# Reminder of Probability

Dr Ahmad Khonsari
ECE Dept.
The University of Tehran
Lect01 prob

# Perf Eval of Comp Systems

OUTLINE:

- Outcomes and events;
- Definitions of probability;
- Probability algebra;
- Measure of dependence between events;

## 1.Outcomes and events

### 1.1. Outcomes of the experiment

**Assume the following:**

- we are given an experiment;
- outcomes are random;
- **example:** we are rolling a die.

**Rolling a die:**

- there are six possible outcomes: 6,5,4,3,2,1;
- the number which we get is outcome of the experiment.
- **A set of possible outcomes is called a sample space and denoted as:**

$$\Omega = \{1,2,3,4,5,6\} \qquad (1)$$

Note: sample space includes all simple results of the experiment.

## 1.2. Events

**What is the event:** 3 points of view

- 1- mathematically: event is a subset of set $\Omega$ ;
- 2- closer to practice: **the set of outcomes of the experiment**;
- 3- well known: phenomenon occurring randomly as a results of the experiment.

**Difference between outcomes and events:**

- outcomes: are given by the experiment itself;
- events: we can define events;
- simplest case: events and outcomes are the same.

**Assume we have rolled a die:**

- set of outcomes is:

$$\Omega = \{1,2,3,4,5,6\} \qquad (2)$$

**Let us define the following events:**

- event $A_1$ consisting in that we got 4 points:

$$A_1 = \{4\} \qquad (3)$$

- this events is the same as outcome 4 of the experiment.

- event $A_2$ consisting in that we got not less than 4 points:

$$A_2 = \{4,5,6\} \qquad (4)$$

- this event is different compared to outcomes.

- event $A_3$ consisting in that we got less than 3 points:

$$A_3 = \{1,2\} \qquad (5)$$

- this event is different compared to outcomes.

**Notes:**

- using the notion of outcomes we can define events;
- usually events and outcomes are different.

# Perf Eval of Comp Systems

## 1.3.Frequency-based computation of probability

**Assume the following:**

- we are given a fair die;
- meaning that we get one of the outcomes from subset $\Omega$ is 1/6.

**Compute probabilities of events:**

- event $A_1$ consisting in that we got 4 points:

$$\Pr\{A_1\} = \frac{1}{6} \qquad (6)$$

- event $A_2$ consisting in that we got not less than 4 points:

$$\Pr\{A_2\} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \qquad (7)$$

- event $A_3$ consisting in that we got less than 3 points:

$$\Pr\{A_3\} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \qquad (8)$$

Note: if we know probabilities of outcomes we can estimate probabilities of events.

## 1.4. Operations with events

**Why we need it when events can be defined in terms of outcomes:**

- we can also define events in terms of other events using set theory.

Let A and B be two sets of outcomes defining events:

- **union** of A and B is the following set:

$$A \cup B = \{x \in A \ \boldsymbol{OR} \ x \in B\} \qquad (9)$$

- **intersection** of A and B is the following set:

$$A \cap B = \{x \in A \ \boldsymbol{AND} \ x \in B\} \qquad (10)$$

- **difference** between A and B is the following set:

$$A - B = \{x \in A \ \boldsymbol{AND} \ x \neq B\} \qquad (11)$$

- **complement** of A is the set:

$$\bar{A} = \{x \in \Omega \ \boldsymbol{AND} \ x \neq A\} \quad (12)$$

Operations can be graphically represented using Venn diagrams.

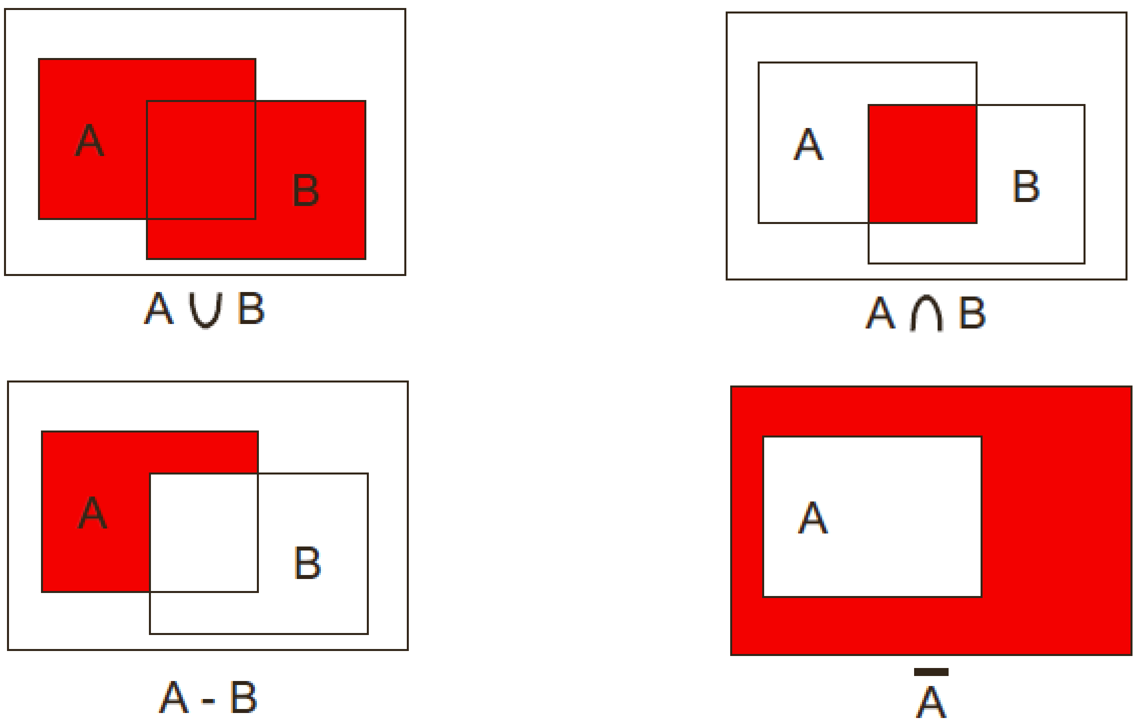



Figure 1: Graphical representation of operations with events.

## 1.5. Events using set theory

**Why we need to use set theory:**

- set theory provide a natural way to describe events.

**Define the following:**

- $\Omega$ is a set of all outcomes associated with an experiment:

- also called sample space;
- for a die the sample space is given by:

$$\Omega = \{1,2,3,4,5,6\} \qquad (13)$$

- $\mathcal{F}$ is a set of subsets (σ-algebra) of called events, such that:

- $\emptyset \in \mathcal{F}$ $and$ $\Omega \in \mathcal{F}$
- $if\ A \in \mathcal{F} then$ the complementary set $\bar{A} \in \mathcal{F}$
- $if\ A_n \in \mathcal{F}, n = 1,2, \dots, then\ \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

# Perf Eval of Comp Systems

## 2.Denfitions of probability

## 2.1. Classic definition

**Assume we have experiment E generating a set of events F such that:**

- events are mutually exclusive: when one occurs, others do not occur!
- events generate a full group: $\bigcup_{i=1}^{n} A_i = \Omega$ !
- events occur with equal chances: $\Pr\{A_1\} = \Pr\{A_2\} = \Pr\{A_n\}$ !

**Note:** these events can be just outcomes.

**Definition:** outcome w favors event A, if w leads to A.

**Definition:** probability of A:

- ratio of the number of outcomes favoring A to all number of outcomes.

**Example:** rolling a die, events A consists in getting even number:

- number of outcomes favoring A: m = 3; all number of outcomes: n = 6;
- probability of A: $\Pr\{A\}$ = 3/6 = 1/2:

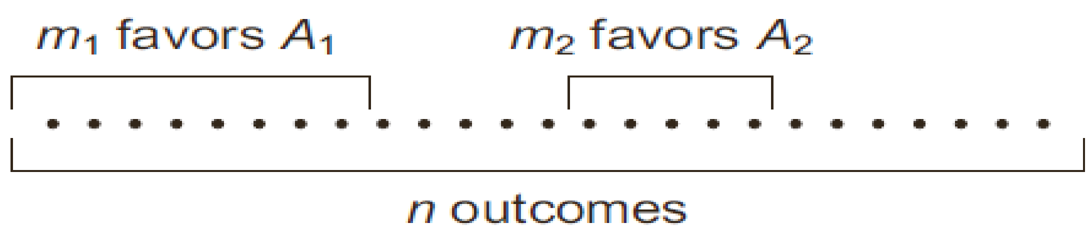
**Properties of classic definition:**

- assume we have n outcomes;
- all outcomes favor event $\Omega(\mathrm{m} = \mathrm{n})$: $\mathrm{Pr}\{\Omega\} = 1$;
- for any event out of F we have: $0 \leq \mathrm{Pr}\{A\} \leq 1, 0 \leq m \leq \mathrm{n}$;
- for complimentary event $\bar{A}$, we have:

$$\mathrm{Pr}\{\bar{A}\} = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - \mathrm{Pr}\{A\} \qquad (14)$$

- sum of exclusive events $A_1$ and $A_2$:

$$\mathrm{Pr}\{A_1 + A_2\} = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = \mathrm{Pr}\{A_1\} + \mathrm{Pr}\{A_2\} \qquad (15)$$



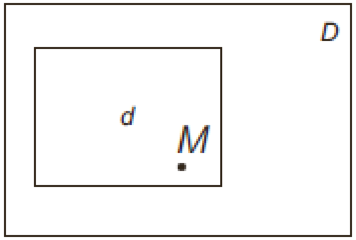
## 2.2. Geometric definition

**Assume we have experiment E consisting in throwing a dot N into a space:**

- space D is some space in $R^N$;
- dot is N-dimensional;
- probability of hitting any subspace d in D is equal.
- **Probability**: of hitting d is equal to:

$$\Pr\{M \in d\} = \frac{measure\ d}{measure\ D} \qquad (16)$$

- measure here depends on $R^N$: if $R^1$ we can use length of D and d.




- Figure 2: Throwing a dot into space $D \in R^2$

## 2.3. Statistical definition

**Why we need one more definition:**

- classic and geometric definition have limited applicability;
- reason: it is not often possible to determine equally probable events.

**Statistical definition:**

- is known for centuries;
- stated by J. Bernoulli in his last work (1713);
- applicable to wide range of events: events with stable relative frequency.

**Relative frequency of event A:**

$$Pr^*\{A\} = \frac{\mu}{\mathrm{n}} \qquad (17)$$

- $\mu$ : number of experiments in which we observed A;
- n : whole number of experiments.

# Perf Eval of Comp Systems

**Definition:** probability of event A is the value to which $Pr^*\{A\}$ converges when $n \to \infty$.

$$\Pr\{A\} \overset{n\to\infty}{=} Pr^*\{A\} = \mu/n \quad (18)$$

**Note:**

- statistical definition have the same properties as classic one;
- the only method to compute approximate probabilities if the experiment is not classic.

**If outcomes are equally likely to occur we can use two methods:**

- classic computation using $\Pr\{A\}$ = m/n;
- using relative frequency: $\Pr\{A\} \overset{n\to\infty}{=} Pr^*\{A\} = \mu/n$

**Button and Pearson compared number of heads in coin tossing:**

$$n = 4040, \qquad \frac{m}{n} = 0.5, \qquad \frac{\mu}{n} = 0.5080$$

$$n = 12000, \qquad \frac{m}{n} = 0.5, \qquad \frac{\mu}{n} = 0.5016$$

$$n = 24000, \qquad \frac{m}{n} = 0.5, \qquad \frac{\mu}{n} = 0.5008$$

## 2.4. Axiomatic definition

**THE EVENT SPACE :**

- We need an event space which is rich enough to enable the computation of the probability for any event of practical interest.
- **Definition:** *A collection $\mathcal{F}$ of subsets of $\Omega$ is a* ***field*** *(or algebra) of subsets of $\Omega$ if the following properties are all satisfied:*

***F1***: $\emptyset \in \mathcal{F}$,
***F2***: *If* $A \in \mathcal{F}$ *then* $A^c \in \mathcal{F}$, *and*
***F3***: *If* $A_1 \in \mathcal{F}$ *and* $A_2 \in \mathcal{F}$ *then* $A_1 \cup A_2 \in \mathcal{F}$

## 2.4. Axiomatic definition

- Let $\mathcal{A}$ be a subset of Pow($\Omega$). Then the intersection of all $\sigma$-algebras containing $\mathcal{A}$ is a $\sigma$-algebras, called the smallest $\sigma$-algebras generated by $\mathcal{A}$. We denote the smallest $\sigma$-algebras generated by $\mathcal{A}$ by $\sigma(\mathcal{A})$ .

**Example : Smallest $\sigma$-field.** The smallest $\sigma$-field associated with $\Omega$ is $F = \{\emptyset, \Omega\}$

**Example :** If $A$ is a subset of $\Omega$, then $F = \{\emptyset, A, \overline{A}, \Omega\}$ is a $\sigma$-field.

**Example :** Let $\Omega = \{a, b, c, d\}$.

A set $C = \{\{a\}, \{b\}\}$ is a subset of , but it is not a field.

Include:

$\{a\}^c = \{b, c, d\}$ , $\{b\}^c = \{a, c, d\}$ , $\{a\} \cup \{b\} = \{a, b\}$ and $(\{a\} \cup \{b\})^c = \{c, d\}$.

Thus the smallest $\sigma$-field containing all the elements of *C is:*

$\{\emptyset, \{a\}, \{b\}, \{a, b\}, \{c, d\}, \{b, c, d\}, \{a, c, d\}, \Omega\}$

## 2.4. Axiomatic definition

- **Example:**

Let $\Omega$={1, 2, 3, 4, 5, 6 } and $\mathcal{A}$={{1, 2}, {2, 3}}.

$\sigma(\mathcal{A})$ = { $\emptyset$, {1 , 2 , 3, 4 ,5, 6} ,
{1 , 2 } , { 3 , 4 , 5 , 6} ,
{ *2 , 3*} , {1 , 4 , 5, 6} ,
{1, 3, 4 , 5 , 6} , {*2*} , {2, 3, 4 , 5 , 6} , {1}, { 1 , 2 , 4 , 5 , 6} , {*3*} ,
{1, *2* , 3} , { 4 , 5 , 6} , {1 , 3} , {2, 4 , 5 , 6}}

*Since:* { 3 , 4 , 5 , 6} ∪ {1 , 4 , 5, 6} = {1, 3, 4 , 5 , 6} , $\{1, 3, 4, 5, 6\}^c$= {*2*} ,
{*2*} ∪ {3, 4 , 5 , 6}= {2, 3, 4 , 5 , 6} , $\{2, 3, 4, 5, 6\}^c$= {*1*} ,
{*2*} ∪ {1, 4 , 5 , 6}= {1, 2, 4 , 5 , 6} , $\{1,2, 4, 5, 6\}^c$= {3} ,
{1} ∪{3}={1,3}

## 2.4. Axiomatic definition

**THE EVENT SPACE :**

- **Example 2.** *In tossing a fair die*
- *(a) How many possible events are there?*
- *(b) Is the collection of all possible subsets of* $\Omega$ *a field?*
- *(c) Consider* $\mathcal{F} = \{\emptyset, \{1, 2, 3, 4, 5, 6\}, \{1, 3, 5\}, \{2, 4, 6\}\}$. *Is* $\mathcal{F}$ *a field?*

**Solution.**

(a) Using the Binomial Theorem, we find that there are $n = \sum_{k=0}^{6} C_{6,k} = (1+1)^6 = 64$ possible subsets of $\Omega$. The number of possible events is 64 compared to only six possible outcomes.

Each of the collections (b) and (c) is a field, by checking F1, F2, and F3 ■

# THE EVENT SPACE

*F3a: If $A_1$, $A_2$, . . . are all in $\mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.*

**Definition 2.**

*A collection $\mathcal{F}$ of subsets of $\Omega$ is a **sigma-field** (or **sigma-algebra**) of subsets of $\Omega$*

*if F1, F2, and F3a are all satisfied.*

**Definition 3.**

The pair $(\Omega, \mathcal{F})$ is called a measurable space ($\mathcal{F}$-measurable).

## 2.4. Axiomatic definition

**Some facts:**

- the most accepted definition;
- includes classic, geometric and statistical as special cases;
- introduced by A.N. Kolmogorov in 1933.
- **Let $\Omega$ be the set of outcomes, $\mathcal{F}$ be σ-algebra:**
- P is a probability measure on $(\Omega, \mathcal{F})$ such that:
  - axiom 1: $\Pr\{A\} \geq 0$
  - axiom 2: $\Pr\{\Omega\} = 1$
  - axiom 3: $\Pr\{\phi\} = 0$
  - axiom 4: $\Pr\{\sum_k A_k\} = \sum_k \Pr\{A_k\}$ for mutually exclusive events.
  - (**another notation:** If $A_1, A_2, \ldots$ are mutually exclusive events in $\mathcal{F}$, then $P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$)

**Note:** P is the a mapping from $\mathcal{F}$ in [0, 1].

**Definition 4.** $(\Omega, \mathcal{F}, \mathrm{P})$ is called the probability space.

## 2.4. Axiomatic definition

More detail of Definition 4.

**Definition 5. Measure Space:** A triplet $(\Omega, \mathcal{F}, \mu)$ is a measure space if $(\Omega, \mathcal{F})$ is a measurable space and $\mu: \mathcal{F} \to [0; \infty)$ is a measure.

**Definition 6. Probability Space:** A measure space is a probability space if $\mu(\Omega)=1$. In this case, μ is a probability measure, which we denote P.

Let P be a probability measure. The *cumulative distribution function* (c.d.f.) of P is defined as: $F(x) = P((-\infty, x])$ , $x \in R$

## 2.4. Axiomatic definition

**Question**

- What is Union Bound?

## 2.4. Axiomatic definition

**Example :** A fair die is tossed once. What is the probability of an even number occurring?

**A:** The sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.
We assign a weight of *w* to each sample point; i.e., $P(i) = w$, $i = 1, 2, . . . , 6$. By the 2nd and 4th axioms we have $P(\Omega) = 1 = 6w$; hence, $w = 1/6$. Letting $A = \{2, 4, 6\}$, $P(A) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 1/6 + 1/6 + 1/6 = 1/2$

## 2.4. Axiomatic definition

**Example**: The sample space of a die is $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{F} = \mathcal{P}ow(\Omega)$, where $\mathcal{P}ow$ is the *power set* of $\Omega$. The probability measure $\mu$ is completely determined by the values of $\mu\{1\}, \mu\{2\}, \ldots, \mu\{6\}$.

For example, suppose:

$\mu\{1\} = 1/12$ $\mu\{4\} = 1/6$

$\mu\{2\} = 1/12$ $\mu\{5\} = 1/6$

$\mu\{3\} = 1/3$ $\mu\{6\} = 1/6$

Then the probability of rolling a 2 or a 3 is $\mu\{2, 3\} = 1/12 + 1/3 = 5/12$.

# 3. Probability algebra

## 3.1. Adding events

**For mutually exclusive (disjoint) events:**

$$\Pr\{\textstyle\sum_k A_k\} = \textstyle\sum_k \Pr\{A_k\} \qquad (20)$$

- holds only when events are exclusive!!!

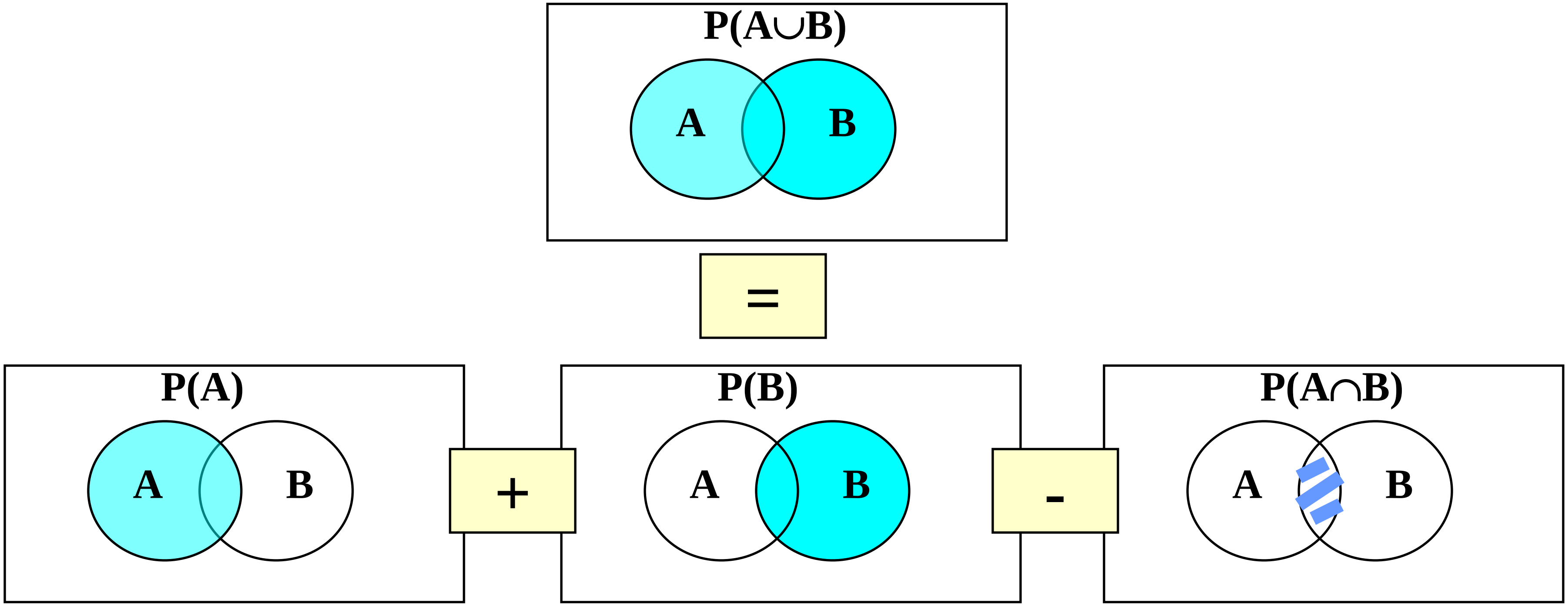
## 3. Probability algebra

### 3.1. Adding events

**For two arbitrary events:**

$$\Pr\{A + B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{AB\} \qquad (21)$$

Thus $\Pr\{A + B\} \leq \Pr\{A\} + \Pr\{B\}$ (one of the bounds)



## 3. Probability algebra

### 3.1. Adding events

**For three arbitrary events: principle of inclusion-exclusion**

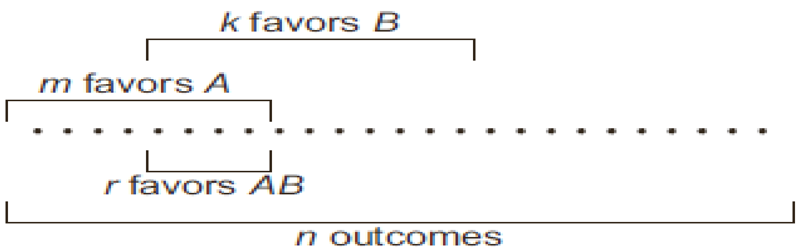
$$\Pr\{A + B + C\} = \Pr\{A\} + \Pr\{B\} - \Pr\{AB\} - \Pr\{BC\} - \Pr\{AC\} + \Pr\{ABC\} \qquad (22)$$

**Note**: one can extend it to n events.

## 3.2. Conditional probability

**Definition**: probability that the event A will occur given that the event B has already occurred.

**Consider the classic experiment with equally probable outcomes:**




- probabilities: $Pr\{A\} = m/n$, $Pr\{B\} = k/n$, $Pr\{AB\} = r/n$;
- if event B already occurred then for event A the number of outcomes decreases to k; reduced probability space
- among k there are r outcomes favoring A: using classic definition $Pr\{A|B\} = r/k$;
- dividing nominator and denominator by n we get:

$$Pr\{A|B\} = \frac{r}{k} = \frac{r/n}{k/n} = \frac{Pr\{AB\}}{Pr\{B\}} \qquad (23)$$

**Result:**

$$Pr\{A|B\} = \frac{Pr\{AB\}}{Pr\{B\}} \qquad (24)$$

**Note the following:**

- one can check that $Pr\{.|B\}$ is a probability measure;
- the conditional probability is not defined if $Pr\{B\}$ = 0.

**Note:**

- Event B with prob. 0 (i.e. $Pr\{B\}$ = 0 ) is different from Impossible event ( i.e. $Pr\{\phi\}$=0 )

**We can change the role of A and B:**

$$Pr\{B|A\} = \frac{Pr\{AB\}}{Pr\{A\}} \qquad (25)$$

Useful notation:

$$p_{X,Y}(x, y \mid A) = \begin{cases} \dfrac{p_{X,Y}(x, y)}{P\{A\}} & (x, y) \in A, \quad P\{A\} > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (2.5)$$

- **Example**: there are three trunks:
- we seize the trunk with probability 1/3;
- what is the probability to seize a given trunk in two attempts (at least once)?
- let: A: we seize the trunk in first attempt, B: we seize the trunk in second attempt:

$$\Pr\{A\} + \Pr\{B\} - \Pr\{AB\} = \frac{1}{3} + \frac{1}{3} - \frac{1}{9} = \frac{5}{9} \qquad (26)$$

put it in another way (first attempt and not the second one (i.e. $\frac{1}{3} * \frac{2}{3}$+ vice versa (i.e . $\frac{2}{3} * \frac{1}{3}$ )+ both (i.e. . $\frac{1}{3} * \frac{1}{3}$)

- Or:

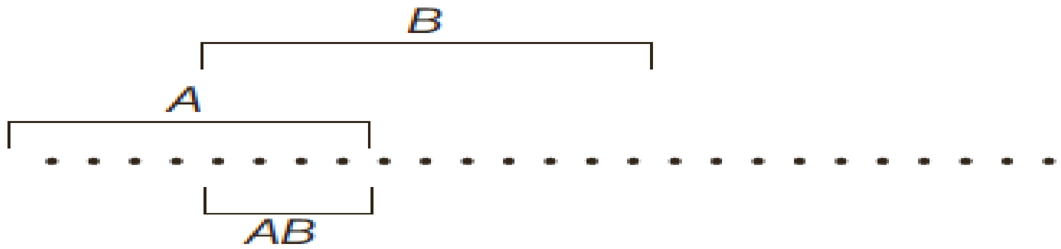
$1 -$ prob. (we do not seize the trunk in both attempts )

$$1 - (1 - \Pr\{A\})(1 - \Pr\{B\}) = 1 - \left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{3}\right) = 1 - \frac{4}{9} = \frac{5}{9}$$

## 3.3. Multiplication of events

- Multiplication of two arbitrary events:

$$\Pr\{AB\} = \Pr\{B\}\ \Pr\{A|B\} = \Pr\{A\}\ \Pr\{B|A\} \qquad (27)$$




- Multiplication ot n arbitrary events:

$$\Pr\{A_1 A_2 \dots A_n\} = \Pr\{A_1\}\ \Pr\{A_2|A_1\} \Pr\{A_3|A_1 A_2\} \dots \Pr\{A_n|\ A_1 \dots A_{n-1}\} \qquad (28)$$

verification:

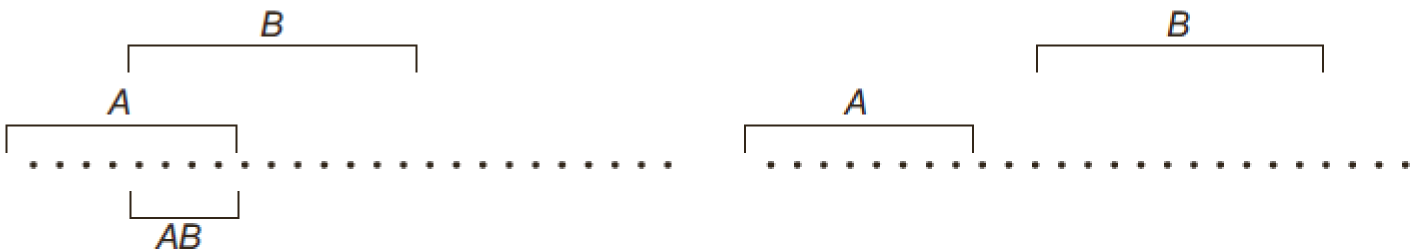
$$\Pr\{A_1 A_2 \dots A_n\} = \Pr\{A_1 A_2\} \Pr\{A_3|A_1 A_2\} \dots \Pr\{A_n|\ A_1 \dots A_{n-1}\}$$
$$\Pr\{A_1 A_2 \dots A_n\} = \Pr\{A_1 A_2 A_3\} \dots \Pr\{A_n|\ A_1 \dots A_{n-1}\}$$

# Perf Eval of Comp Systems

## 3.4. Independent and dependent events

**Definition**: event A is independent of event B if the following holds:

$$Pr\{A|B\} = Pr\{A\} \qquad (29)$$




**Note**: if A is independent of B then B is independent of A.

**Multiplication (chain Rule)**: if events A and B are independent then their product:

$$\Pr\{AB\} = \Pr\{B\}\ \Pr\{A|B\} = \Pr\{A\}\ \Pr\{B\} \qquad (30)$$

**Note**: $\Pr\{AB\} = \Pr\{A\}\ \Pr\{B\}$is sufficient for **2** events to be independent.

## 3.4. disjoint sets (**events**)

**Definition:** If $E1 \cap E2 = \emptyset$, then sets $E1$ and $E2$ are ***pairwise exclusive (disjoint) events:***

Note: ***pairwise exclusive events*** implies ***mutually (collectively) exclusive (disjoint) events.***

**Definition:** If $E_1, E_2, \ldots, E_n$ are sets (events) such that $E_i \cap E_j = \emptyset$, $\forall i, j$, and such that $\bigcup_{i=1}^{n} E_i = \Omega$, then we say that events $\boldsymbol{E_1, E_2, \ldots, E_n}$ ***partition*** set $\Omega$ .

$E_1$

$\Omega=$

| | | | | | |
|---|---|---|---|---|---|
| (1,1) | (1,2) | (1,3) | (1,4) | (1,5) | (1,6) |
| (2,1) | (2,2) | (2,3) | (2,4) | (2,5) | (2,6) |
| (3,1) | (3,2) | (3,3) | (3,4) | (3,5) | (3,6) |
| (4,1) | (4,2) | (4,3) | (4,4) | (4,5) | (4,6) |
| (5,1) | (5,2) | (5,3) | (5,4) | (5,5) | (5,6) |
| (6,1) | (6,2) | (6,3) | (6,4) | (6,5) | (6,6) |

$E_2$

Events $E_1$ and $E_2$ in sample space $\Omega$

# Independence-more

Two events, $E$ and $F$, are **independent** if and only if:
$P(EF) = P(E)P(F)$
Otherwise, they are called **dependent** events.
This property applies regardless of whether or not $E$ and $F$ are from an equally likely sample space and whether or not the events are mutually exclusive.
The independence principle extends to more than two events. In general, $n$ events $E_1, E_2, \ldots, E_n$ are independent if for every subset with $r$ elements (where $r \leq n$) it holds that:
$P(E_a, E_b, \ldots, E_r) = P(E_a)P(E_b)\ldots P(E_r)$

# Independence-more

The general definition implies that for three events $E$, $F$, $G$ to be independent, *all* of the following must be true:

$P(EFG) = P(E)P(F)P(G)$

$P(EF) = P(E)P(F)$

$P(EG) = P(E)P(G)$

$P(FG) = P(F)P(G)$

# Independence-more

Roll two 6-sided dice, yielding values D1 and D2

$EF=\{(1,6)\}$

Let event E: D1=1

event F: D2=6

event G: D1+D2=5 $G=\{(1,4),(2,3),(3,2),(4,1)\}$

$EG=\{(1,4)\}$

1. Are E and F independent?

$P(E)=1/6$

$P(F)=1/6$

$P(EF)=1/36= P(E)P(F)$

Yes

2. Are E and G independent?

$P(E)=1/6$

$P(G)=4/36$

$P(EG)=1/36\neq P(E)P(G)$

No

# Independence-more

Each roll of a 6-sided die is an independent trial

Two rolls: D1 and D2

Let event E: D1=1

event F: D2=6

event G: D1+D2=7 G={(1,6),(2,5),(3,4),(4,3) ,(5,2),(6,1)}

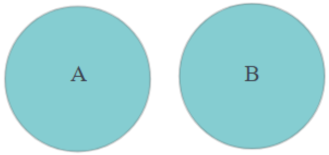
1. Are E and F independent?
   P(E)=1/6 Yes
   P(F)=1/6
   P(EF)=1/36
2. Are E and G independent?
   P(E)=1/6 Yes
   P(G)=1/6
   P(EG)=1/36
3. Are F and G independent?
   P(F)=1/6 Yes
   P(G)=1/6
   P(FG)=1/36
4. Are E,F and G independent?
   No
   P(EFG)=1/36
   ≠(1/6)(1/6)(1/6)

Pairwise independence is not sufficient to prove independence of >2 events!

# Disjoint & Independent Events

**Disjoint events** and **independent events** are different.

Events are considered disjoint if they never occur at the same time; these are also known as *mutually exclusive events*. E.g. In Venn diagram diagram, there is no overlap between event A and event B. These two events never occur together, so they are disjoint events.



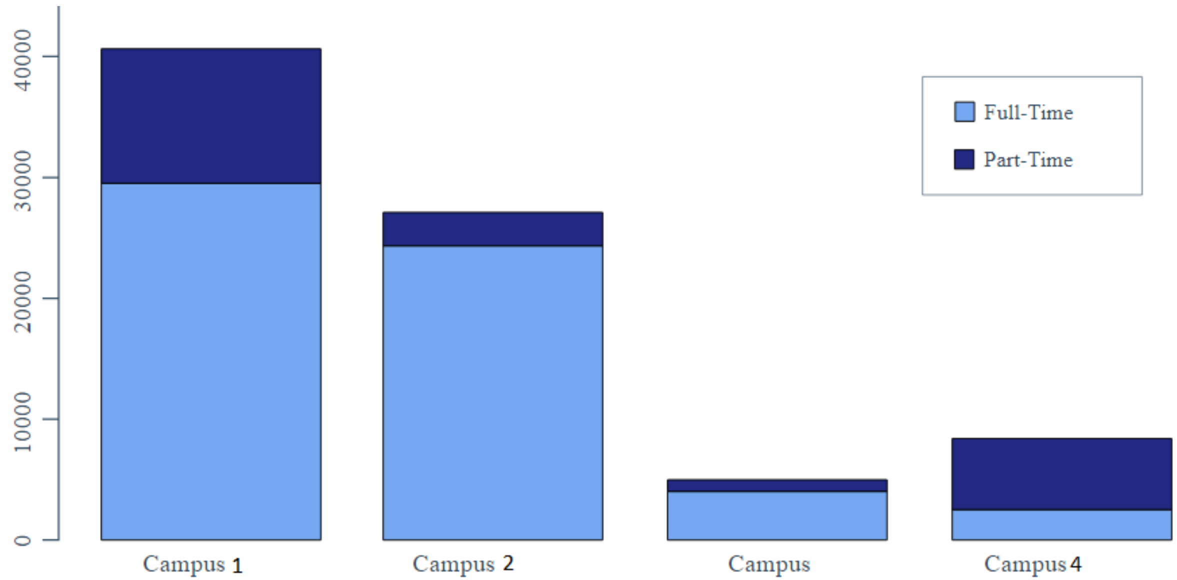

Events are considered independent if they are unrelated. The outcome of one event does not impact the outcome of the other event. Independent events can, and do often, occur together.

# Disjoint & Independent Events

e.g. The proportion of students who are part-time is different at each campus. Enrollment status and primary campus are **not independent**.

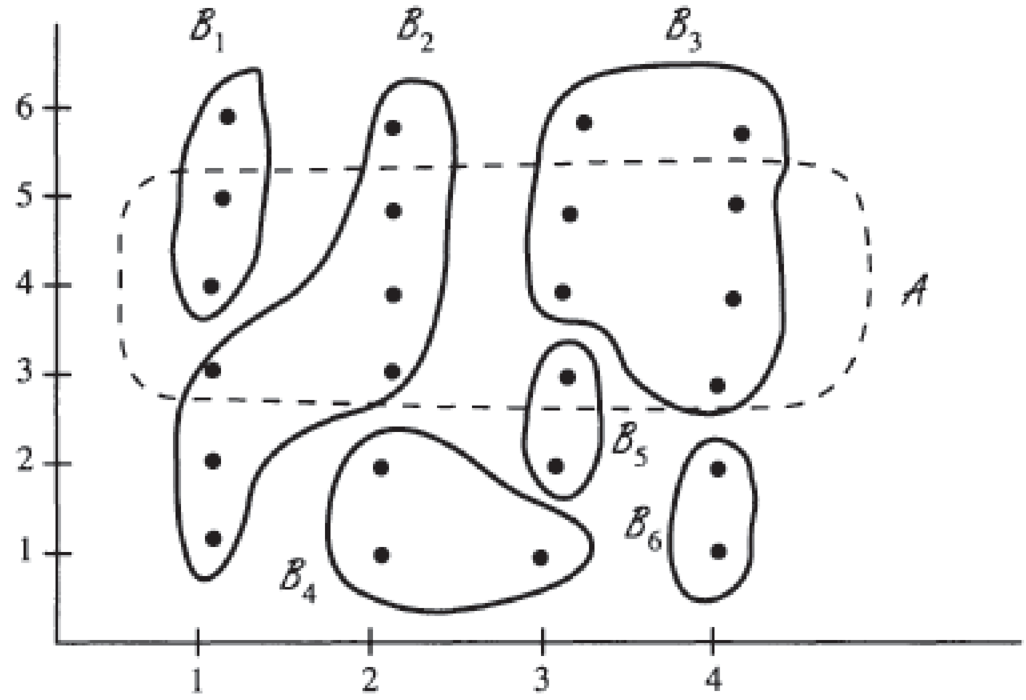
If we know a student's campus, that changes the probability of them being a full- or part-time student.

If we know that a student is full- or part-time, that changes the probability that they came from a specific campus.



# Perf Eval of Comp Systems

## 3.4. disjoint events




Partition of a Set of Pairs

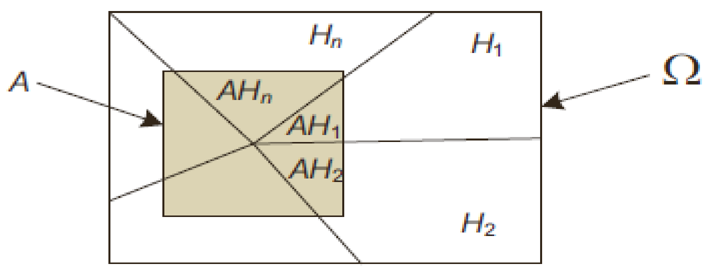
## 3.5. Law of total probability

**Let $H_1, H_2, \dots, H_n$ be the events such that:**

- $H_i \cap H_j = 0 \; if \; i \neq j$ (mutually exclusive events);
- $\Pr\{H_i\} > 0 \; for \; i = 1,2, \dots, n$ (non-zero);
- $H_1 \cup H_2 \cup \dots \cup H_n = \Omega$ (full group)

**Then for any event A the following result holds:**

$$\mathrm{A} = \mathrm{A} \cap \Omega = A(H_1 \cup H_2 \cup \dots \cup H_n) = AH_1 \cup AH_2 \cup \dots + \cup AH_n = \cup_{i=1}^{n} AH_i \qquad (31)$$



**We have:**

$$Pr\{A\} = \Pr\{\sum_{i=1}^{n} AH_i\} = \sum_{i=1}^{n} \Pr\{AH_i\} = \sum_{i=1}^{n} \Pr\{A|H_i\} \Pr\{H_i\} \quad (32)$$

**Law of total probability:**

$$Pr\{A\} = \sum_{i=1}^{n} \Pr\{A|H_i\} \Pr\{H_i\} \quad (33)$$

**Notes:**

- events $H_i$, i = 1, 2,…,n are called hypotheses;
- probabilities $\Pr\{H_1\}, \Pr\{H_2\}, \dots, \Pr\{H_n\}$ are called apriori probabilities.

**Law of total probability helps to find probability of event A if we know:**

- probabilities of hypotheses $H_i$,i = 1,2,…, n;
- probabilities $\Pr\{A|H_i\}$.

**Example**: similar components are made by 3 vendors, we have:

- vendor 1: 50% of components: probability of non-conformance is 0:002;
- vendor 2: 30% of components: probability of non-conformance is 0:004;
- vendor 3: 20% of components: probability of non-conformance is 0:005.

**Question:** if we take 1 component what is the probability that it is non-conformant.

- $H_k$ the chosen detail is made by vendor k = 1,2,3;

$$Pr\{H_1\} = 0.5\,, \Pr\{A|H_1\} = 0.002,$$
$$Pr\{H_2\} = 0.3\,, \Pr\{A|H_2\} = 0.004,$$
$$Pr\{H_3\} = 0.2\,, \Pr\{A|H_3\} = 0.005.$$

- A: the chosen component is non-conformant.

**Using the law of total probability:**

$$Pr\{A\} = \Pr\{H_1\}\Pr\{A|H_1\} + \Pr\{H_2\}\Pr\{A|H_2\} + \Pr\{H_3\}\Pr\{A|H_3\} = 0.0032 \quad (35)$$

## 3.6. Bayes' formula

**Assume:** we carried out experiment and event A occurred:

- we have to re-evaluate probabilities of hypotheses: $\Pr\{H_1\}, \Pr\{H_2\}, \dots, \Pr\{H_n\}$;
- we are looking for $\Pr\{H_1|A\}, \Pr\{H_2|A\}, \dots, \Pr\{H_n|A\}$;

**Use formula for conditional probability** $\Pr\{A|B\} = \frac{\Pr\{AB\}}{\Pr\{B\}}$ **get:**

$$\Pr\{H_k|A\} = \frac{\Pr\{AH_k\}}{\Pr\{A\}} = \frac{\Pr\{H_k\}\Pr\{A|H_k\}}{\Pr\{A\}} \qquad (36)$$

**Use law of total probability** $\Pr\{A\} = \sum_{i=1}^{n} P\{A|H_i\}\Pr\{H_i\}$ **to get:**

$$\Pr\{H_k|A\} = \frac{\Pr\{H_k\}\Pr\{A|H_k\}}{\sum_{i=1}^{n} P\{A|H_i\}\Pr\{H_i\}} \quad k = 1,2,\dots,n \qquad (37)$$

- this formula is known as Bayes's formula.

**Note:** Probabilities $\Pr\{H_1|A\}, \Pr\{H_2|A\}, \dots, \Pr\{H_n|A\}$ are called aposteriori probability.

# Perf Eval of Comp Systems

**Example:** similar components are made by 3 vendors, we get

- vendor 1: 50% probability of non-conformance is 0:002;
- vendor 2: 30% probability of non-conformance is 0:004;
- vendor 3: 20% probability of non-conformance is 0:005;
- if we take 1 component, the probability that it is non-conformant:

$$Pr\{A\} = \Pr\{H_1\}\Pr\{A|H_1\} + \Pr\{H_2\}\Pr\{A|H_2\} + \Pr\{H_3\}\Pr\{A|H_3\} = 0.0032 \quad (38)$$

**Question:** we took 1 component and it is non-conformant, which vendor to blame?

$$\Pr\{H_1|A\} = \frac{Pr\{H_1\}\Pr\{A|H_1\}}{\Pr\{A\}} = \frac{0.5 * 0.002}{0.0032} = \frac{10}{32} = \frac{5}{16}$$
$$\Pr\{H_2|A\} = \frac{Pr\{H_2\}\Pr\{A|H_2\}}{\Pr\{A\}} = \frac{0.3 * 0.004}{0.0032} = \frac{12}{32} = \frac{6}{16}$$
$$\Pr\{H_3|A\} = \frac{Pr\{H_3\}\Pr\{A|H_3\}}{\Pr\{A\}} = \frac{0.2 * 0.005}{0.0032} = \frac{10}{32} = \frac{5}{16}$$

**Answer:** most probably vendor 2.

## 3.7. Measure of dependence between events

**If events A and B are dependent:**

- we can measure the dependence as:

$$Pr\{A|B\} = \frac{Pr\{AB\}}{Pr\{B\}},\ Pr\{B|A\} = \frac{Pr\{AB\}}{Pr\{A\}} \qquad (43)$$

- **the problem:** these metrics are not symmetric.

**Symmetric measure of dependence between events:**

$$\rho_{AB} = \frac{\Pr\{AB\} - \Pr\{A\}\Pr\{B\}}{\sqrt{\Pr\{A\}\Pr\{\bar{A}\}\Pr\{B\}\Pr\{\bar{B}\}}} \qquad (44)$$

**Properties of $\rho_{AB}$ :**

- $\rho_{AB} = 0$ when and only when A and B are independent;
- $-1 \leq \rho_{AB} \leq 1$ ;
- $\rho_{AB} = 1$ when $Pr\{A\} = Pr\{B\} = Pr\{AB\}$ , $\rho_{AB} = -1$ when $A = \bar{B}$;
- $\rho_{\bar{A}B} = -\rho_{AB}$;