

# IASC-LARS Report on Webinar Course on Computational Statistics and Data Science “Statistical Inference in Markov Processes”

## November 23-26, 2020



### IASC- LARS SCHOOL ON COMPUTATIONAL STATISTICS AND DATA SCIENCE

The Latin American Regional Section of the IASC, the IASC-LARS School on Computational Statistics and Data Science, the IASC and the ISI held a first and successful IASC-LARS Webinar on Computational Statistics and Data Science. The webinar course “Statistical Inference in Markov Processes” was team-taught by Professors Verónica González-López and Jesús E. García from University of Campinas, Brazil, from November 23-26, 2020. The course had two collaborators, Mariela Fernández at B3 and Márcio Lanfredi Viola at Federal University of São Carlos, Brazil.

The course was attended by 71 students, 50 students attended on 2 or more days, 56 students submitted their registration forms. Participants came from 24 countries, from Latin America –Argentina, Brazil, Chile, and Uruguay—and from countries as far as away as Algeria, Austria, Bangladesh, Benin, Ethiopia, Germany, Ghana, Iceland, India, Indonesia, Italy, Kenya, Malaysia, Mali, Nigeria, Pakistan, Philippines, South Africa, Sri Lanka, and the USA. Most of the students were undergraduate, master and PhD students (52%), also attending academics and researchers (20%), and professionals coming from government institutions and private industry, 24% of the students were female.

The International Organizing Committee consisted of Luis Firinguetti-Limone IASC Webinar Coordinator & Webinar Host (Chile), David Muñoz-Negron IASC-LARS Chairperson (Mexico), and Alba Martínez-Ruiz IASC-LARS Scientific Secretary (Chile). Katie Junasova, Junior Webmaster at the International Statistical Institute, provided support with the organization and the GoToWebinar platform.

The webinar course was supported by the Latin American Regional Section of the International Association for Statistical Computing (IASC-LARS), the International Association for Statistical Computing (IASC), the International Statistical Institute (ISI), the World Bank, and the Statistical Capacity Building Program.

## Objectives of the IASC-LARS School on Computational Statistics and Data Science

- 1) To spread the knowledge base and advances in Statistical Computing in Latin America and the world,
- 2) To provide an overview of the state-of-the-art of the ongoing research in computational statistics,
- 3) To provide an overall perspective of the application of computational statistics in data science problems,
- 4) To present applications where computational statistics have been crucial to solve problems in real-life applications, and
- 5) To increase the number of researchers and practitioners in computational statistics and data science.

### Agenda and course program.

#### Monday, November 23, 9.00 am.

- Markov processes in finite alphabets with finite memory ([1], [2], [3])
  - ✓ Markov chains with fixed order
  - ✓ Variable length Markov chains
  - ✓ Partition Markov models
  - ✓ Examples from real data

#### Tuesday, November 24, 9.00 am

- Estimation of structures of Markov processes in finite alphabets with finite memory ([4], [2], [3], [5], [6])
  - ✓ Consistent criteria
  - ✓ Procedures and algorithms
  - ✓ Examples from real data

#### Wednesday, November 25, 9.00 am

- Structures in multiple Markov processes in finite alphabets with finite memory ([7], [8])
  - ✓ Insufficient memory to estimate a multivariate Markov process. Estimation through the empirical copula
  - ✓ Partition Markov model for multiple processes
  - ✓ Examples from real data

#### Thursday, November 26, 9.00 am

- Metrics in Markov processes on finite alphabets with finite memory ([3], [9], [10], [11], [12])
  - ✓ How to find parts of a partition
  - ✓ How to establish the discrepancy between processes
  - ✓ How to classify samples
  - ✓ Properties and statistical consistency
  - ✓ Examples from real data

### Number of students by day

Date	Attendance
Monday, Nov. 23	60
Tuesday, Nov. 24	51
Wednesday, Nov. 25	37
Thursday, Nov. 26	34
Total (without replication)	71

Number of students with registration form (country, gender, and occupation): 56.

### Number of students attending 1, 2, 3 or 4 days

Category	Attendance	Percentage (%)
1	21	29.58
2	14	19.72
3	11	15.49
4	25	35.21
Total	71	100.00

50 students attended on 2 or more days.

### Number of students by gender

Gender	Attendance	Percentage (%)
Female	17	23.94
Male	39	54.93
No information	15	21.13
Total	71	100.00

### Number of students by occupation

Occupation	Attendance	Percentage (%)
Undergraduate, master and PhD students	37	52.11
Academic and researchers	14	19.72
Government institutions	2	2.82
Private industry	2	2.82
No information	16	22.54
Total	71	100.00

### Number of students by country

Country	Attendance	Percentage (%)	Country	Attendance	Percentage (%)
Algeria	1	1.41	Italy	3	4.23
Argentina	1	1.41	Kenya	16	22.54
Austria	1	1.41	Malaysia	1	1.41
Bangladesh	1	1.41	Mali	1	1.41
Benin	1	1.41	Nigeria	7	9.86
Brazil	5	7.04	Pakistan	1	1.41
Chile	2	2.82	Philippines	1	1.41
Ethiopia	1	1.41	South Africa	1	1.41
Germany	1	1.41	Sri Lanka	1	1.41
Ghana	1	1.41	Uruguay	1	1.41
Iceland	1	1.41	USA	3	4.23
India	3	4.23	No information	15	21.13
Indonesia	2	2.82	Total	71	100.00

-----  
Report sent on December 11, 2020.

Alba Martinez-Ruiz, IASC-LARS Scientific Secretary.