IAA005 - ESTATÍSTICA APLICADA I
Parte 1

Prof. Arno P. Schmitz

UFPR – Universidade Federal do Paraná

APRESENTAÇÃO

ARNO P. SCHMITZ

Bacharelado em Ciências Econômicas – UFPR

Mestrado em Economia – UFBA

Doutorado em Desenvolvimento Econômico - UFPR

EMENTA

Introdução à Estatística: descritiva e inferencial. Estudo de populações e amostras. Introdução às séries estatísticas. Análise de gráficos. Estudo das medidas de posição e de dispersão. Análise de correlação e regressão linear. Introdução aos testes de hipótese com uma amostra.

BIBLIOGRAFIA

COSTA, H. V. V. da. **Introdução ao R. Recife**: UFPE. 2017.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5 ed. Porto Alegre: AMGH. 2011.

KAZMIER. L. J. **Estatística aplicada à economia e administração**. São Paulo: Pearson Makron Books. 1982.

NAVIDI, W. **Probabilidade e estatística para ciências exatas**. Porto Alegre: AMGH. 2012.

NEDER, H. D. **Amostragem em pesquisas socioeconômicas**. Campinas: Alínea. 2008.

SILVA, B. F. da; DINIZ, J.; BORTOLUZZI, M. A. **Minicurso de estatística básica**: introdução ao software R. Santa Maria: UFSM. 2009.

SOUZA, E. F. M. de; PETERNELLI, L. A.; MELLO, M. P. de. **Software livre R**: aplicação estatística. João Pessoa: UFPR. 2019.

CONTEÚDOS

1. Conceitos básicos
2. Apresentação tabular e gráfica
3. Medidas de posição e dispersão
4. Séries Estatísticas
5. Correlação e regressão

ANÁLISE EXPLORATÓRIA DE DADOS

6. População e Amostra
7. Modelos probabilísticos
8. Modelos estatísticos de inferência

INFERÊNCIA ESTATÍSTICA

CONCEITOS BÁSICOS

O QUE É ESTATÍSTICA

Estatística é uma ciência que estuda e aplica métodos e técnicas que sistematicamente objetivam:

a) Organizar;
b) Descrever;
c) Analisar e;
d) Interpretar dados.

Estes dados são geralmente oriundos de estudos e pesquisas ou de experimentos realizados em quaisquer áreas do conhecimento.

IMPORTÂNCIA DA ESTATÍSTICA

- Interpretação de dados

Exemplos:

- ❖ Questões socioeconômicas e demográficas (renda, consumo, migração, etc.);
- ❖ Questões de saúde e epidemiologia (doenças, tratamentos, etc.);
- ❖ Questões climáticas (chuvas, matas, calor, etc.);
- ❖ Agropecuária (produção, perdas, etc.)
- ❖ Setor público (análise de impactos, projetos, etc.)
- ❖ Política (pesquisas eleitorais e de opinião pública, etc.)
- ❖ Propaganda (pesquisa de mercado, novos produtos, etc.)

- Vasta literatura sobre métodos e aplicações empíricas;
- Aplicações científicas.

ECONOMIA - EXEMPLO

- Como é a variação espacial da renda média das famílias?

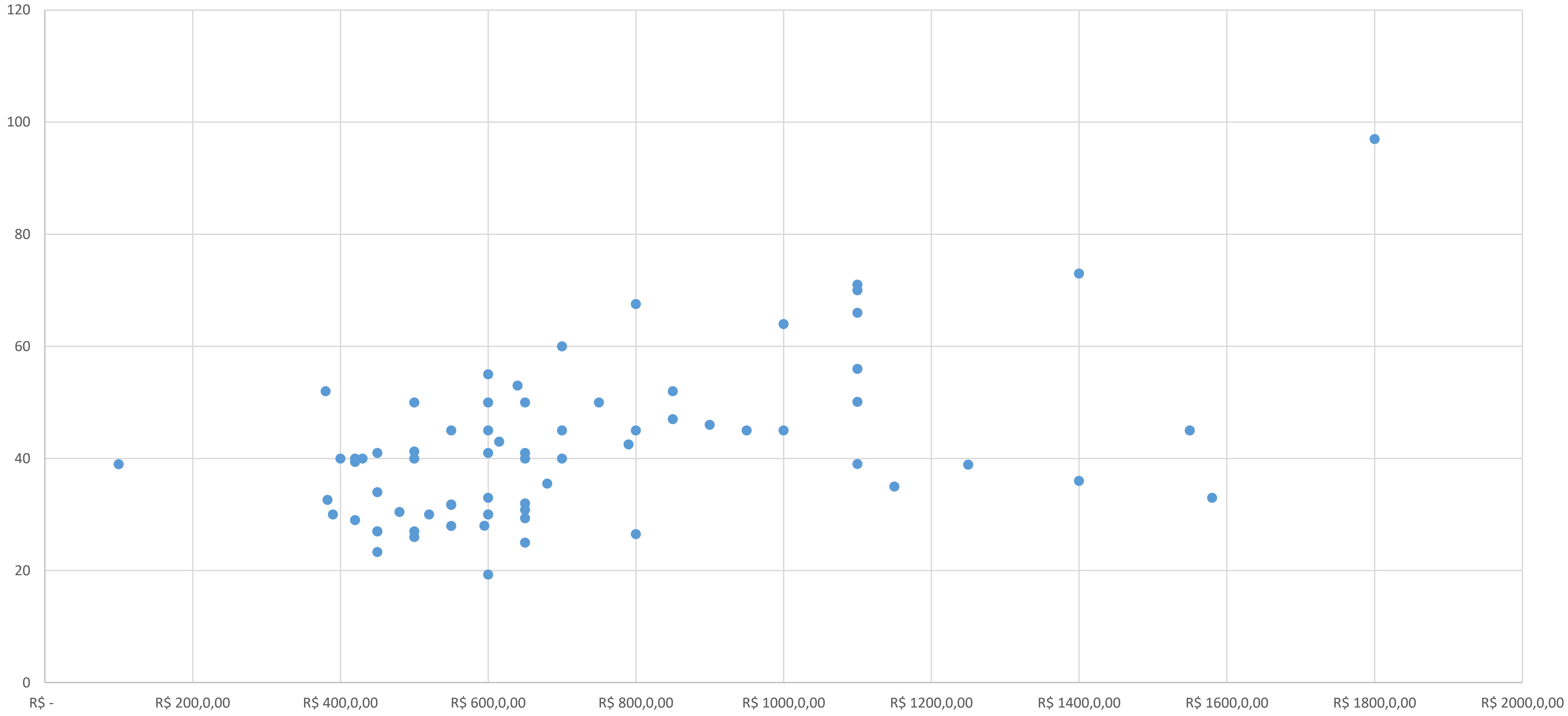
➢ A renda média das famílias está distribuída aleatoriamente ou existe um padrão geográfico?

➢ Como a renda média está relacionada com as características da residência (tamanho, número de moradores, etc.)?

➢ Como as variações da renda média das famílias altera o padrão de consumo?

OUTROS EXEMPLOS

Preço do Aluguel (kitnetes) X Área do Imóvel (m^2) – Centro de Curitiba/PR - 2016



Fatores que explicam o consumo de água

TABELA 10 - RESULTADOS DOS MODELOS DE REGRESSÃO PARA O CONSUMO DE ÁGUA

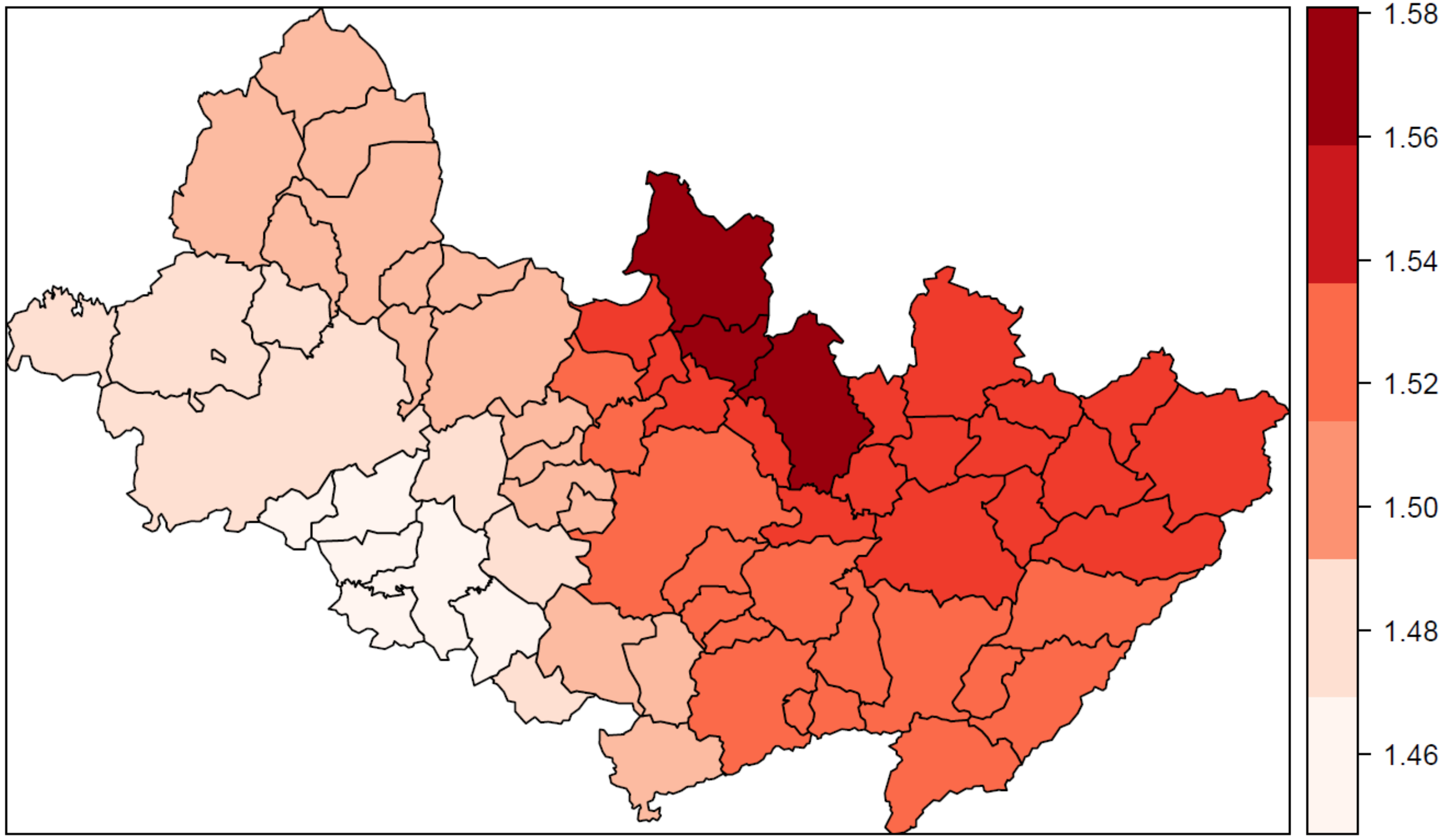
Variáveis	*Pooled*	*Pooled*_ano	EF	EF_ano	EF_*Driscoll-Kraay*
Constante	-3,461***	-4,028***	6,122***	5,282***	5,282***
lnpopurb	1,004***	1,008***	0,097***	0,117***	0,117**
lntrf	-0,204***	-0,197***	-0,229***	-0,223***	-0,223***
lnpib	0,014***	0,008**	0,135***	0,026***	0,026***
lndens	0,007*	0,012***	0,758***	0,261***	0,261***
lntemp	0,623***	0,752***	-0,218**	0,020	0,020
lnprecip	-0,173***	-0,140***	-0,099***	0,044**	0,044**
D	0,004	-0,013	0,032***	0,004	0,004
2006	-	-0,105***	-	-0,138***	-0,138***
2007	-	-0,099***	-	-0,108***	-0,108***
2008	-	-0,086***	-	-0,114***	-0,114***
2009	-	-0,047***	-	-0,104***	-0,104***
2010	-	-0,006	-	-0,042***	-0,042***
2011	-	0,032**	-	-0,028***	-0,028***
2012	-	0,039***	-	0,011	0,011**
2013	-	0,035**	-	0,028***	0,028***
2014	-	-0,011	-	0,042***	0,042***
2015	-	-0,026	-	-0,026***	-0,026***
R^2	98,33	98,42	27,38	39,10	39,10

FONTE: Elaborada pela autora a partir dos *outputs* do *Stata* 15.

NOTA: Todas as variáveis do modelo foram logaritmizadas (*ln*), sendo o volume de água consumido (*lnvolag*, em 1.000 m³/s) a variável dependente. As siglas representam: *lnpopurb*=população urbana atendida pelo abastecimento de água, *lntrf*=tarifa média da água (R$), *lnpib*=PIB municipal (R$ 1.000,00 constantes), *lndens*=densidade demográfica (100 hab/km²), *lntemp*=temperatura média (ºC), *lnprecip*=precipitação média anual (mm), D=cobrança (0=não há, 1=há). O asterisco indica o nível de significância dos parâmetros: * = p<10%, ** = p<5%, *** = p<1%.

Estatística Espacial

Desvio Padrão Local - Disponibilidade Hídrica (metros cúbicos/segundo)



Áreas da Estatística

- Estatística descritiva (análise exploratória de dados);
- Probabilidade;
- Inferência estatística.

Definições

- Dados:

✓ Observações de um evento (pesquisa);

- Informação:

✓ Alguma(s) conclusão(ões) a partir dos dados;

- População (universo):

✓ Conjunto global de pesquisa ao qual se deseja extrair conclusões;

- Amostra:

✓ Subconjunto finito de uma população (sobre a qual são feitas análises e inferências para a população)

Estatística Descritiva

- Objetivo:

➢ Levantar hipóteses e conclusões sobre certas características da população ou amostra

- Contato com os dados e produção de informações básicas;

- Técnicas para descrever e apresentar informações inerentes às bases de dados:

➢ Tabelas

➢ Gráficos

➢ Medidas-resumo informativas

Probabilidade

Estudo da Incerteza

- Fenômenos aleatórios:

➢ Acontecimentos com resultados sem previsão certa (certeza)

➢ Eventos independentes/dependentes

➢ Probabilidade condicional

➢ Funções de distribuição de probabilidade (discretas e contínuas)

a) Distribuição binomial;
b) Distribuição hipergeométrica;
c) Distribuição de Poisson;
d) Distribuição normal (Z, t);
e) Distribuição F;
f) Distribuição Qui-Quadrado.

Inferência Estatística

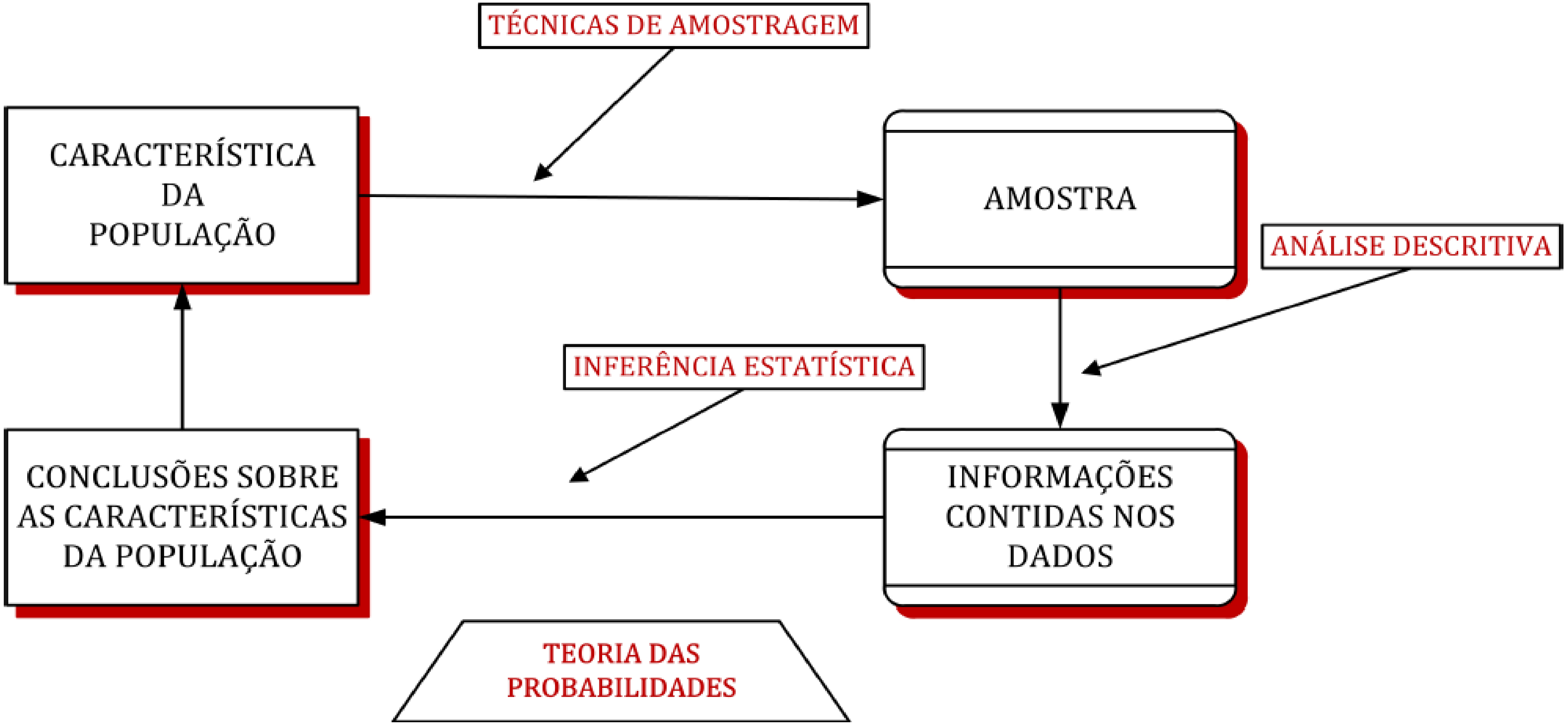
- Técnicas que permitem extrair (inferir) conclusões a partir de uma amostra;

Técnicas:

- Regressão linear simples e múltipla por Mínimos Quadrados Ordinários (cross-section ou séries temporais);
- Regressão com dados longitudinais (espaço e tempo);
- Regressão com variáveis dummy;
- Regressão com equações simultâneas;
- Regressão de modelos com resposta qualitativa (logit, probit, tobit, etc.);
- Modelos de previsão;
- Outros modelos.

População X Amostra

Etapas do Método de Análise Estatística



Fonte: adaptado de [Cancho(2010)]

Problema Fundamental da Estatística

A partir de uma amostra, “como” tirar “CONCLUSÕES” sobre a “População” ?

Resposta: MODELOS DE REGRESSÃO e INFERÊNCIA ESTATÍSTICA

População X Amostra

- População ➔ Censo
- Amostra ➔ Amostragem

Censo X Amostragem

a) Para avaliar toda a população, considerando todos os elementos, realiza-se um censo. P. ex., a cada 10 anos o IBGE realiza o censo demográfico da população brasileira.

b) Caso não haja recursos ou tempo disponíveis para execução do censo, deve-se realizar uma pesquisa por amostragem na população de interesse.

Amostragem – Why?

✓Parâmetros populacionais desconhecidos;

✓ Impossibilidade de realização de um censo;

✓ Amostragem é mais barata e mais rápida;

➔Não existe técnica estatística capaz de ajustar uma amostra mal coletada

➔Plano Amostral (Teoria da Amostragem)

Em geral, uma amostra deve ser um subconjunto representativo da população (de alguma forma).

➢Existem diversas maneiras para se retirar uma amostra de uma população.

Amostras Probabilísticas X Não-probabilísticas

- A amostra é obtida a partir de uma população, por meio de processos definidos pelo pesquisador.

Subdivide-se em dois grupos:

- **Amostragem Probabilística** = Cada elemento da população possui a mesma probabilidade de ser selecionado para compor a amostra → mecanismos de seleção aleatórios;

P. ex., se em uma população de 1.000 peças produzidas existem 20% de peças defeituosas, uma amostra de 100 dessas peças também deve possuir essa mesma proporção de peças defeituosas, ou seja, 20 peças.

- **Amostragem Não Probabilística** = A seleção da amostra depende do julgamento do pesquisador. Há uma escolha deliberada dos elementos para compor a amostra → mecanismos não aleatórios de seleção.

Amostragem com reposição X sem reposição

Quando a população é finita, a amostragem pode ser feita de duas formas:

- **Com reposição** - o elemento amostrado volta a fazer parte da população após ser analisado.

➢ O espaço amostral permanece inalterado.

Exemplo: amostrar a quantidade de veículos que passam em determinada avenida entre as 17:30 e 19:30hs na primeira semana do mês, o veículo amostrado volta para a população.

- **Sem reposição** - o elemento amostrado não volta a compor a população após ser analisado.

➢ O espaço amostral é alterado.

Exemplo: amostrar pessoas para avaliar hábitos de consumo, pessoas amostradas não voltam à população.

População Finita e Infinita

- Em populações infinitas, escolher entre amostragem com reposição ou sem reposição é irrelevante.

- Geralmente, costuma-se empregar o seguinte critério para saber se uma população pode ser considerada finita em relação à amostra:

Se $\frac{n}{N} \geq 0{,}05$, ou seja 5%, a população é considerada finita. Em outras palavras, quando a amostra for menor ou igual a 5% da população, ela deve ser considerada infinita.

- Sendo N e n os tamanhos populacional e amostral, respectivamente.

Plano Amostral

1) **Definição da população alvo** (ex. famílias residentes na região metropolitana de Curitiba, ano 2020 nos meses de fevereiro e março);
2) **Determinação do quadro amostral** (lista, localização ou disponibilidade da população; ex: questionário aplicado em residências sorteadas nos diversos bairros, ponderadas pelo tamanho populacional);
3) **Seleção da técnica de amostragem** (Aleatória simples; Estratificada; Conglomerados; etc. – com uso de questionários, formulário eletrônico, etc.);
4) **Determinar o tamanho da amostra** (Se a população é conhecida e a variável de interesse existe; caso contrário faz-se a determinação do tamanho da amostra baseada em experimento);
5) **Executar o processo de amostragem** (no caso de experimento o processo de amostragem é feito em conjunto com a determinação do tamanho da amostra.

Amostras mais precisas são geralmente amostras maiores (reduz erro amostral)!!

Erros Amostrais

- **Erros amostrais**: Trata-se da diferença entre o que foi apurado na amostra e o verdadeiro valor da população.

✓Ocorrem pois as amostras são aleatórias

- **Erros não amostrais**: Ocorrem quando os dados amostrais são coletados incorretamente, devido a(o):

i) uma amostra tendenciosa;

ii) instrumento de medida defeituoso;

iii) anotações erradas no questionário;

iv) Etc.

- Observação: Os erros não amostrais devem ser minimizados

Dimensionamento do Tamanho da Amostra

- **Para populações "infinitas" (amostras pequenas)**

$$n = \left(\frac{Z_{\propto} \,.\, \sigma_x}{d}\right)^2$$

Em que:

n = tamanho da amostra;

$Z_{\propto}$ = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\propto}{2}$, ou seja, o valor tabelado de Z = 1,96;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

d = erro amostral absoluto admitido;

Dimensionamento do Tamanho da Amostra

- Quando o desvio padrão populacional da variável "x" é desconhecido, opta-se por realizar sua estimativa por experimento, durante o processo de amostragem.

➢ Este experimento inicia-se com os valores das primeiras 50 observações (existem divergências) amostrais selecionadas aleatoriamente (para pesquisas socioeconômicas), eliminados os outliers.

➢ Após efetua-se o cálculo do desvio-padrão amostral da média dos valores de x, $\left(S\bar{x} = s/\sqrt{n}\right)$ e se utiliza este valor em substituição ao valor de "σ_x". Este método é repetido com o crescimento da amostra para verificar a variação do valor do desvio padrão e a suficiência do tamanho da amostra.

Dimensionamento do Tamanho da Amostra

Para estimar o desvio padrão da amostra aplica-se a seguinte expressão:

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}}$$

Em que:

s = desvio padrão da amostra;

X = cada valor da variável X da amostra;

$\bar{X}$ = média dos valores de X da amostra;

n = tamanho da amostra.

Dimensionamento do Tamanho da Amostra

O erro amostral absoluto admitido é dado por:

$$d = Z_{\propto} \cdot \frac{\sigma_x}{\sqrt{n}}$$

Em que:

$d =$ erro amostral absoluto admitido;

$Z_{\propto}$ = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\propto}{2}$, ou seja, o valor tabelado de Z = 1,96;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

n = tamanho da amostra.

Dimensionamento do Tamanho da Amostra

- **Para populações “finitas” (grandes amostras)**

$$n = \frac{1}{\frac{d^2}{Z_{\propto}^2 \, . \, \sigma_x^2} + \frac{1}{N}}$$

Em que:

n = tamanho da amostra;

$d =$ erro amostral absoluto admitido;

$Z_{\propto}$ = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\propto}{2}$, ou seja, o valor tabelado de Z = 1,96;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

N = Tamanho da população.

Escolha do Método de Amostragem

- Para a escolha do método deve-se levar em conta:

✓Tipo de pesquisa que será feita;

✓Acessibilidade e disponibilidade dos elementos da população;

✓Disponibilidade de tempo;

✓Disponibilidade de recursos financeiros e humanos;

✓Outros fatores específicos da pesquisa.

Amostragem Probabilística

- Os métodos mais usuais para se realizar uma amostragem probabilística são: amostragem aleatória simples, amostragem sistemática, amostragem estratificada e amostragem por conglomerados.

Amostragem aleatória simples

- É necessário o conhecimento do tamanho e características da população;
- A amostra é escolhida por sorteio (aleatoriamente) ou pela tabela de números aleatórios;
- Outras formas de acesso aleatório aos elementos da população são admitidas.

Amostragem Aleatória Simples

- **Para uma população infinita**:

➢ Pode-se utilizar faixas de valores.

➢ Porém, é necessário que se garanta que as probabilidades de intervalos de valores que sejam incluídos na amostra sejam iguais às porcentagens da população.

➢ Por exemplo, em uma pesquisa sobre massa corporal, se em uma população 40% dos elementos possuem massa corporal entre 50 e 70Kg, na amostra a porcentagem deve se manter: aproximadamente 40% dos elementos amostrados também devem ter massa corporal entre 50 e 70Kg.

Amostragem Sistemática

- Utilizada quando os elementos estão dispostos de maneira organizada (ex.: fila, lista) e aleatória.

- Escolhe-se um ponto de partida (por sorteio) e seleciona-se cada k-ésimo elemento da população (ex.: o 20º elemento; 40º elemento, etc.)

➢ Exemplo: Em uma fábrica de parafusos, a cada 1000 peças produzidas, uma é retirada para teste.

Amostragem Estratificada

- Indicada quando a população está dividida em grupos distintos, denominados estratos.

➢ Dentro de cada estrato é realizada uma amostragem aleatória simples.

➢ O tamanho da amostra pode ou não ser proporcional ao tamanho do estrato (quando é proporcional diz-se que trata-se de uma **Amostragem Estratificada Proporcional**).

Exemplo: Uma comunidade universitária com 8000 indivíduos está estratificada da seguinte forma:

Estrato	População	Amostra
Professores	800	80
Funcionários	1200	120
Estudantes	6000	600

Amostragem por Conglomerados

- A área da população é dividida em seções (ou conglomerados, ex.: bairros, quarteirões).

- Os conglomerados são selecionados aleatoriamente;
- Dentro de um conglomerado, todos os elementos são amostrados;
- Eventualmente pode-se ter Amostragem por conglomerados em que coleta-se uma amostra e utiliza-se pesos para expandir os resultados para a população;

Exemplo: POF do IBGE - conglomerados (municípios + setores censitários + domicílios = AAS) com estratificação (a partir do censo demográfico).

Amostragem Não Probabilística

- Na amostragem não probabilística, os elementos da população não tem a mesma probabilidade de serem selecionados, portanto, não há garantias da representatividade da população.

Amostragem por conveniência

Os elementos são selecionados por serem imediatamente disponíveis. Exemplo: Uma repórter entrevistando pessoas na rua.

Amostragem por julgamento

Uma pessoa experiente no assunto escolhe intencionalmente os elementos a serem amostrados. Exemplo: Um novo produto a ser “testado” por pessoas.

Amostragem Não Probabilística

Amostragem por quotas

- A amostragem por quotas consiste em um refinamento da amostragem por conveniência. Nela, os elementos selecionados devem obedecer as proporções de características da população.

Exemplo: se em uma população, existem 20% de indivíduos da classe econômica A, 50% da classe B e 30% da classe C, a amostra por conveniência deve respeitar essas proporções (amostragem por quotas), selecionando 20% de indivíduos da classe econômica A, 50% da classe B e 30% da classe C.

Dados de Pesquisas e Variáveis

Dados

- São dados obtidos diretamente em pesquisas

➢Ainda sem qualquer tratamento, processo de síntese ou análise

- Podem estar incluídos em tabelas

➢Porém, não tratados em publicações

Variáveis

- Característica que pode variar em termos de valor entre elementos em uma amostra ou população

➢Elementos diferentes podem ter valores diferentes de uma variável (por exemplo: Renda do indivíduo)

Variáveis

Quantitativa:

- Escala de medida tem valores numéricos (12; 18; 1.253; 1.832,21; 1.326,18; etc.);
- Os valores podem representar magnitudes diferentes para cada variável (R$; kg; metros; etc.).

Qualitativas:

- Escala de medida é um conjunto de categorias (M; F; etc.);
- Categorias distintas diferem em qualidade, não em magnitude numérica.

Escala de Medida

- Valores que a variável expressa

Exemplo:

➢Gênero: masculino, feminino, outro;

➢Número de moradores do domicílio: 0, 1, 2, ...

➢Renda: R$1.200,00; 1.800,00; etc.

➢Altura: 180cm; 1,76m; etc.

- Os métodos estatísticos consideram a escala das variáveis.

Variáveis Quantitativas

- **Variável discreta**:

➢ Os valores possíveis formam um conjunto de números como: 0, 1, 2, 3, ... ; (Ex.: número de moradores, número de veículos, etc.)

- **Variável contínua**:

➢ Os valores são um "contínuo" de valores reais possíveis: 0,01 ; 0,02; 0,03; ...; 1,00; 1,01; ..; (Ex.: Renda do indivíduo; temperatura média; etc.)

Variáveis Qualitativas

- **Variável nominal ou categórica:**

➢Os valores são apresentados em categorias mutuamente exclusivas (Ex.: Gênero, Região, etc.)

- **Variável ordinal:**
- Os valores são apresentados em categorias mutuamente exclusivas que possuem uma ordenação natural (Ex.: Grau de instrução; Classe socioeconômica; etc.)

Séries Estatísticas

- As séries estatísticas são a apresentação de informações (variáveis estatísticas) na forma de tabelas, cujo objetivo é sintetizar as informações estatísticas observadas e torná-las compreensíveis.

➢ Uma tabela ou um gráfico deve ser composto de: **título; corpo e; rodapé**.

➢ **O título** deve conter informações suficientes para compreender as informações expostas:

✓ Fato ou variável de interesse – fenômeno descrito;

✓ Espaço Geográfico – local a que se refere o fato;

✓ Época – tempo em que o fato foi observado.

➢ **O corpo** deve conter o registro dos dados.

➢ **O rodapé** deve conter a identificação da fonte de dados.

Série Temporal

- Série temporal é uma série estatística em que os dados são apresentados segundo a época de ocorrência.

✓O tempo varia, mas o fato - variável(is) de interesse - e o local são fixos.

Tabela 6 - Incentivos à exportação de manufaturados (como % do valor FOB das exportações de manufaturados) - Período 1969-85.

Ano	Isenções			Subsídios				TOTAL GERAL
	ICM	IPI	Draw-back	Crédito Prêmio	Isenção I.R.	Subs. Crédito	Sub-Total	
1969	20,5	7,4	4,0	6,7		4,1	10,8	42,7
1970	20,5	7,2	4,0	13,5		7,5	21,0	52,7
1971	19,8	7,1	4,0	13,2	1,3	7,8	22,3	53,1
1972	19,1	9,0	4,9	16,3	1,3	8,2	25,8	58,8
1973	18,3	8,9	7,2	16,2	1,3	6,5	23,9	58,3
1974	17,7	5,0	12,6	12,0	1,8	6,1	19,9	55,2
1975	17,0	5,4	8,3	12,1	1,7	11,5	25,3	56,0
1976	16,3	5,2	11,8	11,7	1,3	15,9	28,9	62,2
1977	16,3	5,4	12,6	12,4	1,5	19,6	33,5	67,9
1978	16,3	6,1	9,1	12,8	1,8	17,0	31,6	63,1
1979	16,3	6,5	10,5	12,8	2,1	13,9	28,8	62,1
1980	17,7	6,3	9,0	0,0	1,9	2,0	3,9	37,0
1981	18,3	6,8	9,4	6,5	1,8	18,7	27,0	61,6
1982	19,1	7,2	10,3	9,1	1,6	21,7	32,4	69,0
1983	19,1	7,2	8,6	7,8	1,6	9,3	18,7	53,6
1984	20,5	7,0	9,1	7,8	1,6	2,7	12,1	48,7
1985	20,5	7,2	9,1	1,4	1,6	3,6	6,6	43,4

Fonte: Bauman e Moreira (1987) in CLEMENTS, J. C. & McCLAIN, J. S. The political economy of export promotion in Brazil. *In*: GRAHAM, L. S. & WILSON, R. H. *The political economy of Brazil: public policies in an era of transation*. Institute of Latin American Studies Symposia on Latin America Series. Austin: The University of Texas Press, 1990.

Série Geográfica

- Série geográfica é uma série estatística na qual os dados são apresentados segundo a localidade de ocorrência.

➢ O local varia e o tempo e a variável(eis) são fixos

Tabela 1: Tarifa média máxima e aplicada; membros selecionados da OMC, 2013

País (PIB per capita)	Produtos	Média das tarifas máximas	Média das tarifas aplicdas
Estados Unidos (US$ 53000)	Todos	4	3
	Agricultura	5	5
	Vestuário	12	12
Brasil (US$ 11208)	Todos	31	14
	Agricultura	35	10
	Vestuário	35	35
Índia (US$ 1498)	Todos	49	14
	Agricultura	114	34
	Vestuário	38	13

Fonte: World Trade Organization Tariff Profiles e World Bank

Série Específica

- Série Específica – Os dados são agrupados segundo a modalidade de ocorrência. Fato – variável(eis) de interesse, tempo e local fixos.

Taxa (%) de crescimento semestral dos componentes da DA [1]

2011.I

Setores	(%)
Consumo das Famílias	5,7
Consumo da Adm. Pública	2,3
Formação Bruta de Capital Fixo	7,3
Exportação	5,2
Importação	13,9

Fonte: Contas Nacionais Trimestrais - IBGE.

Nota: (1) Refere-se ao crescimento com respeito ao mesmo período do ano anterior

Série Mista

- Série Mista é uma combinação de duas ou mais dos 3 tipos de séries (Temporal, Geográfica e Específica).

Tabela 1
Participação (%) das Economias Fluminense e Paulista por Atividade Econômica em Valor Adicionado Bruto a Preço básico no Brasil (1996 e 2008)

Atividade Econômica	1996		2008	
	RJ	SP	RJ	SP
Agropecuária	1,4	8,6	0,8	7,9
Indústria	8,6	42,9	12,7	33,9
Indústria Extrativa Mineral	18,7	5,0	53,5	1,2
Indústria de Transformação	6,3	46,8	6,7	43,7
Construção	13,9	36,5	10,6	27,6
SIUP*	8,2	45,4	6,0	25,6
Serviços	13,0	34,9	11,6	33,4
Comércio	8,4	41,3	8,3	31,2
Financeiro	10,7	49,9	8,8	51,1
APU**	14,7	20,9	12,7	19,0
Outros	14,2	35,7	12,9	37,7
Total	**11,2**	**35,6**	**11,2**	**32,0**

Fonte: Elaborado pelo Grupo de Economia da Inovação (IE/UFRJ) com base na Retropolação das Contas Nacionais e Regionais, IBGE.
Nota: Preço básico em R$ de 1995.
* Produção e distribuição de eletricidade e gás, água, esgoto e limpeza urbana.
** Administração, saúde e educação públicas e seguridade social.

Gráficos

- A apresentação gráfica de dados estatísticos objetiva expor de forma concisa e simples as séries.
- Isso permite concluir sobre a evolução do fenômeno – variável - ou sobre o comportamento dos valores da série.
- Existem várias maneiras de se representar graficamente os dados estatísticos de acordo com os tipos de séries:

✓Colunas ou Barras;

✓Setores;

✓Linhas;

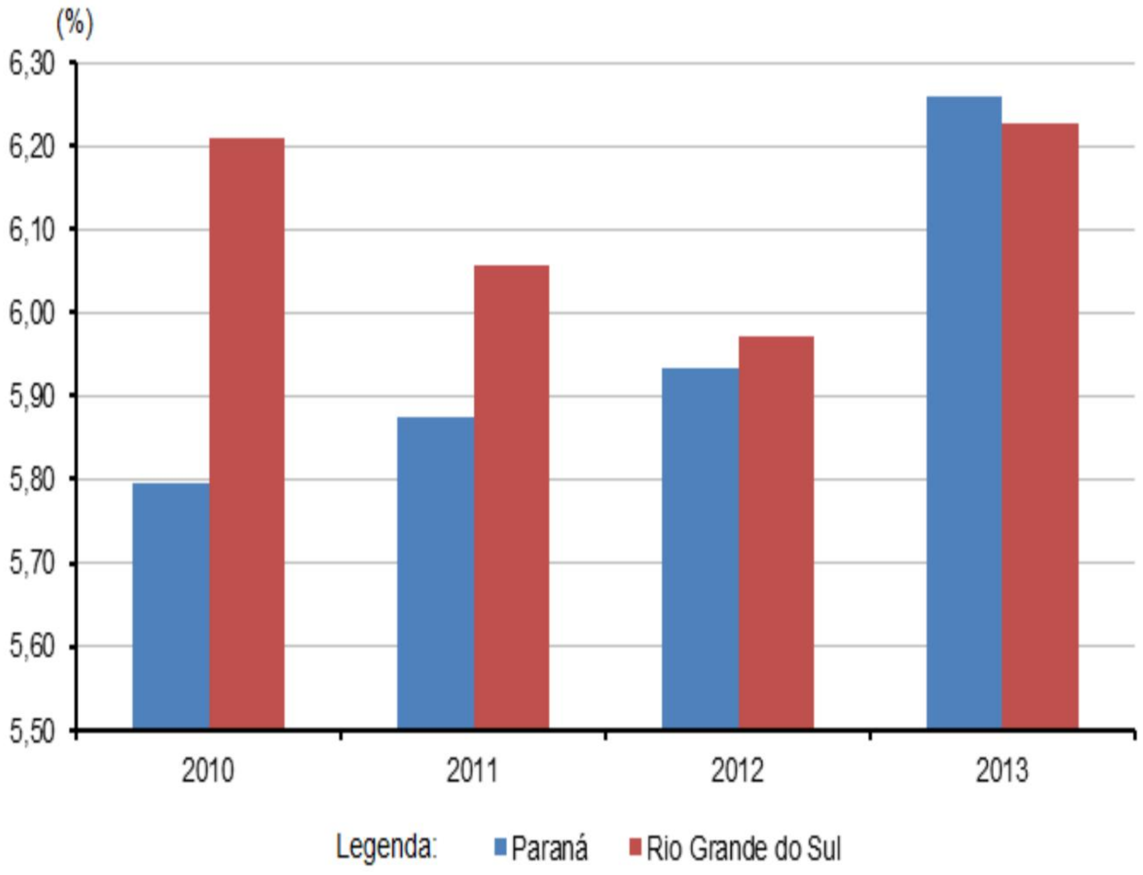
✓Dispersão;

✓Outros.

Gráfico de Colunas ou Barras

Gráfico 1

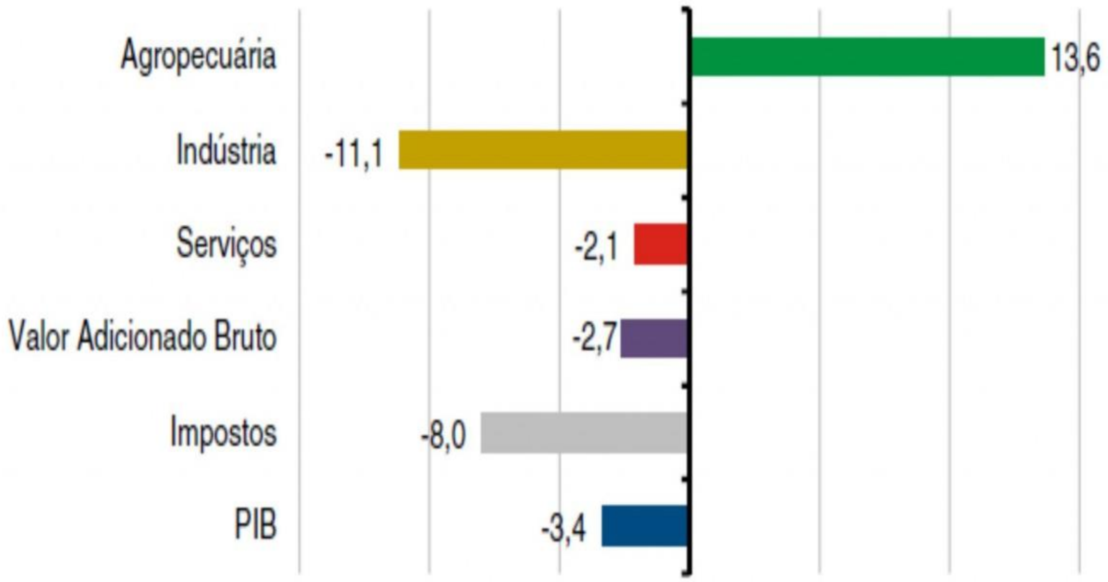
Participação dos Produtos Internos Brutos (PIB) do Rio Grande do Sul e do Paraná no do Brasil — 2010-13



FONTE: IBGE.
FEE.

Gráfico 1

Taxas percentuais de crescimento do Produto Interno Bruto (PIB), acumuladas ao longo do ano, do Rio Grande do Sul — 4.º trim./15

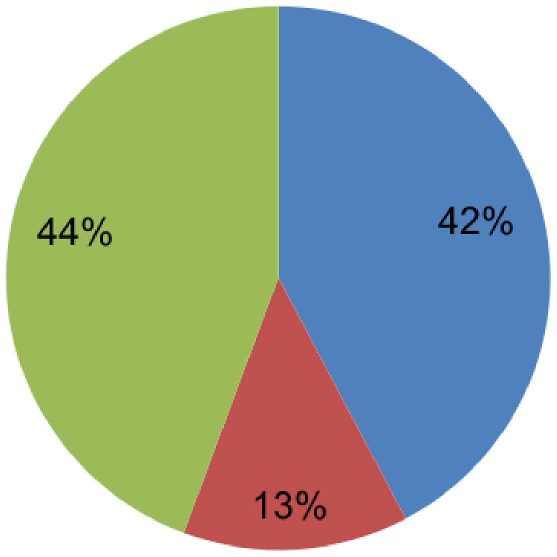


FONTE: FEE/CIES/NCR.

Gráfico de Setores

Gráfico 1

Distribuição dos componentes do Produto Interno Bruto, pela ótica da renda, no Rio Grande do Sul — 2015



- Remunerações
- Impostos, líquidos de subsídios, sobre a produção e a importação
- Excedente Operacional Bruto (EOB) e Rendimento Misto Bruto (RMB)

FONTE: IBGE.
FEE.

Gráfico de Linhas

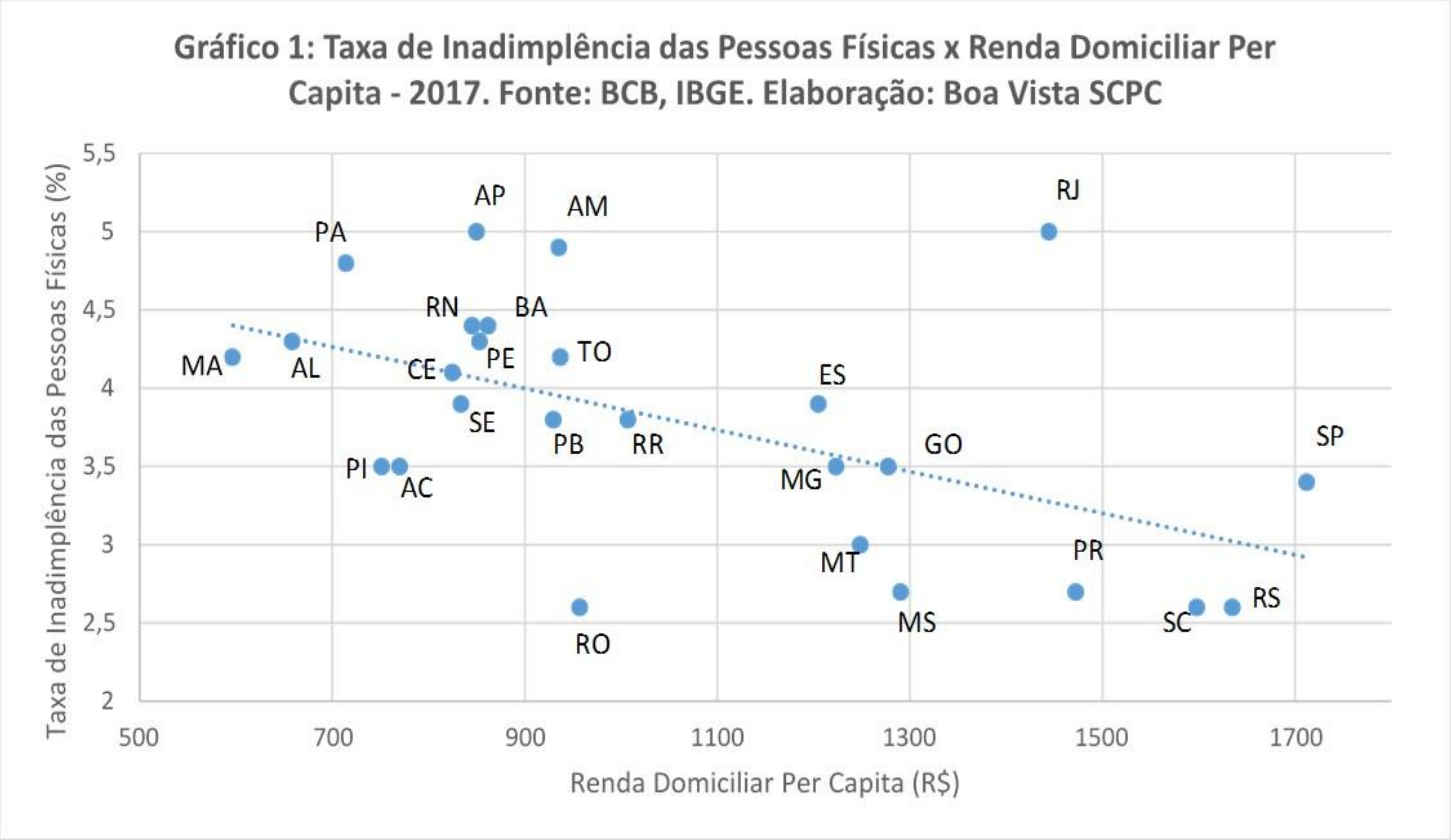
GRÁFICO 5

Dívida bruta do governo geral (DBGG)

(Em % do PIB)

Fonte: Banco Central do Brasil (BCB).

Diagrama de Dispersão



Histograma

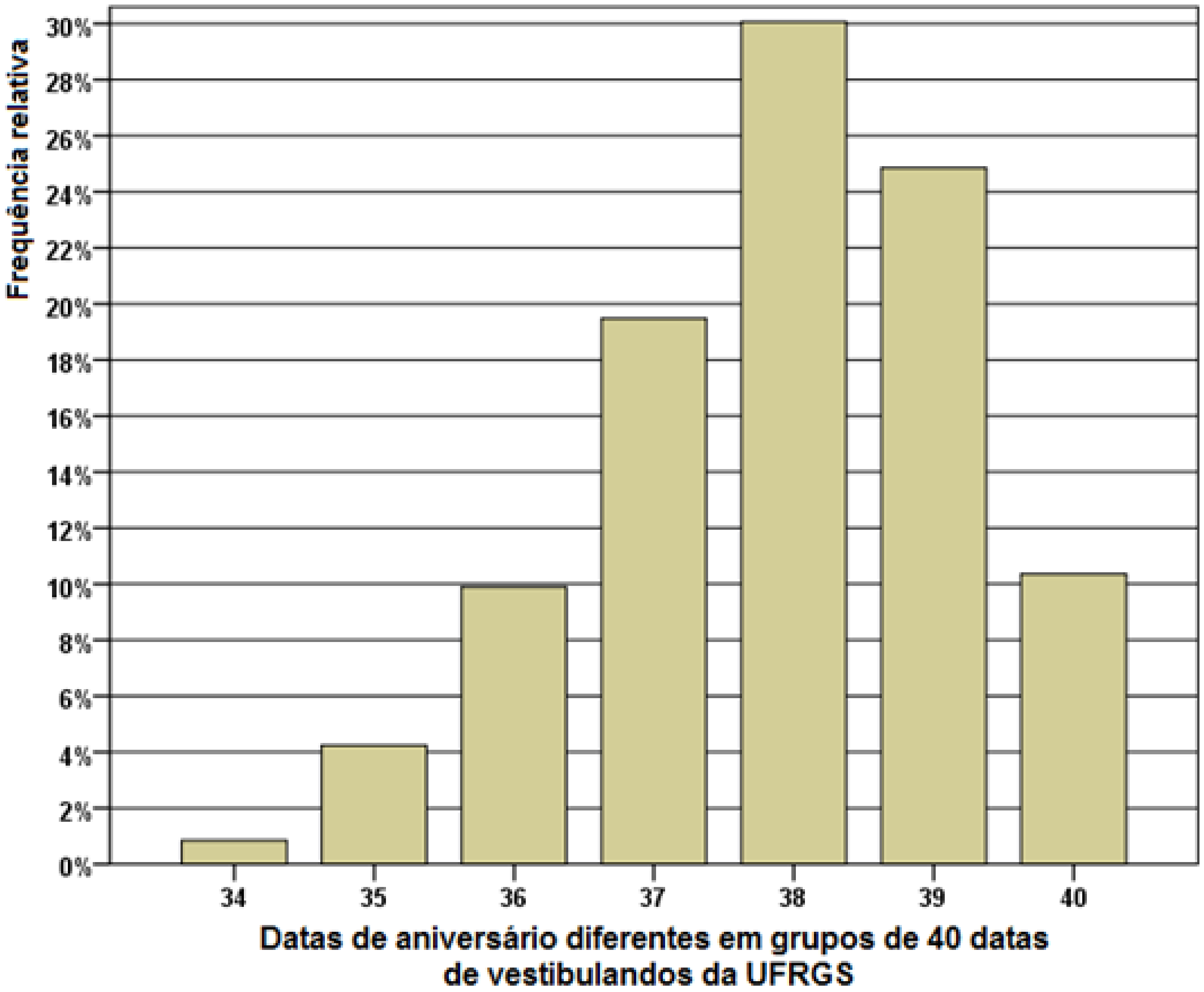
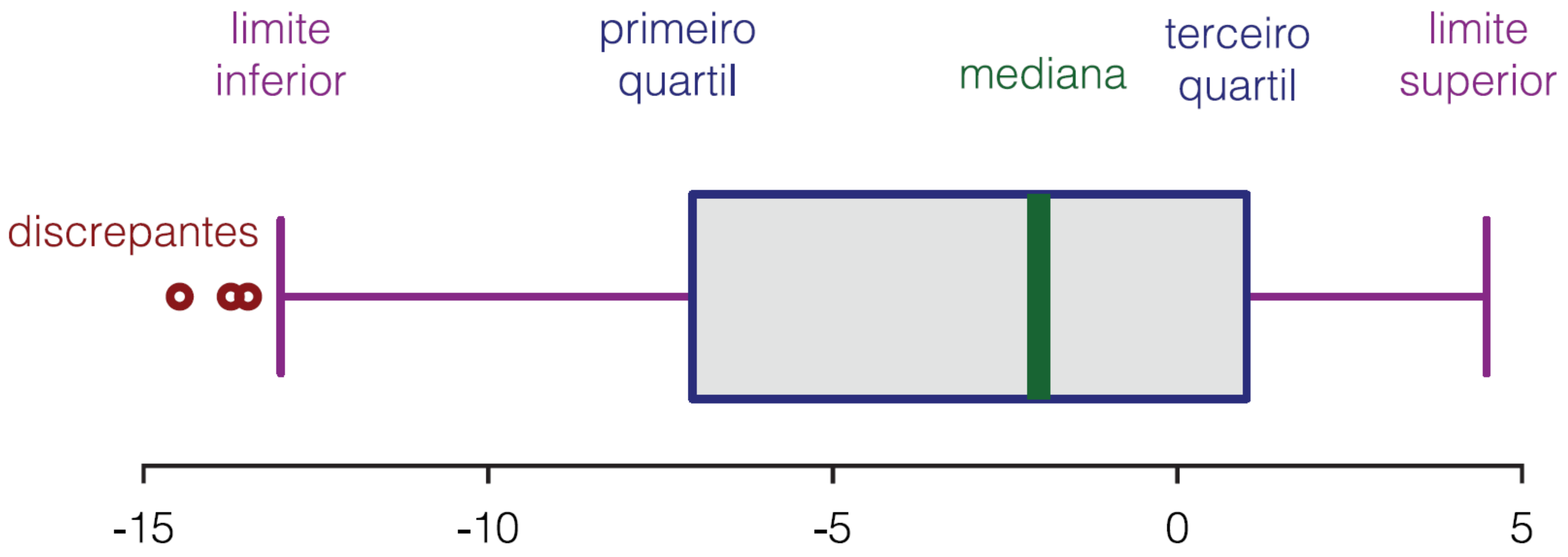


Figura 2 - Distribuição de frequências relativas das datas de aniversário diferentes em 2454 grupos de 40 vestibulandos da UFRGS.

Diagrama de Caixa – Box Plot



Distribuição de Frequência

Tabela 5 - Número de pessoas residentes nos domicílios das crianças e dos adolescentes internados no Hospital João XXIII

Distribuição	Frequência	
	ABS	%
De 2 a 4 pessoas	118	10,2
De 5 a 9 pessoas	65	5,6
Acima de 10 pessoas	4	,3
Sem Informação	965	83,8
Total	1152	100,0

Fonte: SAME-Hospital João XXIII

Distribuição de Frequência

Como construir?

1) Ter em mãos a tabela de dados (amostra ou população);
2) Determinar o número de classes (K)

Alternativas:

a) K = 5 se n < 25

K = $\sqrt{n}$ se n $\geq$ 25;

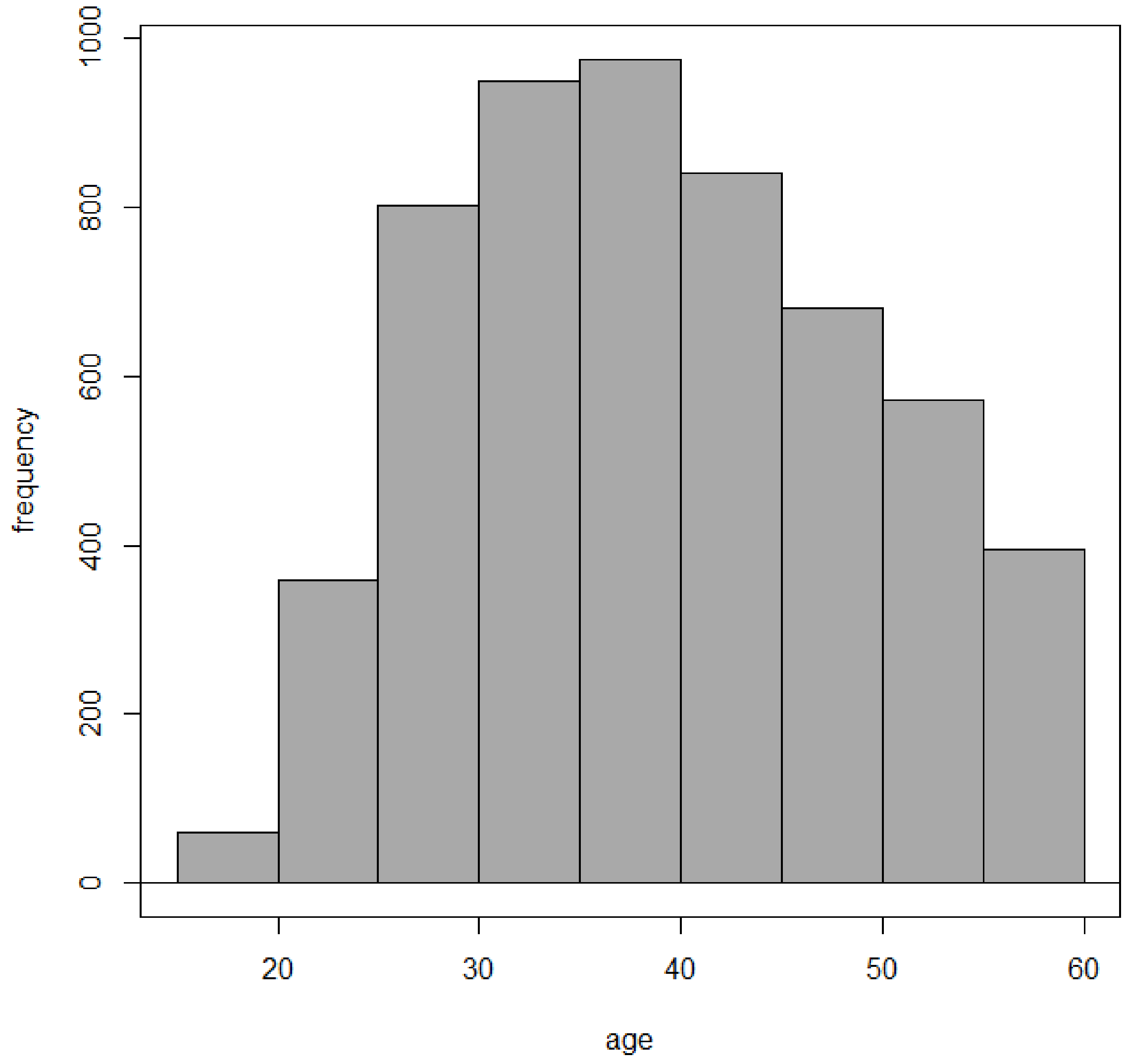
b) Fórmula de Sturges ➔ K = 1 + 3,32.log(n)

em que n = tamanho da amostra ou população;

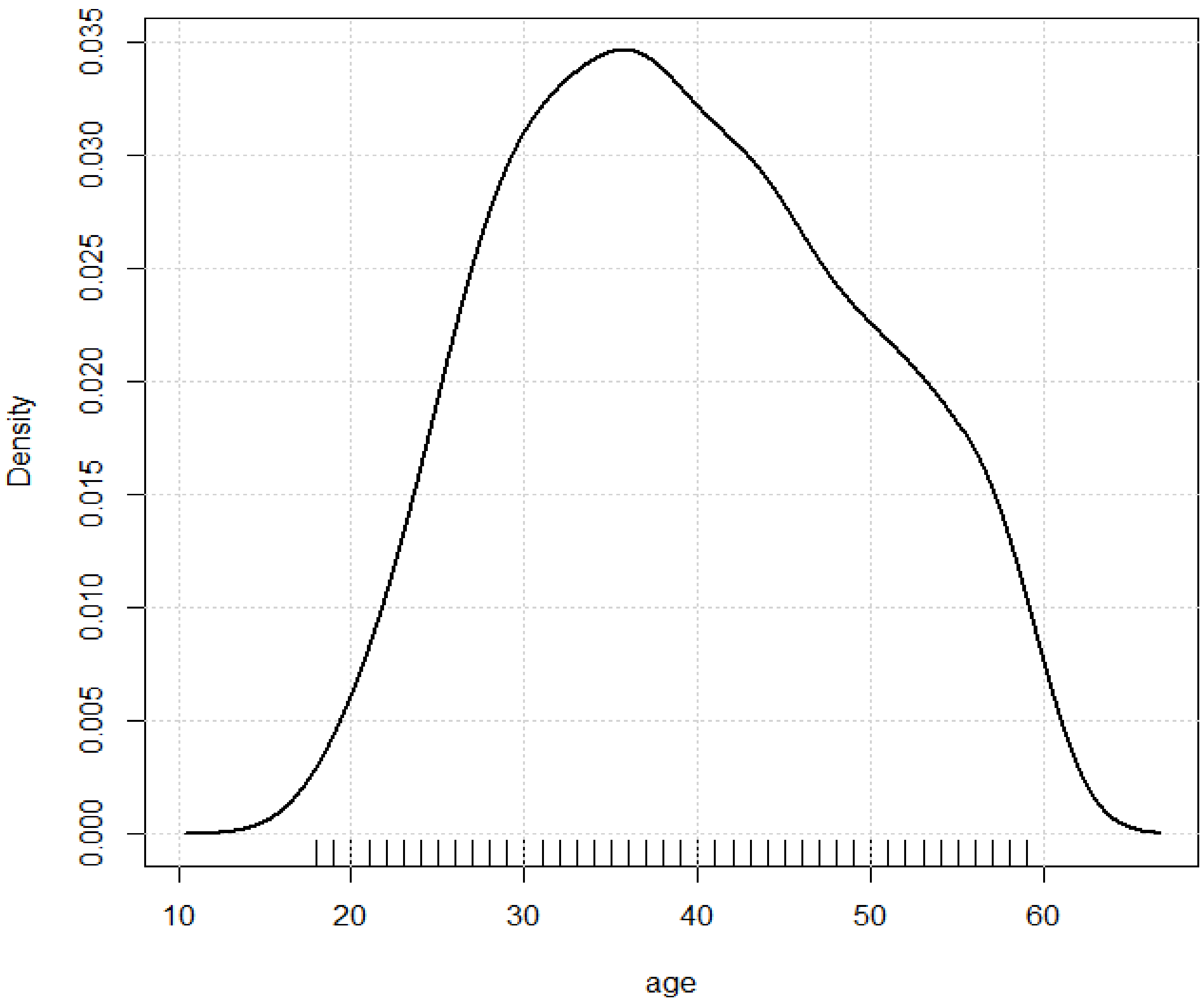
c) Determinação ad hoc (a gosto do pesquisador).

Distribuições de Probabilidade – 5634 pessoas (Idade)

Histograma



Curva de Densidade



Distribuição de Probabilidade Discreta

- As funções de distribuição de probabilidade discretas descrevem as probabilidades das variáveis aleatórias discretas acontecerem (ocorrência do fenômeno estudado).

- Os estatísticos prepararam tabelas de probabilidade, tornando desnecessários os cálculos da probabilidade de ocorrência de determinado evento.

- Dentre as distribuições de probabilidades discretas mais utilizadas estão as distribuições de probabilidade “Binomial”, “Poisson”, “Geométrica”, “Hipergeométrica” e “Binomial Negativa”.

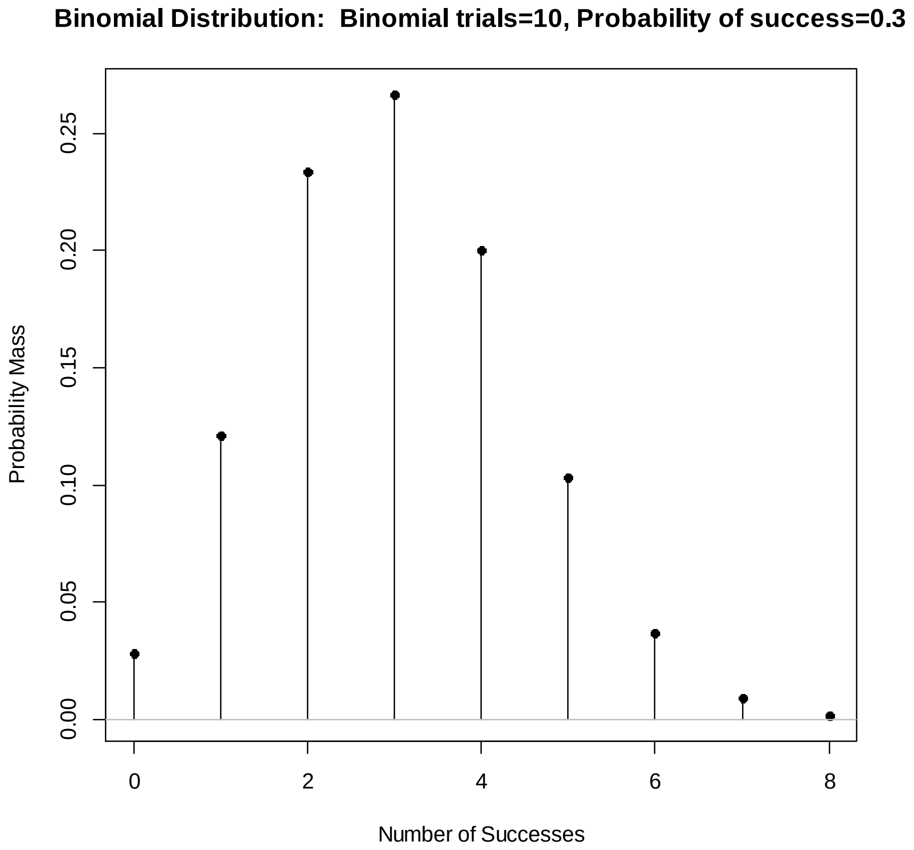
Distribuição Discreta Binomial

Processo de Bernoulli

É um processo de amostragem no qual:

1. Em cada tentativa existem dois resultados possíveis mutuamente exclusivos (sucesso, fracasso);
2. As séries de tentativas, ou observações, são eventos independentes;
3. A probabilidade de sucesso (p), permanece constante de tentativa em tentativa (processo estacionário).

Distribuição Discreta Binomial



$$P(x; n, p) = \ C_n^x \ . \, p^x q^{n-x}$$

Distribuição Discreta Binomial

A Distribuição Binomial é formada a partir da seguinte expressão:

$$P(x; n, p) = C_n^x \,.\, p^x (1-p)^{n-x}$$

Em que:

$P(x; n, p)$ = probabilidade do evento x, dado o número de:

x = número de sucessos do evento;

n = número de tentativas no referido evento;

p = probabilidade de sucesso;

C_n^x = combinação de x, n a n .

Distribuição Discreta Binomial

Dada a expressão:

$$P(x; n, p) = C_n^x \,.\, p^x (1-p)^{n-x}$$

Exemplo:

Durante um ano particular, 70% das ações ordinárias negociadas na bolsa de valores de Nova York tiveram aumentadas suas cotações, enquanto 30% tiveram suas cotações diminuídas ou estáveis. No começo de determinado ano, um serviço de assessoria financeira escolheu 10 ações como sendo "especialmente recomendadas". Se as 10 ações representam uma seleção aleatória, qual a probabilidade de que "nenhuma" das 10 ações tivessem suas cotações aumentadas?

x = 0

n = 10

$$P = \frac{10!}{0!(10-0)!} . 0{,}30^0 . (1-0{,}30)^{\,10-0} = 1 \,.\, 1 .\, 0{,}7^{10} = 0{,}0282$$

p = 0,30

$$P = 2{,}82\%$$

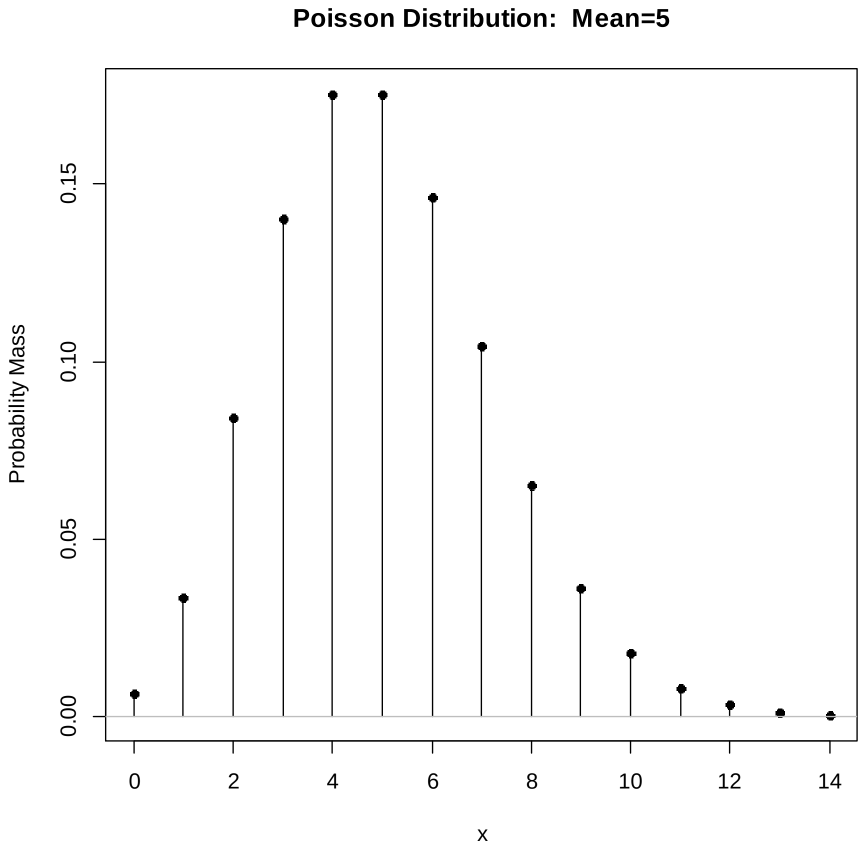
1	DISTRIBUIÇÃO BINOMIAL $X \sim B(n,p)$
	VALORES DA FUNÇÃO MASSA DE PROBABILIDADE $f(x)=P(X=x)$

		p									
n	**x**	**0,05**	**0,1**	**0,15**	**0,2**	**0,25**	**0,3**	**0,35**	**0,4**	**0,45**	**0,5**
1	**0**	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000	0,6500	0,6000	0,5500	0,5000
	1	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000	0,3500	0,4000	0,4500	0,5000
2	**0**	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900	0,4225	0,3600	0,3025	0,2500
	1	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200	0,4550	0,4800	0,4950	0,5000
	2	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900	0,1225	0,1600	0,2025	0,2500
3	**0**	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430	0,2746	0,2160	0,1664	0,1250
	1	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410	0,4436	0,4320	0,4084	0,3750
	2	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890	0,2389	0,2880	0,3341	0,3750
	3	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270	0,0429	0,0640	0,0911	0,1250
4	**0**	0,8145	0,6561	0,5220	0,4096	0,3164	0,2401	0,1785	0,1296	0,0915	0,0625
	1	0,1715	0,2916	0,3685	0,4096	0,4219	0,4116	0,3845	0,3456	0,2995	0,2500
	2	0,0135	0,0486	0,0975	0,1536	0,2109	0,2646	0,3105	0,3456	0,3675	0,3750
	3	0,0005	0,0036	0,0115	0,0256	0,0469	0,0756	0,1115	0,1536	0,2005	0,2500
	4	0,0000	0,0001	0,0005	0,0016	0,0039	0,0081	0,0150	0,0256	0,0410	0,0625
5	**0**	0,7738	0,5905	0,4437	0,3277	0,2373	0,1681	0,1160	0,0778	0,0503	0,0313
	1	0,2036	0,3281	0,3915	0,4096	0,3955	0,3602	0,3124	0,2592	0,2059	0,1563
	2	0,0214	0,0729	0,1382	0,2048	0,2637	0,3087	0,3364	0,3456	0,3369	0,3125
	3	0,0011	0,0081	0,0244	0,0512	0,0879	0,1323	0,1811	0,2304	0,2757	0,3125
	4	0,0000	0,0005	0,0022	0,0064	0,0146	0,0284	0,0488	0,0768	0,1128	0,1563
	5	0,0000	0,0000	0,0001	0,0003	0,0010	0,0024	0,0053	0,0102	0,0185	0,0313

Tabela completa em:
http://www.esac.pt/noronha/estatistica/praticas/Tabela%20Binomial.pdf

Distribuição Discreta Poisson

•



Distribuição Discreta Poisson

A distribuição discreta de Poisson apresenta a probabilidade de Poisson de ocorrência de um determinado número de sucessos.

A principal característica das probabilidades de Poisson é que os eventos ocorrem em um “continuum” de tempo e espaço. É o chamado processo de Poisson.

Exemplos: chamadas em uma central telefônica; navios que atracam em um porto; água que é transportada por um rio, etc.

Distribuição Discreta Poisson

A estimativa da probabilidade de Poisson é dada por:

$$P(X|\lambda) = \frac{\lambda^X e^{-\lambda}}{X!}$$

Em que:

$P(X|\lambda)$ = Probabilidade de Poisson;

λ = número médio de sucessos;

X = Dado número de sucessos de interesse;

Distribuição Discreta Poisson

Exemplo: Um departamento de informática recebe em média 5 chamadas por hora de usuários para solucionar problemas em um dado software. A probabilidade de que, em uma hora selecionada aleatoriamente, sejam recebidas exatamente 3 chamadas é:

$$P(X|\lambda) = \frac{\lambda^X e^{-\lambda}}{X!} = \frac{5^3 e^{-5}}{3!} = \frac{125\,.\,0{,}00673795}{6} = 0{,}1404 = 14{,}04\%$$

Para:

λ = 5;

X = 3;

2

DISTRIBUIÇÃO POISSON $X \sim P(\lambda)$

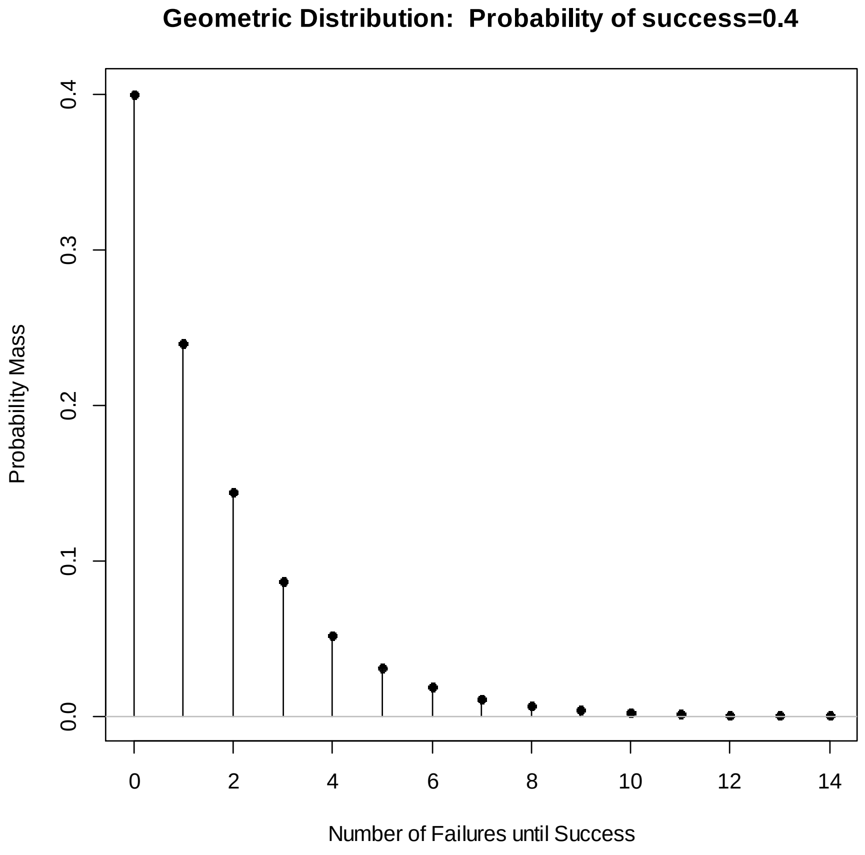
VALORES DA FUNÇÃO MASSA DE PROBABILIDADE

x	λ														
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4	1,5
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679	0,3329	0,3012	0,2725	0,2466	0,2231
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659	0,3679	0,3662	0,3614	0,3543	0,3452	0,3347
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647	0,1839	0,2014	0,2169	0,2303	0,2417	0,2510
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494	0,0613	0,0738	0,0867	0,0998	0,1128	0,1255
4	0,0000	0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111	0,0153	0,0203	0,0260	0,0324	0,0395	0,0471
5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0004	0,0007	0,0012	0,0020	0,0031	0,0045	0,0062	0,0084	0,0111	0,0141
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0003	0,0005	0,0008	0,0012	0,0018	0,0026	0,0035
7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0003	0,0005	0,0008
8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001

x	λ														
	1,6	1,7	1,8	1,9	2	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3
0	0,2019	0,1827	0,1653	0,1496	0,1353	0,1225	0,1108	0,1003	0,0907	0,0821	0,0743	0,0672	0,0608	0,0550	0,0498
1	0,3230	0,3106	0,2975	0,2842	0,2707	0,2572	0,2438	0,2306	0,2177	0,2052	0,1931	0,1815	0,1703	0,1596	0,1494
2	0,2584	0,2640	0,2678	0,2700	0,2707	0,2700	0,2681	0,2652	0,2613	0,2565	0,2510	0,2450	0,2384	0,2314	0,2240
3	0,1378	0,1496	0,1607	0,1710	0,1804	0,1890	0,1966	0,2033	0,2090	0,2138	0,2176	0,2205	0,2225	0,2237	0,2240
4	0,0551	0,0636	0,0723	0,0812	0,0902	0,0992	0,1082	0,1169	0,1254	0,1336	0,1414	0,1488	0,1557	0,1622	0,1680
5	0,0176	0,0216	0,0260	0,0309	0,0361	0,0417	0,0476	0,0538	0,0602	0,0668	0,0735	0,0804	0,0872	0,0940	0,1008
6	0,0047	0,0061	0,0078	0,0098	0,0120	0,0146	0,0174	0,0206	0,0241	0,0278	0,0319	0,0362	0,0407	0,0455	0,0504
7	0,0011	0,0015	0,0020	0,0027	0,0034	0,0044	0,0055	0,0068	0,0083	0,0099	0,0118	0,0139	0,0163	0,0188	0,0216
8	0,0002	0,0003	0,0005	0,0006	0,0009	0,0011	0,0015	0,0019	0,0025	0,0031	0,0038	0,0047	0,0057	0,0068	0,0081
9	0,0000	0,0001	0,0001	0,0001	0,0002	0,0003	0,0004	0,0005	0,0007	0,0009	0,0011	0,0014	0,0018	0,0022	0,0027
10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0004	0,0005	0,0006	0,0008
11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0002
12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001

Tabela completa em:
http://www.esac.pt/noronha/estatistica/praticas/Tabela%20Poisson.pdf

Distribuição Discreta Geométrica



Distribuição Discreta Geométrica

A Distribuição Geométrica considera a probabilidade da ocorrência do número de fracassos até acontecer o primeiro sucesso. O cálculo é dado por:

$$P(Y = k) = p(1 - p)^k$$

Exemplo:

Se a probabilidade de que um certo ensaio de química tenha sucesso é de 0,4; qual será a probabilidade de que ocorram 2 fracassos antes do primeiro sucesso?

$$P(Y = 2) = 0{,}4(1 - 0{,}4)^2 = 0{,}1440$$ = 14,40%

Distribuição Discreta Geométrica

Alternativamente, pode-se obter a probabilidade do número de tentativas necessárias para se obter um sucesso.

$$P(Y = k) = p(1 - p)^{k-1}$$

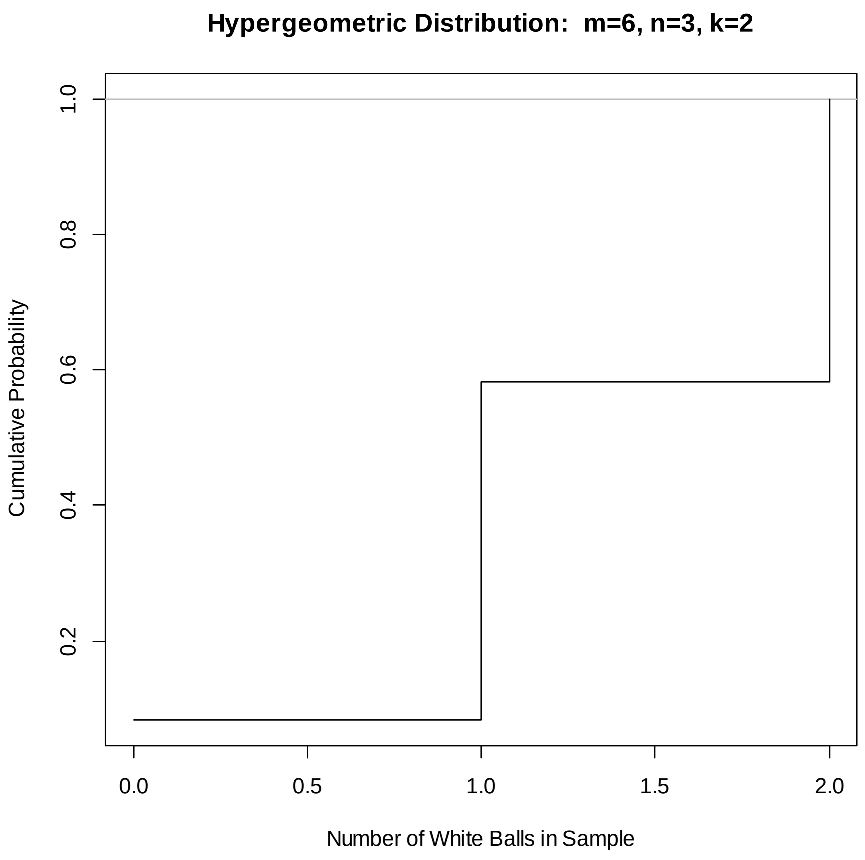
Exemplo:

Na cidade de Curitiba, a probabilidade de ocorrência de chuva de granizo entre os meses de dezembro e janeiro é de 0,10. Admitindo a independência de um dia para o outro, qual é a probabilidade da ocorrência da primeira chuva de granizo acontecer no dia 03 de janeiro?

$$P(Y = 34) = 0{,}1(1 - 0{,}1)^{34-1} = 0{,}003 = 0{,}3\%$$

Distribuição Discreta Hipergeométrica

•



Distribuição Discreta Hipergeométrica

- A amostragem é feita sem reposição de cada item amostrado de uma população finita. Logo, não se pode aplicar o processo de Bernoulli, pois existe mudança na probabilidade de sucesso a medida que os itens são retirados da população.

$$P(X|N, X_t, n) = \frac{\binom{N - X_t}{n - X}\binom{X_t}{X}}{\binom{N}{n}}$$

Em que:

X = número de sucessos;

N = tamanho da população;

X_t = número total de sucessos na população;

n = tamanho da amostra.

Distribuição Discreta Hipergeométrica

Exemplo:

De 6 empregados de uma empresa, 3 trabalham nesta empresa a 5 anos ou mais. Se 4 empregados são aleatoriamente escolhidos deste grupo de 6, a probabilidade de que exatamente 2 estejam na empresa a 5 ou mais anos é:

$$P(X|N,X_t,n) = \frac{\binom{6\ -3}{4\ -2}\binom{3}{2}}{\binom{6}{4}} = \frac{\binom{3}{2}\binom{3}{2}}{\binom{6}{4}} = \frac{\frac{3!}{2!\,1!}\frac{2!}{2!\,1!}}{\frac{6!}{4!\,2!}} =$$

$$P(X|N,X_t,n) = \frac{3\,.\,3}{15} = 0{,}60 = 60\%$$

Em que:

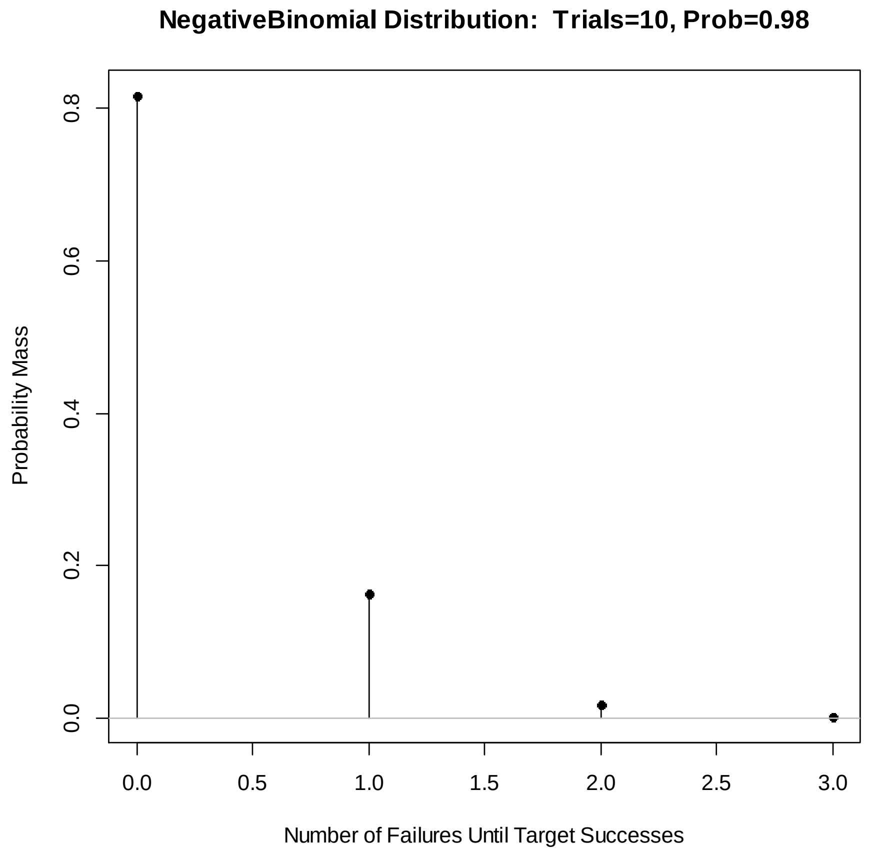
X = 2

N = 6

X_t = 3

n = 4

Distribuição Discreta Binomial Negativa



Distribuição Discreta Binomial Negativa

- Seja "X" uma variável aleatória que conta o número de tentativas necessárias para se obter "k" sucessos, em "n" ensaios de Bernoulli com probabilidade "p" em cada ensaio. Neste caso o último ensaio será o k-ésimo sucesso.

Neste caso, a probabilidade de realizar "x" ensaios é dada por:

$$P(X = x) = \binom{x\ -1}{k\ -1}\ p^k(1-p)^{x-k}$$

Em que:

k = número de sucessos;

x = número de ensaios;

p = probabilidade de sucesso.

Distribuição Discreta Binomial Negativa

- Exemplo: Suponha que em uma fábrica produz resistência para chuveiros, com uma taxa de defeitos de 2%. Qual a probabilidade de que em uma inspeção de 10 resistências tenha-se 3 resistências defeituosas sendo que a terceira defeituosa seja exatamente a décima inspecionada.

$$P(X = 10) = \binom{10\ -1}{3\ -1}\ 0{,}02^{3}(1 - 0{,}02)^{10-3}$$

$$P(X = 10) = \binom{9}{2}.\,0{,}000008\,.0{,}868125$$

$$P(X = 10) = \frac{9!}{2!\,7!}.\,0{,}000008\,.0{,}868125 = 0{,}00025 = 0{,}025\%$$

Em que:

k = 3

x = 10

p = 0,02

Distribuição de Probabilidade Contínua

- Várias distribuições de probabilidade contínuas são aplicáveis a uma ampla variedade de variáveis contínuas, sob certas circunstâncias.

- Os estatísticos prepararam tabelas de probabilidade, tornando desnecessária a integração de área sob a curva de probabilidade para identificar a probabilidade de ocorrência de determinado evento.

- Dentre as distribuições de probabilidade mais utilizadas estão as distribuições de probabilidade "Z - normal", "t de Student", "qui-quadrado", "F de Fisher-Snedecor", "d de Durbin e Watson" e "logística".

Distribuição de Probabilidade Contínua

Distribuição Normal Padronizada (Z)

- A distribuição normal de probabilidade é uma distribuição que é simétrica e mesocúrtica.

Esta distribuição de probabilidade é importante para a inferência estatística por 3 razões distintas:

1. As medidas produzidas em diversos processos aleatórios seguem esta distribuição;
2. A probabilidade normal pode ser usada frequentemente como aproximação de outras distribuições discretas de probabilidade (Poisson e Binomial);
3. As distribuições de estatísticas da amostra tais como a média e a proporção frequentemente seguem a distribuição normal independente da distribuição da população.

Distribuição de Probabilidade Contínua

Distribuição Normal Padronizada (Z)

- A curva de probabilidade, para uma variável normalmente distribuída é dada por:

$$f(X) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\left[\frac{(X-\mu)^2}{2\sigma^2}\right]}$$

- Cada combinação de μ e σ gera uma distribuição normal de probabilidade diferente (simétricas e mesocústicas), as tabelas de probabilidade normal são baseadas em uma distribuição particular: *a distribuição normal padronizada*.
- Esta distribuição normal de probabilidade tem média zero (μ = 0) e desvio padrão igual a 1 (σ = 1).
- Qualquer conjunto de valores (amostras) normalmente distribuído pode ser convertido em valores normais padronizados Z através de:

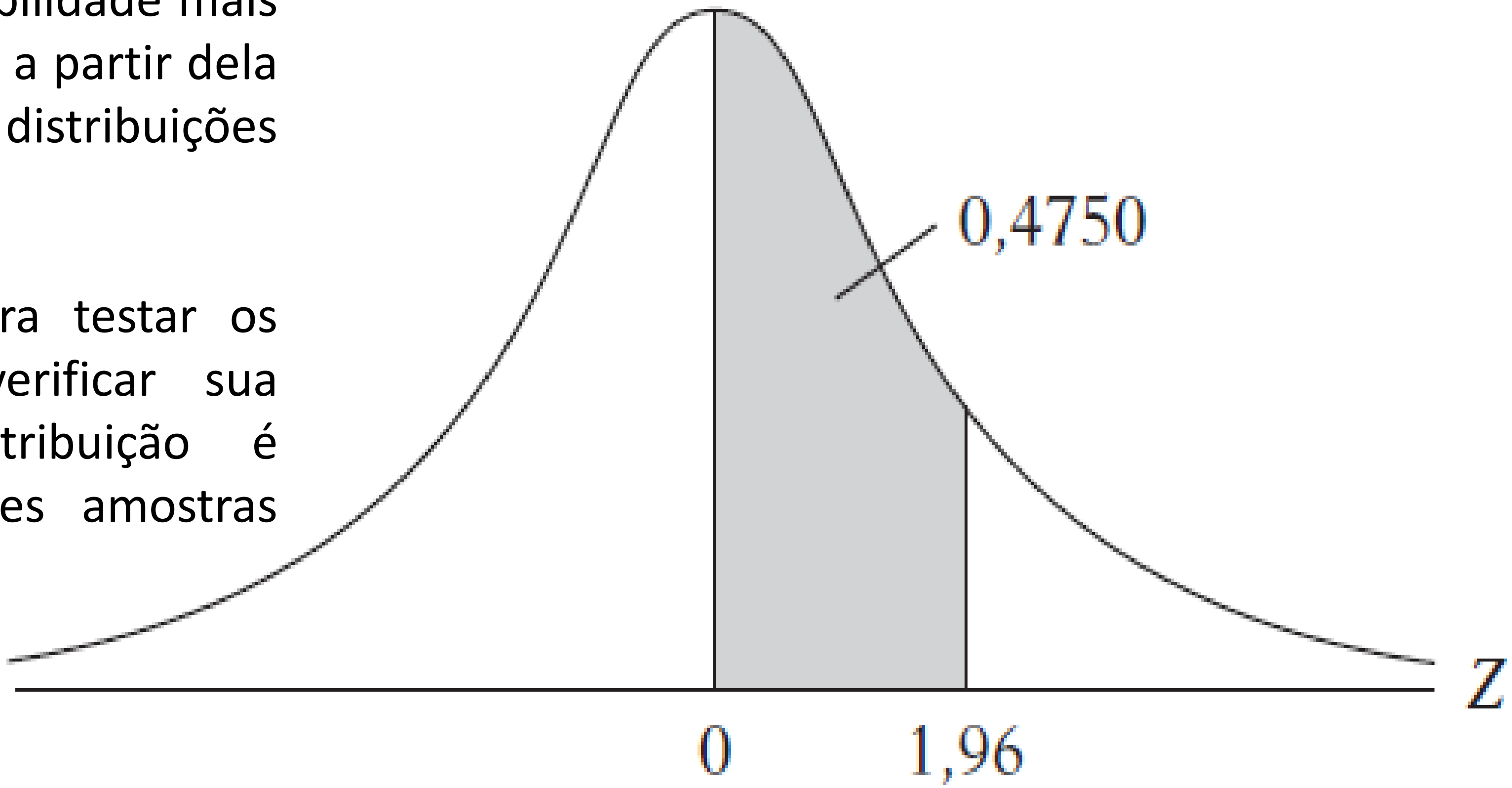
$$Z = \frac{X - \mu}{\sigma}$$

Distribuição de Probabilidade Contínua

Distribuição Normal Padronizada (Z)

Esta é, talvez, a distribuição de probabilidade mais importante para toda estatística. Pois, a partir dela derivaram uma série de outras distribuições estatísticas.

No modelo de MQO ela serva para testar os parâmetros (β_s) calculados e verificar sua significância estatística. Esta distribuição é especialmente utilizada para grandes amostras $(n > 100)$.



Com probabilidade de 95% de confiança
ou 5% de significância

TABELA DA DISTRIBUIÇÃO NORMAL

Tabela A6.2 Distribuição normal — valores de $P(0 \leqslant Z \leqslant z_0)$

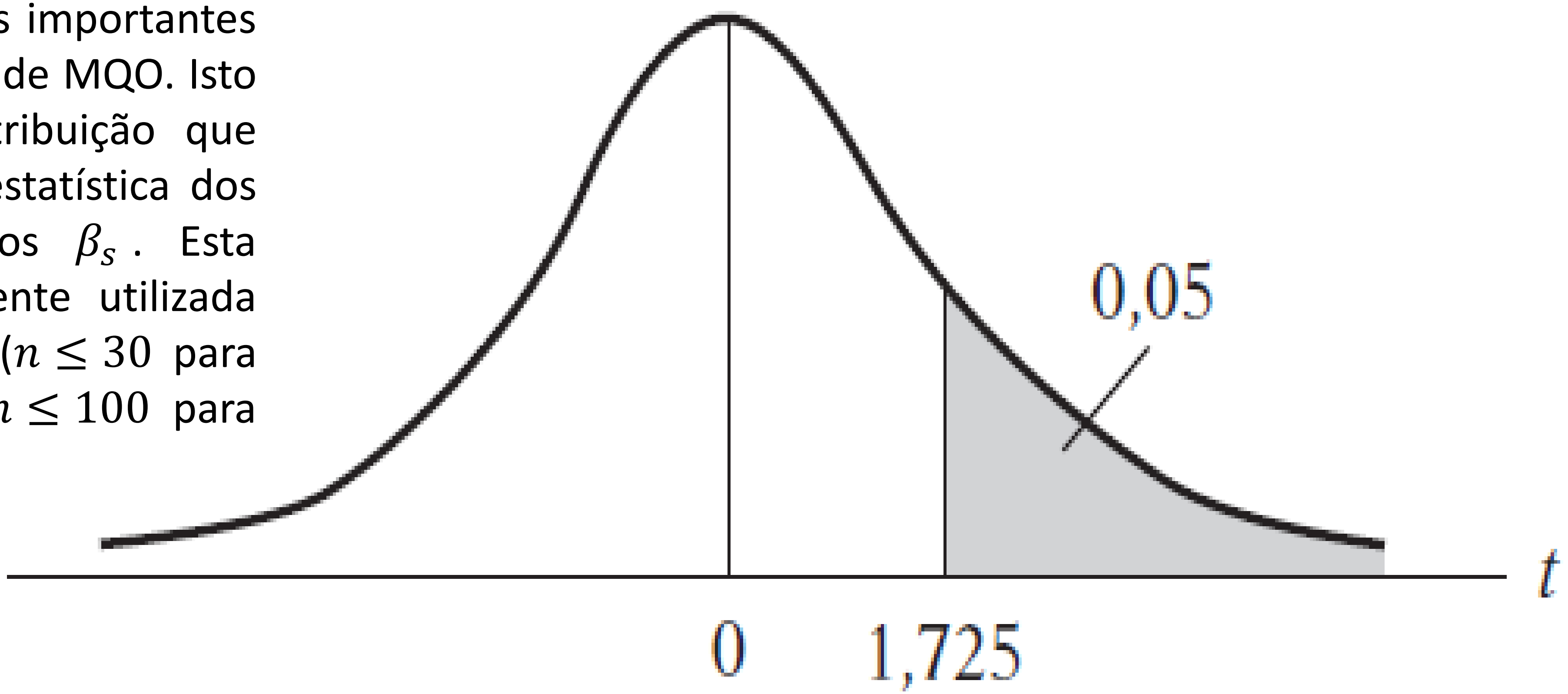
z_0	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4967	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Tabela disponível em:
http://www.dex.ufla.br/thelmasafadi/tabela%20normal.pdf

Distribuição de Probabilidade Contínua

Distribuição t de Student

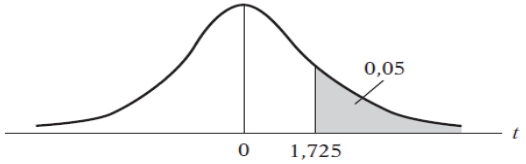
Esta distribuição de probabilidade estatística é uma das mais importantes na aplicação dos modelos de MQO. Isto porque é com esta distribuição que testamos a significância estatística dos parâmetros calculados, os β_s. Esta distribuição é especialmente utilizada para pequenas amostras ($n \leq 30$ para alguns estatísticos; mas $n \leq 100$ para outros estatísticos).



$\Pr(t > 1{,}725) = 0{,}05$ para gl = 20

gl = n - k ➔ n = tamanho amostra; k = número de variáveis explicativas do modelo (desconsidera-se a constante)

Exemplo
$\Pr(t > 2{,}086) = 0{,}025$
$\Pr(t > 1{,}725) = 0{,}05$ para gl = 20
$\Pr(|t| > 1{,}725) = 0{,}10$



Pr/ gl	0,25 0,50	0,10 0,20	0,05 0,10	0,025 0,05	0,01 0,02	0,005 0,010	0,001 0,002
1	1,000	3,078	6,314	12,706	31,821	63,657	318,31
2	0,816	1,886	2,920	4,303	6,965	9,925	22,327
3	0,765	1,638	2,353	3,182	4,541	5,841	10,214
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,683	1,311	1,699	2,045	2,462	2,756	3,396
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090

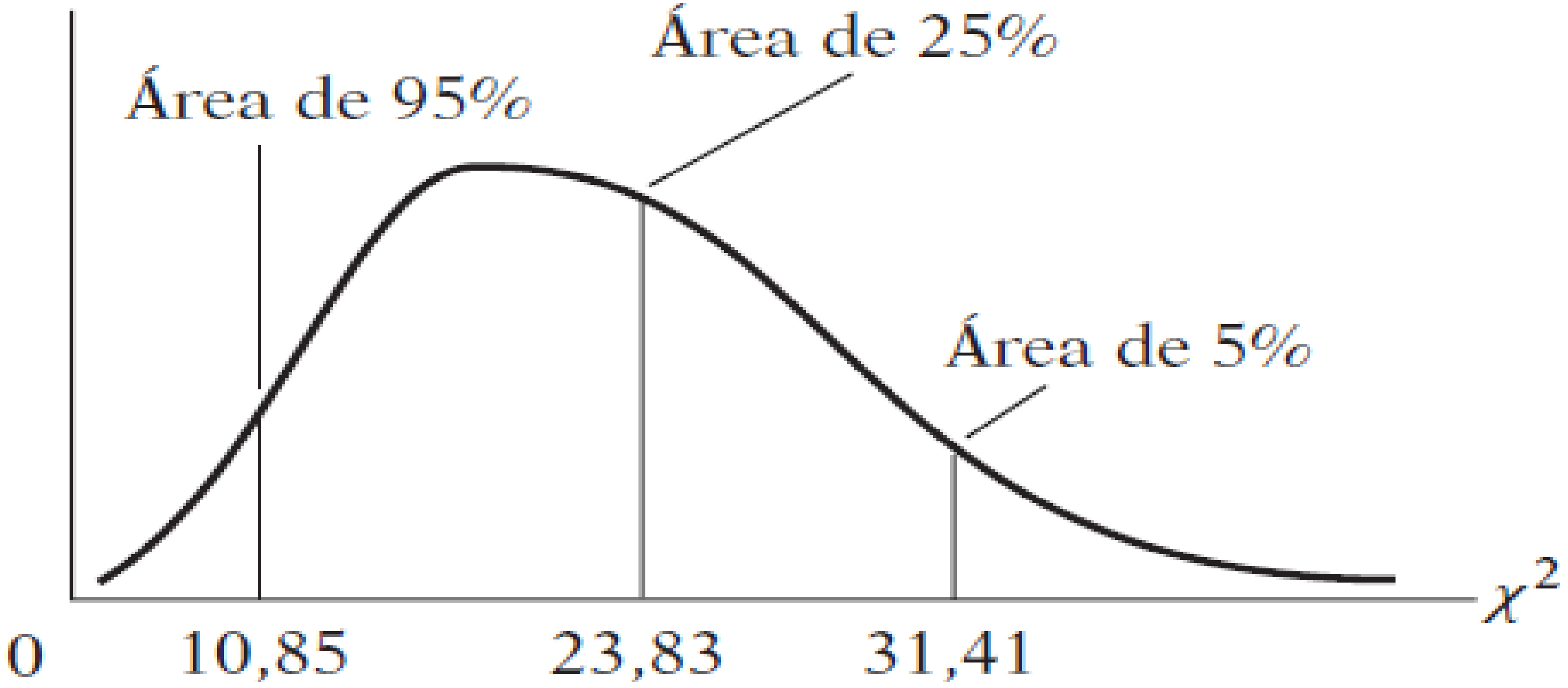
Fonte: PEARSON, E. S.; HARTLEY, H. O. (Eds.). *Biometrika tables for statisticians*. 3. ed. Nova York: Cambridge University Press, 1966. v. 1, tabela 12. Reprodução autorizada pelos editores e curadores da *Biometrika*.
Nota: a menor probabilidade mostrada no título de cada coluna é a área em uma cauda; a probabilidade maior é a área em ambas as caudas.

Distribuição de Probabilidade Contínua

Distribuição Qui-Quadrado

Esta é uma distribuição de probabilidade estatística que tem muitas aplicações em vários testes estatísticos.

Uma das aplicações mais importantes desta distribuição é testar a significância estatística do desvio padrão.



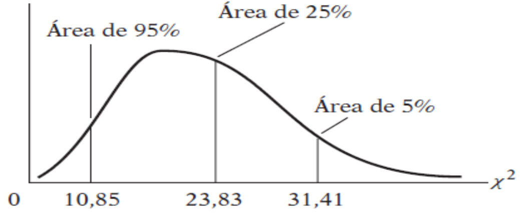
Exemplo

$Pr(\chi^2 > 10{,}85) = 0{,}95$

$Prc(\chi^2 > 23{,}83) = 0{,}25$ para gl = 20

$Pr(\chi^2 > 31{,}41) = 0{,}05$

Exemplo
$Pr(\chi^2 > 10,85) = 0,95$
$Prc(\chi^2 > 23,83) = 0,25$ para gl = 20
$Pr(\chi^2 > 31,41) = 0,05$



Graus de liberdade \ Pr	0,995	0,990	0,975	0,950	0,900
1	392704×10^{-10}	157088×10^{-9}	982069×10^{-9}	393214×10^{-8}	0,0157908
2	0,0100251	0,0201007	0,0506356	0,102587	0,210720
3	0,0717212	0,114832	0,215795	0,351846	0,584375
4	0,206990	0, 297110	0,484419	0,710721	1,063623
5	0,411740	0,554300	0,831211	1,145476	1,61031
6	0,675727	0,872085	1,237347	1,63539	2,20413
7	0,989265	1,239043	1,68987	2,16735	2,83311
8	1,344419	1,646482	2,17973	2,73264	3,48954
9	1,734926	2,087912	2,70039	3,32511	4,16816
10	2,15585	2,55821	3,24697	3,94030	4,86518
11	2,60321	3,05347	3,81575	4,57481	5,57779
12	3,07382	3,57056	4,40379	5,22603	6,30380
13	3,56503	4,10691	5,00874	5,89186	7,04150
14	4,07468	4,66043	5,62872	6,57063	7,78953
15	4,60094	5,22935	6,26214	7,26094	8,54675
16	5,14224	5,81221	6,90766	7,96164	9,31223
17	5,69724	6,40776	7,56418	8,67176	10,0852
18	6,26481	7,01491	8,23075	9,39046	10,8649
19	6,84398	7,63273	8,90655	10,1170	11,6509
20	7,43386	8,26040	9,59083	10,8508	12,4426
21	8,03366	8,89720	10,28293	11,5913	13,2396
22	8,64272	9,54249	10,9823	12,3380	14,0415
23	9,26042	10,19567	11,6885	13,0905	14,8479
24	9,88623	10,8564	12,4011	13,8484	15,6587
25	10,5197	11,5240	13,1197	14,6114	16,4734
26	11,1603	12,1981	13,8439	15,3791	17,2919
27	11,8076	12,8786	14,5733	16,1513	18,1138
28	12,4613	13,5648	15,3079	16,9279	18,9392
29	13,1211	14,2565	16,0471	17,7083	19,7677
30	13,7867	14,9535	16,7908	18,4926	20,5992
40	20,7065	22,1643	24,4331	26,5093	29,0505
50	27,9907	29,7067	32,3574	34,7642	37,6886
60	35,5346	37,4848	40,4817	43,1879	46,4589
70	43,2752	45,4418	48,7576	51,7393	55,3290
80	51,1720	53,5400	57,1532	60,3915	64,2778
90	59,1963	61,7541	65,6466	69,1260	73,2912
100*	67,3276	70,0648	74,2219	77,9295	82,3581

*Para gl maior que 100, a expressão $\sqrt{2\chi^2} - \sqrt{(2k-1)} = Z$ segue a distribuição normal padronizada, em que k representa os graus de liberdade.

TABELA D.4 *(Continuação)*

0,750	0,500	0,250	0,100	0,050	0,025	0,010	0,005
0,1015308	0,454937	1,32330	2,70554	3,84146	5,02389	6,63490	7,87944
0,575364	1,38629	2,77259	4,60517	5,99147	7,37776	9,21034	10,5966
1,212534	2,36597	4,10835	6,25139	7,81473	9,34840	11,3449	12,8381
1,92255	3,35670	5,38527	7,77944	9,48773	11,1433	13,2767	14,8602
2,67460	4,35146	6,62568	9,23635	11,0705	12,8325	15,0863	16,7496
3,45460	5,34812	7,84080	10,6446	12,5916	14,4494	16,8119	18,5476
4,25485	6,34581	9,03715	12,0170	14,0671	16,0128	18,4753	20,2777
5,07064	7,34412	10,2188	13,3616	15,5073	17,5346	20,0902	21,9550
5,89883	8,34283	11,3887	14,6837	16,9190	19,0228	21,6660	23,5893
6,73720	9,34182	12,5489	15,9871	18,3070	20,4831	23,2093	25,1882
7,58412	10,3410	13,7007	17,2750	19,6751	21,9200	24,7250	26,7569
8,43842	11,3403	14,8454	18,5494	21,0261	23,3367	26,2170	28,2995
9,29906	12,3398	15,9839	19,8119	22,3621	24,7356	27,6883	29,8194
10,1653	13,3393	17,1170	21,0642	23,6848	26,1190	29,1413	31,3193
11,0365	14,3389	18,2451	22,3072	24,9958	27,4884	30,5779	32,8013
11,9122	15,3385	19,3688	23,5418	26,2962	28,8454	31,9999	34,2672
12,7919	16,3381	20,4887	24,7690	27,5871	30,1910	33,4087	35,7185
13,6753	17,3379	21,6049	25,9894	28,8693	31,5264	34,8053	37,1564
14,5620	18,3376	22,7178	27,2036	30,1435	32,8523	36,1908	38,5822
15,4518	19,3374	23,8277	28,4120	31,4104	34,1696	37,5662	39,9968
16,3444	20,3372	24,9348	29,6151	32,6705	35,4789	38,9321	41,4010
17,2396	21,3370	26,0393	30,8133	33,9244	36,7807	40,2894	42,7956
18,1373	22,3369	27,1413	32,0069	35,1725	38,0757	41,6384	44,1813
19,0372	23,3367	28,2412	33,1963	36,4151	39,3641	42,9798	45,5585
19,9393	24,3366	29,3389	34,3816	37,6525	40,6465	44,3141	46,9278
20,8434	25,3364	30,4345	35,5631	38,8852	41,9232	45,6417	48,2899
21,7494	26,3363	31,5284	36,7412	40,1133	43,1944	46,9630	49,6449
22,6572	27,3363	32,6205	37,9159	41,3372	44,4607	48,2782	50,9933
23,5666	28,3362	33,7109	39,0875	42,5569	45,7222	49,5879	52,3356
24,4776	29,3360	34,7998	40,2560	43,7729	46,9792	50,8922	53,6720
33,6603	39,3354	45,6160	51,8050	55,7585	59,3417	63,6907	66,7659
42,9421	49,3349	56,3336	63,1671	67,5048	71,4202	76,1539	79,4900
52,2938	59,3347	66,9814	74,3970	79,0819	83,2976	88,3794	91,9517
61,6983	69,3344	77,5766	85,5271	90,5312	95,0231	100,425	104,215
71,1445	79,3343	88,1303	96,5782	101,879	106,629	112,329	116,321
80,6247	89,3342	98,6499	107,565	113,145	118,136	124,116	128,299
90,1332	99,3341	109,141	118,498	124,342	129,561	135,807	140,169

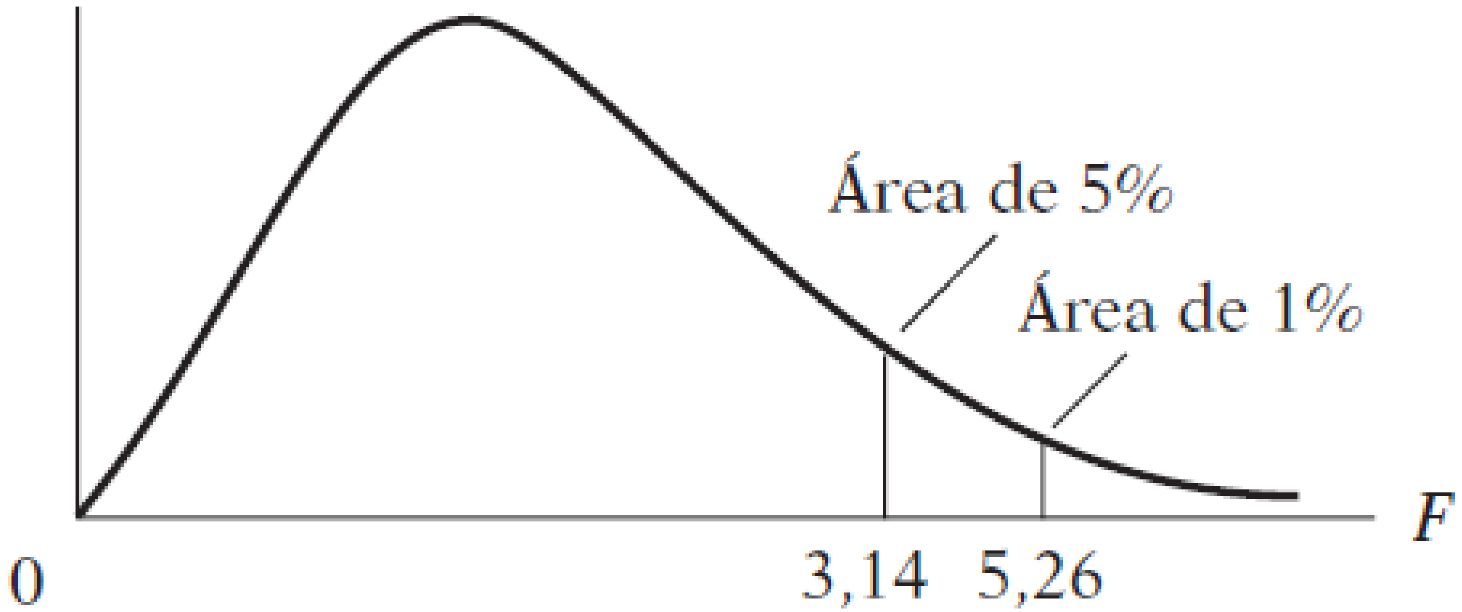
Fonte: resumida de PEARSON, E. S.; HARTLEY, H. O. (Eds.). *Biometrika tables for statisticians*. 3. ed. Nova York: Cambridge University Press, 1966. v. 1, tabela 12.

Distribuição de Probabilidade Contínua

Distribuição F de Fisher-Snedecor

Esta distribuição de probabilidade tem várias aplicações. Uma delas é testar a existência da reta de regressão no MQO.

Outros testes estatísticos se utilizam desta distribuição de probabilidade.

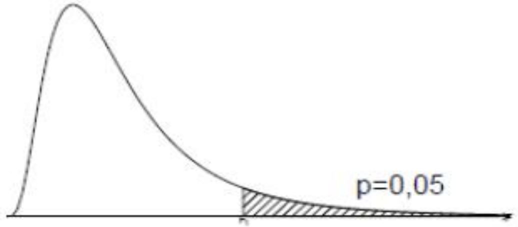


Exemplo: $\Pr(F > 3{,}14) = 0{,}05$ para gl $N_1 = 10$

$\Pr(F > 5{,}26) = 0{,}01$ e $N_2 = 9$

Gl do numerador (N_1) = m = número de restrições lineares (número de parâmetros); gl do denominador (N2) = n-k = n é o tamanho da amostra e k o número de parâmetros da regressão.

Distribuição F de Snedecor a 5% (p=0.05)



	1	2	3	4	5	6	7	8	9	10	12	14	15	16	18	20	30	40	60	120
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.42	19.43	19.43	19.44	19.45	19.46	19.47	19.48	19.49
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.71	8.70	8.69	8.67	8.66	8.62	8.59	8.57	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.87	5.86	5.84	5.82	5.80	5.75	5.72	5.69	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.64	4.62	4.60	4.58	4.56	4.50	4.46	4.43	4.40
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.96	3.94	3.92	3.90	3.87	3.81	3.77	3.74	3.70
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.53	3.51	3.49	3.47	3.44	3.38	3.34	3.30	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.24	3.22	3.20	3.17	3.15	3.08	3.04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.03	3.01	2.99	2.96	2.94	2.86	2.83	2.79	2.75
10	$\boxed{4.96}$	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.86	2.85	2.83	2.80	2.77	2.70	2.66	2.62	2.58
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.74	2.72	2.70	2.67	2.65	2.57	2.53	2.49	2.45
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.64	2.62	2.60	2.57	2.54	2.47	2.43	2.38	2.34
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.55	2.53	2.51	2.48	2.46	2.38	2.34	2.30	2.25
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.48	2.46	2.44	2.41	2.39	2.31	2.27	2.22	2.18
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.42	2.40	2.38	2.35	2.33	2.25	2.20	2.16	2.11
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.37	2.35	2.33	2.30	2.28	2.19	2.15	2.11	2.06
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.33	2.31	2.29	2.26	2.23	2.15	2.10	2.06	2.01
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.29	2.27	2.25	2.22	2.19	2.11	2.06	2.02	1.97
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.26	2.23	2.21	2.18	2.16	2.07	2.03	1.98	1.93
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.22	2.20	2.18	2.15	2.12	2.04	1.99	1.95	1.90
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.20	2.18	2.16	2.12	2.10	2.01	1.96	1.92	1.87
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.17	2.15	2.13	2.10	2.07	1.98	1.94	1.89	1.84
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.15	2.13	2.11	2.08	2.05	1.96	1.91	1.86	1.81
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.13	2.11	2.09	2.05	2.03	1.94	1.89	1.84	1.79
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.11	2.09	2.07	2.04	2.01	1.92	1.87	1.82	1.77
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.09	2.07	2.05	2.02	1.99	1.90	1.85	1.80	1.75
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.08	2.06	2.04	2.00	1.97	1.88	1.84	1.79	1.73
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.06	2.04	2.02	1.99	1.96	1.87	1.82	1.77	1.71
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.05	2.03	2.01	1.97	1.94	1.85	1.81	1.75	1.70
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.04	2.01	1.99	1.96	1.93	1.84	1.79	1.74	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.95	1.92	1.90	1.87	1.84	1.74	1.69	1.64	1.58
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.86	1.84	1.82	1.78	1.75	1.65	1.59	1.53	1.47
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.78	1.75	1.73	1.69	1.66	1.55	1.50	1.43	1.35

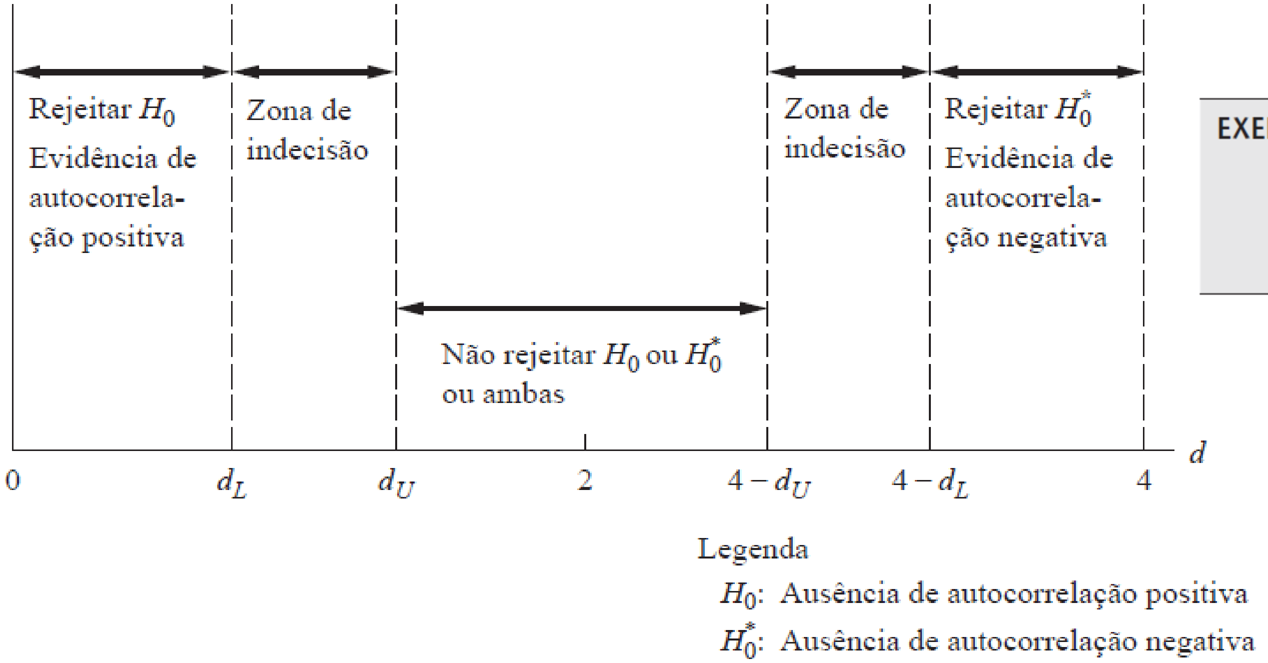
Tabela 5: Quantis da Distribuição F para probabilidade $p = P[F \geq F_t] = 0,05$. Graus de liberdade do numerador dado no topo e do denominador na margem esquerda.

Obtida em :
https://docs.ufpr.br/~niveam/ce071/An%C3%A1lise%20de%20Vari%C3%A2ncia.pdf

Distribuição de Probabilidade de Durbin-Watson (d)

- Serve para examinar a autocorrelação nos resíduos, ou seja, representa uma análise da existência de que as observações estejam correlacionadas. Por exemplo: comportamento dos preços de determinado produto. É um teste para verificar uma das hipóteses dos modelos de MQO.

$$d = \frac{\sum_{t=2}^{t=n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{t=n}\hat{u}_t^2}$$



EXEMPLO 1 Se n = 40 e $k' = 4$, $d_L = 1{,}285$ e $d_U = 1{,}721$. Se um valor calculado de d é menor que 1,285, há evidência de correlação serial positiva de primeira ordem; se é maior que 1,721, não há nenhuma evidência de correlação serial positiva de primeira ordem; mas, se d está entre o limite inferior e o limite superior, a evidência é inconclusiva em relação à presença ou ausência de correlação serial positiva de primeira ordem.

TABELA D.5A Estatística d de Durbin-Watson: pontos de significância de d_L e d_U em níveis de significância de 0,05

	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$		$k'=8$		$k'=9$		$k'=10$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,610	1,400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0,700	1,356	0,467	1,896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0,763	1,332	0,559	1,777	0,368	2,287	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0,824	1,320	0,629	1,699	0,455	2,128	0,296	2,588	—	—	—	—	—	—	—	—	—	—	—	—
10	0,879	1,320	0,697	1,641	0,525	2,016	0,376	2,414	0,243	2,822	—	—	—	—	—	—	—	—	—	—
11	0,927	1,324	0,658	1,604	0,595	1,928	0,444	2,283	0,316	2,645	0,203	3,005	—	—	—	—	—	—	—	—
12	0,971	1,331	0,812	1,579	0,658	1,864	0,512	2,177	0,379	2,506	0,268	2,832	0,171	3,149	—	—	—	—	—	—
13	1,010	1,340	0,861	1,562	0,715	1,816	0,574	2,094	0,445	2,390	0,328	2,692	0,230	2,985	0,147	3,266	—	—	—	—
14	1,045	1,350	0,905	1,551	0,767	1,779	0,632	2,030	0,505	2,296	0,389	2,572	0,286	2,848	0,200	3,111	0,127	3,360	—	—
15	1,077	1,361	0,946	1,543	0,814	1,750	0,685	1,977	0,562	2,220	0,447	2,472	0,343	2,727	0,251	2,979	0,175	3,216	0,111	3,438
16	1,106	1,371	0,982	1,539	0,857	1,728	0,734	1,935	0,615	2,157	0,502	2,388	0,398	2,624	0,304	2,860	0,222	3,090	0,155	3,304
17	1,133	1,381	1,015	1,536	0,897	1,710	0,779	1,900	0,664	2,104	0,554	2,318	0,451	2,537	0,356	2,757	0,272	2,975	0,198	3,184
18	1,158	1,391	1,046	1,535	0,933	1,696	0,820	1,872	0,710	2,060	0,603	2,257	0,502	2,461	0,407	2,667	0,321	2,873	0,244	3,073
19	1,180	1,401	1,074	1,536	0,967	1,685	0,859	1,848	0,752	2,023	0,649	2,206	0,549	2,396	0,456	2,589	0,369	2,783	0,290	2,974
20	1,201	1,411	1,100	1,537	0,998	1,676	0,894	1,828	0,792	1,991	0,692	2,162	0,595	2,339	0,502	2,521	0,416	2,704	0,336	2,885
21	1,221	1,420	1,125	1,538	1,026	1,669	0,927	1,812	0,829	1,964	0,732	2,124	0,637	2,290	0,547	2,460	0,461	2,633	0,380	2,806
22	1,239	1,429	1,147	1,541	1,053	1,664	0,958	1,797	0,863	1,940	0,769	2,090	0,677	2,246	0,588	2,407	0,504	2,571	0,424	2,734
23	1,257	1,437	1,168	1,543	1,078	1,660	0,986	1,785	0,895	1,920	0,804	2,061	0,715	2,208	0,628	2,360	0,545	2,514	0,465	2,670
24	1,273	1,446	1,188	1,546	1,101	1,656	1,013	1,775	0,925	1,902	0,837	2,035	0,751	2,174	0,666	2,318	0,584	2,464	0,506	2,613
25	1,288	1,454	1,206	1,550	1,123	1,654	1,038	1,767	0,953	1,886	0,868	2,012	0,784	2,144	0,702	2,280	0,621	2,419	0,544	2,560
26	1,302	1,461	1,224	1,553	1,143	1,652	1,062	1,759	0,979	1,873	0,897	1,992	0,816	2,117	0,735	2,246	0,657	2,379	0,581	2,513
27	1,316	1,469	1,240	1,556	1,162	1,651	1,084	1,753	1,004	1,861	0,925	1,974	0,845	2,093	0,767	2,216	0,691	2,342	0,616	2,470
28	1,328	1,476	1,255	1,560	1,181	1,650	1,104	1,747	1,028	1,850	0,951	1,958	0,874	2,071	0,798	2,188	0,723	2,309	0,650	2,431
29	1,341	1,483	1,270	1,563	1,198	1,650	1,124	1,743	1,050	1,841	0,975	1,944	0,900	2,052	0,826	2,164	0,753	2,278	0,682	2,396
30	1,352	1,489	1,284	1,567	1,214	1,650	1,143	1,739	1,071	1,833	0,998	1,931	0,926	2,034	0,854	2,141	0,782	2,251	0,712	2,363
31	1,363	1,496	1,297	1,570	1,229	1,650	1,160	1,735	1,090	1,825	1,020	1,920	0,950	2,018	0,879	2,120	0,810	2,226	0,741	2,333
32	1,373	1,502	1,309	1,574	1,244	1,650	1,177	1,732	1,109	1,819	1,041	1,909	0,972	2,004	0,904	2,102	0,836	2,203	0,769	2,306
33	1,383	1,508	1,321	1,577	1,258	1,651	1,193	1,730	1,127	1,813	1,061	1,900	0,994	1,991	0,927	2,085	0,861	2,181	0,795	2,281
34	1,393	1,514	1,333	1,580	1,271	1,652	1,208	1,728	1,144	1,808	1,080	1,891	1,015	1,979	0,950	2,069	0,885	2,162	0,821	2,257
35	1,402	1,519	1,343	1,584	1,283	1,653	1,222	1,726	1,160	1,803	1,097	1,884	1,034	1,967	0,971	2,054	0,908	2,144	0,845	2,236
36	1,411	1,525	1,354	1,587	1,295	1,654	1,236	1,724	1,175	1,799	1,114	1,877	1,053	1,957	0,991	2,041	0,930	2,127	0,868	2,216
37	1,419	1,530	1,364	1,590	1,307	1,655	1,249	1,723	1,190	1,795	1,131	1,870	1,071	1,948	1,011	2,029	0,951	2,112	0,891	2,198
38	1,427	1,535	1,373	1,594	1,318	1,656	1,261	1,722	1,204	1,792	1,146	1,864	1,088	1,939	1,029	2,017	0,970	2,098	0,912	2,180
39	1,435	1,540	1,382	1,597	1,328	1,658	1,273	1,722	1,218	1,789	1,161	1,859	1,104	1,932	1,047	2,007	0,990	2,085	0,932	2,164
40	1,442	1,544	1,391	1,600	1,338	1,659	1,285	1,721	1,230	1,786	1,175	1,854	1,120	1,924	1,064	1,997	1,008	2,072	0,952	2,149
45	1,475	1,566	1,430	1,615	1,383	1,666	1,336	1,720	1,287	1,776	1,238	1,835	1,189	1,895	1,139	1,958	1,089	2,022	1,038	2,088
50	1,503	1,585	1,462	1,628	1,421	1,674	1,378	1,721	1,335	1,771	1,291	1,822	1,246	1,875	1,201	1,930	1,156	1,986	1,110	2,044
55	1,528	1,601	1,490	1,641	1,452	1,681	1,414	1,724	1,374	1,768	1,334	1,814	1,294	1,861	1,253	1,909	1,212	1,959	1,170	2,010
60	1,549	1,616	1,514	1,652	1,480	1,689	1,444	1,727	1,408	1,767	1,372	1,808	1,335	1,850	1,298	1,894	1,260	1,939	1,222	1,984
65	1,567	1,629	1,536	1,662	1,503	1,696	1,471	1,731	1,438	1,767	1,404	1,805	1,370	1,843	1,336	1,882	1,301	1,923	1,266	1,964
70	1,583	1,641	1,554	1,672	1,525	1,703	1,494	1,735	1,464	1,768	1,433	1,802	1,401	1,837	1,369	1,873	1,337	1,910	1,305	1,948
75	1,598	1,652	1,571	1,680	1,543	1,709	1,515	1,739	1,487	1,770	1,458	1,801	1,428	1,834	1,399	1,867	1,369	1,901	1,339	1,935
80	1,611	1,662	1,586	1,688	1,560	1,715	1,534	1,743	1,507	1,772	1,480	1,801	1,453	1,831	1,425	1,861	1,397	1,893	1,369	1,925
85	1,624	1,671	1,600	1,696	1,575	1,721	1,550	1,747	1,525	1,774	1,500	1,801	1,474	1,829	1,448	1,857	1,422	1,886	1,396	1,916
90	1,635	1,679	1,612	1,703	1,589	1,726	1,566	1,751	1,542	1,776	1,518	1,801	1,494	1,827	1,469	1,854	1,445	1,881	1,420	1,909
95	1,645	1,687	1,623	1,709	1,602	1,732	1,579	1,755	1,557	1,778	1,535	1,802	1,512	1,827	1,489	1,852	1,465	1,877	1,442	1,903
100	1,654	1,694	1,634	1,715	1,613	1,736	1,592	1,758	1,571	1,780	1,550	1,803	1,528	1,826	1,506	1,850	1,484	1,874	1,462	1,898
150	1,720	1,746	1,706	1,760	1,693	1,774	1,679	1,788	1,665	1,802	1,651	1,817	1,637	1,832	1,622	1,847	1,608	1,862	1,594	1,877
200	1,758	1,778	1,748	1,789	1,738	1,799	1,728	1,810	1,718	1,820	1,707	1,831	1,697	1,841	1,686	1,852	1,675	1,863	1,665	1,874

(*Continua*)

TABELA D.5A Estatística d de Durbin-Watson: pontos de significância de d_L e d_U em níveis de significância de 0,05 (*Continuação*)

	$k'=11$		$k'=12$		$k'=13$		$k'=14$		$k'=15$		$k'=16$		$k'=17$		$k'=18$		$k'=19$		$k'=20$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0,098	3,503	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0,138	3,378	0,087	3,557	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0,177	3,265	0,123	3,441	0,078	3,603	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0,220	3,159	0,160	3,335	0,111	3,496	0,070	3,642	—	—	—	—	—	—	—	—	—	—	—	—
20	0,263	3,063	0,200	3,234	0,145	3,395	0,100	3,542	0,063	3,676	—	—	—	—	—	—	—	—	—	—
21	0,307	2,976	0,240	3,141	0,182	3,300	0,132	3,448	0,091	3,583	0,058	3,705	—	—	—	—	—	—	—	—
22	0,349	2,897	0,281	3,057	0,220	3,211	0,166	3,358	0,120	3,495	0,083	3,619	0,052	3,731	—	—	—	—	—	—
23	0,391	2,826	0,322	2,979	0,259	3,128	0,202	3,272	0,153	3,409	0,110	3,535	0,076	3,650	0,048	3,753	—	—	—	—
24	0,431	2,761	0,362	2,908	0,297	3,053	0,239	3,193	0,186	3,327	0,141	3,454	0,101	3,572	0,070	3,678	0,044	3,773	—	—
25	0,470	2,702	0,400	2,844	0,335	2,983	0,275	3,119	0,221	3,251	0,172	3,376	0,130	3,494	0,094	3,604	0,065	3,702	0,041	3,790
26	0,508	2,649	0,438	2,784	0,373	2,919	0,312	3,051	0,256	3,179	0,205	3,303	0,160	3,420	0,120	3,531	0,087	3,632	0,060	3,724
27	0,544	2,600	0,475	2,730	0,409	2,859	0,348	2,987	0,291	3,112	0,238	3,233	0,191	3,349	0,149	3,460	0,112	3,563	0,081	3,658
28	0,578	2,555	0,510	2,680	0,445	2,805	0,383	2,928	0,325	3,050	0,271	3,168	0,222	3,283	0,178	3,392	0,138	3,495	0,104	3,592
29	0,612	2,515	0,544	2,634	0,479	2,755	0,418	2,874	0,359	2,992	0,305	3,107	0,254	3,219	0,208	3,327	0,166	3,431	0,129	3,528
30	0,643	2,477	0,577	2,592	0,512	2,708	0,451	2,823	0,392	2,937	0,337	3,050	0,286	3,160	0,238	3,266	0,195	3,368	0,156	3,465
31	0,674	2,443	0,608	2,553	0,545	2,665	0,484	2,776	0,425	2,887	0,370	2,996	0,317	3,103	0,269	3,208	0,224	3,309	0,183	3,406
32	0,703	2,411	0,638	2,517	0,576	2,625	0,515	2,733	0,457	2,840	0,401	2,946	0,349	3,050	0,299	3,153	0,253	3,252	0,211	3,348
33	0,731	2,382	0,668	2,484	0,606	2,588	0,546	2,692	0,488	2,796	0,432	2,899	0,379	3,000	0,329	3,100	0,283	3,198	0,239	3,293
34	0,758	2,355	0,695	2,454	0,634	2,554	0,575	2,654	0,518	2,754	0,462	2,854	0,409	2,954	0,359	3,051	0,312	3,147	0,267	3,240
35	0,783	2,330	0,722	2,425	0,662	2,521	0,604	2,619	0,547	2,716	0,492	2,813	0,439	2,910	0,388	3,005	0,340	3,099	0,295	3,190
36	0,808	2,306	0,748	2,398	0,689	2,492	0,631	2,586	0,575	2,680	0,520	2,774	0,467	2,868	0,417	2,961	0,369	3,053	0,323	3,142
37	0,831	2,285	0,772	2,374	0,714	2,464	0,657	2,555	0,602	2,646	0,548	2,738	0,495	2,829	0,445	2,920	0,397	3,009	0,351	3,097
38	0,854	2,265	0,796	2,351	0,739	2,438	0,683	2,526	0,628	2,614	0,575	2,703	0,522	2,792	0,472	2,880	0,424	2,968	0,378	3,054
39	0,875	2,246	0,819	2,329	0,763	2,413	0,707	2,499	0,653	2,585	0,600	2,671	0,549	2,757	0,499	2,843	0,451	2,929	0,404	3,013
40	0,896	2,228	0,840	2,309	0,785	2,391	0,731	2,473	0,678	2,557	0,626	2,641	0,575	2,724	0,525	2,808	0,477	2,892	0,430	2,974
45	0,988	2,156	0,938	2,225	0,887	2,296	0,838	2,367	0,788	2,439	0,740	2,512	0,692	2,586	0,644	2,659	0,598	2,733	0,553	2,807
50	1,064	2,103	1,019	2,163	0,973	2,225	0,927	2,287	0,882	2,350	0,836	2,414	0,792	2,479	0,747	2,544	0,703	2,610	0,660	2,675
55	1,129	2,062	1,087	2,116	1,045	2,170	1,003	2,225	0,961	2,281	0,919	2,338	0,877	2,396	0,836	2,454	0,795	2,512	0,754	2,571
60	1,184	2,031	1,145	2,079	1,106	2,127	1,068	2,177	1,029	2,227	0,990	2,278	0,951	2,330	0,913	2,382	0,874	2,434	0,836	2,487
65	1,231	2,006	1,195	2,049	1,160	2,093	1,124	2,138	1,088	2,183	1,052	2,229	1,016	2,276	0,980	2,323	0,944	2,371	0,908	2,419
70	1,272	1,986	1,239	2,026	1,206	2,066	1,172	2,106	1,139	2,148	1,105	2,189	1,072	2,232	1,038	2,275	1,005	2,318	0,971	2,362
75	1,308	1,970	1,277	2,006	1,247	2,043	1,215	2,080	1,184	2,118	1,153	2,156	1,121	2,195	1,090	2,235	1,058	2,275	1,027	2,315
80	1,340	1,957	1,311	1,991	1,283	2,024	1,253	2,059	1,224	2,093	1,195	2,129	1,165	2,165	1,136	2,201	1,106	2,238	1,076	2,275
85	1,369	1,946	1,342	1,977	1,315	2,009	1,287	2,040	1,260	2,073	1,232	2,105	1,205	2,139	1,177	2,172	1,149	2,206	1,121	2,241
90	1,395	1,937	1,369	1,966	1,344	1,995	1,318	2,025	1,292	2,055	1,266	2,085	1,240	2,116	1,213	2,148	1,187	2,179	1,160	2,211
95	1,418	1,929	1,394	1,956	1,370	1,984	1,345	2,012	1,321	2,040	1,296	2,068	1,271	2,097	1,247	2,126	1,222	2,156	1,197	2,186
100	1,439	1,923	1,416	1,948	1,393	1,974	1,371	2,000	1,347	2,026	1,324	2,053	1,301	2,080	1,277	2,108	1,253	2,135	1,229	2,164
150	1,579	1,892	1,564	1,908	1,550	1,924	1,535	1,940	1,519	1,956	1,504	1,972	1,489	1,989	1,474	2,006	1,458	2,023	1,443	2,040
200	1,654	1,885	1,643	1,896	1,632	1,908	1,621	1,919	1,610	1,931	1,599	1,943	1,588	1,955	1,576	1,967	1,565	1,979	1,554	1,991

Nota: n = número de observações, k' = número de variáveis explanatórias, excluindo o termo constante.

Fonte: Esta tabela é uma extensão da tabela original de Durbin-Watson, reproduzida de SAVIN, N. E.; WHITE, K. J. "The Durbin-Watson test for serial correlation with extreme small samples or many regressors." *Econometrica,* v. 45, p. 1.989-1996, nov. 1977. Ela foi corrigida por FAREBROTHER, R. W. *Econometrica,* v. 48, p. 1.554, set. 1980. Reprodução autorizada pela Econometric Society.

Distribuição de Probabilidade Contínua

Distribuição Logística

- Consideremos uma base de dados – amostra – composta por exemplo: indivíduos, domicílios, municípios, países, etc.
- Para cada observação da amostra, tem-se uma variável Y_i que assume valores 0 ou 1 (zero se não possui determinada característica e 1 se possui esta característica). Outra possibilidade são variáveis qualitativas não numéricas do tipo: bom:ruim; aprovado:reprovado; etc.
- Tem-se um conjunto de colunas que representam as variáveis explicativas - $X_{1,i}$; $X_{2,i}$; $X_{3,i}$; … … ; $X_{n,i}$.
- O objetivo da regressão logística é assumir que cada valor individual Y_i corresponde a uma variável aleatória de Bernoulli, com probabilidade de sucesso (por exemplo, pessoa identificada como infectada com Covid-19) dada por Prob[Y_i =1] = p_i .
- Logo, a probabilidade de fracasso (por exemplo, pessoa não identificada como infectada com Covid-19) é dada por Prob[Y_i=0] = 1-p_i.

Distribuição de Probabilidade Contínua

Distribuição Logística

O objetivo é fazer com que esta variável Y dependa das variáveis explicativas;

$$Prob[Y_i = 1] = p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})}}$$

- Esta formulação implica que a probabilidade p_i vai se situar no intervalo (0,1).
- Por exemplo, dada a nossa amostra de infectados pelo Covid-19, a resposta desse modelo é a probabilidade de estar infectado pelo Covid-19.
- Além disso o modelo apresentaria quais são as variáveis explicativas que realmente explicam a probabilidade de estar infectado pelo Covid-19.
- Quando os betas são positivos e quando os valores dos X_s aumenta, a probabilidade de sucesso aumenta.
- No “R”, para Machine Learning, experimente usar o pacote H2O.

Distribuição de Probabilidade Contínua

Distribuição Logística