

*Empresa Brasileira de Pesquisa Agropecuária
Secretaria de Administração Estratégica
Ministério da Agricultura e do Abastecimento*

Introdução aos Modelos de Regressão Linear e Não-Linear

Geraldo da Silva e Souza

*Serviço de Produção de Informação
Brasília, DF
1998*

Exemplares desta publicação podem ser adquiridos na:

Embrapa Informação Tecnológica

Parque Estação Biológica (PqEB) Av. W3 Norte (final)
70770-901 Brasília, DF
Tel.: (61) 3340-9999
Fax: (61) 3340-2753
vendas@sct.embrapa.br
www.sct.embrapa.br/liv

Empresa Brasileira de Pesquisa Agropecuária

Secretaria de Gestão e Estratégia
Edifício-sede da Embrapa
Parque Estação Biológica (PqEB) Av. W3 Norte (final)
70770-901 Brasília, DF
Tel.: (61) 3348-4466 – 3448-4194
Fax: (61) 3448-4319 – 3347-4480

Coordenação editorial e gráfica

Embrapa Informação Tecnológica

Tratamento editorial

Francisco C. Martins

Capa

Carlos Eduardo Felice Barbeiro

Revisão de texto

Maria Lúcia Del Fiaco Rocha

Revisão de provas

Ana Carolina Silva Cirotto

Normalização bibliográfica

Rosa Maria e Barros Ferreira

1ª edição

1ª impressão (1998): 1.000 exemplares

2ª impressão (2001): 500 exemplares

3ª impressão (2007): 300 exemplares

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Informação Tecnológica

Souza, Geraldo da Silva e

Introdução aos modelos de regressão linear e não-linear / Geraldo da Silva e Souza. — Brasília: Embrapa-SPI / Embrapa-SEA, 1998.

505 p.

ISBN 85-7383-020-4

1. Análise Estatística. I. Título

CDD 519.536

© Embrapa 1998

Prefácio

A idéia de escrever este livro introdutório em regressão linear e não-linear surgiu da ausência de um texto semelhante em nossa língua. Este material cobre três cursos que tenho a oportunidade de, freqüentemente, ministrar no programa de mestrado em Estatística da Universidade de Brasília: Métodos Estatísticos, Modelos Lineares e Regressão Não-Linear. Para o curso de Métodos, mais voltado para aplicações, omito os Capítulos 1, 3 e 5. No Capítulo 2, a ênfase é dada somente às aplicações do modelo linear de posto completo. Nesse contexto, um curso de um semestre em aplicações do modelo linear pode ser levado a efeito cobrindo a parte relevante do Capítulo 2 (Seção 2.1), o Capítulo 4 e a parte do Capítulo 6 sobre conceitos básicos de regressão não-linear. Tenho usado este material com sucesso nos cursos da UnB em associação com o uso intensivo do computador, particularmente com o SAS (Statistical Analysis System). O curso de Modelos Lineares inclui o Capítulo 3 e todo o material não abordado no curso de Métodos. Os Capítulos 5-8 servem ao propósito de introdução, teórica e aplicada, à regressão não-linear. Desse modo, serve de base para a leitura de textos mais avançados.

Em minha exposição, procuro ilustrar com exemplos os principais conceitos. Ao final de todos os capítulos, apresento um conjunto de problemas. Uns fáceis, outros mais complexos e envolventes e que complementam a teoria desenvolvida no texto. Encorajo o leitor a gastar algum tempo com a solução desses problemas. Para

um bom aproveitamento do texto, o leitor deve estar bem familiarizado com o Cálculo de Várias Variáveis e com a Álgebra Linear. O conhecimento dos fundamentos de programação SAS também é desejável. Todas as ilustrações e rotinas desenvolvidas no texto foram levadas a efeito com o uso do SAS.

Escrevi este livro ao longo de três anos, em diferentes locais. Particularmente importante nesse período, foi o ano que passei como professor visitante do Departamento de Economia da Universidade da Carolina do Norte em Chapel Hill. Muito do que aqui apresento foi resultado de edificantes conversações com A. R. Galant, meu amigo e paciente professor.

Escrever um livro, por mais simples que seja, é uma tarefa árdua que demanda bastante sacrifício e o esforço de muitos. Nesse contexto, sou grato a minha família pela paciência e compreensão pelos inúmeros fins-de-semana perdidos. Também sou grato aos meus estudantes, particularmente a George Freitas von Borries. As rotinas para o cálculo de regressões robustas e de “bootstrap” fazem parte de sua dissertação de mestrado.

Finalmente agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo suporte financeiro parcial e à Embrapa pelo interesse em publicar este livro.

Geraldo da Silva e Souza

Janeiro de 1998
Brasília, DF

Sumário

1	Resultados Preliminares	1
1.1	Tópicos em Álgebra Linear	1
1.2	Extremos de Formas Quadráticas	16
1.3	A Normal Multivariada	17
1.4	Formas Quadráticas	27
1.5	Seqüências de Variáveis Aleatórias	33
1.6	Exercícios	44
1.7	Referências	52
2	Modelos de Regressão Linear	53
2.1	Modelos de Posto Completo	53
2.1.1	Introdução	53
2.1.2	O Princípio dos Mínimos Quadrados	57
2.1.3	Propriedades Estatísticas e Distribucionais	63
2.1.4	Regiões de Confiança	66
2.1.5	Testes de Hipóteses Lineares	72
2.1.6	Análise de Variância na Regressão Linear	77
2.1.7	Determinação da Potência de Testes Estatísticos	80
2.1.8	Estimação e Testes de Funções Não-Lineares	84
2.1.9	Mínimos Quadrados Generalizados	87
2.2	Modelos de Posto Incompleto	95

2.2.1	Propriedades Estatísticas e Distribucionais	98
2.2.2	Regiões de Confiança	100
2.2.3	Testes de Hipóteses Lineares	101
2.2.4	Análise de Variância no Modelo Linear de Posto Incompleto	105
2.2.5	Análise de Ensaios em Blocos Incompletos	105
2.3	Exercícios	127
2.4	Referências	136
3	O Modelo Linear Geral	139
3.1	Introdução	139
3.2	Erros Independentes	139
3.3	Erros com Variância Geral	155
3.3.1	Modelos com Resposta Binária	160
3.3.2	Erros Definidos por Séries Temporais Estacionárias	164
3.4	Variável Dependente Defasada	174
3.4.1	Variáveis Instrumentais	174
3.4.2	Estimador de Hatanaka	175
3.4.3	Estimador de Máxima Verossimilhança	177
3.4.4	Estimador Não-Linear	177
3.5	Exercícios	183
3.6	Referências	186
4	Diagnóstico em Regressão	191
4.1	Introdução	191
4.2	Teste da Especificação $E(Y) = X\beta$	192
4.3	Análise Residual	201
4.3.1	Constância da Variância	203
4.3.2	Diagnóstico de Linearidade da Resposta	218
4.3.3	Correlação Serial e Espacial	220
4.3.4	Observações Influentes	239
4.3.5	Diagnóstico de Normalidade	265

4.4	Multicolinearidade	278
4.4.1	Introdução	278
4.4.2	Diagnóstico	281
4.4.3	Medidas Corretivas	287
4.5	Exercícios	288
4.6	Referências	294
5	Modelos Não-Lineares: Teoria	297
5.1	Introdução	297
5.2	A Regressão Não-Linear	298
5.2.1	Estimação	299
5.2.2	Testes de Hipóteses Estatísticas	305
5.3	Seqüências Cesaro Somáveis	306
5.4	Consistência Forte	313
5.5	Normalidade Assintótica	316
5.6	Os Testes de Wald, ML e RV	329
5.6.1	Teste de Wald	329
5.6.2	Teste dos Multiplicadores de Lagrange	331
5.6.3	Teste da Razão de Verossimilhança	334
5.6.4	Observações Gerais Sobre as Estatísticas W_n , R_n e RV_n	336
5.7	Distribuição do Estimador Restrito	337
5.8	Exercícios	339
5.9	Referências	341
6	Modelos Não-Lineares: Prática	343
6.1	Introdução	343
6.2	Condições de Regularidade	344
6.3	Estimadores de MQO Não-Lineares	346
6.4	A Geometria de MQO Não-Lineares	347
6.5	Propriedades Distribucionais	348
6.6	Cálculo de MQO Não-Lineares	349
6.7	Estimação e Testes de Hipóteses	367
6.7.1	Procedimento de Wald	367

6.7.2	Teste RV	371
6.7.3	Teste ML	374
6.8	Potência de Testes Estatísticos	377
6.9	Resíduos em Séries de Tempo	378
6.10	Exercícios	389
6.11	Referências	393
7	Diagnóstico no Modelo Geral	395
7.1	Introdução	395
7.2	Teste da Especificação $E(Y) = f(\theta^0)$	396
7.3	Constância da Variância	400
7.4	Observações Influentes	412
7.5	Medidas de Curvatura	418
7.5.1	Cálculo de Curvatura	423
7.6	O Viés de Box	436
7.7	Exercícios	442
7.8	Referências	443
8	Métodos Empíricos	445
8.1	Introdução	445
8.2	O Método “Bootstrap”	446
8.3	Intervalos “Bootstrap”	451
8.4	O “Bootstrap” na Regressão	465
8.5	Exercícios	480
8.6	Referências	483