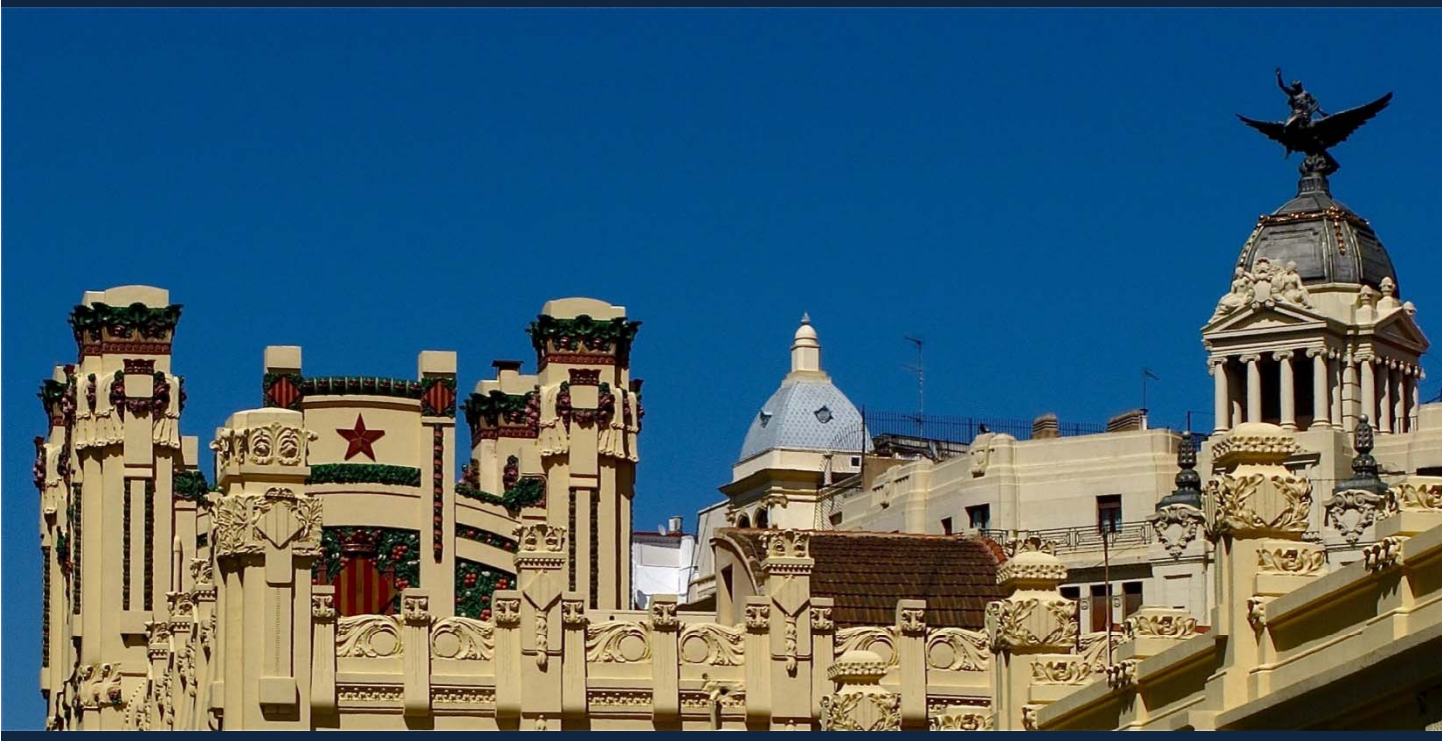
IWSM 2011

Proceedings of the

26th International Workshop on Statistical Modelling

Valencia (Spain), July 11-15, 2011



Editors:
David Conesa
Anabel Forte
Antonio López-Quílez
Facundo Muñoz

Proceedings of the 26th International Workshop on Statistical Modelling

July 11-15, 2011

València

David Conesa, Anabel Forte,
Antonio López-Quílez, Facundo Muñoz
(editors)

Proceedings of the 26th International Workshop on Statistical Modelling.
València, July 11-15, 2011
David Conesa, Anabel Forte, Antonio López-Quílez, Facundo Muñoz, eds.
València 2011.
ISBN 978-84-694-5129-8

Editors:

David Conesa[1], David.V.Conesa@uv.es
Anabel Forte[2], forte@eco.uji.es
Antonio López-Quílez[1], Antonio.Lopez@uv.es
Facundo Muñoz[1], Facundo.Munoz@uv.es

[1] Departament d'Estadística i Investigació Operativa
Universitat de València (Estudi General)
Facultat de Matemàtiques
Dr. Moliner 50, 46100 Burjassot, Spain.

[2] Departamento de Economía
Universitat Jaume I
Facultad de Ciencias Jurídicas y Económicas
Campus del Riu Sec, E-12071 Castelló de la Plana, Spain.

Cover photo: Victor Roda

Printed by Copiformes S.L.

Scientific Programme Committee

- Susie Bayarri (Chair)
  Universitat de València, Spain
- Carmen Armero
  Universitat de València, Spain
- Adrian Bowman
  University of Glasgow, UK
- Charmaine B. Dean
  Simon Fraser University, Canada
- María Durbán
  Universidad Carlos III de Madrid, Spain
- Claire Ferguson
  University of Glasgow, UK
- Herwig Friedl
  Graz University of Technology, Austria
- Gillian Heller
  Macquarie University, Australia
- John Hinde
  University of Galway, Ireland
- Thomas Kneib
  Carl von Ossietzky Universität Oldenburg, Germany
- Arnost Komàrek
  Charles University in Prague, Czech Republic
- Antonio López-Quílez
  Universitat de València, Spain
- Brian Marx
  Louisiana State University, USA
- Pere Puig
  Universitat Autònoma de Barcelona, Spain

Preface

This volume contains all the papers of the 26th International Workshop on Statistical Modelling. Many things have changed since in 1986 an enthusiastic group of statisticians interested in statistical modelling started these series of workshops within a friendly and supportive academic atmosphere. New technologies, more attendants, but always with the same initial spirit: to promote and develop the use of statistical modelling in research and applications.

We are glad to present you these Proceedings, which clearly reflect the aliveness of that spirit. One the one hand, the five invited papers show new advances in theoretical research but always keeping an eye in their applied interest. One the other hand, the great amount of contributions (a total of 140) and their quality demonstrate that the workshop is in good shape. Authors should receive most of the credit for the quality of these Proceedings. Nevertheless, all submissions were carefully reviewed by the members of the Scientific Committee. Their detailed work has been reflected in a big improvement of the preliminary versions jointly with the final selection of contributions.

This 26th edition of the IWSM will be held in Valencia (Spain) in an informal environment (ADEIT- FUNDACIÓ UNIVERSITAT-EMPRESA of the Universitat de València) to encourage discussion and exchange of ideas which could result in future research. Valencia has a great tradition in Statistics and in particular in Bayesian Statistics. This why we are so happy to see that this way of thinking and doing statistics is quite present in these Proceedings reflecting its important role in the Society. We will also like to comment, that many of the contributions in these Proceedings are due to students, which clearly have the future in their hands.

Finally, we wish to acknowledge Carmen Armero, the chair of the local Committee for putting together all the pieces needed in the process of organising this event. Without her interest and passion it would have been impossible.

So welcome to Valencia. Enjoy the city and surroundings and have a great conference.

David Conesa, Anabel Forte, Antonio López-Quílez, Facundo Muñoz
Valencia, June 2011

Contents

Part 1. Invited papers

Part 2. Contributed papers

Modelling the Timing of Marital Dissolution in Italy: censored quantile regression with additive terms

Mariano Porcu[1], Vito M. R. Muggeo[2], Vincenza Capursi[2]

[1] Dipartimento Ricerche Economiche e Sociali, Università di Cagliari, ITALY
email: mrporcu@unica.it
[2] Dipartimento Scienze Statistiche e Matematiche 'S. Vianelli', Università di Palermo, ITALY

Abstract: The analysis of marital dissolution in Italy represents a quite interesting and challenging topic from a substantive standpoint; in fact, despite of the decreasing number of marriages and the increasing number of divorces, the traditional family based on the marriage of heterosexual partners is still considered as a fundamental institution of the society. Here we present a censored quantile regression model with additive terms to investigate the determinants of the timing of marital dissolution on a large and substantial sample from a survey carried on in Italy.

Keywords: censored quantile regression; Timing of Marital Dissolution; Survival data; Smoothing

1 Introduction

It is commonly asserted that the family is the *fundamental institution* of the Italian society. The main consequence of this assertion is a widespread political support addressed to the upholding of the classical family built on the marriage of heterosexual partners. In this context the analysis of the possible determinants of the marital dissolution is a largely debated issue: the study of the factors that could affect the end of the marriage is prominent in the social research. The topic is also of great interest for the policy-makers as the marital dissolution affects some of the key features of the modern societies, such as economy, gender equality and especially fertility. Although the study of the time-to-separation can provide quite useful information and insights to evaluate trends and changes in the formation and dissolution of the marriage, relatively few studies take explicitly into account the time dimension, see Cavanagh and Huston (2008), Gottman and Levenson (2000).
Using a large sample surveyed by the Italian national statistics institute (ISTAT), we aim to model the time-to-marital dissolution in a regression

quantile framework. While the Cox model represents the most used framework to model survival data, censored quantile regression (CQR) offers a more flexible alternative by focusing the attention on narrow slices, lower or upper tails, of the conditional survival distribution of interest (Koenker, 2008).

2 The Data

The data considered in this paper come by from the sample survey on *Families and Social Subjects* (FSS), carried on in Italy by the official statistics institute at the end of 2003 on a sample of over 19,000 Italian families (nearly 50,000 individuals). The survey was addressed to collect broad information on the Italian households, such as the shapes, the network of kinship, the relations among partners, the permanence of young adults in the family, and the working life.

TABLE 1. Some descriptive statistics on data analysed.

	Males ($n = 4633$)		Females ($n = 5235$)	
Covariates	Separated	Non-Separ	Separated	Non-Separ
Area (*obs.*)				
North	335	1640	445	1763
Center	190	1164	223	1269
South	124	1180	151	1384
Education (*obs.*)				
1^{st} stage basic	75	591	63	684
2^{nd} stage basic	253	1529	268	1586
Upper secondary	232	1473	384	1713
Degree	89	391	104	433
Age at Marriage (*Years*)				
Mean (sd)	*26.3(5.7)*	*27.7 (5.8)*	*23.3(5.0)*	*22.2(5.7)*
Childless				
%	*33.4*	*12.8*	*32.1*	*12.9*
Work				
Yes at marr (%)	*80.1*	*86.7*	*51.2*	*45.0*
Yes at separ (%)	*88.0*	—	*64.7*	—

The data from the 2003 FSS survey here analyzed represent the most recent information available on the topic: the data from the last FSS survey carried out in 2010 are not yet available. We have omitted from the sample persons married before the 1970 when the divorce, understood as the 'total dissolution of marital status', was not allowed. Moreover, to keep away from any potential confounding interaction between sex, covariates and timing we have considered females and males independently (e.g., Schoen and Canudas-Romo, 2006). Table 1 summarizes some descriptive statistics for the sample.

3 Methods and Results

We aim to model the time-of-dissolution as a function of the following covariates in a CQR model: the categorical variables Education, Area, and Work at Marriage, and the numerical variables Age at Marriage, Number of Sons and Year of Birth. Until now, CQR has been discussed only with parametric linear terms, however for the aforementioned continuous covariates the linearity assumption is not tenable and more flexible alternatives are requested: we use B-spline bases with quadratic penalties on the coefficients to get smooth estimates of the nonlinear relationships. The additive CQR model with J nonparametric terms for the variables z and linear terms for the variables x, may be written as

$$Q_\tau(Y|x_i) = x_i^T \beta_\tau + \sum_j^J f_{\tau j}(z_{ij}), \tag{1}$$

where the subscript τ points the percentile of interest ($0 < \tau < 1$). Notice that, unlike the usual model for the conditional mean, here the covariate effect (parametric or nonparametric) depends on the percentile τ. The response Y measures the time span of their marriage up to the year of separation; we consider as uncensored the spouses (male or female) which stop living together regardless of the possibility of reconciliation; in fact in Italy only a slight proportion of separations ends with a reuniting of the couple (Castiglioni, 2008). We modify the iterative estimating algorithm described in Bottai and Zhang (2010) to include the additive (spline) terms in the linear predictor and to obtain parameter estimate of the additive CQR model (1). Although QR allows to model every quantile of the response conditional distribution, our analysis focuses on the lowest quantiles ($\tau \leq 0.10$). Indeed, early dissolutions (i.e. the left tail of the survival distribution) are of major interest in the present study since the first years of marriage are known to be crucial for fertility, children social development, changes in lifestyle and also for their influence on the probability of remarriage.

TABLE 2. Point estimates for the parameters of the linear terms in the four CQR models.

	Males		Females	
Linear Terms	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.05$	$\tau = 0.10$
EDUC (2 stage basic vs 1 stage basic)	-2.257	-1.666	-2.190	-3.971
EDUC (upper sec. vs. 1 stage basic)	-1.372	-1.138	-3.444	-5.668
EDUC (degree vs 1 stage basic)	-2.291	-2.474	-4.454	-6.503
Area (center vs. north)	0.197	1.081	0.770	1.116
Area (south vs. north)	2.319	2.730	1.317	1.486
Work at marr (yes vs. no)	1.500	2.157	0.469	0.130

For the four fitted additive CQR models (two quantiles 0.05 and 0.10 for males and females), Table 2 shows the point estimates for the parameters

of the linear terms and Figure 1 reports the fitted smooth effects of the three continuous covariates.

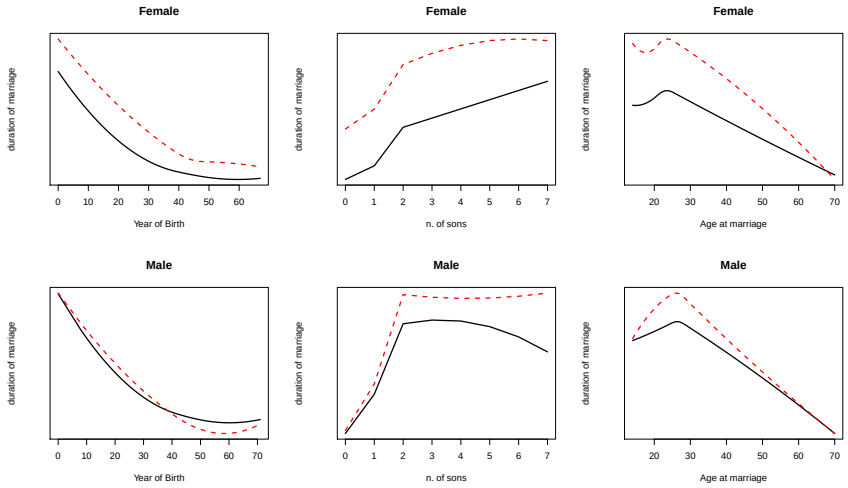


FIGURE 1. Fitted quantiles ($\tau = 0.05$, continuous line; $\tau = 0.10$ dashed line) for males and females.

In short, for the *early* marital dissolutions (i.e., the low percentiles 5% and 10%) we observe strong and somewhat expected effects of the 'area' and of the 'educational level' for both male and female groups. On the other hand, the effect of the 'working status' is somewhat different. The plots in Figure 1 emphasize the nonlinear effects of the continuous covariates, by highlighting a different role of the number of sons variable.

References

Bottai, M., Zhang, J. (2010) Laplace regression with censored data *Biometrical Journal*, **52**, 487-503.

Castiglioni, M., Dalla Zuanna, G. (2008) A Marriage-Cohort Analysis of Legal Separations in Italy. *Population*, **63**, 1, 173-193.

Cavanagh, S.E., Huston, A.C. (2008) The Timing of Family Instability and Children's Social Development. *Journal of Marriage and Family*, **70**, 1258-1269.

Gottman, J.M., Levenson, R.W. (2000) The Timing of Divorce: Predicting When a Couple Will Divorce Over a 14-Year Period. *Journal of Marriage and Family*, **62**, 737-745.

Koenker, R.W. (2008) Censored Quantile Regression Redux *Journal of Statistical Software*, **27**, 6 , 1-25.