

Estatística Computacional

ICMC-USP Projeto Final

O trabalho final deve ser no formato de um artigo científico. Deverá ser escrito em Latex (sugestão: usar o Overleaf) e conter as seções a seguir.

1 Introdução

Descrição do problema, dos dados a serem utilizados e objetivos da análise.

2 Metodologia

Descrição do problema. Descrição da(s) metodologia(s) utilizada(s). Aqui vai um pouco de teoria, fórmulas e descrição dos algoritmos.

3 Resultados/Aplicações

Aplique a(s) metodologia(s) da seção anterior a um ou mais conjuntos de dados. Forneça detalhes numéricos sobre as especificações das distribuições, dos algoritmos utilizados e sobre os custos computacionais.

4 Conclusão/Discussão

Breve resumo do que foi feito e principais resultados. Comente possíveis dificuldades de programação ou convergência se for o caso. Fale sobre possíveis questões em aberto, extensões da metodologia a problemas mais complexos e propostas de possíveis trabalhos futuros.

Observações

- Todas as figuras e tabelas que sejam incluídas no texto precisam ser comentadas.
- Use o R e dados disponibilizados publicamente. Inclua a fonte dos dados.
- Seus resultados devem poder ser reproduzidos por outro pesquisador usando as mesmas ferramentas.
- O trabalho é individual. Cópia direta de códigos computacionais, tabelas e figuras é considerado plágio.
- O formato final deve ser Adobe PDF. Entregar também o arquivo .tex e um arquivo com os comandos utilizados. Os nomes dos arquivos devem conter somente o número USP seguido das extensões “.pdf”, “.tex” e “.r”.
- Não copie figuras e tabelas prontas de outros trabalhos ou páginas da internet, faça suas próprias figuras e tabelas.

Alguns possíveis temas/metodologias (com possíveis interseções)

- Explore outros métodos de amostragem por importância (IS), Adaptive IS, Pareto smoothed importance sampling. Encontre pacotes no R.
- Explore outros métodos de rejeição, e.g. rejeição adaptativa (ARS). Encontre pacotes no R.
- Uso do Stan em problemas estatísticos.
- Inferência em modelos com misturas de distribuições.
- Ilustrando a redução de variância por condicionamento.
- Monte Carlo sequencial em espaço de estados.
- Monte Carlo Hamiltoniano e aplicações.
- Algoritmo de troca (exchange algorithm).

Algumas fontes de dados

- World Bank Open Data <https://data.worldbank.org/>
- European data <https://data.europa.eu>
- Awesome Public Datasets <https://github.com/awesomedata/awesome-public-datasets>
- Organização Mundial da Saúde <https://www.who.int/data/gho/>
- Google Public Data <https://www.google.com/publicdata/directory>
- U.S. Census Bureau <http://www.census.gov/data.html>
- Portal de Dados Abertos do Brasil <https://dados.gov.br/>