



ANÁLISE DE DADOS

MODELO CONTÍNUO DE DISTRIBUIÇÃO DE PROBABILIDADE: DISTRIBUIÇÃO NORMAL

Prof. Dr. Daniel Caetano

2020 - 1

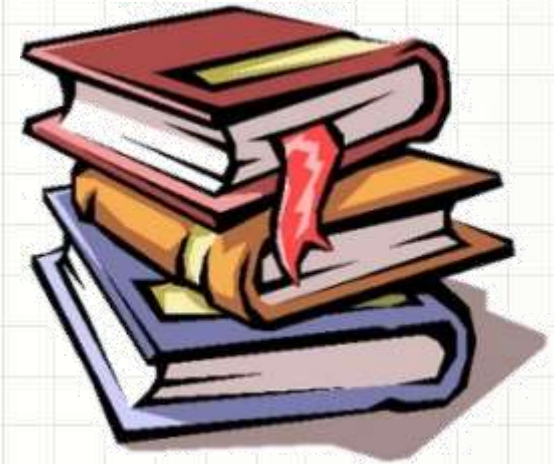
Objetivos

- Compreender os modelos contínuos de probabilidade
- Conhecer e compreender a distribuição normal (ou de Gauss)
- Conhecer o conceito de função densidade de probabilidade

- **Atividade da Aula 9 no SAVA!**



Material de Estudo



Material

Acesso ao Material

Apresentação

<http://www.caetano.eng.br/>
(Análise de Dados – Aula 9)

Material Didático

- Probab. e Estatística Aplicada à Engenharia –
Cap. 6

Minha Biblioteca

Estatística – Teoria e Aplicações usando MS Excel –
Cap. 6
Estatística Aplicada e Probabilidade para Engenheiros
– Cap. 4



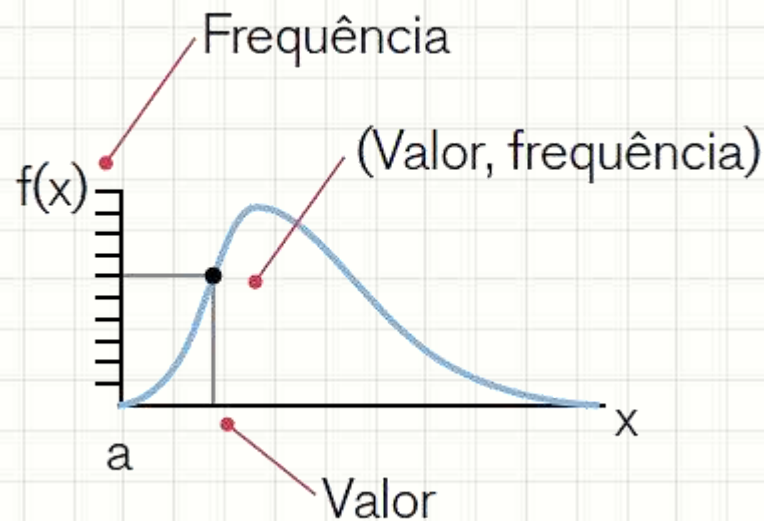
MODELOS CONTÍNUOS DE PROBABILIDADE

Modelos Contínuos de Probabilidade

- O que são?
 - São modelos que descrevem a probabilidade de eventos representados por números reais
- Exemplos:
 - Qual a probabilidade de realizar uma medida e o valor ser 1,305m?
 - Qual a probabilidade de medir o peso de uma pessoa ao acaso e ser 80,75kg?.

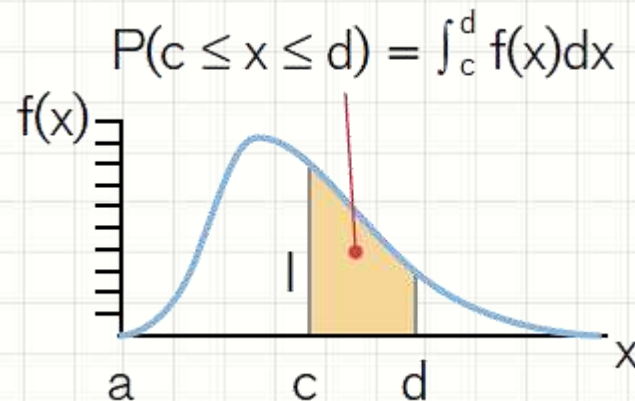
Modelos Contínuos de Probabilidade

- A função de probabilidade...
 - Descreve a frequência associada a cada valor
 - Função Densidade de Probabilidade (FDP)



Modelos Contínuos de Probabilidade

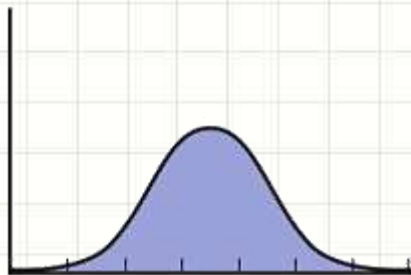
- Função Densidade de Probabilidade (FDP)
 - Se quiser saber a probabilidade de o valor estar entre dois limites (c,d), basta pegar a área:



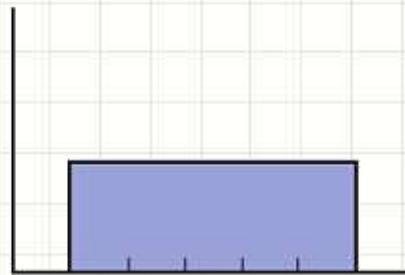
- A área total sob a curva deve valer 1!

Modelos Contínuos de Probabilidade

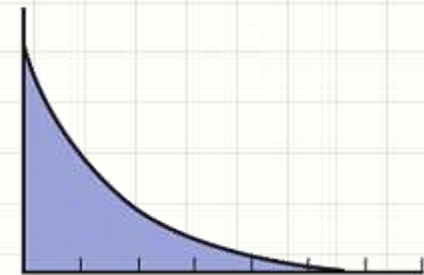
- Modelo mais importante
 - Distribuição Normal
 - Quando existe um valor “real” que estamos tentando medir (a altura de uma pessoa, por exemplo)
 - Replicamos o experimento diversas vezes
 - A distribuição dos resultados será a distribuição normal
 - Outro nome: distribuição Gaussiana.



Valores de X
Painel A
Distribuição Normal



Valores de X
Painel B
Distribuição Uniforme



Valores de X
Painel C
Distribuição Exponencial



DISTRIBUIÇÃO NORMAL

Distribuição Normal

- É a distribuição mais usada na estatística
- Por quê?
 - Inúmeras variáveis contínuas no mundo dos negócios se assemelham à distribuição normal
 - Pode ser usada para aproximar distribuições discretas
 - Proporciona a base para a inferência estatística clássica (Teorema do Limite Central etc., que veremos na próxima aula).

Distribuição Normal

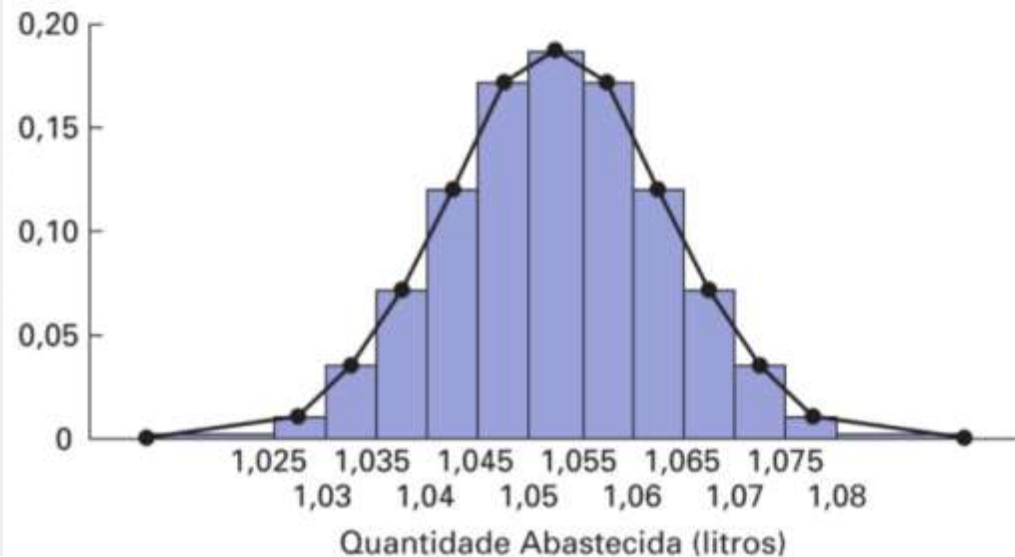
- Exemplo: abastecimento em 10.000 garrafas

Abastecida (litros)	Frequência Relativa
< 1,025	$48/10.000 = 0,0048$
1,025 < 1,030	$122/10.000 = 0,0122$
1,030 < 1,035	$325/10.000 = 0,0325$
1,035 < 1,040	$695/10.000 = 0,0695$
1,040 < 1,045	$1.198/10.000 = 0,1198$
1,045 < 1,050	$1.664/10.000 = 0,1664$
1,050 < 1,055	$1.896/10.000 = 0,1896$
1,055 < 1,060	$1.664/10.000 = 0,1664$
1,060 < 1,065	$1.198/10.000 = 0,1198$
1,065 < 1,070	$695/10.000 = 0,0695$
1,070 < 1,075	$325/10.000 = 0,0325$
1,075 < 1,080	$122/10.000 = 0,0122$
1,080 ou mais	$48/10.000 = 0,0048$
Total	<u>1,0000</u>

Distribuição Normal

- Exemplo: abastecimento em 10.000 garrafas

Abastecida (litros)	Frequência Relativa
< 1,025	$48/10.000 = 0,0048$
1,025 < 1,030	$122/10.000 = 0,0122$
1,030 < 1,035	$325/10.000 = 0,0325$
1,035 < 1,040	$695/10.000 = 0,0695$
1,040 < 1,045	$1.198/10.000 = 0,1198$
1,045 < 1,050	$1.664/10.000 = 0,1664$
1,050 < 1,055	$1.896/10.000 = 0,1896$
1,055 < 1,060	$1.664/10.000 = 0,1664$
1,060 < 1,065	$1.198/10.000 = 0,1198$
1,065 < 1,070	$695/10.000 = 0,0695$
1,070 < 1,075	$325/10.000 = 0,0325$
1,075 < 1,080	$122/10.000 = 0,0122$
1,080 ou mais	$48/10.000 = 0,0048$
Total	1,0000

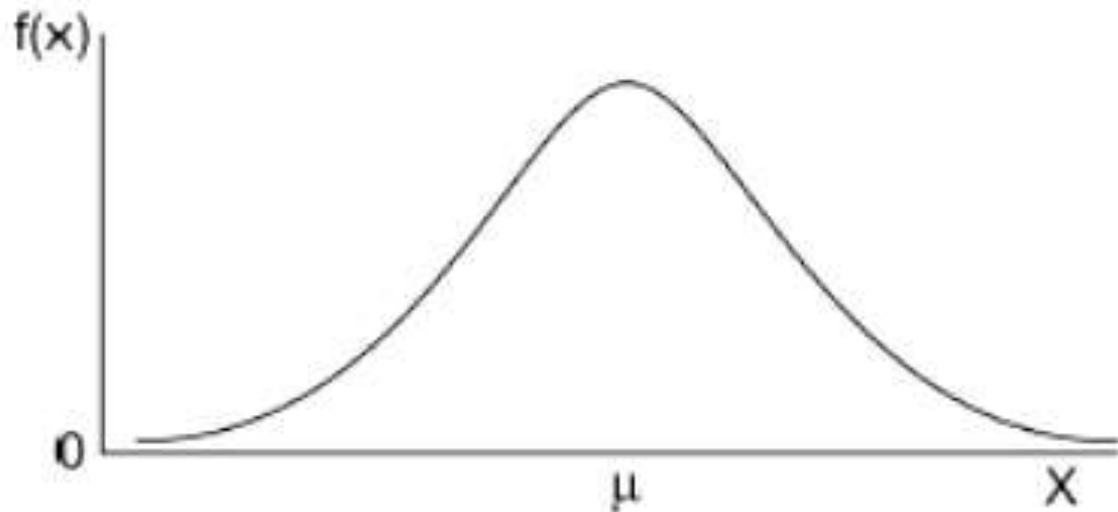


**Observe o
formato!**

Distribuição Normal

- Equação da FDP Normal

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

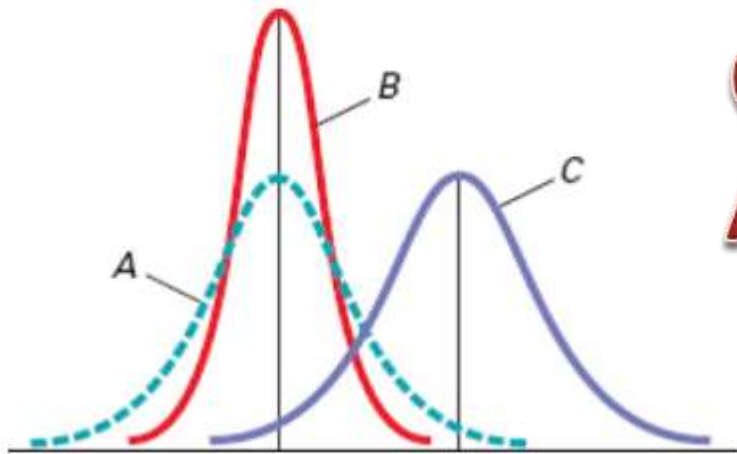


Distribuição Normal

- Equação da FDP Normal

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

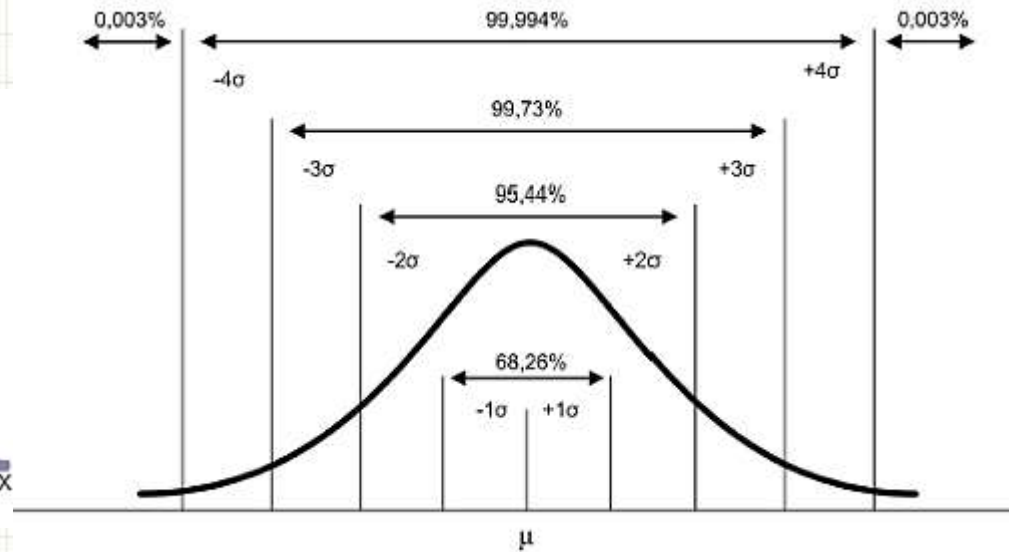
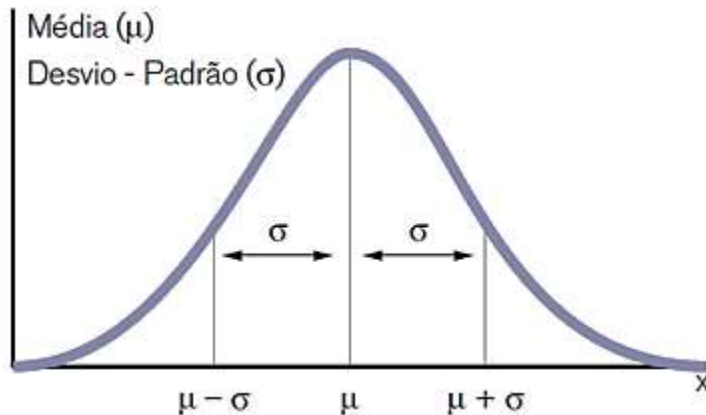
- Variando média μ e desvio padrão σ ...



**O que as curvas
A, B e C têm em
comum?**

Distribuição Normal

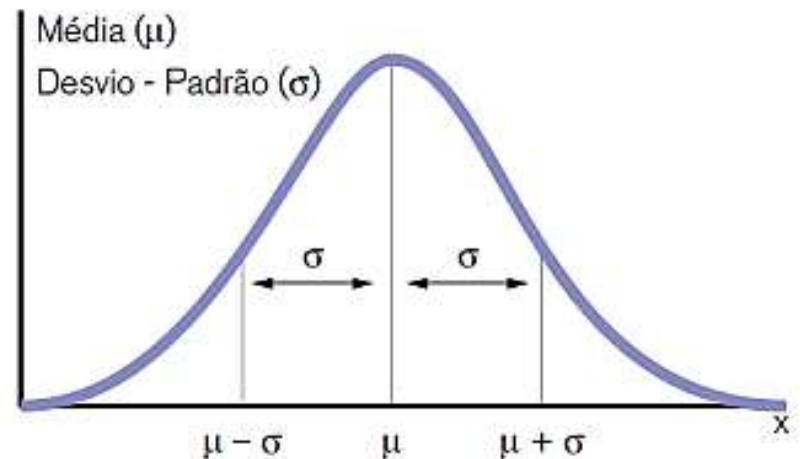
- Equação da FDP Normal $f(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$
 - Para definir a curva, é necessário a média μ e o desvio padrão σ .



- Na falta, estimamos com a média \bar{x} e s

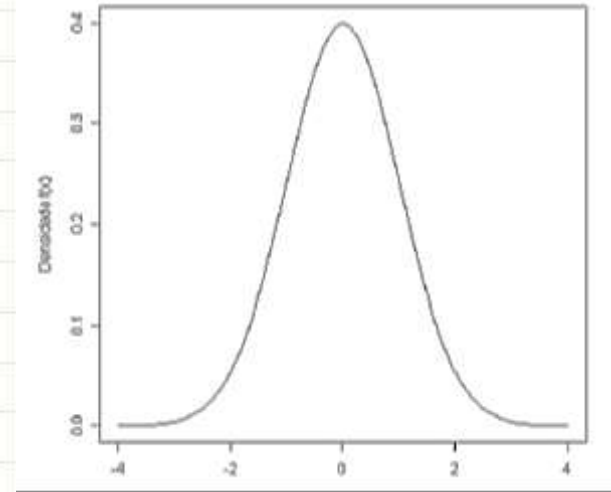
Distribuição Normal

- Propriedades da normal $f(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}$
 - Simétrica em relação a $x = \mu$
 - Único máximo em $x = \mu$
 - Moda = Mediana = μ
 - Tende a 0 quando $x \rightarrow \pm\infty$
 - Duas inflexões em $\mu \pm \sigma$.

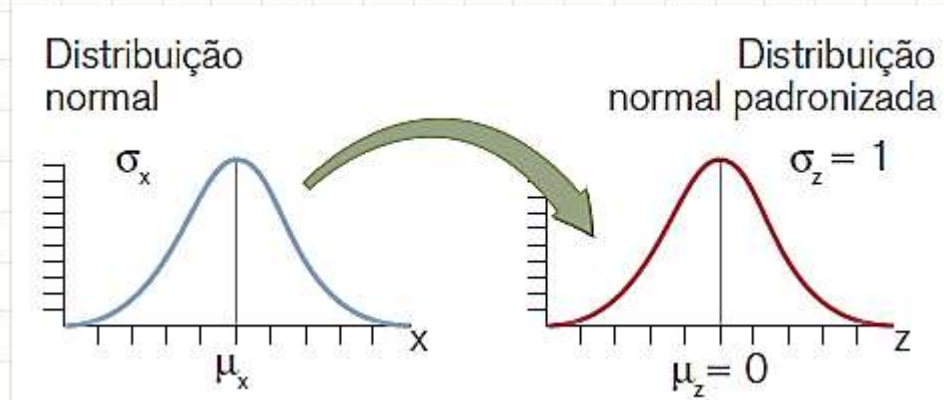


Distribuição Normal Padrão

- É uma distribuição normal...
 - Com média 0
 - Desvio padrão 1

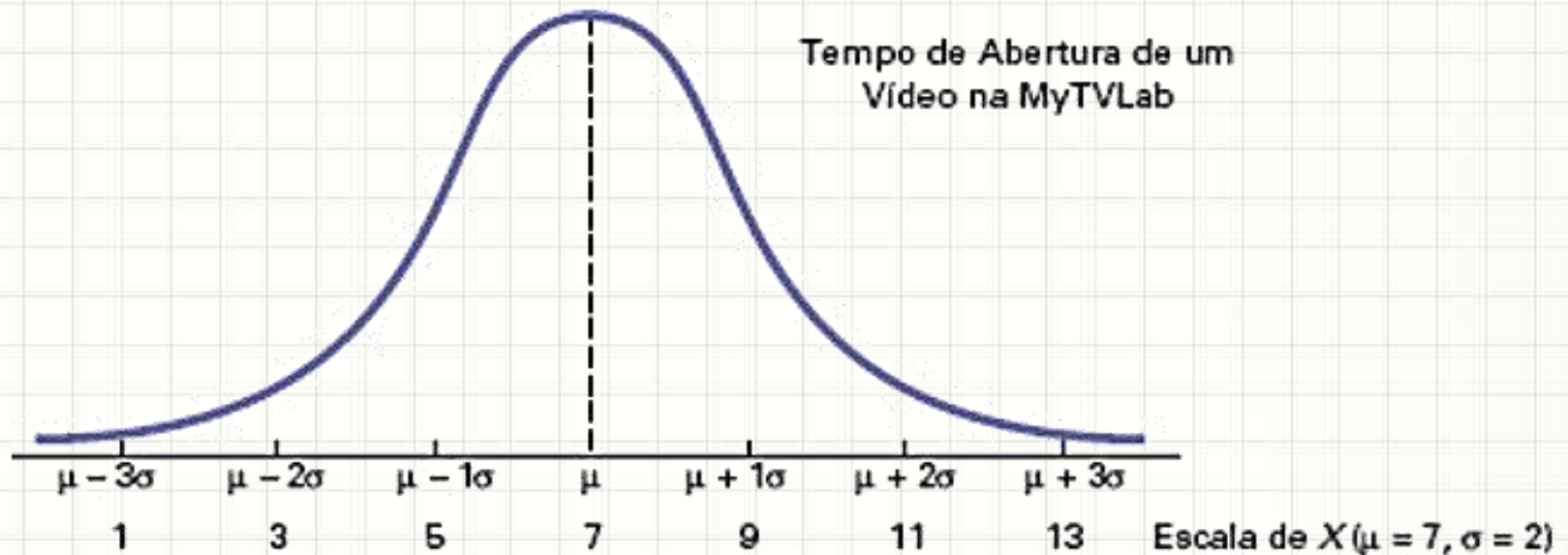


- “Padronizar” uma variável: $Z = \frac{x - \mu}{\sigma}$



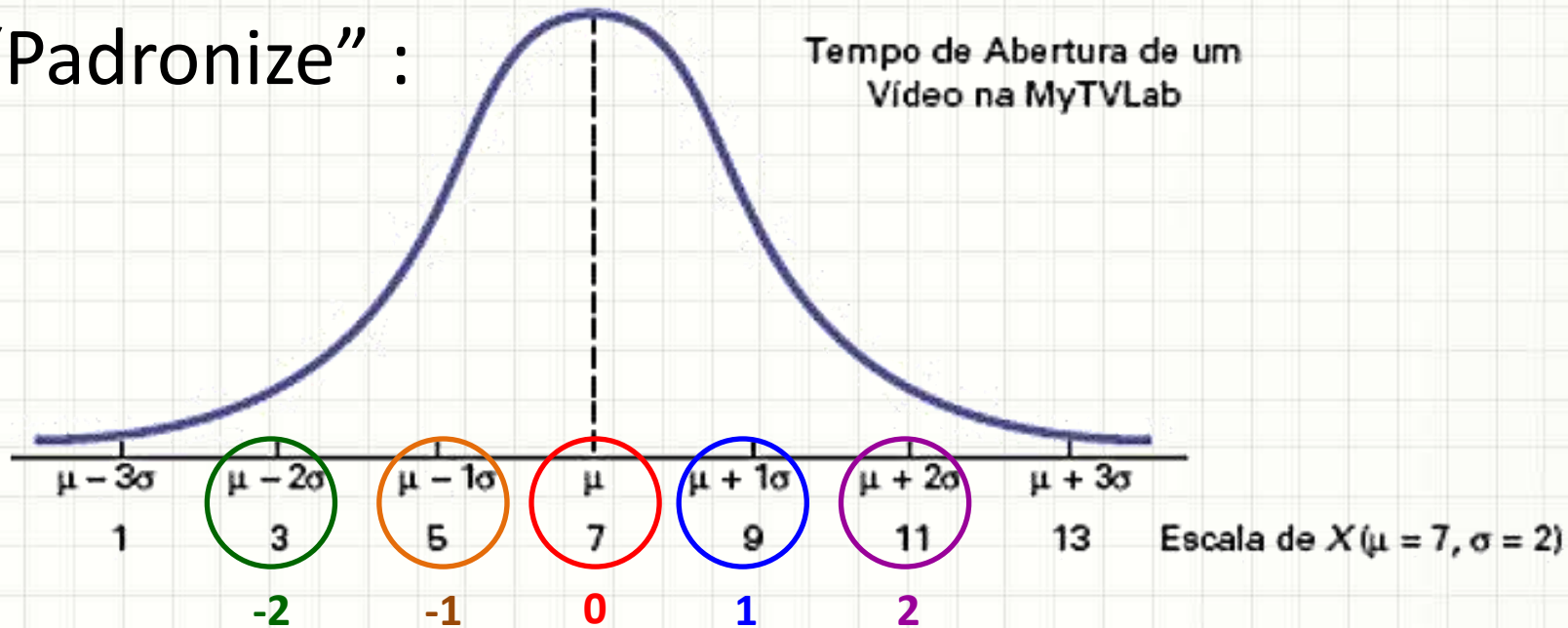
Distribuição Normal Padrão

- Exemplo: “padronize” a distribuição normal relacionada às medidas:



Distribuição Normal Padrão

- “Padronize” :



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

$$Z = \frac{7 - 7}{2} = 0$$

$$Z = \frac{9 - 7}{2} = 1$$

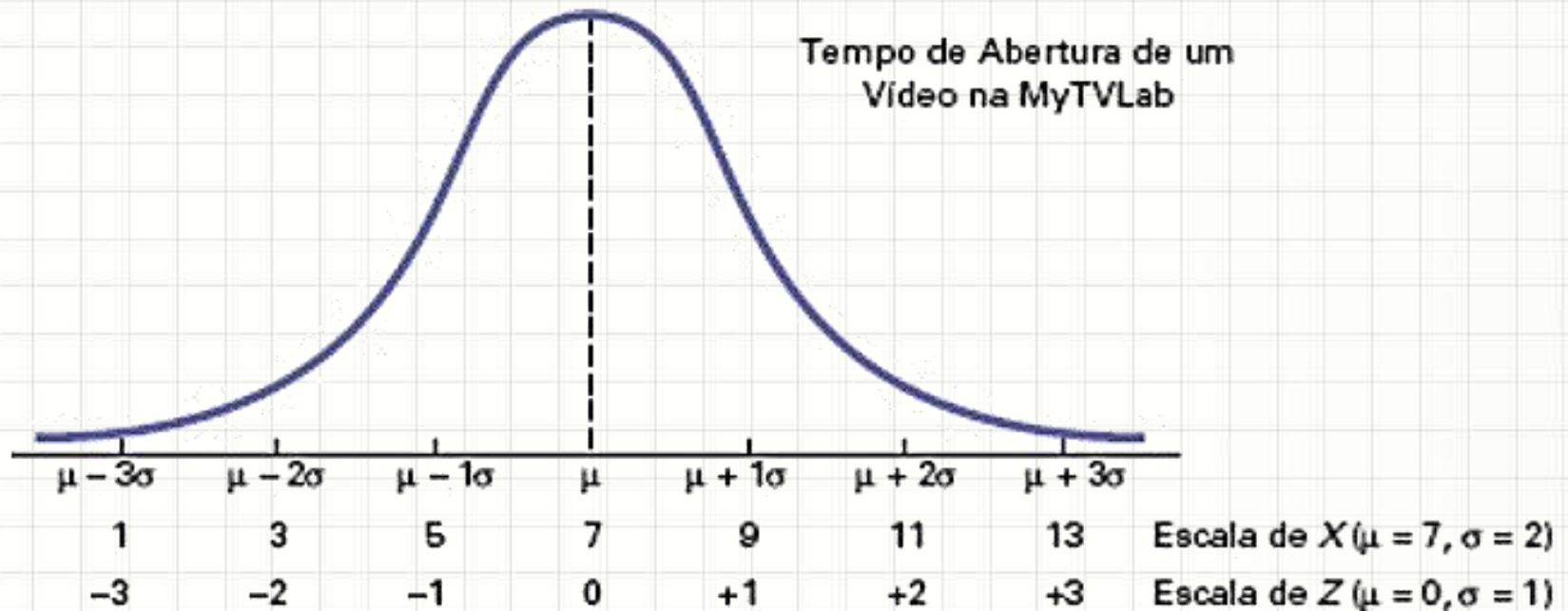
$$Z = \frac{5 - 7}{2} = -1$$

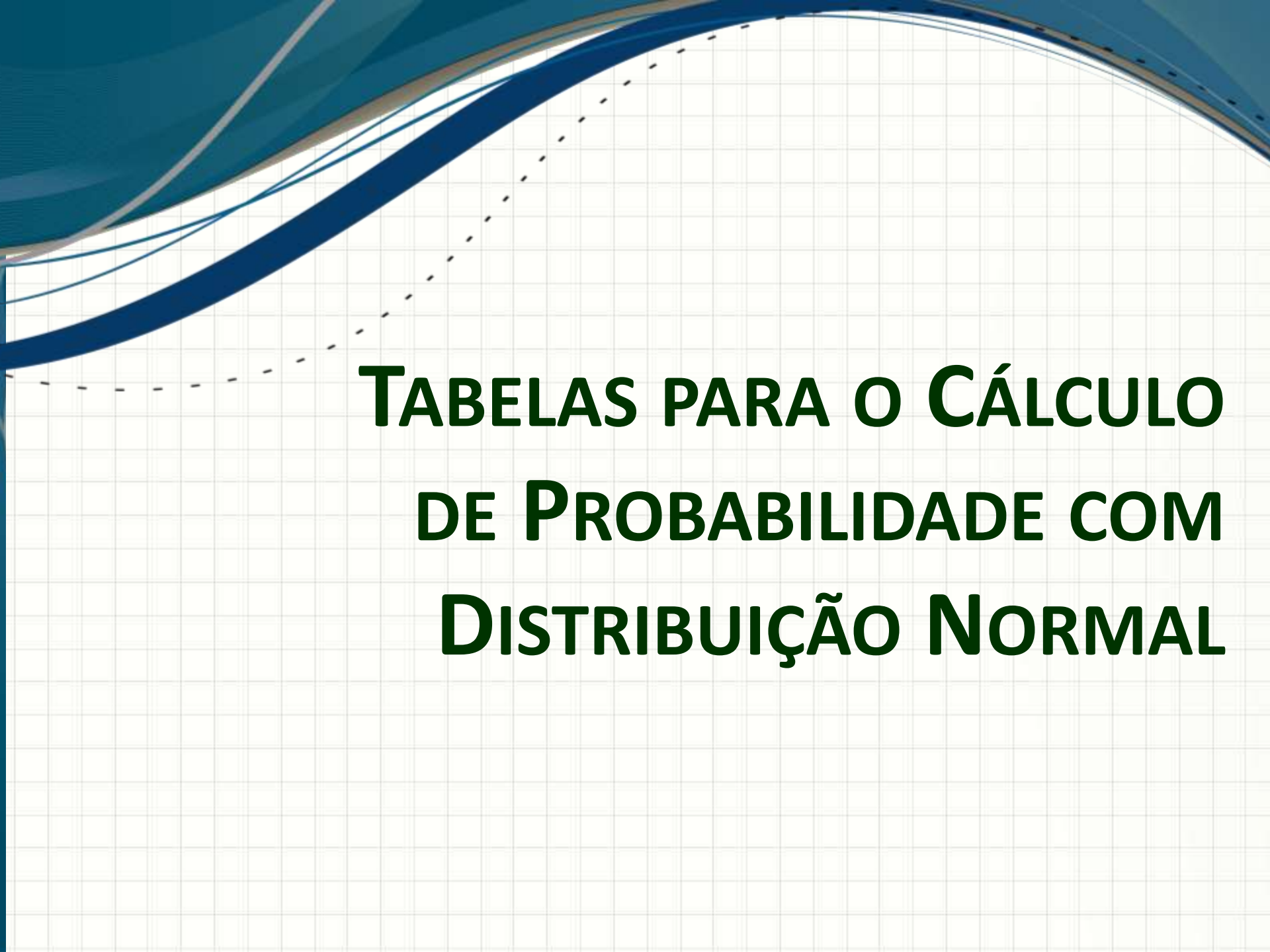
$$Z = \frac{11 - 7}{2} = 2$$

$$Z = \frac{3 - 7}{2} = -2$$

Distribuição Normal Padrão

- “Padronize” :

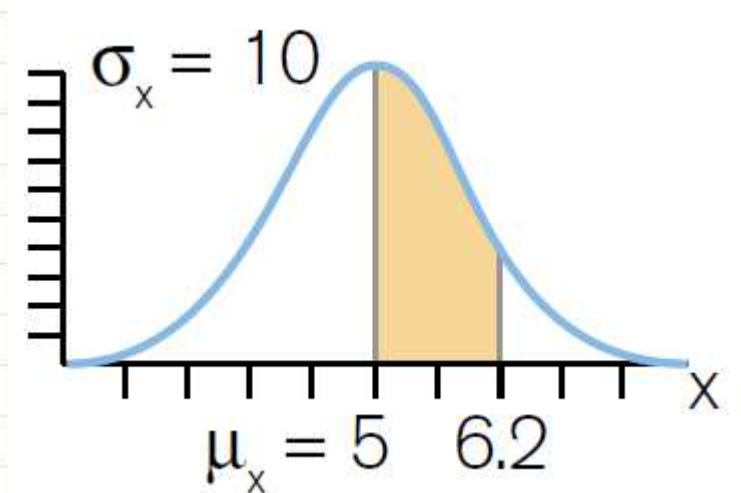




**TABELAS PARA O CÁLCULO
DE PROBABILIDADE COM
DISTRIBUIÇÃO NORMAL**

Determinação de Probabilidades

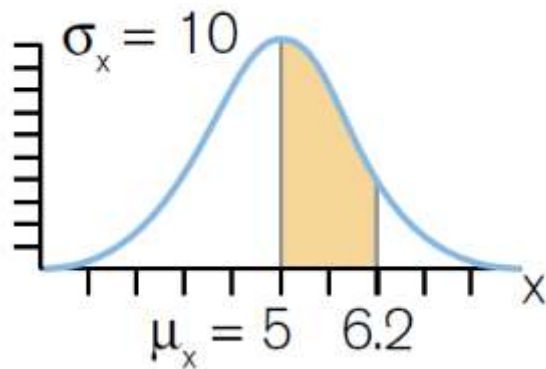
- Área sob a curva da FDP → probabilidade
- Imaginemos que temos o seguinte caso:
 - Queremos saber a probabilidade de o valor medido estar entre 5 e 6,2...



– Como fazer?

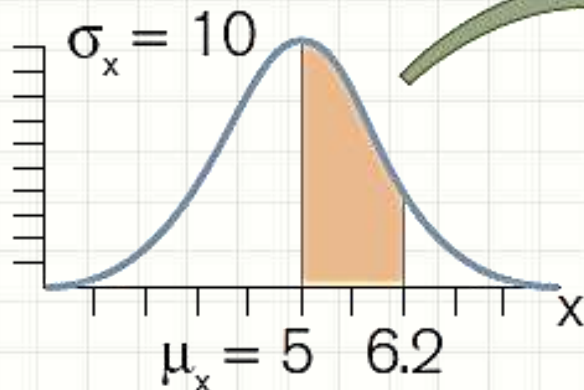
Determinação de Probabilidades

- Primeiro passo: padronizar $Z = \frac{x - \mu}{\sigma}$

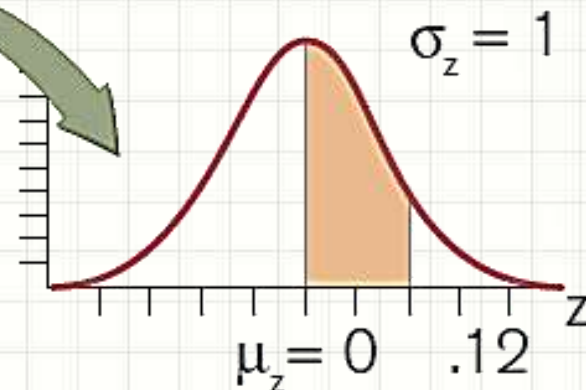


$$Z = \frac{6,2 - 5}{10} = 0,12$$

Distribuição normal

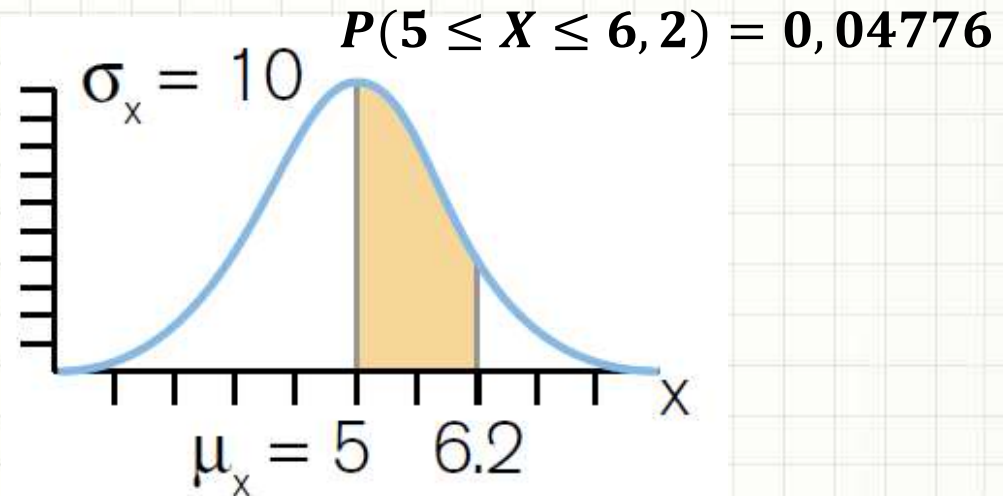


Distribuição normal padronizada



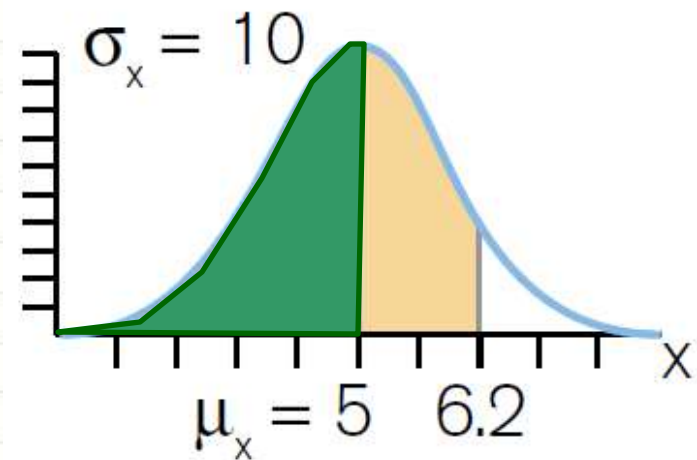
Determinação de Probabilidades

- Sabendo que



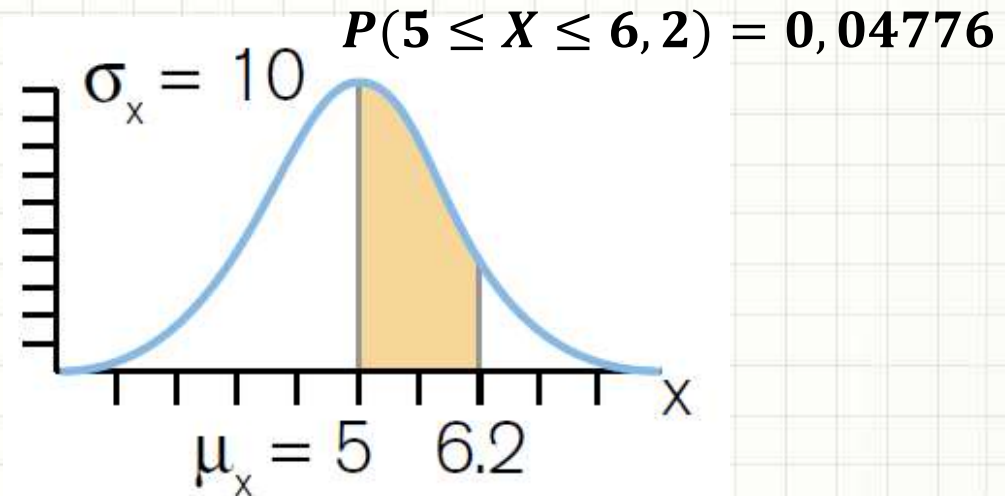
– Qual a $P(0 \leq X \leq 6,2)$?

$$P(0 \leq X \leq 5) = 0,5$$



Determinação de Probabilidades

- Sabendo que

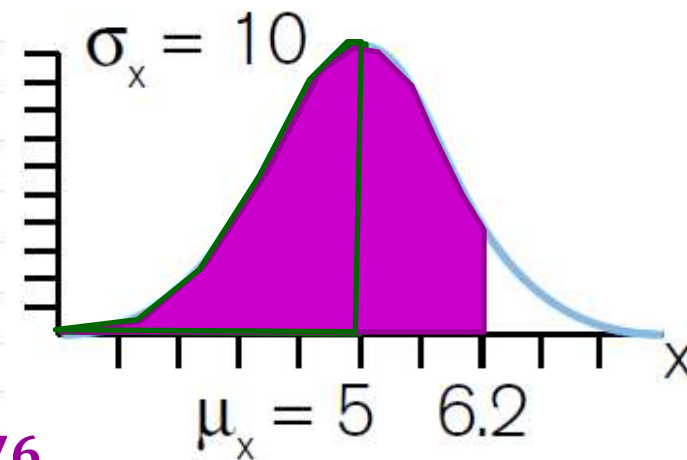


- Qual a $P(0 \leq X \leq 6,2)$?

$$P(0 \leq X \leq 5) = 0,5$$

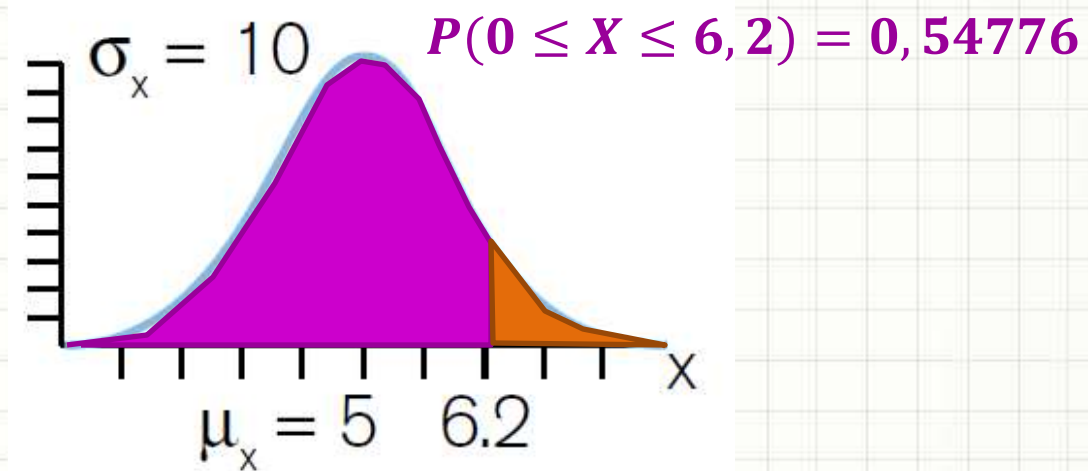
- Logo...

$$P(0 \leq X \leq 6,2) = 0,54776$$



Determinação de Probabilidades

- Sabendo que



– Qual a $P(X \geq 6,2)$?

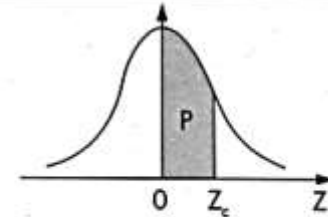
$$P(X \geq 6,2) = 1 - P(0 \leq X \leq 6,2) = 1 - 0,54776 = 0,45224$$

Tabela para Consulta (Parte 1)

Tabela III – Distribuição Normal Padrão

$$Z \sim N(0, 1)$$

Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c										parte inteira e primeira decimal de Z_c
	0	1	2	3	4	5	6	7	8	9	
	$p = 0$										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2



EXERCÍCIO EXEMPLO

Exercício Exemplo

- No controle de qualidade de uma fábrica de lâmpadas, verificou-se que a vida útil das lâmpadas é de, em média, 2000 horas, com um desvio padrão de 200 horas. Para planejamento, solicita-se determinar as seguintes probabilidades:
 - a) Uma lâmpada durar entre 2000 e 2400 horas
 - b) Uma lâmpada durar menos que 1470 horas

Exercício Exemplo

- $\mu=2000$, $\sigma=200$. Determinar probabilidade:
a) Uma lâmpada durar entre 2000 e 2400 horas
- Passo 1: padronizar os valores

$$Z = \frac{x - \mu}{\sigma} \quad Z = \frac{2000 - 2000}{200} = 0,00 \quad Z = \frac{2400 - 2000}{200} = 2,00$$

– Trata-se então do meio da curva até 2,00 desvios

- Passo 2: Uso da tabela

$$P(2000 \leq X \leq 2400) =$$

$$0,47725$$

$$\approx 47,7\%$$

Tabela III – Distribuição Normal Padrão
 $Z \sim N(0, 1)$
Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z)$

parte inteira e primeira decimal de Z	Segunda decimal de Z								
	0	1	2	3	4	5	6	7	8
	$p = 0$								
1,8	46407	46485	46562	46638	46712	46784	46856	46926	4699
1,9	47128	47193	47257	47320	47381	47441	47500	47558	4761
2,0	47725	47778	47831	47882	47932	47982	48030	48077	4812
2,1	48214	48257	48300	48341	48382	48422	48461	48500	4853
2,2	48610	48645	48679	48713	48745	48778	48809	48840	4888

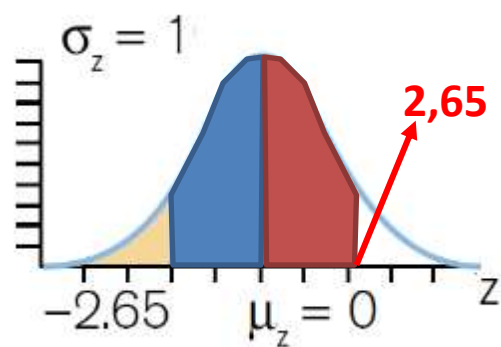
Exercício Exemplo

- $\mu=2000$, $\sigma=200$. Determinar probabilidade:
- b) Uma lâmpada durar menos que 1470 horas

- Passo 1: padronizar os valores

$$Z = \frac{x - \mu}{\sigma} \quad Z = \frac{1470 - 2000}{200} = -2,65$$

– Trata-se do início até -2,65 desvios



- Passo 2: Calculando a probabilidade

$$P(Z \leq -2,65) = 0,5 - P(\text{Azul})$$

$$P(\text{Azul}) = P(0 \leq Z \leq 2,65)$$

$$P(Z \leq -2,65) = 0,5 - 0,49598$$

$$P(Z \leq -2,65) = 0,00402 \approx 0,4\%$$

Tabela III – Distribuição Normal Padrão
 $Z \sim N(0, 1)$
 Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z_0)$

parte inteira e primeira decimal de Z_0	Segunda decimal de Z_0							
	0	1	2	3	4	5	6	7
$p = 0$								
2,5	49379	49396	49413	49430	49446	49461	49477	49492
2,6	49534	49547	49560	49573	49585	49598	49609	49621
2,7	49653	49664	49674	49683	49693	49702	49711	49720



CONCLUSÕES

Resumo

- Modelos Contínuos de Probabilidade
 - Função Densidade de Probabilidade
 - Distribuição Normal (ou Gaussiana)
 - Distribuição Normal Padronizada
 - Uso de tabelas para cálculo de probabilidades
-
- Distribuições de Amostras]
 - Teorema do Limite Central



PERGUNTAS?



EXERCÍCIOS

Exercícios

- 1. A duração de um componente eletrônico é normalmente distribuída com média de 850 dias e desvio padrão de 45 dias. Calcule a probabilidade do componente durar:
 - a) Entre 700 e 1000 dias

Exercícios

- 1. A duração de um componente eletrônico é normalmente distribuída com média de 850 dias e desvio padrão de 45 dias. Calcule a probabilidade do componente durar:
b) Mais que 800 dias

Exercícios

- 1. A duração média de 850 dias; desvio padrão de 45 dias.
b) Mais que 800 dias

$$Z = \frac{x - \mu}{\sigma} \quad Z = \frac{800 - 850}{45} = -1,11$$

$$P(Z \geq -1,11) = 0,5 + P(0 \leq Z \leq 1,11) = 0,5 + 0,36650$$

$$P(Z \geq -1,11) = 0,86650$$

≈ 86,7%

Tabela III – Distribuição Normal Padrão
 $Z \sim N(0, 1)$
Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z_p)$

parte inteira e primeira decimal de Z_p	Segunda decimal de Z_p							
	0	1	2	3	4	5	6	7
	p = 0							
1,0	34134	34375	34614	34850	35083	35314	35543	35769
1,1	36433	36650	36864	37076	37286	37493	37698	37900
1,2	38493	38686	38877	39065	39251	39435	39617	39796
1,3	40320	40490	40658	40824	40988	41149	41309	41466
1,4	41924	42073	42220	42364	42507	42647	42786	42922
1,5	43319	43448	43574	43699	43822	43943	44062	44179
1,6	44520	44630	44738	44845	44950	45053	45154	45254
1,7	45543	45637	45728	45818	45907	45994	46080	46164

Exercícios

- 1. A duração de um componente eletrônico é normalmente distribuída com média de 850 dias e desvio padrão de 45 dias. Calcule a probabilidade do componente durar:
c) Menos que 750 dias

Exercícios

- 1. A duração média de 850 dias; desvio padrão de 45 dias.
c) Menos que 750 dias

$$Z = \frac{x - \mu}{\sigma} \quad Z = \frac{750 - 850}{45} = -2,22$$

$$P(Z \leq -2,22) = 0,5 - P(0 \leq Z \leq 2,22) = 0,5 - 0,48679$$

$$P(Z \leq -2,22) = 0,01321$$

$$\approx \mathbf{1,3\%}$$

Tabela III – Distribuição Normal Padrão
 $Z \sim N(0, 1)$
Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z_p)$

parte inteira e primeira decimal de Z_p	Segunda decimal de Z_p							
	0	1	2	3	4	5	6	7
	p = 0							
2,0	47725	47778	47831	47882	47932	47982	48030	48077
2,1	48214	48257	48300	48341	48382	48422	48461	48500
2,2	48610	48645	48679	48713	48745	48778	48809	48840
2,3	48928	48956	48983	49010	49036	49061	49086	49111
2,4	49180	49202	49224	49245	49266	49286	49305	49324
2,5	49379	49396	49413	49430	49446	49461	49477	49492
2,6	49534	49547	49560	49573	49585	49598	49609	49621
2,7	49653	49664	49674	49683	49693	49702	49711	49720

Exercícios

- 1. A duração de um componente eletrônico é normalmente distribuída com média de 850 dias e desvio padrão de 45 dias. Calcule a probabilidade do componente durar:
d) Exatamente 1000 dias.

Exercícios

- 1. A duração média de 850 dias; desvio padrão de 45 dias.
d) Exatamente 1000 dias.

**A probabilidade é
considerada praticamente
ZERO!**

(Afinal, a área será zero!)