Estatística I

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Universidade Federal de Viçosa
Campus UFV - Viçosa

Sumário

Distribuições Contínuas

Distribuições Contínuas

Iremos estudar três modelos probabilísticos para variáveis aleatórias contínuas, ou seja, variáveis para as quais os possíveis valores pertencem a um intervalo de números reais.

Distribuições Contínuas

Iremos estudar três modelos probabilísticos para variáveis aleatórias contínuas, ou seja, variáveis para as quais os possíveis valores pertencem a um intervalo de números reais.

Definição: Uma função X, definida sobre o espaço amostral Ω e assumindo valores num intervalo de números reais, é dita uma variável aleatória contínua.

Distribuição Uniforme

Distribuição Uniforme

O modelo uniforme é o modelo mais simples para v.a. contínuas.

Distribuição Uniforme

O modelo uniforme é o modelo mais simples para v.a. contínuas.

Definição: Uma v.a. X tem distribuição uniforme no intervalo $[\alpha, \beta]$ se sua f.d.p. é dada por:

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{se} \quad \alpha \leq x \leq \beta; \\ 0, & \text{cc} \end{cases}$$

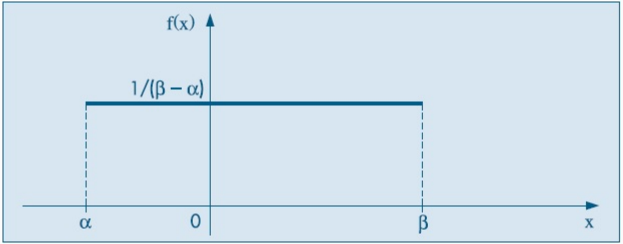


Figura 1: Distribuição uniforme no intervalo $[\alpha, \beta]$

Momentos:

$$E(X) = \frac{\alpha + \beta}{2} \quad \text{e} \quad Var(X) = \frac{(\beta - \alpha)^2}{12}$$

A função de distribuição acumulada da uniforme é fácil de ser encontrada e é dada por:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx = \begin{cases} 0, & \text{se} \quad x < \alpha; \\ \dfrac{x-\alpha}{\beta-\alpha}, & \text{se} \quad \alpha \leq x < \beta; \\ 1, & \text{se} \quad x \geq \beta. \end{cases}$$

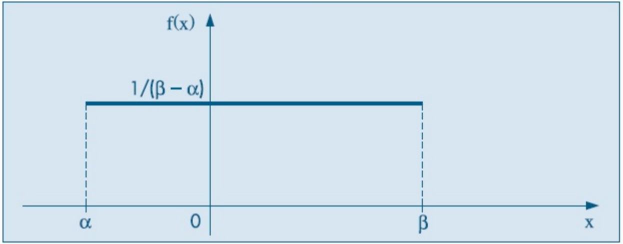


Figura 2: f.d.a. de uma v.a. uniforme no intervalo $[\alpha, \beta]$

Distribuição Exponencial

Distribuição Exponencial

Definição: A variável aleatória contínua X terá distribuição exponencial se sua função densidade de probabilidade, para algum parâmetro $\lambda > 0$, for dada por:

$$f(x;\lambda) = \begin{cases} \lambda e^{(-\lambda x)}, & \text{se} \quad x \geq 0; \\ 0, & cc. \end{cases}$$

$$E(X) = \frac{1}{\lambda} \qquad Var(X) = \frac{1}{\lambda^2}$$

Distribuição Exponencial

1) Mostre que $F(x) = P(X \leq x) = 1 - e^{-x\lambda}$, para $x > 0$

2) Mostre que, para quaisquer $s, t > 0$, $P(X > s + t \mid X > s) = P(X > t)$

3) Suponha que o número de kilômetros que um carro possa trafegar antes da sua bateria estragar tenha distribuição exponencial com um valor médio de 10000 quilômetros. Se uma pessoa deseja realizar uma viagem de pelo menos 5000 km, qual é a probabilidade de que essa pessoa possa terminar a sem precisar trocar a bateria? (obs

Distribuição Exponencial

1) Mostre que $F(x) = P(X \leq x) = 1 - e^{-x\lambda}$, para $x > 0$

2) Mostre que, para quaisquer $s, t > 0$, $P(X > s + t \mid X > s) = P(X > t)$

3) Suponha que o número de kilômetros que um carro possa trafegar antes da sua bateria estragar tenha distribuição exponencial com um valor médio de 10000 quilômetros. Se uma pessoa deseja realizar uma viagem de pelo menos 5000 km, qual é a probabilidade de que essa pessoa possa terminar a sem precisar trocar a bateria? (obs

Distribuição Exponencial

4) A vida média de um satélite é 4 anos, seguindo o modelo exponencial. Seja T a variável definindo o tempo de vida do satélite. Calcule:

a) $P(T > 4)$.

Distribuição Exponencial

4) A vida média de um satélite é 4 anos, seguindo o modelo exponencial. Seja T a variável definindo o tempo de vida do satélite. Calcule:

a) $P(T > 4)$. Resposta: 0,3678

b) $P(5 \leq T \leq 6)$.

Distribuição Exponencial

4) A vida média de um satélite é 4 anos, seguindo o modelo exponencial. Seja T a variável definindo o tempo de vida do satélite. Calcule:

a) $P(T > 4)$. Resposta: 0,3678

b) $P(5 \leq T \leq 6)$. Resposta: 0,0634

c) Se 4 desses satélites forem lançados no mesmo instante, qual é a probabilidade que após 5 anos todos estejam funcionando? Resposta:

Distribuição Exponencial

4) A vida média de um satélite é 4 anos, seguindo o modelo exponencial. Seja T a variável definindo o tempo de vida do satélite. Calcule:

a) $P(T > 4)$. Resposta: 0,3678

b) $P(5 \leq T \leq 6)$. Resposta: 0,0634

c) Se 4 desses satélites forem lançados no mesmo instante, qual é a probabilidade que após 5 anos todos estejam funcionando? Resposta: 0,00674

d) Se 4 desses satélites forem lançados no mesmo instante, qual é a probabilidade que após 5 anos ...ente dois estejam funcionando?

Distribuição Exponencial

4) A vida média de um satélite é 4 anos, seguindo o modelo exponencial. Seja T a variável definindo o tempo de vida do satélite. Calcule:

a) $P(T > 4)$. Resposta: 0,3678

b) $P(5 \leq T \leq 6)$. Resposta: 0,0634

c) Se 4 desses satélites forem lançados no mesmo instante, qual é a probabilidade que após 5 anos todos estejam funcionando? Resposta: 0,00674

d) Se 4 desses satélites forem lançados no mesmo instante, qual é a probabilidade que após 5 anos

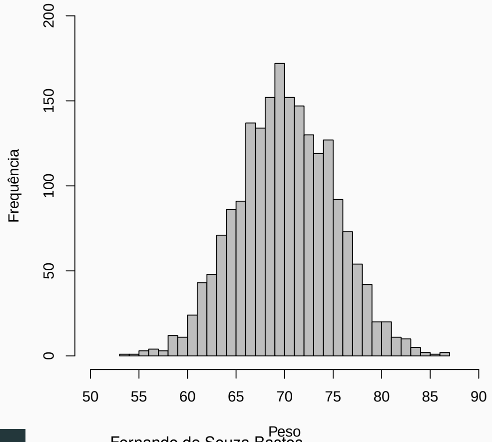
Distribuição Normal

Distribuição Normal

Vamos introduzir, agora, um modelo fundamental em probabilidades e inferência estatística. Suas origens remontam a Gauss em seus trabalhos sobre erros de observações astronômicas, por volta de 1810, donde o nome de distribuição gaussiana para tal modelo.

Distribuição Normal

Considere o peso, em kg, de 2000 jovens selecionados ao acaso entre todos os alunos da UFV, cujo histograma é apresentado:



Introdução

A análise do histograma indica que:

- a distribuição dos valores é aproximadamente simétrica em torno de 70 kg;
- a maioria dos valores encontra-se no intervalo (60;80);
- existe uma pequena proporção de valores abaixo de 55kg e acima de 85kg.

Introdução

A distribuição normal (algumas vezes chamada de distribuição de Gauss) é a distribuição contínua mais habitualmente utilizada no campo da estatística. Ela é de vital importância na estatística, por três razões principais:

- Inúmeras variáveis contínuas comuns no mundo dos negócios possuem distribuições que se assemelham estreitamente à distribuição normal.

Introdução

A distribuição normal (algumas vezes chamada de distribuição de Gauss) é a distribuição contínua mais habitualmente utilizada no campo da estatística. Ela é de vital importância na estatística, por três razões principais:

- Inúmeras variáveis contínuas comuns no mundo dos negócios possuem distribuições que se assemelham estreitamente à distribuição normal.
- A distribuição normal pode ser utilizada para fazer aproximações para várias distribuições de probabilidades discretas.

Introdução

A distribuição normal (algumas vezes chamada de distribuição de Gauss) é a distribuição contínua mais habitualmente utilizada no campo da estatística. Ela é de vital importância na estatística, por três razões principais:

- Inúmeras variáveis contínuas comuns no mundo dos negócios possuem distribuições que se assemelham estreitamente à distribuição normal.
- A distribuição normal pode ser utilizada para fazer aproximações para várias distribuições de probabilidades discretas.

o normal proporciona a base para a

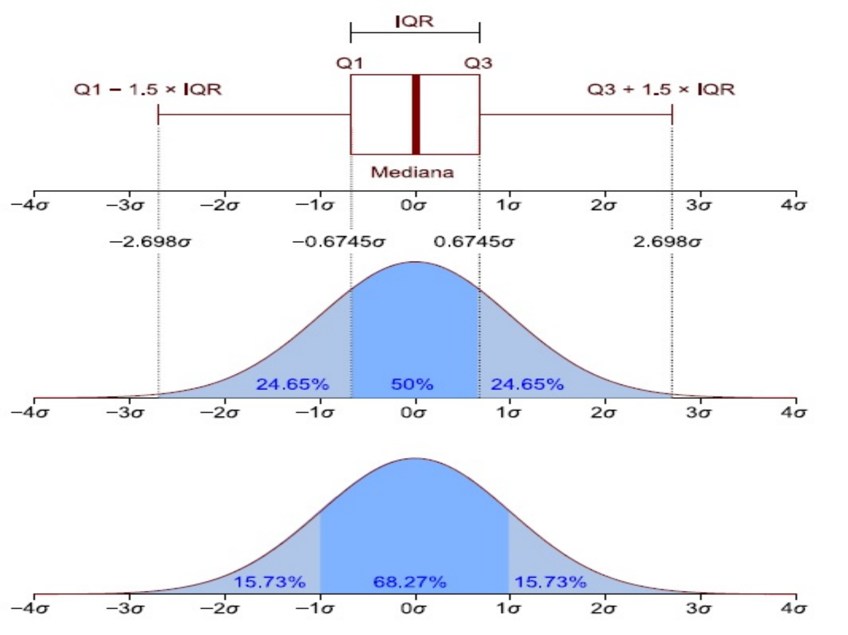
Introdução

A distribuição normal é representada pelo clássico formato de sino. Na distribuição normal, você pode calcular a probabilidade de que venham a ocorrer valores dentro dos limites de determinadas amplitudes ou intervalos. No entanto, uma vez que a probabilidade para variáveis contínuas é mensurada como uma área abaixo da curva, a probabilidade exata de um valor específico, a partir de uma distribuição contínua tal como a distribuição normal, é zero.

Introdução

A distribuição normal possui várias propriedades teóricas importantes:

- Ela é simétrica, sua média aritmética, sua mediana e sua moda são, consequentemente, iguais.
- Em sua aparência, ela tem o formato de um sino.
- Sua amplitude interquartil é igual a 1,349 desvio-padrão. Consequentemente, os 50% valores centrais estão contidos dentro dos limites de um intervalo que tem como valores fronteiriços dois terços de um desvio-padrão abaixo da média aritmética e dois terços de um ...ão acima da média aritmética.


IQR
Q1
Q3
Q1 − 1.5 × IQR
Q3 + 1.5 × IQR
Mediana
−4σ −3σ −2σ −1σ 0σ 1σ 2σ 3σ 4σ
−2.698σ
−0.6745σ
0.6745σ
2.698σ
24.65%
50%
24.65%
−4σ −3σ −2σ −1σ 0σ 1σ 2σ 3σ 4σ
15.73%
68.27%
15.73%
−4σ −3σ −2σ −1σ 0σ 1σ 2σ 3σ 4σ

Função Densidade para a probabilidade normal

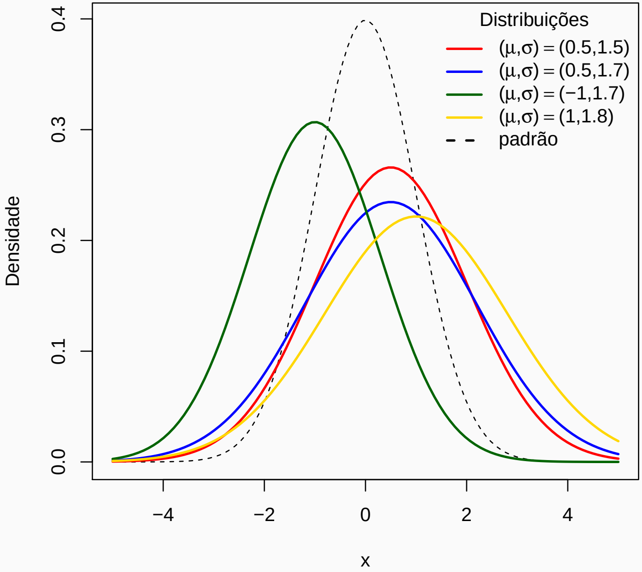
$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(X-\mu)^2}{2\sigma^2}\right], \quad X \in \mathbb{R}, \tag{1}$$

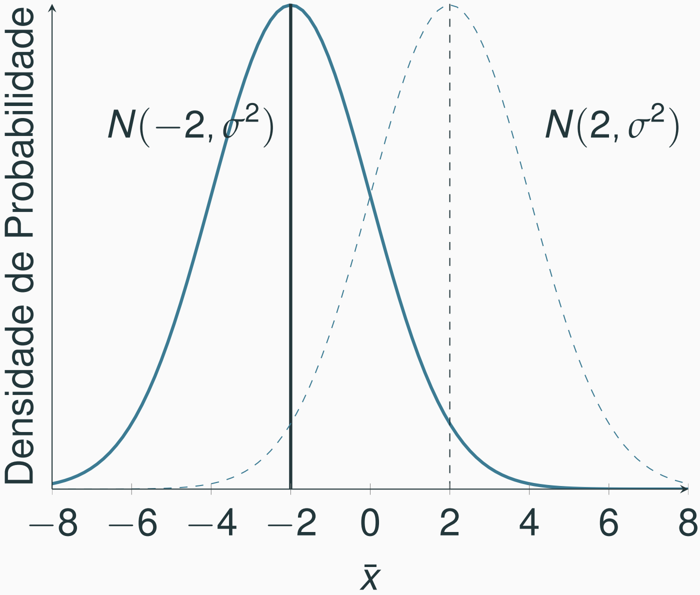
em que μ é a média aritimética e σ é o desvio-padrão.

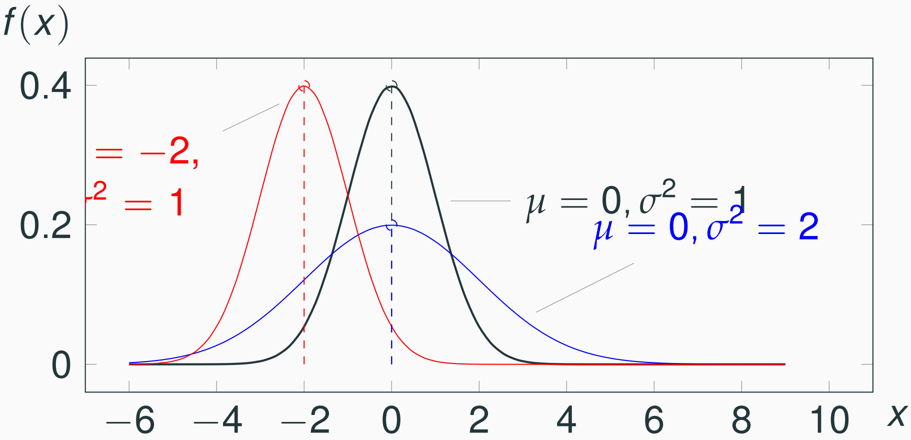
Introdução

Embora a Equação possa parecer complicada, as probabilidades da variável aleatória X dependem somente de dois parâmetro, a média aritmética, μ, e o desvio-padrão, σ. Todas as vezes que você determina valores específicos para μ e σ, é gerada uma distribuição de probabilidade normal diferente.

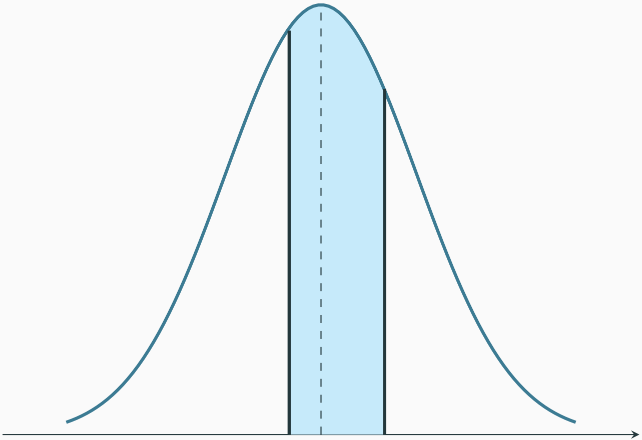
Comparação de distribuições normais




Densidade de Probabilidade
$N(-2, \sigma^2)$
$N(2, \sigma^2)$
−8 −6 −4 −2 0 2 4 6 8
$\bar{x}$


$f(x)$
0.4
0.2
0
$= -2,$
${}^2 = 1$
$\mu = 0, \sigma^2 = 1$
$\mu = 0, \sigma^2 = 2$
-6 -4 -2 0 2 4 6 8 10
x

Cálculo de Probabilidades



$$P(a \leq X \leq b)$$

Se $X \sim N(\mu, \sigma^2)$ então $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$.

A variável $Z \sim N(0, 1)$ denomina-se normal padrão ou normal reduzida. Notemos que:

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$$

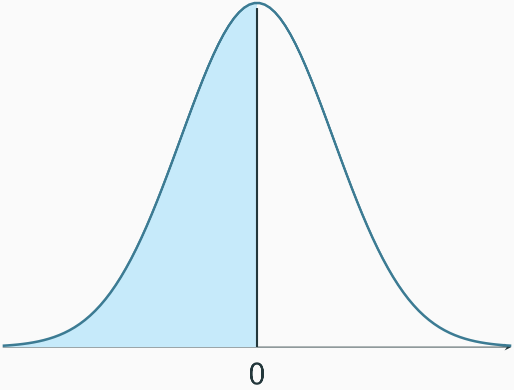
e, dado a variável $Z \sim N(0, 1)$, podemos obter $X \sim N(\mu, \sigma^2)$ através da transformação inversa:

$$X = \mu + Z\sigma$$

Uso da Tabela

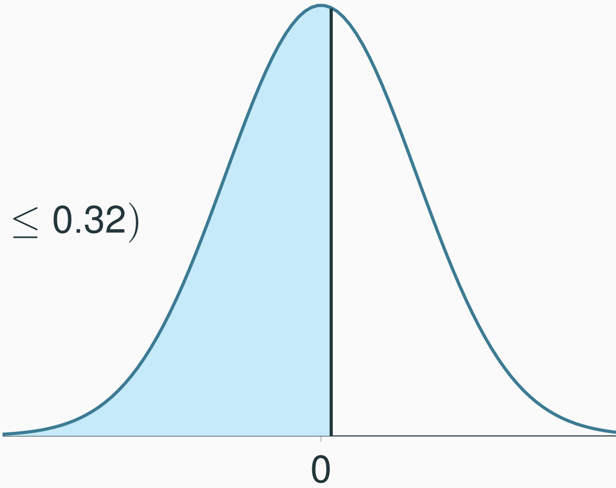
Exemplos

Dado $Z \sim N(0, 1)$, calcule $P(Z \leq 0.5)$



$P(Z \leq 0) = 0.5$

$$P(Z \leq 0.32) = 0.5 + P(0 \leq Z \leq 0.32)$$
$$= 0.5 + 0.1255$$
$$= 0.6255$$



$$P(Z \leq 0) = 0.6255$$

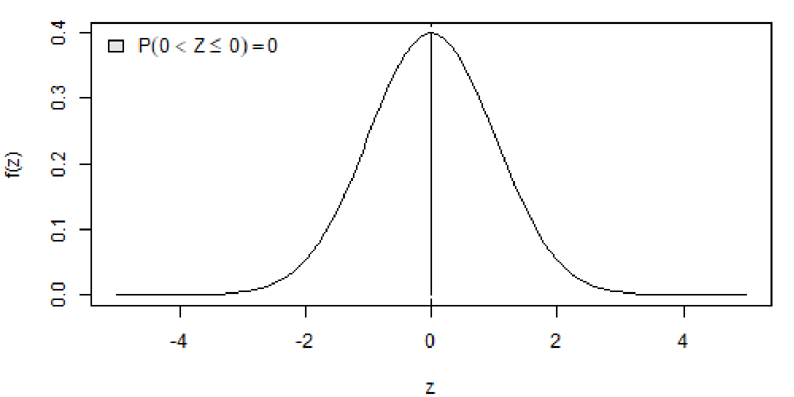
Calcule:

1. $P(0 < Z \leq 1.71) = ?$
2. $P(1.32 \leq Z \leq 1.79) = ?$
3. $P(Z \geq 1.5) = ?$
4. $P(Z \leq -1.3) = ?$
5. $P(-1.5 \leq Z \leq 1.5) = ?$
6. $P(-1.32 < Z < 0) = ?$
7. $P(-2.3 < Z \leq -1.49) = ?$
8. $P(-1 \leq Z \leq 2) = ?$
9. $P(Z \leq z) = 0.975$, qual o valor de z?
10. $P(0 < Z \leq z) = 0.4975$, qual o valor de z?
11. $P(Z \geq z) = 0.3$, qual o valor de z?
12. $P(Z \geq z) = 0.975$, qual o valor de z?

1. $P(0 < Z \leq 1.71) = 0.4564$
2. $P(1.32 \leq Z \leq 1.79) = 0.0567$
3. $P(Z \geq 1.5) = 0.0668$
4. $P(Z \leq -1.3) = 0.0968$
5. $P(-1.5 \leq Z \leq 1.5) = 0.8664$
6. $P(-1.32 < Z < 0) = 0.4066$
7. $P(-2.3 < Z \leq -1.49) = 0.0574$
8. $P(-1 \leq Z \leq 2) = 0.8186$
9. $P(Z \leq z) = 0.975$, o valor de z é 1.96.
10. $P(0 < Z \leq z) = 0.4975$, o valor de z é 2.81.
11. $P(Z \geq z) = 0.3$, o valor de z é 0.53.
12. $P(Z \geq z) = 0.975$, o valor de z é -1.96.

Seja $X \sim N(10, 64)$. Calcule:

1. $P(6 < X \leq 12) =?$
2. k tal que $P(X \geq k) = 0.05$
3. k tal que $P(X \leq k) = 0.025$
4. $P(Z \leq -1.3) =?$
5. $P(-1.5 \leq Z \leq 1.5) =?$
6. $P(-1.32 < Z < 0) =?$
7. $P(-2.3 < Z \leq -1.49) =?$
8. $P(-1 \leq Z \leq 2) =?$
9. $P(Z \leq z) = 0.975$, qual o valor de z?
10. $P(0 < Z \leq z) = 0.4975$, qual o valor de z?
11. $P(Z \geq z) = 0.3$, qual o valor de z?
12. $P(Z \geq z) = 0.975$, qual o valor de z?


P(0 < Z ≤ 0) = 0
0.0
0.1
0.2
0.3
0.4
f(z)
-4
-2
0
2
4
z

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,
0,0	0,00000	0,00399	0,00798	0,01197	0,01595	0,01994	0,02392	0,02790	0,03188	0,035
0,1	0,03983	0,04380	0,04776	0,05172	0,05567	0,05962	0,06356	0,06749	0,07142	0,075
0,2	0,07926	0,08317	0,08706	0,09095	0,09483	0,09871	0,10257	0,10642	0,11026	0,114
0,3	0,11791	0,12172	0,12552	0,12930	0,13307	0,13683	0,14058	0,14431	0,14803	0,151
0,4	0,15542	0,15910	0,16276	0,16640	0,17003	0,17364	0,17724	0,18082	0,18439	0,187
0,5	0,19146	0,19497	0,19847	0,20194	0,20540	0,20884	0,21226	0,21566	0,21904	0,222
0,6	0,22575	0,22907	0,23237	0,23565	0,23891	0,24215	0,24537	0,24857	0,25175	0,254
0,7	0,25804	0,26115	0,26424	0,26730	0,27035	0,27337	0,27637	0,27935	0,28230	0,285
0,8	0,28814	0,29103	0,29389	0,29673	0,29955	0,30234	0,30511	0,30785	0,31057	0,313
0,9	0,31594	0,31859	0,32121	0,32381	0,32639	0,32894	0,33147	0,33398	0,33646	0,338
1,0	0,34134	0,34375	0,34614	0,34849	0,35083	0,35314	0,35543	0,35769	0,35993	0,362
1,1	0,36433	0,36650	0,36864	0,37076	0,37286	0,37493	0,37698	0,37900	0,38100	0,382
1,2	0,38493	0,38686	0,38877	0,39065	0,39251	0,39435	0,39617	0,39796	0,39973	0,401
1,3	0,40320	0,40490	0,40658	0,40824	0,40988	0,41149	0,41309	0,41466	0,41621	0,417
1,4	0,41924	0,42073	0,42220	0,42364	0,42507	0,42647	0,42785	0,42922	0,43056	0,431
1,5	0,43319	0,43448	0,43574	0,43699	0,43822	0,43943	0,44062	0,44179	0,44295	0,444
[illegible]	[illegible]	[illegible]	0,44738	0,44845	0,44950	0,45053	0,45154	0,45254	0,45352	0,454

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,
2,0	0,47725	0,47778	0,47831	0,47882	0,47932	0,47982	0,48030	0,48077	0,48124	0,481
2,1	0,48214	0,48257	0,48300	0,48341	0,48382	0,48422	0,48461	0,48500	0,48537	0,485
2,2	0,48610	0,48645	0,48679	0,48713	0,48745	0,48778	0,48809	0,48840	0,48870	0,488
2,3	0,48928	0,48956	0,48983	0,49010	0,49036	0,49061	0,49086	0,49111	0,49134	0,491
2,4	0,49180	0,49202	0,49224	0,49245	0,49266	0,49286	0,49305	0,49324	0,49343	0,493
2,5	0,49379	0,49396	0,49413	0,49430	0,49446	0,49461	0,49477	0,49492	0,49506	0,495
2,6	0,49534	0,49547	0,49560	0,49573	0,49585	0,49598	0,49609	0,49621	0,49632	0,496
2,7	0,49653	0,49664	0,49674	0,49683	0,49693	0,49702	0,49711	0,49720	0,49728	0,497
2,8	0,49744	0,49752	0,49760	0,49767	0,49774	0,49781	0,49788	0,49795	0,49801	0,498
2,9	0,49813	0,49819	0,49825	0,49831	0,49836	0,49841	0,49846	0,49851	0,49856	0,498
3,0	0,49865	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49896	0,499
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,499
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,499
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,499
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,499
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,499
[illegible]	[illegible]	[illegible]	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,499

Referências

- Bastos, Fernando de Souza (2025). ***Apostila Interativa.*** Disponível online: https://ufvest.shinyapps.io/ApostilaInterativa/.
- Ferreira, Eric Batista e Marcelo Silva de Oliveira (2020). ***Introdução à Estatística com R.*** Editora Universidade Federal de Alfenas. URL: https://www.unifal-mg.edu.br/bibliotecas/wp-content/uploads/sites/125/2021/12/32-EBR_Unifal.pdf.
- Meyer, Paul L (1982). ***Probabilidade: aplicações à estatística.*** Livros Técnicos e Científicos.

Referências ii

- Montgomery, D. C. e G. C Runger (2016). ***Estatística Aplicada E Probabilidade Para Engenheiros.*** 6ª ed. São Paulo: Grupo Gen-LTC.
- Morettin, P.A. e W.O Bussab (2023). ***Estatística básica.*** 10ª ed. São Paulo: Editora Saraiva.
- Peternelli, Luiz Alexandre (s.d.). ***Apostila (EST 106).*** Formato slide Disponível no PVANet - Moodle.