

IASC-LARS WEBINAR ON COMPUTATIONAL STATISTICS AND DATA SCIENCE



IASC- LARS SCHOOL ON COMPUTATIONAL
STATISTICS AND DATA SCIENCE

STATISTICAL INFERENCE IN MARKOV PROCESSES

VERÓNICA GONZÁLEZ-LÓPEZ — JESÚS E. GARCÍA

UNIVERSITY OF CAMPINAS, BRAZIL
UNICAMP

NOVEMBER 23 – 26, 2020



INTERNATIONAL ORGANIZING COMMITTEE

ALBA MARTINEZ-RUIZ (IASC-LARS SCIENTIFIC SECRETARY)
DAVID MUÑOZ-NEGRON (IASC-LARS CHAIRPERSON)
LUIS FIRINGUETTI-LIMONE (IASC OUTREACH OFFICER)

IASC-LARS Webinar on Computational Statistics and Data Science

STATISTICAL INFERENCE IN MARKOV PROCESSES PROGRAM



IASC- LARS SCHOOL ON COMPUTATIONAL STATISTICS AND DATA SCIENCE

The Latin American Regional Section of the International Association for Statistical Computing (IASC-LARS), the IASC-LARS School on Computational Statistics and Data Science, the International Association for Statistical Computing (IASC), and the International Statistical Institute (ISI) are pleased to invite postgraduate and undergraduate students to attend the IASC-LARS Webinar Course “Statistical Inference in Markov Processes”. The course will be team-taught by Professors Verónica González-López and Jesús E. García from the University of Campinas, Brazil, November 23-26, 2020.

The IASC-LARS Courses aim (1) to spread the knowledge base and advances in Statistical Computing in Latin American 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.

IASC-LARS Webinar on Computational Statistics and Data Science

STATISTICAL INFERENCE IN MARKOV PROCESSES PROGRAM

The purpose of the course is to introduce the notions and methods of statistical inference in Markov processes on finite alphabets with finite order. We bring recent literature on how parsimonious structures are designed in the process state space, and how they are consistently estimated. We also develop the steps for the definition of metrics that allow us to estimate that structures and to determine the discrepancy between process samples. Real applications and problems are addressed in each of the stages of the program presented below.

INSTRUCTORS: Verónica González-López (University of Campinas) and Jesús E. García (University of Campinas).

COLLABORATORS: M.L. Lanfredi Viola (Federal University of São Carlos), G.H. Tasca (University of Campinas), M. Fernández (B3), M. Cordeiro (Federal University of Technology).

AGENDA

(Local time: São Paulo – Brazil, UCT/GMT -3 hours)

Monday November 23:

session 1: 10:00 am to 11:00 am

session 2: 11:15 am to 12:15 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

session 1: 10:00 am to 11:00 am

session 2: 11:15 am to 12:15 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

session 1: 09:00 am to 10:15 am

session 2: 10:30 am to 12:15 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

session 1: 09:00 am to 10:15 am

session 2: 10:30 am to 12:15 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

References

- [1] Rissanen J (1983) A universal data compression system. *IEEE Transactions on information theory* 29(5), 656-664.
- [2] Csiszár I, Talata Z (2006) Context tree estimation for not necessarily finite memory processes via BIC and MDL. *IEEE Transactions on Information theory* 52.3: 1007-1016.
- [3] García JE, González-López VA (2017) Consistent Estimation of Partition Markov Models. *Entropy* 19 (4), 160.
- [4] Schwarz G (1978) Estimating the dimension of a model. *The annals of statistics* 6 (2), 461-464.
- [5] Cordeiro MTA, García JE, González-López VA, Mercado-Londoño SL (2019) Stochastic profile of Epstein-Barr virus in nasopharyngeal carcinoma settings. *4open*, 2, 25.
- [6] Csiszár I, Talata Z (2006) Consistent estimation of the basic neighborhood of Markov random fields. *The Annals of Statistics* 123-145.
- [7] Fernández M, García JE, González-López VA (2018), A copula-based partition Markov procedure. *Communications in Statistics-Theory and Methods* 47 (14) 3408-3417.
- [8] Cordeiro MTA, García JE, González-López VA, Mercado-Londoño SL (2020), Partition Markov model for multiple processes. *Mathematical Methods in the Applied Sciences*. 43 (13) 7677-7691.
- [9] García JE, Gholizadeh R, González-López VA (2018), A BIC-based consistent metric between Markovian processes. *Applied Stochastic Models in Business and Industry* 34(6), 868-878.
- [10] Fernández M, García J E, Gholizadeh R, González-López VA (2020), Sample Selection Procedure in Daily Trading Volume Processes. *Mathematical Methods in the Applied Sciences* 43 (13) 7537-7549.
- [11] Cordeiro MTA, García JE, González-López VA, Mercado-Londoño SL (2019), Classification of autochthonous dengue virus type 1 strains circulating in Japan in 2014. *4open*, 2, 20.
- [12] García J E, Gholizadeh R, González-López VA (2018), Stochastic distance between Burkitt lymphoma/leukemia strains. In *Demography and Health Issues* (pp. 143-153). Springer, Cham.

IASC-LARS Webinar on Computational Statistics and Data Science

STATISTICAL INFERENCE IN MARKOV PROCESSES REGISTRATION PROCEDURE

The IASC-LARS Webinar “Statistical Inference in Markov Process” will be held virtually using the platform [GoToWebinar](#) from November 23-26, 2020. The official language is English. The deadline for early registration is **October 20, 2020**. The course is free of charge.

Please complete the [IASC-LARS Webinar Registration Form](#) and send it to iasc.lars.iasc@gmail.com.

To become a IASC-LARS member, please complete the Membership Application Form at <https://www.isi-web.org/index.php/membership/individual-membership/iasc>.

All participants are expected to adhere to the ISI Community Principles and Conduct Policy (<https://www.isi-web.org/index.php/about-isi/policies/community-conduct>).

More information about GoToWebinar please visit <https://www.gotomeeting.com/webinar>. GoToWebinar application is also available for iOS, Android and Windows Phone: <https://support.goto.com/webinar/help/gotowebinar-for-mobile-devices-g2w050033>.

IASC-LARS Webinar on Computational Statistics and Data Science

STATISTICAL INFERENCE IN MARKOV PROCESSES GOTOWEBINAR PLATFORM

To attend this virtual course:

1. Please complete the IASC-LARS Webinar Registration Form and send it to iacsc.lars.iasc@gmail.com as soon as you decide to attend the course.
2. You will receive an email with an URL link to register in the GTW platform to attend the webinar. Please, provide all the requested information. Later you will receive an email confirmation registration.
3. One hour before starting the course, you will receive an invitation link and a webinar's ID to join the course by GoToWebinar.
4. You can also go to <https://www.gotomeeting.com/webinar/join-webinar> and join the webinar entering the 9-digit Webinar ID and your email in the Join a Webinar window.

GOTOWEBINAR

Join a Webinar

Attend for Free. No account required.

5. Please, try to connect about 30 minutes before the start of the course.
6. We recommend joining via a high speed and wired connection and to use a USB headset for best sound quality.
7. Before to attend the course, please visit the webpage <https://www.gotomeeting.com/webinar/join-webinar> and see the video “GoToWebinar Attendee Quick Start”. For more information, you can also visit the YouTube channel <https://www.youtube.com/user/gotowebinar>.
8. If you have any questions or you need any help, please do not hesitate to contact us at iacsc.lars.iasc@gmail.com.