

FLS 6443 - Introdução à Inferência Bayesiana para as Ciências Sociais

Leonardo Sangali Barone
Pós-doutorando no DCP-USP
leonardo.barone@usp.br

1o Semestre de 2019

Conteúdo

1	Apresentação	2
2	Créditos, idioma, data de início, duração, dia, horário e local	3
3	Requisitos	3
4	Dedicação e principais leituras	3
5	Panorama do curso	4
6	Detalhamento das aulas leituras e calendário	5
7	Avaliação	7

1 Apresentação

A **abordagem clássica** em inferência estatística é predominante em cursos de métodos quantitativos em ciências sociais. A **abordagem Bayesiana** é frequentemente ignorada ou mesmo mencionada.

O curso oferece a estudantes de pós-graduação em ciências sociais uma iniciação **breve e inteligível** à inferência estatística Bayesiana com ênfase em aplicações para ciência política, sociologia e economia.

O curso é direcionado para qualquer estudante de pós-graduação interessada na aplicação de modelos estatísticos e em qualquer etapa dos programas de mestrado ou doutorado. Há alguns requisitos desejáveis para participação (ver seção 3) e diferentes possibilidades de aprofundamento nos conteúdos do curso (ler seção 4), mas não há nenhuma exigência relativa ao conhecimento prévio em estatística, computação ou métodos quantitativos.

Diferentemente dos cursos convencionais de inferência estatística, básicos ou avançados, nosso, objetivo **não** será adquirir um repertório de soluções pré-manufaturadas, como teste de hipóteses ou regressão linear, para aplicar a problemas de pesquisa empírica de características diversas. No meio do caminho do aprendizado convencional costuma-se perder o vínculo entre teoria de probabilidade e as técnicas de inferência. O mal uso de estatística nas ciências sociais está diretamente relacionado à aplicação não refletida de tais soluções e à perda do vínculo entre os dois "ramos" da estatística.

Nosso objetivo será compreender os **fundamentos** da construção de modelos estatísticos em ciências sociais sob a ótica Bayesiana. Boa parte do nosso trabalho no curso será reconectar, de maneira inteligível, probabilidade e inferência. A aquisição ferramental teórico e técnico necessário para refazer essa conexão é o que permite, no futuro, construir aplicações inovadoras de estatística para as ciências sociais, simples ou complexos.

Resolver problemas de inferência complexos, porém, não era factível pela abordagem Bayesiana até o final do século XX. O ressurgimento tardio da inferência Bayesiana foi impulsionado pelo aparecimento de algoritmos de simulação – conhecidos como Monte Carlo via Cadeias de Markov (MCMC) – cuja aplicação foi viabilizada pelo crescimento da capacidade computacional disponível para pesquisa. Para tirar proveito da análise Bayesiana, é preciso, pois, aprender a implementar os métodos MCMC. Assim, o objetivo na etapa final do curso será aprender um pouco sobre computação estatística e ler sobre aplicações em modelos complexos.

Finalmente, o aprendizado da abordagem Bayesiana terá dois efeitos colaterais bastante desejáveis que convém notar: (1) a realização uma breve reflexão epistemológica sobre a

construção e aplicação de modelos estatísticos em ciências sociais, (2) aquisição do hábito, repertório e confiança para ler textos intermediários e avançados de estatística.

2 Créditos, idioma, data de início, duração, dia, horário e local

Créditos: 4 créditos

Idioma: Português

Data de Início: 18 de Março de 2019

Duração: 6 aulas

Dia da Semana: Segunda-feira

Horário: 14h às 18h

Local: a definir

3 Requisitos

Formalmente, não há nenhuma disciplina que seja requisito para o curso. Entretanto, é altamente desejável que todas as participantes tenham conhecimento básico de Probabilidade e alguma familiaridade com notação matemática. É suficiente para participar do curso:

- (a) haver cursado uma disciplina de Métodos Quantitativos no DCP-USP ou na IPSA-USP Summer School; ou
- (b) ler as leituras recomendadas para o pré-curso (Aula 0).

Apesar de utilizarmos a linguagem R para a produção de modelos estatísticos durante o curso, **não** é necessário saber programar em R.

4 Dedicção e principais leituras

Vamos ler ao longo do curso capítulos selecionados de 3 livros, abaixo indicados, que serão a base para as aulas expositivas:

- (a) McElreath Richard. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Boca Raton: Chapman and Hall/CRC, XII 2015
- (b) Jackman Simon. Bayesian Analysis for the Social Sciences. Chichester, U.K: Wiley, XII 2009. 1 edition
- (c) Gelman Andrew, Carlin John B., Stern Hal S., Dunson David B., Vehtari Aki, Rubin Donald B. Bayesian Data Analysis. 2013. 3rd. (Texts in Statistical Science)

A **dedicação mínima** exigida no curso é acompanhar **pelo menos um** dos livros acima indicado em todas as aulas, além da realização das atividades de avaliação.

O excelente livro de McElreath (2015) foi escrito em linguagem bastante acessível e com um baixíssimo grau de formalização matemática. É a **leitura recomendada para quem não tem o hábito de ler livros de estatística ou matemática**, ou seja, para um público iniciante ou para iniciados em inferência clássica que queiram uma exposição intuitiva de análise bayesiana. Tem como grande vantagem apresentar a construção dos exemplos do livro em R. Trabalharemos com 6 capítulos deste livro.

O livro de Jackman (2009), que foi a base do curso oferecido pelo próprio autor na IPSA-USP Summer School de 2012, requer um formação em matemática e probabilidade mais consolidada. Tem, por outro lado, a vantagem de ter sido **escrito para a ciência política** e oferecer um contraste interessante com a prática de inferência clássica, dominante no campo. As aulas expositivas terão **rigor matemático igual a ou um pouco menor que este livro**.

Nossa **meta será acompanhar os livros de McElreath (2015) e Jackman (2009)** simultaneamente, embora a leitura exclusiva do primeiro seja suficiente para fazer bom aproveitamento do curso.

Finalmente, o livro de Gelman et al. (2013) é o mais completo dos 3, sobretudo para aplicações avançadas, mas cresce rapidamente em complexidade nos primeiros capítulos e pressupõe uma leitora com domínio mais avançado de cálculo, probabilidade e inferência clássica. Foi escrito para um público de campos científicos diversos e de iniciados. Recorreremos a este livro apenas se e quando o livro de Jackman (2009) estiver suficientemente claro. A sugestão é que estudantes mais avançadas acompanhem também este livro ao longo do curso.

5 Panorama do curso

1. Breve revisão de Probabilidade

2. Introdução à Análise Bayesiana
3. Contrastes entre Inferência Estatística Clássica e Bayesiana
4. Análise Bayesiana para modelos simples
5. Análise Bayesiana para modelos lineares
6. Simulação com Métodos de Monte Carlo
7. Monte Carlo via Cadeias de Markov (MCMC) para Análise Bayesiana
8. Aplicações avançadas em Ciências Sociais

6 Detalhamento das aulas leituras e calendário

(Pré-curso) Aula 0 - Preparação para o Curso (opcional)

Tópicos: Matemática para Cientistas Sociais; Probabilidade e Inferência; Programação em R.

Leituras recomendadas:

Moore, Siegel (2013) - caps. 1 a 4 (matemática)
Imai (2017) - caps. 6 e 7 (probabilidade e inferência)
Imai (2017) - cap. 1 (programação em R)

(18/03) - Aula 1 - Probabilidade e Introdução à Análise Bayesiana

Tópicos: Breve revisão de Probabilidade; Introdução à Análise Bayesiana.

Leituras recomendadas:

McElreath (2015) - cap. 1
Jackman (2009) - cap. 1
Gelman et al. (2013) - cap. 1

Leituras de apoio:

Ross (2014a) - caps. 1 a 8
DeGroot, Schervish (2011) - caps. 1-6

(25/03) Aula 2 - Inferência Bayesiana - Modelos simples

Tópicos: Contrastes entre Inferência Estatística Clássica e Bayesiana; Análise Bayesiana para modelos simples.

Leituras recomendadas:

McElreath (2015) - caps. 2 e 3

Jackman (2009) - caps. 1 e 2

Gelman et al. (2013) - cap. 2

Leituras de apoio

DeGroot, Schervish (2011) - caps. 7 a 9

(01/04) Aula 3 - Inferência Bayesiana - Modelos simples

(Continuação da aula anterior)

(08/04) Aula 4 - Inferência Bayesiana - Modelos lineares

Tópicos: Análise Bayesiana para modelos lineares.

Leituras recomendadas:

McElreath (2015) - caps. 4 e 5

Jackman (2009) - cap. 2

Gelman et al. (2013) - cap. 3 e 14

Leituras de apoio:

DeGroot, Schervish (2011) - cap. 11

Imai (2017) - cap. 4

(15/04) Aula 5 - Monte Carlo via Cadeias de Markov (MCMC) para Análise Bayesiana

Tópicos: Simulação com Métodos de Monte Carlo; Cadeias de Markov; Monte Carlo via Cadeias de Markov (MCMC)

Leituras recomendadas:

McElreath (2015) - cap. 8

Jackman (2009) - parte II (caps 3 a 6)

Gelman et al. (2013) - parte III (cap. 10 a 13)

Leituras de apoio:

Ross (2014b) - cap. 11

DeGroot, Schervish (2011) - cap. 12

(29/04) Aula 6 - Aplicações e Tópicos Avançados de Análise Bayesiana

Tópicos: Aplicações de Análise Bayesiana em Ciências Sociais; Tópicos Avançados.

Leituras para apresentação e resenha (escolher um capítulo, apenas):

Jackman (2009) - parte III (caps 7 a 9)

McElreath (2015) - caps. 9 a 14

Gelman et al. (2013) - parte III (cap. 10 a 13)

jel (2014) - todos os capítulos

7 Avaliação

Faremos 3 atividades **curtas** de aplicação, com construção de modelos a partir da abordagem inferencial Bayesiana. Cada uma delas valerá 20% da nota.

Na última aula do curso cada aluno deverá **expor um capítulo** dentre os indicados para a Aula 6 e entregar (algumas semanas depois) uma **resenha** sobre o capítulo, que corresponderá a 40% da nota.

- (a) Aplicação 1 (20%): construção de um modelo simples com abordagem Bayesiana (em papel e caneta)
- (b) Aplicação 2 (20%): construção de um modelo de regressão linear com abordagem Bayesiana (em R)
- (c) Aplicação 3 (20%): construção de um modelo com aplicação de MCMC (em R)
- (d) Apresentação e resenha (40%): apresentação de texto e resenha de um dos capítulos ou textos sugeridos na Aula 6

Referências

Bayesian Inference in the Social Sciences. Hoboken, New Jersey: Wiley, IX 2014. 1 edition.

DeGroot Morris H., Schervish Mark J. Probability and Statistics. Boston: Pearson, I 2011. 4 edition.

Gelman Andrew, Carlin John B., Stern Hal S., Dunson David B., Vehtari Aki, Rubin Donald B. Bayesian Data Analysis. 2013. 3rd. (Texts in Statistical Science).

Imai Kosuke. Quantitative Social Science: An Introduction. Princeton: Princeton University Press, III 2017.

Jackman Simon. Bayesian Analysis for the Social Sciences. Chichester, U.K: Wiley, XII 2009. 1 edition.

McElreath Richard. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Boca Raton: Chapman and Hall/CRC, XII 2015.

Moore Will H., Siegel David A. A Mathematics Course for Political and Social Research. VII 2013.

Ross Sheldon M. A First Course in Probability. 2014a.

Ross Sheldon M. Introduction to Probability Models. V 2014b.