DCA- 0131: Ciência de Dados

Modelos Estatísticos

Luiz Affonso Guedes - affonso@dca.ufrn.br

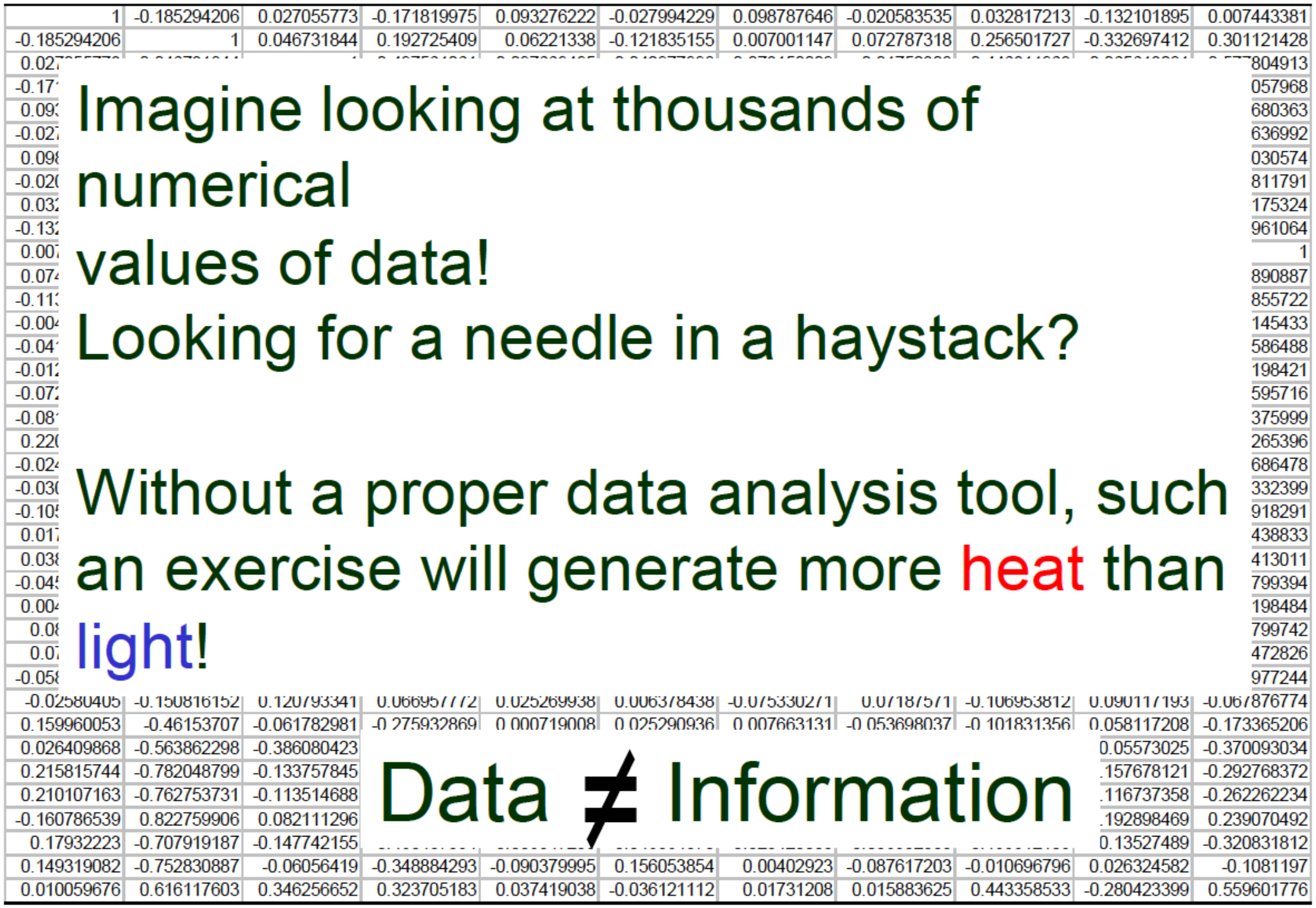
Questões Fundamentais

- Representação
  - Como modelar incertezas em uma forma plausível?
  - Como modelar/codificar o domínio do conhecimento, suposições e restrições?
  - Como fazer inferências no modelo?

Por quê Modelos Estatísticos e Probabilísticos?

- São em geral muito mais compactos (memória)
- São em geral muito mais eficientes (tempo)
- São mais fáceis de aprender a partir de dados ou construir modelos a partir de especialistas.
- Possuem uma base teórica bem estabelecida.

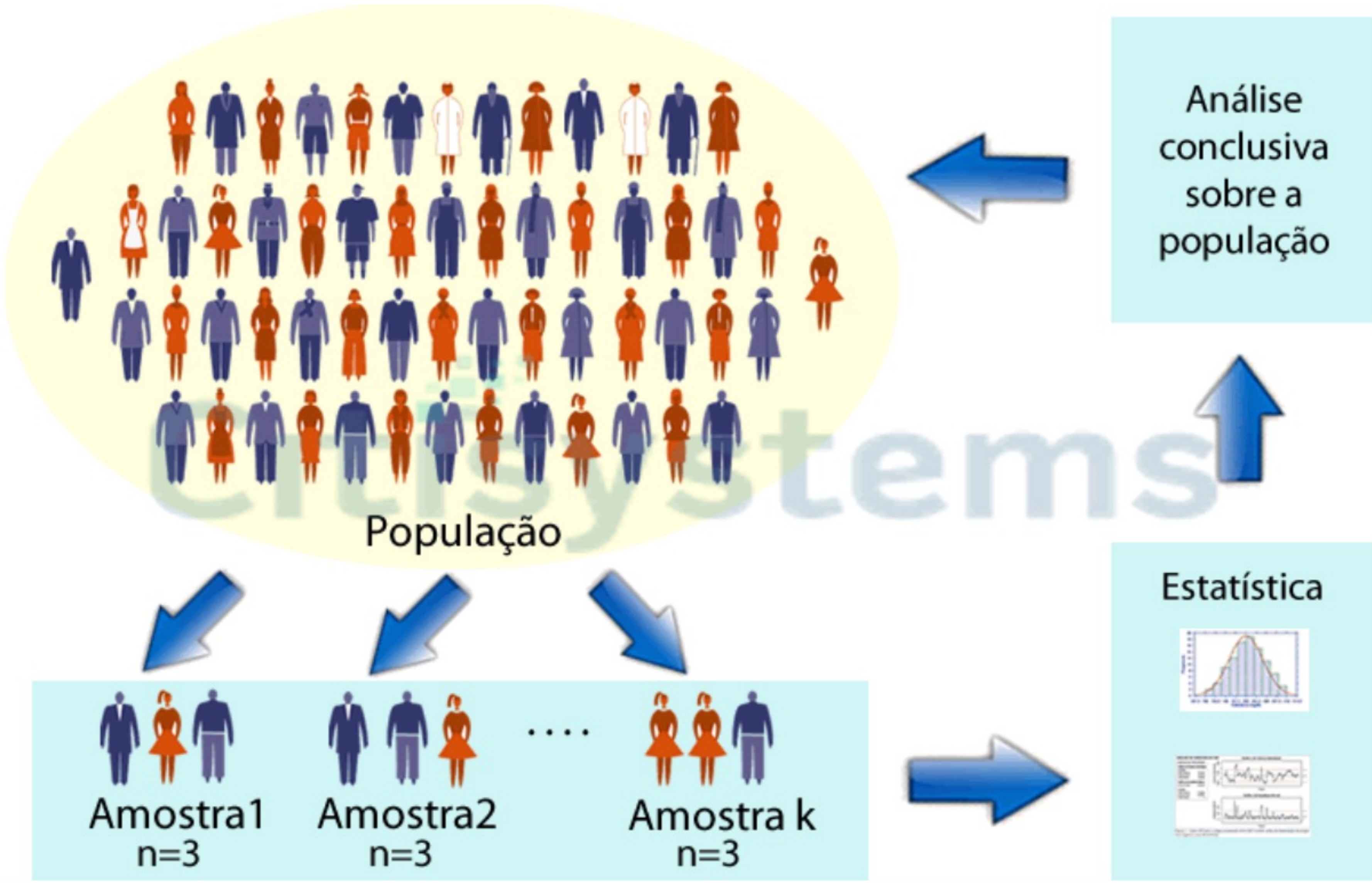
Modelagem a partir de dados



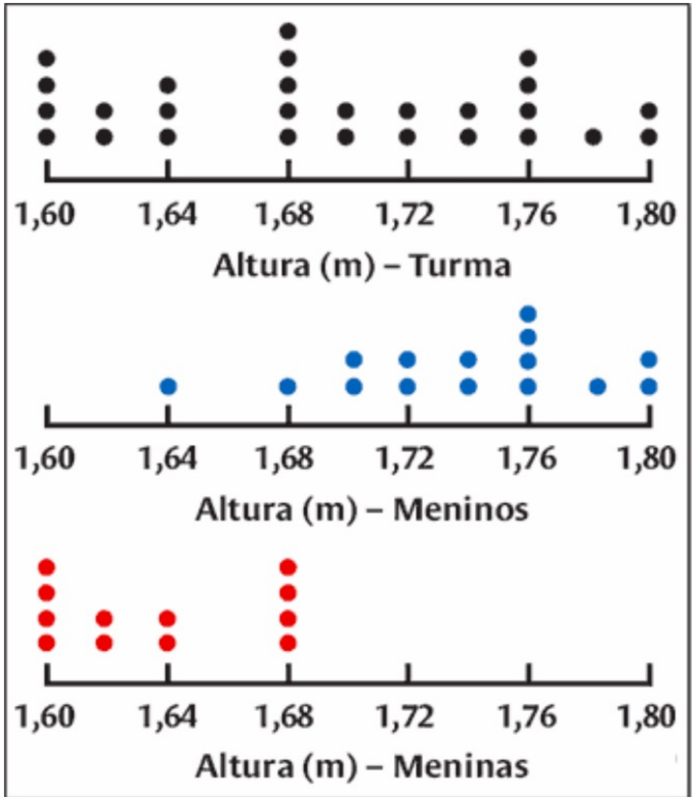
Modelo Estatístico

- Medidas estatísticas
  - Média, mediana, moda
  - Dispersão
  - Quartis
  - Simetria
  - Achatamento
  - Intervalo de confiança
  - p-valor
  - Outliers

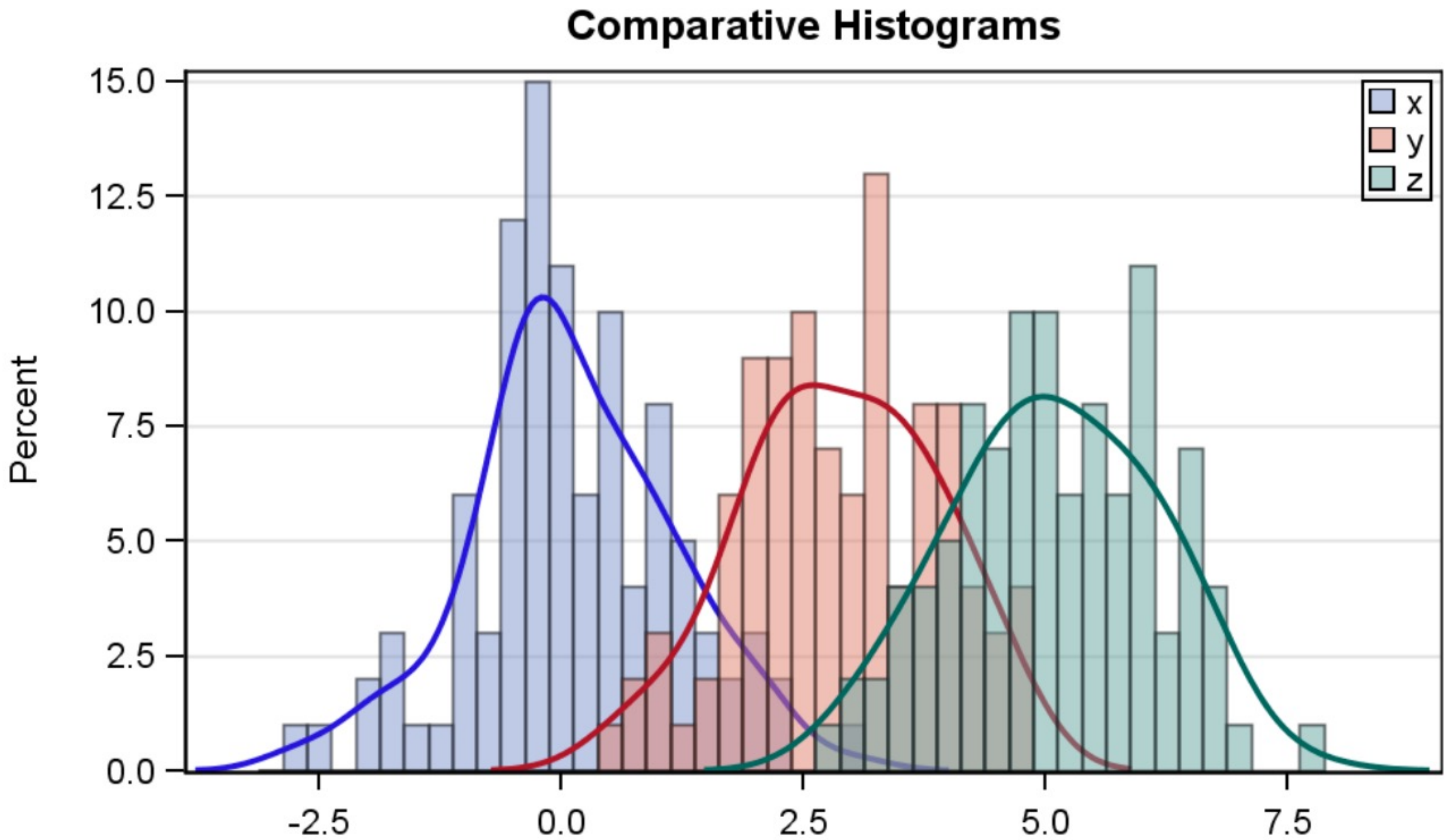
Modelagem Orientada a Dados



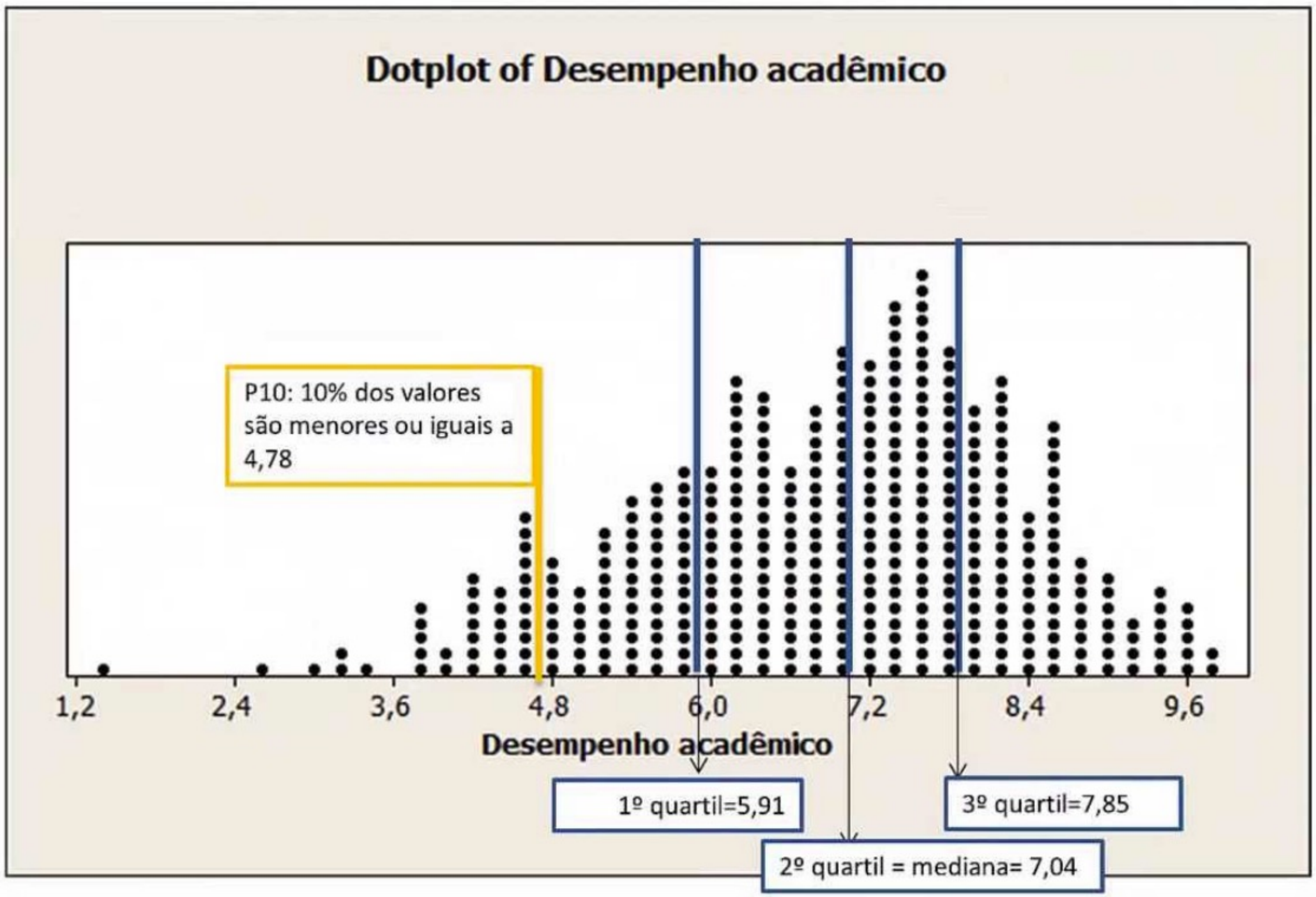
Modelo Baseado em Dados



Gráficos Estatísticos - Histogramas



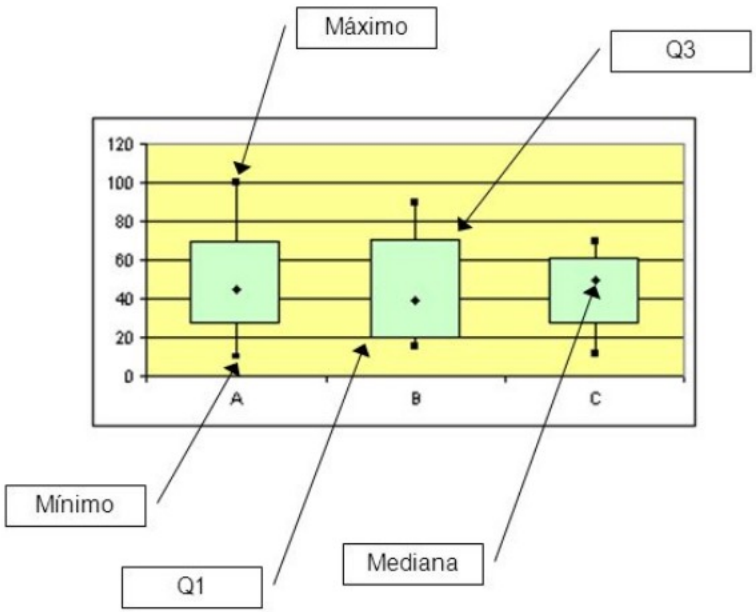
Medidas Estatísticas

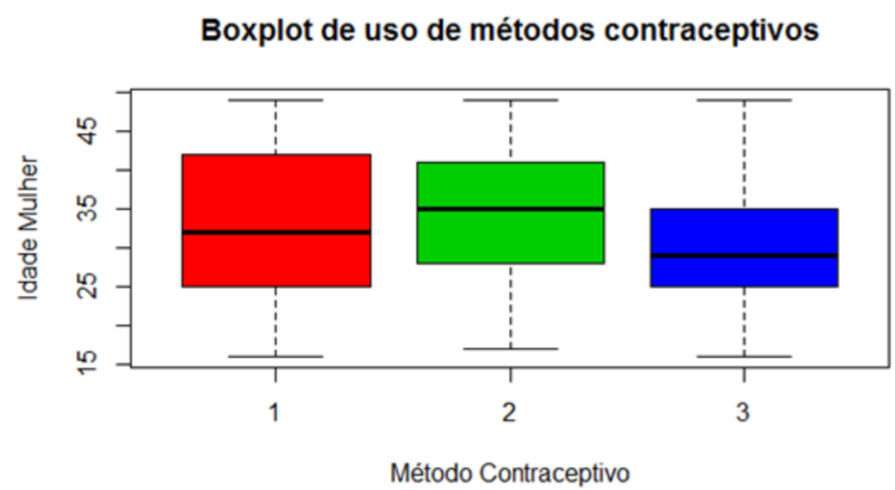


Medidas Estatísticas

Gráficos – BoxPlot

	Experimento		
Coleta	A	B	C
1	10	15	70
2	30	18	52
3	100	75	47
4	25	78	21
5	35	90	55
6	75	43	66
7	70	56	63
8	70	36	42
9	24	25	18
10	55	18	11
Estatística	A	B	C
Mediana	45	40	50
Q1	26	20	26
Mínimo	10	15	11
Máximo	100	90	70
Q3	70	70	61



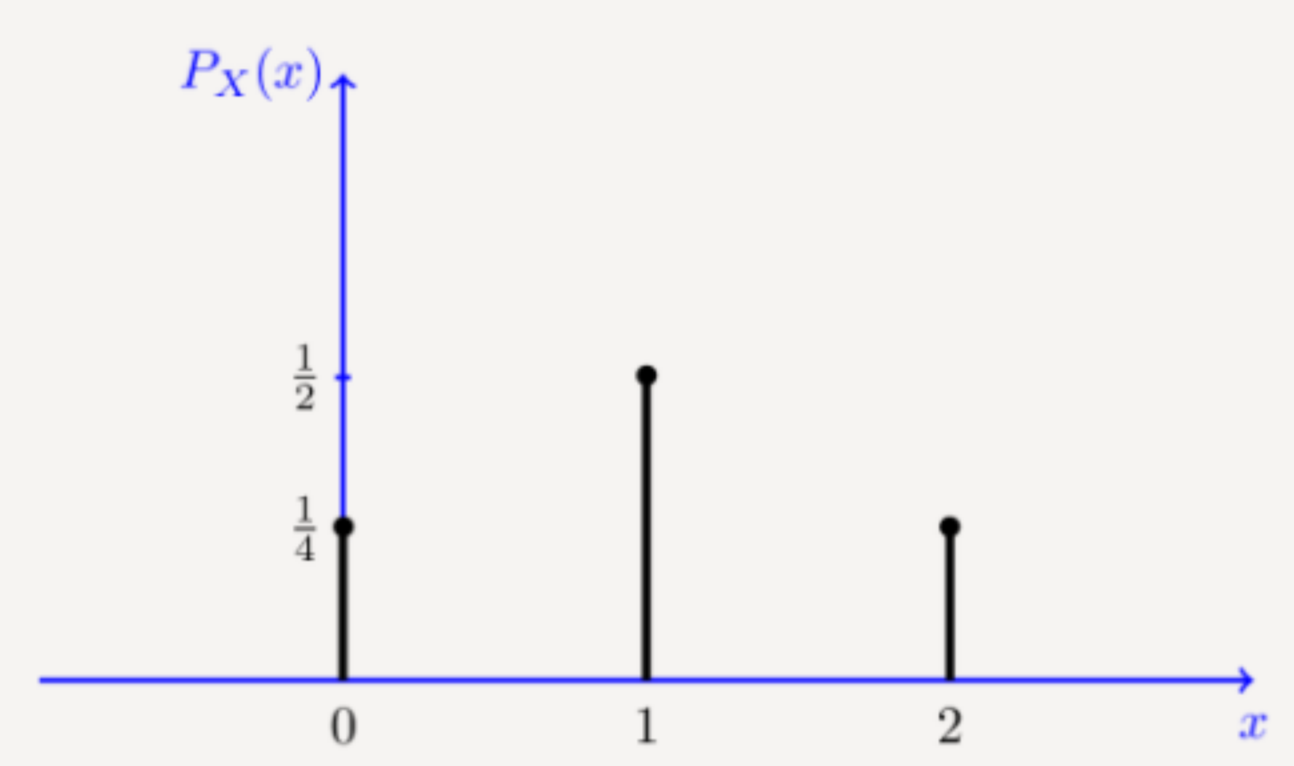


Funções de Probabilidade de V.A. Discretas

- Seja a v.a. discreta $X = \{x_1, x_2, ...x_k\}$
  - A função de probabilidade de massa de x_i é:
    - $P_i(x_i) = P(X=x_i)$
- Exemplo: Jogar duas moedas, considerando que cara = 0 e coroa = 1. A v.a. X é a soma dos dois resultados.

Moeda 1	Moeda 2	X
0	0	0
0	1	1
1	0	1
1	1	2

X	$P_x(X)$
0	1/4
1	1/2
2	1/4



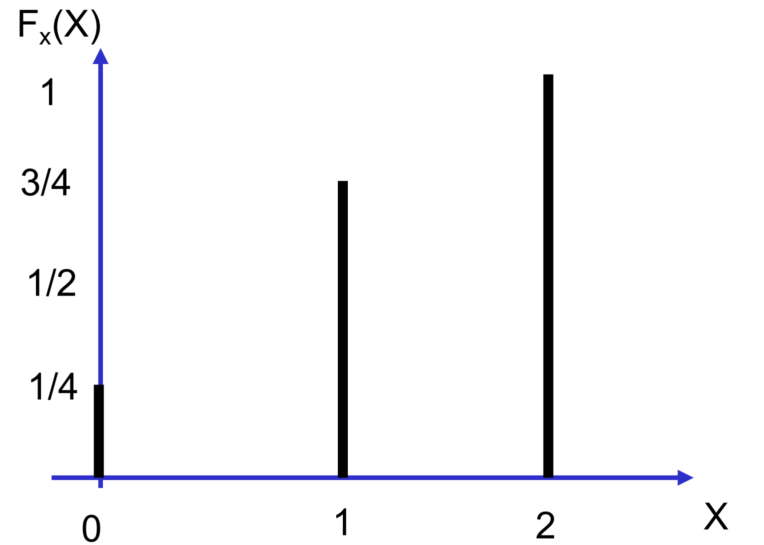
Funções de Probabilidade de V.A. Discretas

- **Função Acumulativa:** é a função acumulativa de uma v.a.:
  - $F_x(X) = P\{X \leq x\} = \sum_{i=1}^{k} P(X = x_i)$, para todo $X \leq x_i$

- **Exemplo:** Jogar duas moedas, considerando que cara = 0 e coroa = 1. A v.a. X é a soma dos dois resultados.

Moeda 1	Moeda 2	X
0	0	0
0	1	1
1	0	1
1	1	2

X	$P_x(X)$	$F_x(X)$
0	1/4	1/4
1	1/2	3/4
2	1/4	1

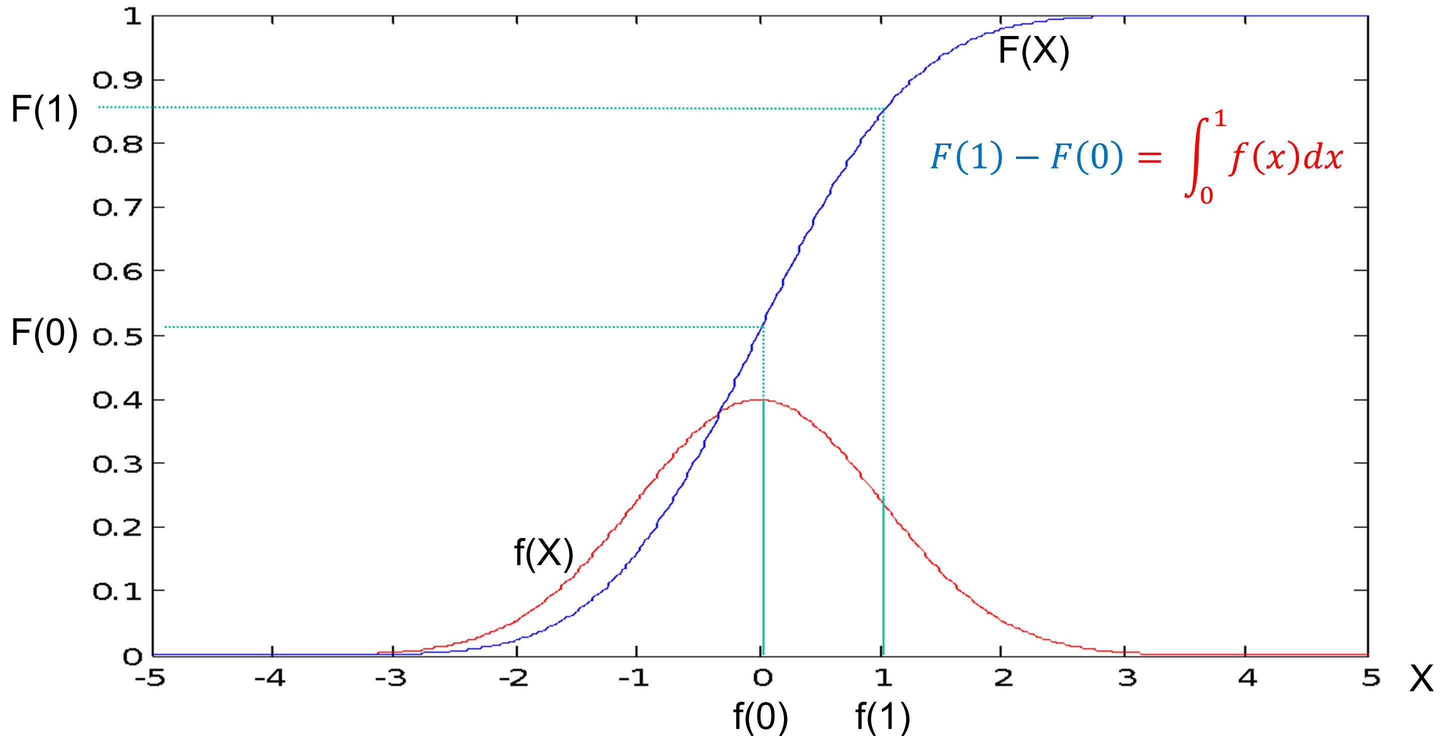


Funções de Probabilidade de V.A. Contínua

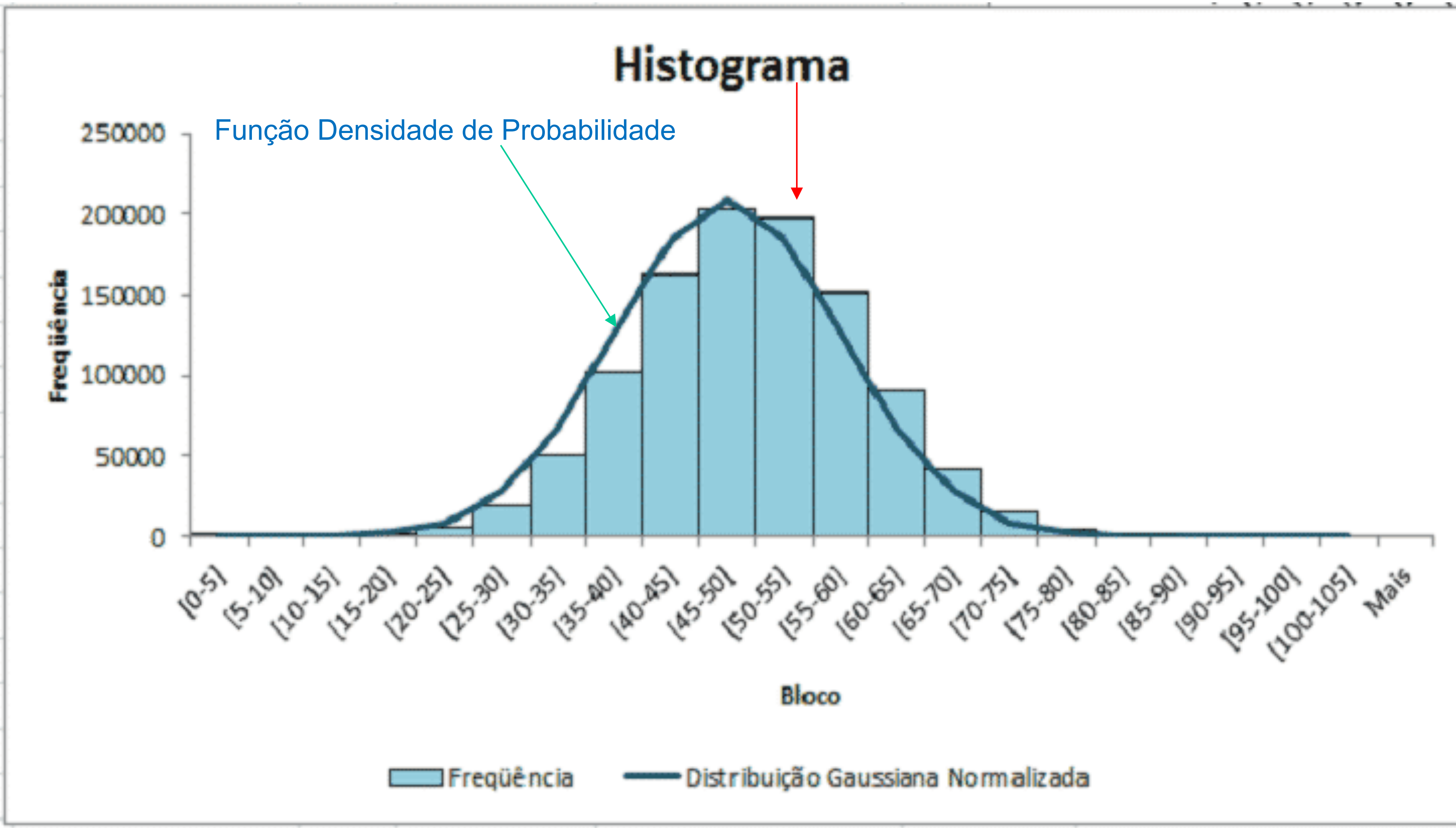
- **Distribuição:** é a função acumulativa de uma v.a.:
  - $F_x(X) = P\{X \leq x\}$, definida para cada x no intervalo de -infinito a +infinito.
- **Densidade:** é a derivada da função distribuição.
  - $f(X) = dF(X)/dX$

Funções Distribuição e Densidade

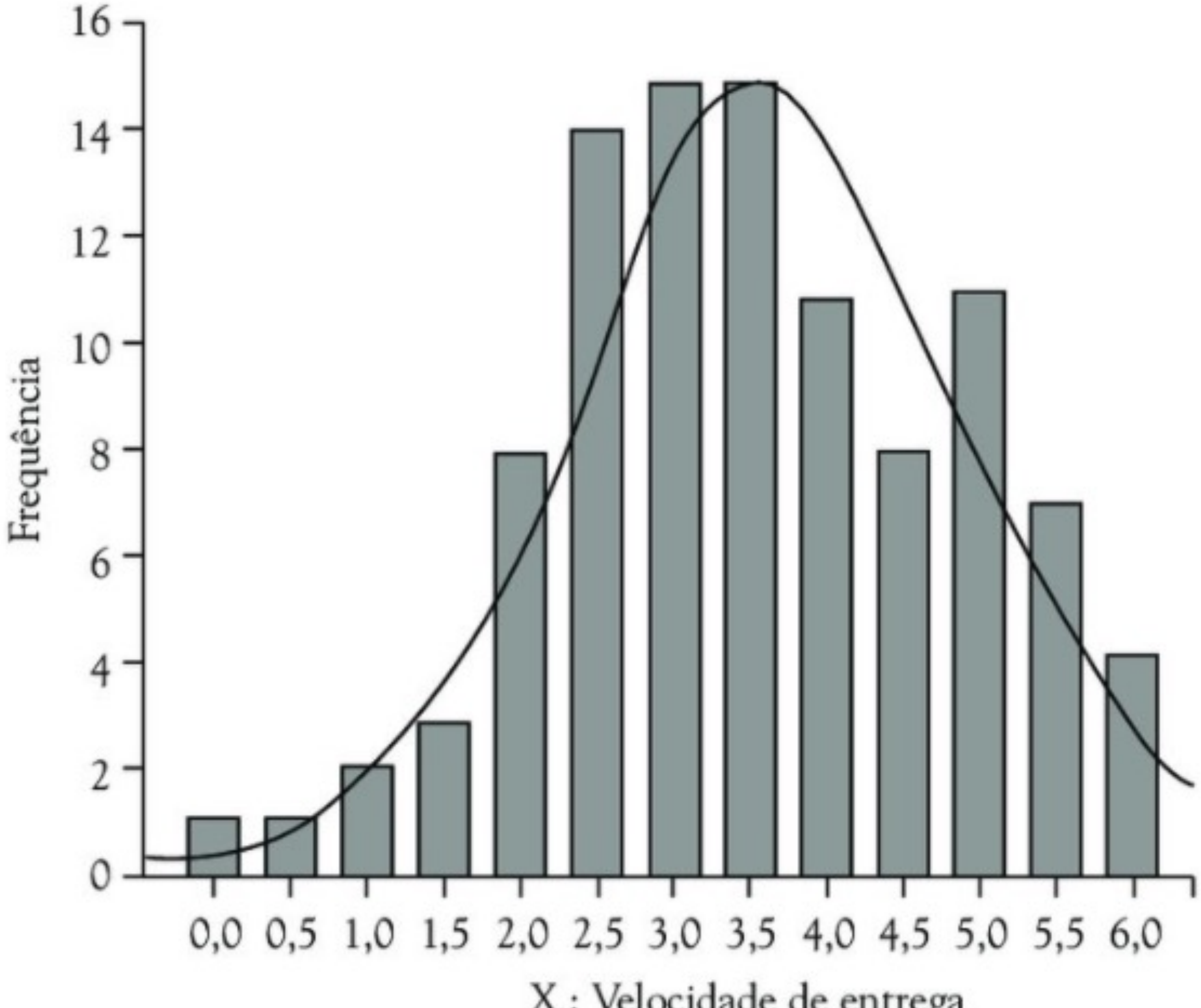
- Propriedades:
  - F(X) é uma função não decrescente;
  - F(-00) = 0 ; F(+00) = 1;
  - P(a<x ≤ b) = F(b) – F(a);
  - f(x) ≥ 0 ; para todo x.



Modelagem Estatística



Modelagem Estatística - Exemplo



Modelos Probabilísticos

Discretos

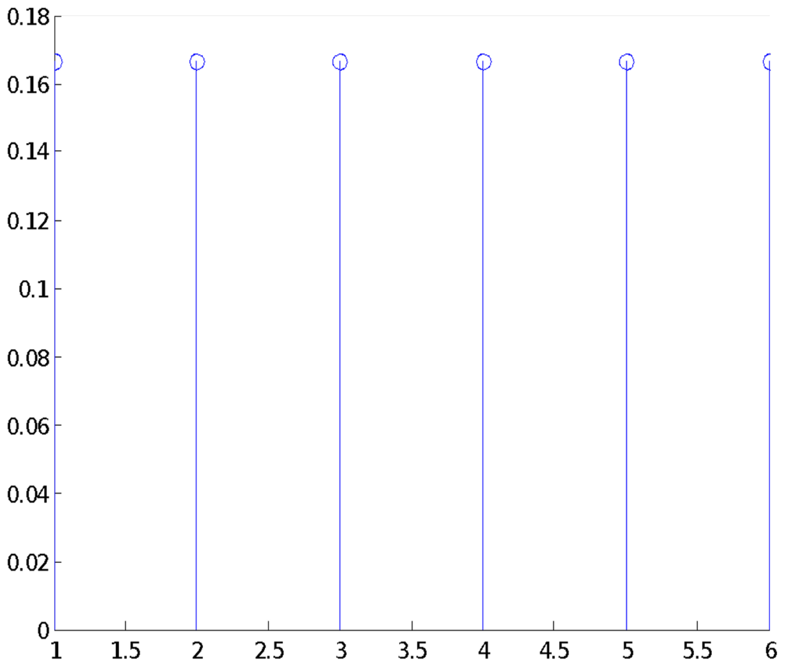
- Uniforme
- Bernoulli
- Binomial
- Hipergeométrico
- Geométrico
- Poisson

Contínuos

- Uniforme
- Exponencial
- Gaussiana – Normal
- Log-Normal
- Gama
- Weibull
- ...

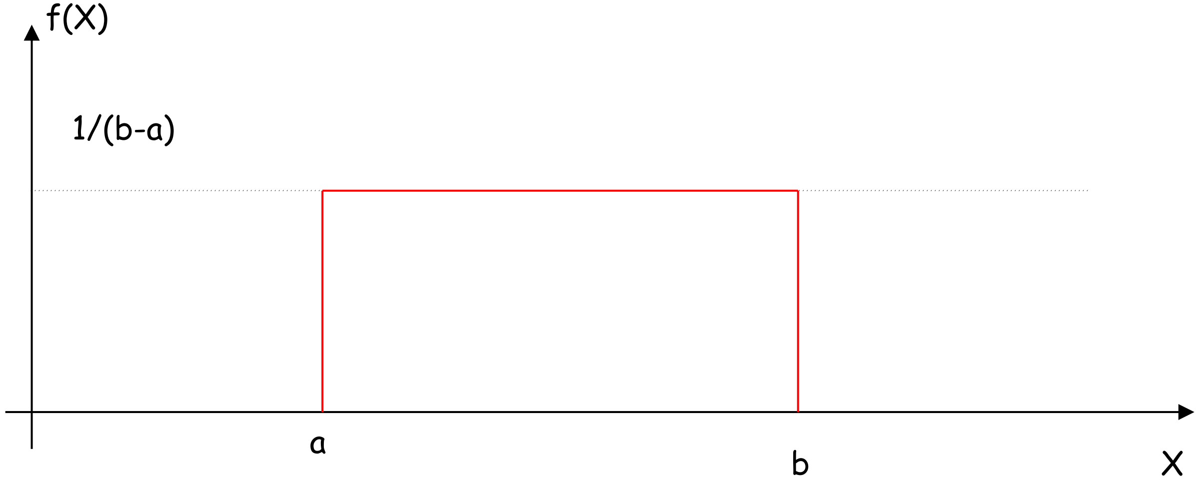
Modelos Discretos

- ## Modelo Uniforme Discreto
  - Seja a V.A. X, cujos possíveis valores são representados por $x_1, x_2, ..., x_k$.
  - Dizemos que X segue o modelo Uniforme Discreto se atribui a mesma probabilidade 1/k para cada um desses k valores.
  - $P(X=x_i) = 1/k, \forall i = 1,2, ..., k.$
  - $E[X] = (1+k)/2$
  - $\sigma^2[X] = (k^2-1)/12$



Modelos Contínuos

- Modelo Uniforme Contínuo – X~U[a,b]
  - A v.a. X tem distribuição uniforme contínua no intervalo [a,b], a<b, se sua função densidade de probabilidade é dada por:
    - $f(x) = 1/(b-a)$, $a \leq x \leq b$
      $= 0$, caso contrário.

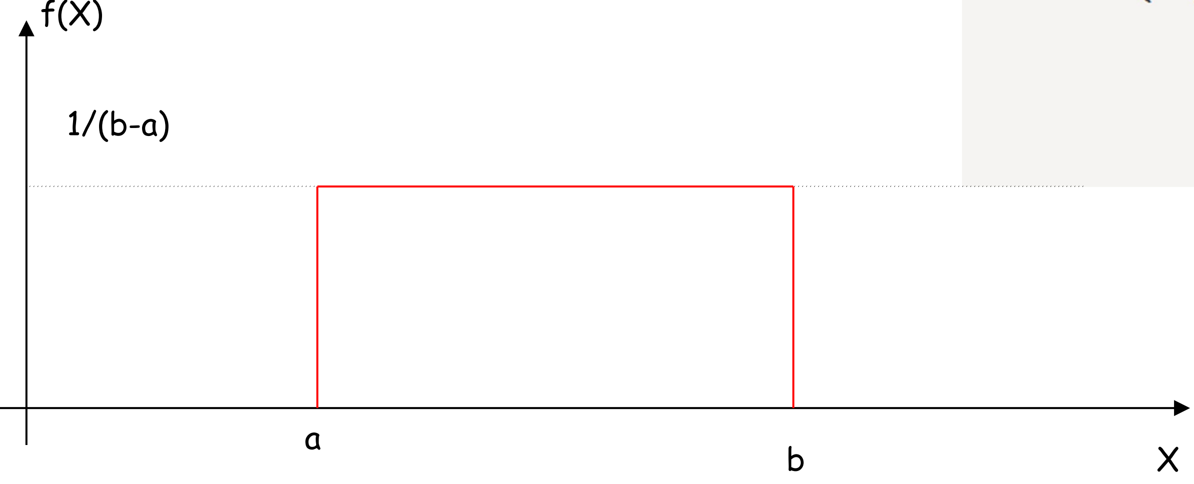


Modelos Contínuos

- Modelo Uniforme Contínuo - X~U[a,b]
  - $E[X] = (a+b) / 2$
  - $E[X^2] = (a^2+ab+b^2)/3$
  - $\sigma^2[X] = (b-a)^2/12$
  - $F(X) = ?$

$$EX^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx$$
$$= \int_a^b x^2 \left(\frac{1}{b-a}\right) dx$$
$$= \frac{a^2+ab+b^2}{3}.$$

$$\mathrm{Var}(X) = EX^2 - (EX)^2$$
$$= \frac{(b-a)^2}{12}.$$



Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - A v.a. contínua X tem distribuição Normal, com parâmetros μ, σ^2, se sua função densidade é dada por:

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

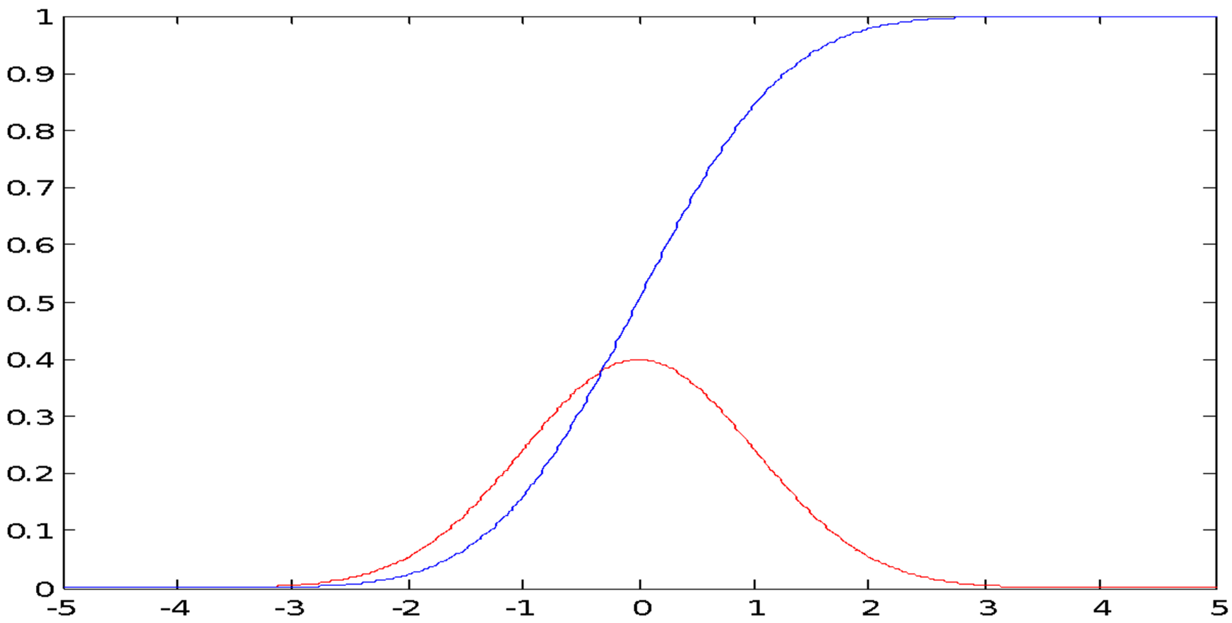
  - Propriedades:
    - $f(x)$ é simétrica em relação à μ;
    - $\int f(x) = 1$, para $-\infty < x < +\infty$;
    - O valor máximo de $f(x)$ se dá para $x = \mu$.
    - $E[X] = \mu$
    - Variância = $E[(X - \mu)^2] = \sigma^2$
    - https://www.youtube.com/watch?v=4HpvBZnHOVI

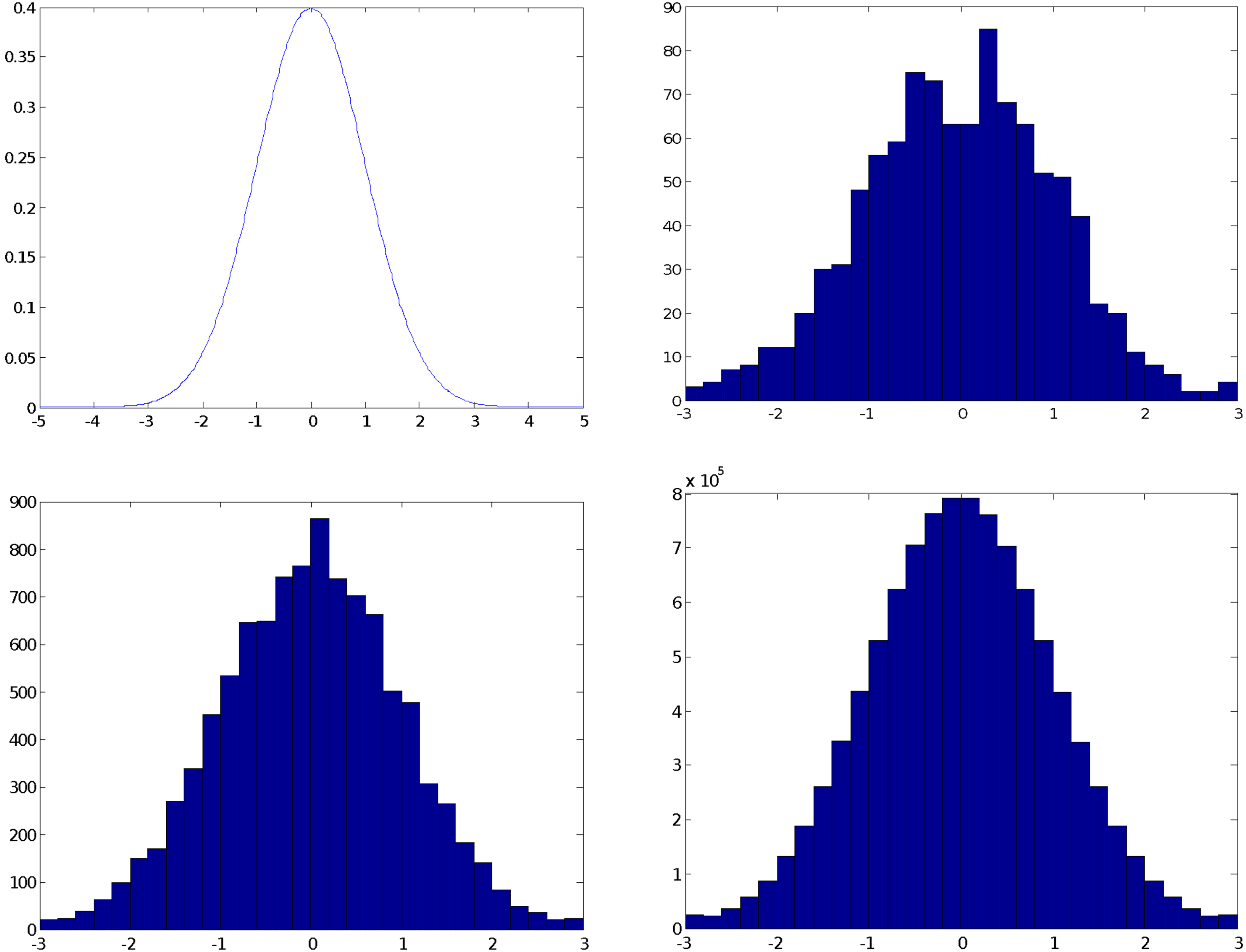
Modelos Contínuos

- Modelo Normal ou Gaussiana - $X \sim N(\mu, \sigma^2)$

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

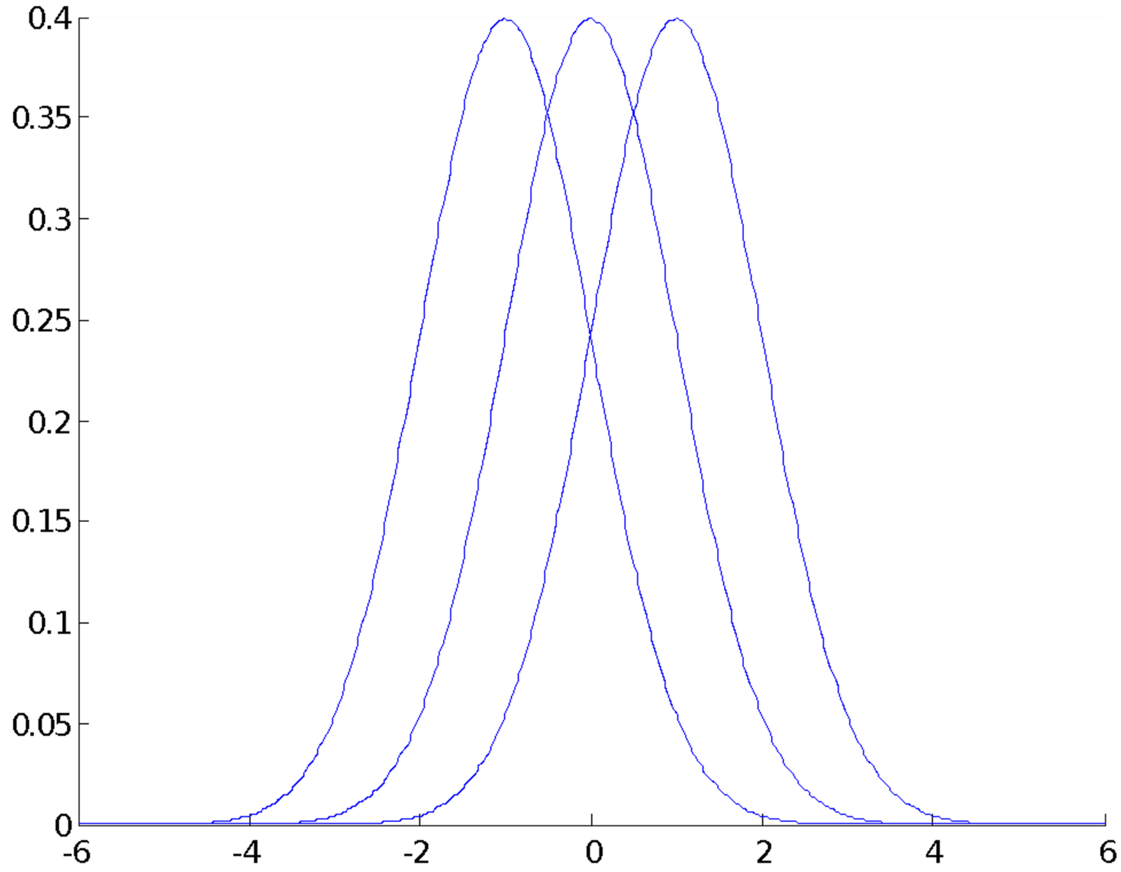
  - $E[X] = \mu$
  - Variância = $E[(X - \mu)^2] = \sigma^2$




0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5 -4 -3 -2 -1 0 1 2 3 4 5
90
80
70
60
50
40
30
20
10
0
-3 -2 -1 0 1 2 3
900
800
700
600
500
400
300
200
100
0
-3 -2 -1 0 1 2 3
x 10^5
8
7
6
5
4
3
2
1
0
-3 -2 -1 0 1 2 3

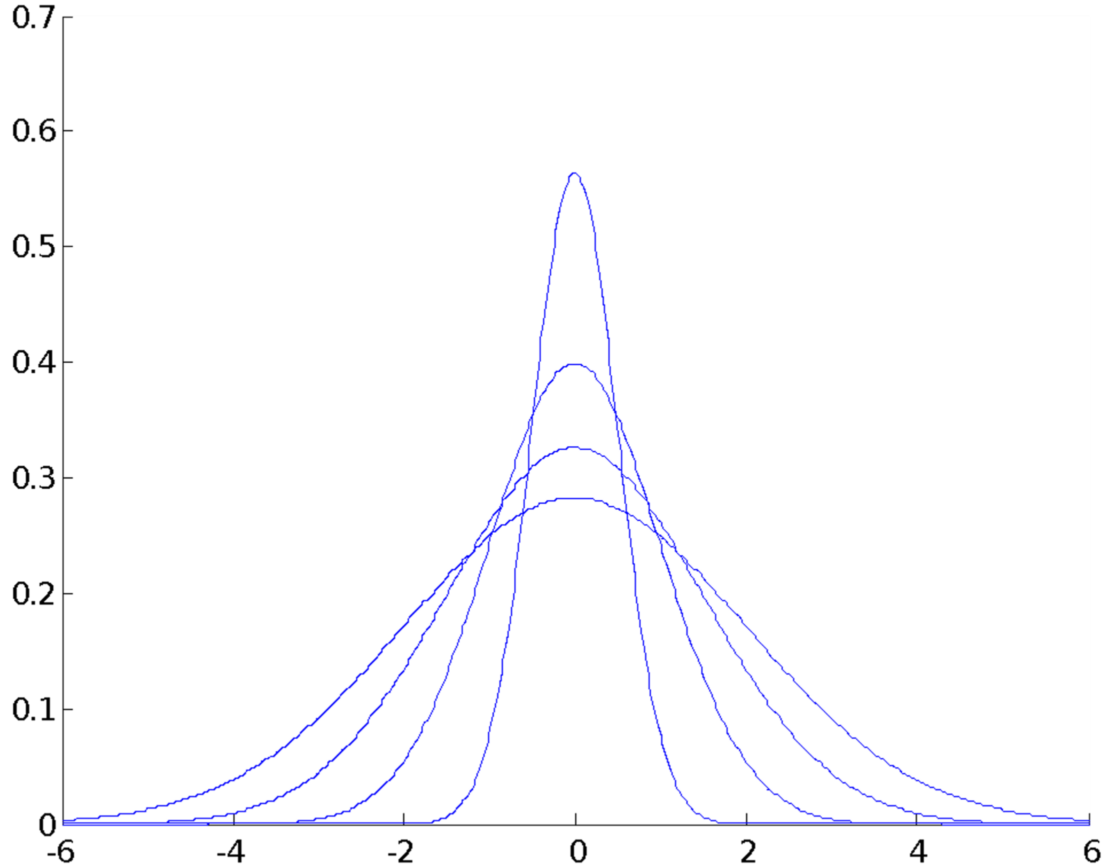
Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$



Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$



Modelos Contínuos

- Modelo Normal ou Gaussiano – $X \sim N(\mu, \sigma^2)$
  - Exercício de Aplicação

Um matemático belga do século XIX pôs na cabeça a idéia de descrever o "homem médio" e, por conta disso, mediu muitas e muitas variáveis[1]. A Tabela 10.1 mostra a distribuição do perímetro torácico[2] que esse matemático mediu em nada menos do que 5.732 soldados escoceses. As medidas estão em polegadas. Como uma polegada vale 2,54 cm, você vê que as medidas variaram entre 83,82 cm e 121,92 cm.[3] Veja o histograma apresentado na Figura 10.1.

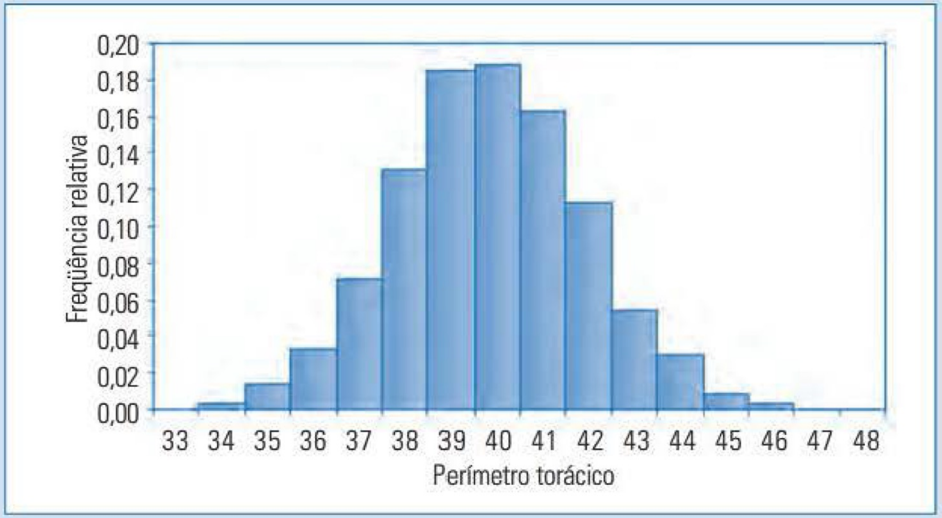


FIGURA 10.1 Histograma para a distribuição de freqüências do perímetro torácico de homens adultos, em polegadas.

Modelos Contínuos

- Modelo Normal ou Gaussiano – $X \sim N(\mu, \sigma^2)$
  - Exercício de Aplicação

Um teste de inteligência[4] foi idealizado pressupondo que quociente de inteligência tem distribuição normal de média $\mu = 100$ e desvio padrão $\sigma = 15$. Veja a Figura 10.3 e note que, de acordo com esse teste:

- As pessoas têm, em média, QI igual a 100.
- Metade das pessoas tem QI igual ou maior do que 100 e metade tem QI igual ou menor do que 100.
- Pessoas com QI muito alto (na cauda à direita da curva) são raras, como também são raras pessoas com QI muito baixo (na cauda à esquerda da curva).

[4] Existem muitas maneiras de "medir" inteligência (embora nenhuma delas explique, exatamente, o que está sendo medido). Mas um dos testes (Weschler) foi idealizado pressupondo que inteligência tem distribuição normal, como mostrado no exemplo. In: MOTULSKY, H. **Intuitive Biostatistics.** Nova York, Oxford Press, 1995. p.38.

- Modelo Normal ou Gaussiano – $X \sim N(\mu, \sigma^2)$
  - Qual é a vantagem de se saber se um problema pode ser modelado por uma distribuição Gaussiana?

- Prova-se, teoricamente[5], que se a variável tem distribuição normal, 34,13% da área sob a curva estão entre a média (μ) e um ponto de abscissa igual à média mais um desvio padrão ($\mu + \sigma$).
- A curva é simétrica em torno da média. Segue-se daí que 34,13% da área sob a curva está entre a média (μ) e um ponto de abscissa igual à média menos um desvio padrão ($\mu - \sigma$).
- Se você somar as porcentagens, terá 68,26%. Então, entre ($\mu - \sigma$) e ($\mu + \sigma$) estão 68,26% da área da curva, como mostra a Figura 10.3.

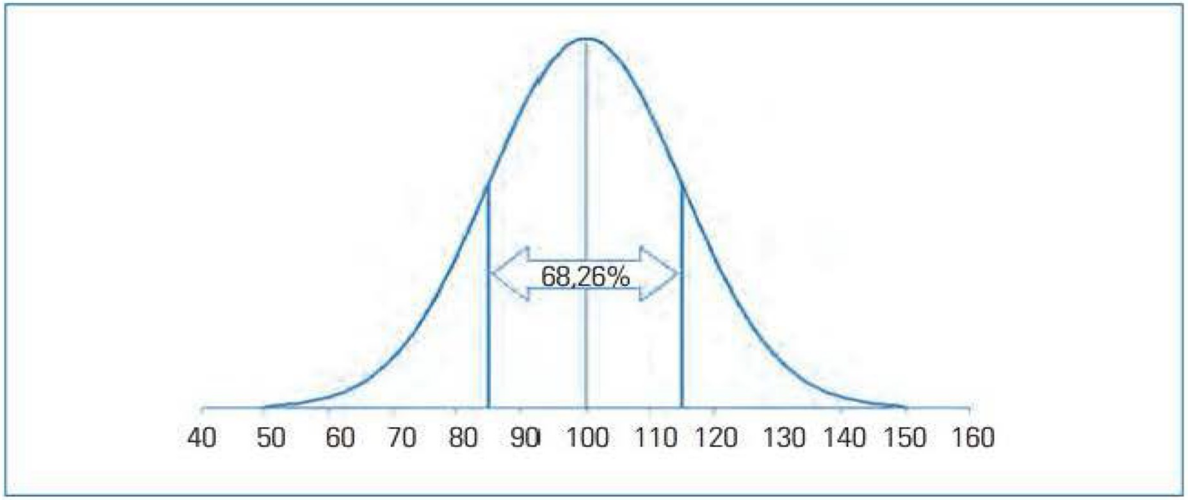


FIGURA 10.3 Distribuição normal: 68,26% dos casos estão entre a média ± 1 desvio padrão.

Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$

- Para o exemplo do teste de QI:
  - Média = 100, Desvio Padrão = 15

Reveja o Exemplo 10.2. Pressupondo que quociente de inteligência tem distribuição normal de média $\mu = 100$ e desvio padrão $\sigma = 15$, então:

- 34,13% das pessoas, segundo o teste, têm quociente de inteligência entre $\mu = 100$ e $\mu + \sigma = 100 + 15 = 115$, ou seja, entre 100 e 115.
- 34,13% das pessoas, segundo o teste, têm quociente de inteligência entre $\mu = 100$ e $\mu - \sigma = 100 - 15 = 85$, ou seja, entre 100 e 85
- 68,26% das pessoas, segundo o teste, têm quociente de inteligência entre 85 e 115.

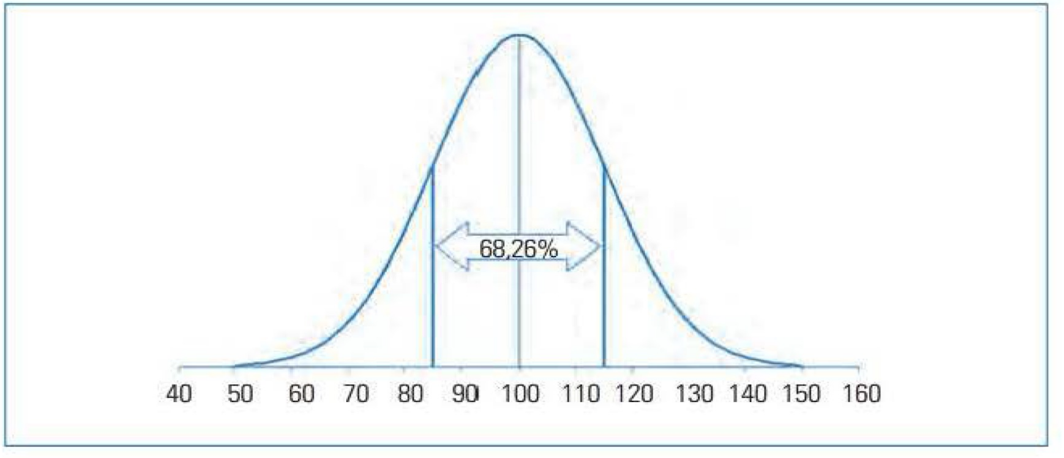
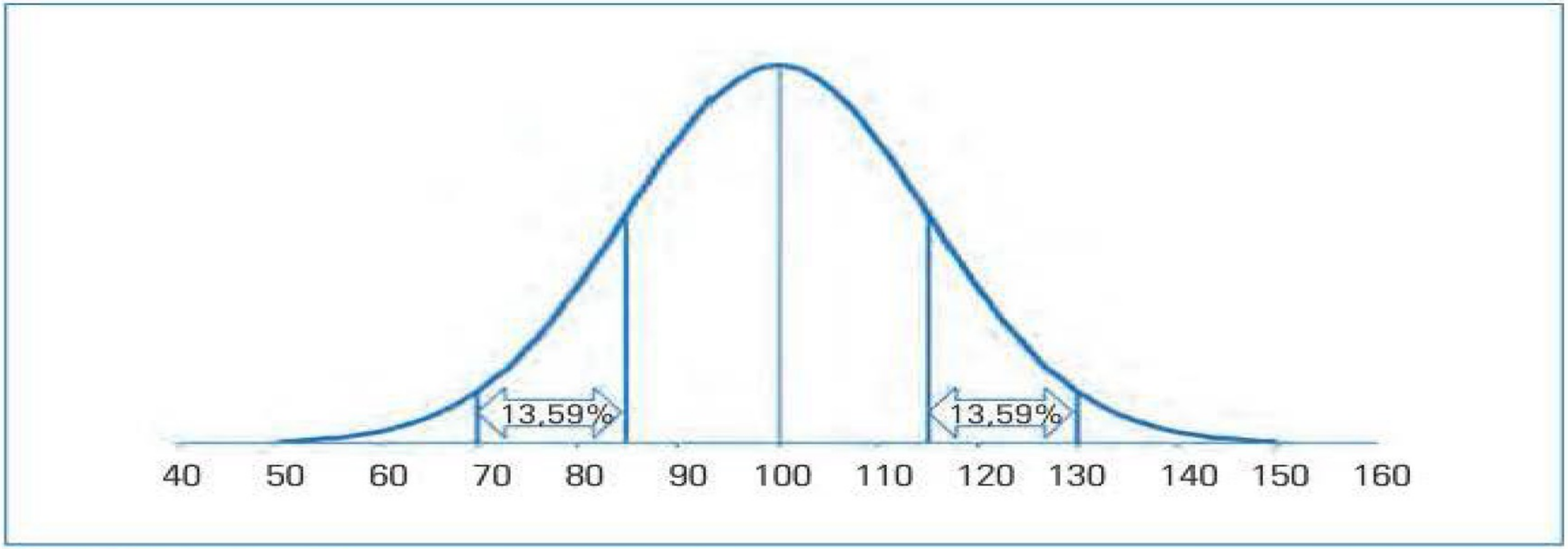


FIGURA 10.3 Distribuição normal: 68,26% dos casos estão entre a média ± 1 desvio padrão.

Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$

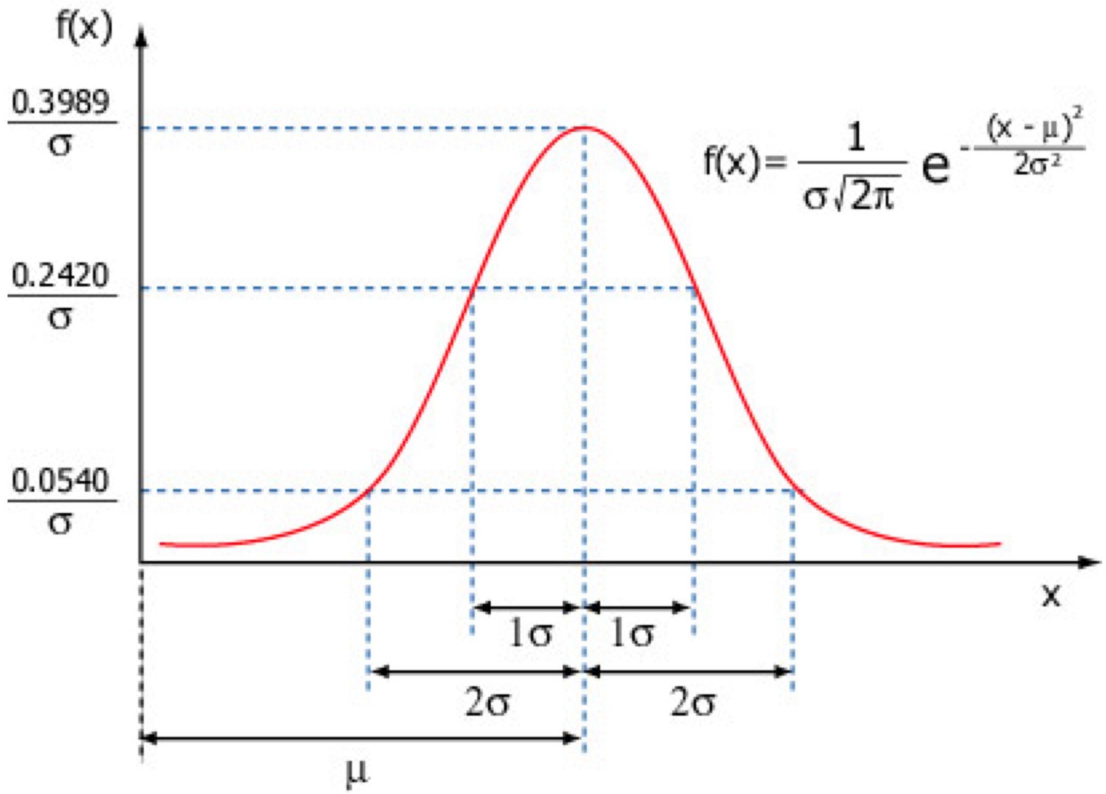
- Para o exemplo do teste de QI:
  - Média = 100, Desvio Padrão = 15

- 13,59% da área sob a curva estão entre a média mais um desvio padrão ($\mu + \sigma$) e um ponto de abscissa igual à média mais dois desvios padrões ($\mu + 2\sigma$).
- A curva é simétrica em torno da média. Segue-se daí que 13,59% da área sob a curva estão entre a média menos um desvio padrão ($\mu - \sigma$) e um ponto de abscissa igual à média menos dois desvios padrões ($\mu - 2\sigma$). Veja a Figura 10.4.



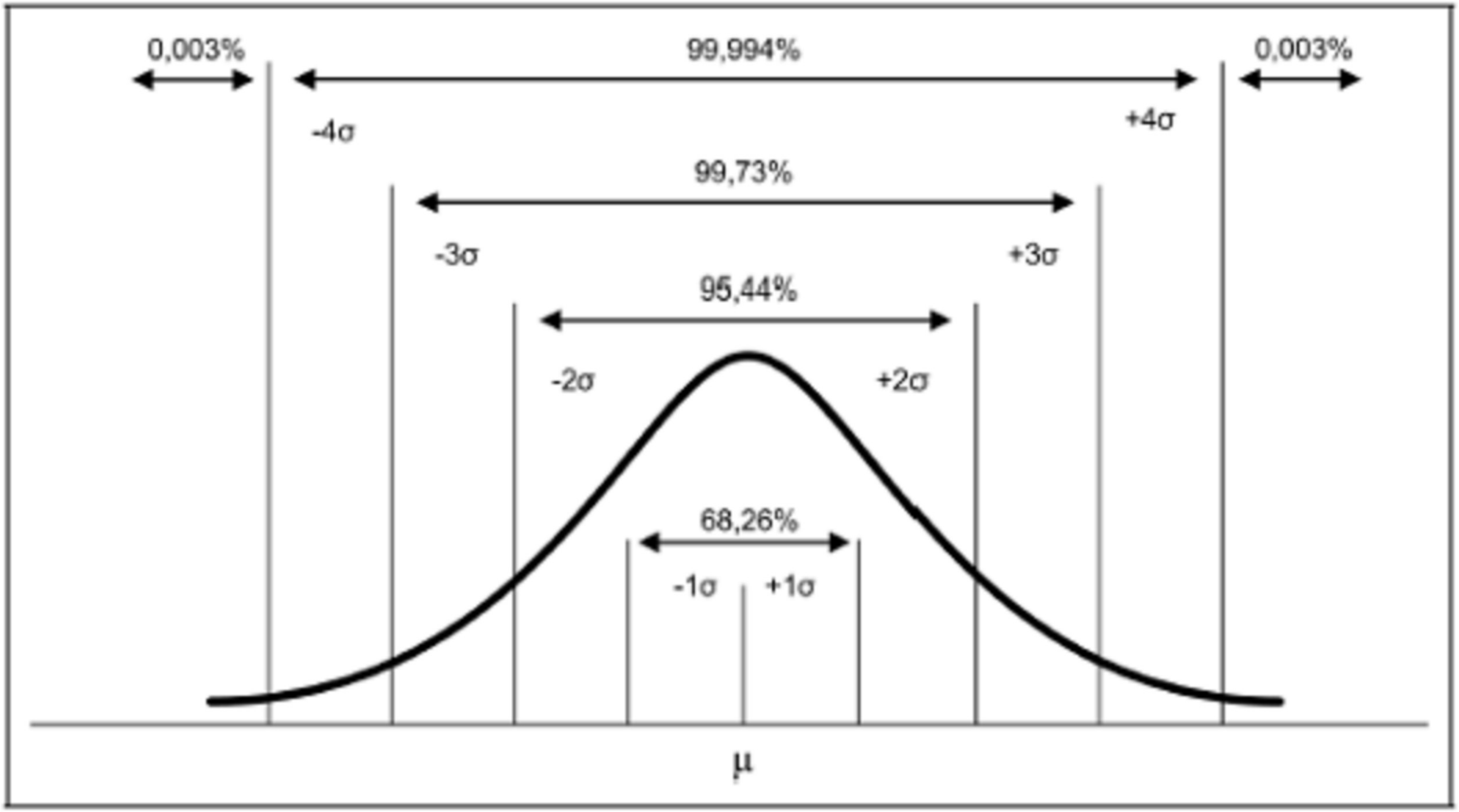
Modelos Contínuos

- Modelo Normal ou Gaussiano – $X \sim N(\mu, \sigma^2)$



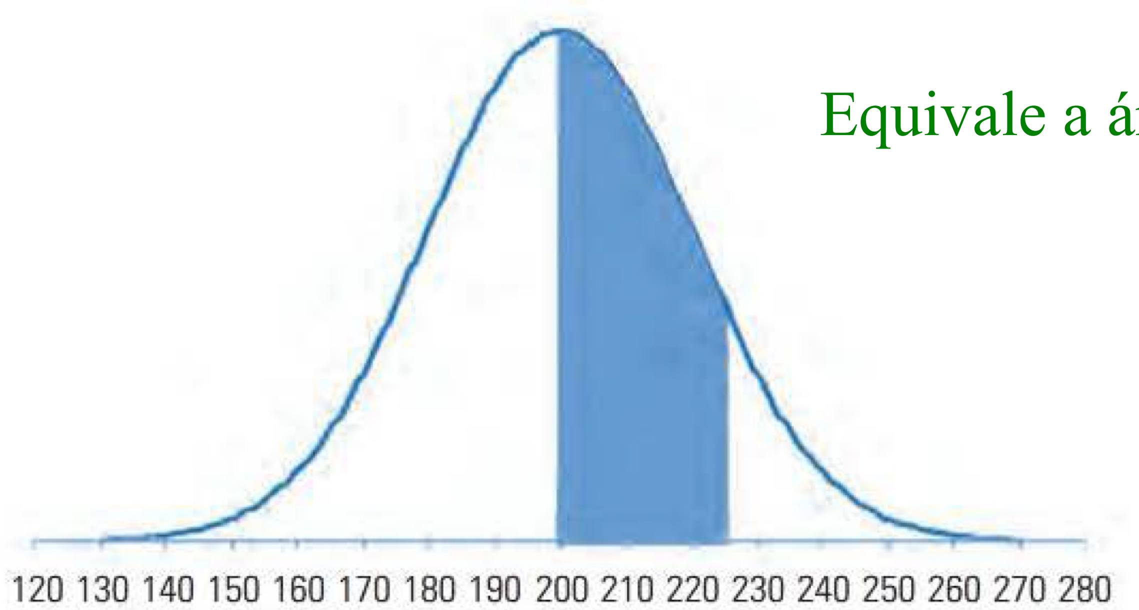
Modelos Contínuos

- Modelo Normal ou Gaussiano – $X \sim N(\mu, \sigma^2)$



Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - $f(x) = \{1/(\sigma \sqrt{2\pi})\} \,.\, e^{-(x-\mu)\,(x-\mu)\,/(2\sigma\sigma)}$, para $-\infty < x < +\infty$
  - $E[X] = \mu$
  - Variância = $E[(X - \mu)^2] = \sigma^2$
  - $F(X) = \int_{-\infty,x} f(X)dx \rightarrow$ não há solução analítica
  - $P(a<x<b) = F(b) - F(a) = \int_{a,b} f(X)dx$ Como se obter isto?



Equivale a área de f(x), entre a e b

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - $f(x) = \{1/(\sigma \sqrt{2\pi})\} \,.\, e^{-(x-\mu)\,(x-\mu)\,/(2\sigma\sigma)}$, para $-\infty < x < +\infty$
  - Dado $X \sim (\mu, \sigma^2)$ $\rightarrow$ $z = (x-u)/\,\sigma$
  - $E[Z] = E[(X-\mu)/\,\sigma] = (1/\,\sigma)\,(E[X] - \mu) = 0$
  - $Var[Z] = Var[(X-\mu)/\,\sigma] = (1/\,\sigma^2)\, Var[(X-\mu)]$

    $= (1/\,\sigma^2)\, Var[X] = 1$

$\rightarrow Z \sim N(0,1)$ $\rightarrow$ Normal padrão!

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - $z = (x-u)/\sigma$
  - $Z \sim N(0,1)$ → Normal padrão!
  - Então:
    - $P(a<X<b) = P[(a-u)/\sigma < Z < (b-u)/\sigma]$
  - Agora, basta saber os valores tabelados de $F(Z)$ para se obter $P(a<X<b)$.

Z = (x-media)/d_padrao

Se, media = 3 e d_padrao = 2
Então: Z = (X-3)/2
X = 3 —> Z = 0
X = 5 —> Z = 1
X = -1 —> Z = ?

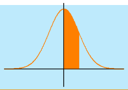
P(Z<z)

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,5	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,6	0,0002	0,0002	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,7	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,8	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
-3,9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Tabela da Distribuição Normal Padrão
P(Z<z)

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Tabela - Normal Padrão de 0 a z



$P(0 \leq Z \leq z)$

Segunda casa decimal de Z (colunas); Parte inteira e primeira casa decimal de Z (linhas)

z	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	0,2704	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	0,4976	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
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$

Tabela (parcial) de distribuição normal reduzida: probabilidade de valor entre zero e 1,25.

	0	1	2	3	4	5	6
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636
0,2	0,0793	0,0832	0,0871	0,0910	0,0946	0,0987	0,1026
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	**0,3944**	0,3962
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - Exemplo: dado $X \sim N(2,9)$, obter $P(2<X<5)$.
  - Procedimento:
    - Transformar X e Z (distribuição padrão) → $z = (x-u)/\sigma$
      - $x=2 \rightarrow z=0$
      - $x=5 \rightarrow z=1$
    - Obter o valor de P(.) na tabela.

      - $P(2<X<5) = P[(2-2)/3 < Z < (5-2)/3]$
        $= P(0<Z<1) = 0{,}3413$

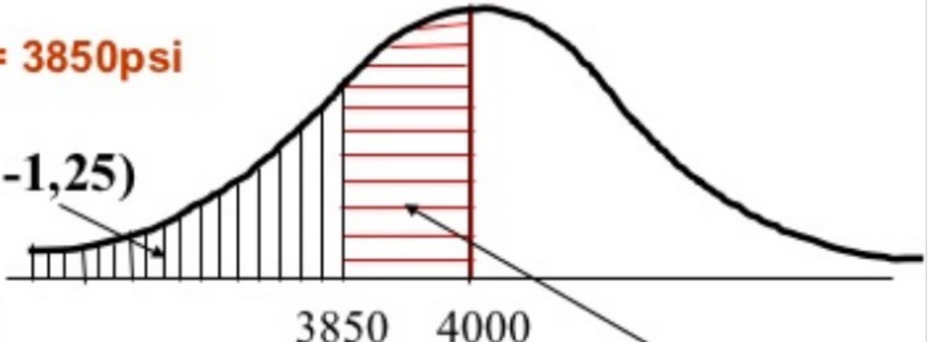
Modelos Contínuos

1) Após 28 dias de curagem, o cimento de uma certa marca tem uma resistência compressiva média de 4000psi. Suponha que a resistência tem uma distribuição normal com desvio-padrão de 120psi. Qual a probabilidade de se comprar um pacote de cimento com resistência compressiva de 28 dias menor que 3850psi?

N(μ;σ) = N(4000,120) psi

X = 3850psi

P(z ≤ -1,25)



$$z = \frac{X-\mu}{\sigma} = \frac{3850-4000}{120} = -1{,}25$$

Área em vermelho = z = -1,25 = 0,3944

-1,25

Área desejada = 0,50 – 0,3944 = 0,1056 = **10,56%**

$$P(Z \leq -1{,}25) = 0{,}1056 = 10{,}56\%$$

INTEGRADO FACULDADE CAMPO MOURÃO/PR

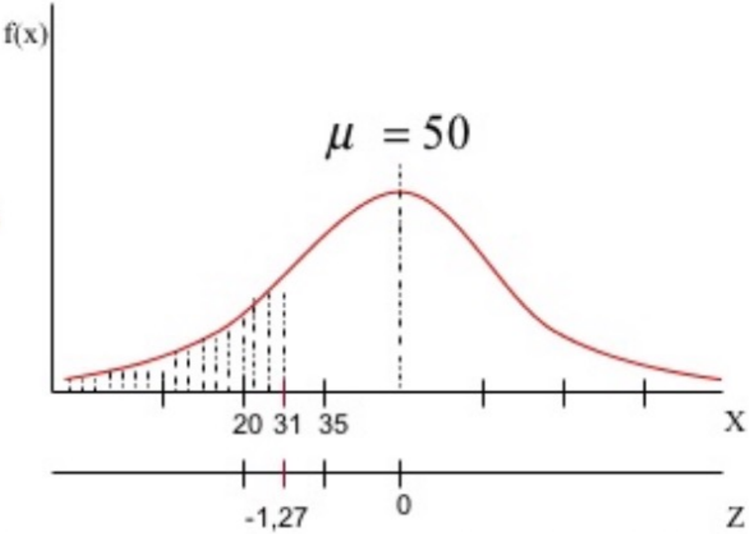
Modelos Contínuos

Distribuição Normal - Exemplos

2) Uma grande empresa faz uso de milhares de lâmpadas elétricas que permanecem acessas continuamente. A vida de uma lâmpada pode ser considerada como uma variável aleatória normal com **vida média de 50 dias e desvio-padrão de 15 dias.** Se no dia 1° de agosto foram instaladas 8000 lâmpadas novas, aproximadamente quantas deverão ser substituídas no dia 1° de setembro?

N(μ,σ) = N(50;15) dias **X = 31 dias**

$$z = \frac{X-\mu}{\sigma} = \frac{31-50}{15} = -1{,}27$$



Consultando tabela:

$$P(Z \leq -1{,}27) = 0{,}3980 \quad \text{logo} \quad 0{,}5000 - 0{,}3980 = 0{,}1020 = 10{,}20\%$$

Deverão ser substituídas um total de (0,1020x 8.000) = **816 lâmpadas**

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
- Exercício:
  - Se o QI de uma população segue uma distribuição gaussiana, com média 100 e desvio padrão 15. Qual é a proporção de pessoas com QI acima de 135?

Modelos Contínuos

- Modelo Normal ou Gaussiana – $X \sim N(\mu, \sigma^2)$
  - $z = (x-\mu)/\sigma$
  - $Z \sim N(0,1)$ → Normal padrão!
- Exercício:
  - Se $X \sim (0,2)$ e $Y = 3X^2$, calcule, $E[Y]$, $\sigma(Y)$ e $f_y(y)$.
  - Dado que X1 e X2 são duas V.A. com distribuição normal, com médias e desvios-padrões, respectivamente, (μ_1, σ_1) e (μ_1, σ_1). Se $X = X_1 + X_2$, obtenha:
    - $E[X]$
    - $Var[X]$


"That's all Folks!"