Distribuição e Densidade de Probabilidade

Aula de hoje

- Tópicos
  - Teoria da Probabilidade
  - Variáveis aleatórias
    - Variáveis discretas
    - Variáveis contínuas
  - Distribuição binomial
  - Distribuição normal

Introdução

- A teoria da probabilidade é a base da inferência estatística.
- Através da inferência estatística podemos buscar inferir características da população a partir de uma amostra de dados obtidos ao acaso.

Introdução

- Modelos probabilísticos são modelos teóricos que nos descrevem a distribuição de frequências de um fenômeno, quando se tem uma amostra representativa de dados obtidos ao acaso.
- Exemplo de um modelo probabilístico para lançamento de um dado:

Face	1	2	3	4	5	6
Frequência Teórica	1/6	1/6	1/6	1/6	1/6	1/6

- Modelos Probabilísticos são descritos através de:
  - Um Espaço Amostral
  - Uma probabilidade para cada ponto amostral

Distribuição de probabilidades

- Algumas propriedades:
- A probabilidade de qualquer evento é maior ou igual a zero e menor ou igual a um.
- A probabilidade de todo o espaço amostral é igual a um.
- A probabilidade do conjunto vazio é igual a zero.

Distribuição conjunta das frequências

- Variáveis grau de instrução (Y) e região de procedência (V) – Exemplo de Bussab-Morettin

V \ Y	Ensino Fundamental	Ensino Médio	Superior	Total
Capital	4	5	2	11
Interior	3	7	2	12
Outro	5	6	2	13
Total	12	18	6	36

Distribuição de probabilidades

- Por definição, a intersecção de dois eventos A e B ocorre quando os dois eventos ocorrem simultaneamente.
  - Em nosso exemplo, qual é a probabilidade de um funcionário da empresa ter completado apenas ensino fundamental e ser da capital?

- Já na reunião de dois eventos A e B, temos que pelo menos um dos eventos deve ocorrer
  - Em nosso exemplo, qual é a probabilidade de um funcionário da empresa ter completado apenas ensino fundamental ou ser da capital?

Distribuição de probabilidades

- Um evento B é complementar ao evento A se ele compreender todos os pontos amostrais do espaço amostral que não pertençam ao evento A.

Variáveis aleatórias

- É uma variável cujo resultado ou valor decorre de um experimento ou fenômeno que envolva um elemento casual.
- São, por exemplo: soma de dois dados, cotação do dólar, precipitação diária de chuva em uma cidade, limite de resistência de uma peça
- Podem ser
  - discretas
  - contínuas
- Notação
  - variáveis aleatórias: X, Y, ... (letras maiúsculas)
  - valores possíveis das variáveis: x, y....(minúsculas)

Variáveis aleatórias discretas

- Assume valores em intervalo de números inteiros
- A função que atribui a probabilidade a cada valor possível de uma variável aleatória discreta é denominada função de probabilidade

  f(x) = P(X = x)

- Exemplo
  - Dado honesto: f(x) = 1/6, para x=1, 2, 3, 4, 5 ou 6

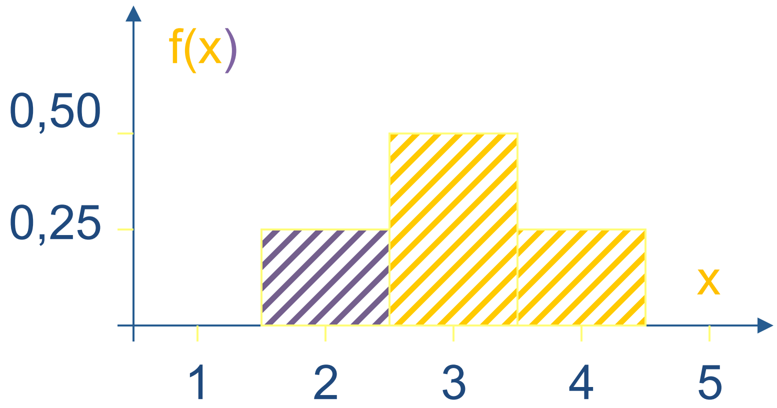
- Propriedades

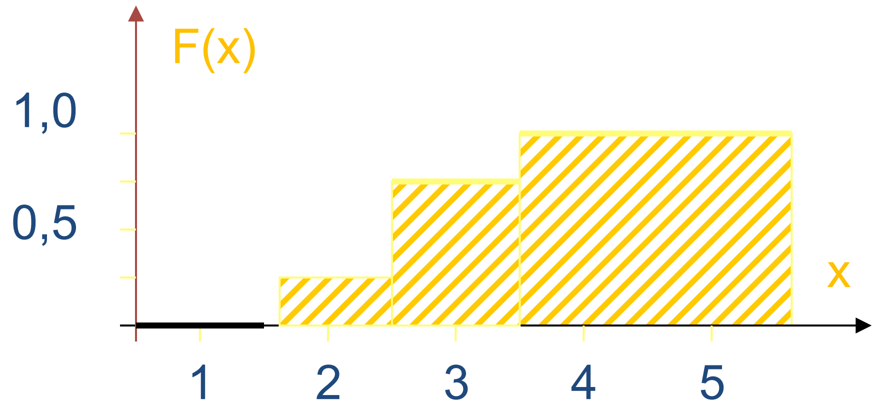
$$f(x) \geq 0 \qquad \sum_{todos\ X} f(x) = 1$$

Função distribuição (acumulada)

- A função distribuição acumulada de uma variável aleatória X associa a cada valor possível de X a probabilidade de ocorrência de um valor menor ou igual a x. Denota-se F(x)

$$F(x) = P(X \leq x)$$





Média e variância de uma distribuição calculada pela distribuição de probabilidades

Média

$$\mu = \sum_{todos\,x} x.f(x) = E(x)$$

Variância

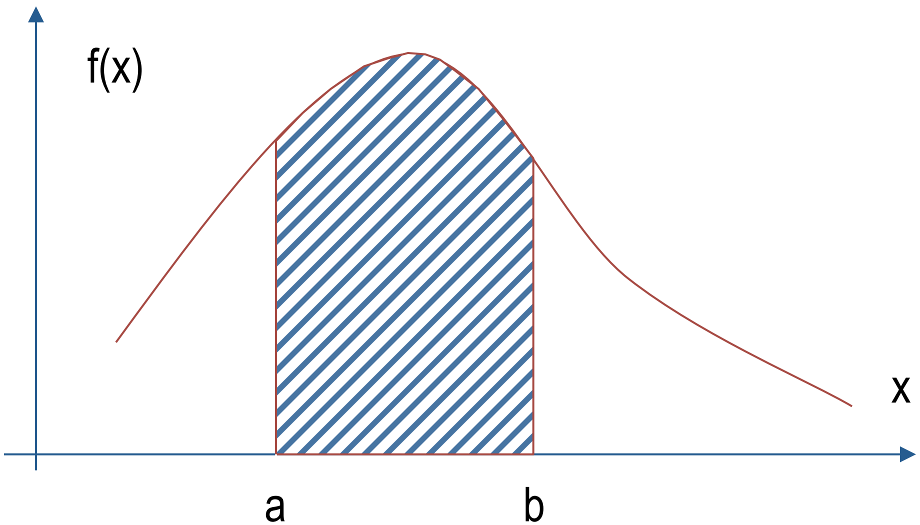
$$\sigma^2 = \sum_{todos\,x} (x-\mu)^2 .f(x)$$

Variáveis aleatórias contínuas

- Assume valores em intervalo de números reais
- Não é possível listar todos os possíveis valores de uma VA contínua
- Associa-se probabilidades a intervalos de valores da VA contínua

Variáveis aleatórias contínuas

- f(x) = função densidade de probabilidade



$$P(X = x) = 0$$

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

Função densidade de probabilidade

- Seja X uma variável aleatória continua. A função de densidade de probabilidade f(x) é uma função que satisfaz as seguintes condições:

1. $f(x)$ é positiva para todo $x \in R_x$

2. Tem-se:

$$P(a < X < b) = \int_a^b f(x)dx \qquad \int_{R_x} f(x) = 1$$

Observações

1. A probabilidade de qualquer ponto é zero

2. $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b)$
   $= P(a < X < b)$.

3. A função integrada entre dois limites a e b (a < b) é a probabilidade, ou seja, a área sob a curva.

4. A função de distribuição é definida como

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

Distribuição binomial

- É distribuição discreta de probabilidade. Ela está associada a um experimento de múltiplas etapas

Experimento binomial

- O experimento consiste de uma sequência de n ensaios idênticos
- Dois resultados são possíveis em cada ensaio: sucesso e fracasso
- P(sucesso)=p   P(fracasso)= 1-p = q

  p+ q=1
- Os ensaios são independentes

Exemplo: jogar 8 vezes um dado

- O experimento consiste em 8 jogadas do dado (ensaios idênticos)
- Cada ensaio resulta em sucesso (sair 6) ou fracasso(não sair 6)
  - P(sucesso)= P(sair 6)=1/6
  - P(fracasso)= P(não sair 6)=5/6
- Os ensaios são independentes

Exemplo: determinar a probabilidade de saírem 3 faces 6, em 8 jogadas de um dado

Ensaios	1	2	3	4	5	6	7	8
Resultados	s	s	s	f	f	f	f	f
Probabilidade	1/6	1/6	1/6	5/6	5/6	5/6	5/6	5/6

$$= C_{8,3}\left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^5 = \frac{8!}{(8-3)!3!}0,17^3.0,83^5$$

$$= 56(0,0049)(0,3939) = 56(0,0019)$$

$$= 0,106$$

Exemplo-lançamento moedas

Solução:

- Número de tentativas n=10
- Número de sucessos desejado k=3
- Probabilidade de sucesso em 1 tentativa p=1/2
- Probabilidade de insucesso em 1 tentativa q=1/2
- Usando estes parâmetros na fórmula da distribuição binomial, temos

$$f(X) = P(X=k) = C_{n,k}\, p^k\, q^{n-k}$$

Exemplo lançamento moedas (resolução com Excel)

Solução:

Se X é a variável aleatória que representa "o número de caras" então

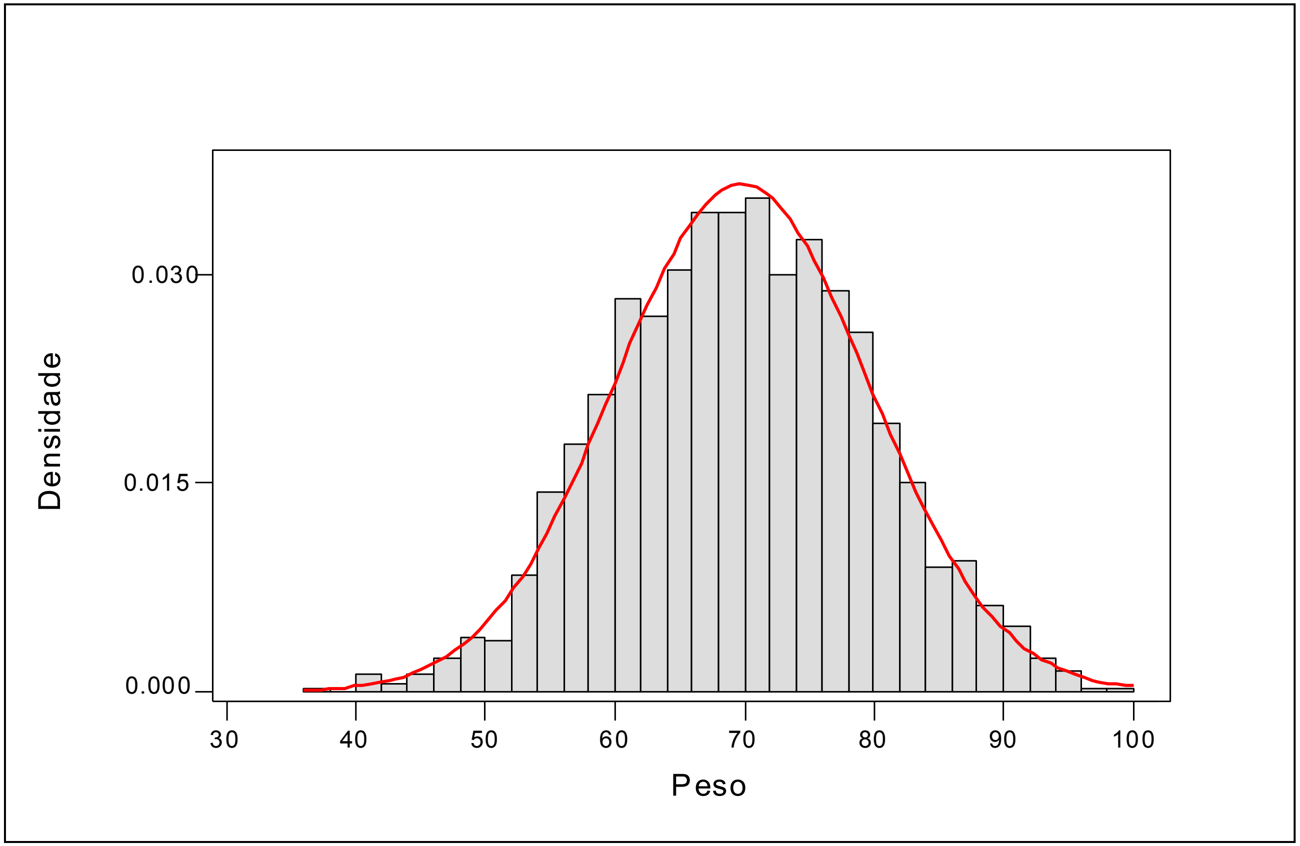
Se quisermos calcular a probabilidade de sair três caras em dez lançamentos

P(X=3)= distrbinom(3 ; 10 ; 0,5 ; 0)= 0,1172

Distribuição normal

X: peso, em kg, de uma pessoa adulta escolhida ao acaso da população.

Como se distribuem os valores da variável aleatória X, isto é, qual a distribuição de probabilidades de X ?



Distribuição normal

A distribuição normal é uma das mais importantes distribuições contínuas de probabilidade pois:

Muitos fenômenos aleatórios comportam-se de forma próxima a essa distribuição. Exemplos:

1. altura
2. pressão sangüínea
3. peso

Pode ser utilizada para calcular, de forma aproximada, probabilidades para outras distribuições.

Distribuição normal

A VA X tem distribuição normal com parâmetros μ e σ^2 se sua função densidade de probabilidade é dada por

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

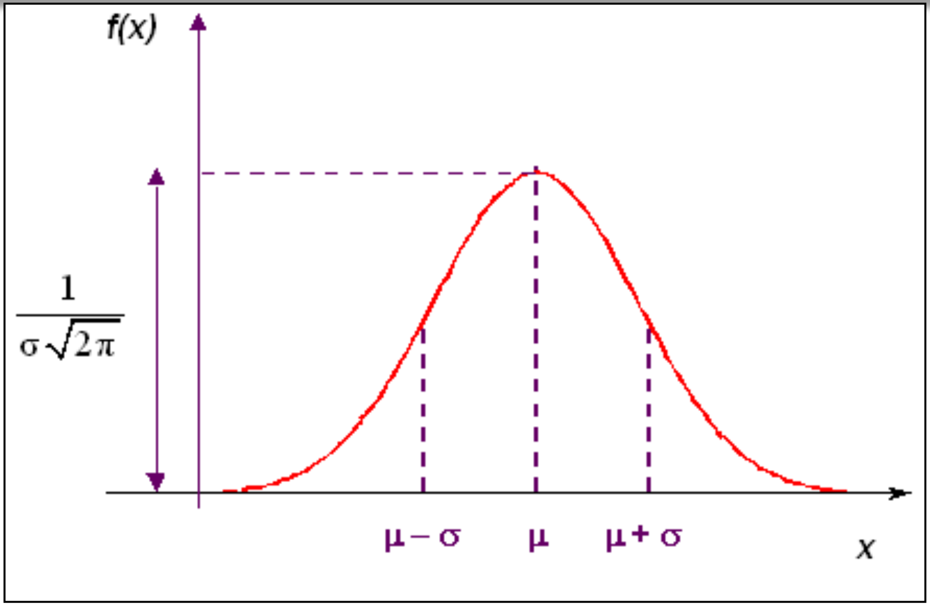
Pode ser mostrado que

1. μ é o valor esperado (média) de X ($-\infty < \mu < \infty$)
2. σ^2 é a variância de X ($\sigma^2 > 0$)

Notação : $X \sim N(\mu ; \sigma^2)$

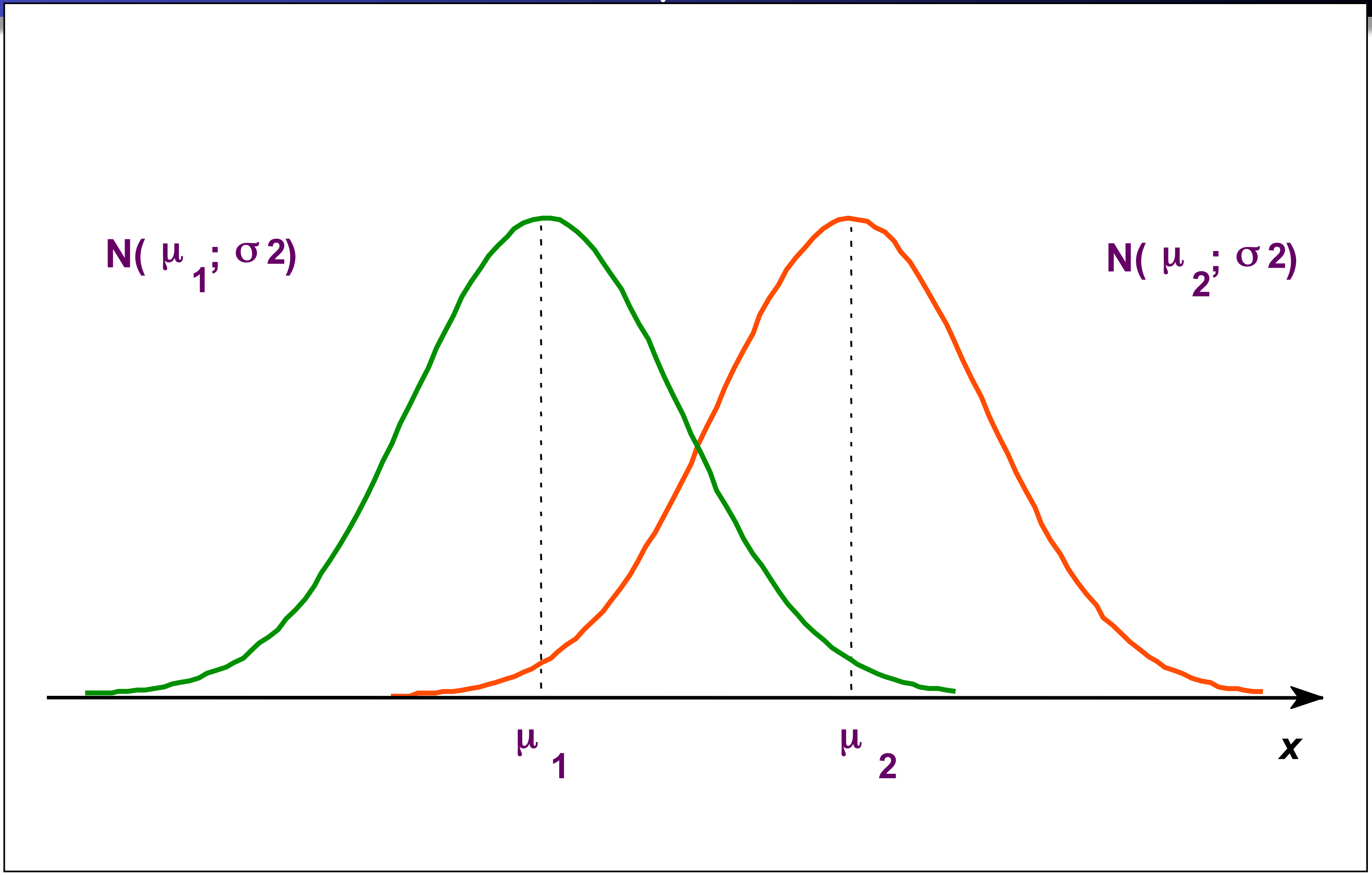
Para cada par de parâmetros μ e σ há uma curva diferente de f(x)

Propriedades da Normal



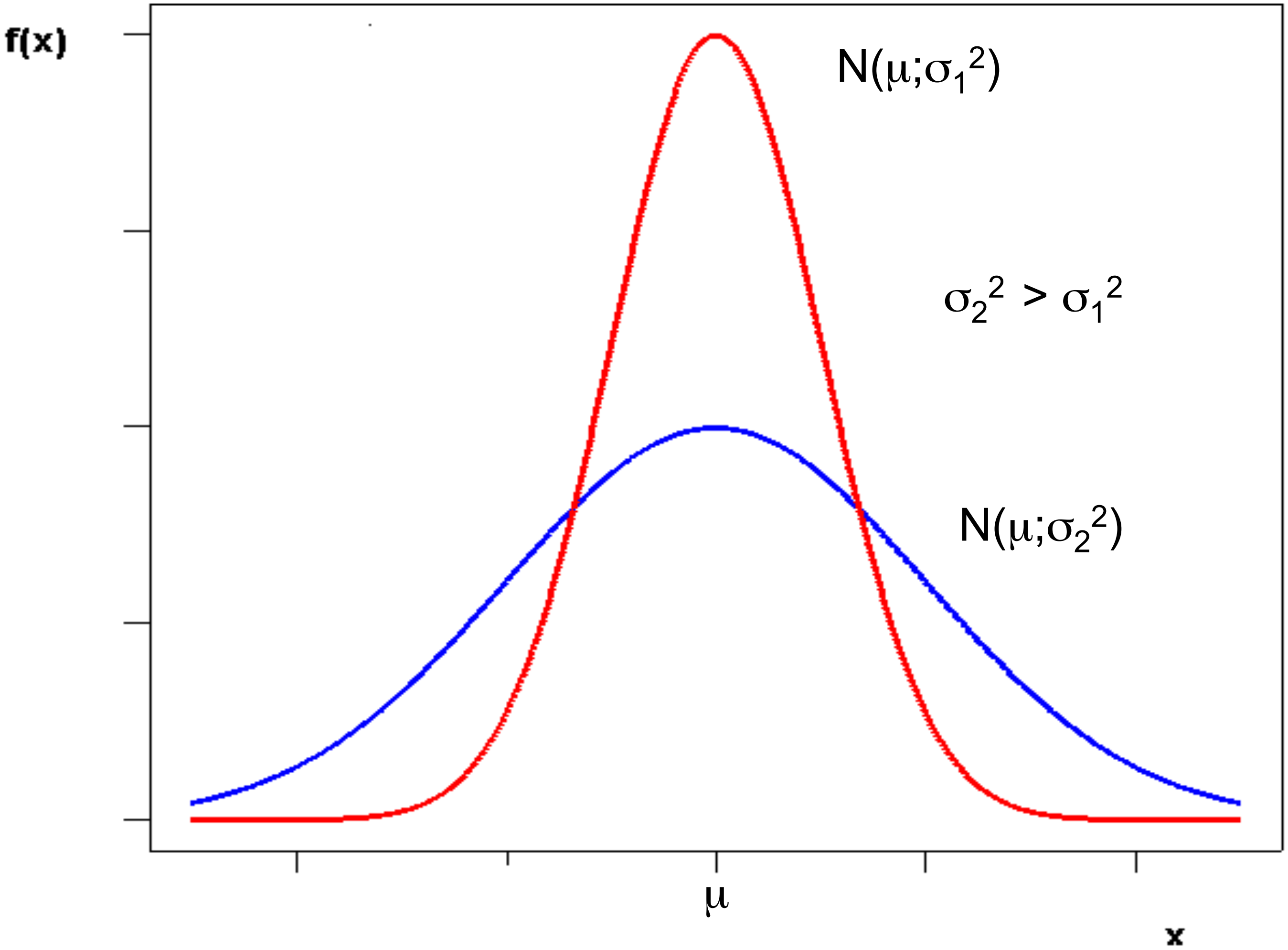
- **E(X) = μ (média ou valor esperado)**
- **Var(X) = σ^2 (e portanto, DP(X) = σ)**
- **$f(x) \to 0$ quando $x \to \pm\infty$**
- **$x = \mu$ é ponto de máximo de $f(x)$**
- **$\mu - \sigma$ e $\mu + \sigma$ são pontos de inflexão de $f(x)$**
- **a curva Normal é simétrica em torno da média μ**

Influência de μ na curva Normal



Curvas Normais com mesma variância σ^2
mas médias diferentes ($\mu_2 > \mu_1$).

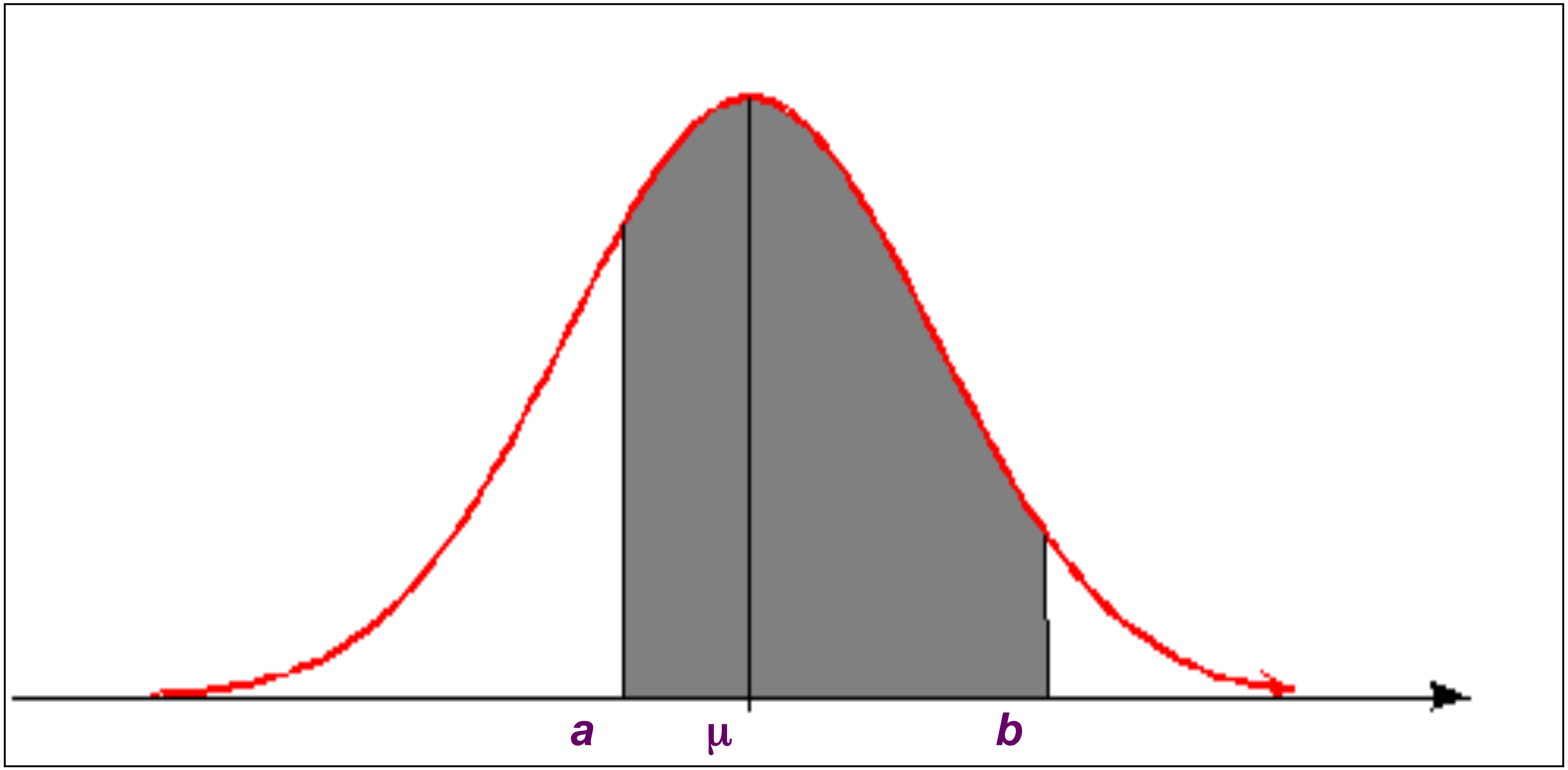
Influência de σ^2 na curva Normal



Curvas Normais com mesma média μ, mas com variâncias diferentes **($\sigma_2^2 > \sigma_1^2$).**

Cálculo de probabilidades

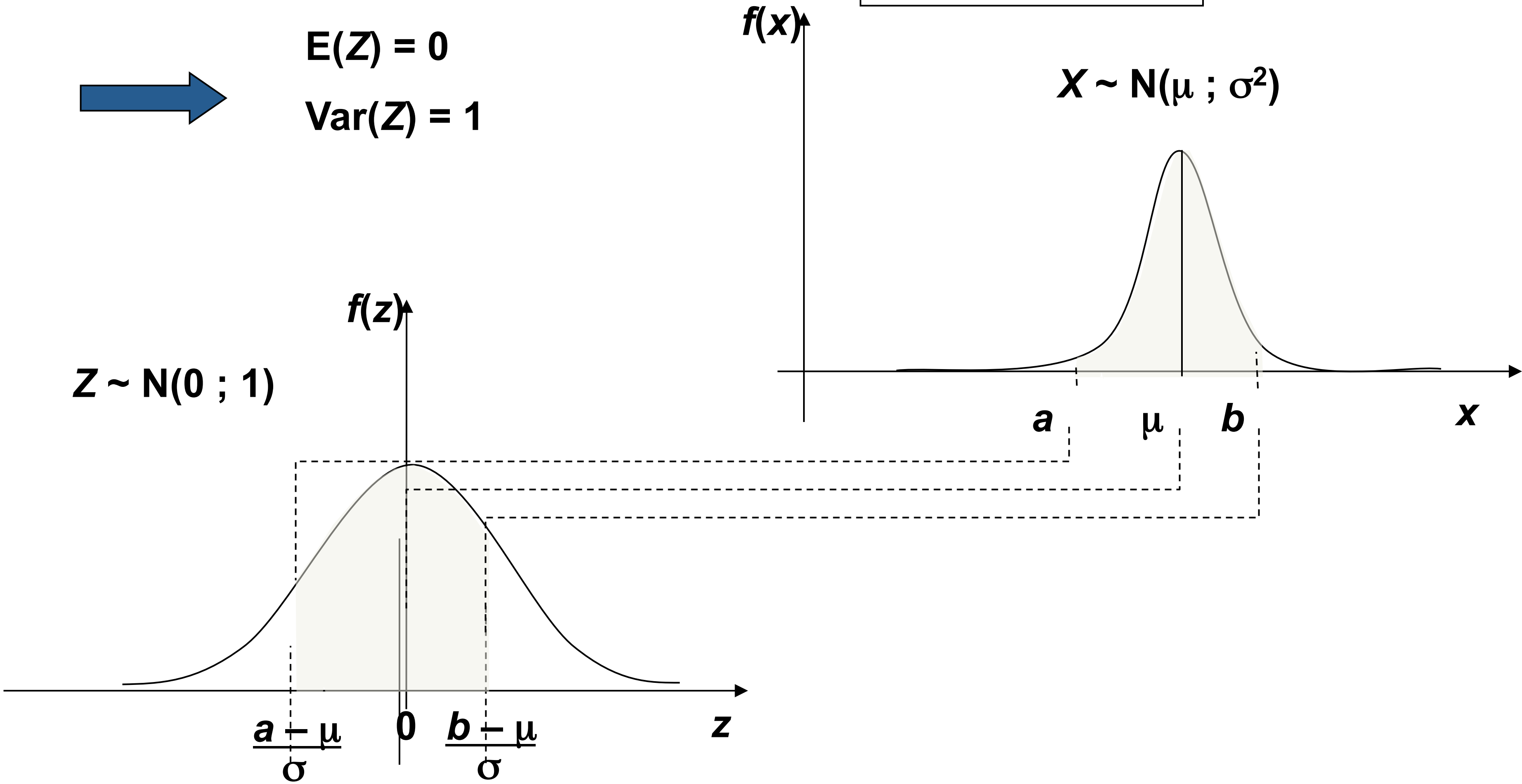
Área sob a curva e acima do eixo horizontal (*x*) entre *a* e *b*.



Se $X \sim N(\mu\ ;\ \sigma^2)$, definimos $$Z = \frac{X - \mu}{\sigma}$$

$\Rightarrow$ $E(Z) = 0$

$Var(Z) = 1$



A v.a. *Z* ~ N(0;1) denomina-se *normal padrão* ou *reduzida.*

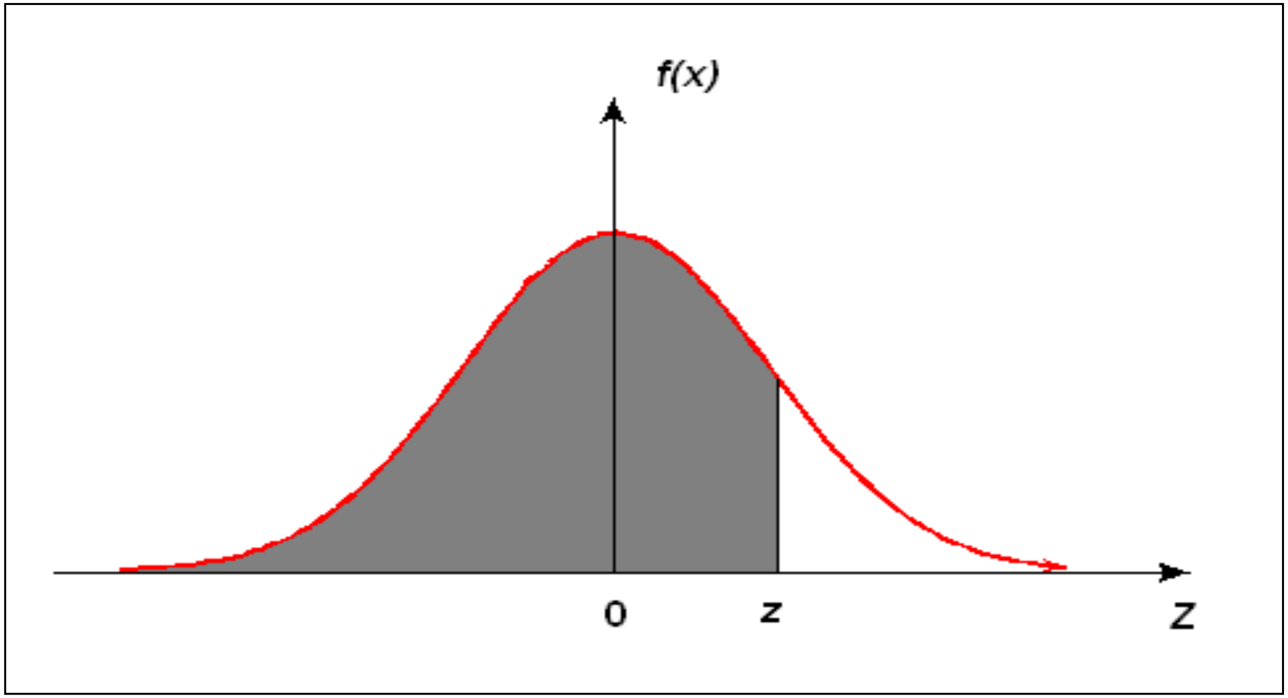
Portanto,

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

Dada a v.a. *Z* ~N(0;1) podemos obter a v.a. *X* ~ N(μ;σ^2) através da transformação inversa

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

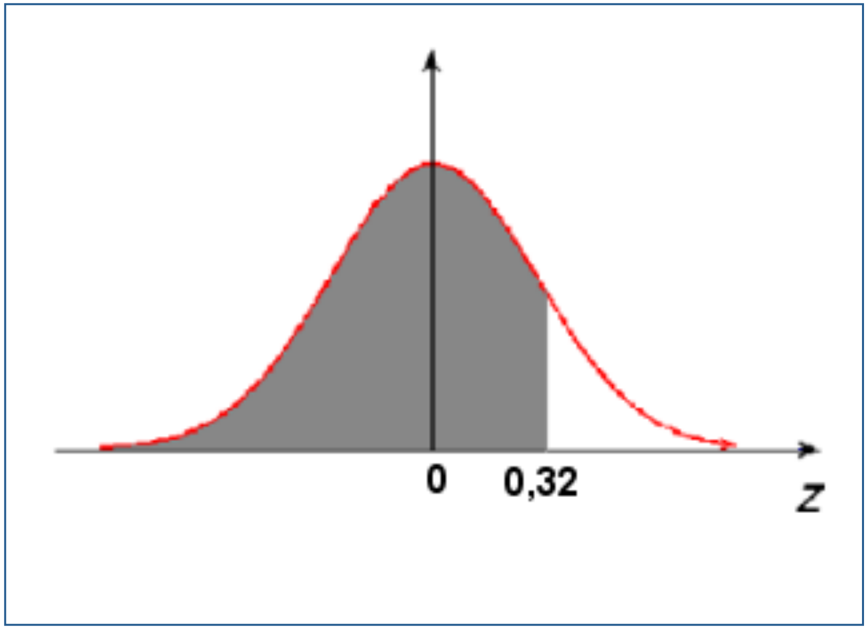
Uso da tabela normal padrão



Denotamos : A(z) = P($Z \leq z$), para $z \geq 0$.

Exemplo: Seja $Z \sim N(0; 1)$, calcular

a) $P(Z \leq 0{,}32)$

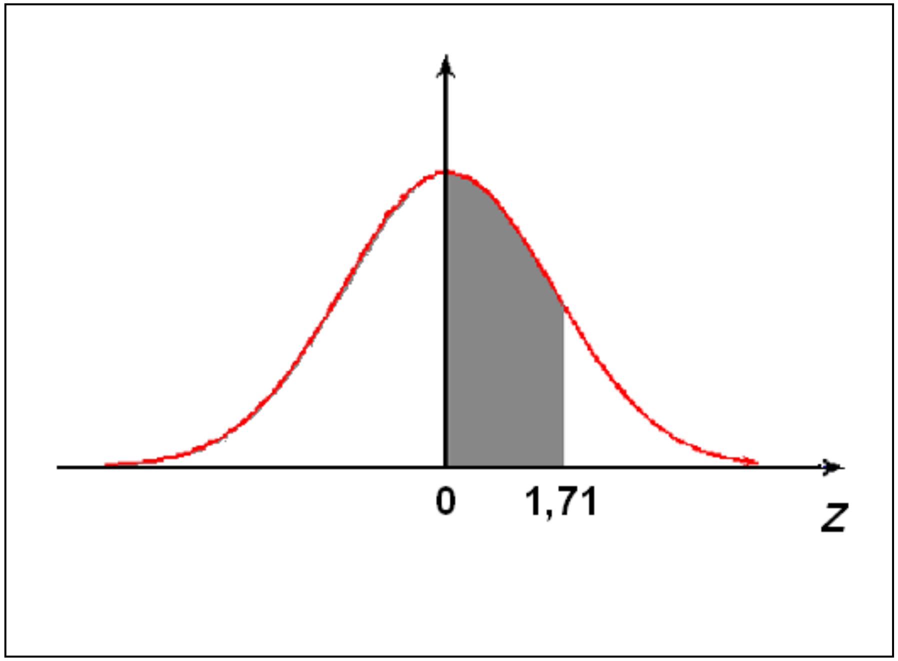


$P(Z \leq 0{,}32) = A(0{,}32) = 0{,}6255.$

Encontrando o valor na Tabela N(0;1)

z	0	1	2
0,0	0,5000	0,5039	0,5079
0,1	0,5398	0,5437	0,5477
0,2	0,5792	0,5831	0,5870
0,3	0,6179	0,6217	0,6255
⋮	⋮	⋮	⋮

b) $P(0 < Z \leq 1{,}71)$



$$P(0 < Z \leq 1{,}71) = P(Z \leq 1{,}71) - P(Z \leq 0)$$

$$= A(1{,}71) - A(0)$$

$$= 0{,}9564 - 0{,}5 = 0{,}4564.$$

Obs.: $P(Z < 0) = P(Z > 0) = 0{,}5.$

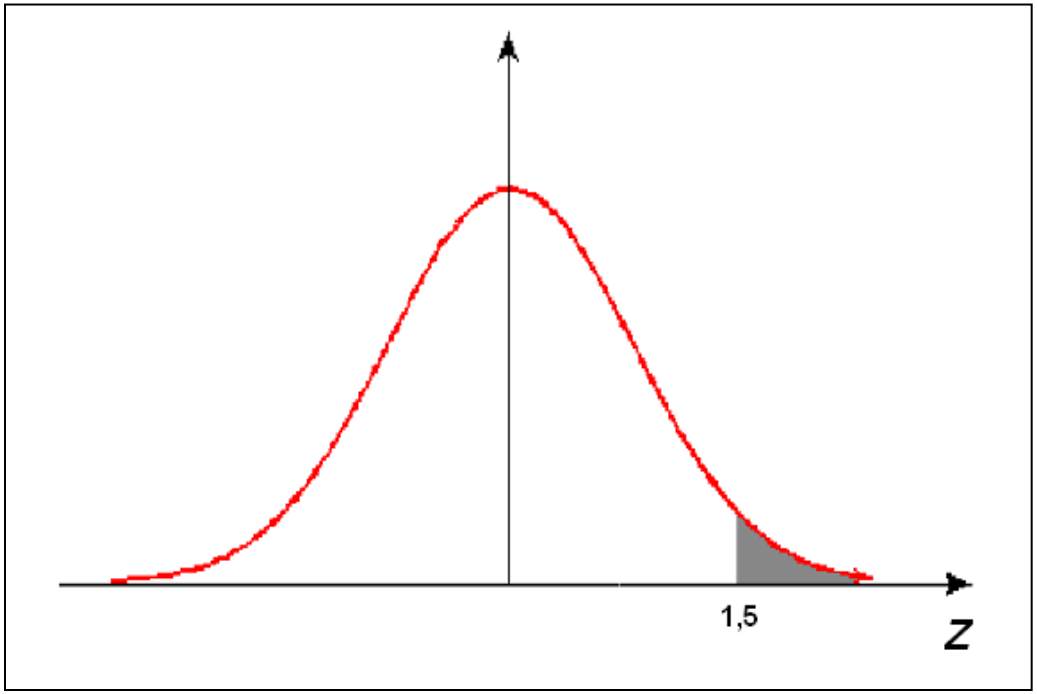
NORMAL DISTRIBUTION TABLE

Entries represent the area under the standardized normal distribution from -∞ to z, Pr(Z<z)
The value of z to the first decimal is given in the left column. The second decimal place is given in the top row.

z	0.00	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

	Values of z for selected values of Pr(Z<z)						
z	0.842	1.036	1.282	1.645	1.960	2.326	2.576
Pr(Z<z)	0.800	0.850	0.900	0.950	0.975	0.990	0.995

d) $P(Z \geq 1,5)$



$P(Z > 1,5) = 1 - P(Z \leq 1,5) = 1 - A(1,5)$

$= 1 - 0,9332 = 0,0668.$

NORMAL DISTRIBUTION TABLE

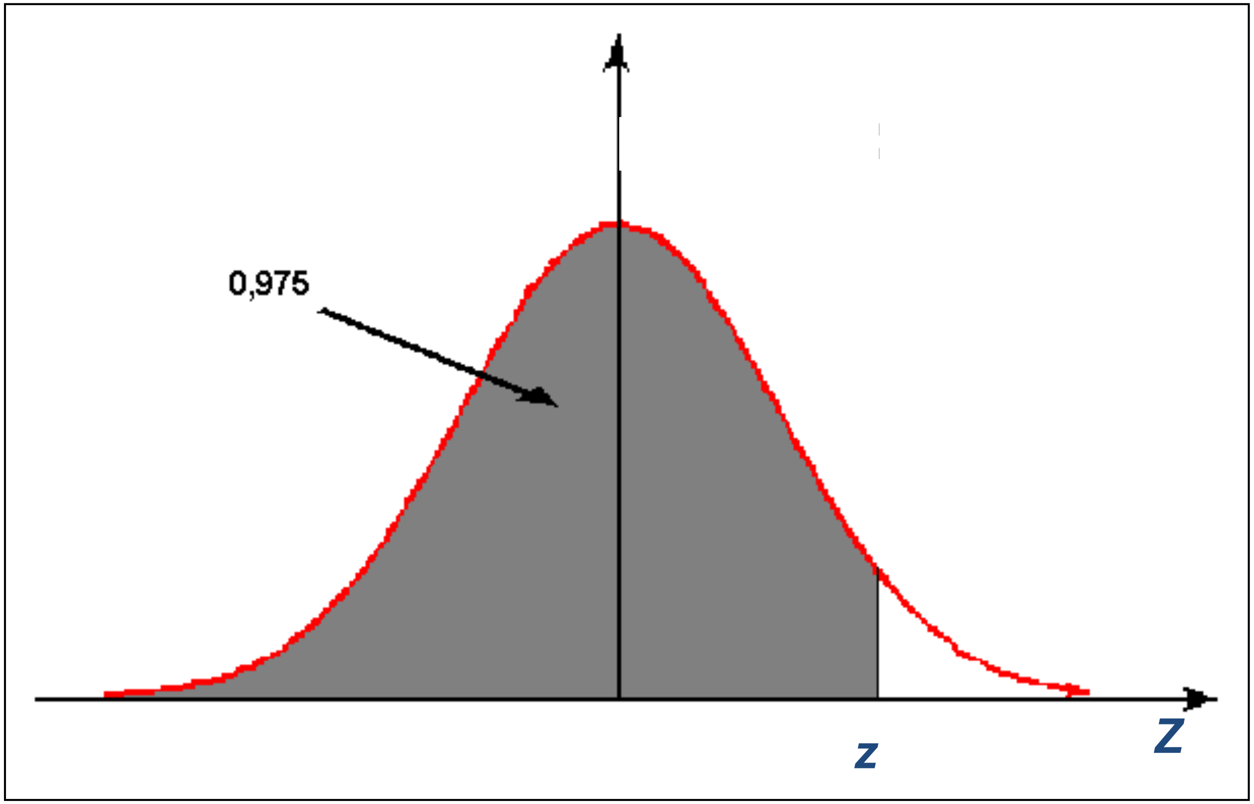
Entries represent the area under the standardized normal distribution from -∞ to z, Pr(Z<z)
The value of z to the first decimal is given in the left column. The second decimal place is given in the top row.

z	0.00	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

Values of z for selected values of Pr(Z<z)							
z	0.842	1.036	1.282	1.645	1.960	2.326	2.576
Pr(Z<z)	0.800	0.850	0.900	0.950	0.975	0.990	0.995

Como encontrar o valor z da distribuição N(0;1) tal que:

(i) $P(Z \leq z) = 0{,}975$



z é tal que $A(z) = 0{,}975$.

Pela tabela, $z = 1{,}96$.