

Investigação Aplicada I

Aula 5

1º Semestre 2016/17

Licenciatura em Ciências Biomédicas Laboratoriais

igrodrigues@ualg.pt; ESSUAlg: gabinete 2.06

Prof. Inês Rodrigues

Protocolo de investigação (estrutura):

- Máximo de 15 páginas escritas a 1,5 espaços, separação de parágrafos 6 pt após em letra Arial 11 pt
 - **Capa :**
 - No topo instituição proponente: Escola Superior de Saúde da Universidade do Algarve
 - A 1/3 da página, 12 pt: Protocolo de investigação e cerca de 3 parágrafos abaixo o título do projecto 14 pt
 - A 4/5 da página, a 9 pt: Área de aplicação do projecto e âmbito (Unidade Curricular de Investigação aplicada I do Curso de Análises Clínicas e Saúde Pública)
 - Em baixo, a 11 pt: Proponentes

Protocolo de investigação (estrutura):

- Índice
- Sem paginação
 - Título
 - Área
 - Resumo (máximo de 300 palavras)

Protocolo de investigação (estrutura):

- Com paginação
 - **Protocolo de Investigação**
 - **Introdução** (Enquadramento do tema com referências bibliográficas; 1-2 páginas)
(Relevância: Devem constar os factos conhecidos – “estado da arte” e a importância dos resultados esperados)
 - **Objectivos** (gerais e específicos)
 - **Material e métodos**
 - **População em estudo**
 - **Tipo, técnica de amostragem e dimensão da amostra**
(Devem ser definidos os critérios de inclusão e de exclusão)
 - **Metodologia e definição das variáveis em estudo**
 - **Recolha e tratamento de dados** (bases de dados utilizadas, software utilizado e metodologia de análise dos dados)

Protocolo de investigação (estrutura):

- Com paginação
 - **Organização do estudo**
 - **Cronograma**
 - Tabela onde na 1ª coluna constam as várias etapas do projecto e na 1ª linha os meses de realização
 - As células da tabela serão preenchidas com sombreado, caso as tarefas estejam atribuídas individualmente, deverá ser indicado o nome do responsável pela execução da tarefa.
 - **Local do estudo**
 - Indicação do local onde se vai realizar

Protocolo de investigação (estrutura):

- Com paginação
 - **Recursos humanos e técnicos**
 - **Recursos humanos:** Indicação de todos os intervenientes no projecto e atribuição das responsabilidades de execução e de coordenação
 - » Poderá ser sob a forma de tabela: onde conste o nome do Investigador, a categoria e funções, e as tarefas que desempenha
 - **Recursos técnicos:** indicação dos serviços /laboratórios envolvidos, respectivos equipamentos utilizados e metodologias/análises a realizar
 - » Poderá ser sob a forma de tabela: onde conste a indicação dos serviços / laboratórios envolvidos, equipamentos a utilizar, e as análises a realizar

Protocolo de investigação (estrutura):

- Com paginação
 - **Questões éticas**
 - Comprometimento da equipa de investigadores em salvaguardar as questões éticas, informando de que modo as salvaguardam
 - **Importância dos resultados esperados**
 - Parágrafo onde é, de forma resumida, demonstrado o interesse prático do trabalho
 - **Verba prevista para a execução do projecto**
 - Tabela onde por etapa do projecto e por tarefa é indicado o custo de realização de forma discriminada em relação aos equipamentos, materiais, reagentes, serviços e material de escritório ou outros a adquirir.
 - **Referências bibliográficas**
 - Em formato APA (American Psychological Association) Style – <http://www.apastyle.org>

Representação dos resultados

Estatística Descritiva

Apresentação dos dados / Resultados

Tabelas: dados organizados segundo valores e categorias

Frequência das variáveis

Gráficos: de linhas, área e de barras, circular, histograma, caixas de bigodes, gráfico de dispersão e fluxogramas.

Os dados devem ser apresentados da forma mais clara e de fácil compreensão

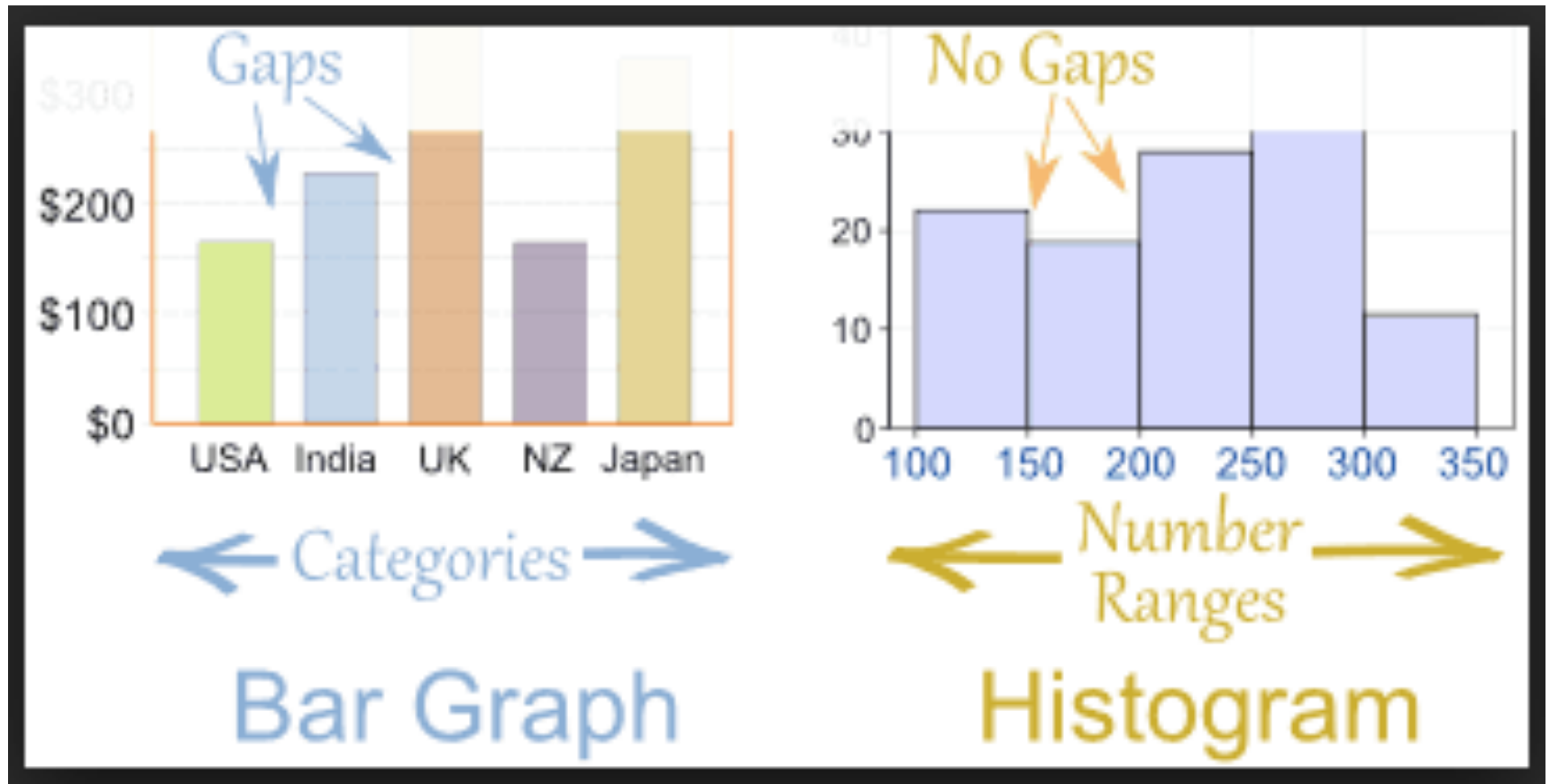
Estadística Descritiva

- Resumo da informação contida na amostra
- Sugestão interpretação dos resultados
- Comunicação dos resultados
- **Não permite verificar hipóteses**

Representação dos resultados

Estatística Descritiva

Histograma vs Gráfico de barras

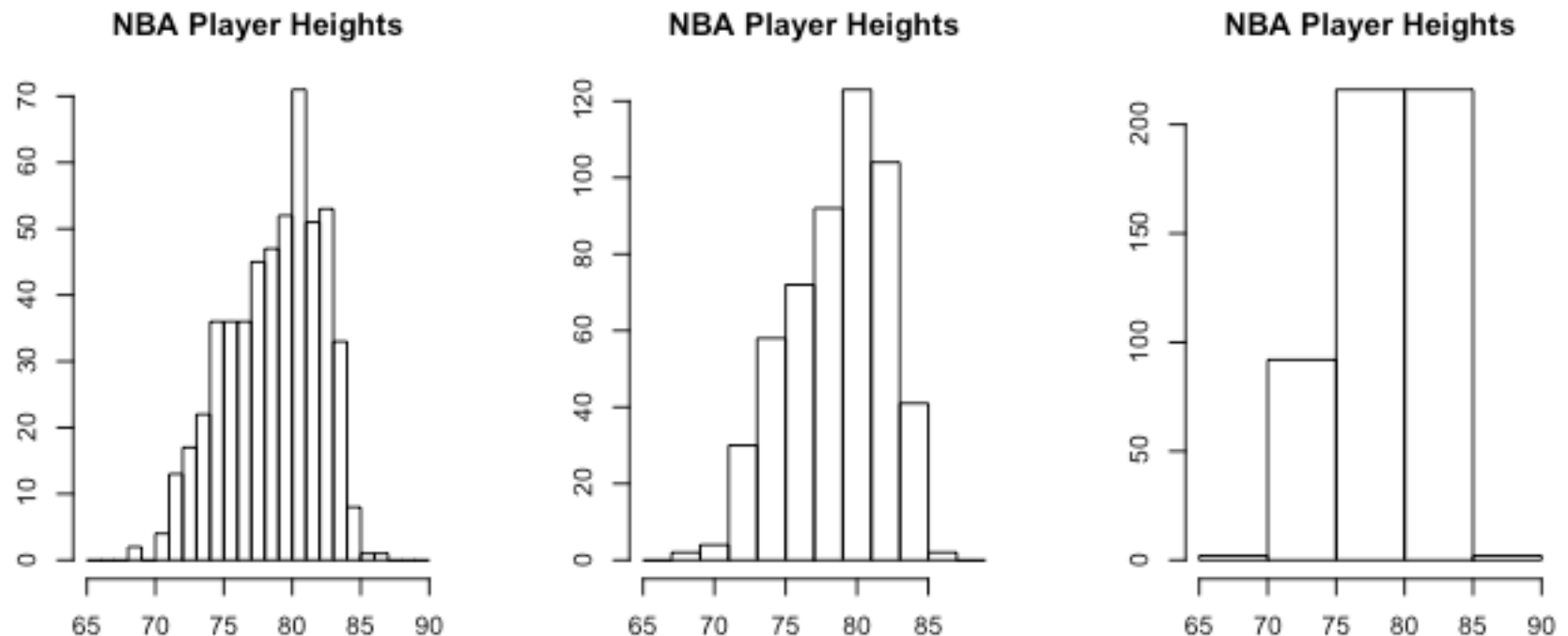


Representação dos resultados

Estatística Descritiva

Importância do número de barras de um histograma

Quanto mais largas são as barras, menos detalhada é a informação representada pelo histograma



Medidas de Tendência central

Média:

Centro da distribuição de uma frequência

Mediana:

O valor mediano numa coleção de números ordenados

Exemplo: 61 70 70 74 85 97 104 112 125 702

Mediana: $(85 + 97)/2=91$

Moda:

- valor mais frequente numa distribuição

Medidas de variabilidade

Amplitude (A= max. – min.)

Variância

$$s^2 = \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Medidas de variabilidade

Desvio padrão

Raiz da variância

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

Coeficiente de variação

(Desvio Padrão / Média) x 100

Medidas de variabilidade medem a dispersão dos dados em torno à média

Quanto maior o CV maior a variação

Exemplo:

SD = 5

X = 100 → Pouca % do desvio padrão faz parte da média

Se:

SD = 50 → Metade da % da média representa desvio dos dados

Coeficiente de variação

$$CV=100(SD/Média)$$

Intervalo (range): diferença entre o valor máximo e o valor mínimo de uma série de dados

Skewness (simetria):

Quando a distribuição de uma variável não é normal, não sendo a media uma medida central dos dados.

Pearson's Skewness coefficient (estimativa sobre a simetria dos dados):

$$\text{Skewness} = (\text{média-mediana})/SD$$

Para uma distribuição simétrica dos dados a média é igual a mediana e o coeficiente de assimetria é 0

Fisher's Kurtosis

Indica se a distribuição dos dados tem a forma de sino perfeito, obedecendo a uma distribuição normal.

Mede se a curva normal é demasiado pequena ou achatada

Skewness

The coefficient of Skewness is a measure for the degree of symmetry in the variable distribution.



Negatively skewed distribution
or Skewed to the left
Skewness < 0



Normal distribution
Symmetrical
Skewness $= 0$



Positively skewed distribution
or Skewed to the right
Skewness > 0

Kurtosis

The coefficient of Kurtosis is a measure for the degree of peakedness/flatness in the variable distribution.



Platykurtic distribution
Low degree of peakedness
Kurtosis < 0

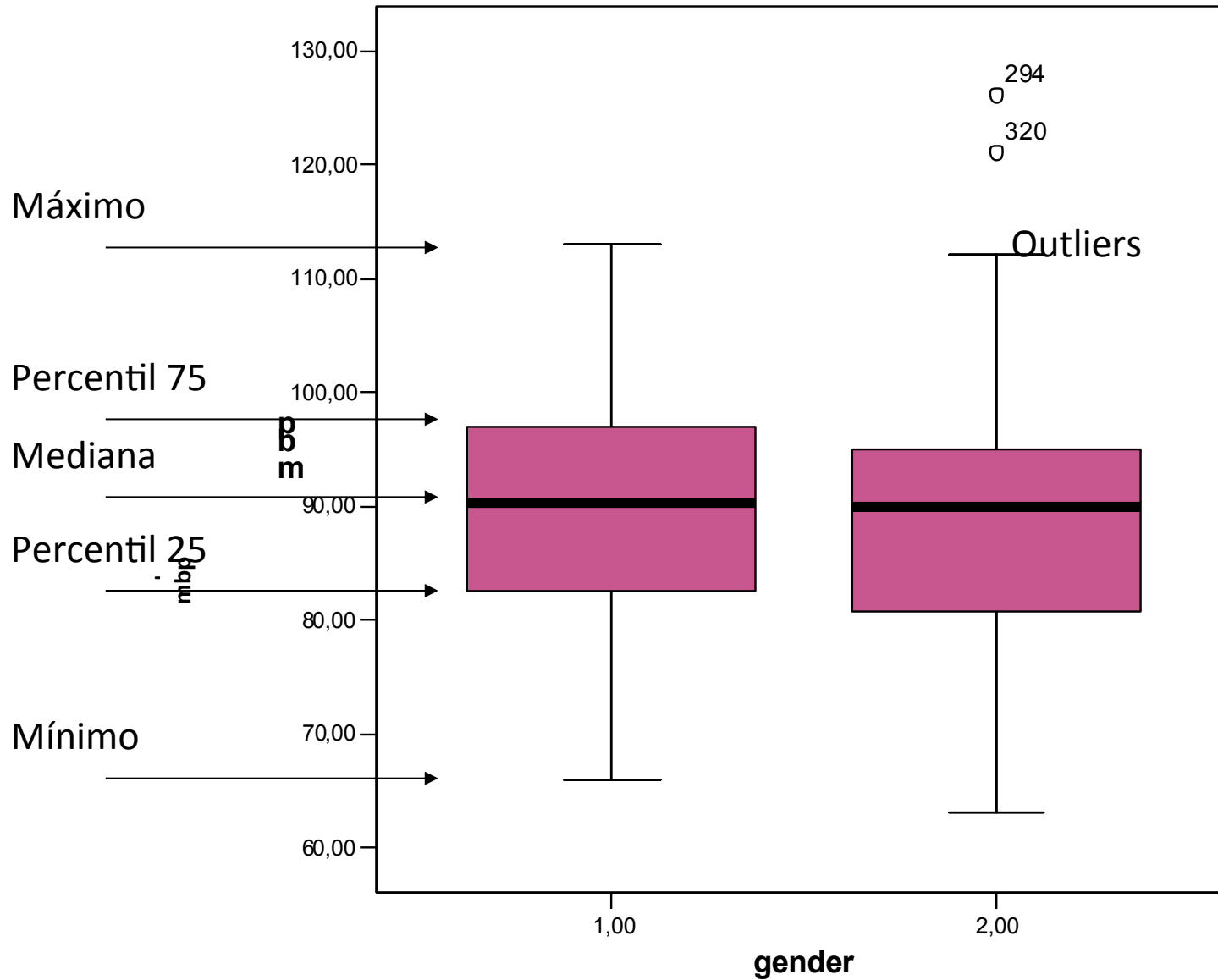


Normal distribution
Mesokurtic distribution
Kurtosis $= 0$

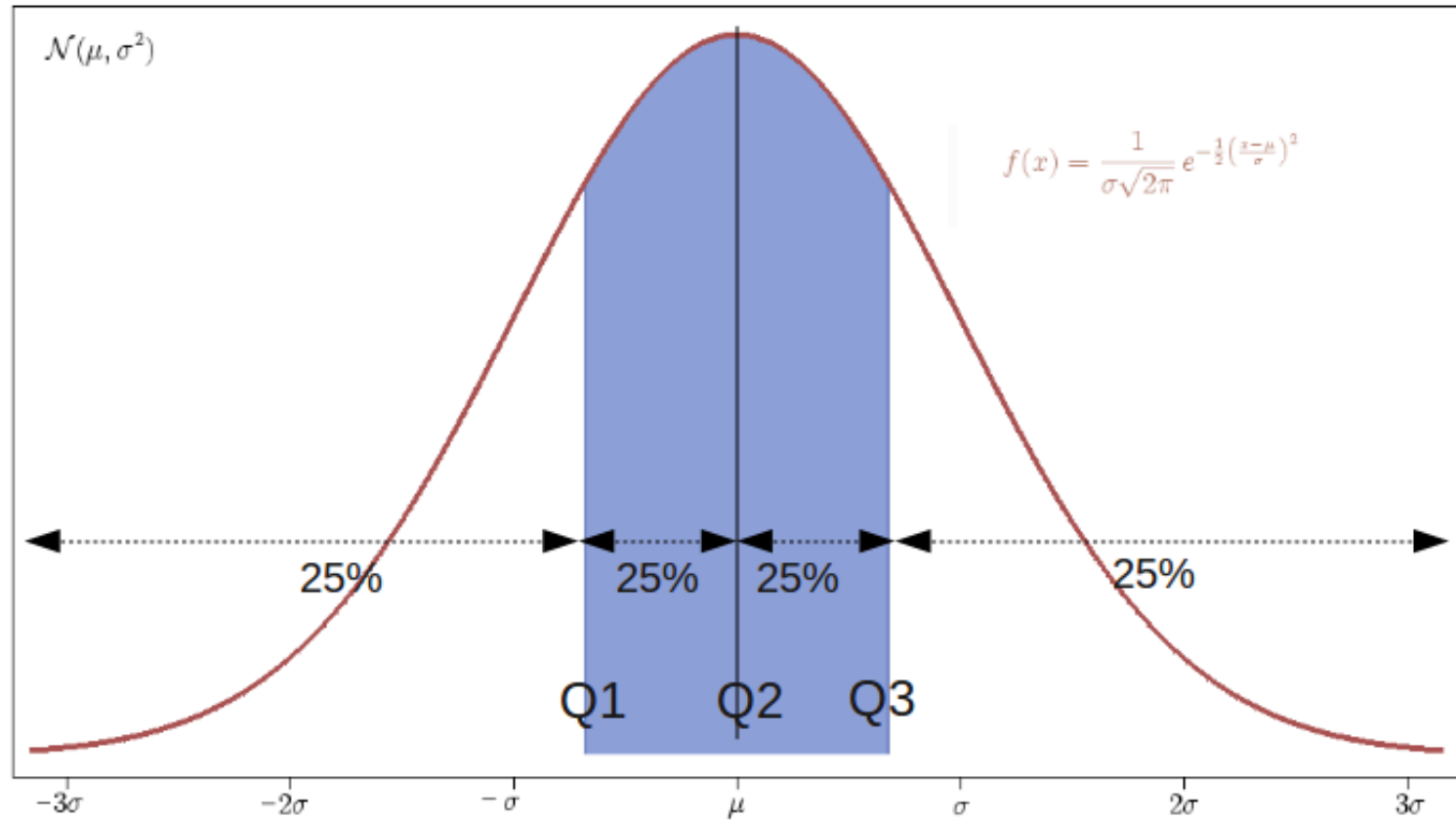


Leptokurtic distribution
High degree of peakedness
Kurtosis > 0

Percentil e Quartil



Percentil e Quartil



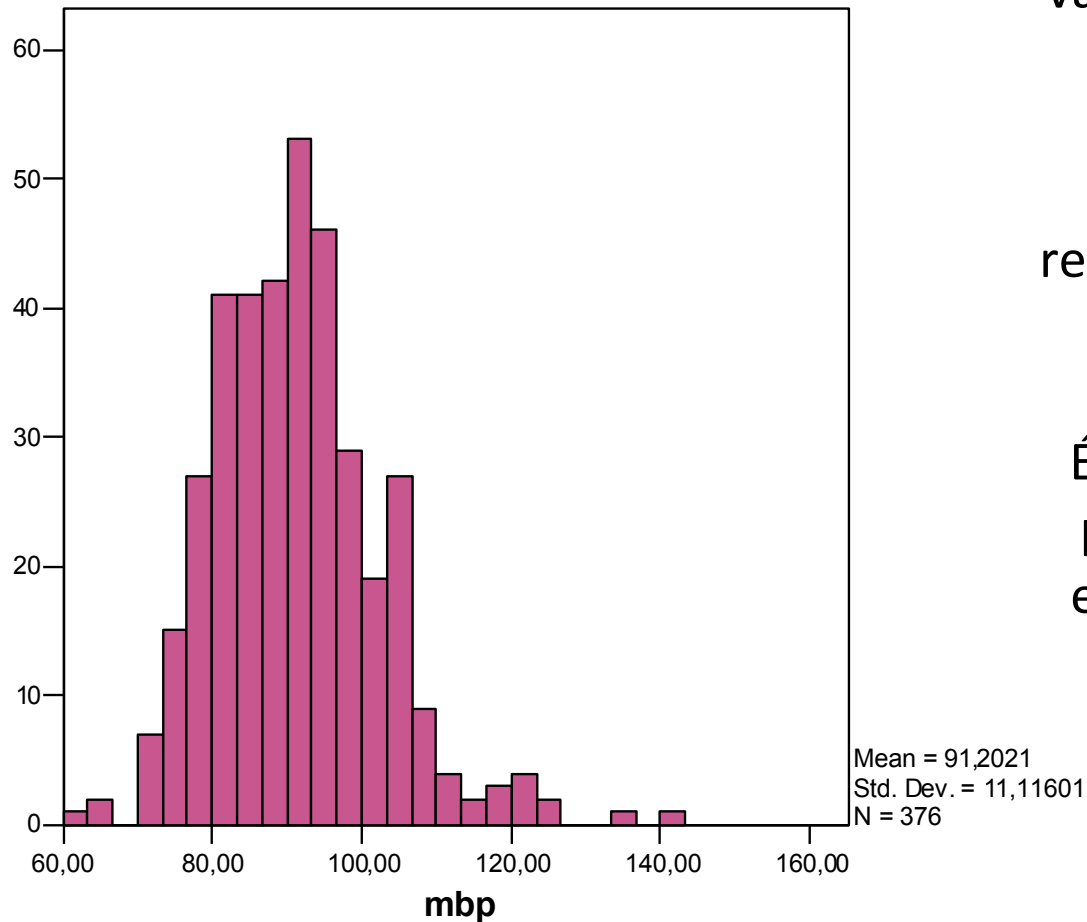
Q2 = Mediana

Histograma

Curva normal

Probabilidade

Histograma



Representação de variáveis
contínuas

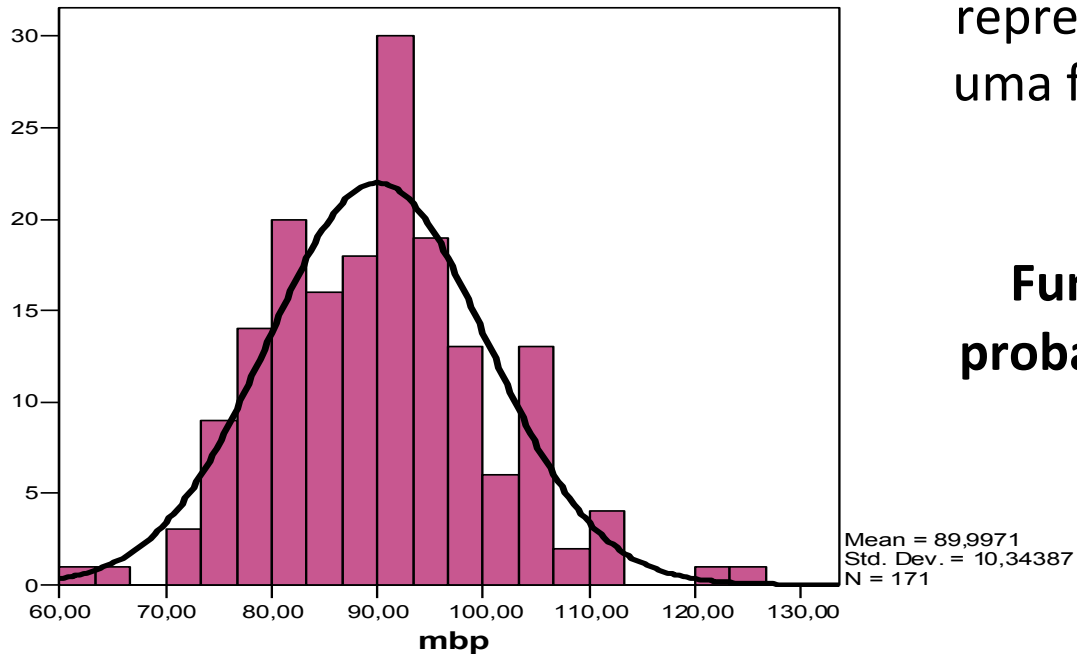
Apropriado para medidas de
variáveis de intervalo, razão e
às vezes variáveis ordinais

A largura das barras
representa o numero de casos
(frequência)

É utilizado para representar
percentagens uma vez que
estas são mais significativas
que os números absolutos

Curva normal Probabilidade

Histograma



Probabilidade empírica =
contagem/ total

Valor esperado de
probabilidade: permite
representar o todo; carece de
uma função que seja aplicável
no geral.

**Função de Densidade de
probabilidade & Distribuição
Normal**

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

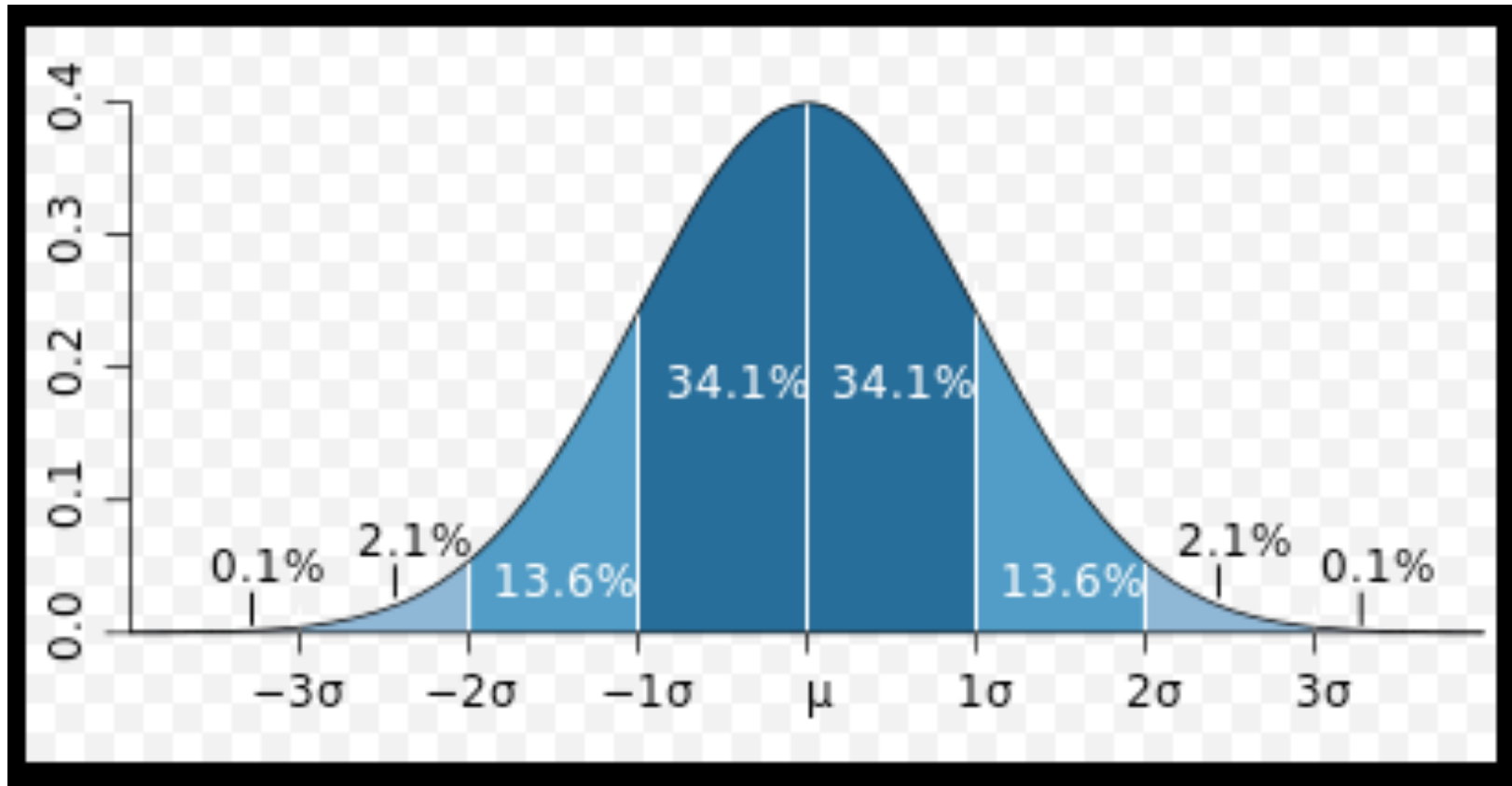
Probabilidade

Curva Normal

Através de médias e desvio padrão conhecidos (variável quantitativa contínua), consegue-se determinar a probabilidade esperada para uma variável através da função de densidade de probabilidade.

Curva normal

Probabilidade

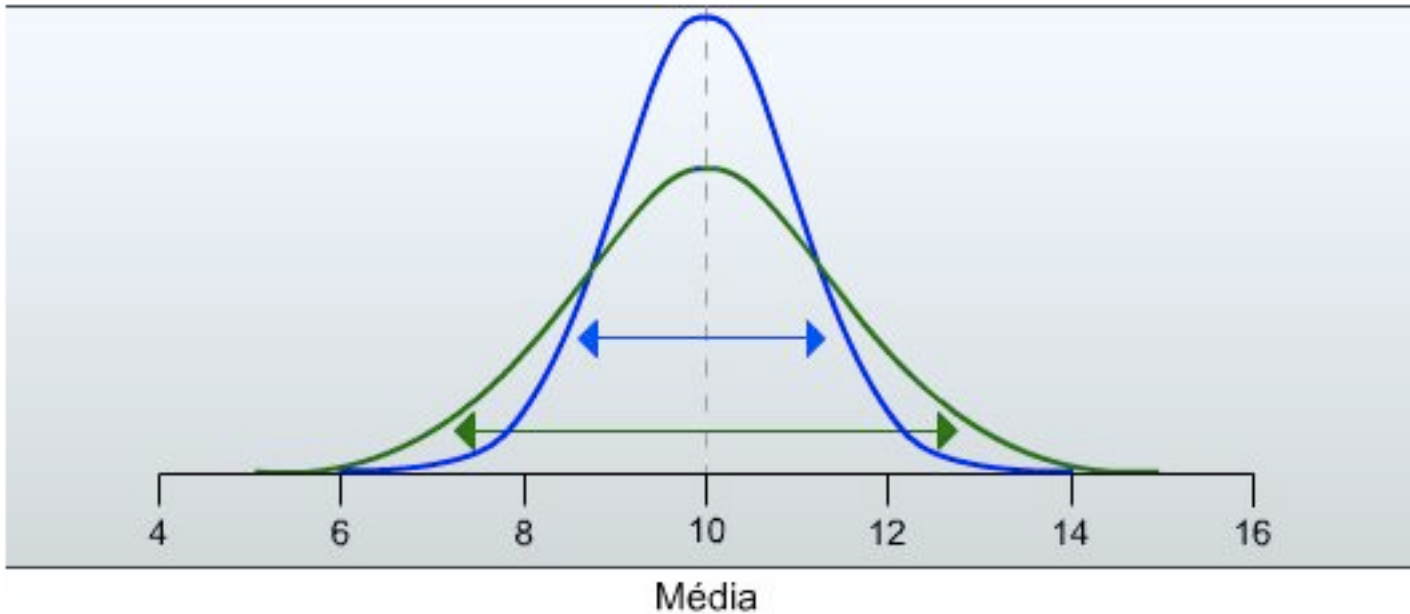


A curva é delimitada por números de desvio padrão iguais que correspondem a Graus de liberdade

A curva é simétrica e tem regiões definidas matematicamente o que torna esta função fácil de trabalhar

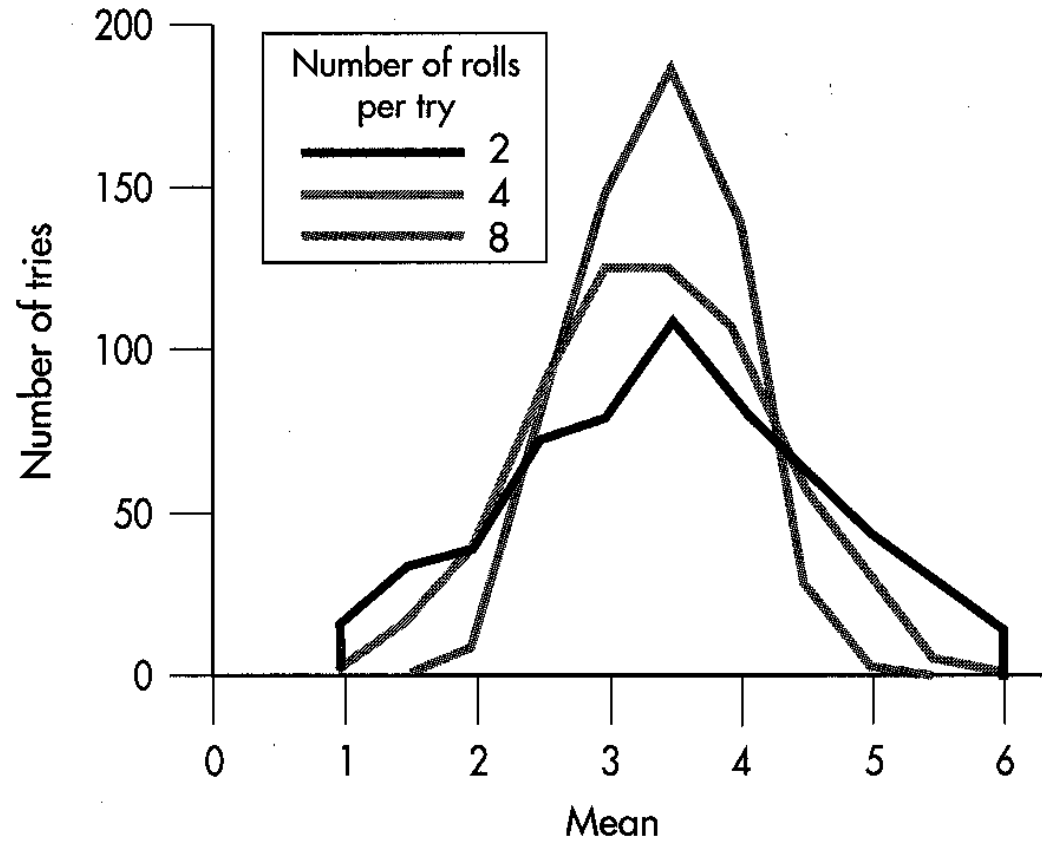
$34 + 34 =$ área para 1 desvio padrão

Curva normal Probabilidade



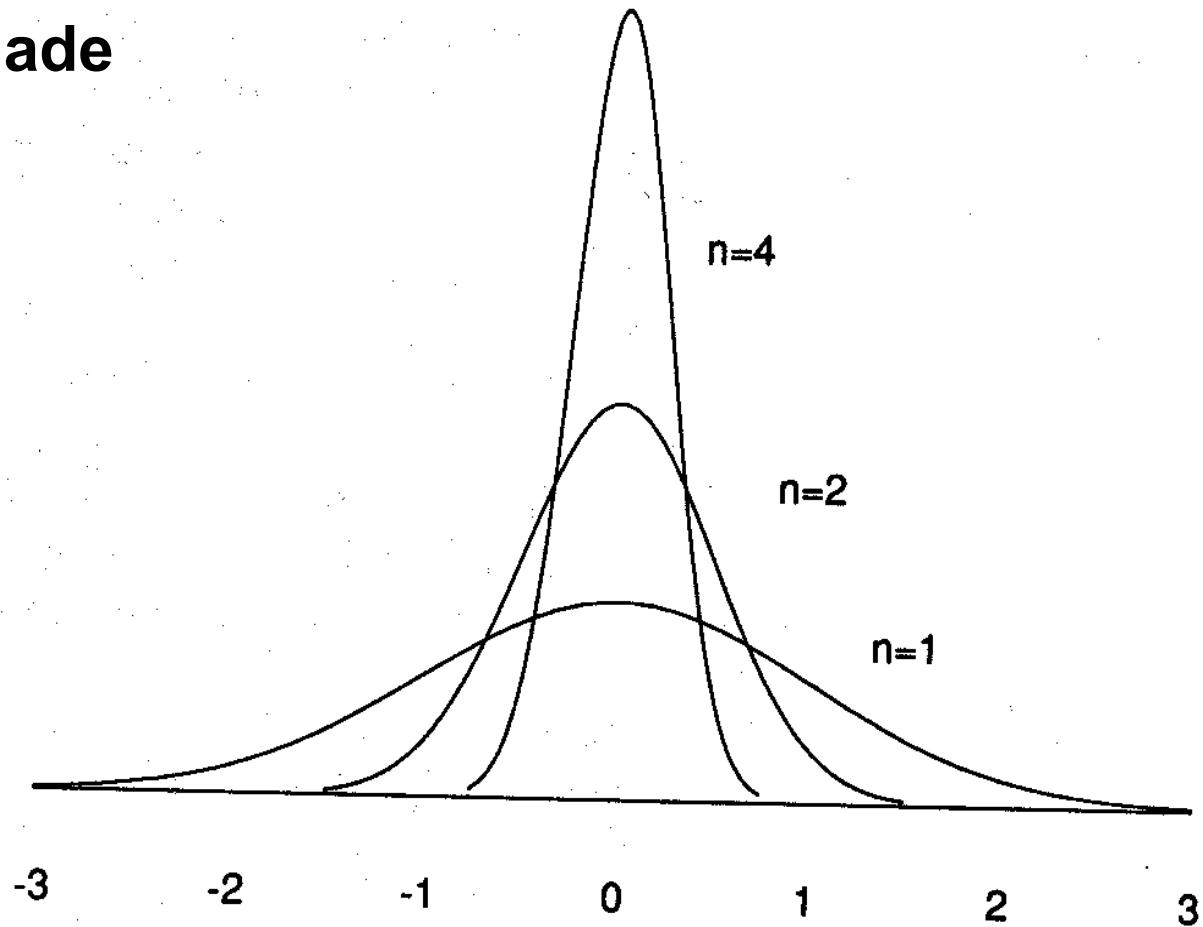
Curvas Normais com médias iguais e desvios padrão diferentes

Curva normal Probabilidade



Curvas Normais com médias iguais e desvios padrão diferentes

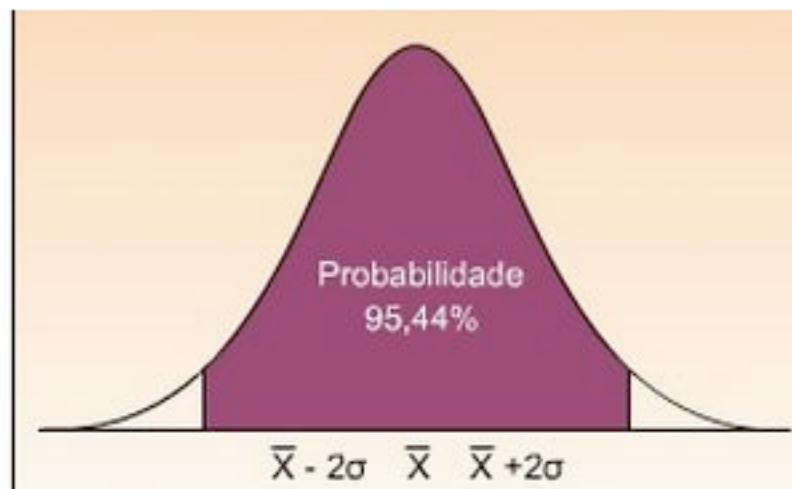
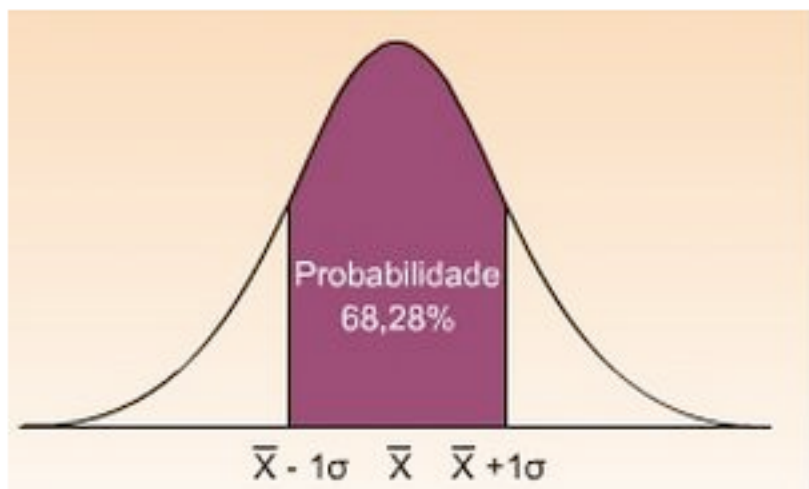
Curva normal Probabilidade



3 distribuições de amostragem para médias de amostras aleatórias com $n = 1$, $n = 2$ e $n = 4$ de uma população $N \sim (0,1)$

Curva normal

Probabilidade



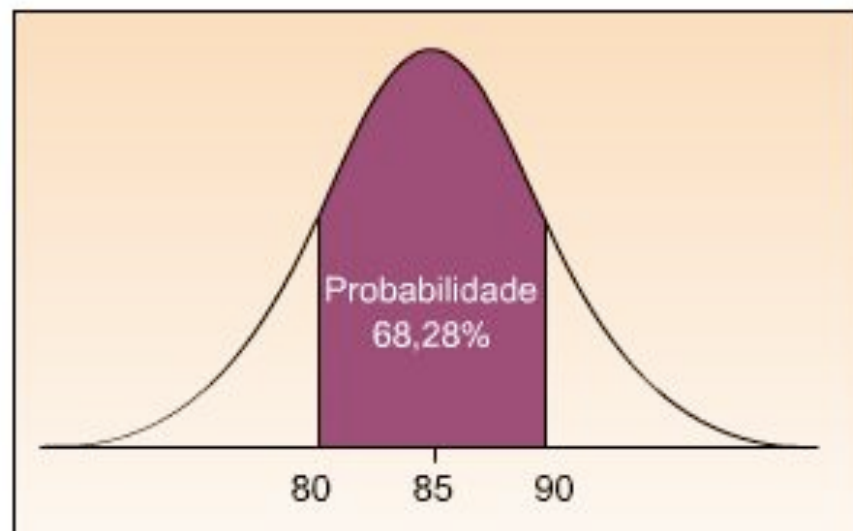
Curva normal

Probabilidade

Sabendo que a média da altura dos pinheiros é 85cm e que o desvio padrão é 5cm:

Neste caso, podemos esperar que **68,28%** dos pinheirinhos tenham alturas contidas no intervalo entre **80cm e 90cm**, uma vez que:

- A média menos 1 desvio-padrão é 80cm ($85 - 5$); e
- A média mais 1 desvio-padrão é 90cm ($85 + 5$).



Curva normal

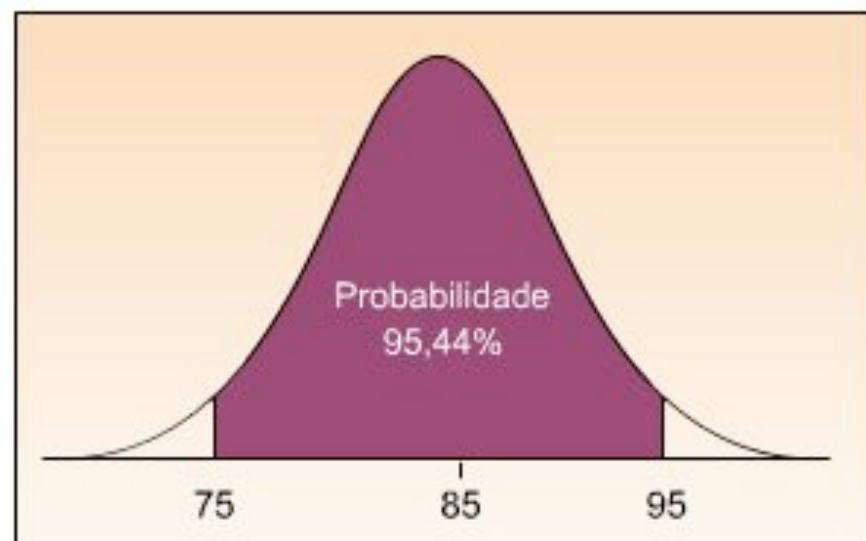
Probabilidade

Sabendo que a média da altura dos pinheiros é 85cm e que o desvio padrão é 5cm:

Também podemos dizer que **95,44%** dos pinheirinhos têm alturas que estão no intervalo de **75cm até 95cm**, uma vez que:

- A média menos 2 desvios-padrão é 75cm ($85 - 2 \times 5$); e
- A média mais 2 desvios-padrão é 95cm ($85 + 2 \times 5$).

Portanto, apenas 4,56% ($100 - 95,44$) restantes estão fora desse intervalo.



Curva normal

Probabilidade

A curva normal de uma distribuição de dados varia segundo o N e a variabilidade da amostra

Permite interpretar os dados corretamente

A curva normal permite ver a dispersão dos dados e conseqüentemente tratamento estatístico dos mesmo.

Através do cálculo integral a curva normal pode ser padronizada:

A distribuição normal de média $\mu=0$ e $\sigma^2=1$ é denominada distribuição normal reduzida

A variável Z (variável contínua) obtém-se da variável X (entre 0 e 1):

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

Curva normal

Probabilidade

A distribuição normal de média $\mu=0$ e $\sigma^2=1$ é denominada distribuição normal reduzida

A variável Z (variável contínua) obtém-se da variável X (entre 0 e 1):

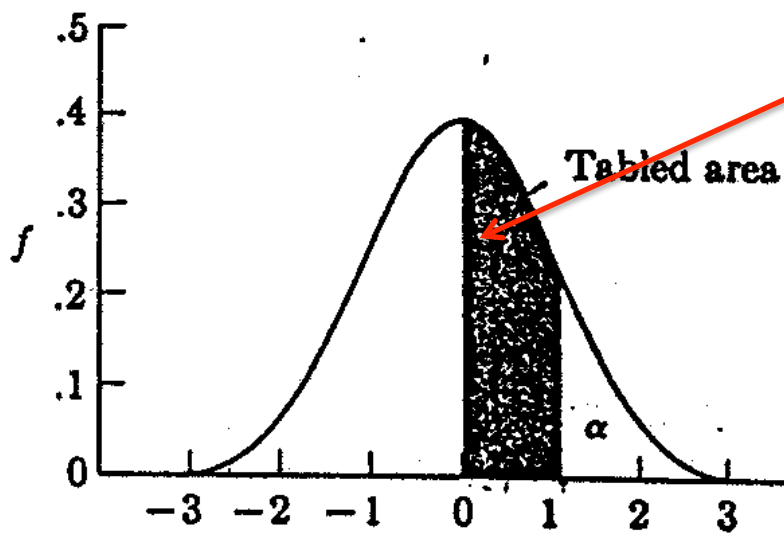
$$Z = \frac{X - \mu}{\sigma}$$

$$Z > 0 \rightarrow X > u$$

$$Z < 0 \rightarrow X < u$$

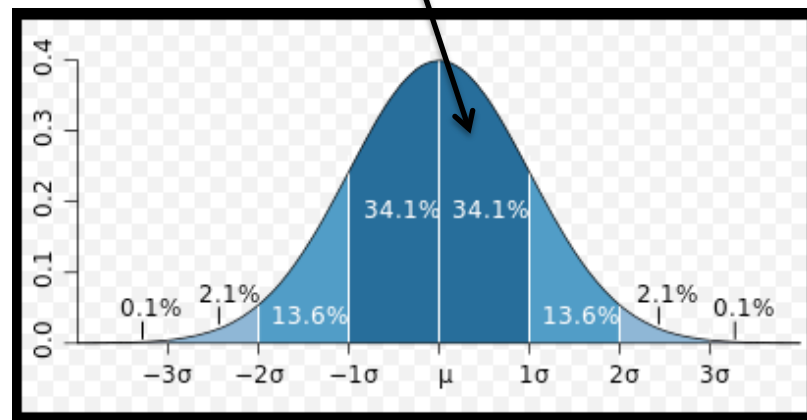
Curva normal

Probabilidade



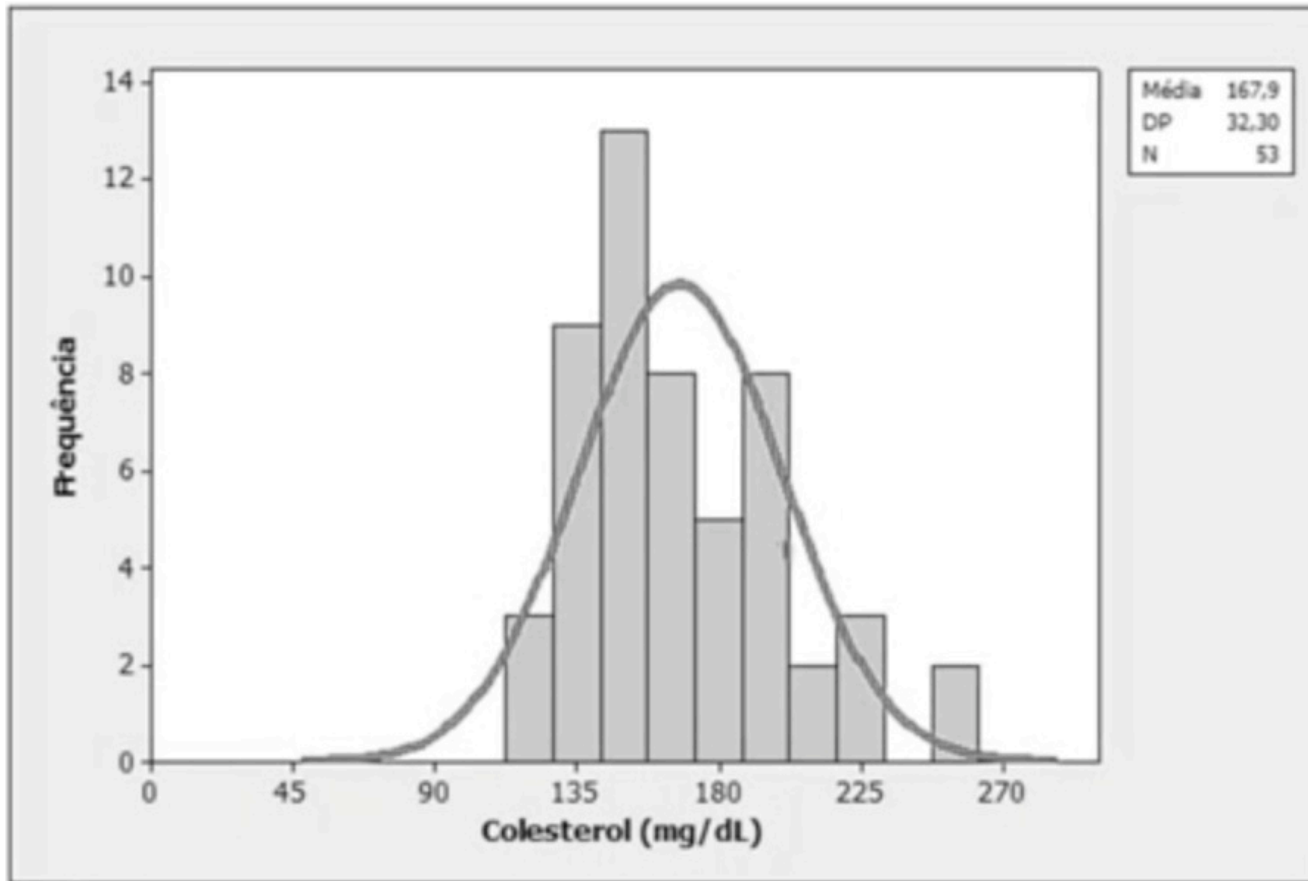
$$\text{Argument} = \frac{Y - \mu}{\sigma}$$

$$P(0 \leq Z \leq 1) = 0,34$$



Curva normal Probabilidade

Média: 167,9
Desvio padrão 32,30
N= 53



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

$$Z = (90 - 167,9) / 32,3 = -2,41$$

Standard deviation

units 0 1 2 3 4 5 6 7 8 9

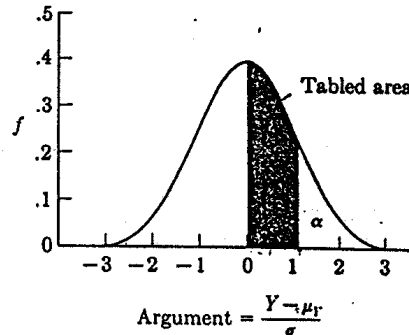
Curve
Prob

0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.499767									
3.6	.499869									
3.7	.499892									
3.8	.499928									
3.9	.499952									
4.0	.499968									
4.1	.499979									
4.2	.499987									
4.3	.499991									
4.4	.499995									
4.5	.499997									
4.6	.499998									
4.7	.499999									
4.8	.499999									
4.9	.500000									

Média: 167,9
Desvio padrão 32,30
N= 53

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

$$Z = (90 - 167,9) / 32,3 = -2,41$$



Standard
deviation
units

	0	1	2	3
0.0	.0000	.0040	.0080	.0120
0.1	.0398	.0438	.0478	.0517
0.2	.0793	.0832	.0871	.0910
0.3	.1179	.1217	.1255	.1293
0.4	.1554	.1591	.1628	.1664
0.5	.1915	.1950	.1985	.2019
0.6	.2257	.2291	.2324	.2357
0.7	.2580	.2611	.2642	.2673
0.8	.2881	.2910	.2939	.2967
0.9	.3159	.3186	.3212	.3238
1.0	.3413	.3438	.3461	.3485
1.1	.3643	.3665	.3686	.3708
1.2	.3849	.3869	.3888	.3907
1.3	.4032	.4049	.4066	.4082
1.4	.4192	.4207	.4222	.4236
1.5	.4332	.4345	.4357	.4370
1.6	.4452	.4463	.4474	.4484
1.7	.4554	.4564	.4573	.4582
1.8	.4641	.4649	.4656	.4664
1.9	.4713	.4719	.4726	.4732
2.0	.4772	.4778	.4783	.4788
2.1	.4821	.4826	.4830	.4834
2.2	.4861	.4864	.4868	.4871
2.3	.4893	.4896	.4898	.4901
2.4	.4918	.4920	.4922	.4925
2.5	.4938	.4940	.4941	.4943
2.6	.4953	.4955	.4956	.4957

Média: 167,9

Desvio padrão 32,30

N= 53

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

$$Z = (90 - 167,9) / 32,3 = -2,41$$

Curva normal
Probabilidade

Distribuição de média na amostragem

- Se a amostra é aleatória, a média da amostra também é aleatória
- Antes de fazer a selecção da amostra a média é uma variável aleatória à qual está associada uma distribuição de probabilidades
- O valor esperado de média na amostragem é
- Variância de média na amostragem é
- O desvio padrão de média na amostragem toma a designação de erro padrão da média e é igual a

$$\bar{X} : E\{\bar{X}\} = \mu$$

$$\bar{X} : Var\{\bar{X}\} \text{ ou } \sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$$

$$\frac{\sigma}{\sqrt{n}}$$

Teorema do limite central

- Para a maioria das populações a distribuição de média na amostragem é aproximadamente normal quando o tamanho da amostra é suficientemente grande
- Se a distribuição da variável X na população é normal:
 - A distribuição de média na amostragem é sempre normal, qualquer que seja o tamanho da amostra.

Estimação da média da população

- Estimação pontual
 - Estimar uma característica da população através de um simples valor
 - **Estimador:** Variável aleatória utilizada para estimar a característica da população
 - **Estimativa:** Valor numérico do estimador

Estimação da média da população

Se a amostra é retirada de uma população infinita, ou com reposição de uma população finita:

- **As estatísticas são estimadores sem vício**

$$E\{\bar{X}\} = \mu$$

$$E\{s^2\} = \sigma^2$$

Estimação por intervalos de confiança

- Se a distribuição de média na amostragem é normal
 - Através da normal reduzida é possível calcular a probabilidade da ocorrer entre 2 valores da distribuição

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Então,

$$\text{Prob} \left\{ -z \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z \right\} = C \quad \text{ou} \quad \text{Prob} \left\{ \bar{x} - z \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z \frac{\sigma}{\sqrt{n}} \right\} = C$$

Ou seja,

$$\text{Prob} \left\{ \bar{x} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \right\} = C$$

C=Confiança=1- α (com α = risco)

Estimação por intervalos de confiança

- O intervalo de confiança C (ou risco α) da estimação de μ é:

$$\left[\bar{x} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \right]$$

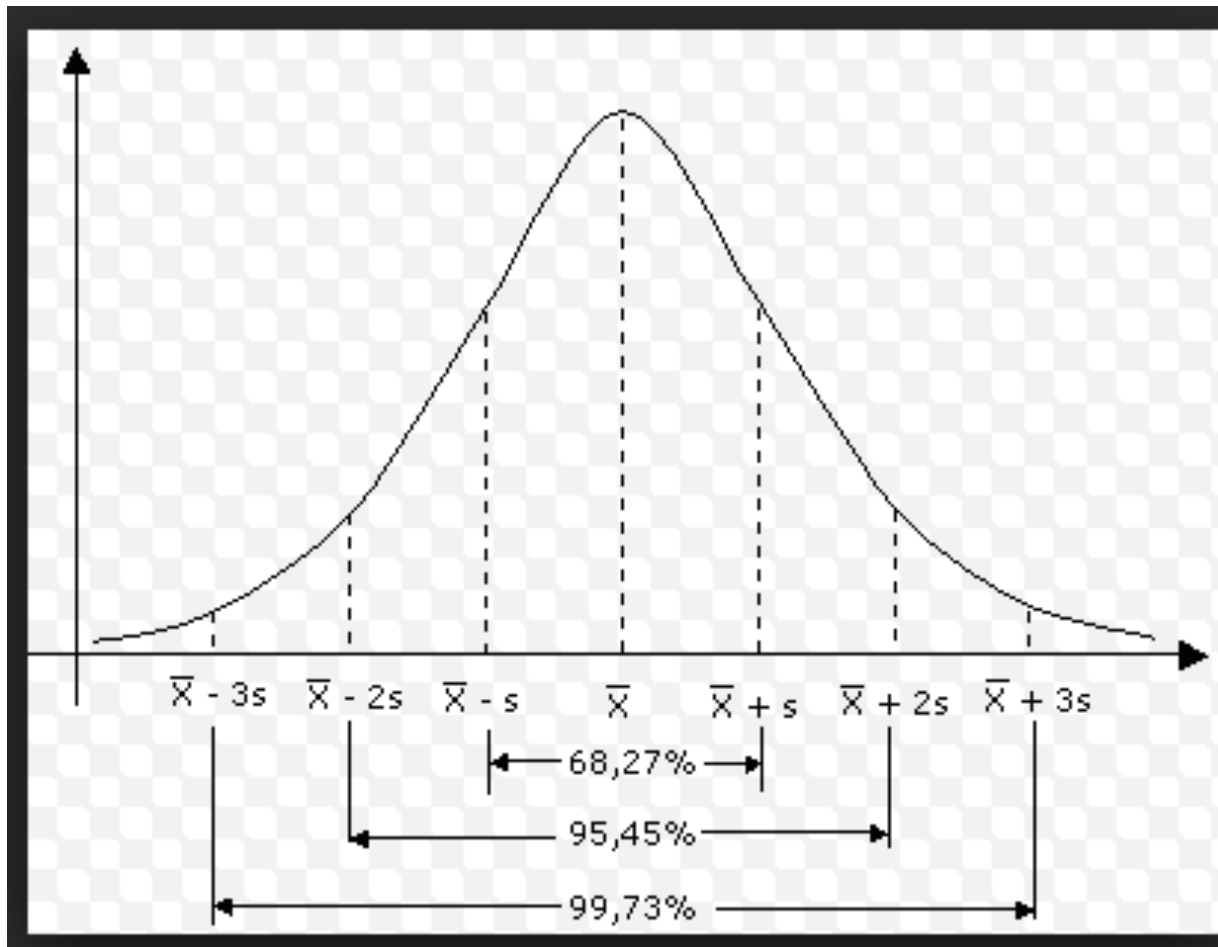
- Se o desvio padrão da população não é conhecido:

como estimador sem vício de

$$\frac{s}{\sqrt{n}}$$

$$\frac{\sigma}{\sqrt{n}}$$

Estimação por intervalos de confiança



$$95\% = \bar{X} \pm 1.96 (SE)$$

$$99\% = \bar{X} \pm 2.58 (SE)$$

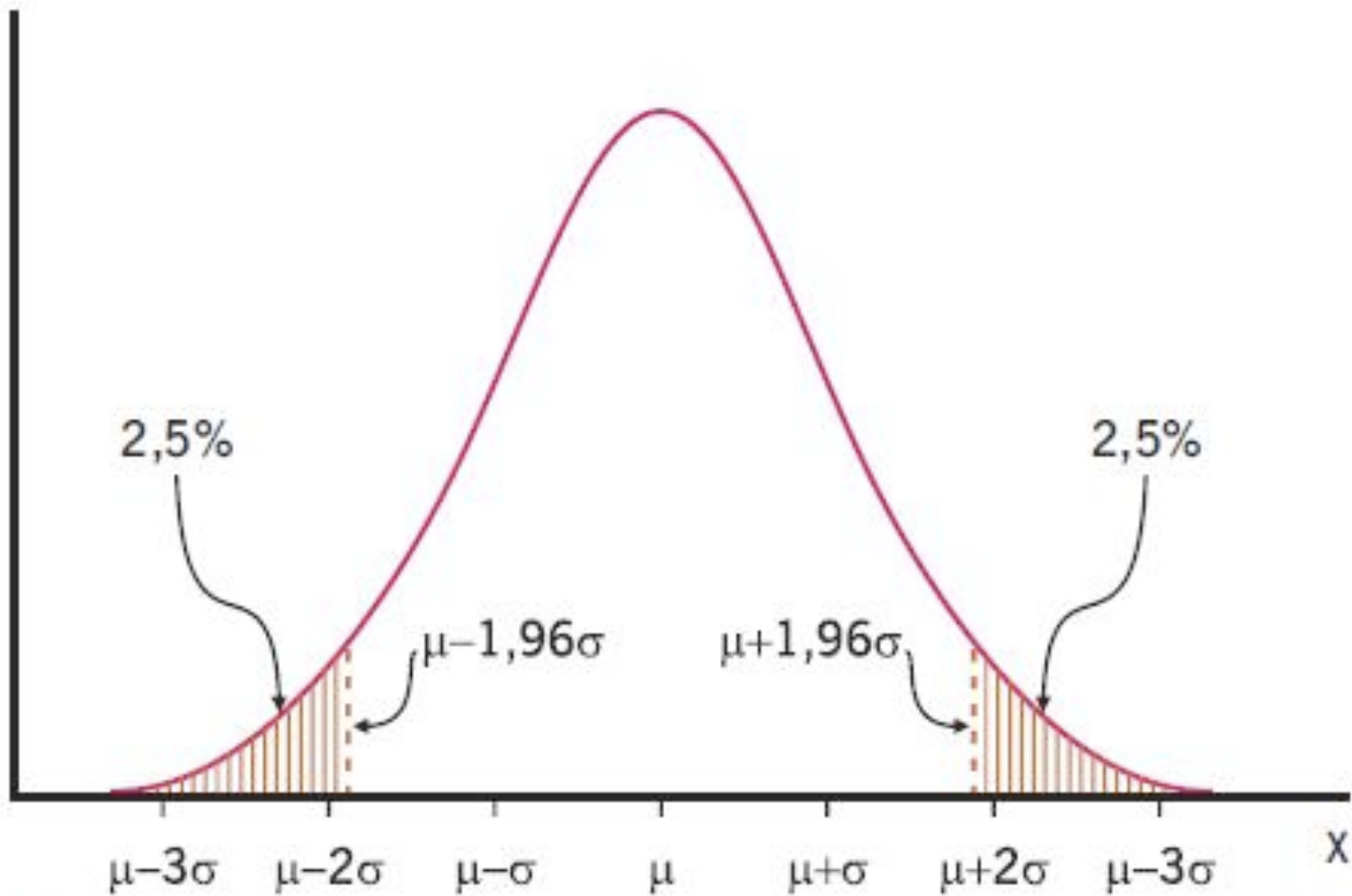
$$SE = SD/\sqrt{n}$$

SE = Standard Error = Erro Padrão

SD = Standard Desviation = Desvio Padrão

Estimação por intervalos de confiança

$\sigma = \text{SE}$
 $U = \text{média}$



Estimação por intervalos de confiança

Exercício:

Quais os valores do intervalo de confiança de 95% estimados para uma média de 100 e desvio padrão de 27, numa amostra com $n=81$?

Cálculo do erro padrão:

$$SE = SD/\sqrt{n}$$

Cálculo do Intervalo de confiança de 95% (95% CI)

$$95\% = \bar{X} \pm 1.96 (SE)$$

Exercícios:

Os valores de uma determinada investigação formam uma distribuição normal.

Média = 70; SD= 15

1. 68% dos valores estão entre _____ e _____
2. 96% dos valores estão entre _____ e _____
3. Numa distribuição assimétrica positiva os valores estendem-se à _____ ou estão representados valores mais _____ da distribuição

(Direita/esquerda & altos/baixos)

Estimação por intervalos de confiança

Exemplo:

- Utilizando os dados das 78 mortes por síndrome de morte súbita:
- X = peso à nascença dos bebés em gramas
- $\bar{X} = 2993,6 \approx 2994$ g
- Sabe-se que $\sigma = 800$ g
- Calcule o intervalo de confiança a 95% para a média dos casos de morte súbita

$$2994 \pm (1,96) \left(\frac{800}{\sqrt{78}} \right) = 2994 \pm (1,96)(90,6) = 2994 \pm 178$$

- Limite inferior: 2816; limite superior: 3172
- O intervalo de confiança para 95% é [2816, 3172]
- Há 95% de confiança de que a média da população do peso à nascença dos bebés com o síndrome da morte súbita se encontra neste intervalo.