Probability Theory and Statistics

Poul Svante Eriksen, Dept. For Mathematical Sciences

svante@math.aau.dk

Literature:

Walpole, Myers, Myers & Ye:
Probability and Statistics for Engineers and Scientists,
Prentice Hall, 8th ed.

Slides and lecture overview:

http://people.math.aau.dk/~svante/K7/

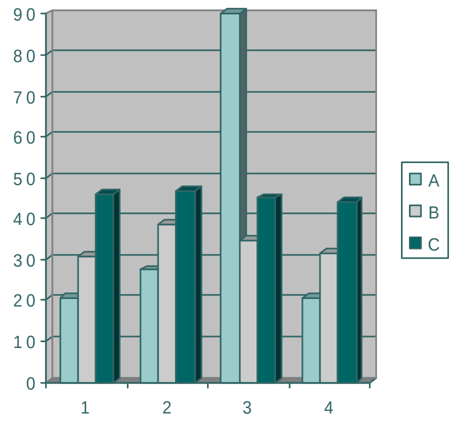
Lecture format:

2x45 min lecturing followed by exercises in group rooms

STATISTICS
What is it good for?

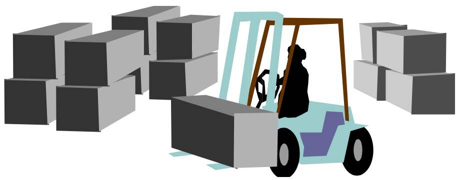
Forecasting:

- Expectations for the future?
- How will the stock markets behave??



Analysis of sales:

- How much do we sell, and when?
- Should we change or sales strategy?



Quality control:

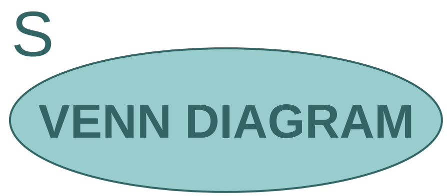
- What is my rate of defective products?
- How can I best manage my production?
- What is the best way to sample?

Probability theory

Sample space and events

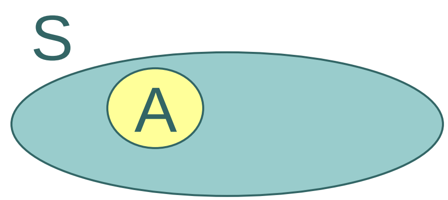
Consider an experiment

Sample space S:



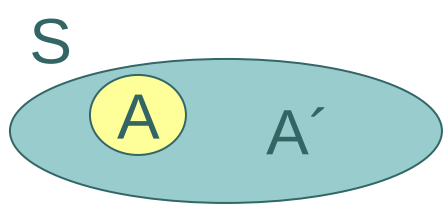
Example:
S={1,2,…,6} rolling a dice
S={head,tail} flipping a coin

Event A:



Example:
A={1,6} when rolling a dice

Complementary event A´:

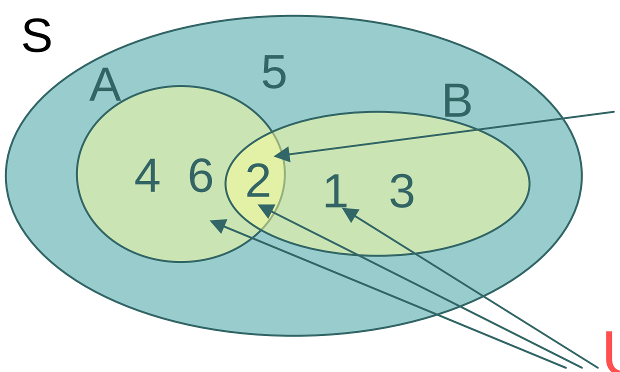


Example:
A´={2,3,4,5} rolling a dice

Probability theory
Events

Example:
Rolling a dice

$S=\{1,2,3,4,5,6\}$
$A=\{2,4,6\}$
$B=\{1,2,3\}$

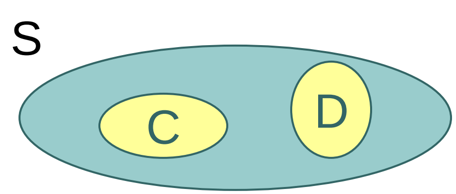


Intersection:
$A \cap B=\{2\}$

Union:
$A \cup B=\{1,2,3,4,6\}$

Disjoint events: $C \cap D = \emptyset$

$C=\{1,3,5\}$ and $D=\{2,4,6\}$ are disjoint



Probability theory

Counting sample points

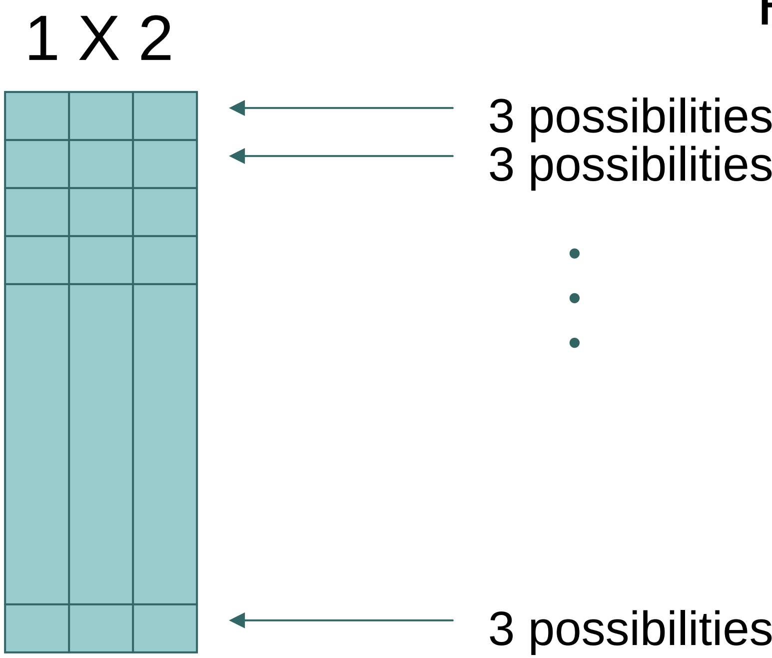
Ways of placing your bets: Guess the results of 13 matches

Possible outcomes:

Home win

Draw

Away win

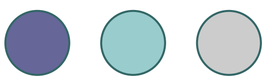


Answer: $3 \cdot 3 \cdot 3 \cdot \cdots \cdot 3 = 3^{13}$

The multiplication rule

Probability theory
Counting sample points

Ordering *n* different objects
Number of permutations ???



There are

- n ways of selecting the first object
- n -1 ways of selecting second object

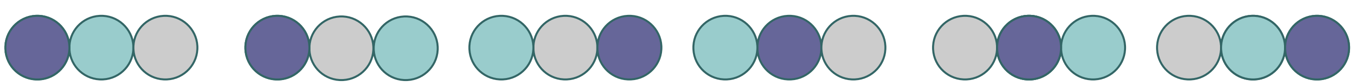
  ⋮

- 1 way of selecting the last object

”n factorial”

$n \cdot (n-1) \cdot \cdots \cdot 1 = n\,!$ ways

The multiplication rule



$3! = 6$

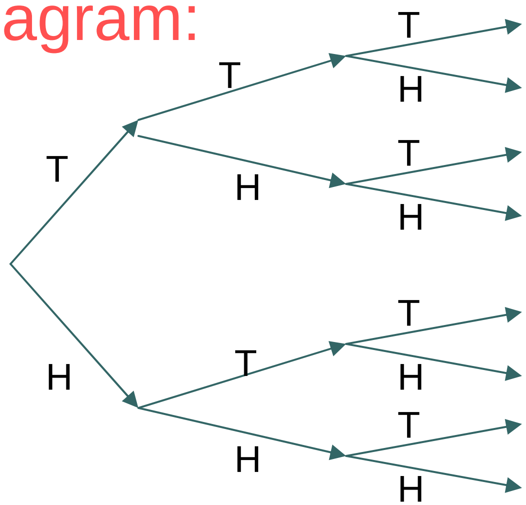
Probability theory
Counting sample points

Multiplication rule:

If k independent operations can be performed in $n_1, n_2, \ldots, n_k$ ways, respectively, then the k operations can be performed in

$$n_1 \cdot n_2 \cdot \cdots \cdot n_k \text{ ways}$$

Tree diagram:



Flipping a coin three times (Head/Tail)

$2^3 = 8$ possible outcomes

Probability theory
Counting sample points

Number of possible ways of selecting r objects from a set of n distinct elements:

	Without replacement	With replacement
Ordered	${}_nP_r = \frac{n!}{(n-r)!}$	n^r
Unordered	$\binom{n}{r} = \frac{n!}{r!(n-r)!}$	-

Probability theory
Counting sample points

Example:
Ann, Barry, Chris, and Dan should from a committee consisting of two persons, i.e. unordered without replacement.

Number of possible combinations:

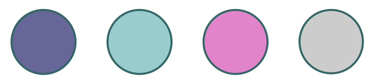
$$\binom{4}{2} = \frac{4!}{2!2!} = 6$$

Writing it out : AB AC AD BC BD CD

Probability theory

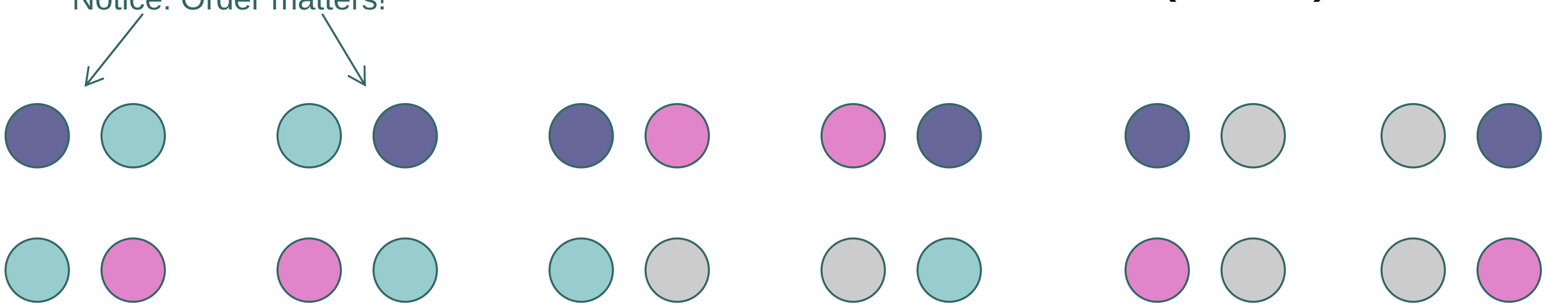
Counting sample points

Example:
Select 2 out of 4 different balls ordered and without replacement



Number of possible combinations: $${}_4P_2 = \frac{4!}{(4-2)!} = 12$$

Notice: Order matters!

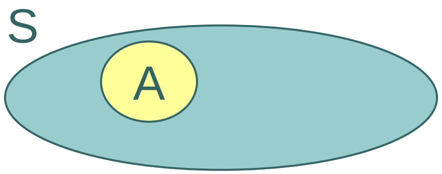


Probability theory

Probability

Let A be an event, then we denote

$P(A)$ the probability for A



It always hold that $0 \leq P(A) \leq 1$ $\quad P(\emptyset) = 0$ $\quad P(S) = 1$

Consider an experiment which has N equally likely outcomes, and let exactly n of these events correspond to the event A. Then

$$P(A) = \frac{n}{N} = \frac{\text{\# successful outcomes}}{\text{\# possible outcomes}}$$

Example:
Rolling a dice

P(even number)

$$= \frac{3}{6} = \frac{1}{2}$$

Probability theory
Probability

Example: Quality control
A batch of 20 units contains 8 defective units.

Select 6 units (unordered and without replacement).

Event A: no defective units in our random sample.

Number of possible samples: $N = \binom{20}{6}$ (# possible)

Number of samples without defective units: $n = \binom{12}{6}$ (# successful)

$$P(A) = \frac{\binom{12}{6}}{\binom{20}{6}} = \frac{12!6!14!}{6!6!20!} = \frac{77}{3230} = 0.024$$

Probability theory
Probability

Example: continued

Event B: exactly 2 defective units in our sample

Number of samples with exactly 2 defective units:

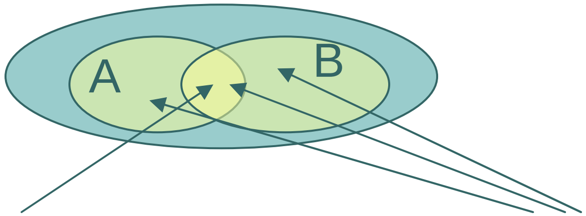
$$n = \binom{12}{4} \cdot \binom{8}{2}$$

(# successful)

$$P(B) = \frac{\binom{12}{4} \cdot \binom{8}{2}}{\binom{20}{6}} = \frac{12!8!6!14!}{4!8!2!6!20!} = 0.3576$$

Probability theory

Rules for probabilities



Intersection: Union:

$A \cap B$ $A \cup B$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(B) = P(B \cap A) + P(B \cap A')$$

If A and B are disjoint: $P(A \cup B) = P(A) + P(B)$

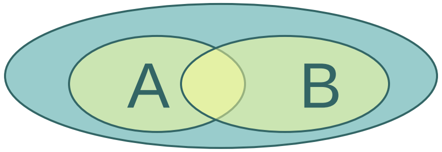
In particular: $P(A) + P(A') = 1$

Probability theory

Conditional probability

Conditional probability for A given B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) > 0$$



Bayes' Rule: $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

Rewriting Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$$

Probability theory
Conditional probability

Example page 59:
The distribution of employed/unemployed amongst men and women in a small town.

	Employed	Unemployed	Total
Man	460	40	500
Woman	140	260	400
Total	600	300	900

$$P(\text{man} \mid \text{employed}) = \frac{P(\text{man \& employd})}{P(\text{employd})} = \frac{460/900}{600/900} = \frac{460}{600} = \frac{23}{30} = 76.7\%$$

$$P(\text{man}|\text{unemployed}) = \frac{P(\text{man \& unemployed})}{P(\text{unemployed})} = \frac{40/900}{300/900} = \frac{40}{300} = \frac{2}{15} = 13.3\%$$

Probability theory
Bayes' rule

Example: Lung disease & Smoking

According to "The American Lung Association" 7% of the population suffers from a lung disease, and 90% of these are smokers. Amongst people without any lung disease 25.3% are smokers.

Events:
A: person has lung disease
B: person is a smoker

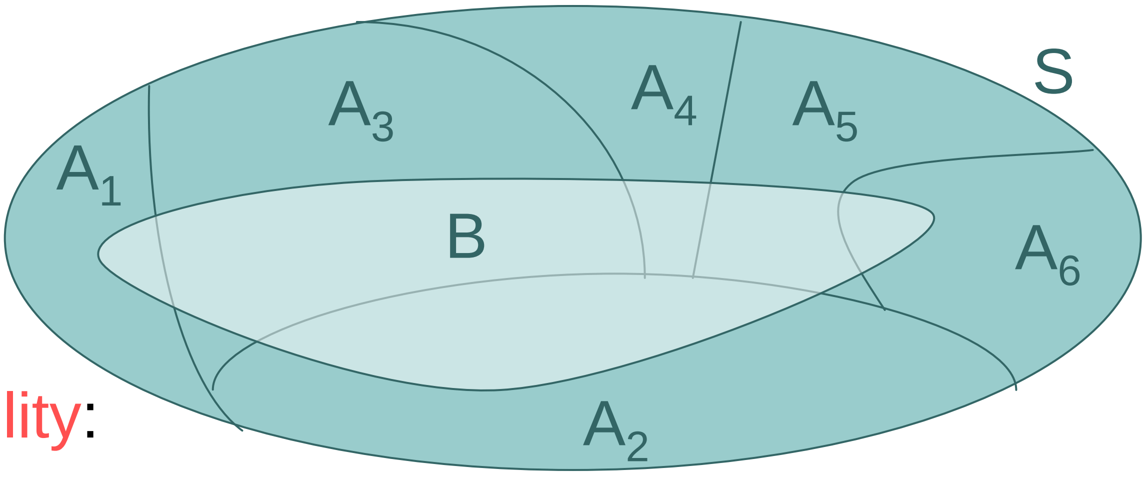
Probabilities:
P(A) = 0.07
P(B|A) = 0.90
P(B|A´) = 0.253

What is the probability that at smoker suffers from a lung disease?

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')} = \frac{0.9 \cdot 0.07}{0.9 \cdot 0.07 + 0.253 \cdot 0.93} = 0.211$$

Probability theory
Bayes' rule – extended version

A_1 , … , A_k is a partitioning of S



Law of total probability:

$$P(B) = \sum_{i=1}^{k} P(B \mid A_i) P(A_i)$$

Bayes' formula extended:

$$P(A_r \mid B) = \frac{P(B \mid A_r) P(A_r)}{\sum_{i=1}^{k} P(B \mid A_i) P(A_i)}$$

Probability theory
Independence

Definition:
Two events A and B are said to be independent if and only if

$$P(B|A) = P(B) \quad \text{or} \quad P(A|B) = P(A)$$

Alternative Definition:
Two events A and B are said to be independent if and only if

$$P(A \cap B) = P(A)P(B)$$

Notice: Disjoint event (mutually exclusive event) are dependent unless one of them has zero probability!

Probability theory
Independence

Example:

	Employed	Unemployed	Total
Man	460	40	500
Woman	140	260	400
Total	600	300	900

$$P(\text{man}|\text{employed}) = \frac{460/900}{600/900} = 76.7\%$$

$$P(\text{man}) = 500/900 = 55.6\%$$

Conclusion: the two events “man” and “employed” are dependent.

Probability theory
Rules for conditional probabilities

Probability of events A and B happening simultaneously

$$P(A \cap B) = P(A \mid B)P(B)$$

Probability of events A, B and C happening simultaneously

$$P(A \cap B \cap C) = P(A \mid B \cap C)P(B \mid C)P(C)$$

Proof:

$$P(A \cap B \cap C) = P(A \mid B \cap C)P(B \cap C) = P(A \mid B \cap C)P(B \mid C)P(C)$$

General rule:

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1 \mid A_2 \cap \cdots \cap A_k) &\cdot \\ P(A_2 \mid A_3 \cap \cdots \cap A_k) &\cdot \\ \cdots P(A_{k-1} \mid A_k) \cdot P(A_k) & \end{aligned}$$