

# **Análise Bayesiana com Excel & R**

AMOSTRA

# Análise Bayesiana com Excel & R

Conrad G. Carlberg



ALTA BOOKS

GRUPO EDITORIAL  
Rio de Janeiro, 2024

## Uma Olhada Rápida no Conteúdo

Prefácio .....	IX
1. Análise Bayesiana e R: Uma Visão Geral .....	1
2. Gerando Distribuições A Posteriori com a Distribuição Binomial .....	19
3. Entendendo a Distribuição Beta .....	41
4. Aproximação de Grade e a Distribuição Beta .....	65
5. Aproximação de Grade com Parâmetros Múltiplos ..	93
6. Regressão Usando Métodos Bayesianos .....	117
7. Manipulando Variáveis Nominais .....	137
8. Métodos de Amostragem MCMC .....	157
Instruções de Instalação para RStan e o Pacote <i>rethinking</i> na Plataforma Windows .....	179
Glossário .....	183
Índice .....	195

## Conteúdo Bônus para Download

Planilhas Excel

Livro: *Statistical Analysis: Microsoft Excel 2016* (PDF)  
[Conteúdo em inglês]

Para acessar materiais de bônus, por favor, registre  
seu livro em [informit.com/register](http://informit.com/register) e insira o ISBN  
9780137580989.

# Análise Bayesiana com Excel & R

Copyright © 2024 ALTA BOOKS.

Copyright © 2023 Conrad Carlberg.

ISBN: 978-85-508-1985-3

Authorized Portuguese translation of the English edition, entitled *Bayesian Analysis with Excel and R*, 1st edition, ISBN 978-0-137-58098-9 by Conrad Carlberg, published by Pearson Education Inc., publishing as Addison-Wesley Professional, which owns or controls all rights to publish and sell the same. Copyright © 2023 Conrad Carlberg. PORTUGUESE language edition published by Grupo Editorial Alta Books Ltda., Copyright © 2024 by STARLIN ALTA EDITORA E CONSULTORIA LTDA.

Impresso no Brasil – 1ª Edição, 2024 – Edição revisada conforme o Acordo Ortográfico da Língua Portuguesa de 2009.

Dados Internacionais de Catalogação na Publicação (CIP)  
(BENITEZ Catalogação Ass. Editorial, MS, Brasil)

C278a  
1.ed. Carlberg, Conrad G.  
Análise Bayesiana com Excel & R / Conrad G. Carlberg ; tradução Leandro Menegaz. – 1.ed. – Rio de Janeiro : Alta Books, 2024.  
224 p.; il.; 15,7 x 23 cm.

Titulo original: Bayesian analysis with excel and R.  
ISBN 978-85-508-2121-4

1. Estatística. 2. Inferência Bayesiana (Probabilidade).  
3. Microsoft Excel (Arquivo de computador). 4. Teorema de Bayes. I. Menegaz, Leandro. II. Título.

10-2024/19 CDD 005.3684

Índice para catálogo sistemático:  
1. Microsoft Excel : Estatística : Inferência Bayesiana : Informática 005.3684  
Aline Grazielle Benitez - Bibliotecária - CRB-1/3129

Todos os direitos estão reservados e protegidos por Lei. Nenhuma parte deste livro, sem autorização prévia por escrito da editora, poderá ser reproduzida ou transmitida. A violação dos Direitos Autorais é crime estabelecido na Lei nº 9.610/98 e com punição de acordo com o artigo 184 do Código Penal.

O conteúdo desta obra fora formulado exclusivamente pelo(s) autor(es).

**Marcas Registradas:** Todos os termos mencionados e reconhecidos como Marca Registrada e/ou Comercial são de responsabilidade de seus proprietários. A editora informa não estar associada a nenhum produto e/ou fornecedor apresentado no livro.

**Material de apoio e errata:** Se parte integrante da obra e/ou por real necessidade, no site da editora o leitor encontrará os materiais de apoio (download), errata e/ou quaisquer outros conteúdos aplicáveis à obra. Acesse o site [www.altabooks.com.br](http://www.altabooks.com.br) e procure pelo título do livro desejado para ter acesso ao conteúdo.

**Suporte Técnico:** A obra é comercializada na forma em que está, sem direito a suporte técnico ou orientação pessoal/exclusiva ao leitor.

A editora não se responsabiliza pela manutenção, atualização e idioma dos sites, programas, materiais complementares ou similares referidos pelos autores nesta obra.

**Produção Editorial:** Grupo Editorial Alta Books  
**Diretor Editorial:** Anderson Vieira  
**Vendas Governamentais:** Cristiane Mutús  
**Gerência Comercial:** Claudio Lima  
**Gerência Marketing:** Andréa Guatiello

**Produtora Editorial:** Isabella Gibara  
**Tradução:** Leandro Menegaz  
**Copidesque:** Rafael Souza  
**Revisão:** Denise Himpel  
**Diagramação:** Natalia Curupana  
**Revisão Técnica:** William Pereira  
(Analista e Desenvolvedor de Sistemas)



Rua Viúva Cláudia, 291 – Bairro Industrial do Jacaré  
CEP: 20.970-031 – Rio de Janeiro (RJ)  
Tels.: (21) 3278-8069 / 3278-8419  
[www.altabooks.com.br](http://www.altabooks.com.br) – [altabooks@altabooks.com.br](mailto:altabooks@altabooks.com.br)  
Ouvidoria: [ouvidoria@altabooks.com.br](mailto:ouvidoria@altabooks.com.br)



---

# Conteúdos

Prefácio .....	IX
<b>1 Análise Bayesiana e R: Uma Visão Geral .....</b>	<b>1</b>
Bayes Está de Volta .....	2
Sobre Estruturar A Priori .....	5
Observando o Jargão .....	6
A Priori, Verossimilhanças e A Posteriori .....	10
A Priori .....	10
A Verossimilhança .....	12
Contrastando uma Análise Frequentista com uma Bayesiana .....	13
A Abordagem Frequentista .....	13
A Abordagem Bayesiana .....	15
Resumo .....	17
<b>2 Gerando Distribuições A Posteriori com a Distribuição Binomial .....</b>	<b>19</b>
Entendendo a Distribuição Binomial .....	21
Compreendendo Algumas Funções Relacionadas .....	26
Trabalhando com Funções Binomiais do R .....	28
Usando a Função <b>dbinom</b> do R .....	29
Usando a Função <b>pbinom</b> do R .....	31
Usando a Função <b>qbinom</b> do R .....	32
Usando a Função <b>rbinom</b> do R .....	36
Lutando com a Matemática .....	36
Resumo .....	39
<b>3 Entendendo a Distribuição Beta .....</b>	<b>41</b>
Estabelecendo a Distribuição Beta no Excel .....	42
Comparando a Distribuição Beta com a Distribuição Binomial .....	44
Decodificando a Documentação de Ajuda do Excel para <b>BETA.DIST</b> ..	51
Replicando a Análise em R .....	53
Entendendo <b>dbeta</b> .....	54
Entendendo <b>pbeta</b> .....	55

Entendendo <b>qbeta</b> .....	58
Sobre Intervalos de Confiança .....	59
Aplicando <b>qbeta</b> a Intervalos de Confiança .....	61
Aplicando <b>BETA.INV</b> a Intervalos de Confiança .....	62
Resumo .....	63
<b>4 Aproximação de Grade e a Distribuição Beta</b> .....	<b>65</b>
Mais sobre Aproximação de Grade .....	65
Definindo A Priori .....	67
Usando os Resultados da Função Beta .....	68
Rastreamento a Forma e a Localização da Distribuição .....	69
Fazendo o Inventário das Funções Necessárias .....	70
Dando uma Olhada nos Bastidores .....	74
Passando das Fórmulas Subjacentes para as Funções .....	85
Comparando Funções Integradas com Fórmulas Subjacentes .....	88
Entendendo A Priori Conjugada .....	89
Resumo .....	90
<b>5 Aproximação de Grade com Parâmetros Múltiplos</b> .....	<b>93</b>
Montando o Cenário .....	95
Opções Globais .....	95
Variáveis Locais .....	96
Especificando a Ordem de Execução .....	96
Curvas Normal, Mu e Sigma .....	97
Visualizando os Arrays .....	99
Combinando Mu e Sigma .....	100
Juntando os Dados .....	101
Calculando as Probabilidades .....	103
Adicionando A Priori .....	106
Fazendo o Inventário dos Resultados .....	108
Visualizando os Resultados a partir de Diferentes Perspectivas .....	109
Resumo .....	116
<b>6 Regressão Usando Métodos Bayesianos</b> .....	<b>117</b>
Regressão à la Bayes .....	117

Análise de Amostra de Regressão .....	120
Métodos de Álgebra Matriciais .....	122
Entendendo <b>quap</b> .....	124
Continuando o Código .....	127
Um Exemplo Completo .....	128
Projetando a Regressão Múltipla .....	130
Organizando uma Regressão Múltipla Bayesiana .....	131
Resumo .....	136
<b>7 Manipulando Variáveis Nominais .....</b>	<b>137</b>
Usando Codificação Fictícia .....	139
Fornecendo Rótulos de Texto no Lugar dos Códigos .....	146
Comparando Médias de Grupo .....	153
Resumo .....	156
<b>8 Métodos de Amostragem MCMC .....</b>	<b>157</b>
Revisão Rápida da Amostragem Bayesiana .....	157
Aproximação de Grade .....	158
Aproximação Quadrática .....	161
MCMC Se Atualiza .....	165
Uma Amostra da Análise MCMC .....	166
Dados de Saída de <b>ulam</b> .....	171
Validando os Resultados .....	172
Obtendo Gráficos de Traço .....	174
Resumo e Considerações Finais .....	175
<b>Instruções de Instalação para RStan e o Pacote <i>rethinking</i> na</b>	
<b>Plataforma Windows .....</b>	<b>179</b>
<b>Glossário .....</b>	<b>183</b>
<b>Índice .....</b>	<b>195</b>

## Conteúdo Bônus para Download

Planilhas Excel

Livro: Statistical Analysis: Microsoft Excel 2016 (PDF) [conteúdo em inglês]

Para acessar os materiais bônus, registre seu livro em [informit.com/register](http://informit.com/register) e insira o ISBN 9780137580989.

## Sobre o Autor

**Conrad Carlberg** é um especialista reconhecido nos EUA em análise quantitativa, análise de dados e aplicativos de gerenciamento, como Microsoft Excel, SAS e Oracle. Ele possui um Ph.D. em estatística pela Universidade do Colorado e recebeu várias vezes a designação Excel MVP da Microsoft. Ele é autor de vários livros, incluindo *Business Analysis with Microsoft Excel*, Quinta Edição, *Statistical Analysis: Microsoft Excel 2016*, *Regression Analysis Microsoft Excel* e *R for Microsoft Excel Users* [todos sem publicação no Brasil].

Carlberg é natural do sul da Califórnia. Após a faculdade, mudou-se para o Colorado, onde trabalhou para uma sucessão de startups e fez pós-graduação. Ele passou dois anos no Oriente Médio, ensinando ciência da computação e esquivando-se de camelos mal-humorados. Depois de terminar a pós-graduação, Carlberg trabalhou no Oeste dos EUA (em uma Baby Bell) em gerenciamento de produtos, e na Motorola.

Em 1995, Conrad abriu uma pequena empresa de consultoria ([www.conradcarlberg.com](http://www.conradcarlberg.com)), que fornece serviços de design e análise para empresas que desejam orientar suas decisões de negócios por meio de análises quantitativas — abordagens que hoje agrupamos sob o termo “analítica”. Ele gosta de escrever sobre essas técnicas e, em particular, como executá-las usando o aplicativo de análise numérica mais popular do mundo, o Microsoft Excel.



---

# Prefácio

Há neste livro inúmeros temas que desejo que você conheça de antemão. Mas, se já estiver familiarizado com a terminologia e conceitos como amostragem Hamiltoniana de Monte Carlo, pares conjugados e distribuições a posteriori, então este livro provavelmente não é para você. Você já sabe muito sobre esses tópicos e, se precisar de mais, sabe onde encontrar.

Por outro lado, se você não se sentir à vontade com o propósito de amostras aleatórias, a interface de usuário do R e por que você pode querer trabalhar com valores corrigidos pela média em vez de valores brutos, então é possível que este livro ofereça algo que lhe interesse. Tanto este livro quanto eu presumimos que você tenha alguma experiência em análise estatística — digamos, no nível introdutório da faculdade, onde você pode esperar estudar um pouco da teoria da probabilidade e como ela se aplica à avaliação de médias, variâncias e correlações amostrais. Especialmente se você estudou esses problemas no passado, estará em melhor posição para entender como a análise Bayesiana difere das abordagens tradicionais e como ela funciona no contexto das funções e pacotes encontrados em R. E se você sentir que precisa de algum trabalho de atualização na análise estatística tradicional, a Editora está disponibilizando para download um e-book intitulado *Statistical Analysis: Microsoft Excel 2016* [sem publicação no Brasil]. Você encontrará detalhes sobre como obter esse livro no final deste Prefácio [conteúdo em inglês].

Você é experiente. Provavelmente sabe algo sobre o conhecimento prévio da análise Bayesiana que eu tinha em mente quando expus os tópicos que gostaria que este livro abordasse. Pareceu-me que o mundo já tinha muitos livros sobre estatística e metodologia experimental: mais um não vai adiantar muito. Algo semelhante pode ser dito sobre o uso de sintaxe e dicção que R reconhece: já temos tantos textos elementares a intermediários em R quanto precisamos.

O que realmente precisávamos, pensei, era uma fonte de informação que conectasse os recursos simplistas do VBA (a linguagem de programação historicamente oferecida pelo Microsoft Excel para dar ao usuário mais controle sobre o aplicativo) com os recursos mais sofisticados de linguagens de programação como R e C.

Da mesma forma, estávamos perdendo informações sobre três tipos básicos de amostragem que vão desde o tipo simplista e univariado de

análise categórica que você encontra em textos de graduação até os complexos métodos de amostragem usados por técnicas como aproximação quadrática e Monte Carlo via Cadeias de Markov [em inglês, *Markov Chain Monte Carlo* — MCMC]. Richard McElreath escreveu e forneceu ao R funções auxiliares que facilitam a tarefa de projetar, escrever e instalar o código que faz o trabalho pesado para você.

Fiz o que pude neste livro para alavancar as habilidades do Excel que você já desenvolveu nas áreas de gerenciamento de funções, manuseio de dados e design de gráficos e plotagens. Chegará ao ponto em que você verá que o Excel também lida com as ferramentas necessárias de cálculo na forma de argumentos de função — embora de forma mais lenta e desajeitada. Logo depois, você verá como as três abordagens fundamentais para construir distribuições a posteriori por amostragem são, de fato, soluções maravilhosamente criativas para o mesmo problema.

Agora, veremos como eu proponho nos levar até lá.

## Capítulo 1: Análise Bayesiana e R: Uma Visão Geral

Quando abordei a Editora pela primeira vez para escrever este livro, saí das discussões um pouco desanimado. Os editores e seus conselheiros foram educados e muito bons em ouvir, mas não achei que recebi muito encorajamento. Em particular, eles queriam saber por que eu gostaria de escrever este livro.

Boa pergunta. Eu tinha vários motivos em mente, mas não era fácil articulá-los. Ainda assim, fiz isso e, aparentemente, fiz com sucesso porque, bem, veja o que você está segurando. E essas razões faziam sentido como um ponto de partida, mas vou me limitar às duas primeiras que me ocorreram:

- Por que você gostaria de lê-lo? Existem vários motivos, mas se você for como a maioria de nós, usará o Microsoft Excel para a maioria das finalidades numéricas, embora o Excel tenha sido projetado como um mecanismo de cálculo de uso geral. Você pode ter ficado longe da análise Bayesiana porque ouviu que o Excel é comparativamente lento. E você está certo: tanto por problemas de software quanto de hardware, houve um tempo em que você teve que esperar e esperar por uma solução para o problema que você apresentava ao software Bayesiano. Não mais. Agora você pode obter uma resposta em um período de tempo razoável e sem fazer suposições com as quais não se sinta confortável.
- As pessoas com quem trabalho estavam usando palavras familiares de maneiras desconhecidas; termos como *a priori*, *verossimilhança* e *parâmetro* em contextos que não pareciam se encaixar. Eu queria saber mais sobre o que elas estavam dizendo. Mas eu precisava de um ponto de partida e, como estava bastante familiarizado com os recursos numéricos do Excel, decidi trabalhar a partir da plataforma do Excel e em direção a uma plataforma baseada em R. É verdade que o Excel é comparativamente lento e não possui muitas funções que você gostaria de ter em uma plataforma orientada à Bayesiana. Mas, para certos problemas, o Excel funciona muito bem e retorna resultados precisos em um curto espaço de tempo. Bem; podemos trabalhar a partir daí.

É sobre isso que se trata o Capítulo 1. Sigamos em frente.