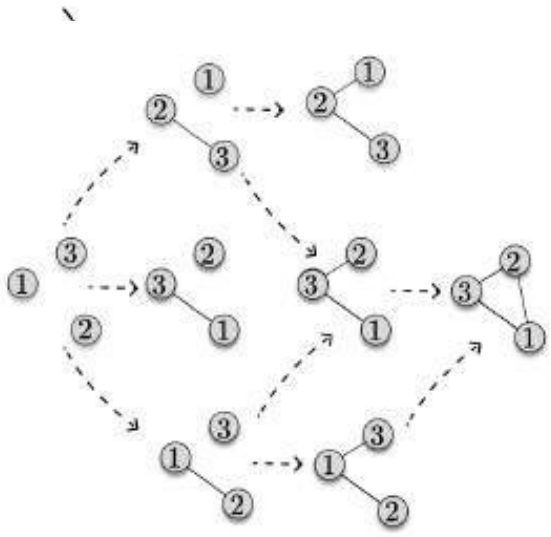
Colóquio Interinstitucional

Modelos Estocásticos e Aplicações

Quarta-feira, 27 de novembro de 2024

Programa

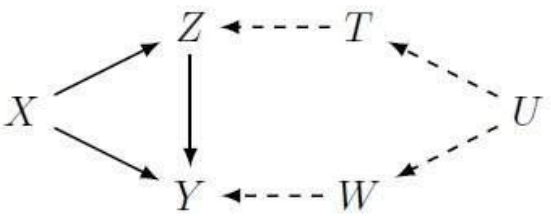
14:00 - 15:20 – **Diego Mesquita (FGV EMAp)**



When do GFlowNets learn the right distribution?

Generative Flow Networks (GFlowNets) are an emerging class of sampling methods for distributions over discrete and compositional objects, e.g., graphs. In spite of their remarkable success in problems such as drug discovery and phylogenetic inference, the question of when and whether GFlowNets learn to sample from the target distribution remains underexplored. To tackle this issue, we first assess the extent to which a violation of the detailed balance of the underlying flow network might hamper the correctness of GFlowNet's sampling distribution. In particular, we demonstrate that the impact of an imbalanced edge on the model's accuracy is influenced by the total amount of flow passing through it and, as a consequence, is unevenly distributed across the network. We also argue that, depending on the parameterization, imbalance may be inevitable. In this regard, we consider the problem of sampling from distributions over graphs with GFlowNets parameterized by graph neural networks (GNNs) and show that the representation limits of GNNs delineate which distributions these GFlowNets can approximate. Lastly, we address these limitations by proposing a theoretically sound and computationally tractable metric for assessing GFlowNets, experimentally showing it is a better proxy for correctness than popular evaluation protocols.

15:40 - 17:00 – **Widemberg Nobre (IM-UFRJ)**



A Bayesian View on Causal Inference Under Mis-specified Models

In this talk, I will present recent advances in Bayesian analysis for causal inference under mis-specied models. A model is said to be correctly specified if the posterior distribution converges to the correct distribution of the data generating process as the sample size increases. Conversely, a model is considered mis-specified if it does not meet this criterion. I will focus on propensity score regression and discuss doubly robust estimation in the contexts of single-level and multi-level models. Synthetic analyses and an application to Tuberculosis data in Brazil will be discussed in order to illustrate the proposed method.

17:00 - 18:00 – **Discussão e lanche**

Local

Escola de Matemática Aplicada – FGV
Sala 408
Edifício Sede
Botafogo

Os participantes dos seminários **não** poderão acessar as dependências da FGV usando bermuda, chinelos, blusa modelo top ou cropped, minissaia ou camiseta regata.

Contatos

Americo Cunha (UERJ)	americo@ime.uerj.br
Evaldo M. F. Curado (CBPF)	evaldo@cbpf.br
João Batista M. Pereira (UFRJ)	joao@dme.ufrj.br
Leandro P. R. Pimentel (UFRJ)	leandro@im.ufrj.br
Maria Eulália Vares (UFRJ)	eulalia@im.ufrj.br
Nuno Crokidakis (UFF)	nuno@mail.if.uff.br
Roberto I. Oliveira(IMPA)	rimfo@impa.br
Simon Griffiths (PUC-Rio)	simon@mat.puc-rio.br
Yuri F. Saporito (FGV-EMAp)	yuri.saporito@fgv.br

Realização: CBPF, FGV EMAp, impa, PUC Rio, UERJ, uff

Apoio: FAPERJ

www.im.ufrj.br/ coloquiomea/