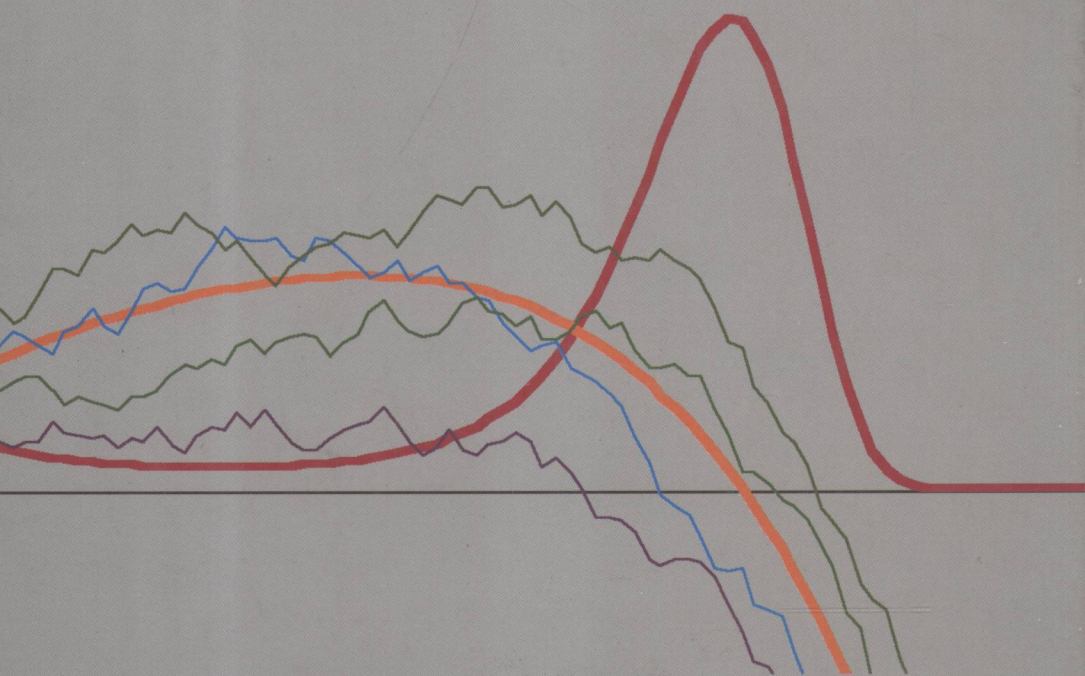


RECENT ADVANCES IN STOCHASTIC MODELING AND DATA ANALYSIS



Christos H Skiadas *editor*

0211.6-53
A836
2007

RECENT ADVANCES IN STOCHASTIC MODELING AND DATA ANALYSIS

Chania, Greece

29 May – 1 June 2007

editor

Christos H Skiadas

Technical University of Crete, Greece



E2009002659

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

RECENT ADVANCES IN STOCHASTIC MODELING AND DATA ANALYSIS

Copyright © 2007 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN-13 978-981-270-968-4

ISBN-10 981-270-968-1

**RECENT ADVANCES IN
STOCHASTIC MODELING
AND DATA ANALYSIS**

Preface

This volume contains a part of the invited and contributed papers which were accepted and presented at the 12nd International Conference on Applied Stochastic Models and Data Analysis in Chania, Crete, Greece, May 29- June 1, 2007. Since 1981, ASMDA aims to serve as the interface between Stochastic Modeling and Data Analysis and their real life applications particularly in Business, Finance and Insurance, Management, Production and Reliability, Biology and Medicine.

Our main objective is to include papers both theoretical and practical, presenting new results having potential for solving real-life problems. Another important objective is to present new methods for solving these problems by analyzing the relevant data. Also, the use of recent advances in different fields will be promoted such as for example, new optimization and statistical methods, data warehouse, data mining and knowledge systems and neural computing.

This volume contains papers on various important topics: Stochastic Processes and Models, Distributions, Insurance, Stochastic Modelling for Healthcare Management, Markov and Semi Markov models, Parametric/ Non-Parametric, Dynamical Systems / Forecasting, Modeling and Chaotic Modeling, Sampling and Optimization problems, Data Mining, Clustering and Classification, Applications of Data Analysis and various other applications. The World Scientific had also published the proceedings in two volumes of the 1993 Sixth ASMDA Conference, held also in Chania, Crete, Greece.

I acknowledge the valuable support of the Mediterranean Agronomic Institute, Chania, Greece, as well as the IBM France. Sincere thanks must be recorded to those whose contributions have been essential to create the Conference and the Proceedings. Finally, I would like to thank Anthi Katsirikou, Mary Karadima, John Dimotikallis and George Matalliotakis for their valuable support.

Chania, July 30, 2007
Christos H. Skiadas
Editor

Contents

Preface	v
1 Stochastic Processes and Models	1
An Approach to Stochastic Process using Quasi-Arithmetic Means <i>Etienne Cuvelier and Monique Noirhomme-Fraiture</i>	2
The Quantum Generator of Translations in a Fraction-Dimensional Manifold <i>Paulius Miškinis</i>	10
Cause of Stock Return Stochastic Volatility: Query by Way of Stochastic Calculus <i>Juho Kannianen</i>	18
On a Class of Backward Stochastic Differential Equations and Applications to the Stochastic Resonance <i>Romeo Negrea</i>	26
2 Distributions	34
An Application of the Extended Waring Distribution to Model Count Data Variables <i>José Rodríguez Avi, Antonio Conde Sánchez, Antonio José Sáez-Castillo and Ma José Olmo Jiménez</i>	35
Estimation of Simple Characteristics of Samples from Skewed and Heavy-Tailed Distributions <i>Zdeněk Fabián</i>	43
Estimating Robustness and Parameter Distribution in Compartmental Models of Neurons <i>Noam Peled and Alon Korngreen</i>	51
On the Stability of Queues with Negative Arrivals <i>Kernane Tewfik</i>	59
Random Multivariate Multimodal Distributions <i>George Kouvaras and George Kokolakis</i>	68
A System Maintained by Imperfect and Perfect Repairs under Phase-Type Distributions <i>Delia Montoro-Gazorla, Rafael Pérez-Ocón and M. Carmen Segovia</i>	76
Asymptotically Robust Algorithms for Detection and Recognition of Signals <i>Veniamin A. Bogdanovich and Aleksey G. Vostretsov</i>	82

Three Parameter Estimation of the Weibull Distribution by Order Statistics <i>Vaida Bartkutė and Leonidas Sakalauskas</i>	91
3 Insurance	101
Stochastic Models for Claims Reserving in Insurance Business <i>Tamás Faluközy, Ildikó Ibolya Vitéz and Miklós Arató</i>	102
Stochastic Risk Capital Model for Insurance Company <i>Gáida Pettere</i>	114
Measuring Demographic Uncertainty via Actuarial Indexes <i>Mariarosaria Coppola, Emilia Di Lorenzo, Albina Orlando and Marilena Sibillo</i>	122
Location as Risk Factor. Spatial Analysis of an Insurance Data-Set <i>Ildikó Vitéz</i>	130
A Hierarchical Bayesian Model to Predict Belatedly Reported Claims in Insurances <i>János Gyarmati-Szabó and László Márkus</i>	137
4 Stochastic Modeling for Healthcare Management	145
Non-Homogeneous Markov Models for Performance Monitoring of Healthcare <i>Sally McClean, Lalit Garg, Brian Meenan and Peter Millard</i>	146
Patient Activity in Hospital using Discrete Conditional Phase-Type (DC-Ph) Models <i>Adele H. Marshall, Louise Burns and Barry Shaw</i>	154
Identifying the Heterogeneity of Patients in an Accident and Emergency Department using a Bayesian Classification Model <i>Louise Burns and Adele H. Marshall</i>	162
Modelling the Total Time Spent in an Accident and Emergency Department and the Associated Costs <i>Barry Shaw and Adele H. Marshall</i>	172
5 Markov and Semi Markov Models	181
Periodicity of the Perturbed Non-Homogeneous Markov System <i>M. A. Symeonaki and P.-C. G. Vassiliou</i>	182
On the Moments of the State Sizes of the Discrete Time Homogeneous Markov System with a Finite State Capacity <i>G. Vasiliadis and G. Tsaklidis</i>	190
Copulas and Goodness of Fit Tests <i>Pál Rakonczai and András Zempléni</i>	198
Discrete Time Semi-Markov Models with Fuzzy State Space <i>Aleka A. Papadopoulou and George M. Tsaklidis</i>	206
An Application of the Theory of Semi-Markov Processes in Simulation <i>Sonia Malefaki and George Iliopoulos</i>	213

On a Numerical Approximation Method of Evaluating the Interval Transition Probabilities of Semi-Markov Models <i>Dimitrios Bitziadis, George Tsaklidis and Aleka Papadopoulou</i>	221
Markov Property of the Solution of the Generalized Stochastic Equations <i>Khalidi Khaled</i>	229
Partially Markov Models and Unsupervised Segmentation of Semi-Markov Chains Hidden with Long Dependence Noise <i>Jérôme Lapuyade-Lahorgue and Wojciech Pieczynski</i>	234
6 Parametric/Non-Parametric	242
Independent χ_1^2 Distributed in the Limit Components of Some Chi-Squared Tests <i>Vassilly Voinov, Mikhail Nikulin and Natalie Pya</i>	243
Parametric Conditional Mean and Variance Testing with Censored Data <i>Wenceslao González Manteiga, Cédric Heuchenne and César Sánchez Sellero</i>	251
Synthetic Data Based Nonparametric Testing of Parametric Mean-Regression Models with Censored Data <i>Olivier Lopez and Valentin Patilea</i>	259
7 Dynamical Systems/Forecasting	267
Application of the Single Index Model for Forecasting of the Inland Conveyances <i>Eugene Kopytov and Diana Santalova</i>	268
Development and Application of Mathematical Models for Internet Access Technology Substitution <i>Apostolos N. Giovanis and Christos H. Skiadas</i>	277
Exploring and Simulating Chaotic Advection: A Difference Equations Approach <i>Christos H. Skiadas</i>	287
8 Modeling and Stochastic Modeling	295
Likelihood Ratio Tests and Applications in 2D Lognormal Diffusions <i>Ramón Gutiérrez, Concepción Roldán, Ramón Gutiérrez-Sánchez and José Miguel Angulo</i>	296
Cartographical Modeling as a Statistical Method for Monitoring of a Spatial Behaviour of Population <i>Irina Pribytkova</i>	304
Learning and Inference in Switching Conditionally Heteroscedastic Factor Models using Variational Methods <i>Mohamed Saidane and Christian Lavergne</i>	312

Correlation Tests Based NARMAX Data Smoother Validation <i>Li Feng Zhang, Quan Min Zhu and Ashley Longden</i>	322
Kernel Based Confidence Intervals for Survival Function Estimation <i>Dimitrios I. Bagkavos, Aglaia Kalamatianou and Dimitrios Ioannides</i>	330
Chaotic Data Analysis and Hybrid Modeling for Biomedical Applications <i>Wlodzimierz Klonowski, Robert Stepien, Marek Darowski and Maciej Kozarski</i>	338
Stochastic Fractal Interpolation Function and its Applications to Fractal Analysis of Normal and Pathological Body Temperature Graphs by Children <i>Anna Soós</i>	342
A Modeling Approach to Life Table Data Sets <i>Christos H. Skiadas and Charilaos Skiadas</i>	350
An Extended Quadratic Health State Function and the Related Density Function for Life Table Data <i>Charilaos Skiadas, George Matalliotakis and Christos H. Skiadas</i>	360
9 Statistical Applications in Socioeconomic Problems	370
Dumping Influence on a Non Iterative Dynamics <i>Cécile Hardouin</i>	371
Firm Turnover and Labor Productivity Growth in the Italian Mechanical Sector <i>Luigi Grossi and Giorgio Gozzi</i>	382
Continuous Sampling Plan under an Acceptance Cost of Linear Form <i>Nicolas Farmakis and Mavroudis Eleftheriou</i>	390
A Dynamic Programming Model of a Machine Tool in Flexible Manufacturing <i>Bernard F. Lamond</i>	398
Particle Filter-Based Real-Time Estimation and Prediction of Traffic Conditions <i>Jacques Sau, Nour-Eddin El Faouzi, Anis Ben Aissa and Olivier de Mouzon</i>	406
Probability of Trend Prediction of Exchange Rate by ANFIS <i>George S. Atsalakis, Christos H. Skiadas and Ilias Braimis</i>	414
The Organizational Structure of Greek Libraries: The State of the Art and the Perspective of Team Working <i>Anthi Katsirikou</i>	423
10 Sampling and Optimization Problems	433
Applicability of Importance Sampling to Coupled Molecular Reactions <i>Werner Sandmann</i>	434

Bispectrum Estimation for a Continuous-Time Stationary Process from a Random Sampling <i>Karim Benhenni and Mustapha Rachdi</i>	442
Search via Probability Algorithm for Engineering Optimization Problems <i>Nguyen Huu Thong and Tran Van Hao</i>	454
Solving the Capacitated Single Allocation Hub Location Problem using Genetic Algorithm <i>Zorica Stanimirović</i>	464
11 Data Mining and Applications	472
Robust Refinement of Initial Prototypes for Partitioning-Based Clustering Algorithms <i>Sami Äyrämö, Tommi Kärkkäinen and Kirsi Majava</i>	473
Effects of Malingering in Self-Report Measures: A Scenario Analysis Approach <i>Massimiliano Pastore, Luigi Lombardi and Francesca Mereu</i>	483
The Effect of Agreement Expected by Chance on Some 2×2 Agreement Indices <i>Teresa Rivas-Moya</i>	491
Qualitative Indicators of Libraries' Services and Management of Resources: Methodologies of Analysis and Strategic Planning <i>Aristeidis Meletiou</i>	499
12 Clustering and Classification	511
Languages Similarity: Measuring and Testing <i>Liviu P. Dinu and Denis Enăchescu</i>	512
On Clustering Romance Languages <i>Liviu P. Dinu and Denis Enăchescu</i>	521
A Clustering Method Associated Pretopological Concepts and k-Means Algorithm <i>T. V. Le, N. Kabachi and M. Lamure</i>	529
Alternatives to the Estimation of the Functional Multinomial Regression Model <i>Manuel Escabias, Ana M. Aguilera and Mariano J. Valderrama</i>	537
A GARCH-Based Method for Clustering of Financial Time Series: International Stock Markets Evidence <i>Jorge Caiado and Nuno Crato</i>	542
13 Applications of Data Analysis	552
Reliability Problems and Longevity Analysis <i>Anatoli Michalski</i>	553

Statistical Analysis on Mobile Applications among City People: A Case of Bangkok, Thailand <i>Pakavadi Sirirangsi</i>	562
Pollution Sources Detection via Principal Component Analysis and Rotation <i>Marie Chavent, Hervé Guégan, Vanessa Kuentz, Brigitte Patouille and Jérôme Saracco</i>	571
Option Pricing and Generalized Statistics: Density Matrix Approach <i>Petr Jizba</i>	578
14 Miscellaneous	588
Inference for Alternating Time Series <i>Ursula U. Müller, Anton Schick and Wolfgang Wefelmeyer</i>	589
Estimation of the Moving-Average Operator in a Hilbert Space <i>Céline Turbillon, Jean-Marie Marion and Besnik Pumo</i>	597
Monte Carlo Observer for a Stochastic Model of Bioreactors <i>Marc Joannides, Irène Larramendy-Valverde and Vivien Rossi</i>	605
Monte Carlo Studies of Optimal Stopping Domains for American Knock Out Options <i>Robin Lundgren</i>	613
SONAR Image Denoising using a Bayesian Approach in the Wavelet Domain <i>Sorin Moga and Alexandru Isar</i>	621
Performance Evaluation of a Tandem Queueing Network <i>Smail Adjabi and Karima Lagha</i>	630
Assessment of Groundwater Quality Monitoring Network Based on Information Theory <i>Malgorzata Kucharek and Wiktor Treichel</i>	636
Improving Type II Error Rates of Multiple Testing Procedures by Use of Auxiliary Variables. Application to Microarray Data <i>Maela Kloareg and David Causeur</i>	645
Author Index	653

CHAPTER 1

Stochastic Processes and Models

An approach to Stochastic Process using Quasi-Arithmetic Means

Etienne Cuvelier and Monique Noirhomme-Fraiture

Institut d'Informatique (FUNDP)
21, rue Grandgagnage, 5000 Namur, Belgium
(e-mail: ecu@info.fundp.ac.be, mno@info.fundp.ac.be)

Abstract. Probability distributions are central tools for probabilistic modeling in data mining. In functional data analysis (FDA) they are weakly studied in the general case. In this paper we discuss a probability distribution law for functional data considered as stochastic process. We define first a new kind of stationarity linked to the Archimedean copulas, and then we build a probability distribution using jointly the Quasi-arithmetic means and the generators of Archimedean copulas. We also study some properties of this new mathematical tool.

Keywords: Functional Data Analysis, Probability distributions, Stochastic Process, Quasi-Arithmetic Mean, Archimedean copulas.

1 Introduction

Probability distributions are central tools for probabilistic modeling in data mining. In functional data analysis, as functional random variable can be considered as stochastic process, the probability distribution have been studied largely, but with rather strong hypotheses, [Cox and Miller, 1965], [Gihman and Skorohod, 1974], [Bartlett, 1978] and [Stirzaker, 2005]. Some processes are very famous like Markov process [Meyn and L, 1993]. Such a process has the property that present is not influenced by all the past but only by the last visited state. A very particular case is the random walk, which has the property that one-step transitions are permitted only to the nearest neighboring states. Such local changes of state may be regarded as the analogue for discrete states of the phenomenon of continuous changes for continuous states. The limiting process is called the Wiener process or Brownian motion. The Wiener process is a diffusion process having the special property of independent increments. Some more general Markov chain with only local changes of state are permissible, gives also Markov limiting process for continuous time and continuous states. The density probability is solution of a special case of the Fokker-Planck diffusion equation.

In preceding work [Cuvelier and Noirhomme-Fraiture, 2005] we used copulas to model the distribution of functional random variables at discrete cutting points. Here, using the separability concept, we can consider the continuous case as the limit of the discrete one. We will use quasi-arithmetic means in order to avoid copulas problem when considering the limit when the number

of cuttings tends to infinity.

In section 2 we define the concept of distribution of functions and recall the notion of separability. In section 3 we propose to use the Quasi-arithmetic mean in conjunction with an Archimedean generator to build a probability distributions appropriate to the dimensional infinite nature of the functional data. And in section 4 we study the properties of this new mathematical tool.

2 Distribution of a functional random variable

Let us recall some definitions that will be useful in the following paper.

Definition 1. Let (Ω, \mathcal{A}, P) a probability space and \mathcal{D} a closed real interval. A *functional random variable (frv)* is any function from $\mathcal{D} \times \Omega \rightarrow \mathbb{R}$ such for any $t \in \mathcal{D}$, $X(t, \cdot)$ is a real random variable on (Ω, \mathcal{A}, P) . Each function $X(\cdot, \omega)$ is called a realization. In the following we will write \underline{X} for $X(\cdot, \omega)$, and \underline{X}_t for $X(t, \cdot)$. \underline{X}_t can be considered as a stochastic process.

We study, here, the measurable and bounded functions.

Definition 2. Let \mathcal{D} a closed real interval, then $\mathcal{L}_2(D)$ is the space of real measurable functions $u(t)$ defined on a real interval \mathcal{D} such that

$$\|u\|_2 = \left\{ \int_{\mathcal{D}} |u(t)|^2 dt \right\}^{1/2} < \infty \quad (1)$$

Definition 3. Let $f, g \in \mathcal{L}_2(D)$. The pointwise order between f and g on \mathcal{D} is defined as follows :

$$\forall t \in \mathcal{D}, f(t) \leq g(t) \iff f \leq_{\mathcal{D}} g \quad (2)$$

Definition 4. The *functional cumulative distribution function (fcdF)* of a frv \underline{X} on $\mathcal{L}_2(D)$ computed at $u \in \mathcal{L}_2(D)$ is given by :

$$F_{\underline{X}, \mathcal{D}}(u) = P[\underline{X} \leq_{\mathcal{D}} u] \quad (3)$$

Definition 5. A frv is called *separable* if there exists in \mathcal{D} an everywhere countable set I of points $\{t_i\}$ and a set N of Ω of probability 0 such that for an arbitrary open set $G \subset \mathcal{D}$ and an arbitrary closed set $F \subset \mathbb{R}$ the two sets

$$\{\omega : X(t, \omega) \in F, \forall t \in G\}$$

$$\{\omega : X(t, \omega) \in F, \forall t \in G \cap I\}$$

differ from each other only on the subset N . The set I is called the separability set [Gihman and Skorohod, 1974].

The space $\mathcal{L}_2(D)$ is a separable Hilbert space. In the following we suppose that any realization of \underline{X} is in $\mathcal{L}_2(D)$.

Definition 6. Two frv $X_1(t, \omega)$ and $X_2(t, \omega)$ ($t \in \mathcal{D}, \omega \in \Omega$) are called stochastically equivalent if for any $t \in \mathcal{D}$

$$P \{X_1(t, \omega) \neq X_2(t, \omega)\} = 0 \quad (4)$$

The interest of separability comes from the following theorem .

Theorem 1 (J.L. Doob). Let \mathcal{X} and \mathcal{Y} be metric spaces, \mathcal{X} be separable, \mathcal{Y} be compact. An arbitrary random function $X(t, \omega)$, $t \in \mathcal{X}$ with values in \mathcal{Y} is stochastically equivalent to a certain separable random function.

3 The QAMM and QAMML distributions

In this section we build a sequence of sets that converge toward a separability set of \mathcal{D} and at each step we define a probability distribution. Let $n \in \mathbb{N}$, and $\{t_1^n, \dots, t_n^n\}$, n equidistant points of \mathcal{D} such that $t_1^n = \inf(\mathcal{D})$ and $t_n^n = \sup(\mathcal{D})$, and $\forall i \in \{1, \dots, n-1\}$ we have $|t_{i+1}^n - t_i^n| = \frac{|\mathcal{D}|}{n} = \Delta_t$. Let the two following sets

$$\mathcal{A}_n(u) = \bigcap_{i=1}^n \{\omega \in \Omega : X(t_i^n, \omega) \leq u(t_i^n)\}$$

$$\mathcal{A}(u) = \{\omega \in \Omega : \underline{X} \leq_{\mathcal{D}} u\}$$

We will use the following distribution to approximate the *fcd*f (3):

$$P[\mathcal{A}_n(u)] = H(u(t_1^n), \dots, u(t_n^n)) \quad (5)$$

where $H(\cdot, \dots, \cdot)$ is a joint distribution of dimension n . In previous works (see [Diday, 2002], [Vrac *et al.*, 2001], [Cuvelier and Noirhomme-Fraiture, 2005]) the Archimedean copulas were used for the approximation with small value of n . Let us recall the definition and property of copulas.

Definition 7. A copula is a multivariate cumulative distribution function defined on the n -dimensional unit cube $[0, 1]^n$ such that every marginal distribution is uniform on the interval $[0, 1]$:

$$C : [0, 1]^n \rightarrow [0, 1] : (u_1, \dots, u_n) \mapsto C(u_1, \dots, u_n)$$

The power of copulas comes from the following theorem (see [Nelsen, 1999]).

Theorem 2 (Sklar's theorem). Let H be an n -dimensional distribution function with margins F_1, \dots, F_n . Then there exists an n -copula C such that for all $x \in \mathbb{R}^n$,

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (6)$$

If F_1, \dots, F_n are all continuous, then C is unique; otherwise, C is uniquely determined on Range of $F_1 \times \dots \times$ Range of F_n .

Before using copulas, we define a function that gives the distribution of the values of \underline{X}_t for a chosen $t \in \mathcal{D}$.

Definition 8. Let \underline{X} a *frv*. We define the *surface of distributions* as follow :

$$G(t, y) = P[\underline{X}_t \leq y] \quad (7)$$

We can use various methods for determining suitable G for a chosen value of t . Thus for example, if \underline{X} is a Gaussian process with mean value $\mu(t)$ and standard deviation $\sigma(t)$, then we can use the *cdf* from $\mathcal{N}(\mu(t), \sigma(t))$. In other cases we can use the empirical cumulative distribution function to estimate \hat{G} :

$$\hat{G}(t, y) = \frac{\#\{X_i(t) \leq y\}}{N} \quad (8)$$

In the following we will always use this function G in conjunction with a function u of $\mathcal{L}_2(\mathcal{D}) : G[t, u(t)]$. So, for ease the notations, we will write $G[t; u] = G[t, u(t)]$. If we use the preceding expression in conjunction with (6), then (5) become :

$$P[\mathcal{A}_n(u)] = C(G[t_1^n; u], \dots, G[t_q^n; u]) \quad (9)$$

An important class of stochastic process is the class of stationary processes. A stochastic process is said to be *strictly stationary* [Burrill, 1972] if its distributions do not change with time; i.e. if for any $t_1, \dots, t_n \in \mathcal{D}$ and for any $h \in \mathcal{D}$, the multivariate distribution function of $(\underline{X}_{t_1+h}, \dots, \underline{X}_{t_n+h})$ does not depend on h . We propose here a more wide stationary property.

Definition 9. A stochastic process is said *copula stationary* if $\forall t_1, \dots, t_n \in \mathcal{D}$ and for any $h \in \mathcal{D}$, the copula of $(\underline{X}_{t_1+h}, \dots, \underline{X}_{t_n+h})$ does not depend on h , i.e. its copula does not change with time.

Let us notice that, if we deal with true functional data, realizations of a stochastic process \underline{X} , we can suppose that there is always the same functional relation between \underline{X}_s and \underline{X}_t for any value $s, t \in \mathcal{D}$. If a *frv* is also a *copula stationary* stochastic process, then we call it a *copula stationary frv*. There is an important class of copulas which is well appropriate for *copula stationary* stochastic processes : the class of Archimedean copulas.

Definition 10. An Archimedean copula is a function from $[0, 1]^n$ to $[0, 1]$ given by

$$C(u_1, \dots, u_n) = \psi \left[\sum_{i=1}^n \phi(u_i) \right] \quad (10)$$

where ϕ , called the generator, is a function from $[0, 1]$ to $[0, \infty]$ such that:

- ϕ is a continuous strictly decreasing function,
- $\phi(0) = \infty$ and $\phi(1) = 0$,

Table 1. Families of completely monotonic generators

Name	Generator	Dom. of θ
Clayton	$t^\theta - 1$	$\theta > 0$
Frank	$-\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$\theta > 0$
Gumbel-Hougaard	$(-\ln t)^\theta$	$\theta \geq 1$

- $\psi = \phi^{-1}$ is completely monotonic on $[0, \infty[$ i.e. $(-1)^k \frac{d^k}{dt^k} \psi(t) \geq 0$ for all t in $[0, \infty[$ and for all k .

Notice that the k -dimension margins of (10) are all the same, and this for any value of $1 \leq k \leq n$. If \underline{X} is a *copula stationary frv* then expression (9) can be written :

$$P[\mathcal{A}_n(u)] = \psi \left(\sum_{i=1}^n \phi(G[t_i^n; u]) \right) \tag{11}$$

Table 1[Nelsen, 1999] shows three important Archimedean generators for copulas. The distribution (11) with the Clayton generator was already used for clustering of functional data coming from the symbolic data analysis framework (see [Vrac *et al.*, 2001] and [Cuvelier and Noirhomme-Fraiture, 2005]). Unfortunately the above limit is almost always null for Archimedean copulas when $n \rightarrow \infty$ (see [Cuvelier and Noirhomme-Fraiture, 2007])!

Proposition 1. *If for $u \in \mathcal{L}_2(D) : G(t; u) < 1, \forall t \in \mathcal{D}$, then*

$$\lim_{q \rightarrow \infty} \psi \left[\sum_{i=1}^q \phi(G[t_i^n; u]) \right] = 0 \tag{12}$$

Another objection to the use of this type of joint distribution is something which we could call *volumetric behavior*.

Definition 11. A function $u \in \mathcal{L}_2(D)$ is called a *functional quantile* of value p , written Q_p , if

$$G(t; Q_p) = p, \forall t \in \mathcal{D} \tag{13}$$

The functional quantile Q_p can be seen as the level curve of value p . Now let us remark that for a functional quantile :

$$P[\mathcal{A}_n(Q_p)] = \psi \left[\sum_{i=1}^n \phi(G[t_i^n; Q_p]) \right] = \psi(n \cdot \phi(p)) < p \tag{14}$$

And it is easy to see that, if $n < m$ then $\psi[n\phi(p)] < \psi[m\phi(p)]$, and thus, the more we try to have a better approximation for a *functional quantile* of value p , the more we move away from reference value p toward zero. A simple way to avoid these two problems is to use the notion of quasi-arithmetic mean, concept which was studied by [Kolmogorov, 1930], [Nagumo, 1930] and [Aczel, 1966].