




PUC Minas
DIRETORIA DE
EDUCAÇÃO CONTINUADA

Pós Graduação Lato Sensu

**Ciência de Dados e Big
Data / Análise
Estatística de Dados**



DIRETORIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas

Quem sou eu?

- Arthur Pereira de Gouveia e Silva
- Casado, 38 anos, pai de uma linda garota de dois anos de idade
- Torcedor “azul grená” do FC Barcelona
- Tolkien, Batman, Poe, Portnoy, Peart, I. Cavallera, Metal, Python, Rubik, Android, Google, Microsoft, Linux, Star Wars, PSOne, PS2, PS3, PS4...
- Telecom → Automação → Manutenção → Saneamento
→ Gestão → Estatística → Data Mining → Data Science



Como me encontrar



@arthurg



gouveia.arthur@gmail.com



www.linkedin.com/in/arthur-gouveia/



@arthur_gouveia



github.com/arthur-gouveia



E quem são vocês?

- Já estudaram estatística?
- Há quanto tempo?
- Gostam?





Programação

1. Introdução

- A importância da Estatística para o cidadão comum
- A importância da Estatística para o Cientista de Dados
- Natureza dos dados
- Amostragem

2. Estatística Descritiva

- Resumo de dados com tabelas de frequências
- Medidas de tendência central
- Medidas de variação
- Medidas de posição
- Correlação
- Representações gráficas



Programação (cont.)

3. Probabilidade

- Fundamentos
- Regra da Adição
- Regra da multiplicação
- Teorema de Bayes

4. Inferência

- Estimativas e tamanhos de amostras
- Testes de hipóteses
- Inferências com base em duas amostras

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Avaliação


Exercícios no www.edmodo.com (disponível também para *mobile*)
<https://edmo.do/j/kpqrtty> ou **Join a Group** e use o código **a8jntt**



Vocês entrarão no grupo AED – Análise Estatística de Dados e deverão completar os *quizzes*. Vocês têm 1h para completar cada *quiz*.

50%

Da nota final

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Avaliação

Trabalho **individual** final de análise de dados

Um exercício simples! Pegue um conjunto de dados e faça uma análise estatística do mesmo.

O que será avaliado:

- Técnicas e ferramentas utilizadas
- Uso correto dos gráficos
- Exatidão dos cálculos
- Qualidade da análise

Sugestões de fontes de dados:

- <https://archive.ics.uci.edu/ml/datasets.html>
- <https://github.com/caesar0301/awesome-public-datasets>
- <http://dados.gov.br/dataset>
- <http://www.portaltransparencia.gov.br/downloads/>

50%

Da nota final

I. Introdução



Conceitos

- **Estatística (ciência):** Coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões
- **População:** Coleção completa de todos os elementos (valores, pessoas, medidas etc.) a serem estudados
- **Amostra:** Subcoleção de elementos extraídos de uma população
- **Parâmetro:** Medida numérica que descreve uma característica de uma população
- **Estatística:** Medida numérica que descreve uma característica de uma amostra
- **Censo:** Coleção de dados relativos a todos os elementos de uma população
- **Pesquisa:** Coleção de dados relativos a uma amostra da população



A origem da Estatística



*Statistik: Análise de
dados do Estado*



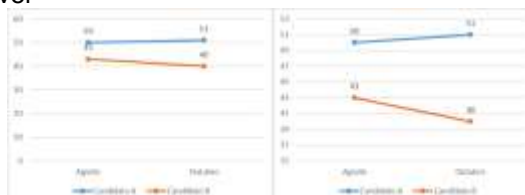
Uso atual da Estatística





A importância da Estatística

- Para o cidadão comum:
 - Nos últimos anos motoristas entre 16 e 19 anos causaram 1,5 milhão de acidentes contra apenas 540 mil causados por motoristas de 70 anos ou mais
 - Melhoramos os resultados em 100%
 - 90% dos carros vendidos nos últimos 20 anos ainda estão em circulação
 - Pesquisa auto selecionável
 - Gráficos tendenciosos



A importância da Estatística

- Para o Cientista de Dados
 - A essa altura do campeonato você ainda tem essa dúvida?!?!?!?



- Começando pelo começo: *"It's an absolute myth that you can send an algorithm over raw data and have insights pop up"*
Jeffrey Heer
- Você terá que preparar e analisar dados, consolidar, comparar e apresentar resultados, preparar experimentos e, claro, aplicar e talvez até desenvolver modelos estatísticos.

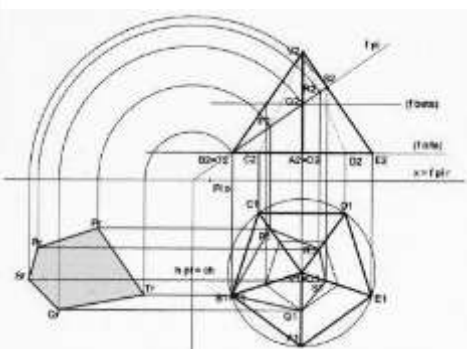


A importância da Estatística

- Se você ainda tem dúvidas...



Áreas da estatística



Estatística Descritiva:

Apenas coleta, organiza, descreve e analisa os dados. Aqui não são tiradas conclusões



Áreas da estatística

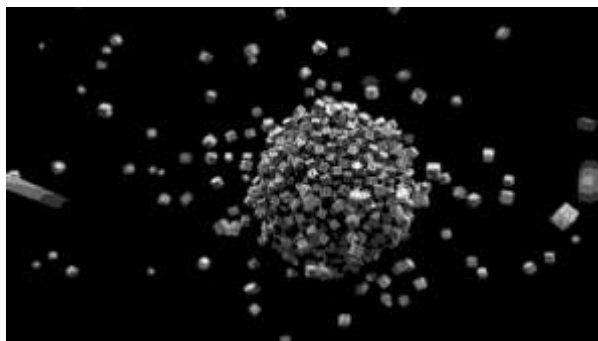
Estatística Inferencial ou Indutiva:


A partir da análise dos dados
são tiradas conclusões e
realizadas inferências sobre os
parâmetros populacionais



Dados!

- Se a Estatística é a ciência para analisar, interpretar, amassar e torturar os dados...
- Vamos falar sobre dados

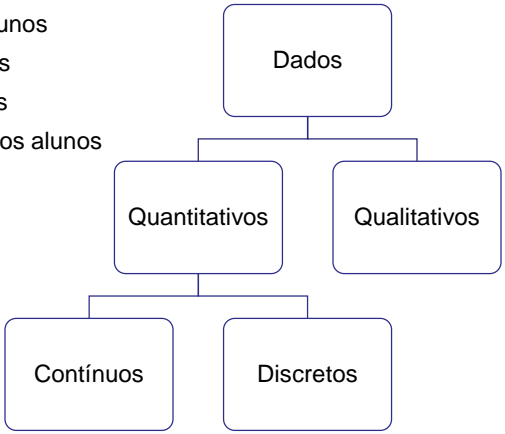


SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Natureza dos dados


Podemos classificar os dados de acordo com seu tipo

- Quantidade de alunos
- Alturas dos alunos
- Idades dos alunos
- Nomes ou sexo dos alunos



```

graph TD
    Dados[Dados] --> Quantitativos[Quantitativos]
    Dados --> Qualitativos[Qualitativos]
    Quantitativos --> Contínuos[Contínuos]
    Quantitativos --> Discretos[Discretos]
  
```

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Natureza dos dados

E também de acordo com o nível de mensuração

- **Nominal:** Os dados não podem ser organizados nem usados em cálculos
 - Marca, tipo, cor, sexo, categoria
- **Ordinal:** Podemos ordenar mas as diferenças não fazem sentido
 - Nota (conceito), Posição, Tamanho (P, M, G)
- **Intervalar:** Análogo ao ordinal mas as diferenças fazem sentido apesar de não haver um *zero absoluto*
 - Ano, Temperatura em °C
- **Racional:** Possui um *zero absoluto* significando ausência
 - Peso, Preço, Altura, Quantidade de pessoa, Tempo*



Amostragem

- Coletas de dados que **representam** a população
- Amostra aleatória
 - Todos os elementos tem a mesma chance de serem escolhidos
- Amostra estratificada
 - Dividimos em grupos com a mesma característica
- Amostragem sistemática:
 - Escolhemos cada $k^{\text{ésimo}}$ elemento
- Amostra por conglomerados
 - Separamos em áreas e pegamos todos os elementos de algumas





Resumo Cap I

- Estatística: Uma ciência **extremamente** importante para todos, especialmente para o Cientista de Dados
- Estatística Descritiva vs Estatística Inferencial
- Cuidado com o tipo de dados e o nível de mensuração!
- Existem diversos tipo de amostragem mas a mais comum é a **estratificada**.

2. Estatística Descritiva





SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas

Estatística Descritiva

- O que diferencia o bom do excelente
- Pode parecer básico...
- ... e **realmente** é básico
- Mas isso não é uma coisa ruim!
- Vamos **resumir** os dados e **descrever** características importantes:
 - Forma da distribuição
 - Posição ou um valor representativo
 - Variação ou dispersão dos dados

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas

Resumo de dados com Tabelas de Frequências

- Grandes conjuntos de dados?
- Conveniente organizá-los e resumi-los.

Tempo Resposta	Nome Fantasia
10	Banco do Brasil
10	Itaú BMG Consignado
9	GVT
9	Tim
9	Banco Cifra
10	Vivo - Telefônica
5	Submarino
7	Banco Santander
10	Smiles
10	Magazineluiza.com
10	Smiles
9	SKY
10	Walmart.com
10	GVT
:	:

Vivo - Telefônica	1000
Oi Fixo	879
Tim	800
Claro Celular	527
SKY	460
Oi Celular	353
Samsung	310
Walmart.com	301
Caixa Econômica Federal	279
Extra.com	258
GVT	234

(0, 3]	1555
(3, 6]	1835
(6, 9]	3095
(9, 12]	1951

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas

Cálculo das classes

- Definir K → Quantidade de classes (*thumb rule*: \sqrt{n} ; Geralmente de 5 a 20 classes)
- Calcular h → Amplitude das classes:
 - $h = (\max - \min) / (K - 1)$
- 1ª classe: $Li_1 = \min - h/2$; $Ls_1 = Li_1 + h$
- 2ª classe: $Li_2 = Ls_1$; $Ls_2 = Li_2 + h$
- 3ª classe: $Li_3 = Ls_2$; $Ls_3 = Li_3 + h$
- E assim por diante...
- Ponto médio da classe: $(Li + Ls) / 2$



Mais Frequências

Classe	Fa	Fac	Fr
(0, 3]	1555	1555	0,1843
(3, 6]	1835	3390	0,2175
(6, 9]	3095	6485	0,3669
(9, 12]	1951	8436	0,2313

Frequência Absoluta
 Frequência Acumulada
 Frequência Relativa



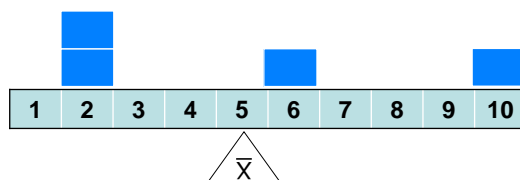
Medidas de tendência central

Uma estatística no
centro do conjunto
 de dados útil para
representa-lo



A média

- A mais importante de todas as estatísticas descritivas;
- O ponto de equilíbrio do conjunto de dados;



- Definição: A **média aritmética** de um conjunto de valores é o valor obtido somando-se todos eles e dividindo-se o total pelo número de valores.



A média

μ = Média Populacional

\bar{X} = Média Amostral

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

- A mais utilizada devido ao fato de:
 - Ser a mais comum;
 - Ser facilmente compreendida;
 - Ser simples de calcular;
 - Ter boas propriedades algébricas.



Propriedades algébricas da média

- A soma dos desvios em relação à média é sempre nula;
- A soma dos quadrados dos desvios em relação a uma constante é mínima se e somente se esta constante é igual à média;
- Somando-se ou subtraindo-se uma constante dos dados, a média fica somada ou subtraída desta constante;
- Multiplicando-se ou dividindo-se os dados por uma constante, a média fica multiplicada ou dividida por esta constante.



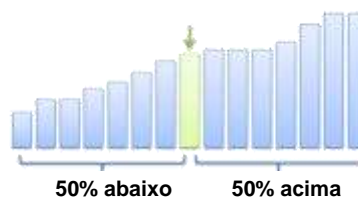
A mediana

- A **mediana** de um conjunto de valores é o valor do meio deste conjunto quando os valores estão ordenados;
- Divide o conjunto de dados em duas metades;
- Ocupa a posição $(n+1)/2$ para dados ordenados;
- Se $(n+1)/2$ não for inteiro, a mediana é a média dos dois dados ao redor da posição $(n+1)/2$

$X = \{1, 2, 3, 4, 5\}$ $Md = 3$

$Y = \{1, 2, 3, 4\}$ $Md = 2,5$

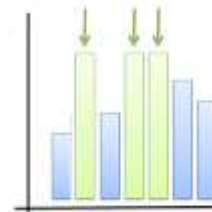
$Z = \{3, 5, 2, 7, 4, 1, 6\}$ $Md = 4$





A moda

- O valor mais frequente;
- Distribuições:
 - Amodais
 - Bimodais
 - Multimodais

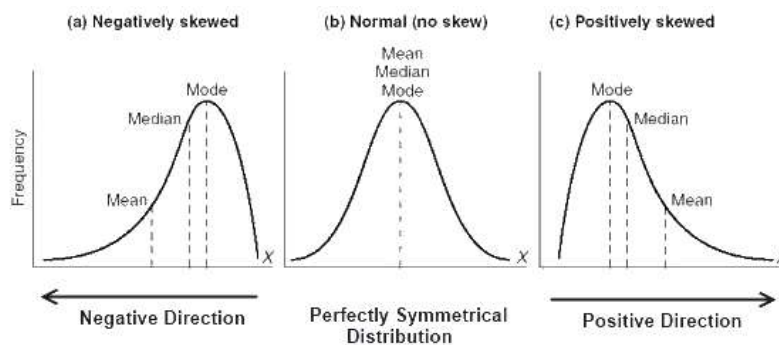


$$X = \{1, 2, 3, 4, 4, 4, 5, 6, 7, 7, 8\}$$

$$Mo = 4$$

$$Y = \{1, 2, 3, 4, 4, 4, 5, 6, 7, 7, 7, 8\}$$

$$Mo = \{4, 7\}$$



Relação entre as medidas de tendência central



E para uma tabela de frequências?

Classe	Fa
[0; 4[2
[4; 8[4
[8; 12[6
[12; 16[6
[16; 20[3
[20; 24[1

Média: Média ponderada
Moda: As classes modais

Categoria	Fa
Azul	81
Amarelo	72
Preto	64
Verde	46
Cinza	93



Calcule as estatísticas abaixo

	A	B	
	56	33	
	56	42	
	57	48	
	58	52	
	61	57	
	63	67	
	63	67	
$X_A = ?$	67	77	$X_B = ?$
$Md_A = ?$	67	82	$Md_B = ?$
$Mo_A = ?$	67	90	$Mo_B = ?$



#podearnaldo #comofaz

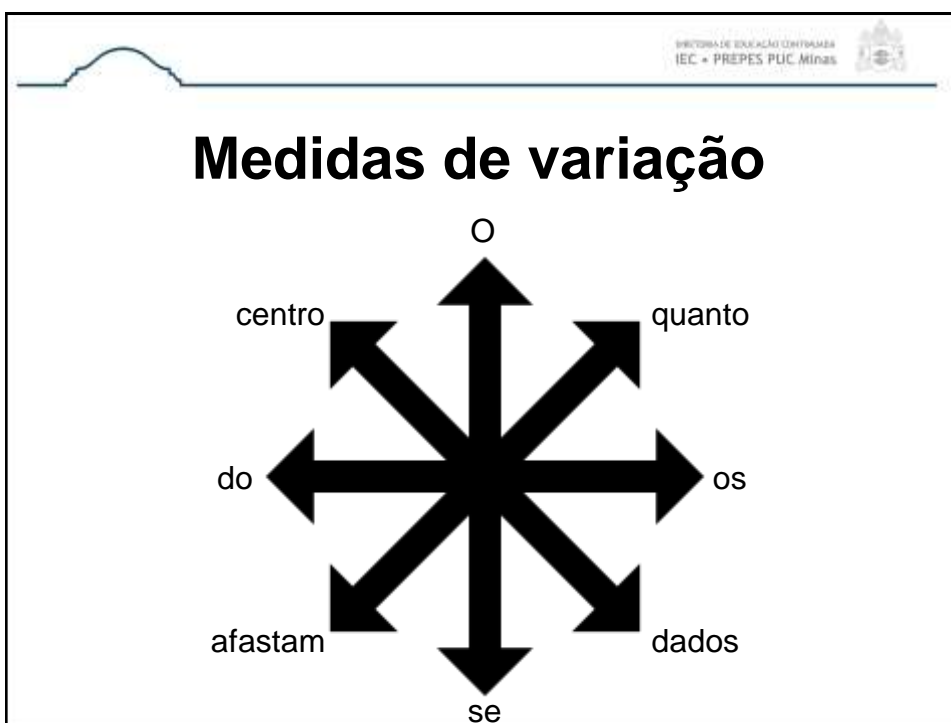
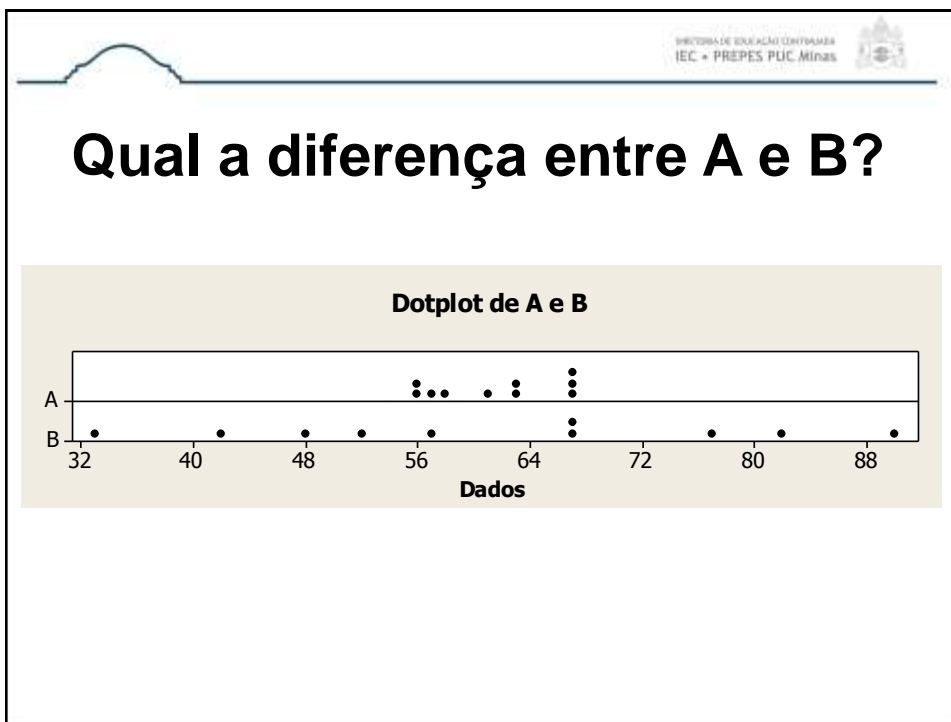
Se os dados
são diferentes,
como as
estatísticas
são iguais?



Pode sim!

As medidas de
posição não
contam toda a
história!







A amplitude

- A distância entre os extremos
- A diferença entre o maior e o menor valores

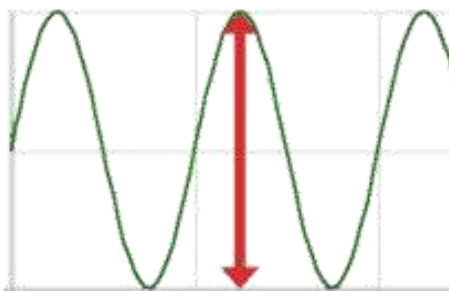
$$R = X_{\max} - X_{\min}$$

$$X = \{1, 3, 5, 2, 5, 8, 2, 8, 5, 7\}$$

$$R_x = 7$$

$$Y = \{10, 21, 32, 14, 24\}$$

$$R_y = 12$$

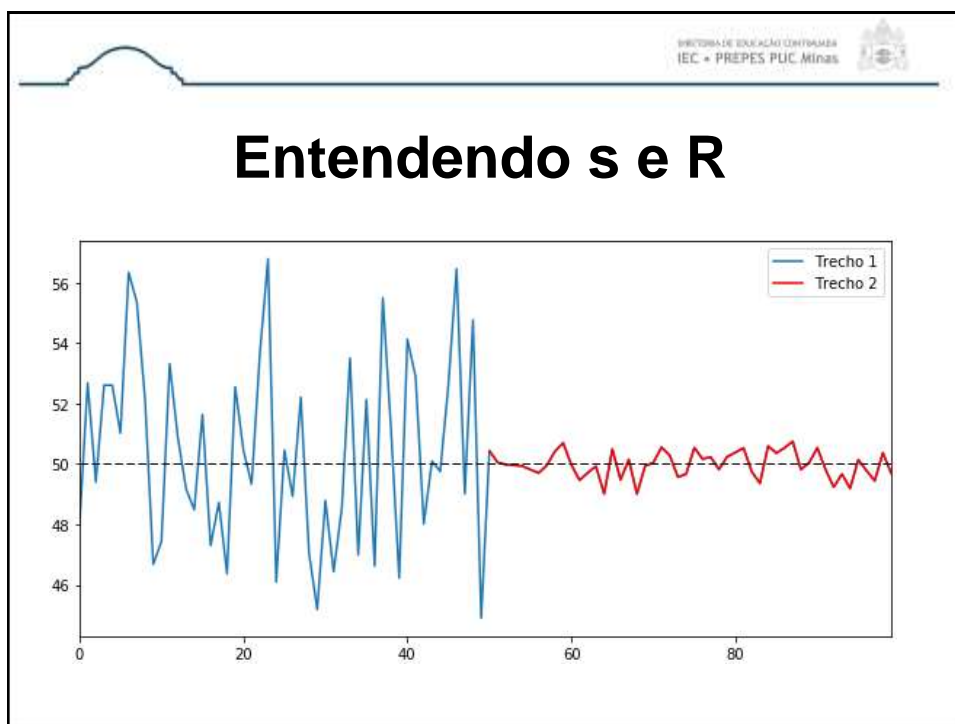


O desvio padrão

- O quanto os dados se afastam de sua média;
- É um valor mínimo de desvio;
- Se somarmos ou subtraímos uma constante aos dados, o desvio não muda
- Se multiplicarmos ou dividirmos os dados por uma constante, o desvio fica multiplicado ou dividido pela constante;

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

Nota de rodapé: A variância é o desvio padrão ao quadrado e seu símbolo é s^2





As separatrizes

- Quartis: Dividem os dados em quatro partes iguais
- Decis: Dividem os dados em dez partes iguais
- Percentis: Dividem os dados em cem partes iguais



As separatrizes

Mediana

Ocupa a posição $(n+1)/2$ para dados ordenados;

Se $(n+1)/2$ não for inteiro, a mediana é a média dos dois dados ao redor da posição $(n+1)/2$

Q1 = P25

Q2 = P50

Q3 = P75

D1 = P10

D2 = P20

D3 = P30

D4 = P40

D5 = P50

D6 = P60

D7 = P70

D8 = P80

D9 = P90

$$P_i = \frac{i \times (n + 1)}{100}$$

Se P_i não for inteiro, será a média dos dois dados ao redor da posição calculada



SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas

Score z

- A média de QI é 100
- Quanto comum é um QI de 104?
- E um QI de 160?
- Isso nos sugere que a distância em relação à média é um indicativo de quanto raro é um valor.
- Mas uma diferença de 4 no QI é quase irrelevante, diferentemente de uma diferença de 4 em uma nota de prova numa escala de 0 a 10!
- Precisamos ter uma medida de distância independente da escala.



Score z

$$z = \frac{x - \bar{x}}{s}$$

- É a resposta que esperamos!
- O *score z* padroniza os dados deixando-os independentes da escala;
- Indica a posição do dado comparando sua distância em relação à média e termos do número de desvios padrão;
- **Alguns modelos de Machine Learning pedem dados na mesma escala. O *score z* pode ser a solução!**

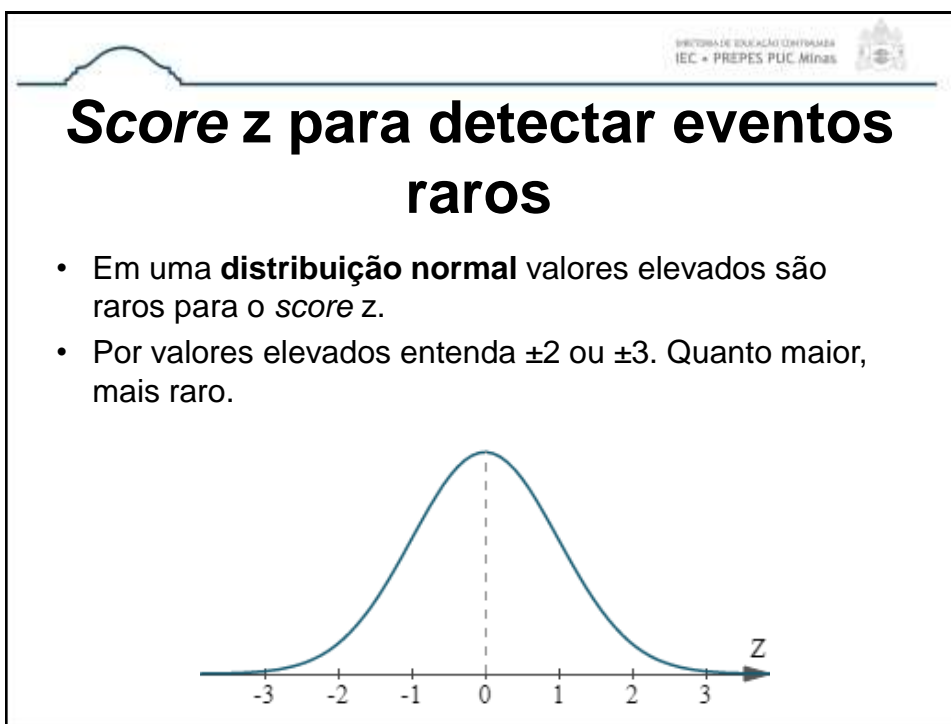


Score z: encontrando dados raros

- José servia à marinha e se encontrou com a esposa em um dia de folga.
- 282 dias depois nasceu seu filho...
- O período médio de gestação é de 268 dias com desvio padrão de 10 dias...



E AGORA JOSÉ?





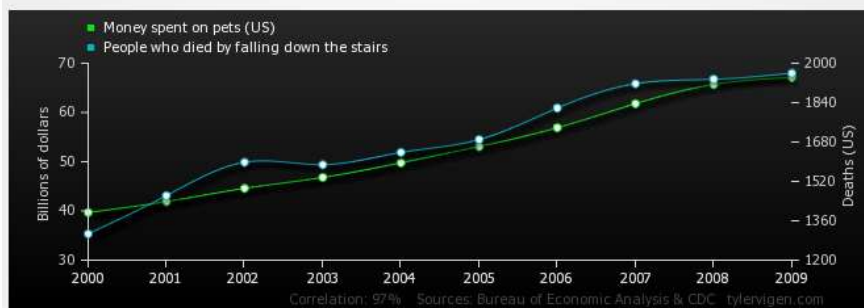
Correlação

- Até aqui vimos como descrever uma única variável;
- Ou um conjunto de variáveis, tratando cada uma individualmente;
- Mas e se quisermos analisar duas variáveis em conjunto?
- E se quisermos saber o quanto uma variável **depende** de outra? O quanto elas variam juntas?



Correlação

Money spent on pets (US)
 correlates with
 People who died by falling down the stairs





Coeficiente de correlação

- No gráfico anterior vimos uma correlação de 97%. Como chegar a esse valor?
- Existem três formas de calcular esse coeficiente

$$r_{XY} = \frac{\sum x_i y_i}{n S_x S_y}$$

$$r_{XY} = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2] \cdot [n \sum Y_i^2 - (\sum Y_i)^2]}}$$



Coeficiente de correlação



`corcoef(x, y) # from numpy import corcoef`
`pearsonr(x, y) # from scipy.stats import pearsonr`



`cor(x, y)`



`=CORREL(x, y)`

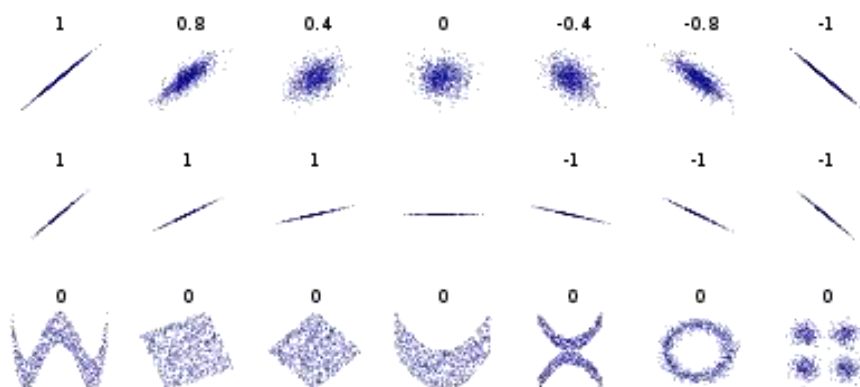


Características de r_{XY}

- Uma grandeza adimensional
- Varia entre -1 e +1
- Quando seu valor absoluto é igual a 1 dizemos que há uma correlação perfeita
- Quando seu valor é igual a zero dizemos que não há correlação linear
- De modo geral quando o valor absoluto é maior que 0,8 dizemos que há uma correlação forte e quando é menor que 0,5 dizemos que há uma correlação fraca.
- O valor de r^2 demonstra o quanto da variância de Y pode ser explicado pela variância de X.



A forma da relação e o valor de r_{XY}





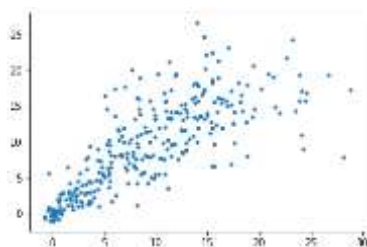
Mas e daí se r é grande?

- Se r_{XY} for significativo podemos criar um modelo de **regressão linear** para prever Y conhecendo o valor de X
- Se as duas variáveis forem fortemente correlacionadas talvez não seja necessário utilizar ambas em um modelo de *Machine Learning*

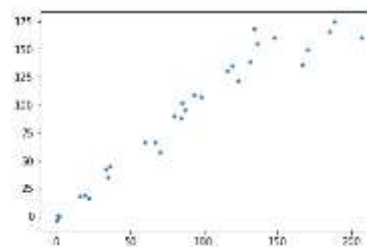


Cuidados com a correlação

- Dados agregados “forjam” correlações



$n = 300$
 $r_{XY} = 0,7818$



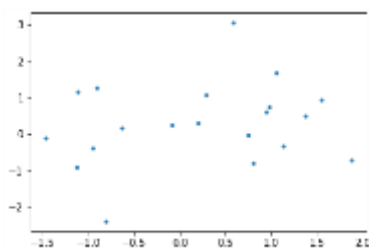
$n = 30$
 $r_{XY} = 0,9643$

Cada ponto é a soma
 de 10 pontos originais

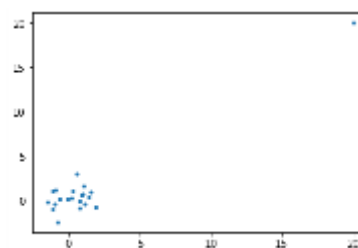


Cuidados com a correlação

- Outliers também “forjam” correlações



$n = 20$
 $r_{XY} = 0,2024$



$n = 21$
 $r_{XY} = 0,9545$

Foi acrescentado o ponto
 (20, 20)

Importante





Correlação x Causalidade

correlação não significa causalidade
correlação não significa causalidade
correlação não significa causalidade
correlação não significa causalidade
correlação não significa causalidade

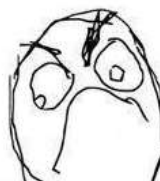
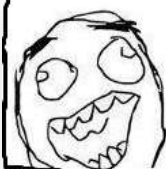


Correlação x Causalidade

Li que existe
correlação forte entre
o número do sapato e
o tamanho da roupa.

Hum... Interessante

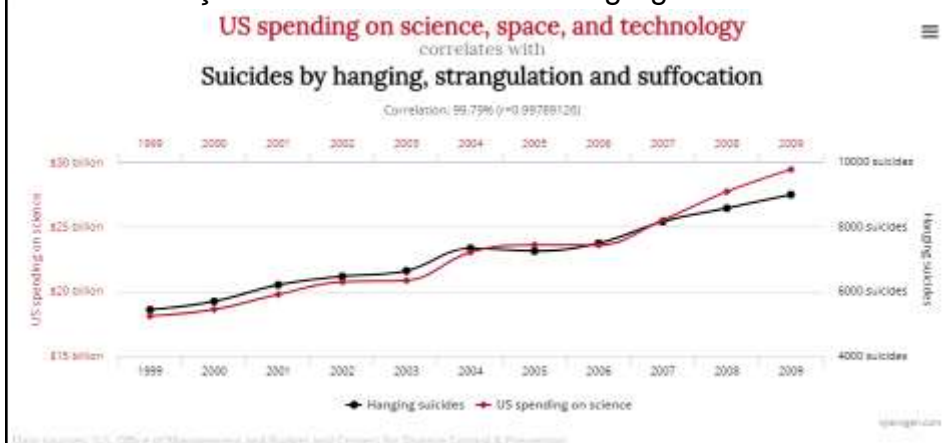
Então se eu comprar
sapatos menores
aquela calça 42 vai
servir!!!!





Cuidados com a correlação

- Correlação x causalidade + dados agregados



Representações Gráficas





Representações Gráficas

- Uma imagem vale mais do que mil dados!
- Cuidados básicos com os gráficos:
 - Escala: Cuidado para não “forjar” escalas a fim de esconder ou evidenciar diferenças
 - 3D: Evite a todo custo gráficos 3D pois eles podem dificultar a análise



Duvida que é importante?

Quer que eu prove?



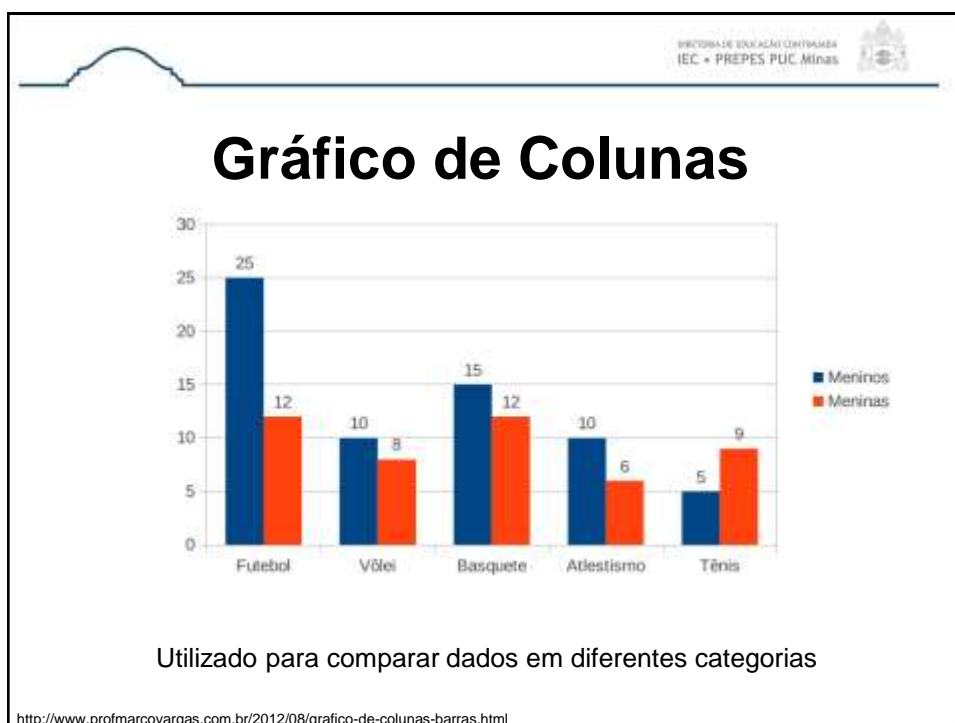
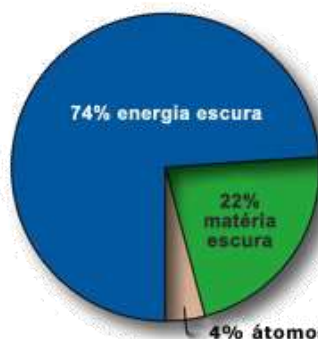




Gráfico de setores (pizza)

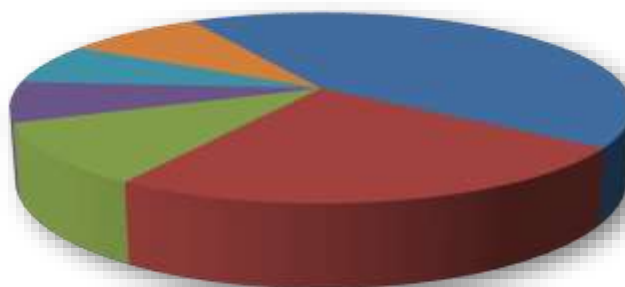


U = Universo

Utilizado para identificar a proporção das partes no todo. Geralmente traz os dados em porcentagem.



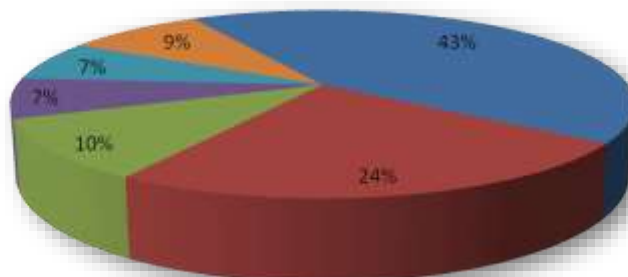
Gráfico de setores (pizza)



Cuidado com o uso dos gráficos 3D. Eles criam ilusões de óptica que podem atrapalhar a análise.



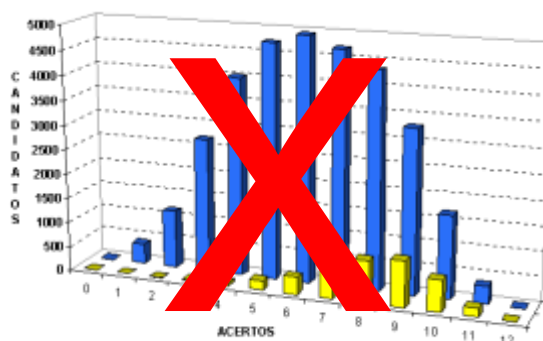
Gráfico de setores (pizza)



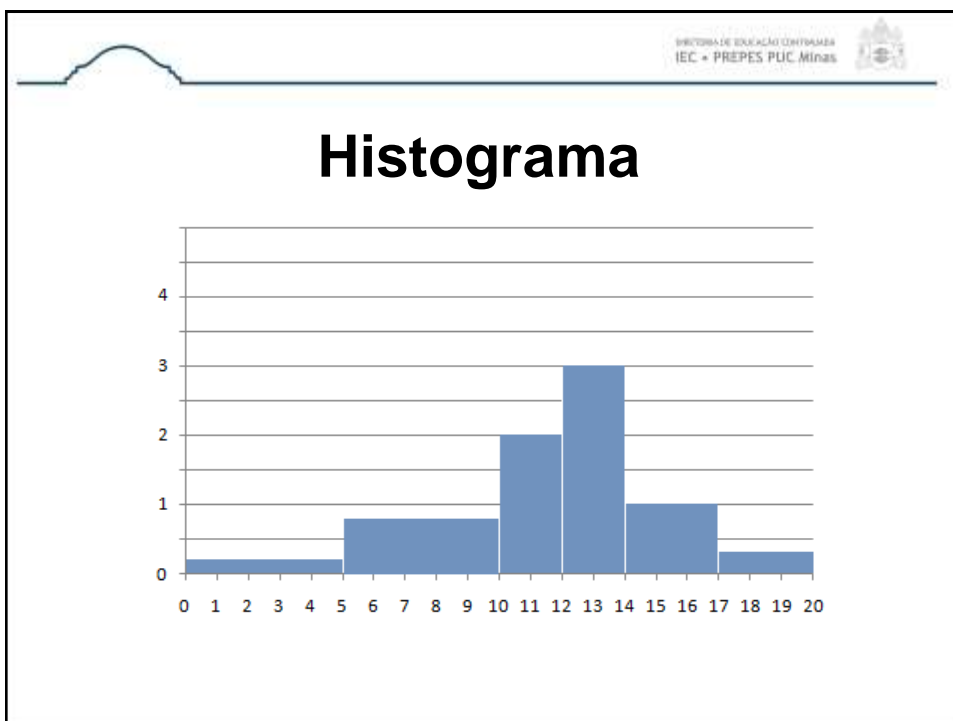
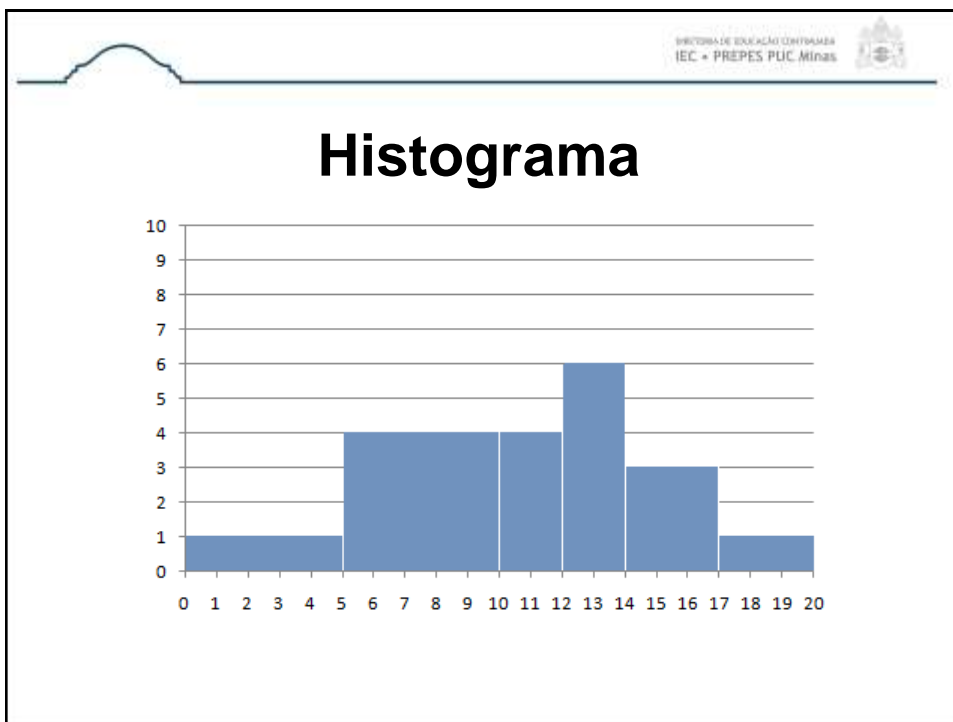
Cuidado com o uso dos gráficos 3D. Eles criam ilusões de óptica que podem atrapalhar a análise.

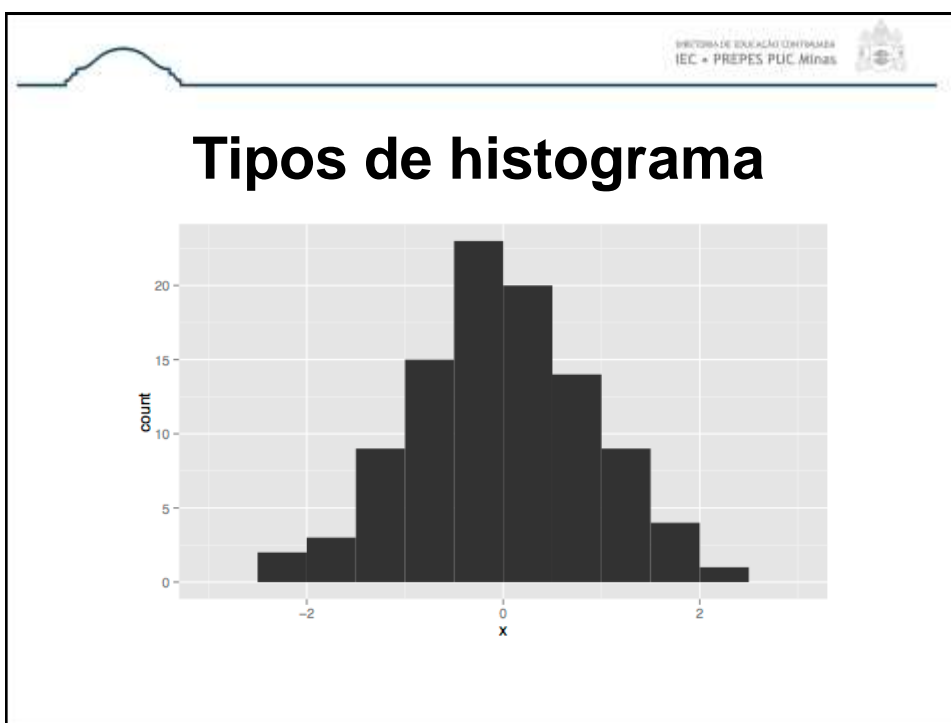
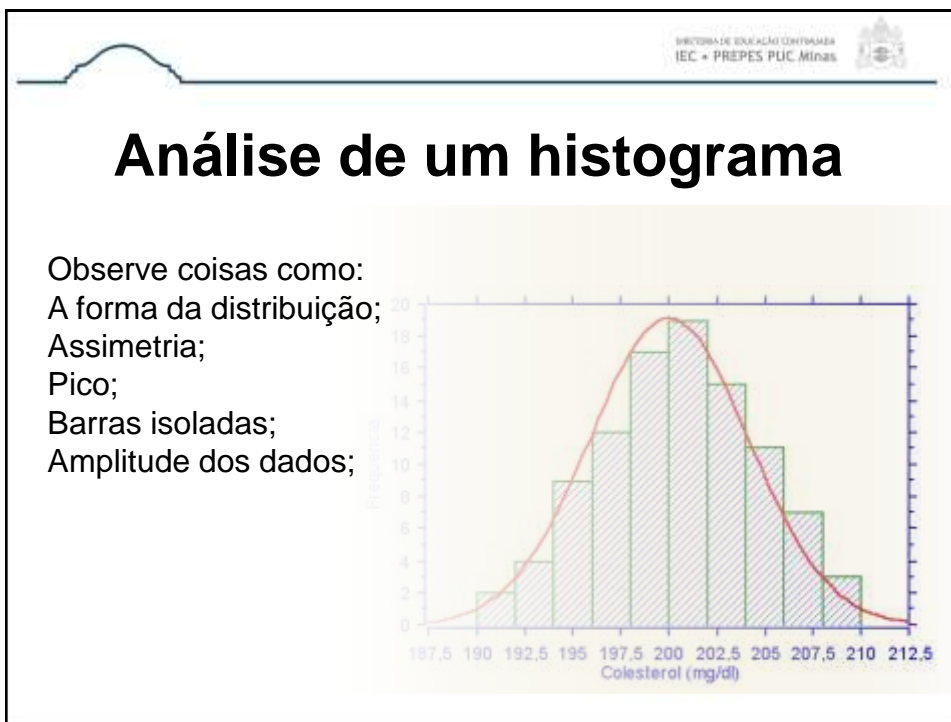


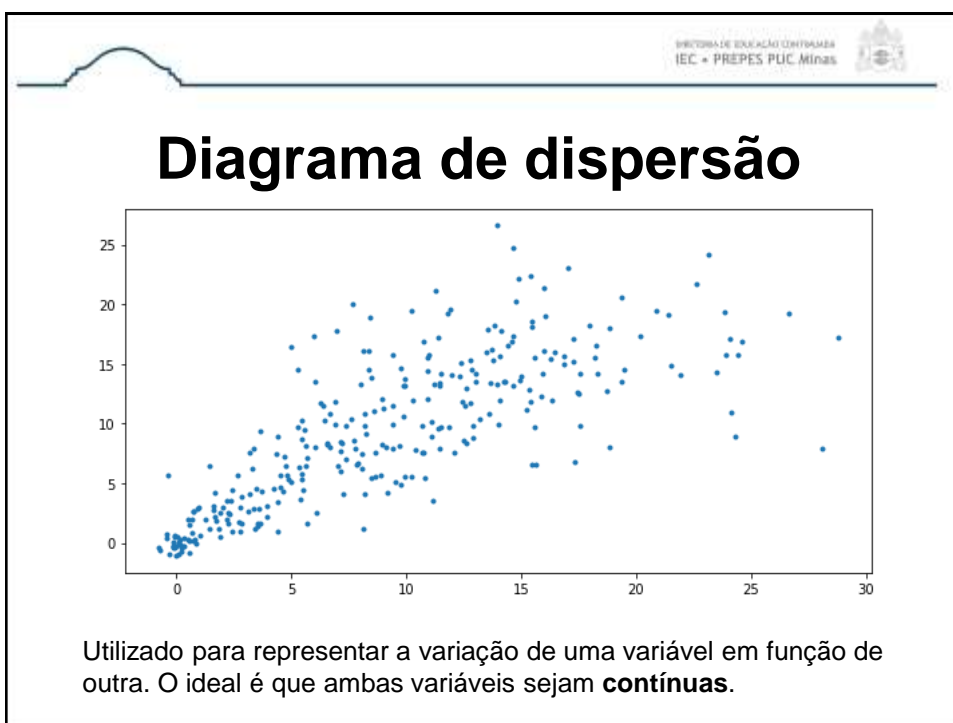
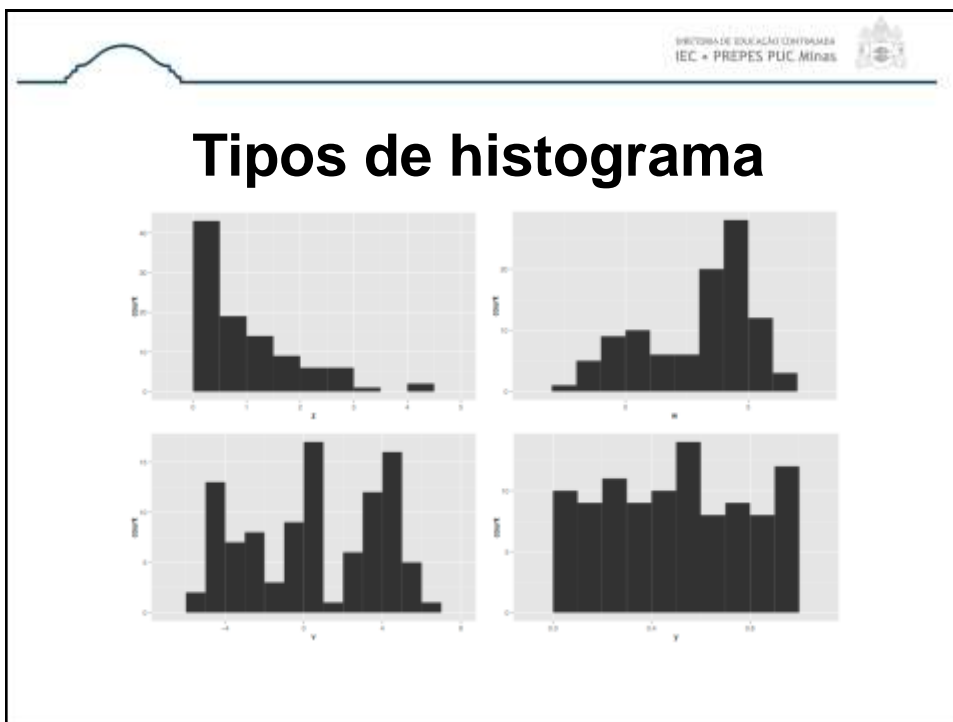
Histograma

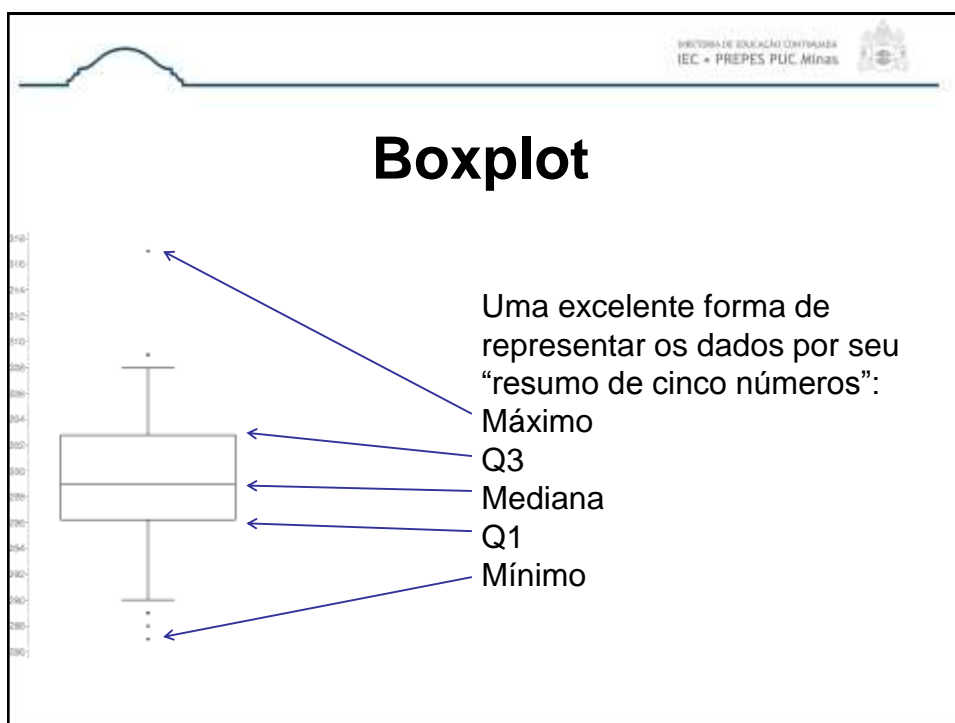


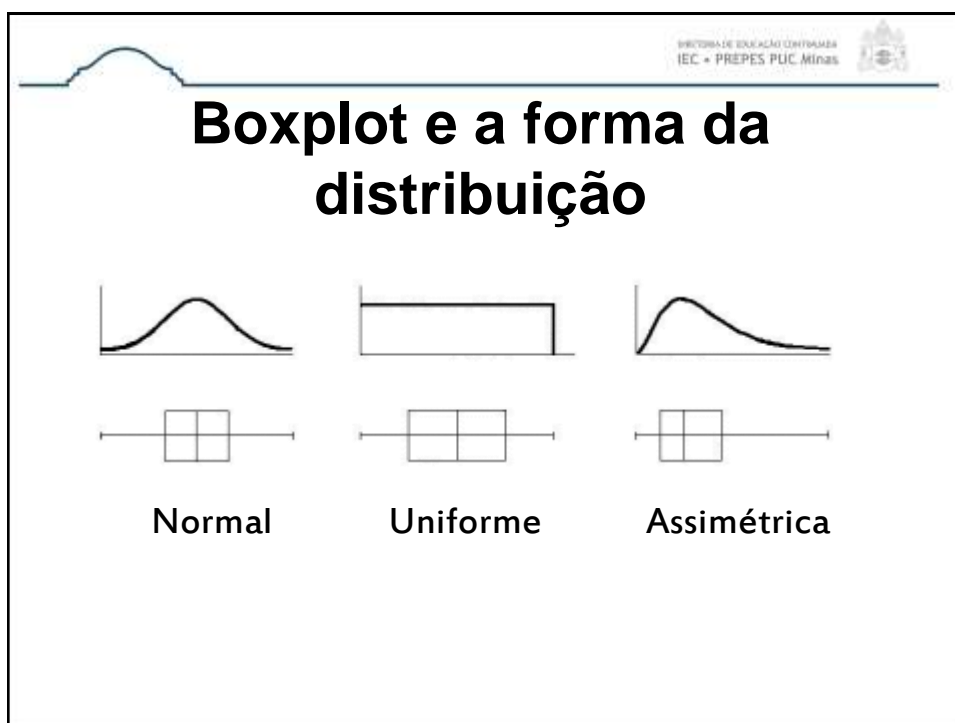
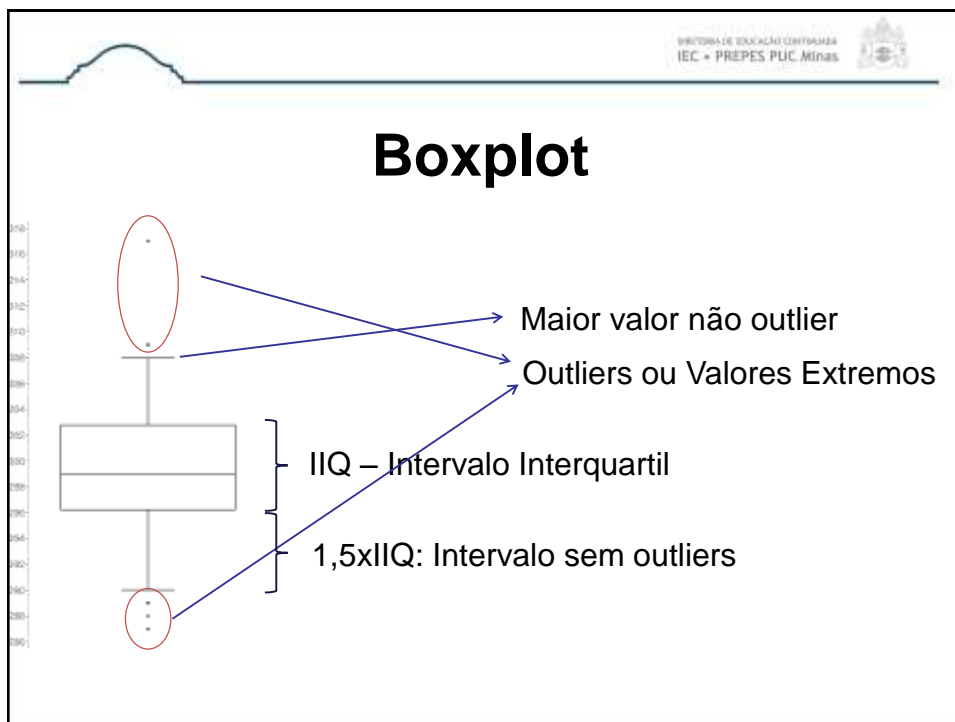
O histograma é um gráfico composto por retângulos **justapostos** em que a base de cada um deles corresponde ao intervalo de classe e a sua área à respectiva frequência.






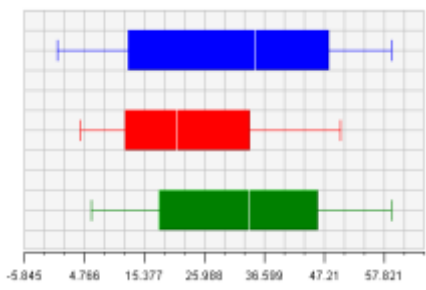







SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Boxplot



- Obtemos o máximo desempenho de um boxplot ao compará-lo com outros.
- A que conclusões chegamos analisando os boxplots acima?

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Resumo Cap II

- Estatística Descritiva: Se quiser se concentrar em um só item, concentre nisso!
- Agrupe os dados em tabelas de frequência
- Resuma os dados com suas medidas de posição, centralização e dispersão
- O score z pode e vai te ajudar bastante. Use para “normalizar” os dados e para verificar se trata-se de um valor extremo
- Correlação é o quanto duas variáveis variam juntas.
- **Correlação não significa causalidade** (repita 500 vezes)



Resumo Cap II

- **Sempre** faça o(s) gráfico(s) de seus dados. Os gráficos mostram relações ocultas nos números “frios”
- Para dados contínuos use histogramas
- Para dados contínuos no tempo use gráficos de linhas
- Para dados discretos use gráficos de colunas
- Cuidado com as “pizzas”. Muitos estatísticos não veem esse gráfico com bons olhos
- Diagramas de dispersão são uma boa para mostrar a relação entre duas variáveis
- Boxplot são big bad fu**ing modafocas. Use sem moderação (mas cuidado para quem você vai mostrar!)

3. Probabilidade





Aspectos Gerais

- Vimos que uma importante área da Estatística é a Inferência
- A Inferência Estatística nos permite, através de uma amostra, inferir os parâmetros da população
- Para tanto nos baseamos em hipóteses e, através da amostra, estimamos a probabilidade de a hipótese estar correta



Aspectos Gerais

- Uma empresa afirma que contrata de forma não tendenciosa tanto homens quanto mulheres
- Analisando os registros é possível observar que as últimas 100 contratações foram apenas de homens
- **“Se sob determinada hipótese (tal como a contratação não tendenciosa) a probabilidade de uma determinada amostra (como 100 homens contratados) é excepcionalmente pequena, concluimos que a hipótese provavelmente não é correta” (TRIOLA, 1999)**



Definições

- **Experimento:** Qualquer processo que permite ao pesquisador fazer observações
- **Evento:** Uma coleção de resultados de um experimento
- **Evento simples:** É um evento que não comporta qualquer decomposição
- **Espaço amostral:** O conjunto de todos os eventos simples possíveis
- **Probabilidade da ocorrência do evento A:** $P(A)$

$$0 \leq P(A) \leq 1$$



Exemplo

- **Experimento:** Arremesso de um dado
- **Evento:** Obtenção do resultado 4
- **Evento simples:** Por não permitir decomposição, a obtenção do resultado 4 é um evento simples
- **Espaço amostral:** 1, 2, 3, 4, 5, 6



Outro Exemplo

- **Experimento:** Arremesso de **dois** dados
- **Evento:** Obtenção do resultado 9
- **Evento simples:** A obtenção do resultado 9 não é mais um evento simples pois pode ser decomposta em 5-4 e 3-6
- **Espaço amostral:** 1-1, 1-2, 1-3, ..., 6-4, 6-5, 6-6




Mais exemplos

- Na tabela abaixo f significa sexo feminino e m sexo masculino

Experimento	Exemplo de evento	Espaço amostral
Um nascimento	1 do sexo feminino (Evento Simples)	{f, m}
Três nascimentos	2 do sexo feminino e 1 sexo masculino (ffm, fmf, mff são eventos simples)	{fff, ffm, fmf, fmm, mff, mfm, mmf, mmm}

- Podemos pensar incorretamente que o evento ffm não é um evento simples pois pode ser decomposto em f, f e m. Porém f ou m não fazem parte do espaço amostral do experimento *Três nascimentos*.

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Definições de probabilidade

- Triola afirma que existem duas definições comuns para a probabilidade de ocorrência do evento A

Realize ou observe um experimento um grande número de vezes e conte a frequência com que o evento A acontece efetivamente


$$P(A) = \frac{\text{Número de ocorrências de A}}{\text{Número de repetições do experimento}}$$

Abordagem de frequência

Suponha que um experimento tenha n eventos simples diferentes, cada um com a *mesma chance* de ocorrer. Se o evento A pode ocorrer em s dentre as n maneiras, então:


$$P(A) = \frac{\text{Número de maneiras como A pode ocorrer}}{\text{Número de eventos simples diferentes}} = \frac{s}{n}$$

Abordagem clássica

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Definições de probabilidade


- Um erro comum é assumirmos que os eventos têm a mesma chance se não conseguimos saber ou identificar a chance real.
- A probabilidade de bater o carro em uma viagem é de 50%?
- A probabilidade de ser aprovado na disciplina é 33,33%?

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Exemplo

- Qual a probabilidade de você ser assassinado este ano?
 - Em 2010 aconteceram 152 homicídios e a população permanente da cidade naquele ano era de 600.747 pessoas
 - Podemos então usar a abordagem pela frequência para calcular que a probabilidade de uma pessoa ser assassinada em Uberlândia é de $152 / 600747 = 0,025 \%$
 - Não podemos usar a abordagem clássica pois os eventos não são equiprováveis (assassinado/não assassinado)

Fontes:
<http://www.datapedia.info/public/cidade/6093/mg/uberlandia#homicidios>
<http://www.datapedia.info/public/cidade/6093/mg/uberlandia#pop>

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


Exemplo

- Qual a probabilidade de termos um presidente da república nascido no Acre?
 - O espaço amostral possui dois eventos simples: o presidente é acreano / o presidente não é acreano
 - Pela abordagem de frequência a probabilidade é zero pois esse evento nunca aconteceu
 - Não podemos usar a abordagem clássica pois os eventos do espaço amostral não são equiprováveis
 - Nos resta, então, fazer uma estimativa grosseira. A população do Acre é 0,4% da população brasileira; considerando o quanto o Acre é um estado remoto, uma estimativa de 0,01% é aceitável.



Eventos complementares

- Definição: O complemento do evento A , denotado por \bar{A} , consiste em todos os resultados nos quais o evento A não ocorre
- $P(A) + P(\bar{A}) = 1$
- $P(A) = 1 - P(\bar{A})$
- $P(\bar{A}) = 1 - P(A)$



A regra da adição

- Regra da adição: Uma forma de obter $P(A \text{ ou } B)$
- $P(A \text{ ou } B)$ é a probabilidade de que aconteça o evento A **ou** o evento B **ou** ambos eventos em **um mesmo experimento**
- Imagine uma série de testes com o “detector de mentiras”

		Real	
		Mentiu	Não mentiu
Teste	Mentiu	42 (Verdadeiro Positivo)	15 (Falso Positivo)
	Não mentiu	9 (Falso Negativo)	32 (Verdadeiro Negativo)



A regra da adição

- Calcule $P(\text{Falso Positivo})$
- Calcule $P(\text{Teste Positivo ou Sujeito Mentiu})$

		Real	
		Mentiu	Não mentiu
Teste	Mentiu	42 (Verdadeiro Positivo)	15 (Falso Positivo)
	Não mentiu	9 (Falso Negativo)	32 (Verdadeiro Negativo)

- Várias formas de calcular *Teste Positivo ou Sujeito Mentiu*: $15 + 42 + 9$ ou $57 + 51 - 42$ ou $57 + 9$
- A palavra **ou** indica **adição** mas tome cuidado para **não somar a mesma coisa duas vezes!**



A regra da adição

- Formalmente: $P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ e } B)$
- Não use a fórmula cegamente. **Entenda** o conceito de $P(A \text{ ou } B)$
- Para encontrar $P(A \text{ ou } B)$ some a quantidade de formas que o evento A pode ocorrer com a quantidade formas que o evento B pode ocorrer **tomando cuidado para não somar a mesma coisa duas vezes**. $P(A \text{ ou } B)$ é a divisão dessa soma com o número total de eventos no espaço amostral.



A regra da multiplicação

- Regra da multiplicação: Uma forma de obter $P(A \text{ e } B)$
- $P(A \text{ e } B)$ é a probabilidade de que o evento **A aconteça em uma prova** do experimento **e** o evento **B aconteça na prova seguinte**.
- Imagine as seguintes questões em uma prova:

Verdadeiro ou Falso:

O fumo é uma das principais causas do câncer.



Assinale a alternativa correta:

O coeficiente de correlação de Pearson tem esse nome em homenagem a:

- a) Karl Marx
- b) Carl Friedrich Gauss
- c) Karl Pearson
- d) Carly Simons
- e) Mario Triola



A regra da multiplicação

- Qual a probabilidade de uma pessoa acertar as duas questões “chutando”?
- Espaço amostral:
 - {V-a; V-b; V-c, V-d, V-e; F-a; F-b; F-c; F-d; F-e}
- Como um “chute” envolve a escolha aleatória das respostas, todos os resultados são equiprováveis.
- V-c é a resposta certa. Então a probabilidade de acertar é:

$$P(\text{Ambas corretas}) = P(V \text{ e } c) = \frac{1}{10} = 0,1$$



A regra da multiplicação

- A probabilidade de V na primeira questão é $\frac{1}{2}$
 - Espaço amostral da primeira questão: {V, F}
- A probabilidade de c na segunda questão é $\frac{1}{5}$
 - Espaço amostral da segunda questão: {a, b, c, d, e}
- Os resultados $P(V \text{ e } c) = 1/10$, $P(V) = 1/2$ e $P(c) = 1/5$ sugerem que $P(V \text{ e } c) = P(V) \cdot P(c)$
- Vejamos outro exemplo antes de afirmar tal generalização.



A regra da multiplicação

- Na extração de duas cartas de um baralho bem embaralhado determine a probabilidade de que a primeira carta seja um Ás e a segunda carta seja um Rei
- Admita que a primeira carta não seja repostada antes da extração da segunda carta





A regra da multiplicação

- Como existem quatro ases entre as 52 cartas temos:

$$P(\text{ás}) = \frac{4}{52}$$

- Assumindo que conseguimos um ás na primeira extração, temos:

$$P(\text{rei}) = \frac{4}{51}$$

- A probabilidade de obter um ás na primeira extração e um rei na segunda é, portanto:

$$P(\text{ás e rei}) = \frac{4}{52} \cdot \frac{4}{51} = 0,00603$$

- Esse exemplo ilustra um princípio muito importante: **P(B) deve levar em conta o fato do evento A já ter ocorrido.**



A regra da multiplicação

- Notação:** $P(B|A)$ representa a probabilidade de ocorrência de B quando se sabe que o evento A já ocorreu. Pode se ler $B|A$ como “ B dado A ”
- Definição:** Dois eventos A e B são **independentes** se a ocorrência de um deles não afeta a probabilidade de ocorrência do outro. Se a ocorrência de um afeta a probabilidade e outros eventos, estes são ditos **dependentes**.
- No exemplo anterior temos:

$$P(\text{ás}) = \frac{4}{52} \text{ e } P(\text{rei}|\text{ás}) = \frac{4}{51}$$



A regra da multiplicação

- Regra Formal da Multiplicação:**

$P(A \text{ e } B) = P(A) \cdot P(B)$ se A e B são independentes

$P(A \text{ e } B) = P(A) \cdot P(B|A)$ se A e B são dependentes

- Regra Intuitiva da Multiplicação:** Para determinar a probabilidade de o evento A ocorrer seguido do evento B devemos multiplicar a probabilidade de A pela probabilidade de B não esquecendo de considerar que a probabilidade de B deve levar em conta a ocorrência de A .



Teorema de Bayes

- Thomas Bayes foi um matemático e pároco inglês e seu trabalho é tido como um novo paradigma da estatística
- Este trabalho, entretando, foi publicado após sua morte
- Assim como Fla x Flu, Barça x Real Madrid, Petralhas x Coxinhas, Biscoito x Bolacha temos atualmente Frequentistas x Bayesianos.





Teorema de Bayes

- O teorema de Bayes descreve a probabilidade de um evento, baseado em um conhecimento a priori que pode estar relacionado ao evento. O teorema mostra como alterar as probabilidade a priori tendo em vista novas evidências para obter probabilidades a posteriori.

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$



Teorema de Bayes

- Imagine duas latas de biscoito. Lata1 tem 30 biscoitos de Baunilha e 10 de Chocolate. Lata2 tem 20 de cada sabor
- Você fecha os olhos e pega um biscoito de Baunilha. Qual a probabilidade de que ele tenha vindo da Lata1?
- Queremos saber $P(L1|B)$
- $P(B|L1) = \frac{3}{4}$ mas $P(B|L1)$ é diferente de $P(L1|B)$
- $$P(L1|B) = \frac{P(B|L1) \cdot P(L1)}{P(B)} = \frac{\frac{3}{4} \cdot \frac{1}{2}}{\frac{5}{8}} = \frac{3}{5} = 60\%$$



Teorema de Bayes

- Imagine um paciente contando seus sintomas para um médico.
- Probabilidade de estar com a doença X é $P(X=1) = 0,6$
- Médico solicita exame
- $P(R = 1 | X = 1) = 0,95$
- $P(R = 1 | X = 0) = 0,10$
- Resultado do exame: $R = 1$
- $$P(X = 1 | R = 1) = \frac{P(R=1|X=1) \cdot P(X=1)}{P(R=1|X=1) \cdot P(X=1) + P(R=1|X=0) \cdot P(X=0)}$$
- $$P(X = 1 | R = 1) = \frac{0,95 \cdot 0,6}{0,95 \cdot 0,6 + 0,1 \cdot 0,4} = 0,9344$$





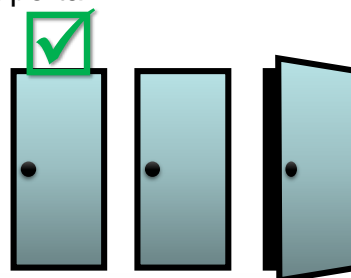
O problema de Monty Hall

- A priori, a probabilidade do carro estar em qualquer porta é de $1/3$
- Mas quando Monty Hall abre uma porta, as probabilidades mudam?
- Se mudam, mudam para quanto?



O problema de Monty Hall

- Vamos assumir que o participante escolhe a porta 1
- $A1$ = O carro está na primeira porta
- $A2$ = O carro está na segunda porta
- $A3$ = O carro está na terceira porta
- C = Monty Hall abre a terceira porta
- $P(C|A1) = 0,5$
- $P(C|A2) = 1,0$
- $P(C|A3) = 0$
- $P(C) = 0,5$



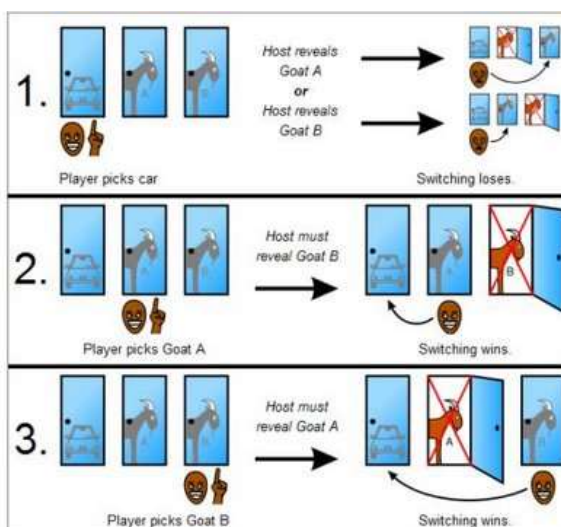


O problema de Monty Hall

- $P(C|A1) = 0,5$
- $P(C|A2) = 1,0$
- $P(C|A3) = 0$
- $P(C) = 0,5$
- $P(A1|C) = \frac{P(C|A1) \cdot P(A1)}{P(C)} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{2}$
- $P(A2|C) = \frac{P(C|A2) \cdot P(A2)}{P(C)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$
- $P(A3|C) = \frac{P(C|A3) \cdot P(A3)}{P(C)} = \frac{0 \cdot \frac{1}{3}}{\frac{1}{2}} = 0$



O problema de Monty Hall





Resumo Cap III

- Probabilidades variam de zero a um
- A adição está relacionada ao **OU**
- A multiplicação está relacionada ao **E**
- Cuidado com os eventos dependentes ao usar a regra da multiplicação
- O teorema de Bayes é uma grande (r)evolução na estatística. Mais do que a formulação do teorema, o modo de pensar!

4. Inferência





Estatística inferencial

- Duas principais aplicações:
- Estimar um **parâmetro populacional**
- Formular conclusões sobre a **população**
- Para inferir dados da população não preciso avaliá-la inteiramente
- Preciso apenas coletar uma **boa amostra**

“Dados coletados de forma imprecisa ou descuidada podem ser totalmente destituídos de valor, mesmo que a amostra seja suficientemente grande”

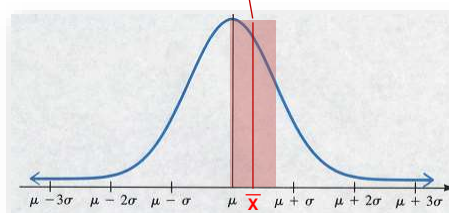
Mario Triola



Inferência

- Qual o melhor estimador para a média μ ?
- É possível demonstrar que é a média \bar{x}
- Mas quão boa é essa estimativa?
- Não tenho a menor ideia
- Pra isso preciso de uma estimativa intervalar

Estimativa pontual





Estimativa pontual x intervalar

- Em média o pão de um Big Mac tem 380 sementes de gergelim



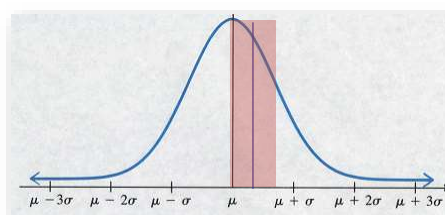
- Um pão de Big Mac tem entre 350 e 410 sementes de gergelim

Fonte: <http://www.ocregister.com/2008/10/03/fresh-buns-how-does-mcdonalds-get-them/>



Intervalo de Confiança

- Estimativa intervalar → Intervalo de confiança
- Não é a probabilidade de o valor estar no intervalo!!!**
- Se fizermos um grande número de intervalos nestas condições, aproximadamente $(1-\alpha)\%$ deles conterão o verdadeiro valor da média (que permanece desconhecido).



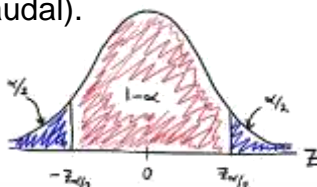


Whaaaaat?

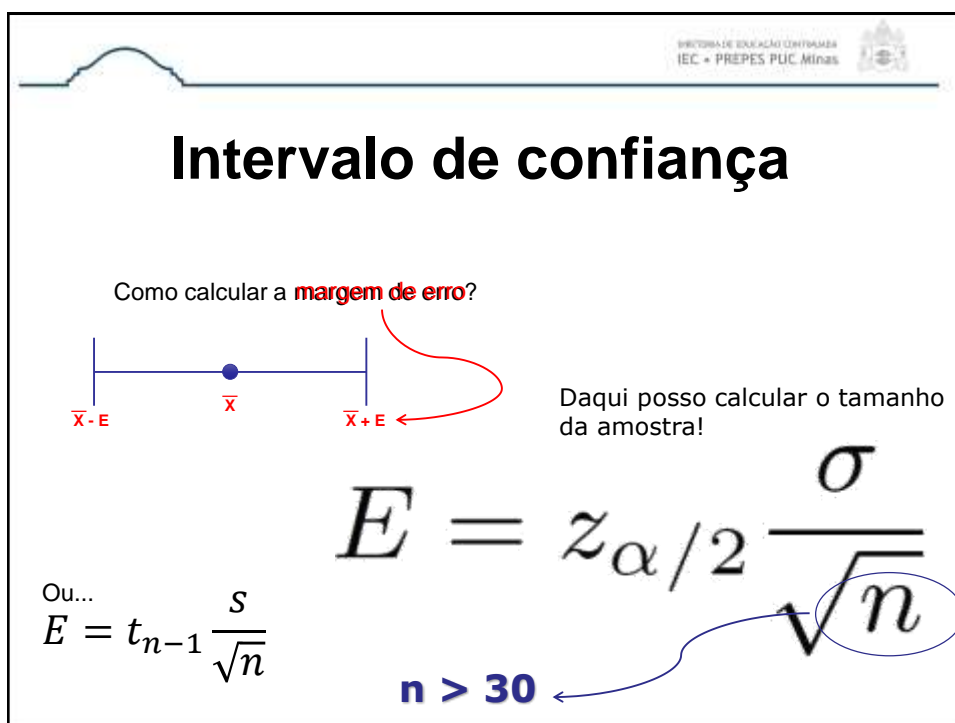
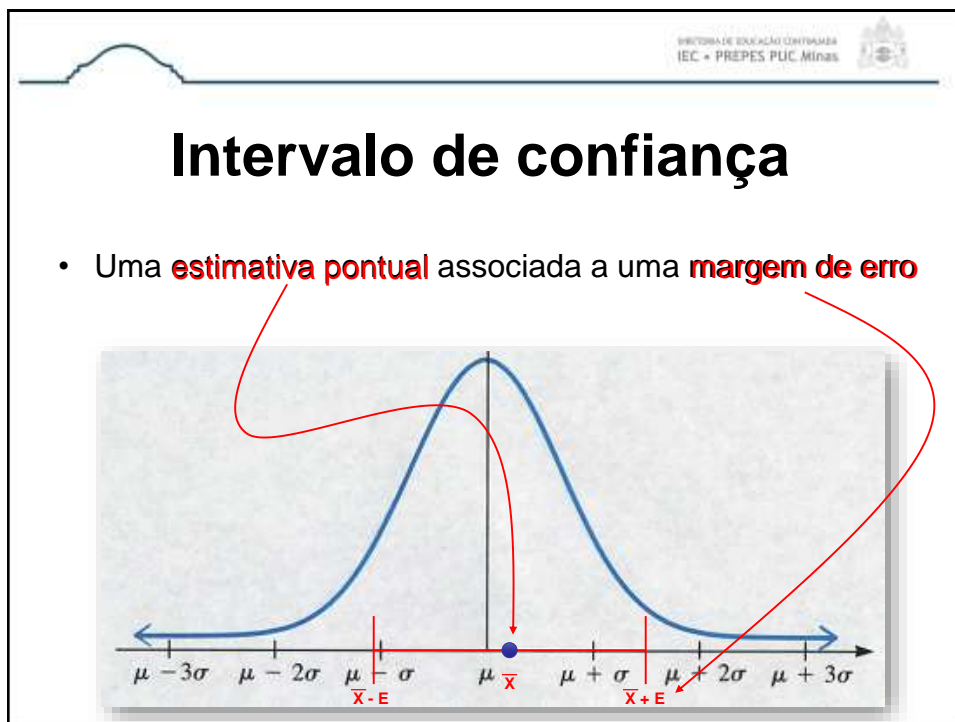


Nível de confiança

- Uma medida do nosso grau de certeza de que o intervalo encontrado contém o parâmetro populacional; o número de intervalos que contém o parâmetro populacional se coletássemos 100 amostras
- Utiliza-se o α para descrever uma probabilidade correspondente a uma área; geralmente dividida igualmente nas duas caudas da distribuição normal (bicaudal).



α	Nível de Confiança	$Z_{\alpha/2}$
10%	90%	1,645
5%	95%	1,96
1%	99%	2,575





Agora que sabemos calcular um intervalo para a média...

- Que tal compararmos **duas médias**?

Aluno	Nota pré	Nota pós
1	8,5	9,0
2	8,5	9,0
3	7,0	8,5
4	9,0	9,5
5	9,5	10,0
6	8,0	9,5
7	8,0	8,5
8	8,5	8,5
Média	8,37	9,06

Teste de hipóteses

A nota média antes do treinamento é 8,37 e após o treinamento é 9,06. A nota aumentou?



MAS HEIN???

Claro que aumentou!

Na Estatística o “diferente” pode ser igual. Precisamos testar essa hipótese.



Fundamentos do Teste de hipóteses

- Uma boa analogia para um teste de hipóteses é a de um julgamento;
- Toda pessoa é inocente até que se prove o contrário;
- Existe a possibilidade de **condenarmos um inocente**;
- Existe a possibilidade de **inocentarmos um culpado**;

Dizer que o treinamento é "culpado" pela diferença nas médias quando na verdade não é.

Dizer que o treinamento não é "culpado" pela diferença nas médias quando na verdade ele é.



Tipos de erros

		Verdade	
		H0: "Inocente"	Ha: "Culpado"
Estatística indica	H0: "Inocente"	Decisão correta	Erro tipo II β
	Ha: "Culpado"	Erro tipo I α	Decisão correta



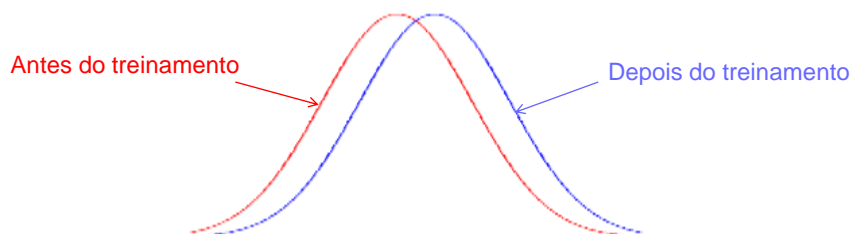
Vocabulário

- **H_0 :** Hipótese Nula – Suposição de que não há diferença. É assumida como verdade até que tenhamos evidências suficientes para rejeitá-la.
- **H_a :** Hipótese Alternativa – Suposição de que há diferença. Se assumirmos que esta hipótese é correta rejeitamos H_0 .
- **Risco (α):** O máximo risco que o pesquisador considera aceitável para rejeitar H_0 . É um valor sempre maior que zero, geralmente 5% ou 10%
- **Diferença Significativa:** Termo usado para descrever uma situação onde a diferença é muito grande para atribuí-la ao acaso.
- **p-value:** É o risco calculado que corremos de errar ao afirmar que há diferença estatística entre os dois grupos de dados quando na verdade não há.



Voltando ao problema das notas...

- **Questão prática:** O treinamento ministrado provocou o aumento das notas?
- **Questão estatística:** A diferença entre as notas antes do treinamento (8,37) e após o treinamento (9,06) é grande o suficiente para afirmarmos que se trata de populações distintas ou essas diferenças se devem ao acaso, a variações naturais do dia a dia?





Traduzindo para o *Estatiquês*

$H_0: \mu_{\text{antes}} = \mu_{\text{depois}}$

$H_a: \mu_{\text{antes}} < \mu_{\text{depois}}$

ou

$H_a: \mu_{\text{antes}} > \mu_{\text{depois}}$

$H_a: \mu_{\text{antes}} \neq \mu_{\text{depois}}$

As notas antes e depois do treinamento são iguais. O treinamento não serviu para nada

As notas depois do treinamento são maiores do que antes.

$> \text{ ou } < \rightarrow$ Unicaudal

$\neq \rightarrow$ Bicaudal



No mundo teórico...

α
 $gl = n_1 + n_2 - 2$


α	0.10	0.05	0.025	0.01	0.005	0.001
1	6.314	6.314	6.314	6.314	6.314	6.314
2	2.878	2.878	2.878	2.878	2.878	2.878
3	2.353	2.353	2.353	2.353	2.353	2.353
4	2.132	2.132	2.132	2.132	2.132	2.132
5	1.960	1.960	1.960	1.960	1.960	1.960
6	1.895	1.895	1.895	1.895	1.895	1.895
7	1.895	1.895	1.895	1.895	1.895	1.895
8	1.858	1.858	1.858	1.858	1.858	1.858
9	1.833	1.833	1.833	1.833	1.833	1.833
10	1.812	1.812	1.812	1.812	1.812	1.812
11	1.796	1.796	1.796	1.796	1.796	1.796
12	1.781	1.781	1.781	1.781	1.781	1.781
13	1.769	1.769	1.769	1.769	1.769	1.769
14	1.759	1.759	1.759	1.759	1.759	1.759
15	1.751	1.751	1.751	1.751	1.751	1.751
16	1.744	1.744	1.744	1.744	1.744	1.744
17	1.738	1.738	1.738	1.738	1.738	1.738
18	1.733	1.733	1.733	1.733	1.733	1.733
19	1.729	1.729	1.729	1.729	1.729	1.729
20	1.726	1.726	1.726	1.726	1.726	1.726
21	1.723	1.723	1.723	1.723	1.723	1.723
22	1.721	1.721	1.721	1.721	1.721	1.721
23	1.719	1.719	1.719	1.719	1.719	1.719
24	1.717	1.717	1.717	1.717	1.717	1.717
25	1.716	1.716	1.716	1.716	1.716	1.716
26	1.715	1.715	1.715	1.715	1.715	1.715
27	1.714	1.714	1.714	1.714	1.714	1.714
28	1.713	1.713	1.713	1.713	1.713	1.713
29	1.712	1.712	1.712	1.712	1.712	1.712
30	1.711	1.711	1.711	1.711	1.711	1.711
31	1.710	1.710	1.710	1.710	1.710	1.710
32	1.709	1.709	1.709	1.709	1.709	1.709
33	1.708	1.708	1.708	1.708	1.708	1.708
34	1.707	1.707	1.707	1.707	1.707	1.707
35	1.706	1.706	1.706	1.706	1.706	1.706
36	1.705	1.705	1.705	1.705	1.705	1.705
37	1.704	1.704	1.704	1.704	1.704	1.704
38	1.703	1.703	1.703	1.703	1.703	1.703
39	1.702	1.702	1.702	1.702	1.702	1.702
40	1.701	1.701	1.701	1.701	1.701	1.701
41	1.700	1.700	1.700	1.700	1.700	1.700
42	1.699	1.699	1.699	1.699	1.699	1.699
43	1.698	1.698	1.698	1.698	1.698	1.698
44	1.697	1.697	1.697	1.697	1.697	1.697
45	1.696	1.696	1.696	1.696	1.696	1.696
46	1.695	1.695	1.695	1.695	1.695	1.695
47	1.694	1.694	1.694	1.694	1.694	1.694
48	1.693	1.693	1.693	1.693	1.693	1.693
49	1.692	1.692	1.692	1.692	1.692	1.692
50	1.691	1.691	1.691	1.691	1.691	1.691

t_{crit}

Compara com...

$$t_{\text{calc}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$




SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


No nosso mundo

```
In [1]: from scipy.stats import ttest_ind, ttest_rel
Out[1]: help(ttest_ind)
Help on function ttest_ind in module scipy.stats.stats:

ttest_ind(a, b, axis=0, equal_var=True, nan_policy='propagate')
    Calculates the T-test for the means of *two independent* samples of
    scores.

    This is a two-sided test for the null hypothesis that 2 independent
    samples
    have identical average (expected) values. This test assumes that
    the
    populations have identical variances by default.
```

SECRETARIA DE EDUCAÇÃO CONTINUADA
IEC • PRÉPES PUC Minas


No nosso mundo

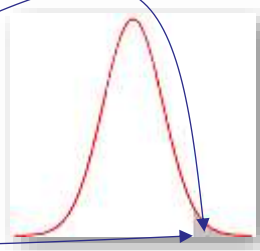
```
In [2]: help(ttest_rel)
Out[2]: Help on function ttest_rel in module scipy.stats.stats:



ttest_rel(a, b, axis=0, nan_policy='propagate')
    Calculates the T-test on TWO RELATED samples of scores, a and b.

    This is a two-sided test for the null hypothesis that 2 related or
    repeated samples have identical average (expected) values.
```

```
In [3]: ttest_ind([9, 8, 7, 7, 8, 4, 8, 9, 7, 7, 8, 4, 8],
                  [4, 5, 7, 8, 5, 4, 2, 4, 5, 6, 8, 9, 5, 4, 7])
Out[3]: Ttest_indResult(statistic=2.5217545079927919,
                        pvalue=0.0181403312541897)
```



p-value: esse é o cara!



Voltando às notas...

Aluno	Nota pré	Nota pós
1	8,5	9,0
2	8,5	9,0
3	7,0	8,5
4	9,0	9,5
5	9,5	10,0
6	8,0	9,5
7	8,0	8,5
8	8,5	8,5
Média	8,37	9,06


Voltando às notas...

```

from scipy.stats import ttest_ind, ttest_rel

nota_pré = [8.5, 8.5, 7, 9, 9.5, 8, 8, 8.5]
nota_pós = [9, 9, 8.5, 9.5, 10, 9.5, 8.5, 8.5]
print(np.mean(nota_pré))
8.375
print(np.mean(nota_pós))
9.0625
ttest_rel(nota_pré, nota_pós)
Ttest_relResult(statistic=-3.6666666666666665,
                 pvalue=0.0079994338096036916)

```





E para dados categóricos?

What flavor of ice cream would you pick?			
	Chocolate	Vanilla	Neither
Children	40	22	15
Teens	12	16	45
Adults	55	54	10
Total	107	92	70

Será que a preferência pelo sabor de sorvete **depende** da faixa etária?



Teste Chi-Quadrado

- Um teste de **independência**
- H_0 : As duas variáveis são **independentes**
- H_a : As duas variáveis são dependentes
- Compara a frequência observada com a “frequência esperada”

What flavor of ice cream would you pick?			
	Chocolate	Vanilla	Neither
Children	40	22	15
Teens	12	16	45
Adults	55	54	10
Total	107	92	70



Teste Chi-Quadrado

```
>>> data = [[40, 22, 15],
             [12, 16, 45],
             [55, 54, 10]]
```

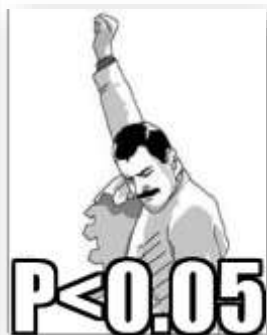
What flavor of ice cream would you pick?			
	Chocolate	Vanilla	Neither
Children	40	22	15
Teens	12	16	45
Adults	55	54	10
Total	107	92	70

```
>>> scipy.stats.chi2_contingency(data)
(73.444270651069871, _____ → Estatística Chi-Quadrado
 4.2498689001693132e-15, _____ → p-valor
 4, _____ → Graus de liberdade (lin-1)x(col-1)
 array([[ 30.62825279,  26.33457249,  20.03717472],
        [ 29.03717472,  24.96654275,  18.99628253],
        [ 47.33457249,  40.69888476,  30.96654275]]))
```

Frequências esperadas



E daí?



Rejeito H_0

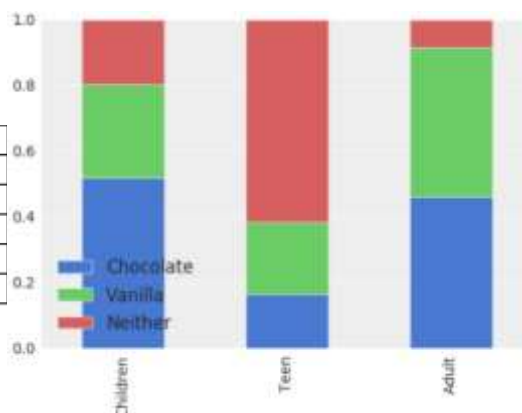


O sabor escolhido **depende**
da faixa etária



Evidenciando a dependência

What flavor of ice cream would you pick?			
	Chocolate	Vanilla	Neither
Children	40	22	15
Teens	12	16	45
Adults	55	54	10
Total	107	92	70



Ufa! Resumão Capítulo IV

- Inferência = Estimar parâmetros populacionais
- Estimativa intervalar x Estimativa pontual
- Estimativa intervalar pressupõe intervalo e nível de confiança
- Intervalo de confiança \approx Margem de erro
- Nível de confiança é o percentual de intervalos que conterá o valor correto do parâmetro populacional e não a probabilidade de o intervalo conter o valor do parâmetro
- Em um teste de hipóteses há a possibilidade de acusarmos um inocente (Erro tipo I) e de inocentarmos um culpado (Erro tipo II)
- A hipótese nula (H_0) é sempre a de igualdade
- O p-value é o risco que corro ao afirmar que há diferença. Se p-value é menor que α rejeito H_0



Concluindo

- Estude estatística. É muito importante para o profissional e para a pessoa.
- Aprofunde-se especialmente na estatística descritiva especialmente na utilização dos gráficos e tabelas de frequência.
- Quando estiver mais confortável desenvolva habilidades nos intervalos de confiança.
- Finalmente fique craque nos testes de hipóteses e definição de tamanhos de amostra!
- Preocupe-se com os conceitos, nem tanto com as fórmulas.
- Utilize Python, R ou outros softwares estatísticos. O computador está aí para isso...
- Me procure se precisar!



Referências

- Introdução à Estatística – Mário F. Triola
- Introdução ao controle estatístico da qualidade – Douglas C. Montgomery
- Teorema de Bayes e inferência Bayesiana
 - <http://greenteapress.com/wp/think-bayes/>
 - http://www.ufjf.br/joaquim_neto/files/2009/09/IB-Slides-v1.1.pdf
- Estatística Básica - Bussab e Morettin

