SDS
Statistica e
Data Science

Palermo, 11-12 April 2024

Proceedings of the Statistics and Data Science 2024 Conference

New perspectives on Statistics and Data Science

Edited by

Antonella Plaia – Leonardo Egidi
Antonino Abbruzzo

Proceedings of the SDS 2024 Conference
Palermo, 11-12 April 2024
Edited by: Antonella Plaia – Leonardo Egidi - Antonino Abbruzzo
-
Palermo: Università degli Studi di Palermo.

ISBN Ebook (Pdf)
978-88-5509-645-4

Contents

Seismic events classification through latent class regression models for point processes

Giada Lo Galbo, Giada Adelfio, Marcello Chiodi

Abstract We are trying to identify sub-processes of seismic events from the point processes' point of view and according to the latent class regression approach. Each seismic event is classified as membership of one of the 4 identified sub-classes of seismic sequences, each defined by particular and well-defined characteristics. So far, seismic sub-sequences have been identified and described according to several declustering methods. In this application, we show how sub-processes can be identified starting from the definition of a spatio-temporal intensity function for point processes, assuming independence of the past.

Key words: Latent class, mixture model, spatio-temporal point process, clustering, earthquake, seismic sequence

1 Introduction

The study of seismic sequences has been approached to identify sub-processes characterizing the background and the induced components [1]. What has been studied so far concerns the characteristics of the magnitude distribution corresponding to the cited above components [11], or the covariates affecting the induced component [3]. What has not been taken into account is the importance of covariates related to the occurrence of earthquakes. This study aims to identify the components of seismic processes, through the application of a latent class regression model, by analyzing the spatio-temporal intensity dependence on covariates.

Giada Lo Galbo · Giada Adelfio · Marcello Chiodi
Department of Economics, Business and Statistics, University of Palermo e-mail: giada.logalbo@unipa.it, e-mail: giada.adelfio@unipa.it, e-mail: marcello.chiodi@unipa.it

2 Methodology

2.1 Spatio-temporal point process

A spatio-temporal inhomogeneous Poisson point process has parametric probability density function, $f_X(X;\boldsymbol{\theta})$, defined within a spatio-temporal bounded window, $|W \times T|$, with volume $W > 0$ and length $T > 0$ as in Eq. (1)[4]:

$$f_X(X;\boldsymbol{\theta}) = e^{\left\{\int_W \int_T [1-\lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t)]dtd\boldsymbol{u}\right\}} \prod_{(\boldsymbol{u},t)\in X} \lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t) \qquad \{(\boldsymbol{u},t)\} \in X \qquad \boldsymbol{\theta} \in \boldsymbol{\Theta} \tag{1}$$

where: $\boldsymbol{\theta}$ is the set of regression parameters; $\boldsymbol{\Theta}$ is the parameter space and $\lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t)$ is the parametric first-order intensity function describing the point pattern X.

Assuming log-linear dependence from a set of covariates, $\boldsymbol{Z}(\boldsymbol{u},t)$, the first-order intensity function of the $(\boldsymbol{u},t)^{th} \in X$ point is specified as in Eq. (2) [10]:

$$\lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t) = \exp\left\{\boldsymbol{\theta}'\boldsymbol{Z}(\boldsymbol{u},t) + \varrho\,(\boldsymbol{u},t)\right\} \tag{2}$$

with $\varrho\,(\boldsymbol{u},t)$ a scalar offset. The parameters, $\boldsymbol{\theta}$, are estimated by maximizing the log-likelihood function, $\ell_X(\boldsymbol{\theta};X)$. Berman and Turner[5] propose a device, which adds a set of dummy points, X_d, within the spatio-temporal window, $|W \times T|$, where there are no observed points; i.e., the latter can be approximated as in Eq. (3):

$$\ell_X\big(\boldsymbol{\theta};X\big) = \sum_{(\boldsymbol{u},t)\in X_D} \ell_X\left(\boldsymbol{\theta};(\boldsymbol{u},t)\right) \approx \sum_{(\boldsymbol{u},t)\in X_D} w_{(\boldsymbol{u},t)} \left\{y_{(\boldsymbol{u},t)} \log\left[\lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t)\right] - \lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t)\right\} \tag{3}$$

where: $\sum_{(\boldsymbol{u},t)\in X_D} w_{(\boldsymbol{u},t)} = |W \times T|$; $X_D = X \cup X_d$; $w_{(\boldsymbol{u},t)}$ is a quadrature weight of the $(\boldsymbol{u},t)^{th}$ point; $y_{(\boldsymbol{u},t)} = \frac{z_{(\boldsymbol{u},t)}}{w_{(\boldsymbol{u},t)}}$, where $z_{(\boldsymbol{u},t)} = 1$ if $(\boldsymbol{u},t) \in X$, $z_{(\boldsymbol{u},t)} = 0$ if $(\boldsymbol{u},t) \in X_d$.

2.2 Finite mixtures of log-linear regression models

A finite mixture model of R components has conditional density function, $g_X(X \mid \boldsymbol{\psi})$, defined by a whole set of parameters $\boldsymbol{\psi} = \{\boldsymbol{\theta}_r, \pi_r\}_{r=1}^R$, as expressed in Eq. (4)[7, 8]:

$$g_X\big(X \mid \boldsymbol{\psi}\big) = \sum_{r=1}^{R} \pi_r f_X(X;\boldsymbol{\theta}_r) \qquad \text{with: } \sum_{r=1}^{R} \pi_r = 1; \quad \pi_r > 0 \tag{4}$$

where: $\{\boldsymbol{\theta}_r, \pi_r\}$ are the regression and weight parameters describing the r^{th} cluster.

The posterior probability of j^{th} cluster membership, is defined as in Eq. (5):

$$\rho_{(\boldsymbol{u},t)_j} = P\big(j \mid (\boldsymbol{u},t), \boldsymbol{\psi}\big) = \frac{\pi_j f_X\big((\boldsymbol{u},t);\boldsymbol{\theta}_j\big)}{\sum_{r=1}^{R} \pi_r f_X\big((\boldsymbol{u},t);\boldsymbol{\theta}_r\big)} \qquad \forall (\boldsymbol{u},t) \in X \tag{5}$$

where: $\{f_X(X;\boldsymbol{\theta}_r)\}_{r=1}^{R}$ is the set of components belonging to the finite mixture.

The parameters estimate, $\boldsymbol{\psi}$, is obtained by maximizing the *'complete data log-likelihood'*, $c\ell_X(\boldsymbol{\psi};X)$, which is expressed as in Eq. (6):

$$c\ell_X(\boldsymbol{\psi};X) = \sum_{(\boldsymbol{u},t)\in X} \log\left[\sum_{r=1}^{R} \pi_r f_X\big((\boldsymbol{u},t);\boldsymbol{\theta}_r\big)\right] \tag{6}$$

The maximization of the complete data log-likelihood function is carried out through the iterative Expectation-Maximization (EM) algorithm [6]. For the r^{th} cluster, at the $(i+1)^{th}$ iteration, given: $\hat{\boldsymbol{\psi}}^{(i)} = \left\{\hat{\boldsymbol{\pi}}^{(i)}, \hat{\boldsymbol{\theta}}^{(i)}\right\}$, the Expectation and the Maximization steps consist of, respectively [9]:

- **$\mathbb{E}$-step** $\quad \mathbb{E}\left[\hat{\pi}_r^{(i+1)}\right] = \dfrac{1}{n(X)} \displaystyle\sum_{(\boldsymbol{u},t)\in X} \hat{\rho}^{(i)}_{(\boldsymbol{u},t)_r} = \dfrac{1}{n(X)} \displaystyle\sum_{(\boldsymbol{u},t)\in X} \hat{P}\left(r \,\middle|\, (\boldsymbol{u},t), \hat{\boldsymbol{\psi}}^{(i)}\right)$
- **M-step** $\quad \hat{\boldsymbol{\theta}}_r^{(i+1)} = \arg\max_{\boldsymbol{\theta}_r} \left\{ \displaystyle\sum_{(\boldsymbol{u},t)\in X} \hat{\rho}^{(i)}_{(\boldsymbol{u},t)_r} \ell_X\left(\boldsymbol{\theta}_r^{(i)};(\boldsymbol{u},t)\right)\right\}$

3 Application to seismic data

We are interested in classifying events characterizing a spatio-temporal clustered point pattern by a probabilistic clustering approach [8]; i.e., according to an estimated probability of latent class membership [2]. The data are provided by the Istituto Nazionale di Geofisica e Vulcanologia (INGV), and belong to the Catalogo delle Localizzazioni ASSolute (CLASS). They refer to Italian earthquakes that occurred between 1980 and 2018, with information on: depth, $D_p(\boldsymbol{u},t)$; root mean square error from P/S arrival time, $R(\boldsymbol{u},t)$; hypocentral error on vertical, $E_v(\boldsymbol{u},t)$, and horizontal, $E_h(\boldsymbol{u},t)$, coordinates; gap azimuth, $G(\boldsymbol{u},t)$; distance from the nearest station, $D(\boldsymbol{u},t)$; distance between probabilistic and expected hypocenter, $D(\boldsymbol{u},t)$; quality location, $Q(\boldsymbol{u},t)$; radius of a sphere with PDF volume, $P(\boldsymbol{u},t)$.

By using the intensity function in Eq. (7), we fit models with a number of latent classes from $R = 2$ to $R = 8$ and choose the model with $R = 4$ latent classes, according to the AIC and BIC criteria:

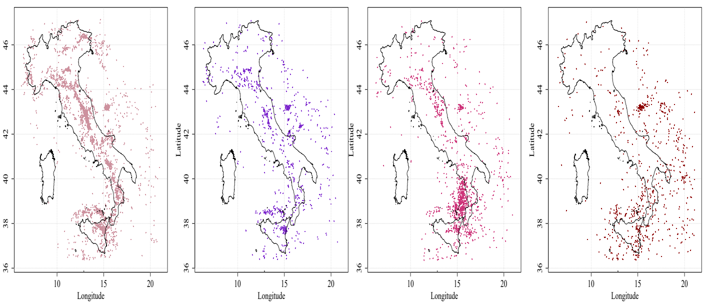
$$\begin{aligned}\lambda_{\boldsymbol{\theta}}(\boldsymbol{u},t) = e^{\{\theta_0+\theta_1 D_p(\boldsymbol{u},t)+\theta_2 R(\boldsymbol{u},t)+\theta_3 E_v(\boldsymbol{u},t)+\theta_4 E_h(\boldsymbol{u},t)+\,\cdots} \\ {}^{\cdots\,+\theta_5 G(\boldsymbol{u},t)+\theta_6 D(\boldsymbol{u},t)+\theta_7 L(\boldsymbol{u},t)+\theta_8 Q(\boldsymbol{u},t)+\theta_9 P(\boldsymbol{u},t)\}}\end{aligned} \tag{7}$$

Tab. 1 shows the results of the latent class regression model fitting. As shown in Tab. 1, the covariates have the same positive (R, E_v, G, D) or negative (D_p, E_h, L, Q, P) effects, respectively, between the 1^{st} and the 2^{nd} latent classes, on the intensity of seismic events, although the different estimates' values; some covariates have similar positive (D_p, G, D) or negative (Q, P) effects, on the intensity of seismic events classified as belonging to the 3^{rd} and the 4^{th} latent classes. According to the regression parameters' estimates, the remaining covariates have null or opposite effects (and different magnitudes), among 3^{rd} and 4^{th} latent classes' intensities.

Table 1: Regression parameters estimates, $\hat{\theta}$, with standard errors, $\hat{\sigma}_{\hat{\theta}}$, and p-values, p, for the model with $R = 4$ latent classes and first-order intensity function defined as in Eq. (7)

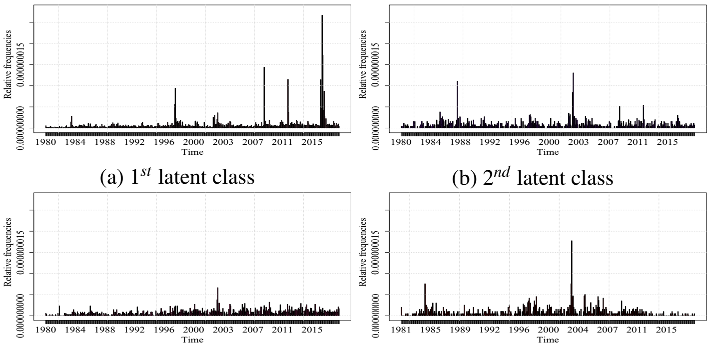
r			D_p	R	E_v	E_h	G	D	L	Q	P
		θ_0	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9
	$\hat{\theta}$	**−14.677**	**−0.057**	**0.541**	**0.118**	−0.003	**0.008**	**0.011**	**−0.330**	**−3.086**	**−0.134**
1	$\hat{\sigma}_{\hat{\theta}}$	0.094	0.004	0.038	0.014	0.008	0.001	0.002	0.021	0.240	0.023
	p	< 0.001	< 0.001	< 0.001	< 0.001	0.718	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	$\hat{\theta}$	**−14.624**	**−0.044**	**0.153**	**0.204**	**−0.301**	**0.006**	**0.012**	**−0.199**	**−2.770**	**−0.170**
2	$\hat{\sigma}_{\hat{\theta}}$	0.094	0.003	0.067	0.010	0.020	0.001	0.002	0.013	0.279	0.030
	p	< 0.001	< 0.001	0.022	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	$\hat{\theta}$	**−14.788**	**0.004**	**0.089**	**0.062**	**0.056**	**0.011**	**0.005**	**−0.275**	**−4.657**	**−0.169**
3	$\hat{\sigma}_{\hat{\theta}}$	0.096	0.001	0.034	0.006	0.007	0.001	0.001	0.018	0.255	0.021
	p	< 0.001	< 0.001	0.009	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	$\hat{\theta}$	**−14.946**	**0.006**	**−2.255**	0.000	0.000	**0.014**	**0.005**	**0.004**	**−4.205**	**−0.059**
4	$\hat{\sigma}_{\hat{\theta}}$	0.115	0.000	0.281	0.001	0.001	0.001	0.001	0.001	0.256	0.007
	p	< 0.001	< 0.001	< 0.001	0.847	0.977	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figs. 1 and 2 show the spatial and the temporal distributions, respectively, of the 1^{st} (2a, 3a), 2^{nd} (2b, 3b), 3^{rd} (2c, 3c) and 4^{th} (2d, 3d) latent classes. As shown in Figs. 1 and 2, the 3^{st} and 4^{th} latent classes are defined by seismic events loosely aggregated and nearer the tectonic zones around the Messina area. The 1^{st} and 2^{nd} latent classes identify seismic sequences which occurred in Central Italy and Northern Apennine chain.



(a) 1^{st} latent class (b) 2^{nd} latent class (c) 3^{rd} latent class (d) 4^{th} latent class

Fig. 1: Spatial distribution of seismic events classified as belonging to the 1^{st} (2a), 2^{nd} (2b), 3^{rd} (2c) and 4^{th} (2d) latent classes, according to the first-order intensity function expressed in Eq. (7)



(a) 1^{st} latent class (b) 2^{nd} latent class

(c) 3^{rd} latent class (d) 4^{th} latent class

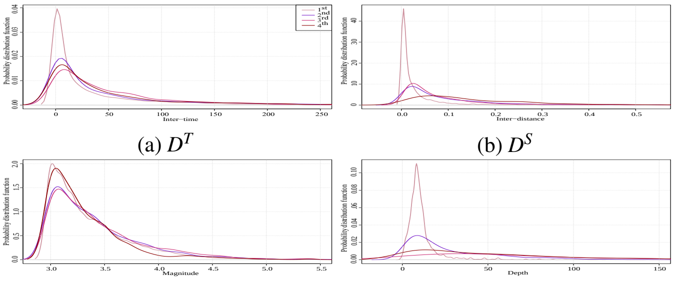
Fig. 2: Temporal distribution of seismic events classified as belonging to the 1^{st} (3a), 2^{nd} (3b), 3^{rd} (3c) and 4^{th} (3d) latent classes, according to the first-order intensity function expressed in Eq. (7)

Fig. 3 and Tab. 2 show, respectively, the density functions of depth (D_p, 4d), magnitude (m, 4c), inter-time (D^T, 4a), inter-distance (D^S, 4b) and the weight parameters estimates, $\{\pi_r\}_{r=1}^{R}$, with the summary statistics of magnitude, m, depth, D_p, inter-time, D^T, and inter-distance, D^S, for each of the $R = 4$ latent classes.

Table 2: Weight parameters, $\{\pi_r\}_{r=1}^{R}$, and summary statistics of magnitude, m, depth, D_p, inter-time (hours), D^T, and inter-distance (km), D^S, for the model with $R = 4$ latent classes and first-order intensity function defined as in Eq. (7)

r	π_r		Quantile					Moment			
			Q_0	Q_1	Q_2	Q_3	Q_4	μ	σ	μ_3	β_2
1	0.65	m	3.00	3.10	3.20	3.50	6.10	3.36	0.40	1.96	8.54
		D_p	−1.40	6.36	8.96	12.08	216.56	11.66	12.60	5.83	54.90
		D^T	0.00	0.62	8.18	38.15	1241.34	32.80	3749.80	4.94	52.49
		D^S	0.00	0.00	0.01	0.03	1.78	0.03	0.00	8.15	132.47
2	0.11	m	3.00	3.10	3.30	3.60	5.70	3.40	0.43	1.71	6.67
		D_p	−0.93	8.00	17.18	41.10	205.69	28.41	29.20	1.87	7.48
		D^T	0.00	2.76	17.15	58.85	1549.91	49.40	8070.56	6.53	87.12
		D^S	0.00	0.02	0.05	0.13	1.24	0.10	0.02	3.33	19.64
3	0.15	m	3.00	3.10	3.30	3.60	5.80	3.44	0.45	1.63	6.17
		D_p	−1.05	28.82	61.96	198.53	615.93	111.35	107.81	1.15	3.74
		D^T	0.00	8.08	29.81	69.92	690.53	52.50	4589.54	2.87	15.87
		D^S	0.00	0.02	0.05	0.11	1.89	0.10	0.02	4.77	38.36
4	0.09	m	3.00	3.10	3.20	3.50	5.40	3.32	0.37	2.11	8.72
		D_p	−0.69	13.98	40.94	82.42	558.78	64.81	75.59	2.48	11.47
		D^T	0.00	3.54	22.63	61.81	574.52	48.78	5057.82	2.71	12.80
		D^S	0.00	0.06	0.11	0.23	2.09	0.17	0.04	4.00	28.7

As shown in Fig. 3 and Tab. 2, the 1^{st} and 2^{nd} latent classes are in that order the most spatio-temporally aggregated and characterized by highest (lowest) maximum values of magnitude (depth). The 1^{st} latent class is characterised by the majority of events belonging to the catalog. The 3^{rd} and 4^{th} latent classes of events are the least clustered and are characterized by events occurring at the highest depths. The latter groups of events correspond to deep earthquakes occurring around the Calabro-Ionian slab.



(a) D^T (b) D^S (c) m (d) D_p

Fig. 3: Density function of inter-time, D^T (4a, hours), inter-distance, D^S (4b, km), magnitude, m (4c), and depth, D_p (4d), conditional to latent class membership

4 Conclusions

Despite the experimental purpose of the analysis, the application of a latent class regression model for point process, assuming independence of the past, allowed us to identify seismic sub-patterns. The results highlighted by the subsequent exploratory analysis on the events characterizing each latent class, are interesting. Further investigation of the identified sub-sequences, or accounting for different subsets of covariates, could lead to in-depth results concerning the seismic phenomenon.

References

1. Adelfio, G., et al.: Earthquakes clustering based on maximum likelihood estimation of point process conditional intensity function. *Proceedings of the "The* 4^{th} *International Workshop on Statistical Seismology" (StatSei4)*, **13**, (2006)
2. Adelfio, G., Chiodi, M.: Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stoch. Environ. Res. Risk. Assess.*, **29**, pp. 443-–450. Springer (2014)
3. Adelfio, G., Chiodi, M.: Including covariates in a space-time point process with application to seismicity. *Stat. Methods. Appl.*, **30**, pp 947 − −971 (2021)
4. Baddeley, A., Rubak, E., Turner, R.: Spatial Point Patterns Methodology and Applications with R. Chapman and Hall/CRC, New York (2015)
5. Berman, M., Turner, T.R.: Approximating point process likelihoods with GLIM. *Appl. Stat.*, **41**:31 − −38, Wiley (1992)
6. Dempster, A. P., Nan M. L., and Donald B. R.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Series B Stat. Methodol.*. **39**:1, pp. 1 − −22 (1977)
7. DeSarbo, W. S., and William L. C.: A maximum likelihood methodology for clusterwise linear regression. *J. Classif.*, 5: pp. 249 − −282 (1988)
8. Fraley, C., Raftery, A. E.: Model-Based Clustering, Discriminant Analysis, and Density Estimation. *J. Am. Stat. Assoc.*, **97**:458, pp. 611–631. Taylor & Francis (2002)
9. Leisch, F.: `FlexMix`: A General Framework for Finite Mixture Models and Latent Class Regression in R. *J. Stat. Softw.*. **11** (2004)
10. Lewis, P., A. W.: Recent results in the statistical analysis of univariate point processes. *N. P. S.* (1971)
11. van Stiphout, T., Zhuang, J., Marsan, D.: Seismic declustering, Community Online Resource for Statistical Seismicity Analysis. (2012)