

Bruno Scarpa

Dipartimento di Scienze Statistiche
Università di Padova
via Cesare Battisti, 241
35121 Padova

tel: 049.8274193
cell: 392.6703028
fax: 049.8274170
email: scarpa@stat.unipd.it
<http://homes.stat.unipd.it/bruno/>

DATI PERSONALI Nato a Venezia, il 29 gennaio 1969.
Residente in Cannaregio 2481, 30121 Venezia. Cittadinanza italiana.

INTERESSI DI RICERCA Modelli bayesiani, big data, data mining, biostatistica, spatiotemporal modeling, applicazioni di metodi statistici a marketing e business, analisi di dati funzionali, modelli nonparametrici frequentisti e bayesiani, modelli di distribuzioni flessibili

FORMAZIONE SCOLASTICA E UNIVERSITARIA **Università di Padova**, Padova
Dottorato di ricerca, Statistica, 8 luglio 1997

- Titolo della tesi: *Modelli per la concentrazione di inquinanti atmosferici combinando i dati di centraline fisse e mobili*
- Supervisore: Guido Masarotto

Università di Padova, Padova

Esame di stato, Discipline statistiche, 1993

Università di Padova, Padova

Laurea, Scienze Statistiche e Demografiche, 19 marzo 1993

- Facoltà: Scienze Statistiche, Demografiche ed Attuariali
- Titolo della tesi: *Confronto della distribuzione del tempo fino all'allarme di alcuni schemi di controllo*
- Relatore: Guido Masarotto
- Voto: 110/110

Liceo Scientifico "Francesco Severi", Venezia - Lido

Maturità scientifica, 1988

POSIZIONE ATTUALE **Dipartimento di Scienze Statistiche, Università di Padova**
Professore ordinario (SECS-S/01-Statistica)

apr. 2019 a oggi,

ESPERIENZA PROFESSIONALE **Dipartimento di Scienze Statistiche, Università di Padova**
Professore associato (SECS-S/01-Statistica)
Ricercatore confermato (SECS-S/01-Statistica),

nov. 2014 a mar. 2019,
gen. 2012 a ott. 2014

Facoltà di Scienze Statistiche, Università di Padova
Ricercatore confermato (SECS-S/01-Statistica),

set. 2006 a dic. 2011

Facoltà di Scienze Politiche, Università di Pavia
Ricercatore confermato (SECS-S/01-Statistica),

set. 2002 ad ago. 2006

H3G s.p.a. - Direzione CRM, Milano

Customer Intelligence Manager,
Coordinamento e guida del gruppo di analisi sulla Customer Base e impostazione delle strutture di dati necessarie per le analisi.

set. 2001 ad ago. 2002

TIN.IT s.p.a. (gruppo TELECOM Italia) – Direzione Marketing, Roma

Marketing Manager CRM&Profiling apr. 2000 ad ago. 2001
Coordinamento e guida del gruppo di una decina di persone, incaricato di gestire le azioni sulla Customer Base e sui navigatori e di effettuare le analisi di marketing.

OMNITEL s.p.a. – Direzione Marketing – Analisi e Ricerche, Corsico - MI

Churn manager - Coordinatore team analisi di previsione del *churn* lug. 1998 ad apr. 2000
Gestione di analisi statistiche orientate alla previsione della domanda e del traffico generato dai diversi gruppi di clienti.

Coordinamento di un gruppo di una ventina di persone (analisti, informatici, product managers) per l'identificazione e predisposizione di azioni per ridurre la disattivazione dei clienti.

RAS–Riunione Adriatica di Sicurtà – Attuariato Rami Danni, Milano

Analista statistico feb. 1997 a lug. 1998
Creazione di nuove tariffe e controllo delle esistenti attraverso modellazione del rischio, analisi della concorrenza, monitoraggio del mercato, stima e previsione di Sinistri e Premi.

Ospedale Civile di Venezia, Venezia

Servizio civile lug. 1993 a lug. 1994
Obiettore di coscienza con il compito di effettuare analisi statistiche di dati clinici ed epidemiologici in collaborazione con medici e biologi.

Dipartimento di Scienze Statistiche, Università di Padova

Titolare di borsa di studio per il *Dottorato di Ricerca* nov. 1993 a nov. 1996

ATTIVITÀ
ISTITUZIONALE
ORGANIZZATIVA
E DI SERVIZIO

Dipartimento di Scienze Statistiche, Università di Padova

- *Presidente del Consiglio di Corso di Studi della Laurea Magistrale in “Scienze Statistiche”,* feb. 2016 a oggi
- *Coordinatore del Gruppo per l'Accreditamento e la Valutazione (GAV) della Laurea Magistrale in “Scienze Statistiche”,* feb. 2016 a oggi
- *Componente della Commissione Didattica,* feb. 2016 a oggi
- *Componente della Commissione Biblioteca,* dic. 2014 a oggi
- *Componente della Commissione Stage,* nov. 2013 a oggi
- *Rappresentante dei ricercatori nella Giunta del Dipartimento,* ott. 2007 a ott. 2010
- *Responsabile della Sezione per l'Elaborazione Automatica dei Dati (SEAD),* ott. 2007 a ott. 2010
- *Responsabile di Techné,* ott. 2007 a ott. 2010
Techné era la struttura del Dipartimento finalizzata alla progettazione e realizzazione di indagini CATI-CAWI, attraverso un *call center* supportato da applicativi informatici per la gestione delle indagini.
- *Membro del Collegio dei docenti della Scuola di Dottorato in Scienze Statistiche,* ott. 2007 a oggi

Facoltà di Scienze Statistiche, Università di Padova

- *Componente della Commissione Stage,* mag. 2010 a set. 2010
- *Responsabile delle Aule e Servizi Informatici per la Didattica (ASID),* ott. 2007 a ott. 2010
Tale attività comportava il coordinamento di 8 tecnici informatici e la gestione delle strutture informatiche della Facoltà, del Dipartimento e della Biblioteca di Scienze Statistiche, 2 aule attrezzate oltre a diverse unità in Biblioteca e nelle aule studio per un totale di un centinaio di macchine, server di posta elettronica, virtual server, domini, servizi di rete LAN e wireless e di infrastruttura, server di calcolo, assistenza ai docenti, supporto alla didattica e alla ricerca

Facoltà di Economia, Università del Sannio in Benevento

- Commissario per la valutazione comparativa per la copertura di un posto di ricercatore universitario
Settore scientifico disciplinare SECS-S/01 Statistica set. 2008 — nov. 2008

ISTAT

- Membro del ‘Big Data Committee’ — ISTAT gen. 2016 — oggi
- Membro della Commissione ISTAT su Big Data dic. 2012 — nov. 2015

ATTIVITÀ DIDATTICA DA RESPONSABILE	Dipartimento di Scienze Statistiche, Università di Padova Metodi Statistici per Big Data, da A.A.2016/17 a A.A.2018/19 <i>Corso di laurea triennale in Statistica per la Tecnologia e le Scienze</i> Statistica Iterazione, A.A.2015/16 <i>Corso di laurea magistrale in Scienze Statistiche</i> Analisi dei dati (data mining), da A.A.2008/09 a A.A.2018/19 <i>Corso di laurea magistrale in Scienze Statistiche</i> Componente del Collegio dei docenti del Dottorato in Statistica, da A.A.2013/14 a A.A.2018/19 Modelli di regressione e classificazione non-parametrica, <i>Dottorato in Statistica</i> A.A.2005/06 Dati, modelli e metodi di analisi della fisiologia e del comportamento riproduttivo, A.A.2005/06 <i>Dottorato in Statistica</i>
	Dipartimento di Matematica, Università di Padova Data mining, da A.A. 2013/14 a A.A.2015/16 <i>Corso di laurea magistrale in Informatica</i>
	Dipartimento di Farmacia, Università di Padova Gli strumenti statistici per l'analisi dei dati, da A.A. 2012/13 a A.A.2013/14 <i>Corso di Alta Formazione in ‘Logistica Farmaceutica’</i>
	Facoltà di Scienze Statistiche, Università di Padova Uno statistico in azienda, da A.A.2008/09 a A.A. 2009/10 <i>Corso di laurea triennale in Statistica e gestione delle imprese</i> Data mining [<i>in inglese</i>], da A.A.2008/09 a A.A.2016/17 <i>Corso di laurea magistrale in Scienze Statistiche</i> Metodi Statistici per le Applicazioni Aziendali, A.A.2007/08 <i>Corso di laurea magistrale in Scienze Statistiche</i> Statistica Descrittiva, da A.A.2006/07 a A.A. 2007/08 <i>Corso di laurea triennale in Statistica e gestione delle imprese</i> Tecniche Statistiche di Classificazione, A.A.2006/07 <i>Corso di laurea triennale in Statistica e tecnologie informatiche</i>
	Facoltà di Economia, Università Commerciale “L.Bocconi”, Milano Statistica per l'Analisi Organizzativa, A.A.2006/07 <i>Corso di laurea specialistica in Organizzazione e Sistemi Informativi</i>
	Facoltà di Sociologia, Università di Milano Bicocca

Statistica per la Ricerca Sociale - Modelli interpretativi I, A.A.2003/04
Corso di laurea in Sociologia

Facoltà di Scienze Politiche, Università di Pavia
Analisi di dati aziendali, da A.A.2003/04 a A.A.2004/05
Master in Ricerca su Società, Mercato e Territorio.

Statistica Economica, da A.A.2002/03 a A.A.2005/06
Corso di laurea triennale in Scienze Sociali per la Cooperazione e lo Sviluppo.

Statistica Economica, A.A.2002/03
Corso di laurea triennale in Scienze Politiche.

Statistica Economica, da A.A.2002/03 a A.A. 2003/04
corso di laurea in Scienze Politiche.

ATTIVITÀ
DIDATTICA
CICLI DI
LEZIONI

Dipartimento di Scienze Statistiche, Università di Padova
Statistica Iterazione, da A.A.2016/17 a A.A.2019/20
Corso di laurea magistrale in Informatica

Facoltà di Scienze Statistiche, Università di Padova
Statistica Laboratorio, da A.A.2007/08 a A.A. 2008/09
Corso di laurea triennale in Statistica e tecnologie informatiche

Metodi Statistici per le Applicazioni Aziendali, A.A.2004/05
Corso di laurea specialistica in Scienze Statistiche, Economiche, Finanziarie e Aziendali

Analisi dei dati (data mining), da A.A.2002/03 a A.A.2007/08
Corso di laurea specialistica in Scienze Statistiche

Facoltà di Scienze Politiche, Università di Pavia
Statistica, A.A.2002/03
corsi di laurea triennale in Scienze Politiche.

Facoltà di Economia, Università Cattolica del Sacro Cuore di Milano
Controllo Statistico della Qualità, da A.A.1997/98 a A.A.1998/99
corso di laurea in Economia.

Facoltà di Economia, Università di Padova
Statistica, A.A.1996/97
corsi di laurea in Economia.

ATTIVITÀ **Biostat2016, Scuola di Alta Formazione Statistica, Asti,** giu. 2016
DIDATTICA: Modulo di Statistical learning

SCUOLE ESTIVE **IASC2007, Statistical Learning, Data Mining and Regression Tools, Procida,** set. 2007
E CORSI BREVI Modulo di Data mining

Corsi per aziende private ed enti pubblici
Assicurazioni Generali, *Segmentation and customer profiling [in inglese]* gen. 2017
Assicurazioni Generali, *Statistical learning [in inglese]* dic. 2016
Eurostat, *Data mining [in inglese]* giu. 2015
Aspiag s.p.a. – Despar, *Regole di associazione* mar. 2015
Eurostat, *Data mining[in inglese]* nov. 2014
Valueteam s.p.a., *Statistical learning* dic. 2011
H3G s.p.a., *Data mining* feb. 2006

RELATORE **Master in Ricerca su Società, Mercato e territorio, Università di Pavia**
 TESI DI LAUREA Laura Ciaffardoni, Renata Virzintaite, Francesca Gugiatti, Maria Vittoria Palermiti, Antonella
 MAGISTRALE, Farina
 SPECIALISTICA **Laurea specialistica in Statistica e Informatica, Università di Padova**
 E MASTER Marco Ruffa, Angela Fraccalvieri, Giacomo Cassol, Giulio Cordioli, Estefanía Hernández Sánchez,
 Stefano Carpenedo, Andrea Pigato, Paolo Moras,
Laurea specialistica in Scienze Statistiche, Economiche, Finanziarie e Aziendali, Uni-
versità di Padova
 Elisa Zanato,
Laurea magistrale in Scienze Statistiche, Università di Padova
 Federico Dolcetto, Francesca Penzo, Taddeo Mauro, Gianluca Barbierato, Daniele Durante, Sab-
 rina Vettori, Lucia Zanotto, Simone Tiberi, Michel Comazzetto, Fabio Fogli, Angelica Zenere,
 Francesco Zulian, Sabrina Vettore, Marco Bono, Mara Conti, Evis Carapuli, Riccardo Grig-
 gio, Massimiliano Russo, Riccardo Corradin, Edoardo Marchi, Sally Paganin, Tommaso Rigon,
 Alex Giarolo, Annalisa Balata, Giulio Guglielmo, Emanuele Aliverti, Luca Martino, Stefano
 Dalla Costa, Alvise Zanardo, Simone Guidolin, Alessandro Aere, Fabio Barbaro, Elia Giacomini,
 Mauro Natali, Anna Zanin, Federico Carli, Jacopo Diquigiovanni, Edoardo Vignotto, Francesco
 Bizzotto, Silvia Roverato, Alberto Bianchi, Irene Sirocich, Valeria Zuccoli

RELATORE **Laurea triennale in Scienze Politiche, Università di Pavia**
 RELAZIONI Giovanni Lavizzari, Elisabetta Mori, Francesca Gugiatti, Victor Bosisio, Sandro Dolcini,
 FINALI **Laurea triennale in Scienze sociali per la Cooperazione e lo Sviluppo, Università di**
 TRIENNALI **Pavia** Fiorella Picchioni, Elena Scali,
Laurea triennale in Statistica e Tecnologie Informatiche, Università di Padova
 Simone Securo, Davide Anima, Luca Brugnaro, Michel Comazzetto, Yabre Yacouba, Emanuele
 Nardo, Roberta Mengato, Marco Masin, Umberto Zandolin,
Laurea triennale in Statistica e Gestione delle Imprese, Università di Padova
 Francesca Penzo, Fabio Cauci, Chiara Noventa,
Laurea triennale in Statistica Economia e Finanza, Università di Padova
 Daniele Durante, Angelo Basile, Pamela Dalla Costa, Alberto Bianchi,
Laurea triennale in Statistica Popolazione e Società, Università di Padova
 Taddeo Mauro, Licia Maron
Laurea triennale in Statistica per la Tecnologia e le Scienze, Università di Padova
 Martina Fumanelli, Laura Masiero, Simone Tosato, Federico Zanghi

RELATORE **Laurea in Scienze Politiche, Università di Pavia**
 TESI DI LAUREA Giacomo Ferrari, Elena Truscelli, Matteo Gabanelli, Cristina Alberti.
 QUADRIENNALE

SUPERVISIONE **Dottorato in Statistica, Università di Padova**
 DI TESI DI
 DOTTORATO

- Emanuele Aliverti (2016 – ad oggi)
- Sally Paganin (2015 – 2019)
 Titolo della tesi: Prior-driven Cluster Allocation in Bayesian Mixture Models
 Posizione attuale: Postdoc alla Berkley University
- Massimiliano Russo (2015 – 2019)
 Titolo della tesi: Bayesian Inference for Tensor Factorization Models
 Posizione attuale: Postdoc alla Harvard University
- Daniele Durante, Università di Padova (2013 – 2015)
 Titolo della tesi: Bayesian Nonparametric Modeling of Network Data
Vincitore del ‘Laplace Award’ e del ‘SBP Grand Data Challenge’,
 Posizione attuale: Assistant professor all’Università Bocconi, Milano
- Ronaldo Rouhver Guedes Silva, Università di Padova (2013 – 2015)
 Titolo della tesi: Nonparametric Models for Dependent Functional Data
 Posizione attuale: Assegno di ricerca all’Università di Verona
- Antonio Canale, Università di Padova (2009 – 2011)

Titolo della tesi: Bayesian Nonparametric Models for Count Data with Applications to Customer Base Management

Posizione attuale: Ricercatore RTD-B all'Università di Padova

- James Mbugua Ciera, Università di Padova (2007 – 2009)

Titolo della tesi: Approximate Bayes random effects models for large datasets

Posizione attuale: Senior data analyst, Twaweza East Africa, Nairobi, Kenya

PhD in “Economics and Techniques for the Conservation of the Architectural and Environmental Heritage”, University of Nova Gorica, Slovenia

- Marco Turlon (2013 – ad oggi)

RESPONSABILE
SCIENTIFICO
ASSEGNI DI
RICERCA
COORDINAMENTO
DI PROGETTI DI
RICERCA

- Daniele Durante, Università di Padova, 2016 – 2017
Bayesian nonparametric models for functional and multidimensional longitudinal data
- Coordinatore del Progetto di ricerca di Ateneo (Università di Padova) “Bayesian inference on brain network data” (Gen. 2016 — Dic. 2017)
- Coordina il Progetto di ricerca di Ateneo (Università di Padova) “Modelli Bayesiani non parametrici per l’analisi di dati funzionali: teoria ed applicazioni in biometria e marketing” (Gen. 2010 — Ago. 2012)
- Responsabile del Progetto “Modelli statistici per lo studio della biometria del ciclo mestruale” finanziato dall’Università di Padova (Gen. 2007 — Dic. 2008)
- Responsabile del Progetto “Biometria del ciclo mestruale” finanziato dall’Università di Pavia (Gen. 2005 — Dic. 2006)
- Responsabile del Progetto “DWH e Data mining per la telefonia di terza generazione. Requisiti utente” all’interno della Divisione CRM di H3G s.p.a., cui partecipavano informatici, analisti e statistici di marketing (Set. 2001 — Set. 2002)
- Responsabile del Progetto di ricerca “Segmentazione e Profiling degli utenti Internet” all’interno di tin.it, cui partecipavano analisti e statistici di marketing (Mag. 2000 — Ago. 2001)
- Responsabile del Gruppo di ricerca “Analisi e previsione del Churn” formato da esperti statistici, informatici e di marketing all’interno di Omnitel s.p.a. (Lug. 1998 — Apr. 2000)

PARTECIPAZIONE
A PROGETTI DI
RICERCA

- Membro del Progetto PRIN “Modern Bayesian nonparametric methods”, coordinato dal prof. I.Pruetner (Feb. 2017 — Feb. 2020)
- Membro del Progetto Marie Skłodowska-Curie ITN “Advanced Multi-Variate Analysis for New Physics Searches at the LHC”, coordinato dal prof. T. Dorigo (Set. 2015 — Ago. 2018)
- Membro del Progetto di Ateneo “Statistical and econometric approach to marketing: applications and developments to customer satisfaction and market segmentation” coordinato dalla prof. F.Bassi (mar 2013—set 2015)
- Membro del Progetto PRIN “New robust methods for the analysis of complex data”, coordinato dal prof. M.Riani (Mar. 2010 — Set. 2012)
- Membro del Progetto PRIN “Aspetti probabilistici e statistici nei recenti sviluppi della teoria delle distribuzioni” coordinato dal prof. A.Azzalini (Feb. 2007 — Mar. 2009)
- Membro del Progetto “Piani sperimentali adattivi per studi clinici” coordinato dal prof. A. Azzalini (Gen. 2007 — Dic. 2007)
- Partecipa al Progetto PRIN “Reti bayesiane causali:distribuzioni iniziali compatibili, identificabilità, aspetti dinamici, applicazioni al marketing ed al web mining” coordinato dal prof. G. Consonni (Nov. 2003 — Dic. 2005)

ORGANIZZAZIONE
DI CONVEGNI

- 2017 Membro del comitato di programma del Convegno intermedio della SIS 2017: “Statistics and Data Science: New Challenges, New Generations” Firenze 28-30 giugno 2017
- 2017 Membro del comitato scientifico di “Start-up research” evento satellite del Convegno intermedio SIS2017
Siena 27–28 giugno 2017.
- 2017 Membro del comitato organizzatore di “Stats under the stars 3” — Una notte per sfidarsi nella soluzione di un problema attraverso l’analisi dei dati
Firenze 27/28 giugno 2017.
- 2016 Membro della giuria di “Stats under the stars 2” — Una notte per sfidarsi nella soluzione di un problema attraverso l’analisi dei dati
Salerno 7/8 giugno 2016.
- 2015 Responsabile del comitato scientifico di “Statistica e data science per il business” riunione satellite del Convegno intermedio SIS2015
Padova 8 settembre 2015.
- 2015 Responsabile del comitato organizzatore di “Stats under the stars” — Una notte per sfidarsi nella soluzione di un problema attraverso l’analisi dei dati
Padova 7/8 settembre 2015.
- 2015 Membro del comitato organizzatore locale del Convegno intermedio della SIS 2015: “Statistica e Demografia: l’eredità di Corrado Gini”
Treviso 9-11 settembre 2015
- 2015 Membro del comitato organizzatore del Convegno: “Corrado Gini, professore di statistica a Padova”
Padova 7-8 settembre 2015
- 2010 Responsabile e presidente del comitato scientifico di “Statistica in azienda, statistici in azienda” riunione satellite della 45^a Riunione scientifica della Società Italiana di Statistica.
Padova 15 giugno 2010
- 2005 Membro del comitato scientifico dell’“International workshop on ongoing research on physiology and behaviour in human reproduction”.
Padova 9-13 dicembre 2005

DISCUSSIONE A
CONVEGNI

- Specialized Session: Statistical Analysis and Big Data
SIS 2014 - 47^a Riunione Scientifica della SIS - Cagliari Giu 2014
- Specialized Session: Latent variable models for marketing research
SIS 2013 - Statistica Conference - Brescia Giu 2013
- Specialized Session: Predictive Learning
SIS 2012 - 46^a Riunione Scientifica della SIS - Roma Giu 2012
- Sessione Specializzata: Predictive Learning
Giornate di studio sulla popolazione, VI edizione - Padova Feb 2005

CONTRO
RELATORE DI
TESI

Commissario per la discussione di tesi di dottorato

Davide De March e Luigi Marangi (Dottorato in Statistica, Università di Firenze)

Juan Francisco Rosco Nieves (Departamento de Matemáticas, Universidad de Extremadura, Spagna)

Paolo Zanini, Susanna Carcano, Mattia Penati, Anna Zanzottera (Dottorato in Modelli e Metodi Matematici per l’Ingegneria, Politecnico di Milano)

Marco Battiston, Silvia Missiroli, Ali Noori Khajavi, Sumeda Nilamani Siriwardena (PhD in Statistics, Università Bocconi)

Marco Bottone, Dario Briscolini, Matteo Serri (Dottorato in Modelli per l'Economia e la Finanza, Sapienza Università di Roma)

Controrelatore di un gran numero di tesi quadriennali, triennali, specialistiche e magistrali alle Università di Pavia e di Padova

RELAZIONI
INVITATE A
CONVEGNI

Relazioni invitate

- Lug. 2018 *Bayesian modelling of networks in business intelligence problems*, XXVII Congreso de Matemáticas Capricornio, Antofagasta, Chile.
- Ott. 2017 *Bayesian nonparametric location-scale-shape mixtures*, Flexible Statistical Models For A Skewed World of Data, Santiago de Chile, Chile.
- Giu. 2017 *Bayes y Big Data*, V Congreso Bayesiano de América Latina (Cobal V), Guanajuato, Messico.
- Set. 2016 *Bayesian non parametric modelling of Higgs pair production*, XIIth Quark Confinement and the Hadron Spectrum, Salonicco, Grecia.
- Oct. 2014 *Bayesian non-parametric model to longitudinally predict churn*, Conference of European Statistics Stakeholders, Roma, Italia.
- Lug. 2013 *Distribuciones a priori informativos para la distribución normal asimétrica*, Primer Simposio Centroamericano de Estadística Bayesiana, San Jose, Costa Rica.
- Mag. 2013 *Non-Gaussian Multivariate Statistical Models and their Applications*, Banff International Research Station for Mathematical Innovation and Discovery, Canada.
- Dic. 2012 *Enriched stick breaking processes for functional data*, 5th International Conference of the ERCIM on Computing and Statistics (ERCIM 2012), Oviedo, Spagna.
- Lug. 2012 *Fitting age-specific fertility rates by a flexible generalized skew-normal probability density function*, Third Palestinian Conference on Modern Trends in Mathematics and Physics, Hebron, Palestina.
- Set. 2011 *Fitting age-specific fertility rates by a skew-symmetric probability density function*, SCo2011, Padova, Italia.
- Set. 2009 *Mining massive data sets from Web*, SIS 2009 Statistical Methods for the analysis of large data-sets, Pescara, Italia.
- Set. 2009 *Fast Bayesian Functional Data Analysis: Application to Basal Body Temperature Data*, SCo2009, Milano, Italia.
- Mar. 2009 *Workshop on Skew Symmetric Distributions: return to ab-normality*, Benevento, Italia.
- Apr. 2008 *Workshop on Skew Symmetric Probability Distributions*, Invited Speaker, Bertinoro, Italia.
- Set. 2007 *Clustering Basal Body Temperature Curves Using Bayesian Functional Data Analysis with Contaminated Informative Priors*, SCo2007, Venezia, Italia.
- Set. 2007 *Structured models for classification and regression*, IASC2007 Statistical Learning, Data Mining and Regression Tools, Procida (NA), Italia.
- Lug. 2006 *Cervical mucus quality and the probability of conception: results from an Italian study*, Annual general meeting of the International Institute for Restorative Reproductive Medicine, Albuquerque - New Mexico (USA).
- Ott. 2005 *The Fertile Window from Studies of Conception*, European Conference of IEEF - EIFLE, Varsavia, Polonia.
- Apr. 2005 *La finestra fertile 'statistica'*, Seminario di Studi della Confederazione italiana dei centri per la regolazione naturale della fertilità, Verona.

- Feb. 2005 *Previsioni sequenziali delle lunghezze del ciclo mestruale*, Giornate di Studio sulla Popolazione, Padova.
- Giu. 2004 *Cervical mucus symptom and daily fecundability*, Statistics in Health Sciences, Nantes, Francia.
- Mag. 2004 *Predicting conception probabilities using daily mucus records*, Expanding Methodologies for Capturing Day-Specific Probabilities of Conception, Rockville, MD, USA.
- Gen. 2003 *The Fertile Window from Studies of Conception*, Inaugurazione XVIII ciclo del dottorato di Padova in ricordo di Antonio C. Capelo, Padova.
- Lug. 2002 *Customer Profiling, Segmentation and Marketing Strategies in Telecommunications*, International Conference on Current Advances and Trends in Nonparametric Statistics, Creta, Grecia.
- Giu. 2001 *I Sistemi di Business Intelligence*, Convegno dell'Institute of International Research - IIR, Milano.
- Set. 1999 *Customer Profiling, Segmentation and Marketing Strategies*, Datawarehousing & Datamining in Telecommunications (IBC Global Conferences), Londra, Regno Unito.
- Apr. 1998 *Analisi spazio-temporale di inquinanti atmosferici e posizionamento di centraline: un approccio non parametrico*, XXXIX Riunione Scientifica della SIS, Sorrento.
- Gen. 1996 *Costruzione di mappe delle concentrazioni di inquinanti atmosferici combinando i dati di centraline fisse e mobili*, L'Informazione statistica per il Governo dell'Ambiente, Roma.

CONTRIBUTI
A CONVEGNI

- Lug. 2011 *Enriched Stick Breaking Processes for Functional Data*, 8th Workshop on Bayesian Nonparametrics, Veracruz, Messico.
- Giu. 2010 *Customer base segmentation of a telecommunication company by using Bayesian semi-parametric analysis of traffic data*, XXXXV Riunione Scientifica della SIS, Padova, Italia.
- Set. 2006 *Optimal rules for timing intercourse to achieve pregnancy* Statistical Latent Variables Models in the Health Sciences, Perugia, Italia.
- Gen. 2005 *Sequential Predictions of Menstrual Cycle Lengths*, Second IMS-ISBA Joint Meeting - MCMSki, Poster Session, Bormio (SO).
- Apr. 1998 *Modelli additivi generalizzati con autocorrelazione tra le osservazioni*, XXXIX Riunione Scientifica della SIS, Sorrento.
- Apr. 1996 *Modelli per la concentrazione di inquinanti atmosferici combinando i dati di centraline fisse e mobili*, Riunione Scientifica dei dottorandi della XXXVIII Riunione Scientifica della SIS, Rimini.

SEMINARI
INVITATI

- 2 Ago. 2018 *Bayes y Big data*, Universidad de Atacama, Copiapó, Chile.
- 6 Ago. 2018 *Bayesian modelling of networks in business intelligence problems*, Pontificia Universidad Católica de Chile, Santiago, Chile.
- 25 Nov. 2016 *Bayesian modelling of networks in business intelligence problems*, Collegio Carlo Alberto, Università di Torino.
- 7 Nov. 2016 *Bayesian modelling of networks in business intelligence problems*, Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers, Brisbane, Australia.
- 25 Ott. 2016 *Enriched Stick Breaking Processes for Functional Data*, School of Mathematics and Statistics, University of Melbourne, Australia.
- 4 Feb. 2015 *Enriched Stick Breaking Processes for Functional Data*, Department of Biostatistics, MDAnderson Cancer Center, Huston (TX), USA.

- 10 Apr. 2013 *Modelli statistici e di datamining per la previsione del Churn*, Dipartimento di Economia, Università Roma Tre.
- 19 Mag. 2008 *Clustering Basal Body Temperature Curves Using Bayesian Functional Data Analysis with Contaminated Informative Priors*, Dipartimento di Scienze Statistiche, Università di Udine.
- 4 Mag. 2006 *Quando concentrare i rapporti sessuali? Metodi Bayesiani per la ricerca di regole ottime per concepire*, Dipartimento di Scienze Economiche e Statistiche, Università di Trieste.
- 24 Ott. 2005 *Modelli di data mining per la previsione del churn*, Istituto di Statistica, Università Cattolica del Sacro Cuore, Milano.
- 17 Mar. 2005 *Data mining: un'esperienza nel Marketing*, Dipartimento di Statistica, Università Ca' Foscari, Venezia.
- 27 Mag. 2004 *Un'esperienza aziendale di data mining*, Facoltà di Economia, Università del Sannio, Benevento.
- 19 Mag. 2004 *Cervical mucus symptom and daily fecundability*, EPI group - NIEHS - National Institute of Health, Research Triangle Park, NC, USA
- 29 Mag. 2003 *Modelli di data mining per la previsione del churn*, Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia.
- 25 Mar. 2003 *Customer Profiling, Segmentazione e Strategie di Marketing nelle Telecomunicazioni*, Facoltà di Scienze della Comunicazione e dell'Economia, Università di Modena e Reggio Emilia.
- 14 Nov. 2002 *Data Mining in Azienda*, Dipartimento di Scienze Statistiche, Università di Padova.
- 28 Mar. 2001 *Data Mining nel Marketing: l'esperienza in una società di internet*, Master in Net Economy, Università di Trieste.

SOGGIORNI
STUDIO E DI
RICERCA

- Facultad de Matemática, Pontificia Universidad Católica de Chile, Santiago, Chile, collaborazione con il Prof R. Arellano Valle. Ago. 2018
- Departamento de Matemática, Universidad de Atacama, Copiapo, Chile, collaborazione con il Prof D. Gallardo. Lug. 2018
- School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, collaborazione con la Prof.ssa K. Mengersen. Nov. 2016
- Department of Mathematics and Statistics, University of Melbourne, Melbourne, Australia, collaborazione con la Prof.ssa A. Delaigle. Ott. 2016
- Department of Biostatistics, MDAnderson Cancer Center, Houston (TX) USA, collaborazione con il Prof. M. Guindani. Gen. 2015
- Department of Statistical Science, Duke University, Durham (NC) USA, collaborazione con il Prof. D. Dunson. Lug. 2011
- Department of Mathematics and Statistics, McMaster University, Hamilton, Canada, collaborazione con il Prof. Balakrishnan. Ago. 2009
- Department of Mathematics and Statistics, McMaster University, Hamilton, Canada, collaborazione con il Prof. Balakrishnan. Ago. 2008
- Biostatistical Branch, NIEHS-National Institute of Health, Research Triangle Park, NC, USA, collaborazione con il Prof. D. Dunson. Lug. 2006
- Biostatistical Branch, NIEHS-National Institute of Health, Research Triangle Park, NC, USA, collaborazione con il Prof. D. Dunson. Mag. 2004
- Biostatistical Branch, NIEHS-National Institute of Health, Research Triangle Park, NC, USA, collaborazione con il Prof. D. Dunson. Aug.2003—Set.2003

- Dept. of Mathematics and Statistics, Lancaster University, Lug.1996—Aug.1996
Inghilterra, collaborazione con i Proff. P. Diggle e J. Tawn.
- Department of Statistics, Stanford University, CA, USA, Dic.1994—Set.1995
collaborazione con il Prof. P. Switzer.

ATTIVITÀ EDITORIALE

- Associate Editor, *Frontiers in Public Health*
- Referee,
Bayesian Analysis, Biostatistics, Annals of the Institute of Statistical Mathematics, Applied Stochastic Models in Business and Industry, Environmetrics, Statistical Methods in Medical Research, Chilean Journal of Statistics, Communications in Statistics - Theory and Methods, Communications in Statistics - Simulation and Computation, Italian Journal of Applied Statistics, Journal of Statistical Planning and Inference, Biometrical Journal, Far East Journal of Theoretical Statistics, Statistical Methods and Applications, Journal of Multivariate Analysis, Test, Statistics and Probability Letters, Computational Statistics and Data analysis, Information Sciences, Journal of the Operational Research Society, Statistics in Medicine, Bioinformatics, Journal of Statistical Computation and Simulation
- Discussant, *Bayesian Analysis*

PREMI E RI- CONOSCIMENTI

- Vincitore della competizione “Statistica, Open Data e Società”, Comune di Milano e Università Cattolica, Milano 2013
- Premio per *Paper Celebrating the 25th Anniversary of Statistics in Medicine* per l’articolo: Scarpa B., Dunson D.B., (2007) Bayesian methods for searching for optimal rules for timing intercourse to achieve pregnancy, *Statistics in Medicine*, 26, pp. 1920–1936. 2007

TEMI DI RICERCA

Metodi statistici per l’analisi di problematiche aziendali

L’innovazione tecnologica permette oggi alle aziende di disporre di masse rilevanti di elementi di informazione che possono venire utilizzate per indirizzarne le scelte strategiche e operative. In particolare molte aziende di servizi (ad esempio banche, assicurazioni, aziende di telecomunicazioni) hanno a disposizione informazioni dettagliate sui singoli clienti e sul loro utilizzo dei prodotti e servizi offerti. La mia attività di ricerca negli anni in cui ho prestato servizio nel settore privato si è concentrata principalmente sull’analisi statistica dei dati sui clienti per indirizzare le scelte strategiche in una prima fase e sul coordinamento di gruppi di ricerca (e operativi) su questi temi in una seconda fase.

Tra le varie attività di analisi di dati quantitativi particolarmente rilevanti sono stati: (i) Analisi dei comportamenti e la relativa modellazione statistica della *Customer Base* mirate all’anticipazione di bisogni di utilizzo dei servizi di telefonia cellulare (in Omnitel). (ii) Definizione di indicatori di propensione al *churn*, e la predisposizione di modelli statistici per la loro previsione e individuazione dei motivi che portano alla disattivazione. (iii) *Profiling* e segmentazione della *Customer Base* e dei visitatori dei siti/portali per ottenere gruppi di utenti omogenei (in tin.it). (iv) Individuazione dei profili principali dei clienti rispetto all’utilizzo dei servizi di internet. (v) Determinazione di azioni di marketing che l’azienda può effettuare per aumentare la fedeltà dei clienti e di azioni di *up-sell* e *cross-sell*. (vi) Supporto quantitativo agli altri settori di marketing per le previsioni (per *economics*, *budget*, *Business Plan* e altro). (vii) (in RAS) Analisi del rischio assicurativo nei rami danni.

Tale attività, nonostante le difficoltà a rendere pubblici i risultati di ricerca, mi ha permesso di produrre i lavori (73) e (74) per quanto riguarda le analisi di marketing e i lavori (85) e (30) nell’ambito della stima del rischio assicurativo. Il lavoro (72) è una rilettura a posteriori e sistematica con qualche esempio reale e concreto dell’attività di uno statistico nel marketing di una azienda di telecomunicazioni, mentre nel lavoro (53) si presentano le principali problematiche statistiche che deve affrontare un’azienda quando si accinge a prevedere la disattivazione della clientela. I lavori (80) e (83) descrivono alcune particolari

specificazioni e applicazioni di modelli statistici per rispondere a problemi originati in contesti aziendali per i quali è necessario l'utilizzo di strumenti inferenziali. Il lavoro (17) utilizza un modello bayesiano non parametrico per segmentare la clientela di una società di telecomunicazioni adottando un approccio funzionale che considera le osservazioni mensili del traffico telefonico come una curva nel tempo.

Data mining, big data e statistica

Nel mondo scientifico è ormai usuale avere a disposizione ingenti quantità di dati sia in termini di unità statistiche sia in termini di variabili disponibili (la rilevazione automatica di molte informazioni e il loro immagazzinamento permette la costituzione di grosse basi di dati utilizzabili ai fini di ricerca). Tale disponibilità richiede allo statistico di individuare strumenti innovativi ma efficaci per il trattamento dei dati. Un problema particolare affrontato riguarda la classificazione in classi rare dal quale sono scaturiti i lavori (64) e (71). Il lavoro (76) invece propone una modifica dei Modelli Additivi Generalizzati che tiene conto della presenza di autocorrelazione tra le osservazioni, problematica spesso presente in questo tipo di dati. I lavori (81) e (82) affrontano un paio di problemi concreti nati in contesti aziendali in cui è necessario analizzare ingenti quantità di dati e per i quali è necessario utilizzare specifici strumenti di analisi e di *data mining*. L'analisi di grandi quantità di dati reperibili utilizzando la rete internet viene affrontato nel lavoro (28).

Il lavoro (20) affronta un problema di analisi di grandi quantità di dati legate da una relazione temporale attraverso l'utilizzo di strumenti bayesiani non-parametrici.

L'utilizzo di specifici strumenti bayesiani non parametrici per l'analisi di reti viene proposto nel lavoro (12).

L'ampia esperienza di analisi di grandi basi di dati attraverso strumenti di *data mining* mi ha spinto a produrre la monografia (3) in cui si presentano, a partire da problemi reali, i principali strumenti di analisi dei dati e *data mining* letti e interpretati dal punto di vista dello statistico. Il lavoro (1) è la seconda edizione dello stesso libro, in inglese e molto ampliata sia nei contenuti, con l'aggiunta di alcune tematiche recentemente sviluppate nel contesto dell'analisi dei *big data*, sia nelle esemplificazioni pratiche, con l'aggiunta di un certo numero di studi di caso relativi a diversi contesti aziendali e di *business*. La voce di enciclopedia (59) propone una rassegna delle principali tematiche di ricerca legate al *data mining*.

Modelli bayesiani non parametrici per dati funzionali

In molte applicazioni in cui sono presenti dati funzionali sono spesso disponibili importanti informazioni a priori sulla forma delle curve e sulla proporzione di curve con specifiche caratteristiche. Non essendo automatico includere queste informazioni nelle procedure usuali per l'analisi di dati funzionali i lavori (19) e (29) propongono due generalizzazioni del processo funzionale di Dirichlet per includere informazioni a priori nel modello. Anche il lavoro (17), in un contesto applicativo completamente diverso, inserisce informazione a priori nell'analisi di dati funzionali.

Il lavoro (13) propone una classe di modelli mistura basati su indici di forma oltre che di posizione e scala in un contesto bayesiano non parametrico.

Il lavoro (20) affronta un problema di serie storiche multivariate come dati funzionali, attraverso l'utilizzo di strumenti bayesiani non-parametrici.

Il problema della modellazione di dati provenienti dal Large Hadron Collider per la produzione di coppie di bosoni di Higgs viene affrontato con strumenti bayesiani non parametrici nel lavoro (56).

Modelli basati su distribuzioni flessibili

La distribuzione normale asimmetrica e le sue generalizzazioni sono ormai largamente studiate ed utilizzate sia per la loro buona capacità di adattamento ai dati che presentano leggere violazioni degli assunti della distribuzione normale, sia per le interessanti rappresentazioni stocastiche che ne forniscono utili caratterizzazioni, sia perché le loro proprietà statistiche le rendono teoricamente molto simile alla distribuzione gaussiana, e quindi facili da analizzare.

Il lavoro (26) propone un nuovo metodo di stima semiparametrico per le distribuzioni *skew-symmetric* basato su equazioni di stima.

I lavori (21) e (23) affrontano il problema della verifica della quantità e del tipo di asimmetria presente nelle distribuzioni normali asimmetriche multivariate, nel primo caso at-

traverso l'analisi di diversi indici di asimmetria e nel secondo caso proponendo un test di skew-normality multivariata.

Il lavoro (13) propone una distribuzione a priori informativa per i tre parametri in caso di verosimiglianza *skew-normal*. Il lavoro (14) utilizza la distribuzione *skew-normal* come elementi per una mistura usata in un contesto bayesiano non parametrico.

Il lavoro (18) utilizza la distribuzione *skew-normal* come elemento che porta flessibilità alla stima della curva di fecondità, mentre in (8) la distribuzione normale asimmetrica è usata per modellare la funzione di mortalità per età in una popolazione.

Un problema di analisi di dati circolari relativo alla direzione dei venti viene affrontato nel lavoro (24) utilizzando la distribuzione skew-normal generalizzata circolare.

Modelli statistici per l'analisi di indicatori biometrici del ciclo mestruale

Il ciclo mestruale è una delle principali funzioni biologiche che regola la riproduzione umana; l'utilizzo di tecnologie sempre più avanzate in campo medico ha permesso, negli ultimi cinquant'anni di studiare e di capire sempre meglio i complessi sistemi biologici del ciclo riproduttivo. L'analisi delle grandezze legate a tale ciclo ha poi molto spesso dei risvolti pratici nel comportamento delle persone per quanto riguarda l'attività sessuale e la vita riproduttiva. Le problematiche di modellazione statistica sono molto complesse in particolare considerando le informazioni rilevabili attraverso strumenti non invasivi e relativamente economici. La voce di enciclopedia (61) presenta una introduzione ai principali modelli statistici usati per studiare i concepimenti e la fecondabilità.

Nell'ambito di ricerche in questo campo ho affrontato alcune tematiche in particolare legate alla stima della probabilità di concepimento in un ciclo quando si hanno a disposizione informazioni ridotte. Il lavoro (31) affronta tale problema attraverso una modellazione frequentista, mentre nei lavori (55), (56) e (70) il tema è affrontato con un approccio bayesiano.

I lavori (29), (54) e (84) affrontano attraverso un approccio decisionale bayesiano il problema di identificazione da parte delle coppie di una decisione che massimizzi la loro utilità nella scelta dei giorni in cui avere rapporti sessuali o dei comportamenti da assumere in termini di variabili epidemiologiche (fumo, alcool, caffè, ecc.). Il lavoro (35) analizza con un approccio frequentista un problema analogo di identificazione della finestra fertile a partire da poche informazioni di calendario. Il confronto di alcune tra le più usate regole per l'identificazione dei giorni fertili nel ciclo è anche affrontato con un approccio frequentista e uno stile divulgativo nel lavoro (57).

Nei lavori (19) e (27) si propongono due diversi modelli bayesiani non-parametrici per analizzare l'andamento della temperatura basale durante il ciclo affrontata come esempio di realizzazione di funzioni in cui sono disponibili rilevanti informazioni a priori, mentre il lavoro (11) utilizza un altro modello bayesiano non parametrico per descrivere l'andamento del tipo di muco. Anche il lavoro (15) affronta l'analisi del tipo di muco, anche se con un approccio frequentista basato sui modelli a classi latenti.

Il lavoro (25) affronta la previsione della lunghezza del ciclo e della sua fase preovulatoria utilizzando un modello *state space* stimato con un approccio bayesiano.

I lavori (10) e (39) affrontano (il primo con approccio frequentista basato sulla verosimiglianza composita e il secondo con un modello bayesiano) il problema di prevedere il sesso dei nascituri a partire dalla collocazione temporale dei rapporti all'interno del ciclo.

Analisi spazio-temporale di dati ambientali

La concentrazione di sostanze chimiche presenti nell'atmosfera costituisce uno degli indicatori principali per la segnalazione di situazioni di inquinamento. L'analisi di tali dati è quindi un tema rilevante per il controllo dell'atmosfera. Nella tesi di dottorato, (4), mi sono concentrato principalmente nella identificazione e nella metodologia di stima di alcuni modelli per la previsione della concentrazione di sostanze inquinanti. A partire da tale lavoro ho trattato alcuni aspetti particolari dell'analisi quali: l'identificazione di mappe spaziali nel lavoro (34), l'utilizzo di centraline mobili e una modellazione completamente stocastica nel lavoro (33), il posizionamento di nuove centraline attraverso un approccio non parametrico nei lavori (75) e (86) e l'effetto della presenza di dati mancanti utilizzando un approccio di stima non parametrico nel lavoro (32).

Applicazioni Biomediche e Demografiche

Le problematiche di carattere epidemiologico e medico sono un campo dove l'applicazione di

modelli statistici ai dati osservati è necessaria per comprendere e interpretare i fenomeni. Il lavoro (57) è una analisi sull'effetto dell'utilizzo della classificazione delle malattie attraverso i DRG in un ospedale e il lavoro (69) consiste in un'analisi sul rischio di nascite con malformazioni congenite, analizzato attraverso modelli non parametrici.

I lavori (38), (40), (43), (47), (48), (51) e (52) affrontano problemi di neurochirurgia in cui alcuni semplici modelli statistici vengono usati per analizzare i dati da elettroencefalogramma e di potenziale evocato somatosensitivo. I lavori (36), (37), (41), (42), (44), (45), (46), (49), (50) e (45) affrontano dei problemi di rinologia usando modelli statistici.

Nel lavoro (18) si propone un nuovo modello demografico per la stima della curva di fecondità che sia sufficientemente flessibile per adattarsi alle diverse forme che tale curva presenta nelle varie nazioni del mondo. Lo stesso problema viene affrontato attraverso un approccio bayesiano non parametrico nel lavoro (16).

Il lavoro (58) presentano tra l'altro una rassegna dei principali modelli demografici e statistici usati per la previsione della fecondità e dei concepimenti.

Applicazioni Didattiche e Divulgative

Il lavoro (2) ha scopo prevalentemente didattico. Intende essere un testo divulgativo per la comprensione da parte di studenti, con alle spalle un solo esame di statistica di base, dei concetti fondamentali del modello lineare e delle sue più semplici modificazioni. Nel lavoro (62) vengono presentati a scopo divulgativo, indirizzato a ostetriche e operatori di consultori familiari, i principali risultati scientifici sulla dimensione e posizionamento della finestra fertile nel ciclo riproduttivo femminile. Il lavoro (63) è una presentazione, a scopo illustrativo, dell'utilizzo di modelli statistici complessi per la valutazione di programmi sociali quando la modellazione più semplice non è sufficiente a cogliere la complessità della realtà. Il lavoro (9) descrive in maniera divulgativa l'evento *Startup Research* presentandone le potenzialità nella formazione dei giovani statistici.

PUBBLICAZIONI Monografie

1. Azzalini A., **Scarpa B.** (2012). *Data Analysis and Data Mining. An Introduction.* Oxford University Press, New York, ISBN 978-0-19-976710-6.
2. **Scarpa B.** (2005). *Il modello lineare: una presentazione a partire da dati reali.* Arvan Editore, ISBN 978-8887801156.
3. Azzalini A., **Scarpa B.** (2004). *Analisi dei dati e data mining.* Springer, Milano, ISBN 88-470-0272-9.
4. **Scarpa B.** (1997). *Modelli per la concentrazione di inquinanti atmosferici combinando i dati di centraline fisse e mobili.* Tesi di dottorato, Università di Padova.
5. **Scarpa B.** (1993). *Confronto della distribuzione del tempo fino all'allarme di alcuni schemi di controllo.* Tesi di laurea, Università di Padova. Arvan, Mira-Venezia, ISBN 88-87801-15-0.

Articoli su rivista di Statistica

6. Russo M., Durante D., **Scarpa B.** (2018) Bayesian inference on group differences in multivariate categorical data. *Computational Statistics and Data Analysis*, **126**, pp. 136–149. DOI: 10.1016/j.csda.2018.04.010.
7. Diquigiovanni J., **Scarpa B.** (2018) Analysis of association football playing styles: an innovative method to cluster networks. *Statistical Modelling: An International Journal, accettato per la pubblicazione*, <https://arxiv.org/abs/1805.10933>.
8. Mazzuco S., **Scarpa B.**, Zanotto L. (2018), A mortality model based on a mixture distribution function. *Population studies*, **72**, pp.191-200. DOI: 10.1080/00324728.2018.1439519.
9. Canale A., Durante D., Paci L., **Scarpa B.** (2018), Connecting statistical brains. *Significance*, **15**, pp. 38-40. DOI: 10.1111/j.1740-9713.2018.01110.x

10. Tiberi S., Sartori N., **Scarpa B.** (2018) A composite likelihood approach to predict the sex of the baby. *Statistical Methods in Medical Research*, **27**, p. 3386-3396.
DOI: 10.1177/0962280217702415.
11. De Bin R., **Scarpa B.** (2017) Non-parametric Bayesian modelling of cervical mucus symptom. *Journal of the Royal Statistical Society (Series C)*, in corso di pubblicazione
12. Durante, D., Paganin, S., **Scarpa, B.**, and Dunson, D.B. (2017) Bayesian modeling of networks in complex business intelligence problems. *Journal of the Royal Statistical Society (Series C)*. (DOI: 10.1111/rssc.12168), **66**, pp. 555-580.
13. Canale A., **Scarpa B.** (2016) Bayesian nonparametric location-scale-shape mixtures. *Test*, DOI: 10.1007/s11749-015-0446-2, **25**, pp. 113-130.
14. Canale A., Kenne Pagui E.C., **Scarpa B.** (2016). Bayesian modeling of university first-year students' grades after placement test. *Journal of Applied Statistics*; DOI: 10.1080/02664763.2016.1157144, **43**, pp. 3015-3029.
15. Bassi F., **Scarpa B.** (2015) Latent class modeling of markers of day-specific fertility. *Metron*, DOI: 10.1007/s40300-015-0066-3, **73**, pp. 263-276.
16. Canale A., **Scarpa B.** (2015) Age Specific Probability of Childbirth. Smoothing via Bayesian nonparametric mixture of rounded kernels. *Statistica*, DOI: 10.6092/issn.1973-2201/5826 , **75**, pp. 101-110.
17. **Scarpa B.** (2015). Bayesian semiparametric customer base segmentation of mobile phone users based on longitudinal traffic data. *Applied Stochastic Models in Business and Industry*; DOI: 10.1002/asmb.2085, **31**, pp. 721-731.
18. Mazzuco S., **Scarpa B.** (2015). Fitting age-specific fertility rates by a flexible generalized skew-normal probability density function. *Journal of the Royal Statistical Society. Series A, Statistics in Society*; DOI: 10.1111/rssa.12053, **178**, pp. 187-203.
19. **Scarpa B.**, Dunson D.B. (2014). Enriched Stick Breaking Processes for Functional Data. *Journal of the American Statistical Association*, DOI: 10.1080/01621459.2013.866564, **109**, pp. 647-660.
20. Durante D., **Scarpa B.**, Dunson D. B. (2014). Locally adaptive factor processes for multivariate time series. *Journal of Machine Learning Research*, **15**, pp. 1493-1522.
21. Balakrishnan N., Capitanio A., **Scarpa B.** (2014). A test for multivariate skew-normality based on its canonical form. *Journal of Multivariate Analysis*, DOI: 10.1016/j.jmva.2014.02.015, **128**, pp. 19-32.
22. **Scarpa B.** (2012). Comment on Flexible Bayesian Human Fecundity Models. *Bayesian Analysis*, **7**, pp. 801-804.
23. Balakrishnan N., **Scarpa B.** (2012). Multivariate Measures of Skewness for the Skew-Normal distribution. *Journal of Multivariate Analysis*, DOI: 10.1016/j.jmva.2011.06.017, **104**, pp. 73-87.
24. Hernández-Sánchez E., **Scarpa B.** (2012). A wrapped flexible generalized skew-normal model for a bimodal circular distribution of wind directions. *Chilean Journal of Statistics*, **3**, pp. 129-141.
25. Bortot P., Masarotto G., **Scarpa B.** (2011) Sequential Predictions of Menstrual Cycle Lengths, *Biostatistics*, **11**, pp. 741-755.
26. Azzalini A., Genton M.G., **Scarpa B.** (2010) Invariance-based estimating equations for skew-symmetric distributions, *Metron*, DOI: 10.1007/BF03263540, **LXVIII**, pp. 353-377.
27. **Scarpa B.**, Dunson D.B., (2009) Bayesian hierarchical functional data analysis via contaminated informative priors, *Biometrics*, DOI: 10.1111/j.1541-0420.2008.01163.x, **65**, pp. 772-780.
28. **Scarpa B.** (2009) Mining massive Data Sets from web, *Statistica applicata*, **21**, pp. 1125-1964.

29. **Scarpa B.**, Dunson D.B., (2007) Bayesian methods for searching for optimal rules for timing intercourse to achieve pregnancy, *Statistics in Medicine*, DOI: 10.1002/sim.2846, **26**, pp. 1920—1936.
Paper Celebrating the 25th Anniversary of Statistics in Medicine
30. **Scarpa B.** (2006), La tariffa incendio e furto per le auto: un modello di Cox, *Statistica*, **LXV**, pp. 437—446.
31. Colombo B., Mion A., Passarin K., **Scarpa B.** (2006) Cervical mucus symptom and daily fecundability: first results from a new data base. *Statistical Methods in Medical Research*, DOI: 10.1191/0962280206sm437oa, **15**, pp. 161—180.
32. **Scarpa B.** (2005). Non-parametric space-time modeling of SO₂ in presence of many missing data, *Statistical Methods and Applications-JISS*, DOI: 10.1007/BF02511575, **14**, pp. 67—82.
33. **Scarpa B.** (2001). Analisi della concentrazione di SO₂ combinando i dati raccolti da centraline fisse e mobili: un modello state space, *Statistica*, **LXI**, pp. 707—730.
34. Capelo A. C., Masarotto G. e **Scarpa B.** (1996). Costruzione di mappe delle concentrazioni di inquinanti atmosferici combinando i dati di centraline fisse e mobili, *Annali di Statistica*, Serie **X**, vol.10, pp. 257—265.
35. Colombo B. e **Scarpa B.** (1996). Calendar methods of fertility regulation: a rule of thumb, *Statistica*, **LVI**, pp. 3—14.

Articoli su riviste non di statistica

34. Del Felice A., Castiglia L., Formaggio E., Cattelan M., **Scarpa B.**, Manganotti P., Tenconi E., Masiero S. (2019) Personalized transcranial alternating current stimulation (tACS) and physical therapy to treat motor and cognitive symptoms in Parkinson's disease: A randomized cross-over trial. *Neuroimage: clinical*, **22**: 101768, DOI: 10.1016/j.nicl.2019.101768
35. Ottaviano G., Savietto E., **Scarpa B.**, Bertocco A., Maculan P., Sergi G., Martini A., Mankato E., Marioni G. (2018), Influence of number of drugs on olfaction in the elderly, *Rhinology*, **56** p.351-357, DOI: 0.4193/Rhin17.152.
36. Cantone E., Ciofalo A., Vodicka J., Iacono V., Mylonakis I, **Scarpa B.**, Russo M., Iengo M., de Vincentiis M., Martini A., Ottaviano G. (2017) Pleasantness of olfactory and trigeminal stimulants in different Italian regions. *European Archives of Oto-Rhino-Laryngology*, **11**, pp. 3907–3913. DOI:10.1007/s00405-017-4722-5.
37. Pendolino A.L., Nardello E., Lund V.J., Maculan P., **Scarpa B.**, Martini A., Ottaviano G. (2018) Comparison between unilateral PNIF and rhinomanometry in the evaluation of nasal cycle. *Rhinology*, **56**, p. 122-126, DOI: 10.4193/Rhin17.168.
38. Del Felice A. , Bargellesi S., Linassi F., **Scarpa B.**, Formaggio E., Boldrini P., Masiero S., Zanatta P. (2017) The potential role of pain-related SSEPs in the early prognostication of long-term functional outcome in post-anoxic coma. *European Journal of Physical and Rehabilitation Medicine*, **53**, pp.883-891. DOI: 10.23736/S1973-9087.17.04303-9.
39. **Scarpa B.** (2016) Bayesian Inference on Predictors of Sex of the Baby. *Frontiers in Public Health*, DOI: 10.3389/fpubh.2016.00102, **4**:102.
40. Feletti A., Anglani M., **Scarpa B.**, Schiavi F., Boaretto F, Zovato S., Taschin E., Gardi M., Zanoletti E., Piermarocchi S., Murgia A., Pavesi G., Opocher G. (2016) Von Hippel-Lindau disease: an evaluation of natural history and functional disability. *Neuro-Oncology*, DOI: 10.1093/neuonc/nov313, **18**, pp. 1011-1020.
41. Ottaviano G., Scadding G.K., Iacono V., **Scarpa B.**, Martini A., Lund V.J: (2016), Peak nasal inspiratory flow and peak expiratory flow. Upright and sitting values in an adult population. *Rhinology*, DOI: 10.4193/Rhin15.180, **54**, pp. 160-163.
42. Ottaviano G., Cantone E., D'Errico A., Salvalaggio A., Citton V, **Scarpa B.**, Favaro A., Sinisi A.A., Liuzzi R., Bonanni G., Di Salle F., Elefante A., Manara R., Staffieri A., Martini A., Brunetti A. (2015). Sniffin' Sticks and olfactory system imaging in patients with Kallmann syndrome *International Forum of Allergy & Rhinology*; DOI: 10.1002/alr.21550, **5**, pp. 855-861.

43. Zanatta P., Linassi F., Mazzarolo A.P., Aricò M., Bosco E., Bendini M., Sorbara C., Ori C., Carron M., **Scarpa B.** (2015) Pain-related Somato Sensory Evoked Potentials: a potential new tool to improve the prognostic prediction of coma after cardiac arrest. *Critical Care*, DOI: 10.1186/s13054-015-1119-y, **19**:403.
44. Ottaviano G., Lund V.J., Nardello E., **Scarpa B.**, Frasson G., Staffieri A., Scadding G.K. (2014). Comparison between unilateral PNIF and rhinametry in healthy and obstructed noses. *Rhinology*, **52**, pp. 25-30.
45. Ottaviano G., Lund V.J., Nardello E., **Scarpa B.**, Mylonakis I, Frasson G, Iacono V, Manzato E, Marioni G, Staffieri A. (2014). Peak nasal inspiratory flow: a useful and handy tool for the diagnosis of nasal obstruction in the elderly. *European Archives of Oto-Rhino-Laryngology*, **1/2014**.
46. Ottaviano G., Cosmi E., Iacono V., **Scarpa B.**, Staffieri A., Scadding G.K. (2014). Does the contraceptive pill influence peak nasal inspiratory flow values? *Rhinology*, DOI: 10.4193/Rhin, **52**, pp. 355-359.
47. Bosco E., Zanatta P., Ponzin D., Marton E., Feletti A., **Scarpa B.**, Longatti P., Paolin A. (2014). Prognostic value of somatosensory-evoked potentials and CT scan evaluation in acute traumatic brain injury. *Journal of neurosurgical anesthesiology* , DOI: 10.1097/ANA.000000000000040, **26**, pp. 299-305.
48. Bosco E., Zanatta P., Marton E., Feletti A., **Scarpa B.**, Longatti P., Sorbara C. (2013). Prognostic value of SEPs in acute phase of traumatic brain injury patients: Comprehensive analysis of early cortical components. *Clinical Neurophysiology*, DOI: 10.1016/j.clinph.2013.06.111, **124**:e208.
49. Ottaviano G., Zuccarello D., Frasson G., **Scarpa B.**, Nardello E., Foresta C., Marioni G., Staffieri A. (2013). Olfactory sensitivity and sexual desire in young adult and elderly men: An introductory investigation. *American Journal of Rhinology and Allergy*, DOI: 10.2500/ajra.2013.27.3879, **27**, pp. 157-161.
50. Ottaviano G., Scadding G.K., **Scarpa B.**, Accordi D., Staffieri A., Lund V.J.. (2012). Unilateral Peak Nasal Inspiratory flow. Normal values in adult population, *Rhinology*, DOI: 10.4193/rhin12.071, **50**, pp. 386-392.
51. Basaldella L., Marton E., Fiorindi A., **Scarpa B.**, Badreddine H., Longatti P. (2012). External ventricular drainage alone versus endoscopic surgery for severe intraventricular hemorrhage: a comparative retrospective analysis on outcome and shunt dependency. *Neurosurgical Focus*, DOI: 10.3171/2012.1.FOCUS11349, **32**, pp. E4-1-E4-8.
52. Bosco E., Marton E., Feletti A., **Scarpa B.**, Longatti P., Zanatta P., Giorgi E., Sorbara C. (2011). Dynamic monitors of brain function: a new target in neurointensive care unit, *Critical Care*, **15**, R170.
53. **Scarpa B.** (2007) La previsione del churn, *Quaderni di Management*, **25**, pp. 1920—1936.
54. **Scarpa B.**, Dunson D.B., Giacchi E. (2007) Bayesian Selection of Optimal Rules for Timing Intercourse Using Calendar and Mucus, *Fertility and Sterility*, DOI: 10.1016/j.fertnstert.2006.12.017, **88**, pp. 915—924.
55. **Scarpa B.**, Dunson D.B., Colombo B. (2006) Cervical mucus secretions on the day of intercourse: An accurate marker of highly fertile days, *European Journal of Obstetrics & Gynecology and Reproductive Biology*, DOI: 10.1016/j.ejogrb.2005.07.024, **125**, pp. 72—78.
56. **Scarpa B.**, Dunson D.B., (2006) Bayesian selection of predictors of day-specific conception probabilities, *Paediatric and Perinatal Epidemiology*, DOI: 10.1111/j.1365-3016.2006.00768.x, **20**, pp. 30—37.
57. Selle V. e **Scarpa B.** (1995). La qualità dei dati delle dimissioni ospedaliere nell’Ospedale Civile di Venezia: aspetti metodologici e valutativi ai fini dell’utilizzo dei DRG, *Epidemiologia e prevenzione*, **19**, pp. 342—348.

Contributi in volumi

56. **Scarpa B.** e Dorigo T. (2017) Bayesian non parametric modelling of Higgs pair production *EPJ Web Conf.* **137**, pp. 11011 DOI: 10.1051/epjconf/201713711011
57. **Scarpa B.** (2012) Optimizing Fertility by Timing Intercourse during the Menstrual Cycle. In: Fehring R. J. and Notare T., editors, *Human Fertility: Where Faith and Science Meet*, Marquette University Press, pp. 37-49.
58. **Scarpa B.** (2010) Probabilistic and Statistical Models for Conception. In: N. Balakrishnan (curatore) *Methods and Applications of Statistics in the Life and Health Sciences John.* Wiley & Sons, pp.336-339.
59. **Scarpa B.** (2011) Data mining. In: M. Lovirc (curatore) *International Encyclopedia of Statistical Sciences.* Springer, Berlino.
60. Bellio R., **Scarpa B.** (2010). Strumenti statistici per il controllo della qualità. In: Mazzaro V.; Sartor M.. *La guida del Sole24ore alla qualità*, p. 55-76.
61. **Scarpa B.** (2010) Probabilistic and Statistical Models for Conception. In: Kotz S. (Editor-in-Chief), Read C.B., Balakrishnan N., Vidakovic B. *Encyclopedia of Statistical Sciences*, Wiley & Sons.
62. **Scarpa B.** (2006) La finestra fertile: evidenze statistiche, In: E.Giacchi, S.Girotto, G.Bozzo, *Il Periodo Fertile*, Edizioni Libreria Cortina, Verona, pp. 13—30.
63. **Scarpa B.** (2004) Networks, Uncertainty and Stochastic models. In: E. Tezza, *Evaluating social programmes. The relevance of relationships*, ILO, Torino, pp. 57—61.
64. **Scarpa B.**, Torelli N. (2004) Selecting the training set in classification problems with rare events. In: M. Vichi, P. Monari, S. Mignani, A. Montanari, (curatori) *New Developments in Classification and Data Analysis*, Springer-Verlag, pp. 39—46.

Atti di convegni

65. Canale A., **Scarpa B.** Manzano P., Dall’Osso M., Dorigo T., Finos L., Kotkowski G., Menardi G., **Scarpa B.** Hemisphere Mixing: a Fully Data-Driven Model of QCD Multijet Backgrounds for LHC Searches. *Proceedings of the European Physical Society Conference on High Energy Physics. 5-12 July, 2017 Venice, Italy (EPS-HEP2017)*. Online at <http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=314>, id. 370.
66. Canale A., **Scarpa B.** (2013) Age Specific Probability of Childbirth, In *Concorso: “Statistica, Open data e Società”*, Comune di Milano e Università Cattolica del Sacro Cuore, (Milano, 13 Dicembre 2014). **Premiato come ‘miglior prodotto’**
67. Ciera J., **Scarpa B.**, Dunson D.. (2005) Fast Bayesian Functional Data Analysis: Application to Basal Body Temperature Data. In *Proceedings SCo2009 Complex Data Modelling and Computationally Intensive Statistical Methods for Estimation and Prediction*, (Milano, 14–16 Settembre 2009).
68. Bortot P., **Scarpa B.**, Masarotto G. (2005) Previsioni sequenziali delle lunghezze del ciclo mestruale, In *Atti delle Giornate di Studio sulla Popolazione*, (Padova, 16–18 Gennaio 2005).
69. Rivellini G., Rosina A., **Scarpa B.** (2004) Late maternal age and risk of congenital malformation, In *Atti della XLII Riunione Scientifica della Società Italiana di Statistica*, Sessioni spontanee (Bari, 9–11 Giugno 2004), pp.19—22.
70. **Scarpa B.** (2004) Cervical mucus symptom and daily fecundability. In *Proceedings of the International Conference on Statistics in Health Sciences* (Nantes, June 23–25, 2004)
71. **Scarpa B.**, Torelli N. (2003) Selecting the training set in classification problems with rare events. In *Atti del convegno CLADAG 2003* (Bologna, settembre 2003).
72. Scarpa B. (2003). Il ruolo dello statistico di marketing nelle aziende di telecomunicazioni. In *Atti della Giornata di Inaugurazione XVIII ciclo del dottorato di Padova in ricordo di Antonio C. Capelo* (Padova, 9 gennaio 2003), pp. 97—106.
73. Scarpa B. (2001). Fidelizzare i clienti: come passare dai dati alle azioni. In *I sistemi di Business Intelligence (Convegno IIR)* (Milano, 26–27 giugno 2001).

74. **Scarpa B.** (1999). Customer Profiling, Segmentation and Marketing Strategies. In *Datawarehousing & Datamining in Telecommunications (IBC Global Conferences)* (Londra, 14–15 settembre 1999).
75. Masarotto, G, **Scarpa B.** (1998). Analisi spazio-temporale di inquinanti atmosferici e posizionamento di centraline: un approccio non-parametrico, In *Atti della XXXIX Riunione Scientifica della Società Italiana di Statistica*, Sessioni plenarie e specializzate (Sorrento, 15–18 aprile 1998), pp. 195–206.
76. **Scarpa B.** (1998). Modelli Additivi Generalizzati con autocorrelazione tra le osservazioni. In *Atti della XXXIX Riunione Scientifica della Società Italiana di Statistica*, Sessioni spontanee (Sorrento, 15–18 aprile 1998), pp. 857–864.

Rapporti Tecnici e Manoscritti

77. Manzano P., Dall’Osso M., Dorigo T., Finos L., Kotkowski G., Menardi G., **Scarpa B.** Hemisphere Mixing: a Fully Data-Driven Model of QCD Multijet Backgrounds for LHC Searches. <https://arxiv.org/abs/1712.02538>.
78. Canale A., **Scarpa B.** (2014) Informative Bayesian inference with the skew-normal distribution. <http://arxiv.org/abs/1305.3080>
79. De Bin R., Scarpa (2014) Non-parametric Bayesian modeling of cervical mucus symptom. Technical Report 170, Department of Statistics, University of Munich, Germania. *Sottoposto a revisione*
80. **Scarpa B.** (2005), Lo stress in azienda. Modelli di analisi di un’indagine per l’identificazione delle cause di stress, Università degli Studi di Pavia, Dipartimento di Statistica ed Economia Applicate, *Working Paper*, n.23.
81. **Scarpa B.** (2004), La previsione del *churn*: problemi e possibili soluzioni di *data mining* nella realtà della telefonia cellulare, Università degli Studi di Pavia, Dipartimento di Statistica ed Economia Applicate, *Working Paper*, n.20.
82. **Scarpa B.** (2004), Confronto tra modelli di *data mining* per la previsione della domanda di energia elettrica, Università degli Studi di Pavia, Dipartimento di Statistica ed Economia Applicate, *Working Paper*, n.21.
83. **Scarpa B.** (2004), La *Customer satisfaction* per un’azienda di servizi informatici. Impostazione e analisi di un’indagine via web, Università degli Studi di Pavia, Dipartimento di Statistica ed Economia Applicate, *Working Paper*, n.22.
84. **Scarpa B.** (2004), Choice of an ”optimal rule” for avoid conception using calendar and mucus data, Università degli Studi di Pavia, Dipartimento di Statistica ed Economia Applicate, *Working Paper*, n.19.
85. **Scarpa B.** (2001), Il grado di invalidità permanente nella polizza infortuni: una valutazione di variabili predittive mediante modello a rischi proporzionali di Cox, Università degli Studi di Padova, Dipartimento di Scienze Statistiche, *Working Paper*, n. 2001.6.
86. Arbia G., Lafratta G., **Scarpa B.** (1996), Evaluating and updating the sample design for the concentration of SO₂ in Padua, Università degli Studi ”G. d’Annunzio” di Pescara, Dipartimento di Metodi Quantitativi e Teoria Economica, *Quaderni di statistica*, n. 4/1996.

Il sottoscritto

COGNOME: Scarpa
NOME: Bruno
CODICE FISCALE: SCRBRN69A29L736I
NATO A: Venezia
IL: 29/01/1969 SESSO: M

consapevole che, secondo quanto previsto dagli artt. 46 e 47 D.P.R. n. 445/00, le dichiarazioni mendaci sono punite ai sensi del codice penale e delle leggi speciali in materia,

DICHIARA:

che quanto indicato nel presente curriculum corrisponde a verità.

Venezia, 30 aprile 2019

Bruno Scarpa