

PROGETTO DI UNITÀ DI RICERCA - MODELLO B
Anno 2007 - prot. 20077P5AWA_003

1 - Area Scientifico-disciplinare

13: Scienze economiche e statistiche 100%

2 - Durata del Progetto di Ricerca

24 Mesi

3 - Coordinatore Scientifico

BATTAGLIA FRANCESCO

Professore Ordinario

Università degli Studi di ROMA "La Sapienza"

Facoltà di SCIENZE STATISTICHE

Dipartimento di STATISTICA, PROBABILITÀ E STATISTICHE APPLICATE

4 - Responsabile dell'Unità di Ricerca

POLI IRENE

Professore Ordinario

04/11/1947

PLORNI47S44G388V

Università "Ca' Foscari" di VENEZIA

Dipartimento di STATISTICA

041/2347418
(Prefisso e telefono)

041/2347444
(Numero fax)

irenpoli@unive.it

5 - Curriculum scientifico

Testo italiano

Irene Poli è professore di Statistica presso l'Univerista' Ca' Foscari di Venezia (www.unive.it). Dopo la laurea conseguita presso l'Università di Pavia ha continuato i suoi studi presso l'Imperial College of Science and Technology di Londra (UK) e successivamente presso il Centre for Non Linear Studies del Laboratorio di Los Alamos (N.M., USA). Ha insegnato come professore nelle univerista' di Milano, Bologna, Modena e Venezia. I suoi più importanti interessi di ricerca sono stati nell'ambito dell'inferenza statistica Bayesiana, dei modelli lineari e non lineari per serie storiche, delle reti neurali predittive, del calcolo evolutivo e degli algoritmi genetici, e più recentemente del disegno sperimentale evolutivo. È membro della Società Italiana di Statistica, Fellow della New York Academy of Science e della Bernoulli Society (Institute of Mathematical Statistics), inoltre è membro eletto della Royal Statistical Society. È stata Direttore di Dipartimento, Presidente di Corso di Laurea e membro del Senato Accademico dell'Univerista' Ca' Foscari di Venezia. Attualmente è membro del Comitato Scientifico del Programma di Dottorato della Scuola di Studi Avanzati di Venezia. È anche componente del Comitato Scientifico di Civen, un Consorzio di Università dedicato alla ricerca nel campo delle nanotecnologie (www.civen.org), dove collabora all'International Master in Nanotechnologies e allo sviluppo di un progetto concernente i DNA microarray presso il laboratorio NanoFab del Parco Scientifico e Tecnologico Vega di Venezia. È stata coordinatrice di numerosi progetti di ricerca e attualmente è coordinatrice di un gruppo internazionale di ricerca presso l'univerista' Ca' Foscari di Venezia dedicato al progetto europeo Programmable Artificial Cell Evolution (PACE) nell'area IST-Future and Emergent Technologies (www.istpace.org). È CoDirettore del European Centre for Living Technology (ECLT, www.ecltech.org), un Centro di ricerca dedicato alle bionanotecnologie, finanziato dalla Commissione Europea (FP6).

Testo inglese

Irene Poli is professor of Statistics at the University Ca' Foscari of Venice (www.unive.it). After graduating from the University of Pavia, she pursued postgraduate research at Imperial College of Science and Technology of London (U.K.), and later at the Center for Non-linear Science, Los Alamos National Laboratory (N.M., U.S.A.). She served as professor at the Universities of Milano, Bologna, Modena and Venezia. Her major research interests have been in Bayesian nonparametric inference, non linear time series models, in predictive neural networks, in evolutionary computation and genetic algorithms, and more recently in evolutionary design of experiments. She is member of the Italian Statistical Society, Fellow of the New York Academy of Science and of the Bernoulli Society (Institute of Mathematical Statistics), and an elected member of the Royal Statistical Society. She has served as department chair and statistical university programs coordinator, and has been member of the Academic Senate of Ca' Foscari University. Currently she is member of the Scientific Board of the Doctoral Program of the School of Advanced Studies in Venice. She is also a component of the Scientific Committee of CIVEN, a university network devoted to the research in the fields of nanotechnologies (www.civen.org), where she collaborates in the International Masters Program in Nanotechnologies, and in a large research project on DNA-microarray, at the

NanoFabrication Facility Laboratory (NFF) associated with the Scientific and Technological Park Vega of Venice. She has been a principal investigator for several scientific research projects, and currently she heads an international research group at Ca' Foscari university, devoted to the EU integrated project Programmable Artificial Cell Evolution (PACE) in the area IST - Future and Emergent Technologies (www.protocell.org). She is CoDirector of the European Center for Living Technology (ECLT, www.ecltech.org), a research Centre dedicated to the Bionanotechnologies, supported by the European Commission (FP6).

6 - Pubblicazioni scientifiche più significative del Responsabile dell'Unità di Ricerca

1. M. FORLIN, POLI I., D. DE MARCH, N. PACKARD, G. GAZZOLA, R. SERRA. (2007). Evolutionary Experiments for self-assembling amphiphilic systems.. CHEMOMETRICS AND INTELLIGENT LABORATORY SYSTEMS. ISSN: 0169-7439.
2. M. THEIS, G. GAZZOLA, M. FORLIN, POLI I., M.M. HANCZYC, N.H. PACKARD, M. BEDAU. (2006). Optimal Formulation of Complex Chemical Systems with a Genetic Algorithm. COMPLEXUS. ISSN: 1424-8492.
3. ROBERTO SERRA, TIMOTEO CARLETTI, POLI I. (2006). Synchronization phenomena in surface-reaction models of protocells. ARTIFICIAL LIFE. ISSN: 1064-5462.
4. Brusco S., Minerva T., POLI I., Solinas G. (2002). Un automa cellulare per lo studio del distretto industriale. POLITICA ECONOMICA. vol. 2, pp. 147-192 ISSN: 1120-9496.
5. Biganzoli E., Boracchi P., POLI I. (2000). Reti neurali artificiali per lo studio di fenomeni complessi: limiti e vantaggi delle applicazioni in biostatistica. STATISTICA. vol. LX, pp. 12 ISSN: 0390-590X.
6. MINERVA T., PATERLINI S., POLI I. (2000). A genetic algorithm for neural network design. ECONOMICS & COMPLEXITY. ISSN: 1398-1706.
7. Minerva T., Paterlini S., POLI I. (1999). Algoritmi ibridi per l'analisi di serie storiche finanziarie. SCIENZA & BUSINESS. vol. 3-4, pp. 56-77 ISSN: 1594-1418.
8. POLI I., Roverato A. (1998). A genetic algorithm for graphical model selection. JOURNAL OF THE ITALIAN STATISTICAL SOCIETY. vol. 7, pp. 197-208 ISSN: 1121-9130.
9. POLI I., JONES R.D. (1994). A neural net model for prediction. JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION. vol. 89 ISSN: 0162-1459.
10. T. MINERVA, POLI I., G. SOLINAS. (2006). Un automa cellulare per lo studio del distretto industriale. In: A. NATALI, M. RUSSO, G. SOLINAS. Distretti industriali e sviluppo locale. BOLOGNA: IL MULINO (ITALY).
11. M. BEDAU, A. BUCHANAN, G. GAZZOLA, M. HANCZYC, N. PACKARD, POLI I. (2005). Evolutionary design of a DDPD model of ligation., In: M. BEDAU. Lecture Notes in Computer Science. in corso di stampa. BERLIN: Springer (GERMANY).
12. Minerva T., POLI I. (2001). A neural net model to predict high tides in Venice. In: Borra S., Rocci R., Vichi M., Schader M. Advances in Classification and Data Analysis. BERLIN: Springer (GERMANY).
13. Minerva T., POLI I. (2001). Building ARMA Models with Genetic Algorithms. In: Boers E.J.W., et al. Applications of Evolutionary Computing. (pp. 516). BERLIN: Springer (GERMANY).
14. M. FORLIN, POLI I., D. DE MARCH, N. PACKARD, R. SERRA. (2007). Evolving the experimental design for amphiphilic systems. European Symposium on Nature-inspired Smart Information Systems, (www.nisis.de).
15. R. SERRA, T. CARLETTI, POLI I., M. VILLANI, A. FILISETTI. (2007). Protocell Models and Synchronization. Conference on Systems Science.
16. D. SLANZI, M. TREVISAN, L. VILLANOVA, G. PALU', POLI I. (2007). Bayesian Analysis of Citomegalovirus Infection from microarray data., Risk and Prediction, SIS.
17. D. SLANZI, POLI I., D. DE MARCH, M. FORLIN. (2007). Bayesian Networks for Detecting Relevant Variable Interaction for Biochemical Experiments. Risk and Prediction, SIS.
18. R. SERRA, T. CARLETTI, POLI I., M. VILLANI, A. FILISETTI. (2007). Conditions for emergent synchronization in protocells. EUROPEAN CONFERENCE ON COMPLEX SYSTEMS.
19. R. SERRA, T. CARLETTI, POLI I. (2006). Surface Reaction Models of Protocells. International Symposium on Mathematical and Computational Biology. 25-30 November. MANAUS, AM: (BRAZIL).
20. AGOSTINELLI, C., POLI I. (2003). Evolving classification and regression trees. Proceedings of the 18th International Workshop on Statistical Modelling, Leuven, Belgium. 7-11 July, 2003.
21. AGOSTINELLI C., CUCINA D., POLI I. (2005). A stochastic electric field particle dynamics (EFPD) model for a microfluidic system. Rapporto di Ricerca, Dipartimento di Statistica, Università Ca' Foscari, Venezia. (vol. 1).
22. MINERVA T., POLI I. (2005). Evo-Stat: Evolutionary Statistical Modeling. Research Report, Dep. of Social, Cognitive and Quantitative Science, Un. of Modena and Reggio Emilia. (vol. 36).
23. POLI I., R. SERRA, M. ROMANAZZI, C. AGOSTINELLI. (2005). Evolutionary Designs For Microfluidic Experiments. Rapporto di ricerca, Dipartimento di Statistica, Università Ca' Foscari, Venezia. (vol. 10).

7 - Elenco dei partecipanti all'Unità di Ricerca

7.1 - Componenti

Componenti della sede dell'Unità di Ricerca

n°	Cognome	Nome	Università/Ente	Qualifica	Impegno	
					1° anno	2° anno
1.	PARPINEL	Francesca	Università "Ca' Foscari" di VENEZIA	Professore Associato confermato	10	10
2.	PIZZI	Claudio	Università "Ca' Foscari" di VENEZIA	Professore Associato non confermato	9	6
3.	POLI	Irene	Università "Ca' Foscari" di VENEZIA	Professore Ordinario	10	10
TOTALE					29	26

Componenti di altre Università / Enti vigilati

Nessuno

Titolari di assegni di ricerca

n°	Cognome	Nome	Università/Ente	Impegno	
				1° anno	2° anno
1.	SLANZI	Debora	Università "Ca' Foscari" di VENEZIA	9	9
TOTALE				9	9

Titolari di borse

Nessuno

7.2 - Altro personale

n°	Cognome	Nome	Università/Ente	Dipartimento	Qualifica	Impegno 1° anno	Impegno 2° anno
1.	DE MARCH	DAVIDE	Ca' Foscari	Statistica	laureato	6	6
2.	FORLIN	MICHELE	Ca' Foscari	Statistica	laureato	6	6
TOTALE						12	12

7.3 - Personale a contratto da destinare a questo specifico Progetto

n°	Tipologia di contratto	Costo previsto	Impegno 1° anno	Impegno 2° anno	Note
1.	Assegnista	19.000	6	6	6 Si ritiene utile la presenza di un assegnista dedicato alla modellazione per l'evoluzione delle popolazioni degli esperimenti
TOTALE		19.000	6	6	

7.4 - Dottorati a carico del PRIN da destinare a questo specifico Progetto

Nessuno

8 - Titolo specifico del Progetto svolto dall'Unità di Ricerca

Testo italiano

Disegni sperimentali evolutivi

Testo inglese

Evolutionary Experimental Designs

9 - Abstract del Progetto svolto dall'Unità di Ricerca

Testo italiano

In questo progetto di ricerca l'Unità di Venezia si propone di elaborare nuovi strumenti metodologici per il disegno statistico sperimentale utilizzando come approccio il calcolo evolutivo.

Il disegno sperimentale anziché essere definito a priori evolverà con la costruzione di algoritmi genetici. La modellazione statistica, nella forma di modelli di regressione polinomiale, di regressione quantilica, reti Bayesiane, reti neurali, consentirà una trasformazione da una generazione all'altra dell'algoritmo, in modo tale che siano i modelli a guidare un'evoluzione intelligente del disegno.

Testo inglese

In this research project the Unit of Venice would participate in developing new methodological tools for Statistical Design of Experiments adopting the evolutionary computational approach.

We will not define an initial Design of Experiments but we will evolve it with the construction of genetic algorithms.

The statistical modelling, in the form of polynomial regression models, quantile regressions, Bayesian networks, neural networks, will allow a transformation between two successive generations of the algorithm, so that the models drive an intelligent evolution of the Design.

10 - Parole chiave

n°	Parola chiave (in italiano)	Parola chiave (in inglese)
1.	DISEGNO DEGLI ESPERIMENTI	DESIGN OF EXPERIMENTS
2.	CALCOLO EVOLUTIVO	EVOLUTIONARY COMPUTATION
3.	ALGORITMI GENETICI	GENETIC ALGORITHMS
4.	ALTA DIMENSIONALITÀ	HIGH DIMENSIONALITY
5.	MODELLAZIONE DATI SPERIMENTALI	EXPERIMENTAL DATA MODELLING

11 - Stato dell'arte

Testo italiano

In numerosi settori della ricerca scientifica emerge con crescente frequenza il problema di sviluppare studi in spazi ad alta dimensionalità e con scarse conoscenze a priori. Si vuole infatti capire la dinamica di sistemi complessi indagando sulle numerose componenti che caratterizzano tali sistemi e sulla rete d'interazione tra tali componenti che porta all'emergere di nuove funzionalità o nuove entità. La caratterizzazione di tali sistemi può risultare però estremamente difficile a causa del numero eccessivamente elevato di variabili che possono intervenire nella dinamica del sistema.

In questa ricerca si vuole affrontare il tema della costruzione di strumenti metodologici per realizzare disegni sperimentali in presenza di alta dimensionalità, con riferimento particolare al settore biochimico e ai sistemi molecolari. Lo studio e la progettazione di tecnologie innovative richiede infatti la definizione di disegni sperimentali e di modelli fisici e statistici in grado di condurre i ricercatori nelle aree più informative dello spazio multidimensionale di ricerca. Più specificamente l'oggetto di studio riguarda il disegno sperimentale per la costruzione di vescicole artificiali; questo tema rappresenta attualmente un settore emergente delle bionanotecnologie con numerosi campi di applicazione che includono la tecnologia dell'informazione, la medicina, il controllo dell'ambiente. Gli esperimenti, che sono disegnati per studiare la formazione e l'evoluzione di entità molecolari complesse in ambienti mutevoli, possono essere condotti con un grado crescente di integrazione e controllo dell'ambiente chimico, dai pipetting systems per piattaforme a micropozzetti, alle reti microfluidiche totalmente integrate e programmabili. Lo scopo principale di tali esperimenti concerne l'identificazione della composizione chimica che in un particolare ambiente costruito conduce alla formazione e all'evoluzione di vescicole. Questo processo richiede l'esplorazione di uno spazio combinatoriale complesso: quale tipo di molecole scegliere da una libreria composta da diverse decine differenti di molecole; con quali livelli di concentrazione devono essere considerate nella composizione; quali parametri sperimentali adottare, come tempi di reazione, livelli di temperatura, sequenzialità delle fasi sperimentali. Tutti questi elementi (fattori e livelli) su cui indagare producono un'esplosione combinatoriale di variabili caratterizzanti il processo. Nel disegnare gli esperimenti il ricercatore deve proporre un insieme piccolo di variabili e livelli e un numero piccolo di esperimenti (gli esperimenti in questo settore sono costosi e impegnativi). Disegnare questi esperimenti rappresenta perciò una sfida scientifica che richiede strumenti metodologici nuovi (Cawse).

Il calcolo evolutivo rappresenta un approccio computazionale che può essere adottato per disegnare esperimenti ad alta complessità combinatoriale. Questo approccio consente di evolvere il disegno sperimentale (anziché impegnarsi in un'unica scelta) e l'evoluzione avviene in popolazioni di disegni che sequenzialmente si avvicinano al disegno con esperimenti ottimali. In letteratura si stanno sviluppando alcuni approcci al disegno sperimentale ad alta dimensionalità costruiti per specifiche applicazioni e basati su analisi di tipo simulativo e non sperimentale. Inoltre in questi approcci non è presente la modellazione statistica per evolvere le popolazioni di esperimenti. La modellazione statistica, articolata con diverse tipologie di modelli (lineari di regressione e non lineari di regressione o con reti neurali, o con reti Bayesiane) potrebbe rappresentare un elemento estremamente importante nella costruzione dell'approccio evolutivo al disegno sperimentale che si propone per questa ricerca.

Testo inglese

In several sectors of the scientific research frequently emerges the problem of developing studies in high dimensional settings and with poor initial knowledge. One would like to understand the dynamics of complex systems studying the large number of components that characterize these systems and in particular the network of interactions among the components that leads to the emergence of new functionalities or new entities. Characterizing these systems however may result very difficult because of the excessive number of variables that can affect their dynamics.

In this research we would like to tackle the theme of the construction of methodological tools to realize experimental designs in presence of high dimensionality, with particular interest for the biochemical sector and for molecular systems. The study and the creation of innovative technologies requires in fact the formulation of experimental designs and physical and statistical models able to lead the experimentalist in the most informative area of the multidimensional space of research. More specifically the object of this study concerns the experimental design for the formation of artificial vesicles; this theme represents at the moment an emergent sector of the Bionanotechnologies with several sectors of applications which include the technology of information, medicine, and environment control. The experiments, which are designed to investigate the formation and the evolution of complex entities in given engineered environments, can be conducted with increasing degrees of integration and control over the chemical environment from pipetting systems for microwell plates to fully integrated programmable microfluidic networks.

The experiments are conducted for the identification of the chemical composition that in a particular environment allow the formation and the evolution of vesicles. This process requires the exploration of a combinatorial complex space: which kind of molecule we should choose in a library composed of almost hundred molecules; with which level of concentration they should be considered in the composition, which experimental parameters we should adopt, such as reaction time, temperature levels, order in the experimental phases. All these elements on which to investigate create what is known as the combinatorial explosion of variables characterizing the process.

In designing the experiments the researcher must propose a small number of variables and levels and a small number of experiments (experiments in this sector are particularly expensive and require special competence). Designing these experiments represent then a scientific challenge that requires new methodological tool (Cawse, 2003).

The evolutionary computation represents an approach that may be appropriate to design experiments with high combinatorial complexity. This approach allows to evolve the experimental design (instead to commit to a single choice) and the evolution is realized in populations of designs that sequentially approach the design with optimal experiments. In the literature have been developing few approaches to the design of experiments in high dimensional settings, built for particular applications or conducted with simulation studies instead of real experiments. Moreover in these approach there is any statistical modelling to evolve the population of experiments. The statistical modelling developed with several different structures (linear and non linear regression models, neural networks, Bayesian networks) might represent an extremely important element in the construction of the evolutionary approach to the experimental design.

12 Riferimenti bibliografici

Beck-Sickinger A., Weber P. (2002) "Combinatorial Strategy in Biology and Chemistry", Wiley.

Bedau M. A., Buchanan A., Gazzola G., Hanczyc M., Maeke T., McCaskill J.S., Poli I., Packard N. H. (2005) "Evolutionary design of a DDPD model of ligation" Lecture Notes in Computer Science, 3871 pp. 201-212.

Broudiscou A., Leardi R., Phan-Tan-Luu R. (1996) "Generic algorithms as a tool for selection of D-optimal design", Chemometr. Intell. Lab. Syst, 35, pp. 105-116.

Cawse J. N. (Ed.) (2003) "Experimental Design for Combinatorial and High Throughput Materials Development", Wiley, New York.

Cornell J. A. (2002) "Experiments with Mixtures: Designs, Models and the Analysis of Mixture Data", Wiley, New York.

- Cox D. R., Reid N.(Eds.) (2000) "The theory of the design of experiments", Chapman & Hall, London.
- Ernst J., Nau G.J., Bar-Joseph Z. (2005) "Clustering short tie series gene expression data", *Bioinformatics*, 21, pp. 159-168.
- Forlin M., Poli I., De March D., Packard N., Gazzola G., Serra R. (2007) "Evolutionary experiments for self-assembling amphiphilic systems", in *Chemometrics and Intelligent Laboratory Systems*.
- Goldberg D. E. (1989) "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley, MA.
- Hanczyc M., Szostak J. (2004) "Replicating vesicles as models of primitive cell growth and division", *Current Opinion in Chemical Biology* 8, pp. 660-664.
- Hereida-Langner, Carlyle W.M., Montgomery D.C., Borrer C.M., Runger G.C. (2003) "Genetic algorithms for the construction of D-optimal designs", *Journal of Quality Technology*, 35, pp. 28-46.
- Holland J. H. (1992) "Adaptation in Natural and Artificial Systems", MIT Press, Cambridge, MA, 1992.
- Jensen F.V. (2001) "Bayesian Network and Decision Graphs", Springer, Verlag, New York.
- Kirsten G., Maier W. F. (2004) "Strategies for the discovery of new catalysts with. combinatorial chemistry", *Applied Surface Science*, 223, pp. 87-101.
- Koehler R. (2005) "Quantile regression", Cambridge University Press.
- Lewis S.M., Dean A.M. (2001) "Detection of interactions in experiments on large numbers of factors", *J.R. Statist. Soc. B*, 63, pp. 633-672.
- Luisi P. (2007) "The emergence of life" Cambridge University Press.
- Minerva T., Poli I. (2001) in J.W. Boers et al. (Eds.), "Applications of Evolutionary Computing" Springer, Berlin.
- Montgomery D. C. (2005) *Design and Analysis of Experiment*, Wiley & Son, NY.
- Parpinel F., Mestriner A. (2007) "Quantile regression for Italian reference growth charts", in *Book of short papers S.Co. 2007, fifth Conference*, edited by P.Mantovan, A.Pastore, S.Tonellato, pp.368-373
- Patri A.K., Kukowska-Latalo J.F., Baker J.R. (2005) "Targeted drug delivery with dendrimers: Comparison of the release kinetics of covalently conjugated drug and non-covalent drug inclusion complex". *Advanced Drug Delivery Reviews*, 57, pp. 2203-2214.
- Potyrailo R.A., Amis E.J. (2003) "High-Throughput analysis", Kluwer Academic, NY.
- Serra R., Carletti T., Poli I. (2007) "Synchronization phenomena in surfacereaction models of protocells" *Artificial Life*, 13, 2, pp. 123-138.
- Sexton C.J., Anthony D. K., Lewis S.M., Please C. P., Keane A.J. (2006) "Design of experiment algorithms for assembled products", *Journal of Quality Technology*, 38, pp. 298-308.
- Sipper M. (2002) "Machine Nature", McGraw-Hill, New York.
- Theis M., Gazzola G., Forlin M., Poli I., Hanczyc M., Bedau M. A. (2006) *Optimal Formulation of Complex Chemical Systems with a Genetic Algorithm*. ECCS Proceedings, Oxford.

13 - Descrizione del programma e dei compiti dell'Unità di Ricerca

Testo italiano

Nell'ambito di questo progetto di ricerca, che si propone di riformulare alcune aree della Statistica adottando il paradigma dell'evoluzione dei sistemi naturali, l'Unità di Venezia intende sviluppare alcune metodologie nel settore del disegno sperimentale.

Il disegno statistico sperimentale, inteso come approccio scientifico alla pianificazione degli esperimenti (Montgomery, 2005), rappresenta un complesso processo di analisi e modellazione che richiede l'identificazione e la scelta di un insieme limitato di valori in uno spazio multidimensionale. I valori devono essere in grado di informare il ricercatore sul processo generatore e essere quindi rappresentativi dello spazio di definizione del fenomeno oggetto di studio. Nel disegno classico degli esperimenti la rappresentatività dei punti sperimentali viene in genere ottenuta seguendo alcuni principi fondamentali come la casualizzazione, la replicazione, la struttura a blocchi e l'ortogonalità. La modellazione realizzata generalmente con forme polinomiali di vario ordine consente quindi di interpretare il diverso ruolo esercitato dalle variabili e dai loro livelli nel determinare la risposta del sistema.

Quando il numero delle variabili cresce il disegno statistico classico è inadeguato a rispondere soddisfacentemente alle esigenze del ricercatore che spesso si trova con scarse conoscenze a priori o intuizioni sul meccanismo generatore dei dati. L'alta dimensionalità (HD) del fenomeno può essere affrontata con un approccio che non tende a trasformare la struttura stessa del fenomeno in modo da ridurre le dimensioni (come l'approccio delle componenti principali) ma sceglie le dimensioni (variabili di processo e di composizione, livelli e interazioni) e perciò la struttura dei singoli esperimenti secondo un criterio di ottimizzazione evolutiva.

Più specificamente questa ricerca si propone di costruire un approccio metodologico al disegno sperimentale che consente di evolvere la struttura di singoli esperimenti verso la regione dello spazio multidimensionale ove si posizionano valori ottimali di risposta del sistema. Gli esperimenti che si considereranno sono di tipo biochimico e la ricerca si sviluppa nel settore noto come "chimica combinatoriale". Considereremo esperimenti mistura (Cornell, 2002) caratterizzati dalla miscelazione di un insieme di molecole in un ambiente acquoso, secondo un particolare protocollo sperimentale (livello di temperatura, tempi di riscaldamento, ecc.). Le molecole sono di tipo anfifilico e in un particolare ambiente tendono ad auto-organizzarsi in forme specifiche e con diverse funzionalità. L'interesse in questo studio è sulla formazione di vescicole (Luisi, 2007), intese come nanocontenitori in grado per esempio di rilasciare un farmaco in aree di un organismo alterate da elementi patogeni (Hanczyc, Szostak, 2004; Patri et al., 2005; Thesis, 2006).

Il problema più rilevante che si pone allo sperimentatore concerne la scelta del tipo di molecole (variabili o fattori dell'esperimento) che possono essere selezionate in una libreria con circa un centinaio di tipi, la scelta del livello di concentrazione (livelli delle variabili), il grado della temperatura e di acidità della soluzione. La tecnologia con cui si realizzano gli esperimenti è nota come microwell plate, data da griglie di 96 micropozzetti non comunicanti, in ciascuno dei quali si realizza un esperimento. L'insieme degli esperimenti viene replicato alcune volte per avere informazioni anche sul disturbo casuale. Il livello di torbidità della soluzione rappresenta la risposta del sistema ed è misurata da un microscopio fluorescente: un più alto livello di torbidità indica una presenza maggiore di vescicole nel sistema. L'elevata onerosità degli esperimenti in termini sia economici sia time consuming richiedono un'adeguata progettazione degli esperimenti. Lo scopo principale del disegno sperimentale concerne quindi la scelta della miscela molecolare che conducendo ai livelli più alti di torbidità massimizza la risposta del sistema.

In questa ricerca noi intendiamo ottenere il disegno sperimentale con un approccio evolutivo basato su algoritmi genetici (Goldberg, 1989; Holland, 1991; Sipper, 2002; Lewis, Dean, 2001). Secondo il paradigma evolutivo il disegno sperimentale verrà concepito come il risultato di un processo dinamico dove una popolazione di possibili candidati di diverse misture è generata ed evoluta in differenti generazioni per effetto di un insieme di operatori genetici, come la selezione naturale, l'incrocio, la mutazione, l'innovazione. Con questo approccio noi costruiamo un disegno stocastico "population-based" dove ciascun esperimento (mistura), codificato come una sequenza genetica, rappresenterà un membro delle diverse popolazioni. Ogni sequenza genetica verrà quindi valutata in relazione al livello di risposta del sistema misurato come livello di torbidità della miscela. L'algoritmo inizierà costruendo una prima popolazione di esperimenti (sequenze genetiche) scegliendo in modo casuale le variabili da considerare e i loro livelli di concentrazione. Le sequenze verranno quindi valutate con esperimenti realizzati presso il laboratorio Nanofab del Parco Scientifico e Tecnologico di Venezia. La risposta di ogni esperimento verrà considerata come fitness dell'algoritmo e sulla base di questa valutazione si procederà all'applicazione degli operatori genetici che trasformeranno la struttura della popolazione ottenendo così la seconda generazione. Questa popolazione di esperimenti verrà quindi sottoposta ai ricercatori chimici per l'esecuzione degli stessi. L'algoritmo procederà per un insieme limitato di

generazioni, costruendo una sorta di dialogo tra disegno ed esperimento reale, e realizzando in tal modo un'evoluzione del disegno sperimentale verso la formulazione ottimale. I primi risultati di questo approccio, presentati in Forlin, et al. (2007), sono stati estremamente incoraggianti: scegliendo solo 5 variabili da un insieme di 16 e livelli di concentrazione non superiori a 6, il disegno statistico classico noto come {16,5} Simplex Lattice Design (Cornell, 2002; Cox, Reid, 2000) avrebbe richiesto l'esecuzione di 15.504 esperimenti. In questo studio scegliendo con l'algoritmo soltanto 180 differenti esperimenti si sono ottenuti risultati molto soddisfacenti in relazione all'obiettivo da ottimizzare.

L'algoritmo genetico semplice costruito rivela però un punto di debolezza legato alla casualità del processo evolutivo che se da un lato consente di esplorare lo spazio sperimentale in modo efficace dall'altro può portare alla distruzione di elementi di conoscenza (frammenti di sequenze genetiche) estremamente utili per la scoperta della mistura ottimale. In questa ricerca intendiamo costruire una nuova classe di algoritmi genetici, la cui evoluzione viene guidata da modelli statistici. Questa classe, che chiameremo Model-based Genetic Algorithms (MGA), prevede ad ogni generazione dell'algoritmo la costruzione di modelli statistici per i dati ottenuti dagli esperimenti: con i modelli potremo scoprire le componenti (molecole e interazioni tra molecole) più significative e influenti nel determinare i livelli alti di risposta del sistema e ottenere previsioni sullo spazio sperimentale non esplorato da quella generazione di esperimenti. Dai modelli otterremo così informazioni molto preziose sul ruolo di certe componenti nell'esperimento e svilupperemo vie appropriate per trasmettere tali informazioni nelle generazioni successive. Considereremo modelli sia di tipo lineare che non lineare: le strutture polinomiali con diversi ordini di interazione verranno sviluppate e confrontate con modelli di reti neurali e reti Bayesiane. Il confronto concernerà la loro capacità di rappresentazione e di previsione. L'evoluzione dell'algoritmo, che può essere considerata una realizzazione di un processo temporale, verrà quindi valutata con modelli di regressione quantitativa (Koenker, 2005, Parpinel e Mestriner, 2007) che ci informeranno sulla distribuzione degli esperimenti in termini di risposta del sistema: stimeremo la posizione della risposta degli esperimenti nelle diverse generazioni, e ciò consentirà una valutazione della variazione nelle proporzioni dei migliori e dei peggiori esperimenti. Inoltre la proporzione di variazione con differenti operatori genetici consentirà un efficace monitoraggio della struttura evolutiva del processo.

I dati generati dall'evoluzione genetica degli esperimenti possono essere studiati come un panel di serie storiche brevi. Lo studio di tali serie storiche brevi sarà condotto seguendo sia l'approccio parametrico sia quello non parametrico al fine di distinguere pattern di geni reali da quelli casuali nell'evoluzione della struttura genetica (Ernst et al. 2005).

L'approccio evolutivo che proponiamo è, a nostra conoscenza, nuovo e innovativo nel settore del disegno degli esperimenti. La complessità combinatoriale dello spazio di ricerca caratterizza ormai molti settori scientifici (Potyrailo and Amis, 2003; Beck-Sickinger, Weber, 2002) ove la capacità computazionale acquisita conduce a esplorare spazi di ricerca sempre più ampi e inaspettati, ma richiede strumenti metodologici in grado di affrontare la crescente complessità dei sistemi. La collaborazione con le altre Unità del progetto sarà estremamente utile nello sviluppo di questi nuovi strumenti metodologici.

Testo inglese

In the framework of the national research project, that aims to reformulate some areas of Statistics with the paradigm of the evolution of the natural systems, the Unit of Venice would participate in developing some methodologies for the "Design of Experiments".

The statistical design of experiments, regarded as a scientific approach to planning the experiments (Montgomery, 2005), represents a complex process of analyses and modelling which requires the identification and the selection of a limited number of values in a multidimensional space. These values should be able to inform the researcher on the generating process and be representative of the space of definition of the phenomenon under study.

In the classical design of experiments the capacity of the experimental points to be representative is generally achieved following some fundamental principles, such as randomization, replication and structures by blocking. Modelling, generally realized with polynomial forms of different orders allow then to interpret the different role played by the variables and their levels in determining the answer of the system.

When the number of the variables involved grows the classical statistical design doesn't seem able to give a satisfactory answer to the request of the researcher, who frequently is with a poor initial knowledge or lacking in intuition on the generating process. The high dimensionality (HD) of the phenomenon can be addressed with an approach that doesn't transform the structure of the phenomenon to reduce the dimensionality (as with the principal components analysis), but selects the dimensions (process and composition variables, levels and interactions) and thus the structure of the single experiments according to a criterion of evolutionary optimization.

More specifically the goal of this research is to build a methodological approach for the design of experiments that evolves the structure of the single experiments toward the region of the multidimensional space where are located the optimal values of the systems response. We will study biochemical experiments and the research can be regarded in the area of "combinatorial chemistry". We will consider Mixture Experiments (Cornell, 2002) characterized by mixing a set of molecules in an aqueous composition, according to a particular experimental protocol (degree of temperature, heating time, etc.). The molecules will be amphiphilic molecules: in a particular solution they tend to self-organize and form specific entities with different functionalities. In this study the main interest is in the formation of vesicles (Luisi, 2007), regarded as nanocontainers, able for example to deliver a drug in a particular area of an organism altered by a pathogenic component (Hanczyc and Szostak, 2004; Patri et al. 2005, Theis, 2006).

The scientific question to the experimenter regards the choice of the kind of molecules, that may be selected in a library with almost one hundred kinds (the variables or the factors of the experiments) the choice of the concentration levels (the variables levels), the degree of the temperature, the degree of acidity of the solution. The technology on which the experiments are developed is known as "microwell plate", and consist of an array with 96 microwells not communicating; in any microwell an experiment is conducted. The set of experiments is replicated few times to achieve information also on the random noise.

The answer of the system is a measure taken by a fluorescent microscope and represents the level of turbidity of the system (higher is the turbidity and higher is expected the presence of vesicles in the systems). The experiments are very expensive and very demanding in terms of scientific competence and time. The principal aim of the experimental design concerns the selection of the molecular mixture that leading to the highest levels of turbidity maximizes the response of the system. In this research we plan to achieve the experimental design with an evolutionary approach based on genetic algorithms (Goldberg, 1989; Holland, 1992; Sipper, 2002; Lewis and Dean, 2001). According to this approach the experimental design will be seen as a result of a dynamical process where a population of possible candidates of different mixtures will be generated and evolved in different generations through a set of genetic operators, e.g. selection, crossover, mutation, innovation. With these algorithms we will build a population-based stochastic design, whereby each member of the populations, representing an experiment (e.g. a mixture), is coded as a genetic sequence, and evaluated in terms of the system response: the turbidity level. The algorithm will begin by building a first population of experiments (genetic sequences) designed with a random choice of the variables and their level of concentration. Each sequence will then be evaluated conducting a real experiment according to the structure derived by the algorithm. The experiments will be conducted at the Scientific and Technological Park of Venice in the NanoFab Laboratory. The response of each experiment will be regarded as the fitness function of the algorithm, and on the basis of this evaluation we will proceed with the application of genetic operators that will transform the structure of the population of experiments achieving what we will regard as the population of second generation of experiments. This population of candidate mixtures will then be proposed to the chemical experimentalists to conduct the real experiments. The algorithm will proceed for a limited number of generations creating a sort of dialog between design and real experiments and achieving in this way an evolution of the design towards the optimal formulation.

First results of this approach have been extremely encouraging (Forlin, et al, 2007). Selecting just 5 variables among a set of 16, and a number of levels of concentrations not higher than 6, the classical statistical design, known as {16,5}Simplex Lattice Design (Cornell, 2002; Cox and Reid, 2000), would have required the execution of 15,504 experiments. In this study we executed just 180 different experiments (mixtures achieved with the evolution of a simple genetic algorithm) achieving very satisfactory results in term of the objective to optimize. The simple genetic algorithm presented and evaluated in this first study reveals however a weak point related to the randomness of the evolutionary process that if from one side it lets to explore the experimental space in an efficient way, from another side it can lead to the destruction of pieces of knowledge (fragments of genetic sequences) extremely useful for the discovery of the optimal mixture. In this research we intend to build a new class of genetic algorithms whose evolutionary process is guided by statistical models. This class, that we will call Model-based Genetic Algorithms (MGA), will involve at each generation of the algorithm the construction of statistical models for the data achieved from the experiments: with the models we will be able to discover the components (molecules and interactions among molecules) more relevant and influent in determining the levels of high response of the systems and obtain predictions on the unexplored experimental space.

From models we will then obtain very precious information on the role of some components in the experiment and will develop methodologies to transfer this information to the next generations. We will consider both linear and nonlinear models. Polynomial structures with different order of interaction will be developed and confronted with neural network models and with Bayesian networks. The comparison will concern their capacity to describe and to predict the system. The evolution of the algorithm, that can be regarded as the realization of a time process, will be evaluated with models of quantile regression (Koenker, 2005; Parpinel e Mestriner, 2007) that will inform on the distribution of the experiments in terms of the response to the system: we will estimate the location of the experiments in the different generations, the variation in the proportion of the best and the worst experiments. Moreover the proportion of variation due to the different genetic operators will allow a monitor action on the generation development of the experiments for an adaptive evolution. The data generated from the genetic evolution of the experiments will also be studied as a panel of short time series. The study of such short time series will be conducted using both the parametric and the nonparametric approach, with the aim of discriminating real and relevant patterns from noise in the evolution of the genetic structure (Ernst et al., 2005).

The evolutionary approach that we propose in this research is, at our knowledge, new and innovative in the area of Design of Experiments. The combinatorial complexity of the space of research characterizes in fact several scientific sectors (Potyrailo and Amis, 2003; Beck-Sickinger, Weber, 2003), where the achieved computational capacities leads to explore larger and larger experimental spaces, but it requires methodological tools able to face the growing complexity of the systems

The collaboration with the other Units of research of this project will be extremely useful in the development of these new methodological tools.

14 - Descrizione delle attrezzature già disponibili ed utilizzabili per la ricerca proposta

Testo italiano

Nessuna

Testo inglese

Nessuna

15 - Descrizione delle Grandi attrezzature da acquisire (GA)

Testo italiano

Nessuna

Testo inglese

Nessuna

16 - Mesi persona complessivi dedicati al Progetto

		Numero	Impegno 1° anno	Impegno 2° anno	Totale mesi persona
<i>Componenti della sede dell'Unità di Ricerca</i>		3	29	26	55
<i>Componenti di altre Università/Enti vigilati</i>		0			
<i>Titolari di assegni di ricerca</i>		1	9	9	18
<i>Titolari di borse</i>	<i>Dottorato</i>	0			
	<i>Post-dottorato</i>	0			
	<i>Scuola di Specializzazione</i>	0			
<i>Personale a contratto</i>	<i>Assegnisti</i>	1	6	6	12
	<i>Borsisti</i>	0			
	<i>Altre tipologie</i>	0			
<i>Dottorati a carico del PRIN da destinare a questo specifico progetto</i>		0	0	0	0
<i>Altro personale</i>		2	12	12	24
TOTALE		7	56	53	109

17 - Costo complessivo del Progetto dell'Unità articolato per voci

Voce di spesa	Spesa in Euro	Descrizione dettagliata (in italiano)	Descrizione dettagliata (in inglese)
Materiale inventariabile	9.000	Acquisto personal computer, stampanti e altre periferiche. Libri e riviste.	Personal Computers, printers and other devices. Books and journals.
Grandi Attrezzature	0		
Materiale di consumo e funzionamento	6.000	Acquisto materiali di cancelleria, CDR, DVDR, spese postali e telefoniche. Prodotti chimici per esperimenti	Writing materials, paper, CDR, DVDR, telephone and mail expenses. Chemical products for experiments.
Spese per calcolo ed elaborazione dati	500	Acquisto software, elaborazione dati.	Software and computing services
Personale a contratto	19.000	Assegno di ricerca dedicato alla modellazione per l'evoluzione delle popolazioni degli esperimenti	Research fellowship for modelling the evolution of the populations of the experiments
Dottorati a carico del PRIN da destinare a questo specifico progetto	0		
Servizi esterni			
Missioni	3.500	Partecipazione a riunioni e convegni.	Participation to workshop and conference
Pubblicazioni			
	2.000	Partecipazione e organizzazione di convegni nazionali ed	National and international scientific conferences

Partecipazione / Organizzazione convegni		<i>internazionali, organizzazione riunioni intermedie gruppo di ricerca.</i>	<i>and workshops. Organization of intermediate workshops.</i>
Altro			
TOTALE	40.000		

18 - Prospetto finanziario dell'Unità di Ricerca

Voce di spesa	Importo in Euro
a.1) finanziamenti diretti, disponibili da parte di Università/Enti vigilati di appartenenza dei ricercatori dell'unità operativa	2.000
a.2) finanziamenti diretti acquisibili con certezza da parte di Università/Enti vigilati di appartenenza dei ricercatori dell'unità operativa	10.000
b.1) finanziamenti diretti disponibili messi a disposizione da parte di soggetti esterni	
b.2) finanziamenti diretti acquisibili con certezza, messi a disposizione da parte di soggetti esterni	
c) cofinanziamento richiesto al MUR	28.000
Totale	40.000

19 - Certifico la dichiarata disponibilità e l'utilizzabilità dei finanziamenti a.1) a.2) b.1) b.2)

SI

Firma _____

(per la copia da depositare presso l'Ateneo e per l'assenso alla diffusione via Internet delle informazioni riguardanti i programmi finanziati e la loro elaborazione necessaria alle valutazioni; D. Lgs, 196 del 30.6.2003 sulla "Tutela dei dati personali")

Firma _____

Data 24/10/2007 ore 19:25

ALLEGATO

Curricula scientifici dei componenti il gruppo di ricerca

Testo italiano

1. **PARPINEL Francesca**

Curriculum:

Francesca Parpinel è professore associato di Statistica presso l'Università Ca' Foscari di Venezia (www.unive.it). Laureata presso l'Università di Padova in Scienze Statistiche ed Economiche, è Dottore di Ricerca in Statistica (Università di Padova). È stata incaricata con un contratto di studio e di ricerca dalla Commissione Europea Centro Comune di Ricerca-Istituto di Ingegneria dei Sistemi e dell'Informatica, ISPRA, su "Diagnostica di processi non lineari". Ha trascorso un periodo di studio presso l'Università Parigi VI.

Ha continuato i suoi studi presso il Dipartimento di Scienze Statistiche dell'Università di Padova con una borsa di studio per l'attività di ricerca post-dottorato.

È stata chiamata in qualità di ricercatore di Statistica presso la Facoltà di Economia e Commercio dall'Università di Venezia

I suoi più importanti interessi di ricerca sono stati nell'ambito dell'inferenza statistica per processi stocastici a tempo continuo retti da equazioni differenziali stocastiche, dei modelli lineari e non lineari per serie storiche, dell'inferenza non parametrica, e più recentemente della regressione quantilica. Nell'ambito della formazione statistica ha curato diverse pubblicazioni per una delle quali ha ricevuto il premio come miglior lavoro per la didattica della statistica dalla Società Italiana di Statistica.

È membro della Società Italiana di Statistica.

Attualmente è membro del Collegio Didattico del Dottorato d'Eccellenza in Economia e Organizzazione dell'Università Ca' Foscari

Ha partecipato a numerosi progetti di ricerca.

Pubblicazioni:

◆ PARPINEL F., PIZZI C. (2002). *Iterative estimation procedure for option pricing with stochastic volatility models*. METRON. vol. LX, n.1-2, pp. 211-223 ISSN: 0026-1424.

◆ PARPINEL F., PROVASI C. (2004). *Elementi di Probabilità e Statistica per le Scienze Economiche*. ISBN: 88-348-4378-9. TORINO: GIAPPICHELLI (ITALY).

◆ FIORIN S, PARPINEL F. (2007). *Simulation study for strong consistency in the weak topology of LS estimates for a functional linear model*. S.Co.2007 Complex models and computational intensive methods for estimation and prediction. 6-8 Settembre 2007. (pp. 225-230). ISBN/ISSN: 978-88-6129-114-0. PADOVA: Cleup (ITALY).

◆ PARPINEL F., MESTRINER A. (2007). *Quantile regression for Italian reference growth charts*. S.Co.2007 Complex models and computational intensive methods for estimation and prediction. 6-8 Settembre 2007. (pp. 368-373). ISBN/ISSN: 978-88-6129-114-0. PADOVA: cleup (ITALY).

◆ PARPINEL F., PIZZI C. (2006). *Nonparametric strategies applied to time series analysis*. SER2006. 18-19 aprile 2006. (pp. 203-206).

◆ PARPINEL F. (2002). *Testing unit roots in time series analysis through permutation methods*. SIS, XLI riunione scientifica, Milano. 5-7 giugno 2002. (pp. 635-638).

2. **PIZZI Claudio**

Curriculum:

Claudio Pizzi è professore associato in statistica economica presso la Facoltà di Economia dell'Università degli Studi di Venezia - Ca' Foscari dal 2004.

Ha svolto attività didattica principalmente nei corsi di statistica economica, controllo statistico della qualità e statistica per il mercato monetario e finanziario ed è membro dei collegi didattici del corso di laurea in Economia e Finanza e in quello di Statistica e Informatica per la Gestione delle Imprese.

I suoi principali interessi di ricerca riguardano lo studio delle serie storiche in particolare serie storiche nonlineari. I suoi contributi sono stati presentati a convegni e i suoi lavori dedicati principalmente alle serie storiche finanziarie sono stati pubblicati in riviste internazionali e nazionali.

È membro della Società italiana di Statistica e della Società Italiana di Economia Demografia e Statistica.

A partecipato a numerosi Progetti di Ricerca di Interesse Nazionale e nel 2004 ha coordinato un'unità locale di ricerca di un progetto PRIN.

Pubblicazioni:

◆ PIZZI C. (2007). *The asymmetric threshold model ASETAR(2,1,1)*. S.Co.2007 Complex models and computational intensive methods for estimation and prediction. 6-8 settembre 2007. (pp. 392-397). ISBN/ISSN: 978-88-6129-114-0. PADOVA: CLEUP.

◆ PIZZI C., PARPINEL F. (2006). *Nonparametric strategies applied to time series analysis*. SER2006. 18-19 Aprile. (pp. 203-206).

◆ GEROLIMETTO M, PIZZI C., E PROCIDANO I. (2005). *A bootstrap test for hidden cointegration with an application to real data*. Atti convegno Statistical Inference on Linear and Non-linear Dynamics in Time Series.

◆ PIZZI C., PROCIDANO I., RIGATTI LUCHINI S. (2003). *The wild bootstrap for unit root test in presence of outliers*. STATISTICA. vol. 3, pp. 581-588 ISSN: 0390-590X.

◆ PIZZI C., PELLIZZARI P. (1998). *Adaptive local linear model for financial time series*. JOURNAL OF COMPUTATIONAL INTELLIGENCE IN FINANCE. vol. 6, pp. 32-39.

◆ PIZZI C., DE NADAI G. (1996). *Modelli a Soglia vs Reti neurali artificiali nella previsione dei valori del mercato mobiliare Italiano*. QUADERNI DI STATISTICA E MATEMATICA APPLICATA ALLE SCIENZE ECONOMICO-SOCIALI. pp. 175-193.

◆ PELLIZZARI P, PIZZI C., E SALMASI L. (2005). *Cointegrazione dinamica tra serie storiche economiche*. In: PROVASI C. Modelli complessi e metodi computazionali intensivi per la stima e la previsione. (pp. 389-394). PADOVA: Cleup (ITALY).

◆ PIZZI C., PARPINEL F., LISI, F. (2001). *Combining Forecast for Nonlinear Testing*. In: PROVASI CORRADO EDS. Modelli Complessi e Metodi Computazionali Intensivi Per la Stima e la Previsione. (pp. 193-198). PADOVA: CEDAM (ITALY).

3. **SLANZI Debora**

Curriculum:

Debora Slanzi è assegnista di ricerca presso il Dipartimento di Statistica dell'Università Ca' Foscari di Venezia. Laureata nel 2002 in Scienze Statistiche ed Economiche presso l'Università degli Studi di Padova, nel 2006 ha conseguito il titolo di Dottore di Ricerca in Statistica presso l'Università degli Studi di

Padova. Ha trascorso un periodo di studio presso il dipartimento di Computer Science dell'Università di Aalborg (DK) in qualità di "visiting Ph.D. Student". Dal 2006 è assegnista di ricerca presso il Dipartimento di Statistica dell'Università Ca' Foscari nell'ambito del progetto europeo "Programmable Artificial Cell Evolution" (PACE). Ha svolto attività di supporto alla didattica principalmente in corsi di Statistica e Serie Storiche presso l'Università di Padova, Venezia e Pavia. È socio della Società Italiana di Statistica, dal 2003 al 2005 ha partecipato a progetti di Ateneo ex 60% presso l'Università di Padova, e nel 2006 ha collaborato ad un progetto PRIN. I suoi principali interessi di ricerca riguardano lo studio delle reti Bayesiane e dei modelli grafici, in particolare nell'ambito del disegno sperimentale e della biostatistica, e lo studio dei metodi di calcolo evolutivo.

Publicazioni:

- ◆ FORLIN M, DE MARCH D, SLANZI D., POLI I. (2007). A Predictive Evolutionary Approach to Design Biochemical Experiments.
- ◆ SLANZI D. (2007). A Bayesian network learning algorithm for complex domain. (pp. 463-468).
- ◆ SLANZI D., POLI I, DE MARCH D, FORLIN M. (2007). Bayesian Networks for Detecting Relevant Variable Interactions for Biochemical Experiments.
- ◆ SLANZI D., TREVISAN M, VILLANOVA L, PAL G, POLI I. (2007). Bayesian Analysis of Cytomegalovirus Infection from Microarray Data.
- ◆ BOLZAN M., BROGINI A., SLANZI D. (2005). I MODELLI GRAFICI ESPLORATIVI PER L'ANALISI DEL FABBISOGNO DI ASSISTENZA OSPEDALIERA INFORMALE. EFFICACIA ESTERNA DELLA FORMAZIONE UNIVERSITARIA: IL PROGETTO OUTCOMES. (pp. 375-388).
- ◆ BROGINI A., SLANZI D. (2005). UNSUPERVISED VS SUPERVISED LEARNING IN A REAL COMPLEX SYSTEM. MODELLI COMPLESSI E METODI COMPUTAZIONALI INTENSIVI PER LA STIMA E LA PREVISIONE. S.CO. 2005. (pp. 467-472).
- ◆ BOLZAN M., BROGINI A., SLANZI D. (2005). APPRENDIMENTO DI MODELLI GRAFICI ESPLORATIVI PER LA VALUTAZIONE IN AMBITO SOCIO-SANITARIO: IL CASO DELL'ASSISTENZA INFORMALE. NON PROFIT. vol. 1, pp. 207-224 ISSN: 1122-9322.
- ◆ BROGINI A, SLANZI D. (2007). L'uso di reti Bayesiane per l'analisi di dati statistici multivariati. In: Demografia e Statistica: un ricordo di Enzo Lombardo tra scienza e cultura.
- ◆ BROGINI A., BOLZAN M., SLANZI D. (2004). IDENTIFYING A BAYESIAN NETWORK FOR THE PROBLEM HOSPITAL AND FAMILIES. THE ANALYSIS OF PATIENT SATISFACTION WITH THEIR STAY IN HOSPITAL. In: DI BACCO M., D'AMORE G., SCALFARI F. APPLIED BAYESIAN STATISTICAL STUDIES IN BIOLOGY AND MEDICINE. (pp. 41-72).
- ◆ BROGINI A, SLANZI D. (2007). Several computational studies about variable selection for Bayesian networks.

Testo inglese

1. **PARPINEL Francesca**

Curriculum:

Francesca Parpinel is associate professor of Statistics at the University Ca' Foscari of Venice (www.unive.it). Graduated from the University of Padua in Statistical and Economic Science, she is Ph. Doctor in Statistics (University of Padua). She got a collaboration for study and research with the European Commission of Ispra, on "Diagnostics of non linear processes". She passed a period of study at the University Paris VI. She pursued her post-doctoral research at the Department of Statistical Sciences of the University of Padua. Researcher in Statistics at the Faculty of Economics by the University of Venice. Her major research interests have been in statistical inference for continuous stochastic processes driven by stochastic differential equations, in linear and non linear models for time series, in nonparametric inference, and more recently in quantile regression. In the context of Statistics education she cured several publications, for one of which she received the prize as the best work for the Statistics teaching by the Italian Statistical Society. She is member of the Italian Statistical Society. Currently she is a member of the Scientific Board of the Doctoral Program in Economics and Organization of the University Ca' Foscari. She joined to several research projects.

Publications:

- ◆ PARPINEL F., PIZZI C. (2002). Iterative estimation procedure for option pricing with stochastic volatility models. METRON. vol. LX, n.1-2, pp. 211-223 ISSN: 0026-1424.
- ◆ PARPINEL F., PROVASI C. (2004). Elementi di Probabilità e Statistica per le Scienze Economiche. ISBN: 88-348-4378-9. TORINO: GIAPPICHELLI (ITALY).
- ◆ FIORIN S, PARPINEL F. (2007). Simulation study for strong consistency in the weak topology of LS estimates for a functional linear model. S.Co.2007 Complex models and computational intensive methods for estimation and prediction. 6-8 Settembre 2007. (pp. 225-230). ISBN/ISSN: 978-88-6129-114-0. PADOVA: Cleup (ITALY).
- ◆ PARPINEL F., MESTRINER A. (2007). Quantile regression for Italian reference growth charts. S.Co.2007 Complex models and computational intensive methods for estimation and prediction. 6-8 Settembre 2007. (pp. 368-373). ISBN/ISSN: 978-88-6129-114-0. PADOVA: cleup (ITALY).
- ◆ PARPINEL F., PIZZI C. (2006). Nonparametric strategies applied to time series analysis. SER2006. 18-19 aprile 2006. (pp. 203-206).
- ◆ PARPINEL F. (2002). Testing unit roots in time series analysis through permutation methods. SIS, XLI riunione scientifica, Milano. 5-7 giugno 2002. (pp. 635-638).

2. **PIZZI Claudio**

Curriculum:

Claudio Pizzi is Associate Professor of Economic Statistics (SECS/S03) at University Ca' Foscari of Venice. He has done teaching activity mainly in the courses of "Economic Statistics", "Statistical quality control" in the Treviso unit of the University of Venice, where he is also responsible of the computer science section, and "Statistics for monetary and financial market" in Venice. He is member of the didactic committee of the undergraduate course of "Economics and Finance" and of "Statistics and computer science for management". His main research interests are the analysis of non linear time series. His contributions treat the problem of non linearity both from the point of view of the identification of the non linearity in time series and the definition of non parametric models based on the local approach. These models have been applied mostly to financial time series. His works have been presented in several conferences and his main papers have been published in international and national reviews. He is member of the Italian Statistical Society and of the Italian Society of Economics, Demography and Statistics. He has been researcher in several scientific research projects and in 2004 he has been the coordinator of a local unit of a PRIN project.

Publications:

- ◆ PIZZI C. (2007). *The asymmetric threshold model ASETAR(2,1,1)*. S.Co.2007 *Complex models and computational intensive methods for estimation and prediction*. 6-8 settembre 2007. (pp. 392-397). ISBN/ISSN: 978-88-6129-114-0. PADOVA: CLEUP.
- ◆ PIZZI C., PARPINEL F. (2006). *Nonparametric strategies applied to time series analysis*. SER2006. 18-19 Aprile. (pp. 203-206).
- ◆ GEROLIMETTO M, PIZZI C., E PROCIDANO I. (2005). *A bootstrap test for hidden cointegration with an application to real data*. Atti convegno *Statistical Inference on Linear and Non-linear Dynamics in Time Series*.
- ◆ PIZZI C., PROCIDANO I., RIGATTI LUCHINI S. (2003). *The wild bootstrap for unit root test in presence of outliers*. STATISTICA. vol. 3, pp. 581-588 ISSN: 0390-590X.
- ◆ PIZZI C., PELLIZZARI P. (1998). *Adaptive local linear model for financial time series*. JOURNAL OF COMPUTATIONAL INTELLIGENCE IN FINANCE. vol. 6, pp. 32-39.
- ◆ PIZZI C., DE NADAI G. (1996). *Modelli a Soglia vs Reti neurali artificiali nella previsione dei valori del mercato mobiliare Italiano*. QUADERNI DI STATISTICA E MATEMATICA APPLICATA ALLE SCIENZE ECONOMICO-SOCIALI. pp. 175-193.
- ◆ PELLIZZARI P, PIZZI C., E SALMASI L. (2005). *Cointegrazione dinamica tra serie storiche economiche*. In: PROVASI C. *Modelli complessi e metodi computazionali intensivi per la stima e la previsione*. (pp. 389-394). PADOVA: Cleup (ITALY).
- ◆ PIZZI C., PARPINEL F., LISI, F. (2001). *Combining Forecast for Nonlinear Testing*. In: PROVASI CORRADO EDS. *Modelli Complessi e Metodi Computazionali Intensivi Per la Stima e la Previsione*. (pp. 193-198). PADOVA: CEDAM (ITALY).

3. SLANZI Debora

Curriculum:

Debora Slanzi is research fellow at the Department of Statistics of the University Ca' Foscari of Venice. Graduated in 2002 in Statistics at the University of Padua, in 2006 she has got the Ph.D. in Statistics at the University of Padua. She was a visiting Ph.D. student at the Department of Computer Science of the Aalborg University. Since 2006 she is research fellow at the Department of Statistics in the European project "Programmable Artificial Cell Evolution" (PACE). She has done teaching activity (exercise lessons) mainly in the courses of Statistics and Time Series at the University of Padua, Venice and Pavia. She is a member of the Italian Statistical Society, from 2003 to 2005 she joined to several research projects. Her main research interests regard Bayesian networks and graphical models, in particular in the design of experiments and biostatistics, methods of evolutionary computation.

Publications:

- ◆ FORLIN M, DE MARCH D, SLANZI D., POLI I. (2007). *A Predictive Evolutionary Approach to Design Biochemical Experiments*.
- ◆ SLANZI D. (2007). *A Bayesian network learning algorithm for complex domain*. (pp. 463-468).
- ◆ SLANZI D., POLI I, DE MARCH D, FORLIN M. (2007). *Bayesian Networks for Detecting Relevant Variable Interactions for Biochemical Experiments*.
- ◆ SLANZI D., TREVISAN M, VILLANOVA L, PAL G, POLI I. (2007). *Bayesian Analysis of Cytomegalovirus Infection from Microarray Data*.
- ◆ BOLZAN M., BROGINI A., SLANZI D. (2005). *I MODELLI GRAFICI ESPLORATIVI PER L'ANALISI DEL FABBISOGNO DI ASSISTENZA OSPEDALIERA INFORMALE. EFFICACIA ESTERNA DELLA FORMAZIONE UNIVERSITARIA: IL PROGETTO OUTCOMES*. (pp. 375-388).
- ◆ BROGINI A., SLANZI D. (2005). *UNSUPERVISED VS SUPERVISED LEARNING IN A REAL COMPLEX SYSTEM. MODELLI COMPLESSI E METODI COMPUTAZIONALI INTENSIVI PER LA STIMA E LA PREVISIONE. S.CO. 2005*. (pp. 467-472).
- ◆ BOLZAN M., BROGINI A., SLANZI D. (2005). *APPRENDIMENTO DI MODELLI GRAFICI ESPLORATIVI PER LA VALUTAZIONE IN AMBITO SOCIO-SANITARIO: IL CASO DELL'ASSISTENZA INFORMALE. NON PROFIT. vol. 1, pp. 207-224 ISSN: 1122-9322*.
- ◆ BROGINI A, SLANZI D. (2007). *L'uso di reti Bayesiane per l'analisi di dati statistici multivariati*. In: *Demografia e Statistica: un ricordo di Enzo Lombardo tra scienza e cultura*.
- ◆ BROGINI A., BOLZAN M., SLANZI D. (2004). *IDENTIFYING A BAYESIAN NETWORK FOR THE PROBLEM HOSPITAL AND FAMILIES. THE ANALYSIS OF PATIENT SATISFACTION WITH THEIR STAY IN HOSPITAL*. In: DI BACCO M., D'AMORE G., SCALFARI F. *APPLIED BAYESIAN STATISTICAL STUDIES IN BIOLOGY AND MEDICINE*. (pp. 41-72).
- ◆ BROGINI A, SLANZI D. (2007). *Several computational studies about variable selection for Bayesian networks*.