



# Análise Estatística com Excel

Para  
**leigos**

Tradução da  
3ª Edição

**Joseph Schmuller**

PhD



ALTA BOOKS  
EDITORA  
Rio de Janeiro, 2018

# Sumário Resumido

Introdução .....	1
<b>Parte 1: Começando a Análise Estatística com Excel .....</b>	<b>7</b>
<b>CAPÍTULO 1:</b> Avaliando Dados no Mundo Real .....	9
<b>CAPÍTULO 2:</b> Compreendendo as Capacidades Estatísticas do Excel.....	31
<b>Parte 2: Descrevendo Dados.....</b>	<b>61</b>
<b>CAPÍTULO 3:</b> Mostre e Informe: Colocando Dados em um Gráfico .....	63
<b>CAPÍTULO 4:</b> Encontrando seu Centro .....	93
<b>CAPÍTULO 5:</b> Desviando da Média .....	109
<b>CAPÍTULO 6:</b> Seguindo Padrões e Constantes.....	127
<b>CAPÍTULO 7:</b> Resumindo Tudo .....	143
<b>CAPÍTULO 8:</b> O Que é Normal? .....	167
<b>Parte 3: Tirando Conclusões a Partir dos Dados .....</b>	<b>179</b>
<b>CAPÍTULO 9:</b> O Jogo da Confiança: Estimativas .....	181
<b>CAPÍTULO 10:</b> Teste de Hipótese com Uma Amostra.....	195
<b>CAPÍTULO 11:</b> Teste de Hipótese com Duas Amostras .....	211
<b>CAPÍTULO 12:</b> Testando Mais de Duas Amostras .....	243
<b>CAPÍTULO 13:</b> Testes Um Pouco Mais Complicados.....	273
<b>CAPÍTULO 14:</b> Regressão: Linear e Múltipla .....	287
<b>CAPÍTULO 15:</b> Correlação: Ascensão e Queda das Relações .....	325
<b>Parte 4: Probabilidade.....</b>	<b>347</b>
<b>CAPÍTULO 16:</b> Introdução à Probabilidade .....	349
<b>CAPÍTULO 17:</b> Mais sobre Probabilidade.....	373
<b>CAPÍTULO 18:</b> Carreira de Modelo .....	387
<b>Parte 5: A Parte dos Dez.....</b>	<b>407</b>
<b>CAPÍTULO 19:</b> Dez Dicas e Armadilhas Estatísticas e Gráficas .....	409
<b>CAPÍTULO 20:</b> Dez Assuntos (Treze, Na Verdade) Que Não Se Encaixaram em Nenhum Outro Capítulo .....	417
<b>APÊNDICE A:</b> Quando Sua Planilha é um Banco de Dados.....	449
<b>APÊNDICE B:</b> A Análise de Covariância.....	465
<b>Índice.....</b>	<b>481</b>

# Sumário

INTRODUÇÃO .....	1
Sobre Este Livro .....	2
O Que Pode Ser Ignorado com Segurança .....	2
Penso que.....	2
Como Este Livro Está Organizado .....	3
Parte 1: Começando a Análise Estatística Com o Excel .....	3
Parte 2: Descrevendo Dados .....	3
Parte 3: Tirando Conclusões a Partir dos Dados .....	3
Parte 4: Probabilidade .....	4
Parte 5: A Parte dos Dez. ....	4
Apêndice A: Quando Sua Planilha É um Banco de Dados .....	4
Apêndice B: A Análise de Covariância .....	4
Apêndice Bônus 1: Quando Seus Dados Vivem em Outro Lugar .....	5
Apêndice Bônus 2: Dicas para Professores (E Alunos) .....	5
Além Deste Livro .....	5
Ícones Usados Neste Livro .....	5
De Lá para Cá, Daqui para Lá. ....	6
PARTE 1: COMEÇANDO A ANÁLISE ESTATÍSTICA COM EXCEL .....	7
CAPÍTULO 1: <b>Avaliando Dados no Mundo Real</b> .....	9
As Noções Estatísticas (e Relacionadas) Que Você Precisa Conhecer .....	10
Amostras e Populações .....	10
Variáveis: Dependentes e independentes. ....	11
Tipos de dados .....	12
Uma pequena probabilidade. ....	13
Estatística Inferencial: Testando Hipóteses .....	14
Hipóteses nulas e alternativas. ....	15
Dois tipos de erro .....	16
O Que Há de Novo no Excel 2013? .....	18
O Que Há de Velho no Excel 2013? .....	21
Conhecendo os Fundamentos .....	23
Células com preenchimento automático .....	23
Referenciando células. ....	26
O Que Há De Novo Nesta Edição? .....	28

<b>CAPÍTULO 2: Compreendendo as Capacidades Estatísticas do Excel</b> .....	31
Começando.....	32
Preparação para as Estatísticas.....	34
Funções de planilha no Excel.....	34
Acessando rapidamente as funções estatísticas.....	37
Funções matriciais.....	38
O que há em um nome? Um leque de possibilidades.....	41
Criando suas próprias fórmulas matriciais.....	49
Utilizando ferramentas de análise de dados.....	50
Acessando as Funções Mais Usadas.....	54
Para Usuários de Mac.....	55
A Faixa de Opções.....	55
Ferramentas de análises de dados.....	57
<b>PARTE 2: DESCREVENDO DADOS</b> .....	61
<b>CAPÍTULO 3: Mostre e Informe: Colocando Dados em um Gráfico</b> .....	63
Por que Usar Gráficos?.....	64
Alguns Princípios Básicos.....	65
As Capacidades Gráficas do Excel.....	66
Inserindo um gráfico.....	66
Transformando-se em um Colunista.....	67
Empilhando as Colunas.....	71
Mais uma coisa.....	72
Fatiando a Pizza.....	72
Uma palavra do sábio.....	74
Desenhando a Linha.....	75
Adicionando um Minigráfico (Sparkline).....	78
Levantando a Barra.....	80
A Coisa se Complica.....	82
Encontrando Outro Uso para o Gráfico de Dispersão.....	86
Power View!.....	87
Para Usuários de Mac.....	89
<b>CAPÍTULO 4: Encontrando seu Centro</b> .....	93
Média: A Sabedoria dos Promédios.....	94
Calculando a média.....	94
MÉDIA e MÉDIAA.....	95
MÉDIASE e MÉDIASES.....	97
MÉDIA.INTERNA.....	100
Outros meios para um fim.....	102
Medianas: Pegos no Meio.....	103
Encontrando a mediana.....	104
MED.....	104

Estatísticas À La Moda . . . . .	105
Encontrando a moda . . . . .	105
MODO.ÚNICO e MODO.MULT . . . . .	106
<b>CAPÍTULO 5: Desviando da Média . . . . .</b>	<b>109</b>
Medindo a Variação . . . . .	110
Calculando a média de desvios quadrados: Variação e como calculá-la . . . . .	110
VAR.P e VARPA . . . . .	113
Variação da amostra . . . . .	115
VAR.A e VARA . . . . .	116
De Volta às Raízes: Desvio Padrão . . . . .	116
Desvio padrão da população . . . . .	117
DESVPAD.P e DESVPADPA . . . . .	117
Desvio padrão de amostra . . . . .	118
DESVPAD.A e DESVPADA . . . . .	118
Funções faltantes: DESVPADSE e DESVPADSES . . . . .	119
Funções Relacionadas . . . . .	123
DESVQ . . . . .	123
Desvio médio . . . . .	123
DESV.MÉDIO . . . . .	125
<b>CAPÍTULO 6: Seguindo Padrões e Constantes . . . . .</b>	<b>127</b>
Conhecendo Alguns Zs . . . . .	128
Características dos valores z . . . . .	128
Bonds versus Bambino . . . . .	129
Notas de provas . . . . .	130
PADRONIZAR . . . . .	130
Onde Você Fica? . . . . .	133
ORDEM.EQ e ORDEM.MÉD . . . . .	133
MAIOR e MENOR . . . . .	135
PERCENTIL.INC e PERCENTIL.EXC . . . . .	136
ORDEM.PORCENTUAL.INC e ORDEM.PORCENTUAL.EXC . . . . .	138
Ferramenta de análise de dados: Ordem e Percentil . . . . .	140
Para Usuários de Mac . . . . .	142
<b>CAPÍTULO 7: Resumindo Tudo . . . . .</b>	<b>143</b>
Contando . . . . .	144
CONT.NÚM, CONT.VALORES, CONTAR.VAZIO, CONT.SE, CONT.SES . . . . .	144
O Longo e o Curto . . . . .	146
MÁXIMO, MÁXIMOA, MÍNIMO e MÍNIMOA . . . . .	147
Quase Esotérico . . . . .	148
DISTORÇÃO e DISTORÇÃO.P . . . . .	148
CURT . . . . .	150
Ajustando a Frequência . . . . .	152
FREQUÊNCIA . . . . .	152

	Ferramenta de análise de dados: Histograma	154
	Você Pode Me Dar Uma Descrição?	156
	Ferramenta de análise de dados: Estatística Descritiva	156
	Seja Rápido!	158
	Estatística Instantânea	161
	Para Usuários de Mac	162
	Estatística descritiva	163
	Histograma	164
	Estatísticas instantâneas	166
<b>CAPÍTULO 8:</b>	<b>O Que é Normal?</b>	167
	Atingindo a Curva	168
	Aprofundando	169
	Parâmetros de uma distribuição normal.	169
	DIST.NORM.N	171
	INV.NORM.N	172
	Um Membro Distinto da Família	173
	DIST.NORMP.N	175
	INV.NORMP.N	175
	PHI e GAUSS.	176
<b>PARTE 3: TIRANDO CONCLUSÕES A PARTIR DOS DADOS</b>		
		179
<b>CAPÍTULO 9:</b>	<b>O Jogo da Confiança: Estimativas</b>	181
	Entendendo a Distribuição Amostral	182
	Uma Ideia EXTREMAMENTE Importante: o Teorema do Limite Central	183
	Simulando o Teorema do Limite Central.	184
	Os Limites da Confiança	189
	Calculando limites de confiança para uma média.	189
	INT.CONFIANÇA.NORM.	191
	Inserindo um t	192
	INT.CONFIANÇA.T.	194
<b>CAPÍTULO 10:</b>	<b>Teste de Hipótese com Uma Amostra</b>	195
	Hipóteses, Testes e Erros	196
	Testes de hipótese e distribuições amostrais	197
	Mais Alguns Zs	199
	TESTE.Z	201
	t para Um.	202
	DIST.T, DIST.T.CD e DIST.T.BC	204
	INV.T e INV.T.BC	205
	Testando uma Variação	206
	DIST.QUIQUA e DIST.QUIQUA.CD.	207
	INV.QUIQUA e INV.QUIQUA.CD.	209

<b>CAPÍTULO 11: Teste de Hipótese com Duas Amostras</b> .....	211
Hipóteses para Dois .....	212
Revisão das Distribuições Amostrais .....	213
Aplicando o Teorema do Limite Central .....	213
Ainda Mais Zs .....	215
Ferramenta de Análise de Dados: Teste Z: Duas Amostras para Médias .....	217
t para Dois .....	219
Cara de um, focinho de outro: variações iguais .....	220
Como p's e q's: variações desiguais .....	221
TESTE.T .....	222
Ferramenta de Análise de Dados: Teste t: Duas Amostras .....	223
Um Conjunto Harmonioso: Teste de Hipótese para Amostras Pareadas .....	226
TESTE.T para amostras combinadas .....	228
Ferramenta de análise de dados: teste t: Duas Amostras em Par para Médias .....	229
Testando Duas Variações .....	232
Utilizando F em combinação com t .....	234
TESTE.F .....	234
DIST.F e DIST.F.CD .....	236
INV.F e INV.F.CD .....	238
Ferramenta de Análise de Dados: teste F: Duas Amostras para Variâncias .....	239
Para Usuários de Mac .....	240
<b>CAPÍTULO 12: Testando Mais de Duas Amostras</b> .....	243
Testando Mais de Duas .....	244
Um grande problema .....	245
Uma solução .....	245
Relações significativas .....	249
Depois do teste F .....	250
Ferramenta de análise de dados: Anova: Fator Único .....	253
Comparando as médias .....	255
Outro Tipo de Hipótese, Outro Tipo de Teste .....	257
Trabalhando com a ANOVA de medidas repetidas .....	257
Seguindo tendências .....	260
Ferramenta de análise de dados: Anova: Fator Duplo Sem Repetição .....	264
Analisando tendências .....	266
Para Usuários de Mac .....	267
Análise de Variância com Fator Único .....	268
Medidas Repetidas .....	269
<b>CAPÍTULO 13: Testes Um Pouco Mais Complicados</b> .....	273
Decifrando as Combinações .....	274
Dividindo as variações .....	274

	Ferramenta de análise de dados: Anova: Fator Duplo	
	Sem Replicação	276
	Decifrando as Combinações Mais Uma Vez	278
	Linhas e colunas	278
	Interações	279
	A análise	279
	Ferramenta de análise de dados: Anova: Fator Duplo	
	Com Replicação	281
	Para Usuários de Mac	283
<b>CAPÍTULO 14:</b>	<b>Regressão: Linear e Múltipla</b>	287
	O Gráfico da Difusão	288
	Linhas de Gráfico	289
	Regressão: Que Linha!	291
	Utilizando a regressão para prever	293
	Variação em torno da linha de regressão	293
	Testando hipóteses sobre regressão	295
	Funções de Planilha para Regressão	300
	INCLINAÇÃO, INTERCEPÇÃO, EPADYX	301
	PREVISÃO	303
	Função matricial: TENDÊNCIA	303
	Função matricial: PROJ.LIN	307
	Ferramenta de análise de dados: Regressão	309
	Resultados tabelados	310
	Resultados em gráficos	312
	Lidando com Diversas Relações ao Mesmo Tempo:	
	Regressão Múltipla	313
	Ferramentas do Excel para Regressão Múltipla	315
	Revisão de TENDÊNCIA	315
	Revisão de PROJ.LIN	316
	Revisão da ferramenta de análise de dados Regressão	319
	Para Usuários de Mac	321
<b>CAPÍTULO 15:</b>	<b>Correlação: Ascensão e Queda das Relações</b>	325
	Gráficos de Dispersão, Mais Uma Vez	326
	Compreendendo a Correlação	326
	Correlação e Regressão	329
	Testando Hipóteses de Correlação	331
	Um coeficiente de correlação é maior que zero?	332
	Dois coeficientes de correlação são diferentes?	333
	Funções de Planilha para Correlação	334
	CORREL e PEARSON	334
	RQUAD	336
	COVARIANÇA.P e COVARIANÇA.S	336
	Ferramenta de Análise de Dados: Correlação	337
	Resultados tabelados	338
	Ferramenta de Análise de Dados: Covariância	341



Testando Hipóteses Sobre Correlação .....	342
Funções de Planilha: FISHER, FISHERINV .....	342
Para Usuários de Mac .....	344
<b>PARTE 4: PROBABILIDADE .....</b>	<b>347</b>
<b>CAPÍTULO 16: Introdução à Probabilidade .....</b>	<b>349</b>
O que é Probabilidade? .....	350
Experimentos, tentativas, eventos e espaços amostrais .....	350
Espaços amostrais e probabilidade .....	350
Eventos Compostos .....	351
União e intersecção .....	351
Mais uma vez: Intersecção .....	352
Probabilidade Condicional .....	353
Trabalhando com as probabilidades .....	354
A base para os testes de hipótese .....	354
Espaços Amostrais Grandes .....	354
Permutas .....	355
Combinações .....	356
Funções de Planilha .....	356
FATORIAL .....	357
PERMUT e PERMUTONA .....	357
COMBIN e COMBINA .....	358
Variáveis Aleatórias: Discretas e Contínuas .....	359
Distribuições de Probabilidade e Funções de Densidade .....	359
A Distribuição Binomial .....	361
Funções de Planilha .....	363
DISTR.BINOM e INTERV.DISTR.BINOM .....	363
IST.BIN.NEG.N. ....	365
Teste de Hipótese com a Distribuição Binomial .....	366
INV.BINOM .....	366
Mais sobre testes de hipótese .....	367
A Distribuição Hipergeométrica .....	369
DIST.HIPERGEOM.N. ....	369
<b>CAPÍTULO 17: Mais sobre Probabilidade .....</b>	<b>373</b>
Descobrimo o Beta .....	374
DIST.BETA .....	375
INV.BETA .....	377
Poisson .....	378
DIST.POISSON .....	379
Trabalhando com Gama .....	381
A função Gama e GAMA .....	381
A Distribuição Gama e DIST.GAMA .....	381
INV.GAMA .....	384
Exponencial .....	384
DISTR.EXPON .....	385

<b>CAPÍTULO 18: Carreira de Modelo</b> .....	387
Modelando uma Distribuição .....	388
Conhecendo melhor a distribuição de Poisson .....	388
Utilizando DIST.POISSON .....	390
Testando o caimento .....	391
Uma palavra sobre TESTE.QUIQUA .....	393
Jogando bola com um modelo .....	394
Uma Discussão Simulada .....	397
Arriscando: O método de Monte Carlo .....	397
Trapaceando .....	397
Simulando o Teorema do Limite Central .....	401
Para Usuários de Mac .....	405

## PARTE 5: A PARTE DOS DEZ .....

407

<b>CAPÍTULO 19: Dez Dicas e Armadilhas Estatísticas e Gráficas</b> .....	409
Significativo Nem Sempre Quer Dizer Importante .....	410
Tentar Não Rejeitar uma Hipótese Nula Tem Diversas Implicações .....	410
A Regressão Nem Sempre É Linear .....	411
Não É Uma Boa Ideia Extrapolar um Gráfico de Dispersão de Amostra .....	411
Analise a Variação em Torno de uma Linha de Regressão .....	412
Uma Amostra Pode Ser Grande Demais .....	412
Consumidores: Conheça Seus Eixos .....	412
É Errado Representar Graficamente uma Variável Categórica Como Se Ela Fosse uma Variável Quantitativa .....	413
Sempre que Possível, Inclua Variação em Seu Gráfico .....	414
Cuidado Ao Relacionar Conceitos de Livros de Estatística com o Excel .....	414

<b>CAPÍTULO 20: Dez Assuntos (Treze, Na Verdade) Que Não Se Encaixaram em Nenhum Outro Capítulo</b> .....	417
Técnicas de Previsão .....	418
Uma experiência móvel .....	418
Como ser ajustável, mas de modo exponencial .....	420
Representando Graficamente o Erro Padrão da Média .....	422
Probabilidades e Distribuições .....	425
PROB .....	425
DIST.WEIBULL .....	425
Criando Amostras .....	426
Testando a Independência: A Verdadeira Função de TESTE.QUIQUA .....	427
Esoterismo Logarítmico .....	430
O que é um logaritmo? .....	430
O que é e? .....	432

	DIST.LOGNORMAL.N	435
	INV.LOGNORMAL	436
	Função matricial: PROJ.LOG	437
	Função matricial: CRESCIMENTO	440
	Os Registros de Gama	444
	Classificando Dados	445
	Para Usuários de Mac	446
<b>APÊNDICE A:</b>	<b>Quando Sua Planilha é um Banco de Dados</b>	449
	Introdução aos Bancos de Dados em Excel	449
	O banco de dados Satélites	450
	O intervalo de critérios	451
	O formato de uma função de banco de dados	452
	Contando e Recuperando	454
	BDCONTAR e BDCONTARA	454
	BDEXTRAIR	455
	Aritmética	455
	BDMÁX e BDMÍN	455
	BDSOMA	456
	BDMULTIPL	456
	Estatística	456
	BDMÉDIA	457
	BDVAREST e BDVARP	457
	BDEST e BDESVPA	457
	Segundo o Formulário	458
	Tabelas Dinâmicas	459
<b>APÊNDICE B:</b>	<b>A Análise de Covariância</b>	465
	Covariância: Observação Detalhada	465
	Por que Analisar a Covariância	466
	Como Analisar a Covariância	467
	ANCOVA no Excel	468
	Método 1: ANOVA	469
	Método 2: Regressão	473
	Depois da ANCOVA	476
	E Mais Uma Coisa	477
<b>ÍNDICE</b>		481

# Introdução

O quê? Mais um livro sobre estatística? Bem... este é um livro sobre estatística, mas em minha humilde (e totalmente influenciada) opinião, não é *apenas* mais um livro sobre estatística.

O quê? Outro livro sobre Excel? Mesma opinião influenciada — não é apenas mais um livro sobre Excel. O quê? Outra edição de um livro que não é apenas mais um livro sobre estatística nem apenas mais um livro sobre Excel? Bem... sim. Você me pegou.

Aqui está a ideia — para as duas edições anteriores e para esta. Muitos livros de estatística ensinam os conceitos, mas não oferecem uma maneira de aplicá-los. Isso geralmente leva a uma falta de entendimento. Com o Excel, você tem um pacote pronto para aplicar os conceitos de estatística.

Observando pelo ponto de vista contrário, muitos livros sobre Excel apresentam as capacidades do programa, mas não falam sobre os conceitos por trás delas. Antes de falar sobre uma ferramenta de estatística do Excel, falarei sobre o fundamento estatístico no qual ela é baseada. Dessa forma, você entenderá a ferramenta ao usá-la — e também poderá usá-la de maneira mais eficaz.

Eu não queria escrever um livro que dissesse apenas “selecione este menu” e “clique neste botão”. É claro que isso também é necessário em muitos livros que mostram como usar um pacote de software. Mas meu objetivo é ir muito além disso.

Também não queria escrever um “livro de receitas” estatísticas: Quando se deparar com o problema nº 310, utilize o procedimento estatístico nº 214. Mais uma vez, meu objetivo era ir muito além disso.

Conclusão: Este livro não fala apenas sobre estatística ou apenas sobre Excel — ele fica bem na interseção desses dois assuntos. Enquanto falo sobre estatística, falo sobre todas as características estatísticas do Excel. (Bem, quase todas. Deixei uma de fora. Ela também ficou de fora das duas primeiras edições. Chama-se “Análise Fourier”. Toda a matemática necessária para compreendê-la demandaria um livro inteiro, e talvez você nunca use essa ferramenta.)

# Sobre Este Livro

Embora estatística envolva uma progressão lógica de conceitos, organizei este livro para que você possa abri-lo em qualquer capítulo e comece a ler. A ideia é que você encontre o que está procurando em um momento de emergência e utilize o conceito imediatamente — seja ele um conceito estatístico ou uma ferramenta do Excel.

Por outro lado, se você quiser ler do início ao fim, não tem problema algum. Se você é novato em estatística e precisa usar o Excel para análise estatística, recomendando que comece pelo começo — mesmo que você conheça muito bem o Excel.

# O Que Pode Ser Ignorado com Segurança

Qualquer livro de referência traz um monte de informações e este não é uma exceção. Minha intenção é que ele seja útil, mas nem todos os assuntos foram abordados no mesmo nível. Portanto, se você não quer muitos detalhes, poderá evitar parágrafos marcados com o ícone Papo de Especialista.

De vez em quando, você encontrará uma caixa de texto. Elas fornecem informações que desenvolvem um pouco mais um assunto, mas não fazem parte do caminho principal. Se estiver com pressa, poderá ignorá-las.

Como este livro foi escrito de modo que você possa abri-lo em qualquer parte e começar a usá-lo, instruções passo a passo aparecem em toda parte. Muitos dos procedimentos descritos têm passos em comum. Depois de executar alguns procedimentos, você provavelmente começará a pular os primeiros passos quando chegar a um procedimento que ainda não foi seguido.

# Penso que...

Este não é um livro introdutório sobre Excel nem sobre Windows, então, suponho o seguinte:

- » Você sabe como trabalhar com o Windows. Não falarei sobre detalhes de como apontar, clicar, selecionar e assim por diante.
- » Você tem o Excel instalado em seu Windows (ou no seu Mac) e pode trabalhar com os exemplos. Não falarei sobre os passos de instalação do Excel.

» Você já trabalhou com Excel antes e compreende os fundamentos das planilhas e das fórmulas.

Se você não conhece muito bem o Excel, sugiro que leia o *Excel 2016 For Dummies*, por Greg Harvey (em inglês), e para ter noção dos fundamentos das fórmulas, temos o *Excel Fórmulas e Funções Para Leigos*, por Ken Bluttman (Edição Alta Books).

## Como Este Livro Está Organizado

Organizei este livro em cinco partes e quatro apêndices (dois deles são apêndices bônus e podem ser encontrados para download no site da editora — [www.altabooks.com.br](http://www.altabooks.com.br), digitando o título do livro ou ISBN).

### Parte 1: Começando a Análise Estatística Com o Excel

Na Parte 1, faço uma introdução geral à estatística e às capacidades estatísticas do Excel. Discuto conceitos estatísticos importantes e descrevo técnicas úteis. Se já faz muito tempo que você fez um curso de estatística ou se você nunca fez um curso de estatística, comece por aqui. Se você nunca trabalhou com as funções integradas do Excel (quaisquer que sejam), comece por aqui, com certeza.

### Parte 2: Descrevendo Dados

Parte da estatística é pegar conjuntos de números e resumi-los de diversas maneiras. Nesta parte, você descobre como fazer isso. Todos conhecemos médias e como calculá-las, mas isso é apenas uma parte da história. Aqui, falarei sobre outros tipos de estatísticas, que preenchem os espaços vazios, e mostrarei como usar o Excel para trabalhar com estas estatísticas. Também apresentarei os gráficos do Excel nesta parte.

### Parte 3: Tirando Conclusões a Partir dos Dados

A Parte 3 fala sobre o objetivo principal da análise estatística: Ir além dos dados e ajudar a tomar decisões. Geralmente, os dados são medidas de uma amostra retirada de uma grande população. O objetivo é usar esses dados para descobrir o que está acontecendo com a população.

Isso gera muitas questões: O que é uma média? O que significa a diferença entre duas médias? Os dois assuntos estão associados? Estas são apenas algumas das questões sobre as quais falarei na Parte 3. Também falarei sobre as funções e ferramentas do Excel que irão ajudá-lo a responder essas perguntas.

## **Parte 4: Probabilidade**

A probabilidade é a base da análise estatística e da tomada de decisões. Na Parte 4, falarei sobre probabilidade. Mostrarei como aplicá-la, principalmente na área de modelagem. O Excel oferece um rico conjunto de capacidades integradas que o ajudarão a compreender e aplicar a probabilidade. É aqui que você as encontra.

## **Parte 5: A Parte dos Dez**

A Parte 5 tem duas finalidades. Primeiro, eu tenho a chance de ficar no púlpito e discursar sobre pegadinhas estatísticas e dar dicas úteis. As pegadinhas e dicas somam dez. Além disso, discuto sobre dez (está bem, 13) pontos do Excel que não consegui encaixar em nenhum outro capítulo. Todos eles vêm do mundo da estatística. Se tem a ver com Excel e com estatística, e se você não encontrar em nenhuma outra parte do livro, certamente encontrará aqui.

Como disse nas duas primeiras edições — esta Parte dos Dez é bem útil.

## **Apêndice A: Quando Sua Planilha É um Banco de Dados**

Além de fazer cálculos, o Excel tem outra função: Manter registros. Embora não seja um banco de dados dedicado, o Excel oferece algumas funções de banco de dados. Algumas delas são estatísticas por natureza. Apresento as funções de banco de dados do Excel no Apêndice A, juntamente com as tabelas dinâmicas, que permitem que você vire seu banco de dados do avesso e veja seus dados de maneiras diferentes.

## **Apêndice B: A Análise de Covariância**

A Análise de Covariância (ANCOVA) é uma técnica estatística que combina outras duas técnicas — análise de variação e análise de regressão. Se você conhece a relação entre duas variáveis, pode usar esse conhecimento de maneiras inesperadas, e a análise de covariância é uma delas. A questão é que o Excel não tem uma ferramenta integrada para a ANCOVA, — mas mostrarei como usar o que o Excel oferece para conseguir realizar seu trabalho.

## Apêndice Bônus 1: Quando Seus Dados Vivem em Outro Lugar

Este Apêndice é sobre importar dados para o Excel — da web, de bases de dados e de textos.

## Apêndice Bônus 2: Dicas para Professores (E Alunos)

O Excel é ótimo para gerenciar, manipular e analisar dados. Também é uma ótima ferramenta para ajudar as pessoas a entender conceitos estatísticos. Este Apêndice aborda algumas maneiras de usar o Excel para fazer exatamente isso.

## Além Deste Livro

Os Apêndices Bônus, a Folha de Cola Online e o download do pacote StatPlus Mac Le podem ser encontrados no site da editora, em [www.altabooks.com.br](http://www.altabooks.com.br) (procure pelo título do livro ou pelo ISBN).

## Ícones Usados Neste Livro

Como em todos os livros da série *Para Leigos*, os ícones estão em toda parte. Cada um deles traz uma figura na margem que o ajuda a saber algo especial sobre o parágrafo ao lado do qual o ícone está localizado.



DICA

Este ícone aponta uma dica ou um atalho que o ajuda a trabalhar e faz com que você seja um ser humano melhor.



LEMBRE-SE

Este ícone refere-se à sabedoria eterna, que você deverá levar consigo depois que terminar de ler o livro, pequeno gafanhoto.



CUIDADO

Preste atenção neste ícone. É um lembrete para evitar algo que possa atrapalhar seu trabalho.



PAPO DE ESPECIALISTA

Como mencionei em “O Que Pode Ser Ignorado Com Segurança”, este ícone indica que você pode passar reto se estatística e Excel não são suas paixões.



# De Lá para Cá, Daqui para Lá

Você pode começar o livro por onde quiser, mas gostaria de dar algumas dicas. Quer aprender os fundamentos da estatística? Vire a página. Quer conhecer os recursos estatísticos do Excel? Capítulo 2. Quer começar pelos gráficos? Capítulo 3. Para qualquer outro assunto, consulte o Sumário ou o Índice e vá em frente.

# 1

**Começando  
a Análise  
Estatística  
com Excel**

## **NESTA PARTE . . .**

Descubra as capacidades estatísticas do Excel

---

Explore como trabalhar com populações e amostras

---

Teste suas hipóteses

---

Entenda os erros em tomadas de decisão

---

Determine variáveis independentes e dependentes

## NESTE CAPÍTULO

**Introduzindo os conceitos estatísticos**

**Generalizando a partir de amostras a populações**

**Conhecendo a probabilidade**

**Tomando decisões**

**Novos e antigos recursos no Excel 2013**

**Compreendendo importantes fundamentos do Excel**

## Capítulo 1

# Avaliando Dados no Mundo Real

O campo da estatística está intimamente relacionado à tomada de decisões — tomada de decisões com base em grupos de números. Os estatísticos fazem perguntas o tempo todo: O que os números nos dizem? Quais são as tendências? Quais previsões podemos fazer? A que conclusões podemos chegar?

Para responder a essas perguntas, os estatísticos desenvolveram um conjunto impressionante de ferramentas analíticas. Essas ferramentas nos ajudam a entender as montanhas de dados que nos aguardam e também a compreender os números que geramos durante nosso próprio trabalho.

# As Noções Estatísticas (e Relacionadas) Que Você Precisa Conhecer

Como o cálculo intensivo geralmente faz parte das ferramentas usadas pelo estatístico, muitas pessoas têm a ideia errada de que estatística é apenas o processamento de números. Entretanto, o processamento numérico é apenas uma pequena parte do caminho para decisões racionais.

Ao lidar com a carga de processamento numérico, os programas aumentam nossa velocidade por esse caminho. Alguns pacotes de software são especializados em análise estatística e possuem muitas das ferramentas usadas pelos estatísticos. Embora não seja comercializado especificamente como um pacote estatístico, o Excel oferece muitas dessas ferramentas, por isso, escrevi este livro.

Eu disse que o processamento numérico é uma pequena parte do caminho para decisões racionais. A parte mais importante são os conceitos com os quais os estatísticos trabalham, e é sobre eles que falarei no restante deste capítulo.

## Amostras e Populações

Em uma noite de eleição, os comentaristas de TV geralmente preveem o resultado das eleições antes do encerramento da votação. Na maioria das vezes, eles acertam. Como fazem isso?

O truque é entrevistar uma amostra de eleitores logo após a votação. Supondo que os eleitores digam a verdade sobre seu voto e supondo que a amostra realmente represente a população, os analistas de rede utilizam os dados de amostra para generalizar a população de eleitores.

Este é o trabalho de um estatístico — usar o que foi descoberto em uma amostra e tomar uma decisão sobre a população da qual saiu a amostra. Mas, às vezes, essas decisões não são iguais ao que os números previram. Os fãs de história provavelmente conhecem a memorável imagem do presidente Harry Truman segurando uma cópia do Chicago Daily Tribune com a famosa, porém errada, manchete: “Dewey derrota Truman” após as eleições de 1948. Parte do trabalho de um estatístico é expressar quanta confiança ele tem em sua decisão.

Outro exemplo relacionado a eleições refere-se à ideia da confiança na decisão. As pesquisas de pré-eleição (novamente supondo uma amostra que represente os eleitores) informam a porcentagem de eleitores pesquisados que preferem cada candidato. A organização da pesquisa acrescenta a precisão de sua pesquisa. Quando você ouve um jornalista dizendo “com margem de erro de três por cento”, está ouvindo um julgamento sobre a confiança.

Mais um exemplo. Suponha que você seja o responsável por determinar a média de velocidade de leitura de todas as crianças do sexto ano dos Estados Unidos, mas não tem tempo nem dinheiro para testar todas elas. O que você faria?

A melhor aposta é pegar uma amostra de alunos de sexto ano, medir sua velocidade de leitura (em palavras por minuto) e calcular a média da velocidade de leitura da amostra. Você pode, então, usar a média da amostra como uma estimativa da média da população.

Estimar a média da população é um tipo de *inferência* que os estatísticos usam com dados de amostra. Falarei com mais detalhes sobre inferência na seção “Estatística Inferencial: Testando Hipóteses”.



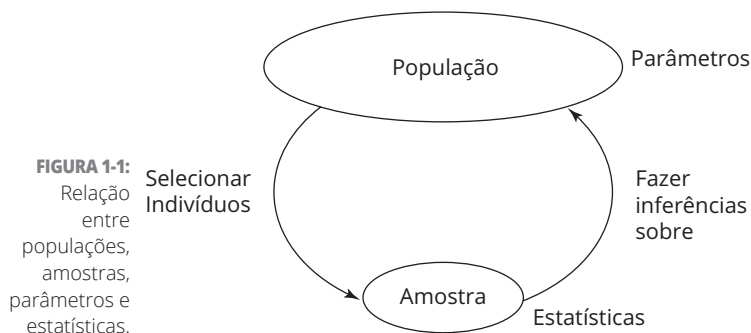
LEMBRE-SE

Alguns termos que você deve conhecer: As características de uma população (como a média da população) são chamadas de *parâmetros*, e as características de uma amostra (como a média da amostra) são chamadas de *estatísticas*. Ao restringir seu campo de visão às amostras, suas estatísticas tornam-se *descritivas*. Quando você amplia seus horizontes e passa a se preocupar com as populações, suas estatísticas são *inferenciais*.



LEMBRE-SE

Uma convenção de notação que você deve conhecer: Os estatísticos usam letras gregas ( $\mu$ ,  $\sigma$ ,  $\rho$ ) para representar parâmetros, e letras de nosso alfabeto ( $\bar{x}$ ,  $s$ ,  $r$ ) para representar estatísticas. A Figura 1-1 resume a relação entre populações e amostras e entre parâmetros e estatísticas.



## Variáveis: Dependentes e independentes

Para simplificar, uma *variável* é algo que pode aceitar mais de um valor. (Algo que só pode ter um valor é chamado de *constante*.) Algumas variáveis com as quais você deve estar familiarizado são a temperatura do dia, o índice BOVESPA, sua idade e o valor do dólar.

Os estatísticos se preocupam com dois tipos de variáveis, *independentes* e *dependentes*. Cada tipo de variável surge em qualquer estudo ou experimento, e os estatísticos avaliam a relação entre elas.

Por exemplo, imagine uma nova maneira de ensinar a ler, que tem a intenção de aumentar a velocidade de leitura de alunos do sexto ano. Antes de colocar esse novo método nas escolas, seria uma boa ideia testá-lo. Para fazer isso, um pesquisador colocaria uma amostra de alunos do sexto ano, de maneira aleatória, em dois grupos: Um grupo é ensinado com o novo método, e o outro é ensinado por métodos tradicionais. Antes e depois de os dois grupos receberem as instruções, o pesquisador mede a velocidade de leitura de todas as crianças que participam do estudo. O que acontece a seguir? Falarei sobre isso na seção intitulada “Estatísticas Inferenciais: Testando Hipóteses”.

Por enquanto, entenda que a variável independente aqui é o Método de Ensino. Os dois valores possíveis dessa variável são Novo e Tradicional. A variável dependente é a velocidade de leitura — que poderia ser medida em palavras por minuto.



LEMBRE-SE



LEMBRE-SE

Em geral, a ideia é tentar descobrir se as mudanças na variável independente estão associadas com as mudanças na variável dependente.

Nos exemplos que aparecem no livro, mostrarei como usar o Excel para calcular diversas características de grupos de pontuações. Tenha em mente que sempre que mostrar um grupo de pontuações, estou falando sobre os valores de uma variável dependente.

## Tipos de dados

Existem quatro tipos de dados. Quando trabalhamos com uma variável, o modo como trabalhamos com ela depende do tipo de dado que ela representa.

A primeira variedade é chamada de dados *nominais*. Se um número é um dado nominal, ele é apenas um nome. Seu valor não significa nada. Um bom exemplo é o número da camisa de um jogador. Ele é apenas um modo de identificar o atleta e diferenciá-lo do restante do time. O número não indica o nível de habilidade do atleta.

Em seguida, temos os dados *ordinais*. Os dados ordinais representam ordem, e os números começam a adquirir um significado além de serem apenas identificadores. Um número maior indica a presença em maior quantidade de determinado atributo do que um número menor. Um exemplo é a Escala de Moh. Utilizada desde 1822, é uma escala cujos valores estão entre 1 e 10. Os mineralogistas utilizam essa escala para classificar a dureza das substâncias. O diamante, classificado como 10, é o mais duro. O talco, classificado como 1, é o mais mole. Uma substância com determinada classificação pode arranhar qualquer substância que tenha uma classificação inferior.

O que falta na Escala de Moh (e em todos os dados ordinais) é a ideia de intervalos iguais e diferenças iguais. A diferença entre uma dureza de 10 e uma dureza de 8 não é a mesma que a diferença entre uma dureza de 6 e uma dureza de 4.

Os dados *intervalados* fornecem diferenças iguais. As temperaturas em Fahrenheit são um exemplo de dados intervalados. A diferença entre 60 graus e 70 graus é a mesma que entre 80 graus e 90 graus.

Eis algo que poderá surpreendê-lo com relação às temperaturas em Fahrenheit: Uma temperatura de 100 graus não é o dobro de uma temperatura de 50 graus. Para que as relações (duas vezes, metade) sejam válidas, é preciso que zero signifique ausência absoluta do atributo que se está medindo. Uma temperatura de 0 grau Fahrenheit não significa ausência de calor — é apenas um ponto qualquer na escala Fahrenheit.

O último tipo de dado, os dados *relacionais*, inclui um ponto zero significativo. No caso das temperaturas, a escala de Kelvin nos fornece dados relacionais. Cem graus Kelvin é duas vezes mais quente do que 50 graus Kelvin. Isso porque o ponto zero da escala Kelvin é o *zero absoluto*, onde toda movimentação molecular (a base do calor) para. Outro exemplo é uma régua. Oito centímetros é o dobro de quatro centímetros. O comprimento zero significa ausência total de comprimento.



Qualquer um desses tipos pode ser a base de uma variável independente ou de uma variável dependente. As ferramentas analíticas utilizadas dependem do tipo de dados com o qual você está lidando.

## Uma pequena probabilidade

Quando os estatísticos tomam decisões, eles expressam sua confiança nessas decisões em termos de probabilidade. Eles nunca podem ter certeza sobre o que decidem. Eles só podem informar quão prováveis são suas conclusões.

Então, o que é probabilidade? A melhor maneira de falar sobre isso é utilizando alguns exemplos. Se você jogar uma moeda para cima, qual a probabilidade de conseguir tirar cara? Intuitivamente, você sabe que, se a moeda for honesta, existe uma chance de 50% para cara e 50% para coroa. Em termos de tipos de números associados à probabilidade, ela é de  $\frac{1}{2}$ .

E com um dado? Qual a probabilidade de obter um 3? Hmm... um dado tem seis faces e uma delas é o 3, portanto, deve ser  $\frac{1}{6}$ , certo? Certo.

Mais um. Você tem um baralho comum. Selecione uma carta aleatória. Qual a probabilidade de a carta ser de paus? Bem, um baralho tem quatro naipes, então a resposta é  $\frac{1}{4}$ .

Acho que você está entendendo. Se você quiser saber a probabilidade de algo acontecer, descubra de quantas maneiras esse evento pode acontecer e divida pelo número total de eventos que podem acontecer. Em cada um dos três exemplos, o evento no qual estávamos interessados (cara, 3 e paus) só pode acontecer de uma maneira.