

Instituto de Estudos Sociais e Políticos -
IESP-UERJ

LEGO II - Modelos de Regressão Linear e
extensões

Fernando Guarnieri

2º semestre 2026

1 Ementa

Trata-se de um curso sobre regressão linear e modelos lineares generalizados. Ao final do curso, os alunos serão capazes de construir, ajustar, representar graficamente, entender e avaliar modelos lineares e lineares generalizados, utilizando os ajustes para fazer previsões e inferências sobre quantias de interesse, incluindo efeitos causais. O curso tem caráter prático, permitindo ao aluno não só compreender os conceitos, como também se apropriar das técnicas por meio do uso de softwares de computação estatística (leia-se R).

Seguiremos o livro *Applied Regression and Other Stories*, de Andrew Gelman, Jennifer Hill e Aki Vehtari. Neste livro, a ênfase não está tanto na matemática, mas sim na computação. Isto é, os alunos aprenderão os principais pontos da regressão e da inferência estatística por meio do uso do computador como ferramenta experimental, simulando, por exemplo, dados "falsos" a partir de distribuições de probabilidade, ajustando um modelo a esses dados e verificando se os parâmetros do modelo são recuperados.

Usaremos muito o R, portanto, alguma familiaridade com sua linguagem é altamente recomendada.

2 Plano de Trabalho

Regressão é o tipo de coisa que só se aprende praticando. A expectativa é que o aluno dedique aproximadamente 10 horas semanais ao curso, entre leitura, aulas e exercícios. Os exercícios ocuparão pelo menos metade desse tempo e, durante as aulas, os alunos discutirão e trabalharão juntos na resolução de problemas.

Os alunos devem ler o texto e trazer para a aula uma pequena "resenha guiada" sobre ele e a lista de exercícios da semana resolvida. Na primeira metade da aula será apresentado um caso relacionado ao conteúdo e serão discutidas questões que surgirem sobre o texto e os exercícios. Na segunda metade, farei uma breve demonstração de análise de dados no R relacionada ao tópico tratado no texto e, em seguida, apresentarei pequenos problemas para os alunos resolverem, a fim de reforçar suas habilidades computacionais. No final da aula, os alunos discutirão, em equipe, uma questão substantiva relacionada ao tema.

O leitor atento notou que não haverá aula expositiva. Trata-se do que se chama de aprendizado ativo, em que, nas interações em classe, os alunos fazem coisas, conversam com os colegas e resolvem problemas coletivamente. As atividades pré-classe têm grande importância, e as atividades em sala de aula servem para esclarecer, aprofundar e consolidar o que foi estudado.

3 Avaliação

A avaliação será baseada em listas de exercícios (em média três problemas por aula), "resenhas guiadas" (duas ou três questões sobre o texto), participação em aula e um exame final. Os alunos devem entregar a lista de exercícios e a resenha na aula seguinte. Cada resenha entregue vale um ponto. Cada problema da lista de exercícios respondido corretamente vale um ponto; um problema respondido, mas errado, vale meio ponto; e um problema não respondido vale zero. Os alunos que participaram ativamente das atividades receberão um ponto. O exame final, que será realizado em sala de aula, consiste em questões de múltipla escolha com cada resposta correta valendo um ponto.

4 Leituras

GELMAN, Andrew; HILL, Jennifer; VEHTARI, Aki. Regression and other stories. Cambridge University Press, 2021.

Disponível em [<https://avehtari.github.io/ROS-Examples/index.html>]

Outras leituras serão indicadas conforme o tema das aulas.

5 Programa de Aulas

1. Aula 1: Ter noção dos objetivos e desafios da regressão.
2. Aula 2 Explorar dados e estar atento a questões de mensuração e ajuste; representar graficamente uma reta e conhecer algumas ferramentas matemáticas básicas e distribuições de probabilidade.
3. Aula 3: Compreender a estimação estatística e a avaliação da incerteza, bem como os problemas dos testes de hipótese na estatística aplicada. Simular modelos probabilísticos e a incerteza associada a inferências e previsões.
4. Aula 4 Distinguir interpretações descritivas e causais da regressão, compreendendo essas diferenças em seu contexto histórico.
5. Aula 5: Compreender e trabalhar com regressão linear simples com um único preditor.
6. Aula 6: Desenvolver uma compreensão conceitual do ajuste por mínimos quadrados e ser capaz de realizar esses ajustes no computador.
7. Aula 7: Realizar e compreender previsões probabilísticas e formas simples de agregação bayesiana de informação, além de ser introduzido às distribuições a priori e à inferência bayesiana.
8. Aula 8: Construir, ajustar e compreender modelos lineares com múltiplos preditores.
9. Aula 9: Compreender a importância relativa dos diferentes pressupostos dos modelos de regressão e ser capaz de verificar modelos e avaliar seu ajuste aos dados.
10. Aula 10: Aplicar a regressão linear de forma mais eficaz por meio da transformação e combinação de preditores.
11. Aula 11: Ajustar, compreender e apresentar modelos de regressão logística para dados binários.
12. Aula 12: Construir, compreender e avaliar regressões logísticas com interações e outras complexidades.
13. Aula 13: Ajustar, compreender e apresentar modelos lineares generalizados, incluindo regressão de Poisson, regressão binomial negativa, regressão logística ordenada e outros modelos.
14. Aula 14: Utilizar teoria da probabilidade e simulação para orientar decisões

de coleta de dados, sem cair na armadilha de exigir níveis irrealistas de certeza.

15. Aula 15: Utilizar pós-estratificação para generalizar resultados da amostra para a população e usar modelos de regressão para imputar dados ausentes.