

Statistics Lecture

Claudio Ortelli ¹

¹Finance Institute
USI

Advanced Learning and Research Institute, 2012

Basic definitions

Random Experiment

Definition

A random experiment is an experiment whose outcome can not predicted with certainty.

Examples

- Observing if a system component is functioning properly or has failed at a given point in time in the future.
- Determining the execution time of a program.
- Determining the response time of a server request.

Definition

The result of the experiment is called the outcome of the experiment. The totality of the possible outcomes of a random experiment is called the sample space of the experiment and it will be denoted by the letter S .

Basic definitions

Sample space

- The definition of the sample space S is determined by the experiment and the purpose for which the experiment is carried out. When observing the status of two components of a running system, it may be sufficient to know if zero, one or two components have failed without having to exactly identify which component has failed.
- We classify the sample spaces w.r.t. the number of elements they contain.
 - *Finite sample space*: the set of possible outcomes of the experiment is finite;
 - *Countably infinite sample space*: the outcomes of the experiment are in a one-to-one relationship with \mathbb{N} ;
 - Otherwise the sample space is called *uncountable* or *nondenumerable*.
- A finite or countably finite sample space is called a *discrete* sample space. *Continuous* sample spaces, such as all the points on a line or interval are examples of uncountable sample spaces.

Basic definitions

Event

Definition

Given a random experiment and the corresponding sample space S , a collection of certain outcomes is called an event E . E is a subset of the sample space $E \subset S$. Equivalently, any statement of conditions identifying E is called an event.

Example

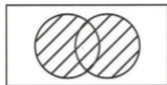
In the random experiment “Toss of a die”, we define the sample space $S = \{1, \dots, 6\}$. The event $E =$ “The outcome is an even number” is equivalent to $E = \{2, 4, 6\}$.

Definition

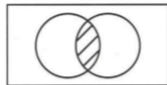
Given a random experiment with sample space S we call a single performance of the experiment a *trial*.

Basic definitions

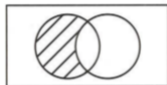
Basic Set Operations



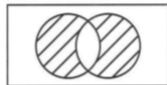
Union: $A \cup B$



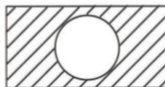
Intersection: $A \cap B$



Difference: $A \setminus B$



Symmetric Difference: $A \Delta B$



Complementation: A^c

Basic definitions

Sequential sample space

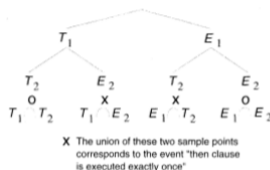


Figure 1.8. Tree diagram of a sequential sample space

The set of all leaves of the tree is the sample space of interest.

Basic definitions

Two-dimensional sample space

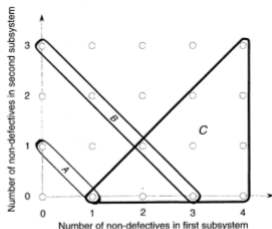


Figure 1.9. A two-dimensional sample space

- ① $A :=$ "the system has exactly one non-defective component"
- ② $B :=$ "the system has exactly three non-defective components"
- ③ $C :=$ "the first subsystem has more non-defective components than the second subsystem"

Basic definitions

Events

- Let E be an event of S and let denote the outcome of a specific trial by s . If $s \in E$ we then say that the event E has *occurred*. Only one outcome $s \in S$ can occur on any trial. However, every event including s will occur.
- Let E , E_1 and E_2 be events. We define the event \bar{E} to be *the complement* of E , i.e. $\bar{E} = \{s \in S \mid s \notin E\}$.
- Let E_1 and E_2 be two events. We define the event
 - $E_1 \cup E_2 := \{s \in S \mid s \in E_1 \text{ or } s \in E_2\}$ *the union* of E_1 and E_2 .
 - $E_1 \cap E_2 := \{s \in S \mid s \in E_1 \text{ and } s \in E_2\}$ *the intersection* of E_1 with E_2 .
 - $E_1 \setminus E_2 := \{s \in S \mid s \in E_1 \text{ and } s \notin E_2\}$ *the difference* of E_1 and E_2 .
- If $E_1 \cap E_2 = \emptyset$ we say that the two events E_1 and E_2 are *mutually exclusive* or *disjoint*.
- We denote by $|E|$ *the cardinality* of E , i.e. the number of elements (outcomes) in E .

Basic definitions

Algebra of events

Definition

Let S denotes the sample space of a given experiment and \mathcal{F} a collection of events. We say that \mathcal{F} is an *algebra* over S if the following two conditions are fulfilled:

- ① S must be an element of \mathcal{F} , i.e. $S \in \mathcal{F}$.
- ② If $E_1 \in \mathcal{F}$, $E_2 \in \mathcal{F}$, the sets $E_1 \cup E_2$, $E_1 \cap E_2$ must also belong to \mathcal{F} .
- ③ If $E_1 \in \mathcal{F}$, $E_2 \in \mathcal{F}$, $E_1 \setminus E_2$ must also belong to \mathcal{F} .

Interpretation: an algebra \mathcal{F} over S is a family of subsets of S which is closed with respect to the three binary operators \cup , \cap and \setminus .

Remark: condition 3 of the previous definition can be replaced by the following equivalent condition:

- 3^{bis} If $E \in \mathcal{F}$ then $\overline{E} \in \mathcal{F}$.

Example

- a) The collection given by $\mathcal{F} = \{S, \emptyset\}$ is an algebra (the so called *trivial* algebra over S).
- b) The collection $\mathcal{F} = \{E, \bar{E}, S, \emptyset\}$ is the algebra generated by $E \subset S$.
- c) The power set of S , denoted by $\mathcal{P}(S)$, is defined to be the collection consisting of all subsets of S (including the empty set \emptyset).

Exercise

- ① Show by mathematical induction that if S is a finite set with $|S| = n$ elements, then the power set of S contains $|\mathcal{P}(S)| = 2^n$ elements.
- ② Show that condition 3^{bis} is equivalent to condition 3.
- ③ Given $S = \{1, 2, 3, 4, 5, 6\}$ construct at least two different algebras on S .
- ④ De Morgan's law. Let A and B two events. Show that

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \text{ and } \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

Exercise

- ① *De Morgan's law. Let I be a non empty, possibly uncountable set and $A_i, i \in I$ a family of sets indexed by I . Show that*

$$\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \bar{A}_i \text{ and } \overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \bar{A}_i.$$

Definition

The indicator of the set $A \subseteq S$ is the function on S given by

$$1_A(s) = \begin{cases} 1 & \text{if } s \in A \\ 0 & \text{if } s \notin A. \end{cases}$$

Interpretation: the function 1_A “indicates” whether A occurs.

Exercise

Prove the following equalities:

$$1_{A \cup B} = \max\{1_A, 1_B\}; \quad 1_{A \cap B} = \min\{1_A, 1_B\} = 1_A 1_B$$
$$1_{\bar{A}} = 1 - 1_A; \quad 1_{A \Delta B} = |1_A - 1_B|$$

Basic definitions

σ –Algebra of events

Definition

Let S denotes the sample space of a given experiment and \mathcal{G} a collection of events. We say that \mathcal{G} is a σ –*algebra* over S if

- 1 \mathcal{G} is an *algebra* over S .
- 2 If $E_n \in \mathcal{G}$, $n = 1, 2, \dots$, then

$$\bigcup_{n=1}^{\infty} E_n \in \mathcal{G}.$$

Interpretation: a σ –algebra \mathcal{G} over S is a family of subsets of S which is closed with respect to

- the difference operator \setminus (or, equivalently, complementation),
- the *countable* union and intersection of its elements.

Basic definitions

Algebra and σ -Algebra generation

Theorem

Let \mathcal{E} be a collection of subsets of S . Then there are a smallest algebra $\alpha(\mathcal{E})$ and a smallest σ -algebra $\sigma(\mathcal{E})$ containing all the sets that are in \mathcal{E} .

Proof.

$\mathcal{P}(S)$ is a σ -algebra on S . Therefore it exists at least one σ -algebra and one algebra containing \mathcal{E} . We define $\alpha(\mathcal{E})$ (or $\sigma(\mathcal{E})$) to consist of all sets that belong to every algebra (or σ -algebra) containing \mathcal{E} . It is easy to verify that this system is an algebra (or σ -algebra) and indeed the smallest. \square

Basic definitions

Algebra and σ -Algebra generation

Examples

- ① Let \mathcal{C} be the family of open intervals on the real line, i.e.

$$\mathcal{C} := \{(a, b) \mid a, b \in \mathbb{R} \text{ and } a < b\}.$$

$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{C})$ is called the Borel sigma algebra over \mathbb{R} .

- ② Let $f : S \rightarrow \mathbb{R}$ be a real valued function. The family of preimages:

$$\{f^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$$

is a σ -algebra on S , the σ -algebra generated by f [denoted $\sigma(f)$].

Basic definitions

Measurable space

Definition

The pair (S, \mathcal{G}) where S is the sample space and \mathcal{G} a σ -algebra on S is called a measurable space and the elements of \mathcal{G} are called events.

Remark: a subset A of the sample space S is an event if and only if $A \in \mathcal{G}$.

Example

Suppose we toss a die once. We choose $S = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{G} := \{S, \emptyset, \{1, 2, 3\}, \{4, 5, 6\}\}$. The subset $A = \{2, 4, 6\}$ of S is *not* an event of \mathcal{G} . Let define the sigma algebra $\mathcal{H} := \sigma(\{\{1, 2, 3\}, A\})$ and the new measurable space (S, \mathcal{H}) . $A \in \mathcal{H}$ so that A is now an event.

Exercise

Complete the sigma algebra \mathcal{H} by enumerating its elements.

Basic definitions

Probability Law

Definition

Let (S, \mathcal{H}) be a measurable space. A probability measure or probability law is a positive real-valued function $P : \mathcal{H} \rightarrow [0, 1]$ such that the following axioms hold

- 1 $P(S) = 1$.
- 2 For every sequence $\{E_n\}_{n \in \mathbb{N}}$ of pairwise disjoint events ($E_i \cap E_j = \emptyset$, $\forall i \neq j$) it must hold

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Remark: a probability law is a function defined on a sigma algebra. The arguments of P are events, i.e. subsets of S and elements of \mathcal{H} .

A probability law has many useful relations, see Trivedi pp. 15-16.

Example

Probability of union of events version 1 (Trivedi, page 15).

If A_1, A_2, \dots, A_n are any events, then

$$\begin{aligned}P\left(\bigcup_{i=1}^n A_i\right) &= P(A_1 \cup A_2 \cup \dots \cup A_n) \\&= \sum_i P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) \\&\quad + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) + \dots \\&\quad + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n)\end{aligned}$$

where the successive sums are over all possible events, pairs of events, triples of events, and so on.

Example

Probability of union of events, version 2 (Trivedi, page 16).

If A_1, A_2, \dots, A_n are any events, then

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= P(A_1) + P(\bar{A}_1 \cap A_2) + P(\bar{A}_1 \cap \bar{A}_2 \cap A_3) + \dots \\ &\quad + P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_{n-1} \cap A_n). \end{aligned}$$

Exercise

Using the properties of P prove previous equality.

Definition

A probability space is a the triple (S, \mathcal{G}, P) where

- S is the sample space
- \mathcal{G} is a σ -algebra over S
- P is a probability measure

Remark: if the sample space is finite or countable it is possible to define a probability measure on $\mathcal{G} = \mathcal{P}(S)$. In that case every subset E of S is an event. However, if S is an uncountable sample space this is no longer true. The σ -algebra on which to define the probability law P must be smaller than $\mathcal{P}(S)$ in order to consistently define P (see Trivedi, page 17).

Exercise

- ① Let $S = \{s_1, s_2, \dots, s_n\}$ be a finite sample space and $\mathcal{G} = \mathcal{P}(S)$. Assume that each sample point s is equally likely. Show that the function P defined for all events $E \in \mathcal{G}$ by

$$P(E) = \frac{|E|}{n}$$

satisfies the axioms of a probability law.

Exercise

- ① (Trivedi, p. 19). Show that if event B is contained in event A , then $P(B) \leq P(A)$.

Exercise

- ① (Trivedi, p. 19). Consider a pool of six I/O buffers. Assume that any buffer is just as likely to be available (or occupied) as any other.

Compute the probabilities of the events

- ① $A =$ "at least 2 but no more than 5 buffers occupied".
- ② $B =$ "at least 3 but no more than 5 occupied".
- ③ $C =$ "all buffers available or an even number of buffers occupied".

Also determine the probability that at least one of the events A , B and C occurs.

Basic definitions

Finite probability space

If the sample space S of an experiment is finite, then the computation of probabilities is often simple. Assume (as it is almost always the case) that \mathcal{G} contains all elementary events, i.e. $\{s_i\} \in \mathcal{G} \forall i = 1, \dots, n$ and that $P(s_i) = p_i \geq 0$ and

$$\sum_{i=1}^n p_i = 1.$$

Because any event E consists of a certain collection of sample points, $P(E)$ can be computed by axiom 2 as the sum of the probabilities of the elementary events the union of which make up E .

Example

Let us assume that we toss a loaded die where the probabilities are given by $p_1 = 0.2$, $p_i = 0.16$ $i = 2, \dots, 6$. The probability of the event $E =$ "the outcome is ≤ 4 " is

$$P(E) = P(\{1\} \cup \{2\} \cup \{3\}) = 0.2 + 0.16 + 0.16 = 0.52$$

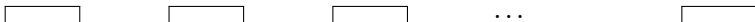
Combinatorial problems

Ordered samples of size k , with replacement

We are interested in the number of ways we can select k distinct objects from among n objects where

- the selection order is important,
- the same object can be selected any number of times.

We call this problem a permutations with replacement problem. Alternatively, we could express the same problem as the number of ordered sequences $s_{i_1}, s_{i_2}, \dots, s_{i_k}$, where each s_{i_r} belongs to s_1, \dots, s_n .



The solution is n^k . See example 1.4 at page 21.

Combinatorial problems

Ordered samples of size k , without replacement

We are interested in the number of ways we can select k distinct objects from among n objects where

- the selection order is important,
- object can be selectet at most one time.

We call this problem a permutations without replacement problem. The solution, called the number of permutations of n distinct objects taken k at a time, is denoted by $P(n, k)$ and is equal

$$P(n, k) = \frac{n!}{(n - k)!}$$



...



See example 1.5 at page 21.

Combinatorial problems

Unordered samples of size k , without replacement

We are interested in the number of ways we can select k distinct objects from among n objects where

- the selection order is *not* important,
- object can be selectet at most one time.

The solution is called the number of combinations of n distinct objects taken k at a time, is denoted by $C(n, k)$ or $\binom{n}{k}$ and is equal

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$



See example 1.6 at page 22.

Combinatorial problems

Exercises

Exercise

Solve Problems 1, 3 and 5 of Trivedi, page 23.

Conditional probability

Conditional probability of events

- Probability theory models the way we should measure uncertainty.
- When we run a random experiment we only know that the outcome will belong to the sample space S . According to this information (no information) we have a probability space and a law P which assigns probabilities to events: (S, \mathcal{G}, P) .
- What happens to these probabilities if we acquire new information about the outcome of the experiment?
- Suppose that we don't observe the outcome of the experiment but we know that $s \in B \in \mathcal{G}$, i.e. the event B is realised.
- Then, what is the probability of event A given that event B is realised (we simply say “given B ”), denoted $P(A \mid B)$?

Conditional probability

Conditional probability of events

- Assume $B \in \mathcal{G}$ with $P(B) \neq 0$. The conditional probability of B given B is ... $P(B | B) = 1$. There is no uncertainty regarding the realisation of event B once we know it is realised. We have *normalized* the probability of event B such that now it is equal to 1.
- Assume that S is finite and $\mathcal{G} = \mathcal{P}(S)$. If $P(B) \neq 0$ it makes then sense to define for all $s \in S$

$$P(\{s\} | B) = \begin{cases} \frac{P(\{s\})}{P(B)} & s \in B \\ 0 & s \notin B \end{cases}.$$

In fact, recall that $B = \bigcup_{s \in B} \{s\}$ and from axiom 2 of a (conditional) probability law

$$P(B | B) = P\left(\bigcup_{s \in B} \{s\} | B\right) = \sum_{s \in B} P(\{s\} | B) = \sum_{s \in B} \frac{P(\{s\})}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

For any $A \in \mathcal{G}$ recall that $A = (A \setminus B) \cup (A \cap B)$.

Conditional probability

Conditional probability of events

(continued)

$$\begin{aligned}P(A \mid B) &= \sum_{s \in A} P(\{s\} \mid B) \\&= \sum_{s \in A \setminus B} P(\{s\} \mid B) + \sum_{s \in A \cap B} P(\{s\} \mid B) \\&= 0 + \sum_{s \in A \cap B} \frac{P(\{s\})}{P(B)} \\&= P(A \cap B) / P(B)\end{aligned}$$

- All sample points not in B can be disregarded because the outcome of the experiment must be in B . For each event $A \in \mathcal{G}$ the relevant sample points are those common with B , i.e. the points in $A \cap B$.

Conditional probability

Conditional probability of events

Definition

The conditional probability of event A given event B is therefore defined by

$$P(A | B) = \begin{cases} \frac{P(A \cap B)}{P(B)} & \text{if } P(B) > 0 \\ \text{undefined} & \text{if } P(B) = 0 \end{cases}.$$

Multiplication rule (MR): rearranging the terms in the definition of $P(A | B)$ we obtain the following equalities

$$P(A \cap B) = \begin{cases} P(B)P(A | B) & \text{if } P(B) \neq 0 \\ P(A)P(B | A) & \text{if } P(A) \neq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Exercise

- ① Solve Problem 1 Trivedi, page 25.
- ② Let (S, \mathcal{G}, P) be a probability space and $P(A) \neq 0$. Show that the set function defined on \mathcal{G}

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

- ① is a probability law on (S, \mathcal{G}) ,
- ② satisfies $P_A(\bar{A}) = 0$.
- ③ Let P be the uniform distribution on a finite sample space S and let A be a event on S . Prove that $P(\cdot \mid A)$ is the uniform distribution on A .

Conditional probability

Independence

If the probability of the occurrence of an event A does not change regardless of whether event B has occurred, we are likely to conclude that the two events are independent. Formally we have the following definition.

Definition

(Independent events). We define two events A and B to be independent if and only if

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

Remarks:

- 1 Do not confuse independent events with mutually exclusive (disjoint) events.
- 2 If A and B are mutually exclusive events, then $A \cap B = \emptyset$, which implies $P(A \cap B) = 0$. Now, if they are independent as well, then $P(A) = 0$ or $P(B) = 0$.
- 3 If an event A is independent of itself, that is, if A and A are independent, then $P(A) = 0$ or $P(A) = 1$. Why? Try to prove it!

Conditional probability

Remarks (continued):

- 1 The relation of independence, denoted by \perp , is not transitive. If $A \perp B$ and $B \perp C$ it does not follow that $A \perp C$.
- 2 If events A and B are independent, then so are events \bar{A} and B , events A and \bar{B} and events \bar{A} and \bar{B} . Prove it!

The concept of independence can be extended to a list of n events:

Definition

Independence of a set of events. A list of n events A_1, A_2, \dots, A_n is mutually independent if and only if for each set of k ($2 \leq k \leq n$) distinct indices i_1, i_2, \dots, i_k which are elements of $\{1, 2, \dots, n\}$ we have

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}).$$

Remark: It is possible to have $P(A \cap B \cap C) = P(A)P(B)P(C)$ but $P(A \cap B) \neq P(A)P(B)$, $P(A \cap C) \neq P(A)P(C)$, $P(B \cap C) \neq P(B)P(C)$. Under these conditions A, B and C are not mutually independent. (Trivedi, Example 1.10).

System's reliability

Reliability of a component

Let us consider an electronic system with n independent components. Define the event

$A_i :=$ “The i – th component is functioning properly”.

Definition

The reliability of component i is defined as

$$R_i = P(A_i),$$

i.e. it is the probability that the component is functioning properly.

A series system is a system such that the entire system fails if any one of its components fails.

System's reliability

Parallel system

Definition

A parallel system is a system such that the entire system fails only if all its components fail.

Theorem

Product law of reliabilities for series systems. The reliability of a series system decreases “quickly” with an increase in complexity (number of components).

Proof.

Let us consider the event A = “The system functions properly”. The reliability of a series system of n components is then

$$R = P(A) = P(A_1 \cap A_2 \cdots \cap A_n) = \prod_{i=1}^n P(A_i)$$



System's reliability

Parallel redundancy

Example

Let a series system have $n = 5$ components and $P(A_i) = 0.970$ for all components.

The system reliability is then equal to

$$R = P(A_1)^5 = 0.97^5 = 0.859.$$

If we increase $n = 10$ the system reliability decreases to 0.738!

What if $n = 1'000'000$?

System's reliability

Parallel redundancy

In order to mitigate the problem one possible solution is to implement parallel redundancy.

Example

Consider a parallel system of n independent components. The system runs correctly if at least one of its components runs properly, i.e.

$$A = (A_1 \cup A_2 \cdots \cup A_n).$$

But this means that

$$P(\bar{A}) = P(\overline{(A_1 \cup A_2 \cdots \cup A_n)}) = P(\bar{A}_1 \cap \bar{A}_2 \cdots \cap \bar{A}_n) = \prod_{i=1}^n P(\bar{A}_i)$$

Example

(continued) Applying the identity $P(\overline{B}) = 1 - P(B)$ to both sides of the equality and solving w.r.t. $P(A)$, we obtain the final formula

$$P(A) = 1 - \prod_{i=1}^n (1 - P(A_i))$$

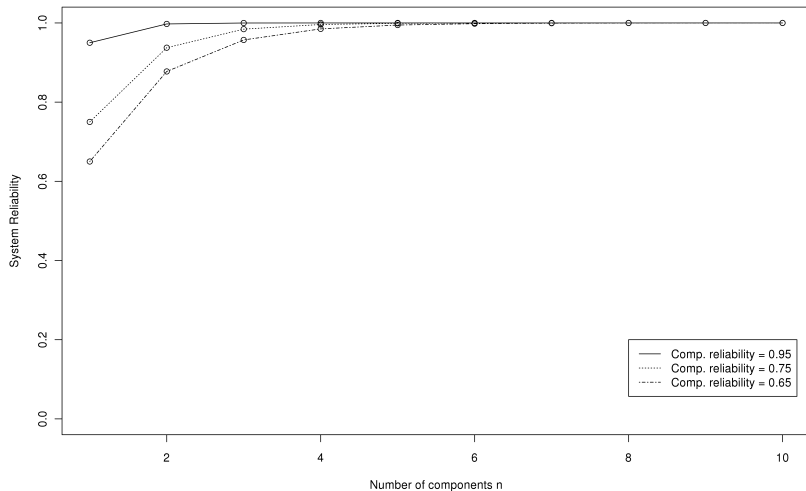
or equivalently, in terms of R and R_i

$$R = 1 - \prod_{i=1}^n (1 - R_i).$$

The next picture shows the so called Product Law of Unreliabilities.

System's reliability

Product Law of Unreliabilities



System's reliability

Law of Diminishing Returns

From the previous picture it is evident one characteristic of parallel redundancy: the marginal increase in reliability decreases with increasing number of parallel components. This behaviour is called the **Law of Diminishing Returns**.

Definition

A system with both series and parallel parts is called a *series-parallel system*.

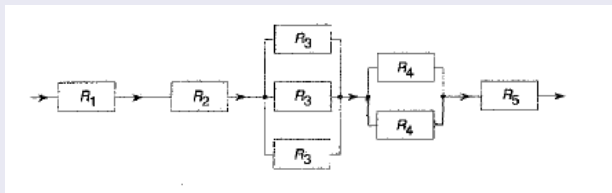


Figure: series-parallel reliability block diagram

System's reliability

Structure function

Definition

Let X be a state vector of a system with n components so that $X = (x_1, \dots, x_n)$ where

$$x_i = \begin{cases} 1 & \text{if component } i \text{ is functioning,} \\ 0 & \text{if component } i \text{ has failed.} \end{cases}$$

The structure function $\Phi(X)$ is defined by

$$\Phi(X) = \begin{cases} 1 & \text{if system is functioning,} \\ 0 & \text{if system has failed.} \end{cases}$$

The reliability of the system is then

$$R = P(\Phi(X) = 1).$$

System's reliability

Example communication network

Example

Communication network.

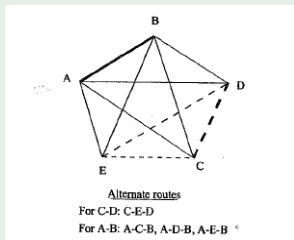


Figure: Communication network with five nodes

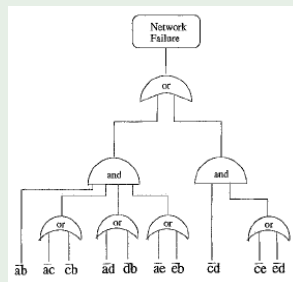


Figure: Fault tree for the communication network

Exercises: solve Problems 1,2,3,4 of Trivedi, page 35.

Definition

A list of n events B_1, B_2, \dots, B_n that are collectively exhaustive (i.e. $\bigcup B_i = S$) and mutually exclusive (i.e. $B_i \cap B_j = \emptyset$ for $i \neq j$) form an event space $S' = \{B_1, B_2, \dots, B_n\}$.

Let B an event and $S' = \{B, \bar{B}\}$ the corresponding event space. For any event A

$$A = (A \cap B) \cup (A \cap \bar{B}).$$

It follows (why?)

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \bar{B}) \\ &= P(A | B)P(B) + P(A | \bar{B})P(\bar{B}). \end{aligned}$$

The same formula can be generalized with respect to the event space $S' = \{B_1, B_2, \dots, B_n\}$:

Bayes' Rule

Theorem

Theorem of total probability. Let $S' = \{B_1, B_2, \dots, B_n\}$ be an event space and A an event. Then

$$P(A) = \sum_{i=1}^n P(A \mid B_i)P(B_i).$$

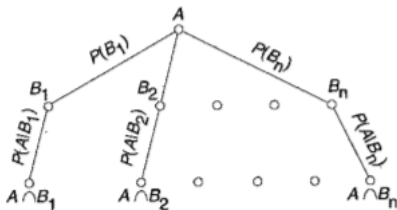


Figure 1.15. The theorem of total probability

Bayes' Rule

Bayes' Theorem

Let us assume that the event A is known to have occurred but it is not known which of the mutually exclusive and collectively exhaustive events B_1, \dots, B_n has occurred. In this situation we want to use the available information in an efficient way in order to evaluate $P(B_i | A)$.

Theorem

Bayes' Theorem. Let $S' = \{B_1, B_2, \dots, B_n\}$ be an event space and A and event. Then

$$\begin{aligned} P(B_i | A) &= \frac{P(B_i \cap A)}{P(A)} \\ &= \frac{P(A | B_i)P(B_i)}{\sum_{j=1}^n P(A | B_j)P(B_j)}. \end{aligned}$$

Example

Trivedi, page 39. Measurements on a certain day indicated that the source of incoming jobs at the North Carolina Super Computing Center (NCSC) is 15% from Duke, 35% from University of North Carolina (UNC) and 50% from North Carolina State (NCS).

- Probabilities that a job initiated from these universities is a multitasking job are 0.01, 0.05 and 0.02, respectively.
- Find the probability that a job chosen at random at NCSC is a multitasking job.
- Find the probability that a randomly chosen job comes from UNC, given that it is a multitasking job.

Let B_i = "job is from university i " and A = "job uses multitasking" ...

Exercises: solve Problems 1,2,3,4 of Trivedi, pp. 43-44.