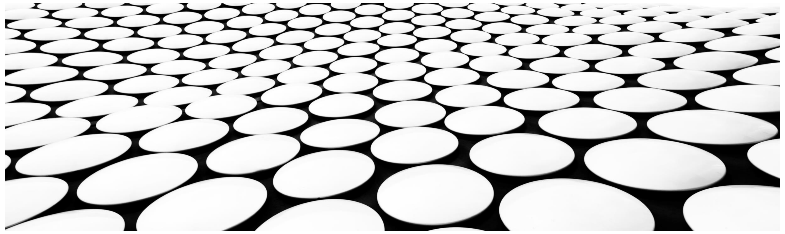
p(A) = p(A | D1) * P(D1) + p(A | D2) * P(D2) + p(A | D3) * P(D3)
p(A) = 0.2 * 0.2 + 0.4 * 0.4 + 0.3 * 0.4 = 0.32

<u>Hypothesis evaluation</u>

p(D2|A) = (p(A|D2) * p(D2)) / p(A)

= 0.4 * 0.4 / 0.32
= 0.16 / 0.32
= 0.5

NAIVE BAYES CLASSIFIER



Only 3 sellers: D1, D2, D3

p(D1) = 0.2, P(D2) = 0.4, P(D3) = 0.4

P(A | D1) = 0.2

p(A | D2) = 0.4

p(A | D3) = 0.3

<u>Space partitioning</u>

p(E) = p(A | D1) * P(D1) + p(A | D2) * P(D2) + p(A | D3) * P(D3)

p(E) = 0.2 * 0.2 + 0.4 * 0.4 + 0.3 * 0.4 = 0.32

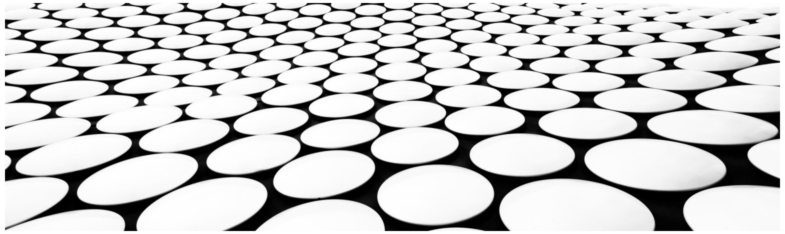
Assuming you bought bicycle A, where did you buy it from?

P(D1|A) — p(D1|A) = (p(A|D1) * p(D1)) / p(E) = 0.04/0.32 = 0.125

P(D2|A) — p(D2|A) = (p(A|D2) * p(D2)) / p(E) = 0.16/0.32 = 0.5

P(D3|A) — p(D3|A) = (p(A|D3) * p(D3)) / p(E) = 0.12/0.32 = 0.375

NAIVE BAYES CLASSIFIER



Assuming you bought bicycle A, where did you buy it from?

P(D1 | A)	p(D1|A) = (p(A|D1) * p(D1)) / p(E) = 0.04/0.32 =	0.125
P(D2 | A)	p(D2|A) = (p(A|D2) * p(D2)) / p(E) = 0.16/0.32 =	0.5
P(D3 | A)	p(D3|A) = (p(A|D3) * p(D3)) / p(E) = 0.12/0.32 =	0.375

LET’S CONTINUE…

Next video:
Conditional independence