

**ABDELKADER BENHARI**



**PROBABILITY, STATISTICS AND RANDOM PROCESSES**

*This course is an introduction to probability, statistics  
and random processes*

Ce livre a été publié sur [www.bookelis.com](http://www.bookelis.com)

ISBN : 979-10-227-1493-8

© Abdelkader BENHARI

Tous droits de reproduction, d'adaptation et de traduction,  
intégrale ou partielle réservés pour tous pays.

L'auteur est seul propriétaire des droits et responsable du contenu de ce livre.

## Contents

I. PROBABILITY .....	6
Basic Ideas of Probability .....	7
1. Probability Spaces .....	7
1.1. Discrete Probability Spaces.....	8
1.2. Continuous Probability Spaces .....	9
1.3. Properties of Probability .....	9
2. Conditional Probability and Statistical Independence .....	11
2.1. Conditional Probability.....	11
2.2. Composite Probability Formulae.....	11
2.3. Bayes Formulae.....	12
2.4. Statistical Independence .....	12
Appendix Combinatorics .....	13
Random Variables and Distributions .....	15
1. Random Variables .....	15
1.1. Discrete Random Variables.....	15
1.2. Continuous Random Variables .....	18
1.3. Distributions of Functions of Random Variables .....	21
2. Random Vectors (Multidimensional Random Variables) .....	23
2.1. Discrete Random Vectors .....	24
2.2. Continuous Random Vectors.....	24
2.3. Marginal Distributions/Probabilities/Densities .....	24
2.4. Conditional Distributions/Probabilities/Densities .....	25
2.5. Independence of Random Variables .....	26
2.6. Distributions of Functions of Random Vectors.....	26
Mathematical Expectations (Statistical Average) of Random Variables .....	31
1. Mathematical Expectations (Statistical Average) .....	31
1.1. Definitions.....	31
1.2. Properties.....	32
1.3. Moments .....	32
1.4. Holder Inequality.....	34
2. Correlation Coefficients and Linear Regression (Approximation) .....	35
3. Conditional Expectations and Regression Analysis .....	37
4. Generating and Characteristic Functions .....	38
5. Normal Random Vectors .....	40
Memo .....	42
Definition .....	42
Examples .....	42
Properties .....	42
Linear Regression.....	43
Regression .....	43
Normal Distribution .....	43
Limit Theorems .....	44

1. Inequalities .....	44
2. Convergences of Sequences of Random Variables .....	45
3. The Weak Laws of Large Numbers .....	46
4. The Strong Laws of Large Numbers .....	47
5. The Central Limit Theorems .....	49
Conditioning. Conditioned distribution and expectation. ....	51
1. The conditioned probability and expectation. ....	51
2. Properties of the conditioned expectation. ....	53
3. Regular conditioned distribution of a random variable. ....	59
Transition Probabilities .....	67
1. Definitions and notations. ....	67
2. The product between a probability and a transition probability. ....	68
3. Contractivity properties of a transition probability. ....	70
4. The product between transition probabilities. ....	73
5. Invariant measures. Convergence to a stable matrix .....	74
Disintegration of the probabilities on product spaces .....	75
1. Regular conditioned distributions. Standard Borel Spaces .....	75
2. The disintegration of a probability on a product of two spaces .....	78
3. The disintegration of a probability on a product of n spaces .....	79
The Normal Distribution .....	83
1. One-dimensional normal distribution .....	83
2. Multidimensional normal distribution .....	83
3. Properties of the normal distribution .....	86
4. Conditioning inside normal distribution .....	88
5. The multidimensional central limit theorem .....	91
<b>II. STATISTICS</b> .....	95
Basic Concepts .....	96
1. Populations, Samples and Statistics .....	97
2. Sample Distributions .....	99
2.1. $\chi^2$ (Chi-Square)-Distribution .....	99
2.2. t(Student)-Distribution .....	100
2.3. F-Distribution .....	100
3. Normal Populations .....	103
Parameter Estimation .....	104
1. Point Estimation .....	105
1.1. Point Estimators .....	105
1.2. Method of Moments (MOM) .....	105
1.3. Maximum Likelihood Estimation (MLE) .....	106
2. Interval Estimation .....	108
Tests of Hypotheses .....	111
1. Parameters from a Normal Population .....	112
2. Parameters from two Independent Normal Populations .....	115

III. RANDOM PROCESSES .....	118
Introduction .....	119
1. Definition .....	120
2. Family of Finite-Dimensional Distributions .....	121
3. Mathematical Expectations .....	122
4. Examples .....	123
4.1. Processes with Independent, Stationary or Orthogonal Increments .....	123
4.2. Normal Processes .....	124
Markov Processes (1).....	125
1. General Properties .....	126
2. Discrete-Time Markov Chains .....	128
2.1. Transition Probabilities .....	128
2.2. Classification of States .....	130
2.3. Stationary & Limit Distributions.....	135
2.4. Examples: Simple Random Walks .....	136
Appendix Eigenvalue Diagonalization .....	138
Markov Processes (2).....	140
1. Continuous-Time Markov Chains .....	141
1.1. Transition Rates.....	141
1.2. Kolmogorov Forward and Backward Equations .....	142
1.3. Fokker-Planck Equations .....	144
1.4. Ergodicity .....	145
1.5. Birth and Death Processes.....	146
1.6. Poisson Processes .....	147
Appendix Queuing Theory .....	153
2. Continuous-Time and Continuous-State Markov Processes .....	155
2.1. Basic Ideas.....	155
2.2. Wiener Processes.....	156
Hidden Markov Models .....	159
1. Definition of Hidden Markov Models .....	160
2. Assumptions in the theory of HMMs .....	161
3. Three basic problems of HMMs✓ .....	163
3.1. The Evaluation Problem .....	163
3.2. The Decoding Problem .....	163
3.3. The Learning Problem.....	163
4. The Forward/Backward Algorithm and its Application to the Evaluation Problem	165
5. Viterbi Algorithm and its Application to the Decoding Problem .....	167
6. Baum-Welch Algorithm and its Application to the Learning Problem .....	169
6.1. Maximum Likelihood (ML) Criterion.....	169
6.2. Baum-Welch Algorithm .....	169
Second-Order Processes and Random Analysis.....	172
1. Second-Order Random Variables and Hilbert Spaces .....	173

2. Second-Order Random Processes .....	174
2.1. Orthogonal Increment Random Processes.....	174
3. Random Analysis .....	176
3.1. Limits .....	176
3.2. Continuity .....	176
3.3. Derivatives .....	177
3.4. Integrals.....	178
Stationary Processes.....	179
1. Strictly Stationary Processes .....	180
2. Weakly Stationary Processes .....	181
2.1. Definition .....	181
2.2. Properties of Correlation/Covariance Functions .....	181
2.3. Periodicity .....	182
2.4. Random Analysis .....	182
2.5. Ergodicity (Statistical Average = Time Average).....	183
2.6. Spectrum Analysis & White Noise .....	184
3. Discrete Time Sequence Analysis: Auto-Regressive and Moving-Average (ARMA)	
Models .....	186
3.1. Definition .....	186
3.2. Transition Functions .....	186
3.3. Mathematical Expectations.....	188
3.4. Parameter Estimation .....	189
4. Problems .....	193
Martingales .....	196
1. Simple properties .....	197
2. Stopping times .....	199
3. An application: the ruin problem. ....	205
Convergence of martingales .....	207
1. Maximal inequalities .....	207
2. Almost sure convergence of semimartingales .....	210
3. Uniform integrability and the convergence of semimartingales in $L^1$ ....	214
4. Singular martingales. Exponential martingales. ....	218
<b>Bibliography:</b> .....	221

# **I. POBABILITY**

# Basic Ideas of Probability

## 1. Probability Spaces

There are two definitions of probabilities for random events: classical and modern. The modern definition of probability is based on the measure theory in which a random event is nothing but a set and its probability is the measure of the set.

**Definition (Sigma-Algebra)** Let  $\Omega$  be a set and  $\Pi$  a class  $\Pi$  of subsets of  $\Omega$ , i.e., a subset of  $2^\Omega$ ,  $\Pi$  is said to be a  $\sigma$ -algebra of  $\Omega$  if

- (1)  $\Omega \in \Pi$
- (2) if  $A \in \Pi$ , then  $\bar{A} = \Omega - A \in \Pi$  (which implies that  $\emptyset \in \Pi$ )
- (3) if  $A_i \in \Pi$ , where  $i \in I$  and  $I$  is at most a countable index set, then  $\bigcup_{i \in I} A_i \in \Pi$  (which means that the class  $\Pi$  is closed with respect to union)

**Remark 1:**  $2^\Omega$  is the power set of  $\Omega$ , i.e., the set of all subsets of  $\Omega$ .

**Remark 2:** In measure theory,  $(\Omega, \Pi)$  is called a *measurable space*.

**Remark 3:** Since  $\bigcap_{i \in I} A_i = \overline{\bigcup_{i \in I} \bar{A}_i} = \bigcup_{i \in I} \bar{\bar{A}_i} \in \Pi$ ,  $\Pi$  is also closed with respect to intersection.

**Example** Let  $\Omega = \{\omega_1, \omega_2\}$ ,  $\Pi = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}\}$ , where  $\emptyset$  stands for empty set,  $\Pi$  is then a  $\sigma$ -algebra.

**Definition (Probability Space)** Let  $\Omega$  be a set,  $\Pi$  a  $\sigma$ -algebra of  $\Omega$  and  $P$  a real-valued function defined on  $\Pi$ , the triplet  $(\Omega, \Pi, P)$  is called a *probability space* if  $P$  satisfies the following conditions

- (1)  $P(A) \geq 0$  for all  $A \in \Pi$



$$(2) \quad P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i) \quad \text{for all } A_1, A_2, \dots, A_n, \dots \in \Pi \text{ such that } A_i \cap A_j = \emptyset \text{ when } i \neq j$$

$$(3) \quad P(\Omega) = 1 \quad (\text{which implies that } P(\emptyset) = 0)$$

**Remark 1:** Usually,  $\Omega$  is often called *sample space*,  $\Pi$  the field of *random events* and for all  $A \in \Pi$ ,  $P(A)$  the *probability* of occurrence of  $A$ .

**Remark 2:** In measure theory, the probability space  $(\Omega, \Pi, P)$  is also called *measured space*.

**Remark 3:** Two random events  $A$  and  $B$  are said to be *incompatible* if  $AB = \emptyset$ . In this case,  $P(AB) = 0$ .

## 1.1. Discrete Probability Spaces

The number of all possible occurrences in a random experiment is countable.

**Definition** A probability space  $(\Omega, \Pi, P)$  is called a *discrete probability space* if the sample space  $\Omega$  is a countable (finite or denumerable infinite) set and  $\Pi = 2^\Omega$ .

**Remark 1:** To specify a discrete probability  $P$ , it suffices to specify a mapping  $p: \Omega \rightarrow [0, 1]$  such that  $p(\omega) \geq 0$  for all  $\omega \in \Omega$  and  $\sum_{\omega \in \Omega} p(\omega) = 1$ . Then, for all  $A \in \Pi$ ,  $P(A) = \sum_{\omega \in A} p(\omega)$ .

**Remark 2:** If  $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$  and  $p(\omega_i) = \frac{1}{N}$ , where  $i = 1, 2, \dots, N$ , then the resulting triple  $(\Omega, \Pi, P)$  is called *classical probability space*.

**Example** Let  $\Omega = \{\omega_1, \omega_2\}$ ,  $\Pi = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}\}$ , and

$$(1) \quad p(\omega_1) = \frac{1}{3}, \quad p(\omega_2) = \frac{2}{3}, \quad \text{then } (\Omega, \Pi, P) \text{ is a discrete probability space}$$

$$(2) \quad p(\omega_1) = p(\omega_2) = \frac{1}{2}, \quad \text{then } (\Omega, \Pi, P) \text{ is a classical probability space}$$

**Example** Let  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ ,  $\Pi = 2^\Omega$  and  $p(\omega_n) = \frac{\frac{1}{n^2}}{\sum_{k=1}^{+\infty} \frac{1}{k^2}} = \frac{6}{(n\pi)^2}$ ,  $n = 1, 2, \dots$ , then

$(\Omega, \Pi, P)$  is a discrete probability space.

## 1.2. Continuous Probability Spaces

The number of all possible occurrences in a random experiment is uncountable.

**Definition** A probability space  $(\Omega, \Pi, P)$  is called a *continuous probability space* if the sample space  $\Omega$  is a continuum.

**Example (Geometric Probability)** Assume that the sample  $\Omega$  is an interval, an area or a volume, then the probability of a point falling into a part of  $\Omega$  is given by

$$P = \frac{\text{Measure of the part of } \Omega}{\text{Measure of } \Omega}$$

## 1.3. Properties of Probability

**Theorem (Finite Measure)** Let  $(\Omega, \Pi, P)$  be a probability space, then for all  $A \in \Pi$ ,

$$P(A) + P(\bar{A}) = P(\Omega) = 1 \Rightarrow P(A) \leq 1$$

**Theorem (Monotonicity)** Let  $(\Omega, \Pi, P)$  be a probability space, then for all  $A, B \in \Pi$ ,

$$A \subseteq B \Rightarrow P(A) \leq P(A) + P(B - A) = P(B)$$

**Theorem (Union)** Let  $(\Omega, \Pi, P)$  be a probability space, then for all  $A, B \in \Pi$ ,

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

**Theorem (Union)** Let  $(\Omega, \Pi, P)$  be a probability space, then for all  $A_1, A_2, \dots, A_n \in \Pi$ ,

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k})$$

**Hint:**

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^n (A_i A_{n+1})\right) \\ &= \sum_{k=1}^n \left\{ (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k}) \right\} + P(A_{n+1}) - \sum_{k=1}^n \left\{ (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k} A_{n+1}) \right\} \\ &= \sum_{1 \leq i_1 \leq n} P(A_{i_1}) + P(A_{n+1}) \\ &\quad + \sum_{k=2}^n \left\{ (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k}) \right\} + \sum_{k=1}^{n-1} \left\{ (-1)^k \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k} A_{n+1}) \right\} \\ &\quad + (-1)^n \sum_{1 \leq i_1 < \dots < i_n \leq n} P(A_{i_1} \dots A_{i_n} A_{n+1}) \\ &= \sum_{1 \leq i \leq n+1} P(A_i) \\ &\quad + \sum_{k=2}^n (-1)^{k-1} \left\{ \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k}) + \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} P(A_{i_1} \dots A_{i_{k-1}} A_{n+1}) \right\} \\ &\quad + (-1)^n P(A_1 \dots A_n A_{n+1}) \\ &= \sum_{1 \leq i \leq n+1} P(A_i) + \sum_{k=2}^n \left\{ (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n+1} P(A_{i_1} \dots A_{i_k}) \right\} + (-1)^n P(A_1 \dots A_n A_{n+1}) \\ &= \sum_{k=1}^{n+1} \left\{ (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n+1} P(A_{i_1} \dots A_{i_k}) \right\} \end{aligned}$$

## 2. Conditional Probability and Statistical Independence

### 2.1. Conditional Probability

**Definition** Let  $(\Omega, \Pi, P)$  be a probability space and  $A, B \in \Pi$ , the *conditional probability* of B, given that A has occurred, is defined as  $P(B/A) = \frac{P(AB)}{P(A)}$ , where  $P(A) > 0$ .

**Theorem** Let  $(\Omega, \Pi, P)$  be a probability space and  $A \in \Pi$  with  $P(A) > 0$ , the triplet  $(\Omega_A, \Pi_A, P_A)$  is also a probability space, where  $\Omega_A = \Omega \cap A$ ,  $\Pi_A = \{AB | B \in \Pi\}$  and  $P_A(AB) = P(B/A)$ .

### 2.2. Composite Probability Formulae

**Theorem (Composite Probability Formula)** Let  $(\Omega, \Pi, P)$  be a probability space, and  $A \in \Pi$ , if  $A \subseteq \bigcup_k E_k$ , where  $E_k \in \Pi$  with  $P(E_k) > 0$  and  $E_i \cap E_j = \emptyset$  for all  $i \neq j$ , then

$$P(A) = \sum_k P(A/E_k)P(E_k).$$

**Proof:**

$$P(A) = P\left(A \left( \bigcup_k E_k \right)\right) = P\left(\bigcup_k (AE_k)\right) = \sum_k P(AE_k) = \sum_k P(A/E_k)P(E_k) \quad \#$$

**Remark:**

$$A \subseteq B \Rightarrow A \cap B = A, \quad A \cap \left(\bigcup_k E_k\right) = \bigcup_k (A \cap E_k)$$

## 2.3. Bayes Formulae

**Theorem (Bayes Formula)** Let  $(\Omega, \Pi, P)$  be a probability space and  $A \in \Pi$  with  $P(A) > 0$ , if  $A \subseteq \bigcup_k E_k$ , where  $E_k \in \Pi$  with  $P(E_k) > 0$  and  $E_i \cap E_j = \emptyset$  for all  $i \neq j$ , then

$$P(E_i/A) = \frac{P(E_i)P(A/E_i)}{\sum_k P(E_k)P(A/E_k)}.$$

**Proof:**

$$P(E_i/A) = \frac{P(AE_i)}{P(A)} = \frac{P(E_i)P(A/E_i)}{\sum_k P(E_k)P(A/E_k)} \quad \#$$

## 2.4. Statistical Independence

**Definition** Let  $(\Omega, \Pi, P)$  be a probability space and  $A, B \in \Pi$ ,  $A$  and  $B$  are said to be *statistically independent* if  $P(AB) = P(A)P(B)$ .

**Remark 1:** If  $A$  and  $B$  are independent, then  $P(A/B) = \frac{P(AB)}{P(B)} = P(A)$ .

**Remark 2:** Recall that two events  $A$  and  $B$  are said to be incompatible if  $AB = \emptyset$ . In this case,  $P(AB) = 0$ .

**Definition** Let  $(\Omega, \Pi, P)$  be a probability space and  $\Pi'$  a subset of  $\Pi$ ,  $\Pi'$  is said to be *statistically independent* if for all finite subsets  $\Pi''$  of  $\Pi'$ ,  $P\left(\bigcap_{A \in \Pi''} A\right) = \prod_{A \in \Pi''} P(A)$ .

**Remark:** The statistical independence of any two events of  $\Pi'$  can not guarantee the statistical independence of  $\Pi'$ . For example,  $\Pi' = \{A, B, C\}$ ,  $\Pi'$  is statistically independent if

$$P(AB) = P(A)P(B), P(AC) = P(A)P(C), P(BC) = P(B)P(C), P(ABC) = P(A)P(B)P(C)$$

are established at the same time.

## Appendix Combinatorics

**Sample Selection** Suppose there are  $m$  distinguishable elements, how many ways there are in which one can select  $r$  elements from these  $m$  distinguishable elements?

Order counts?	Repetitions are allowed? (With/Without replacement)	The number of ways to choose the samples	Remarks
Yes	Yes	$m^r$	Permutation
Yes	No	$\frac{m!}{(m-r)!}$	Permutation
No	Yes	$\frac{(m+r-1)!}{r!(m-1)!}$	Combination
No	No	$\frac{m!}{r!(m-r)!}$	Combination

**Balls into Cells** There are eight different ways in which  $n$  balls can be placed into  $k$  cells:

Distinguish the balls?	Distinguish the cells?	Can cells be empty?	The number of ways to place $n$ balls into $k$ cells
Yes	Yes	Yes	$k^n$
Yes	Yes	No	$k! \left\{ \begin{matrix} n \\ k \end{matrix} \right\}$
No	Yes	Yes	$\frac{(k+n-1)!}{n!(k-1)!}$
No	Yes	No	$\frac{(n-1)!}{(k-1)!(n-k)!}$

Yes	No	Yes	$\sum_{r=1}^k \left\{ \begin{matrix} n \\ r \end{matrix} \right\}$
Yes	No	No	$\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$
No	No	Yes	$\sum_{r=1}^k p_r(n)$
No	No	No	$p_k(n)$

where  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{r=1}^k (-1)^{k-r} \binom{k}{r} r^n$  is the Stirling cycle number and  $p_k(n)$  the number of partition of the number  $n$  into exactly  $k$  integer pieces.