ANALYSIS OF ENGINEERING & SCIENTIFIC DATA

STAT2201

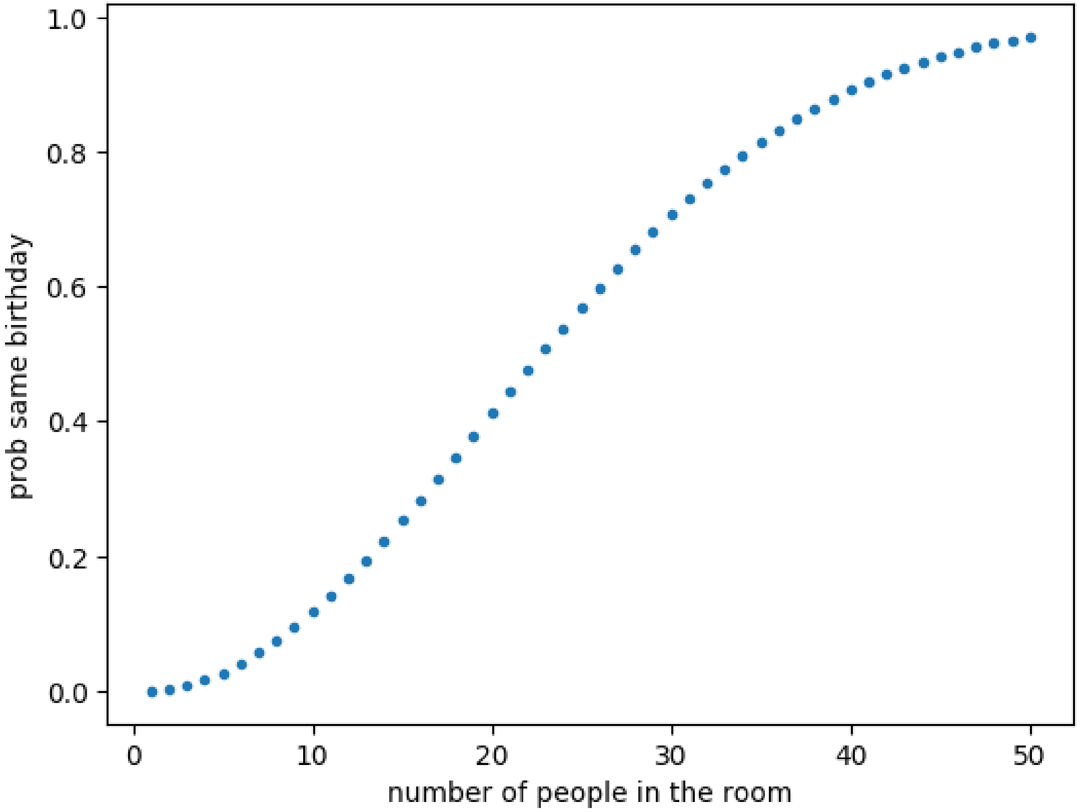
Slava Vaisman

THE UNIVERSITY OF QUEENSLAND

UNIT 2 PROBABILITY AND MONTE CARLO

- **The birthday problem**
- **Sample Space, Outcomes and Events**
- **Probability**
- **Conditional Probability and Independence**
- **The birthday problem and Monte Carlo**

THE BIRTHDAY PROBLEM



For 23 people, the probability that someone share birthday is more than 50%!

RANDOM EXPERIMENT

An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a random experiment. Some examples of random experiments are:

- tossing a coin;
- tossing a die;
- counting the number of calls arriving at a Vodafone call center between 9am and 3pm;
- select a ball from an urn containing white and black balls;
- count the number of packets to arrive at a node in a communications network during a one-minute period;
- the amount of snowfall in Moscow in February.

What are the ingredients of a random experiment?

THE SAMPLE SPACE

- Since random experiments do not consistently yield the same result, it is necessary to determine the set of possible results.
- We define an outcome or sample point of a random experiment as a result that cannot be decomposed into other results.
- When we perform a random experiment, one and only one outcome occurs. Thus outcomes are mutually exclusive in the sense that they cannot occur imultaneously.
- The sample space S (or Ω) of a random experiment is defined as the set of all possible outcomes.

SAMPLE SPACE EXAMPLES

- Tossing a coin $\Omega = \{H, T\}$.
- Tossing a die $\Omega = \{1,2,3,4,5,6\}$.
- Counting the number of calls arriving at a Vodafone call center (between 9am and 3pm) $\Omega = \mathrm{N}$, where $\mathrm{N} = 0,1,2, \ldots$
- Select a ball from an urn containing white and black balls $\Omega = \{\mathrm{W}, \mathrm{B}\}$.
- I Count the number of packets to arrive at a node in a communications network during a one-minute period) $\Omega = \mathrm{N}$.
- The amount of snowfall in Moscow in February) $\Omega = \mathrm{R}^{+}$.

TYPES OF SAMPLE SPACES

For historical (and didactic) reasons we often distinguish between two types of sample spaces:

- A sample space is discrete if it consists of a finite or countable infinite set of outcomes.
- A sample space is continuous if it contains an interval of real numbers, vectors or similar objects.

EVENTS

- Often we are not interested in a single outcome but in whether or not one of a **group** of outcomes occurs. Such subsets of the sample space are called events.
- An event is any set of outcomes (any **subset** of the sample space Ω).
- Events are usually denoted by upper case letters, A, B, C, et cetera.
- We say that event A occurs if the outcome of the experiment is one of the elements in A.
- Two events of special interest are the **certain** event, Ω, which consists of all outcomes and hence always occurs,
- and the **impossible** or **null** event, which contains no outcomes and hence never occurs.

EXAMPLES OF EVENTS

- Suppose we select the natural numbers as the sample space, i.e. $\Omega = N$. The following are events:
  - The outcome is even. $A = \{2,4,6, \ldots\}$.
  - The outcome is less or equal than 4. $B = \{1,2,3,4\}$.
- The event that the sum of two dice is 10 or more:
$$E = \{(4;\ 6), (5;\ 5), (5;\ 6), (6;\ 4), (6;\ 5), (6;\ 6)\}$$
- E is the event that the lifetime of a plain engine is at least 400 hours: $E = [400, \infty)$.
- Let E be the event that between 10000 and 20000 packets arrive in a one-minute period: $E = [10000, 10001, \cdots, 20000]$.

MORE EXAMPLES OF EVENTS

- Suppose that a coin is tossed 3 times, and that we "record" every head and tail (not only the number of heads or tails). The sample space can then be written as

$$\Omega = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$$

where, for example, HTH means that the first toss is heads, the second tails, and the third heads.

- An alternative sample space is the set $\{0,1\}^3$ of binary vectors of length 3, e.g., HTH corresponds to $(1, 0, 1)$, and THH to $(0, 1, 1)$.
- The event A that the third toss is heads is $A = \{HHH, THH, HTH, TTH\}$.

Events are sets, so we can apply the usual set operations to events.
Basic Set Operations - Union

The set $A \cup B$ (A union B) is the event that A or B or both occur.

BASIC SET OPERATIONS - INTERSECTION

The set $A \cap B$ (A intersection B) is the event that A and B both occur.

BASIC SET OPERATIONS - COMPLEMENT

The event A^c (A complement) is the event that A does not occur.

BASIC SET OPERATIONS - SUBSETS

If $A \subset B$ (A is a subset of B) then event A is said to imply event B. Clearly this means that if A occurs then B must occur.

BASIC SET OPERATIONS – DISJOINT SETS

Two events A and B which have no outcomes in common, that is, $A \cap B = \emptyset$, are called disjoint events.

VENN DIAGRAMS

Venn diagrams are used to illustrate events and the relationship between them (complement, union, intersection, partition). They were popularized by John Venn (1834-1923).

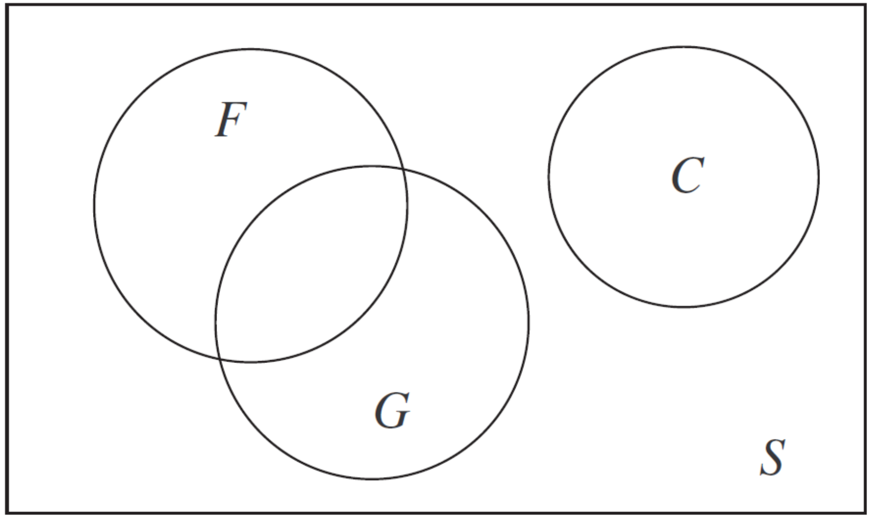


Fig 1. Venn diagram illustrating the relationship between three events, F, G and C.

VENN DIAGRAMS

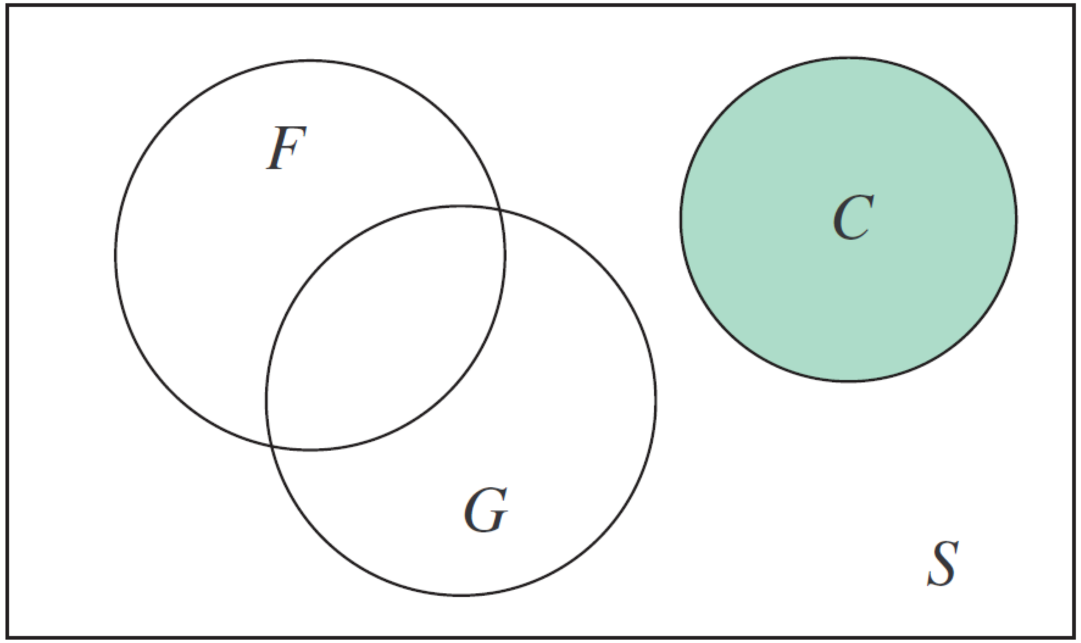


Fig 2. The event C is shaded

VENN DIAGRAMS

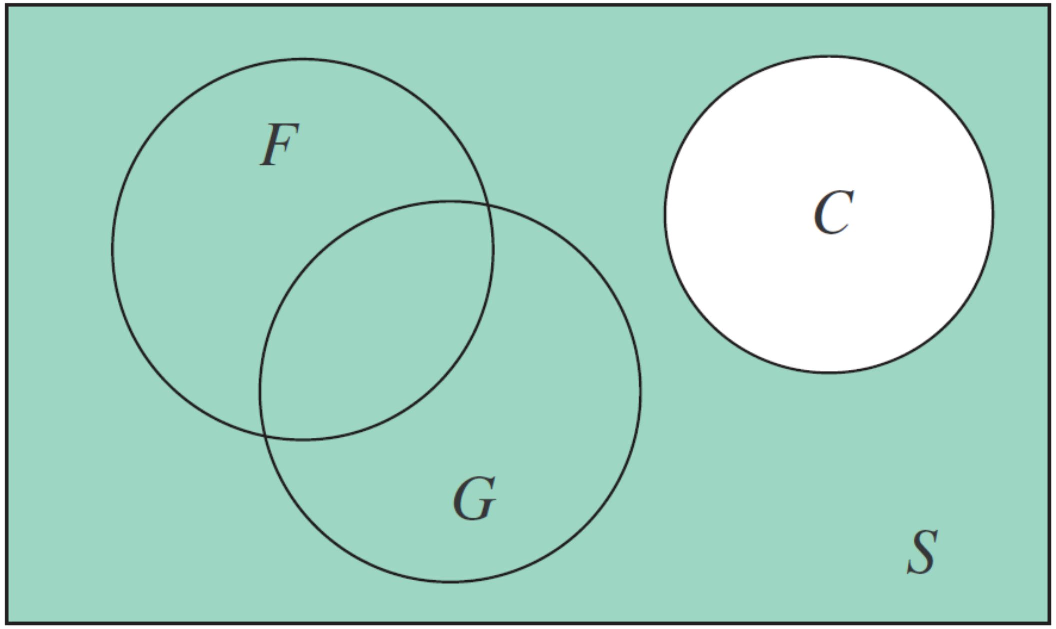


Fig 3. The event C^c is shaded

VENN DIAGRAMS

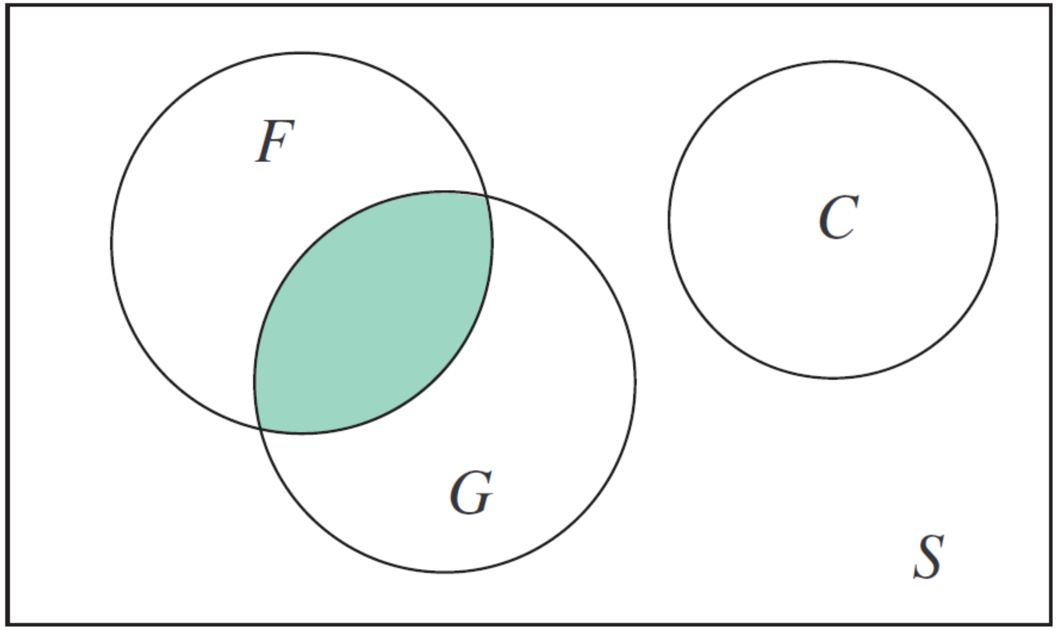


Fig 4. The event $F \cap G$ is shaded

VENN DIAGRAMS

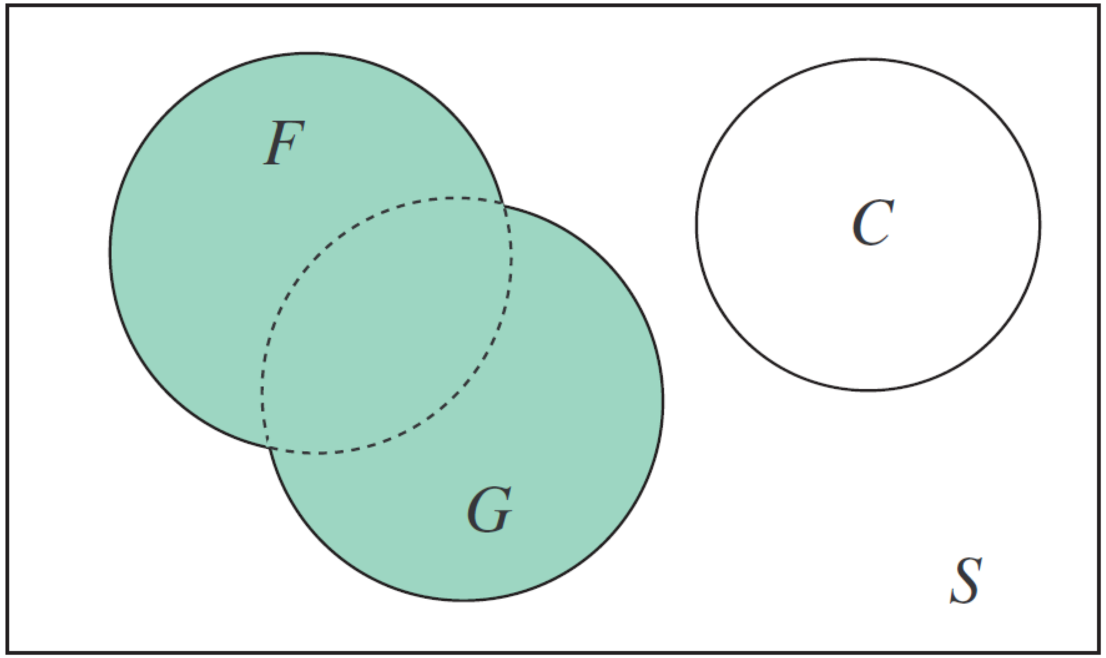


Fig 5. The event $F \cup G$ is shaded

VENN DIAGRAMS

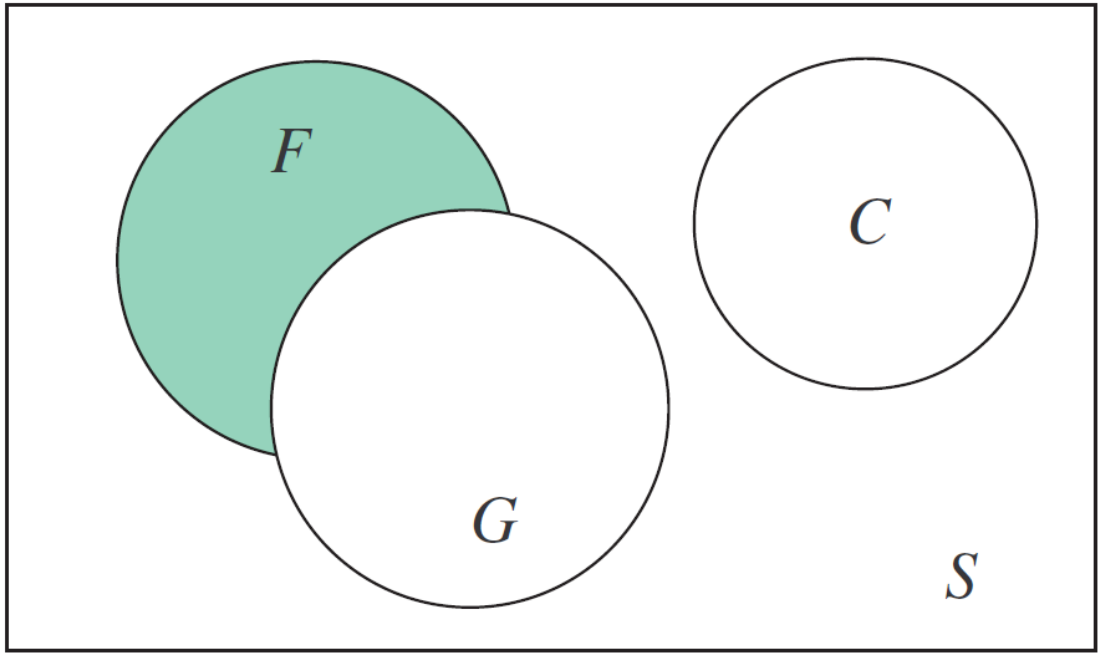


Fig 6. The event $F \cap G^c$ is shaded

VENN DIAGRAMS

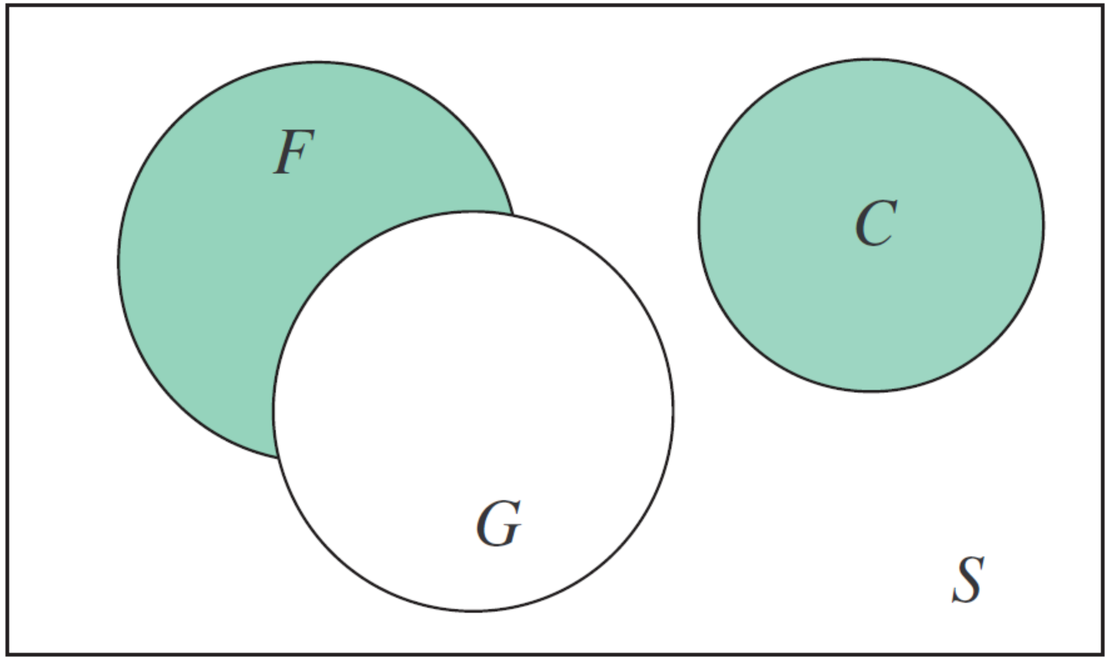


Fig 7. The event $F \cap G^c \cup C$ is shaded

THE DISTRIBUTIVE LAW FOR SET OPERATIONS

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$

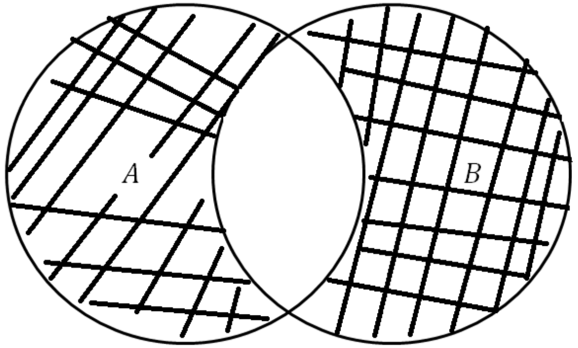
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

VENN DIAGRAMS - EXAMPLE

- Let A and B be events. Find an expression for the event "exactly one of the events A and B occurs".

  In words: (A and not B) or (B and not A)

$$(A \cap B^c) \cup (A^c \cap B)$$



- Draw a Venn diagram for this event.

DEMORGAN'S LAWS

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

MORE EXAMPLES

Example 2-7 As in Example 2-1, camera recycle times might use the sample space $S = R^+$, the set of positive real numbers. Let

$$E_1 = \{x \mid 10 \le x < 12\} \quad \text{and} \quad E_2 = \{x \mid 11 < x < 15\}$$

Then,

$$E_1 \cup E_2 = \{x \mid 10 \le x < 15\}$$

and

$$E_1 \cap E_2 = \{x \mid 11 < x < 12\}$$

Also,

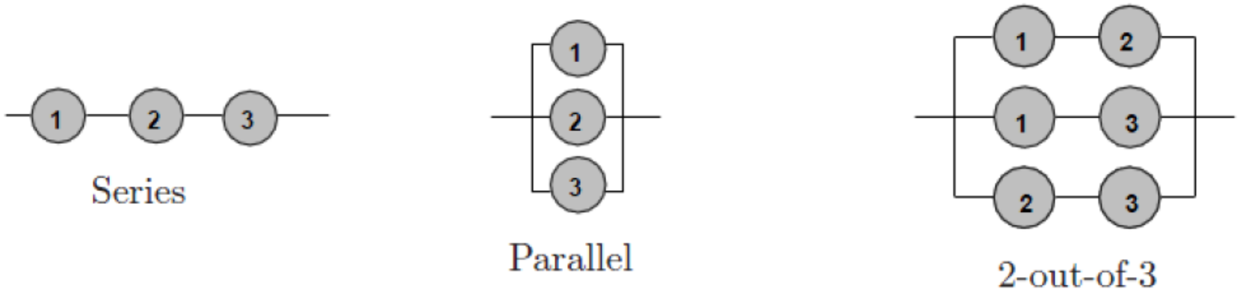
$$E_1' = \{x \mid x < 10 \quad \text{or} \quad 12 \le x\}$$

and

$$E_1' \cap E_2 = \{x \mid 12 \le x < 15\}$$

NETWORK RELIABILITY (1)

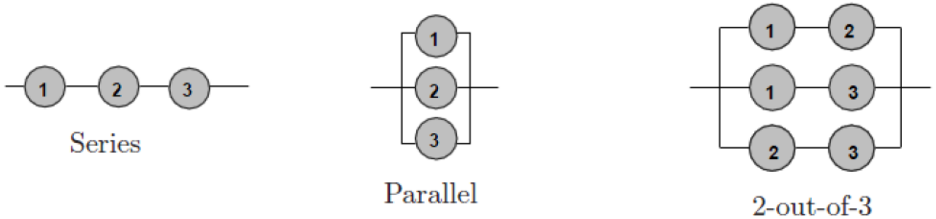
- Consider three systems, each consisting of 3 unreliable components.



Three unreliable systems

- The series system works if and only if (abbreviated as *iff*) all components work; the parallel system works if at least one of the components works; and the 2-out-of-3 system works if at least 2 out of 3 components work.

NETWORK RELIABILITY (2)



Three unreliable systems

Let A_i be the event that the i-th component is functioning, $i = 1, 2, 3$; and let D_a, D_b, D_c be the events that respectively the series, parallel and 2-out-of-3 system is functioning.

$$D_a = A_1 \cap A_2 \cap A_3.$$
$$D_b = A_1 \cup A_2 \cup A_3.$$
$$D_c = (A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3).$$

PROBABILITY

Probability is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.

- Recall that we discussed the sample space Ω (sometimes called S),
- and the set of events (subsets of Ω). The events are denoted by F .
- The third and the final ingredient in the model for a random experiment is to seek a **"measure"** which tells us how likely it is that a particular event will occur. This measure is a function from events to probabilities.

PROBABILITY FORMAL DEFINITION

Definition:

A probability $P: F \to [0,1]$, is a rule (or function) which assigns a number between 0 and 1 to each event, and which satisfies the following axioms:

1. $0 \leq P(A) \leq 1$ for any event $A \in F$.
2. $P(\Omega) = 1$
3. If $\{A_i\}$ are disjoint, $P(\cup A_i) = \sum P(A_i)$.

SOME ADDITIONAL CONSEQUENCES

If P has the properties of a probability measure then automatically fulfils the following:

1. $P(\emptyset) = 0$
2. If $A \subset B$ then $P(A) \leq P(B)$.
3. $P(A^c) = 1 - P(A)$.
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

PROOF OF 3. $P(A^c) = 1 - P(A)$

Note that A and A^c are mutually exclusive, that is $A \cap A^c = \emptyset$.

From axiom 3: $\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$.

Since $A \cup A^c = \Omega$ Axiom 2 yields

$$\mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1 = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A).$$

$\square$

VENN DIAGRAM "PROOF" OF

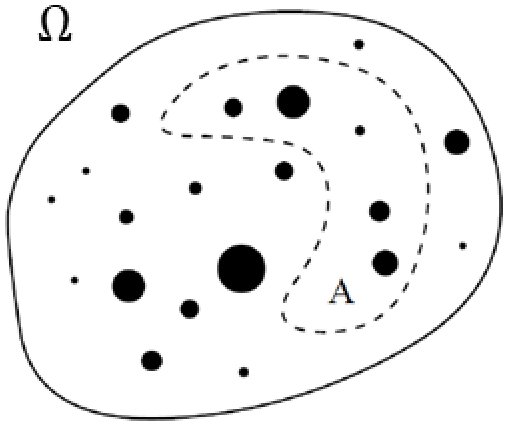
$$P\,(A \cup B) = P(A) + P(B) - P(A \cap B)$$

DISCRETE SAMPLE SPACES

- Let Ω be a discrete sample space, e.g. $\Omega = \{\mathrm{a}_1, \mathrm{a}_2, \dots \mathrm{a}_\mathrm{n}\}$.
- The easiest way to specify a probability P on a discrete sample space is to specify first the probability p_i of each elementary event $\{a_i\}$ and then to define:

$$P(A) = \sum_{i:a_i \in A} p_i \qquad for \;\; A \subset \Omega$$

This idea is represented in the figure below. To find the probability of the set A we have to sum up the weights of all the elements in A.



THE EQUILIKELY PRINCIPLE

- It is up to the modeler to properly specify these probabilities p_i.
- Fortunately, in many applications all elementary events are equally likely, and thus the probability of each elementary event is equal to 1 divided by the total number of elements in Ω.
- Specifically, If Ω has a finite number of outcomes, and all are equally likely, then the probability of each event A is defined as

$$P(A) = \frac{|A|}{|\Omega|}$$

DISCRETE SAMPLE SPACES - EXAMPLE

- We draw two cards from a full deck of 52 cards (no jokers). What is the probability of drawing at least one Ace? Draw the cards one by one. Each card has the same probability to be drawn.
- Solution:
  - $|\Omega| = 52 \cdot 51$
  - Let A be the event: "at least one Ace". Then,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{52 \cdot 51}$$

- We need to count how many elements are in A, however, finding A^c is simpler. Specifically, $A^c = 48 \cdot 47$, and therefore:

$$P(A) = 1 - P(A^c) = 1 - \frac{48 \cdot 47}{52 \cdot 51} \approx 0.15.$$

CONTINUOUS SAMPLE SPACES

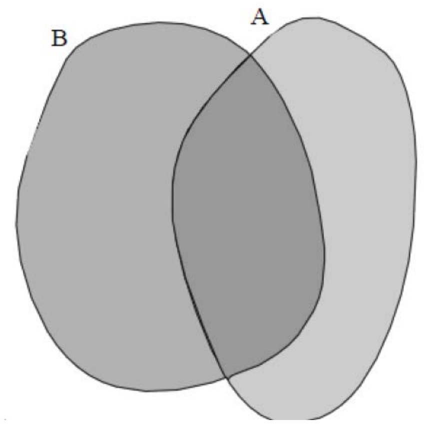
- When the sample space is not countable, for example $\Omega = R$, it is said to be continuous.
- **Example**: We draw at random a point in the interval $[0,1]$. Each point is equally likely to be drawn. How do we specify the model for this experiment?
- The sample space is obviously $\Omega = [0,1]$, which is a continuous sample space.
- We cannot define P via the elementary events $\{x\}, x \in [0,1]$ because each of these events must have probability 0!
- However, we can define P as follows: For each $0 \leq a \leq b \leq 1$, let

$$P(a, b) = b - a$$

In particular, we can find the probability that the point falls into any a set A as the length of that set.

CONDITIONAL PROBABILITY

- How do probabilities change when we know some event $B \subset \Omega$ has occurred?



- Suppose B has occurred. Thus, we know that the outcome lies in B.
- Then, A will occur if and only if $A \cap B$ occurs, and the relative chance of A occurring is therefore $\mathrm{P}(A \cap B)/P(B)$.

CONDITIONAL PROBABILITY

- This leads to the definition of the conditional probability of A given B :

Example: We throw two dice. Given that the sum of the eyes is 10, what is the probability that one 6 is cast?

- Let B be the event that the sum is 10, that is

$$B = \{(4,6),(6,4),(5,5)\}.$$

- Let A be the event that one 6 is cast:

$$A = \{(6,1),(1,6),(6,2),(2,6),\dots,(6,5),(5,6)\}.$$

- In this case $A \cap B = \{(6,4),(4,6)\}$, and

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{2/36}{3/36} = \frac{2}{3}.$$

THE MULTIPLICATION RULE FOR PROBABILITIES

$$P(A\,|B) = \frac{\mathrm{P}(A \cap B)}{P(B)} \;\Rightarrow\; \mathrm{P}(A \cap B) = P(A\,|B)P(B)$$

This can generalize this to n intersections of $A_1 \cap A_2 \cap \cdots \cap A_n$ via:

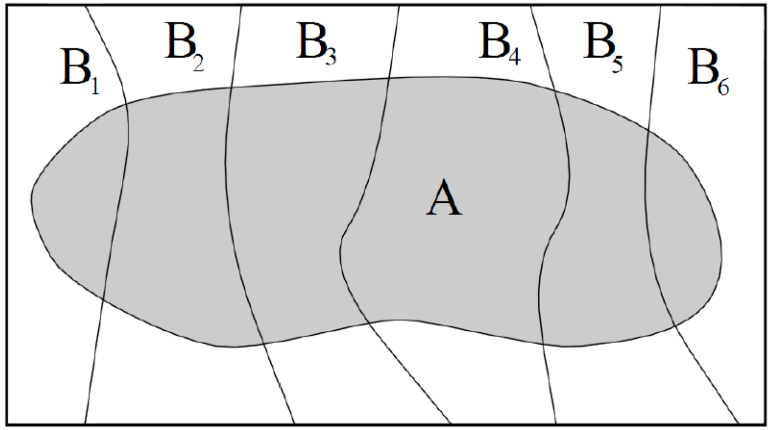
$$P(\,A_1 \cap A_2 \cap \cdots \cap A_n\,) = P(A_1)P(A_2\,|A_1)\,\cdots P(A_n\,|A_1, A_2, \dots, A_{n-1})$$

Example: We draw consecutively 3 balls from a bowl with 5 white and 5 black balls, without putting them back. What is the probability that all balls will be black?

Solution: Let A_i be the event that the ith ball is black. We wish to find the probability of $A_1A_2A_3$, which by the product rule

$$\mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_2, A_1) = \frac{5}{10}\frac{4}{9}\frac{3}{8} \approx 0.083.$$

THE LAW OF TOTAL PROBABILITY



- Suppose that $B_1, B_2, \ldots, B_n$ is a partition of Ω.
- From the third axiom of probability:

  If $\{A_i\}$ are disjoint, $P(\cup A_i) = \sum P(A_i)$

$$P(A) = \sum P(A \cap B_i)$$

- and by the multiplication rule:

$$P(A) = \sum P(A \cap B_i) = \sum P(A\,|B_i)P(B_i)$$

Example 2-28

Semiconductor Failures Continuing with semiconductor manufacturing, assume the following probabilities for product failure subject to levels of contamination in manufacturing:

Probability of Failure	Level of Contamination
0.10	High
0.01	Medium
0.001	Low

In a particular production run, 20% of the chips are subjected to high levels of contamination, 30% to medium levels of contamination, and 50% to low levels of contamination. What is the probability that a product using one of these chips fails? Let

- H denote the event that a chip is exposed to high levels of contamination
- M denote the event that a chip is exposed to medium levels of contamination
- L denote the event that a chip is exposed to low levels of contamination

Then,

$$P(F) = P(F \mid H)P(H) + P(F \mid M)P(M) + P(F \mid L)P(L)$$
$$= 0.10(0.20) + 0.01(0.30) + 0.001(0.50) = 0.0235$$

The calculations are conveniently organized with the tree diagram in Fig. 2-17.

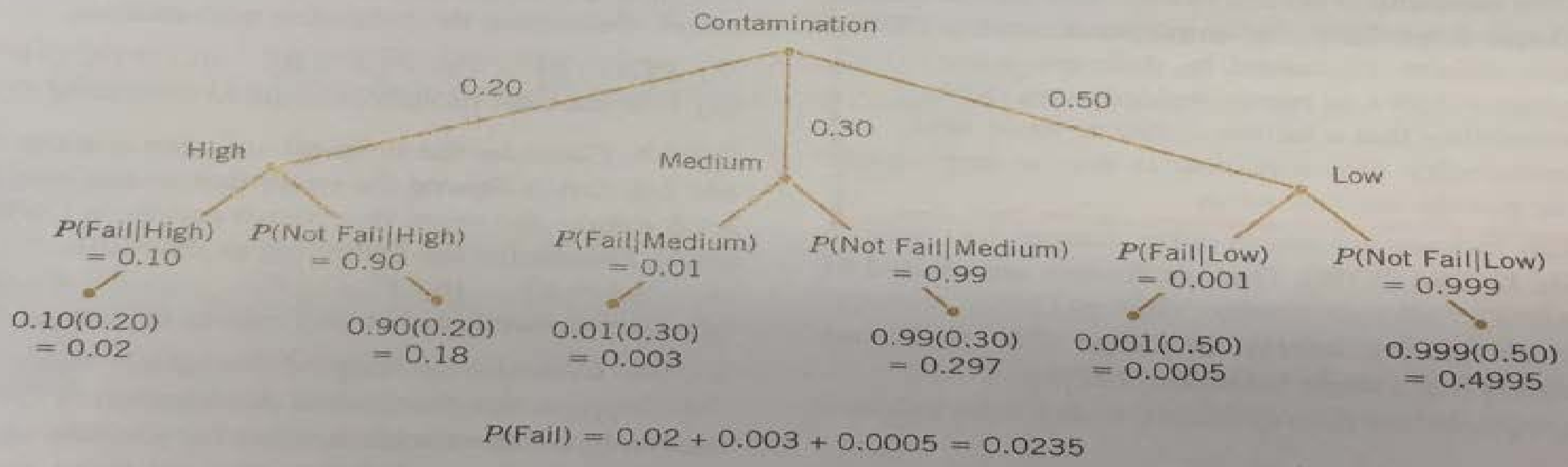


FIGURE 2-17 Tree diagram for Example 2-28.

INDEPENDENCE

Two events are independent if any one of the following equivalent statements is true:

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

The definition of independence can be extended to n events via

$$P(\, A_1 \cap A_2 \cap \cdots \cap A_n \,) = \prod P(A_i)$$

INDEPENDENT AND MUTUALLY EXCLUSIVE EVENTS – ARE NOT THE SAME!

- We toes a dice.
- Let $A = \{ 2 \}$ and $B = \{ 6 \}$.
- Clearly, A and B are mutually exclusive (I cannot get both 2 and 6).
- However, we know that given A, the probability of B happening is 0. That is,

$$P(A \mid B) = 0 \neq P(A) = \frac{1}{6},$$

- that is, A and B are **not independent**.

BACK TO THE BIRTHDAY PROBLEM

- n = Number of students in class
- $E \ = \{Two\ or\ more\ shared\ birthdays\}$.
- $P(E) \ =?$
- It is easier to calculate $P(E^c)$, where $E^c \ = \{No\ one\ has\ the\ same\ birthday\}$.
- For $n = 3$,

  $|\Omega| = 365 \cdot 365 \cdot 365 = 365^3$, since for each student, we have 365 possibilities.

  Now, the birthday of the first student is not important, however, the birthday of the second student has only $365 - 1 = \ 364$ possibilities (such that the match will not occur), similarly, the third student is left with $365 - 2 = 363$ possible placements. That is,

$$P(E^c) = \frac{365 \cdot 364 \cdot 363}{365^3}$$

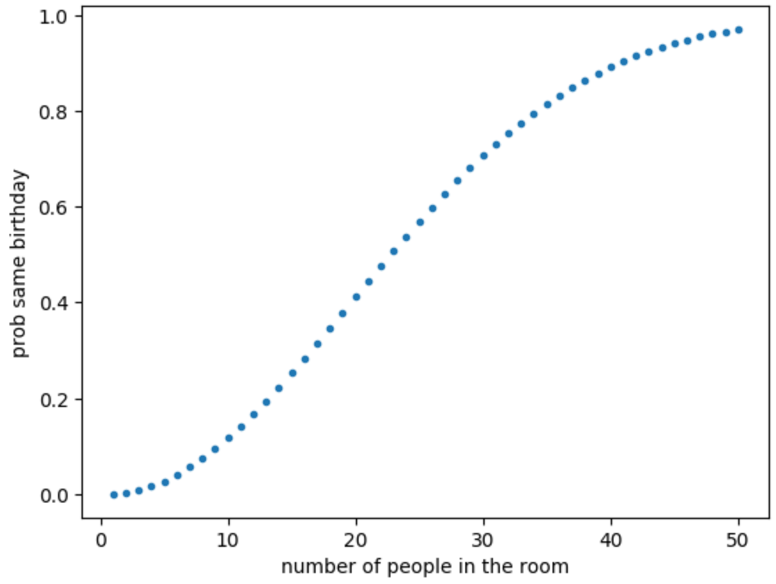
BACK TO THE BIRTHDAY PROBLEM

- For general $n \leq 365$, we have

$$
\begin{aligned}
P(E^c) &= P(\text{No same birthday}) \\
&= \frac{\text{number outcomes without same birthday}}{\text{number of possible birthday outcomes}} \\
&= \frac{365 \cdot 364 \cdot \ldots \cdot (365 - n + 1)}{365^n} \\
&= \frac{365!/(365 - n)!}{365^n}
\end{aligned}
$$

Exact Solution

In [1]:
```
using Combinatorics, PyPlot

function sameBirthDayChance(n)
    return 1 - factorial(365,365-n) / (365^n);
end

grid = 1:50
chances = []

for i=1:50
    push!(chances, sameBirthDayChance(big(i)))
end

chances

xlabel("number of people in the room")

ylabel("prob same birthday")

PyPlot.plot(grid,chances,".");
```



In [18]:
```
sameBirthDayChance(big(23))
```

Out[18]: 5.072972343239854072254172283370325002359718452929878099019740020018841839181 27e-01

MONTE CARLO AND THE BIRTHDAY PROBLEM (1)

Monte Carlo

In [26]:

```
using StatsBase

function birthdayMonteCarlo(n)
    # sample size
    N = 10000
    ell = []

    # number of students
    #n = 23
    dates = 1:365

    for i=1:N
        # generate n uniform dates
        datessample = sample(dates,n)

        countUnq = unique(datessample)

        if(size(datessample,1)>size(countUnq,1))
            push!(ell,1)
        else
            push!(ell,0)
        end
    end

    return mean(ell)
end

birthdayMonteCarlo(30)
```

Out[26]: 0.7072

MONTE CARLO AND THE BIRTHDAY PROBLEM (2)

In [21]:
```
mcchances = []
for i=1:50
    push!(mcchances, birthdayMonteCarlo(i))
end

PyPlot.plot(grid,chances,".",mcchances,"x");
```