

VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÕES

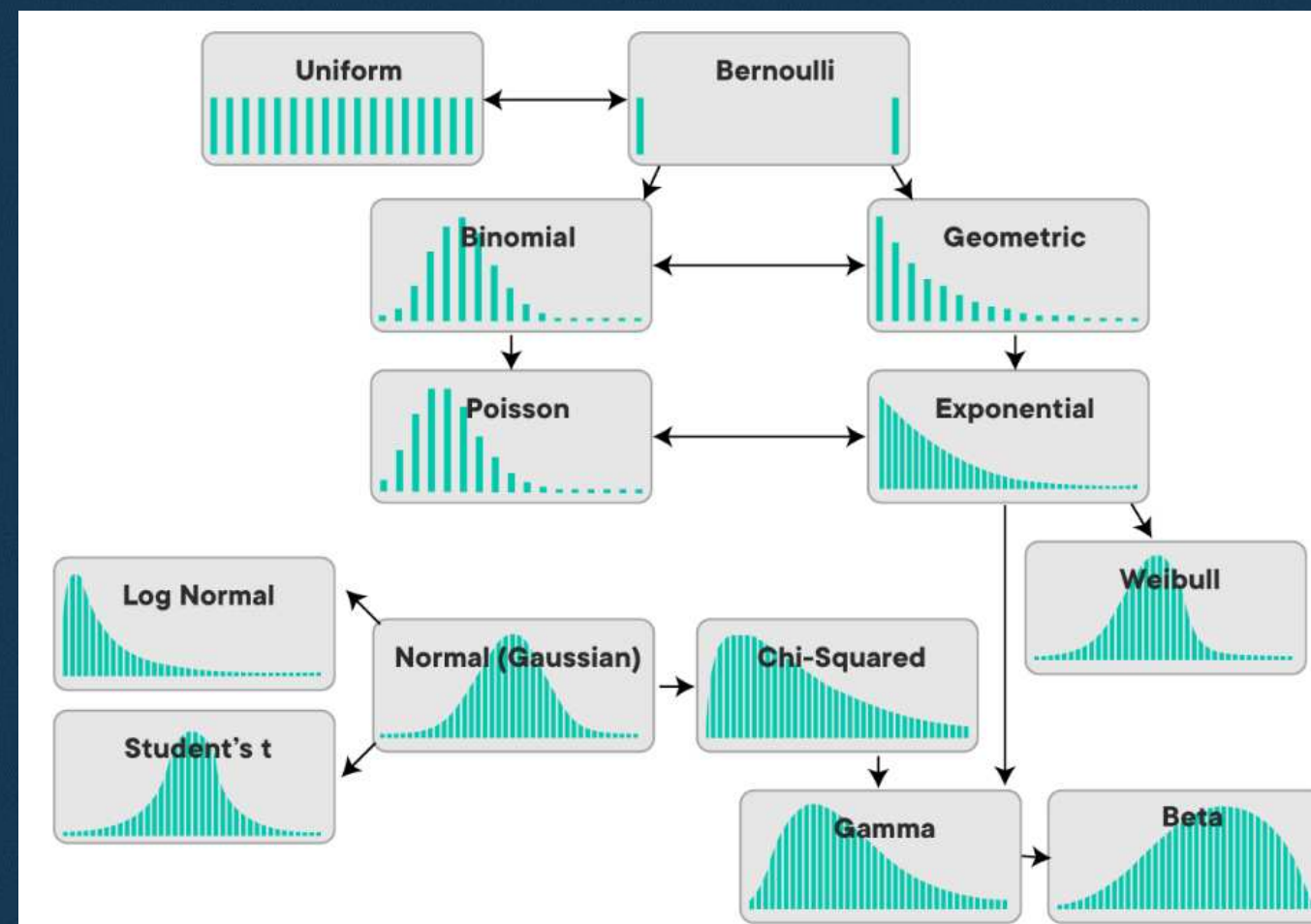
Por que é importante?

- Criar modelos que representam com precisão os dados e prever comportamentos futuros.
- Otimizar os processos e melhorar a eficiência
- Entender a distribuição dos tempos de espera em um serviço ao cliente.
- Análise de riscos
- Retornos de investimentos e a avaliação de riscos de crédito.

Distribuições

Discretas e contínuas

- Uniforme
- Bernoulli
- Binomial
- Poisson
- Exponencial
- Normal



VARIÁVEIS ALEATÓRIAS

Variáveis Aleatórias

Uma variável aleatória é uma variável quantitativa, cujo resultado depende de fatores aleatórios.

Exemplo: No lançamento de um dado, embora saibamos os possíveis resultados, o resultado depende de fatores de sorte (por isso é aleatório).

Variáveis Aleatórias

Uma variável aleatória é uma função que atribui valores numéricos aos resultados de um experimento.

São representadas por letras maiúsculas (X).

Seus valores são representados por letras minúsculas (x).

As variáveis aleatórias podem ser discretas ou contínuas.

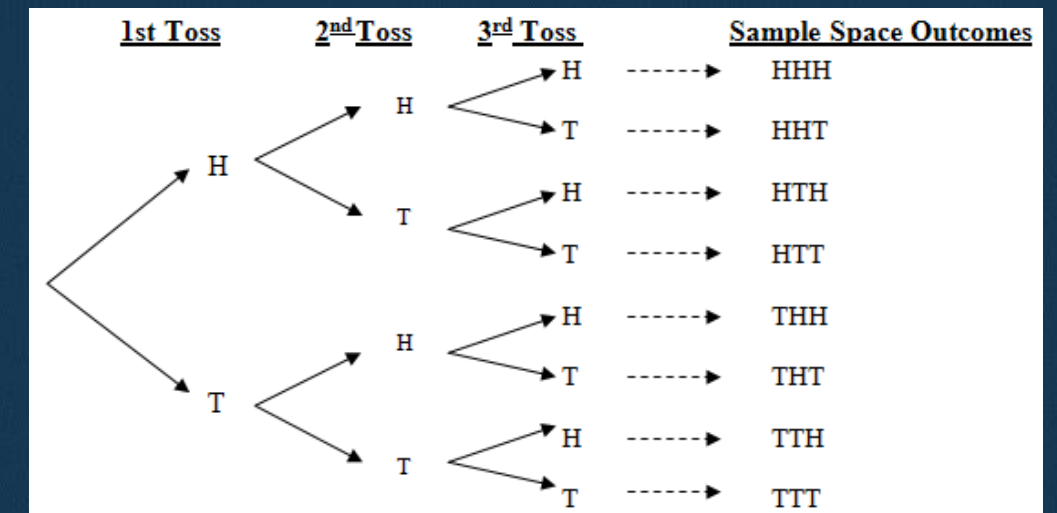
Variável Aleatória Discreta

A variável aleatória discreta assume um número contável de valores distintos.

Variáveis aleatórias discretas são definidas por uma função massa de probabilidade: $P(X = x)$.

Exemplo

Suponha que uma moeda seja jogada 3 vezes.



A variável aleatória X conta o número de caras nos lançamentos. Os valores possíveis de X são 0, 1, 2 ou 3.

Temos que $P(X=0) = 1/8$.

A distribuição de probabilidade de X é:

x	0	1	2	3
$P(X=x)$	1/8	3/8	3/8	1/8

Exemplo

A distribuição de probabilidade de X é:

x	0	1	2	3
P(X=x)	1/8	3/8	3/8	1/8

$$P(X < 1) = P(X=0) = 1/8$$

$$P(X \leq 1) = P(X=0) + P(X=1) = 1/8 + 3/8 = 4/8$$

$$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 1/8 + 3/8 + 3/8 = 7/8$$

$$P(X \leq 2) = 1 - P(X=3) = 1 - 1/8 = 7/8$$

Variável Aleatória Contínua

A variável aleatória contínua é caracterizada por (infinitos) valores dentro de um intervalo.

Variáveis aleatórias contínuas são definidas por uma função densidade de probabilidade: $P(X \leq x)$

Variável Aleatória Contínua

Para uma variável aleatória contínua: $P(X=x) = 0$.

Para uma variável aleatória contínua X , as probabilidades são calculadas para intervalos: $P(X < a)$, $P(a < X < b)$, $P(X \geq b)$

Como $P(X = a)$ e $P(X = b)$ são iguais a zero, o seguinte é válido para variáveis aleatórias contínuas:

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

DISTRIBUIÇÕES

DISTRIBUIÇÕES DISCRETAS

Distribuições Discretas

Uniforme

Bernoulli

Binomial

Geométrica

Binomial Negativa

Hipergeométrica

Poisson

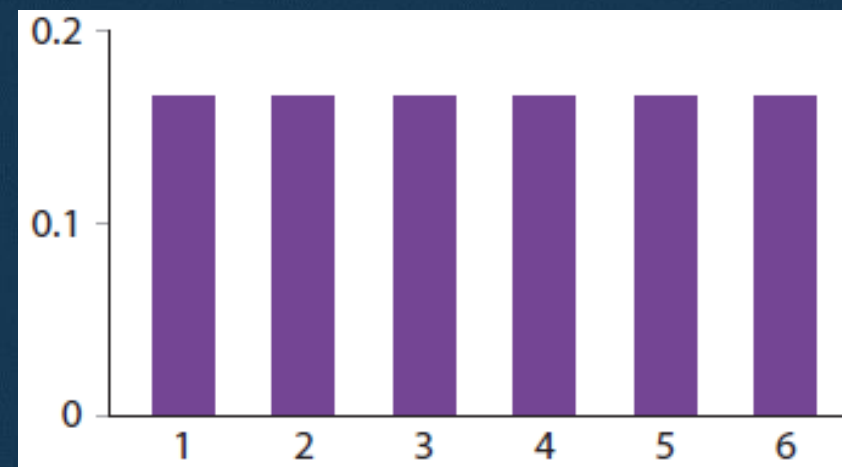
UNIFORME

Uniforme

Todos os resultados têm a mesma probabilidade de ocorrer.

$$P(X=x) = 1/N$$

Exemplo: Ao jogar um dado, qualquer resultado (face do dado) tem a probabilidade igual a $1/6$.



BERNOULLI

Bernoulli

Um experimento Bernoulli é o que tem apenas dois resultados possíveis.

Os chamamos de sucesso e fracasso, onde “sucesso” é o que se deseja observar.

Exemplo: você quer observar a ocorrência de defeito em peças. Uma peça selecionada aleatoriamente, com defeito, seria um resultado do tipo “sucesso” (o resultado “sucesso” não está associado a algo bom, necessariamente).

Aplicações

- Uma venda pode ser efetuada ou não
- Um cliente pode ser do tipo adimplente ou inadimplente
- Uma peça fabricada por uma indústria pode ser perfeita ou defeituosa
- Um consumidor pode devolver ou não um produto comprado
- Um exame médico pode ter como resultado positivo ou negativo

Bernoulli

Suponha que p = probabilidade de sucesso e $1 - p$ = probabilidade de fracasso.

$$P(X=x) = p^x (1 - p)^{(1-x)}$$

Exemplo: A jogada de uma moeda, resultado “cara”.

Nesse caso, $p = 1/2$ e $1-p = 1/2$

$$P(X=1) = 1/2^1 (1 - 1/2)^{(1-1)} = 1/2$$

Exemplo

A partir de uma pesquisa, foi verificado que no período de vendas de Natal, cada cliente que entra no site de determinada loja tem 60% de chance de comprar um produto qualquer. Qual a probabilidade de compra e de não compra?

A probabilidade de sucesso (o cliente adquirir um produto qualquer) é de 0,6 e a probabilidade de não comprar produto algum é de $1 - 0,6 = 0,4$.

BINOMIAL

Binomial

Uma variável aleatória binomial é definida como o número de sucessos em n tentativas de um processo de Bernoulli.

$$P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Exemplo: Jogue uma moeda 15 vezes, X é o número de caras.

$p = 1/2$, $n = 15$

$$P(X=5) = \frac{15!}{5!(15-5)!} \left(\frac{1}{2}\right)^5 \left(1 - \frac{1}{2}\right)^{15-5} = 0,09$$

Binomial

Exemplo: Jogue uma moeda 15 vezes, X é o número de caras.

$p = 1/2$, $n = 15$

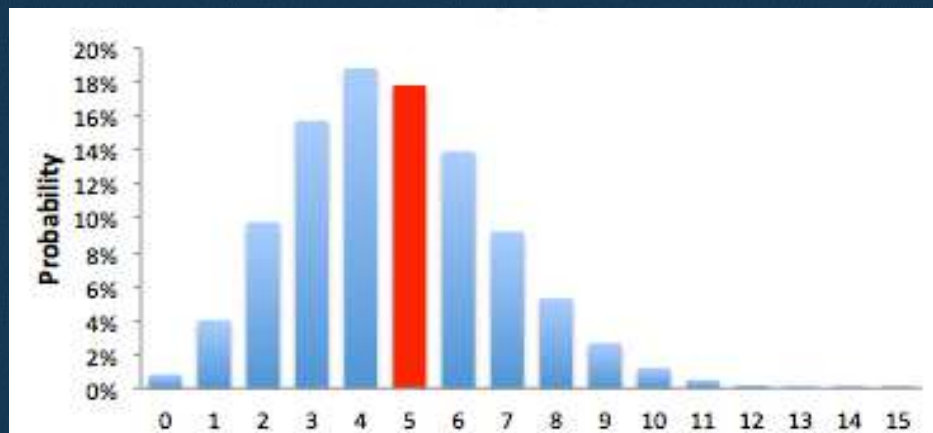
$$P(X=5) = \frac{15!}{5!(15-5)!} \left(\frac{1}{2}\right)^5 \left(1 - \frac{1}{2}\right)^{15-5} = 0,09$$

$$P(X < 5) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$$

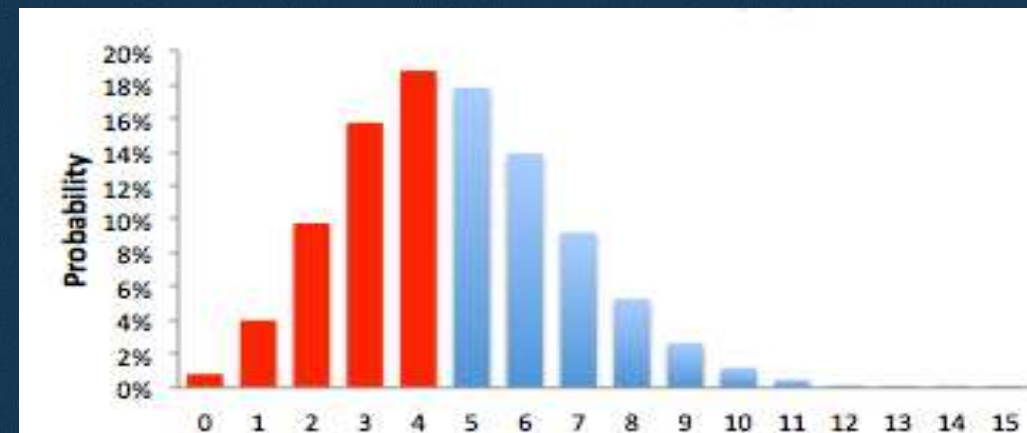
$$P(X > 5) = 1 - P(X \leq 5)$$

Binomial

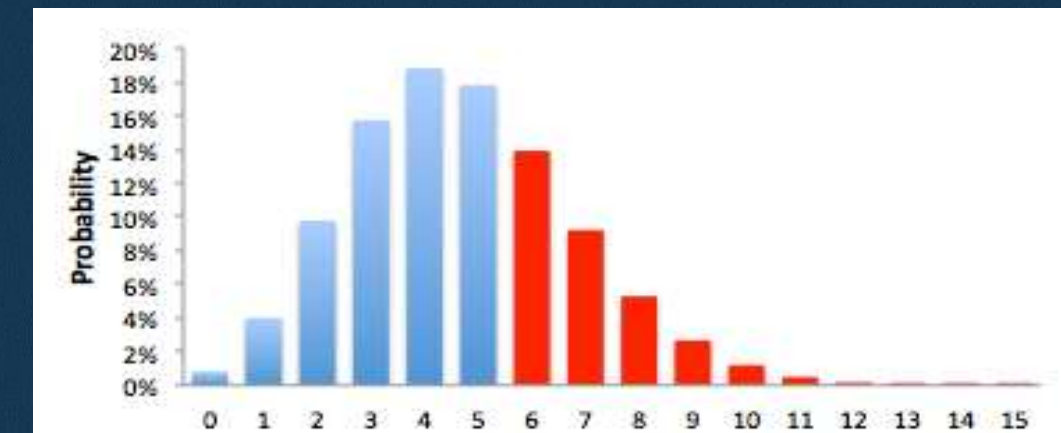
$$P(X = 5)$$



$$P(X < 5)$$



$$P(X > 5)$$



Exemplo

Suponha que uma fábrica produz peças eletrônicas e que cada peça tem uma probabilidade de 5% de ser defeituosa. A empresa realiza inspeções diárias de 100 peças escolhidas aleatoriamente para verificar a qualidade. Qual a probabilidade de 10 peças serem defeituosas?

- $n = 100$ (o número de peças inspecionadas)
- $p = 0,05$ (a probabilidade de encontrar uma peça defeituosa)

$$P(X=10) = \frac{100!}{10!(100-10)!} 0,05^{10} (1 - 0,05)^{100-10} = 0,02$$

Aplicações

- Testes de Medicamentos: Avaliar a eficácia de um novo medicamento, onde “sucesso” é definido como o paciente respondendo positivamente ao tratamento.
- Marketing: Determinar a taxa de resposta a uma campanha publicitária, onde “sucesso” é definido como um cliente realizando uma compra após ver um anúncio.
- Pesquisa Eleitoral: Estimar o número de eleitores que apoiarão um candidato específico, onde “sucesso” é definido como um eleitor indicando suporte ao candidato.

POISSON

Poisson

Uma variável aleatória binomial conta o número de sucessos em um número fixo de tentativas de Bernoulli.

Por outro lado, uma variável aleatória de Poisson conta os sucessos em um determinado intervalo de tempo ou espaço.

Aplicações

Sucessos em relação ao tempo:

- Número de carros que cruzam a ponte Rio-Niterói entre 9h e 10h em uma segunda de manhã.
- Número de emails recebidos por hora.

Sucessos em relação ao espaço:

- Número de defeitos em um rolo de 50 metros de tecido.
- Número de terremotos em uma região.

Poisson

Para uma variável aleatória de Poisson X , a probabilidade de x sucessos em um determinado intervalo de tempo ou espaço é:

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

Sendo $e = 2,718$, $\lambda = \text{média}$

Exemplo

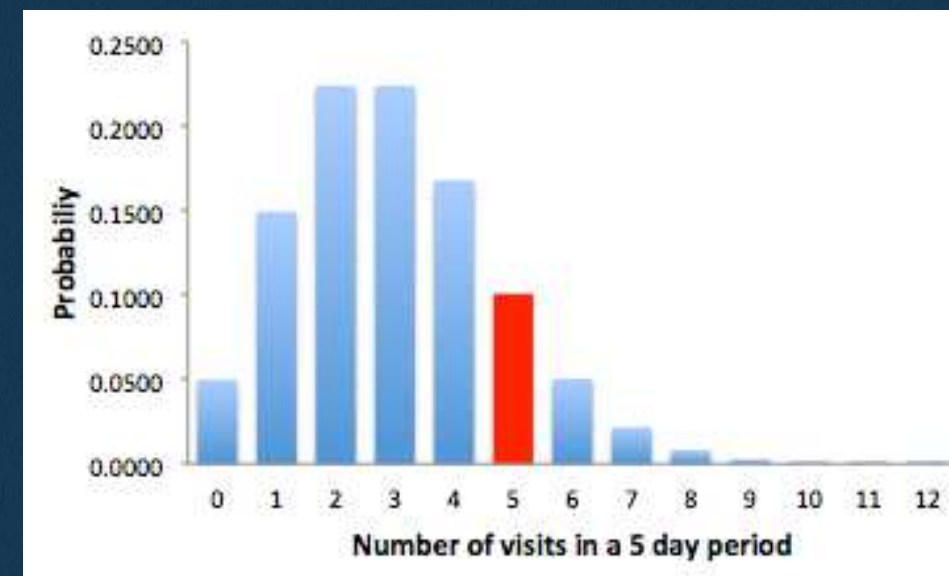


Suponha que um cliente típico da Starbucks faça em média 3 visitas em um período de 5 dias.

Qual é a probabilidade de um cliente visitar a rede cinco vezes em um período de 5 dias?

$$\lambda = 3, x = 5$$

$$P(X = 5) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-3} 3^5}{5!} = 0,1008$$



Exemplo



Suponha que um cliente típico da Starbucks faça em média 3 visitas em um período de 5 dias.

Qual é a probabilidade de um cliente visitar a rede 8 vezes em um período de 10 dias?

Se a média para 5 dias é de 3 visitas, podemos estimar que, para 10 dias, a média é de 6 visitas.

$$\lambda = 6, x = 8$$

$$P(X = 8) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-6} 6^8}{8!} = 0,1033$$

DISTRIBUIÇÕES CONTÍNUAS

Distribuições Contínuas

Uniforme

Normal

Exponencial

Gama

Qui-quadrado

t de Student

F de Snedecor

EXPONENCIAL

Exponencial

A distribuição exponencial é usada para modelar o tempo entre eventos que ocorrem de forma independente e a uma taxa constante.

Ela é caracterizada por uma única taxa λ , que é a taxa média de ocorrências dos eventos por unidade de tempo.

$$P(X < x) = 1 - e^{-\lambda x}$$

Aplicações

- Tempo de Filas: Modelar o tempo de serviço ou atendimento em um sistema de filas, como em um banco ou uma loja.
- Manutenção: Modelar o tempo até a próxima intervenção de manutenção preventiva em um equipamento.
- Intervalos de Chamadas: Modelar o tempo entre chamadas recebidas por um operador em um call center.

Exemplo

Imagine que, em uma estação de ônibus, os ônibus chegam em média a cada 15 minutos. Qual a probabilidade de um passageiro que chega à estação esperar menos do que 10 minutos.

Taxa $\lambda = 1/15$

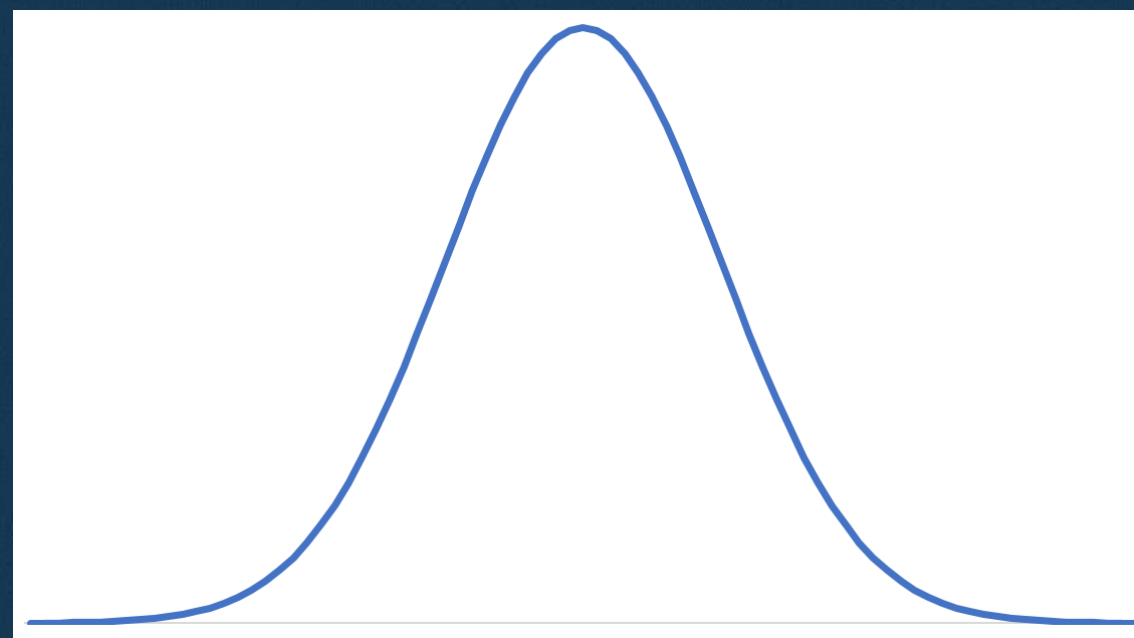
$$P(X < x) = 1 - e^{-\lambda x}$$

$$P(X < 10) = 1 - e^{-\frac{1}{15} \cdot 10} = 0,4866$$

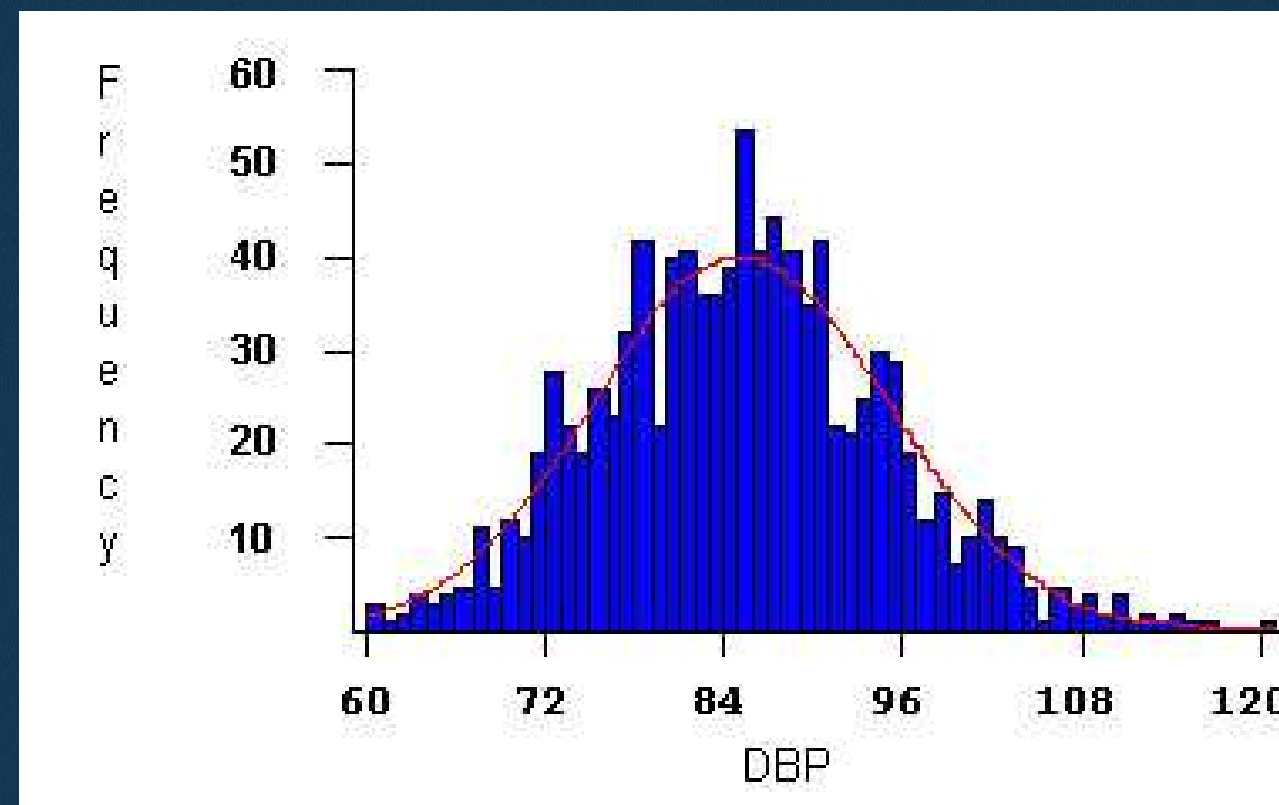
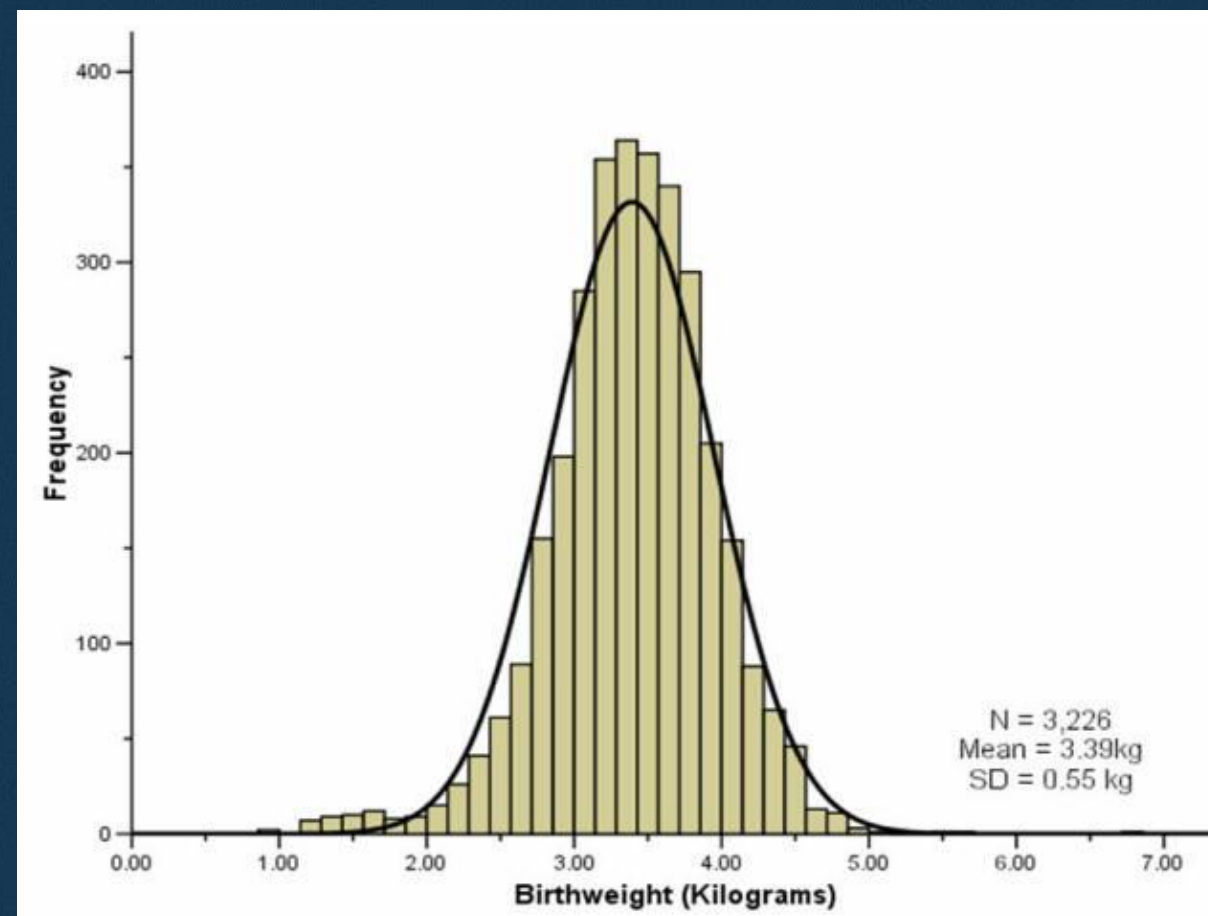
NORMAL

Distribuição Normal

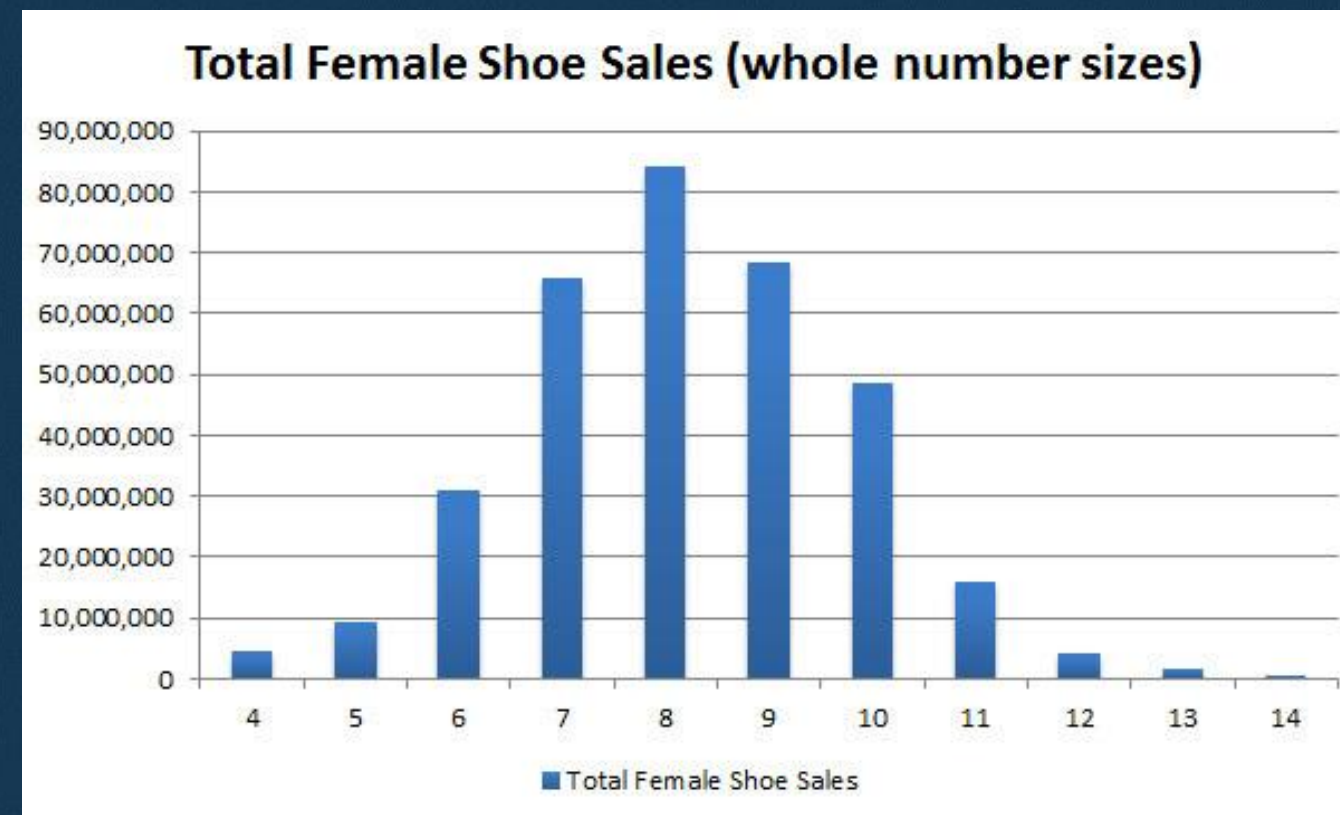
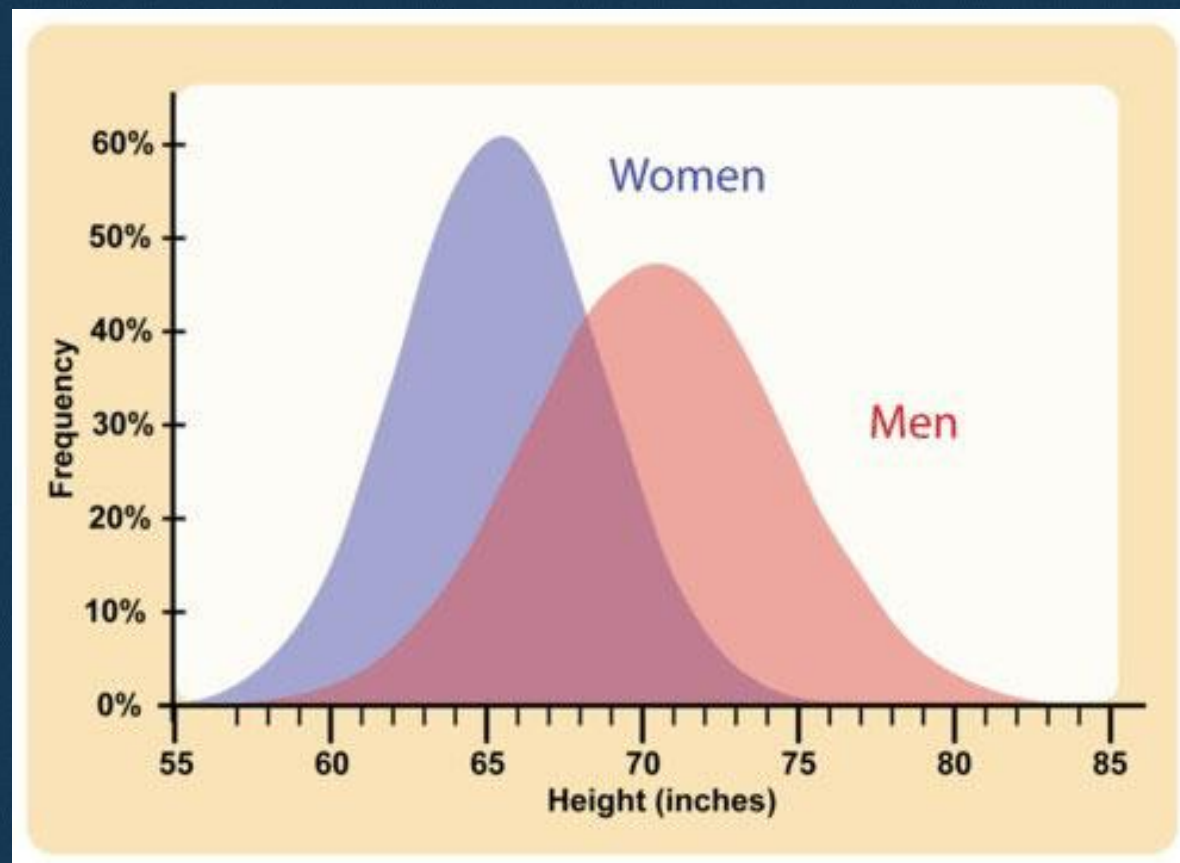
A distribuição Normal (Gaussiana) é uma distribuição simétrica em forma de sino.



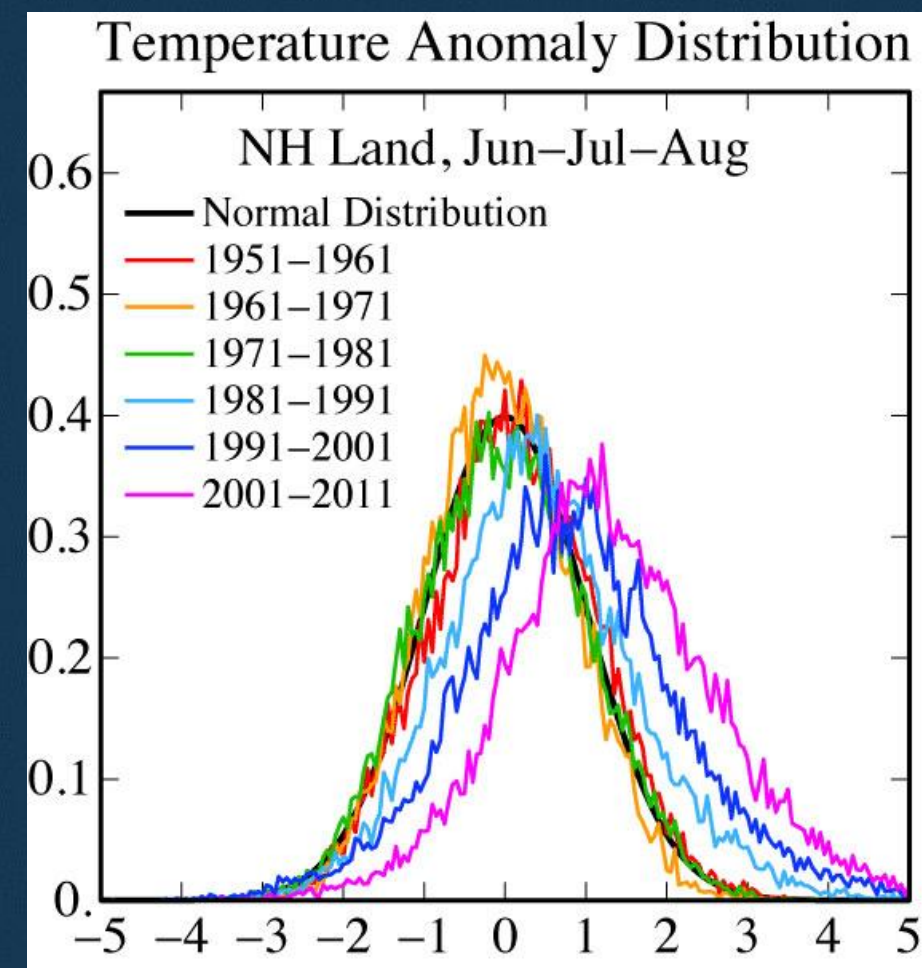
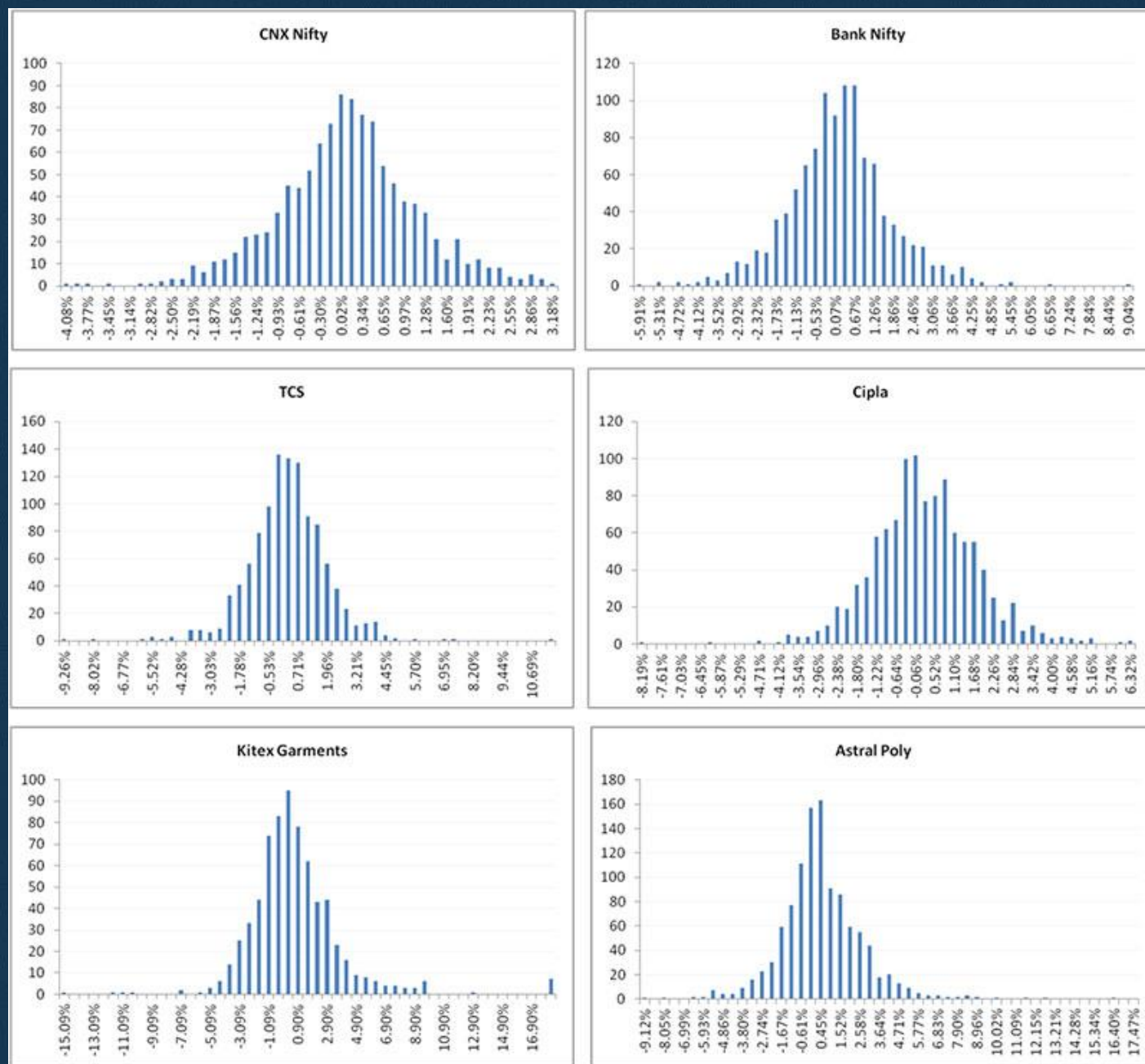
Distribuição Normal



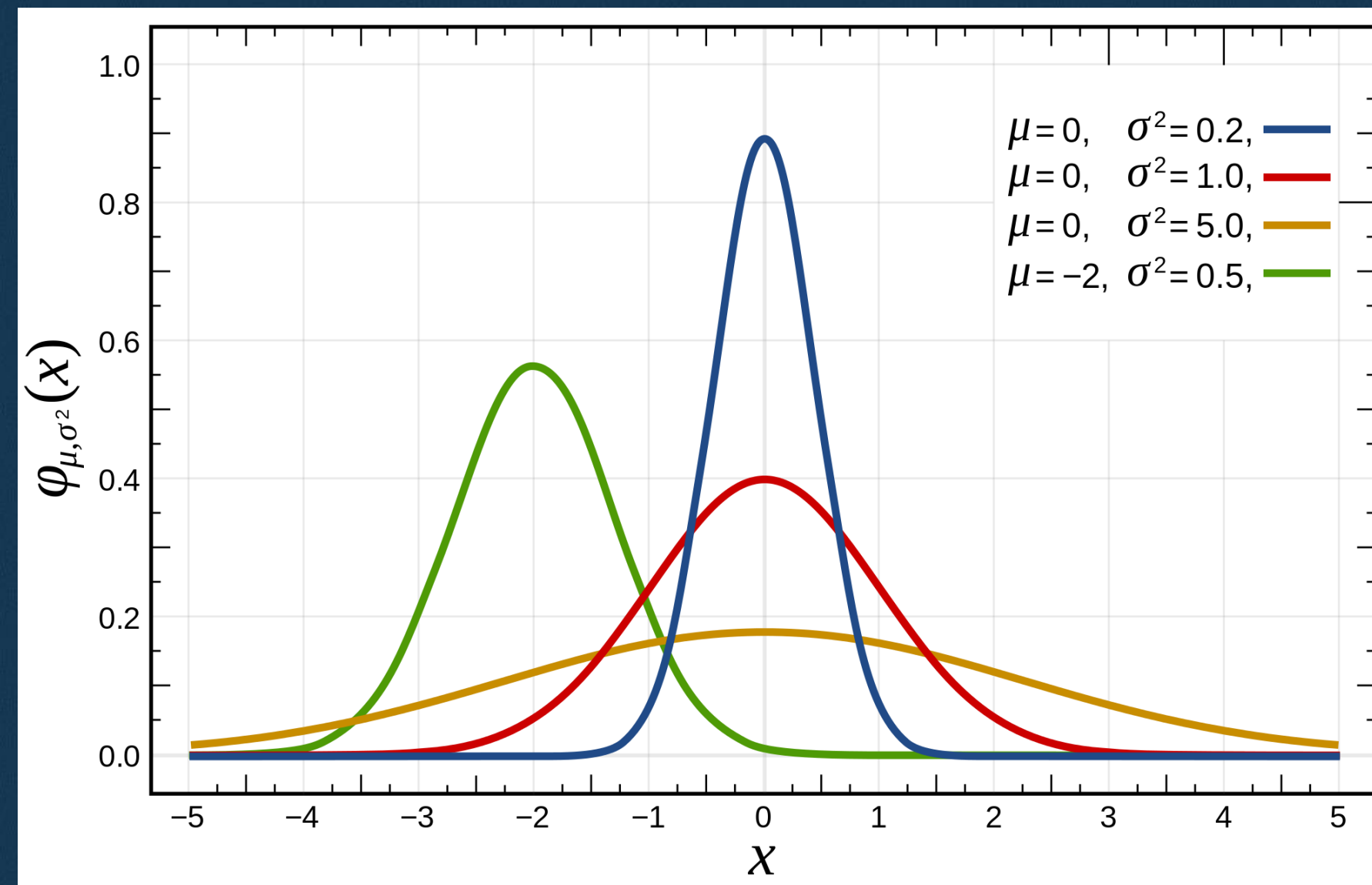
Distribuição Normal



Distribuição Normal



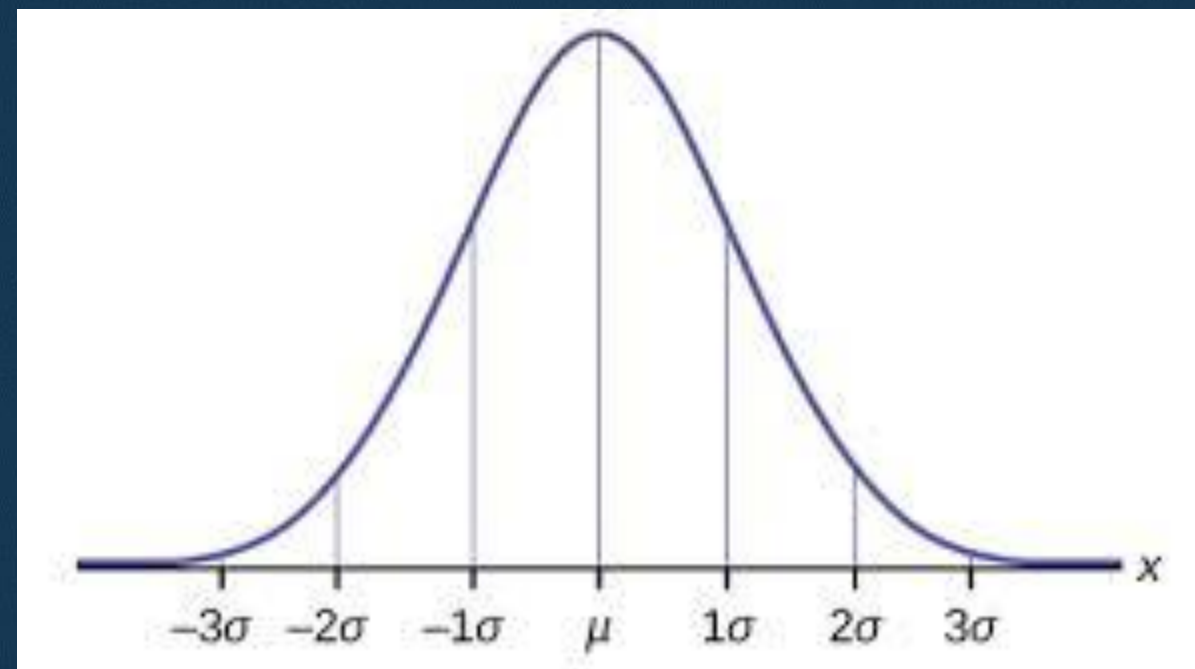
Curvas



Característica

A distribuição normal é descrita por somente dois parâmetros: μ e σ .

- μ é a média da população
- σ é o desvio padrão da população



Exemplo

As pontuações de QI são normalmente distribuídas com média 100 e desvio padrão 15. A variável aleatória X é a pontuação de QI de um adulto selecionado aleatoriamente. Modele a distribuição normal das pontuações de QI.

Fórmula

A distribuição normal segue a seguinte fórmula.
Note que só precisamos de μ e σ .

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

Calculando probabilidades

Para uma distribuição Normal, as probabilidades podem ser calculadas usando:

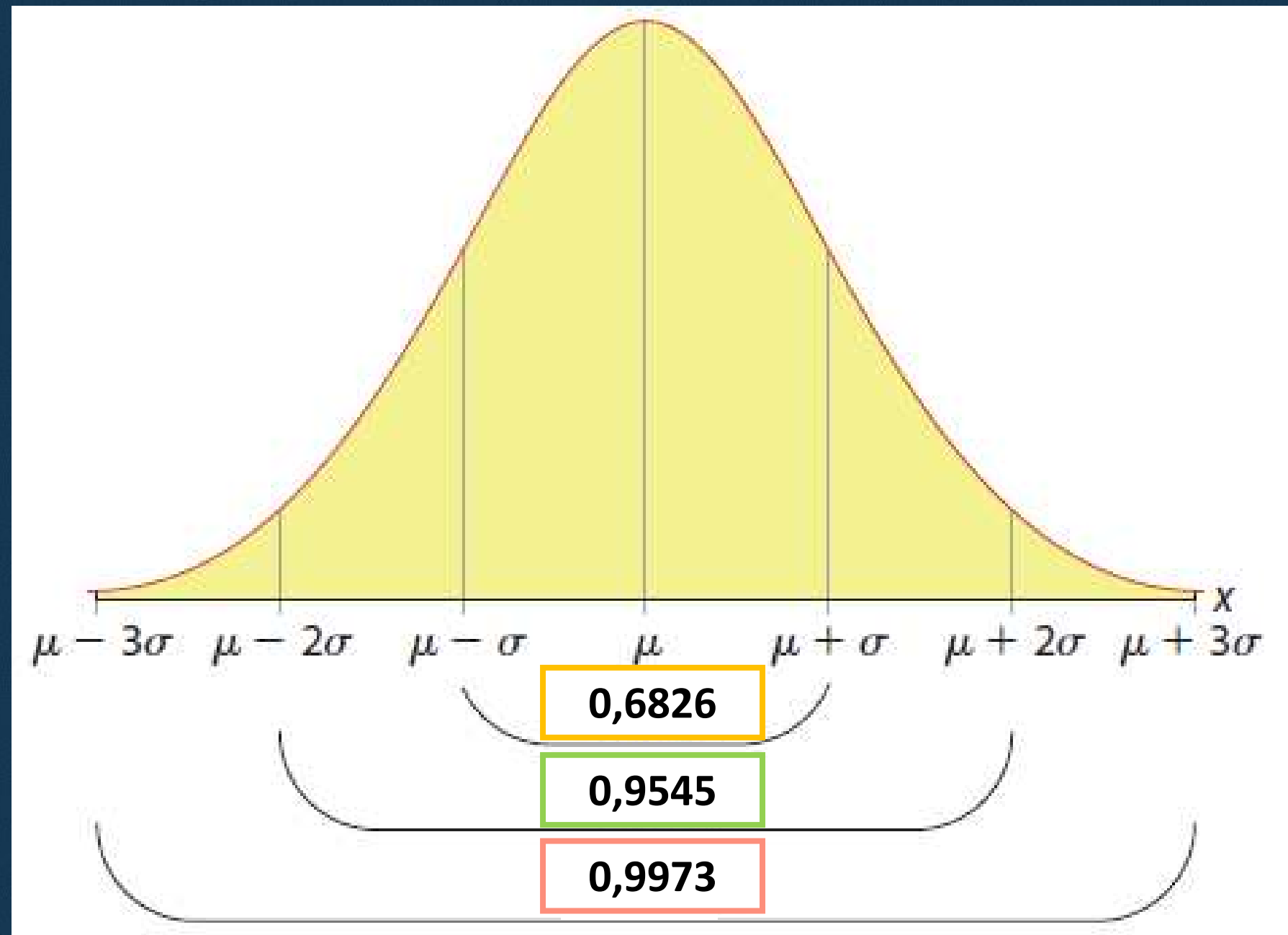
1. A regra empírica
2. A distribuição Normal padrão
3. Diretamente em um software ou linguagem de programação

Regra empírica

A regra empírica diz que, quando temos uma distribuição normal:

- 68% das observações estão entre 1 desvio padrão
- 95% das observações estão entre 2 desvios padrão
- 99,7% das observações estão entre 3 desvios padrão

Regra empírica



Exemplo

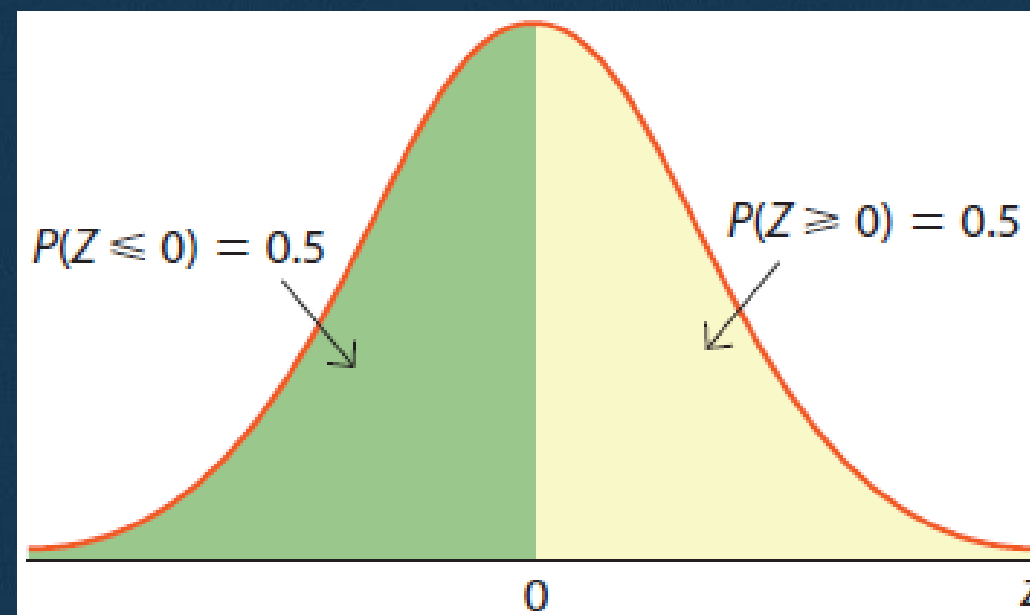
As pontuações de QI são normalmente distribuídas com média 100 e desvio padrão 15. A variável aleatória X é a pontuação de QI de um adulto selecionado aleatoriamente. Encontre a probabilidade de que seu QI seja:

- (a) Entre 85 e 115
- (b) Acima de 115
- (c) Abaixo de 70

Normal Padrão

A distribuição normal padrão (Z) é um caso especial da distribuição normal.

- A média (μ) é igual a zero
- O desvio padrão (σ) é igual a um



Z-score

Seja X uma variável aleatória normal com média μ e desvio padrão σ . Para um determinado valor x , o z-score é:

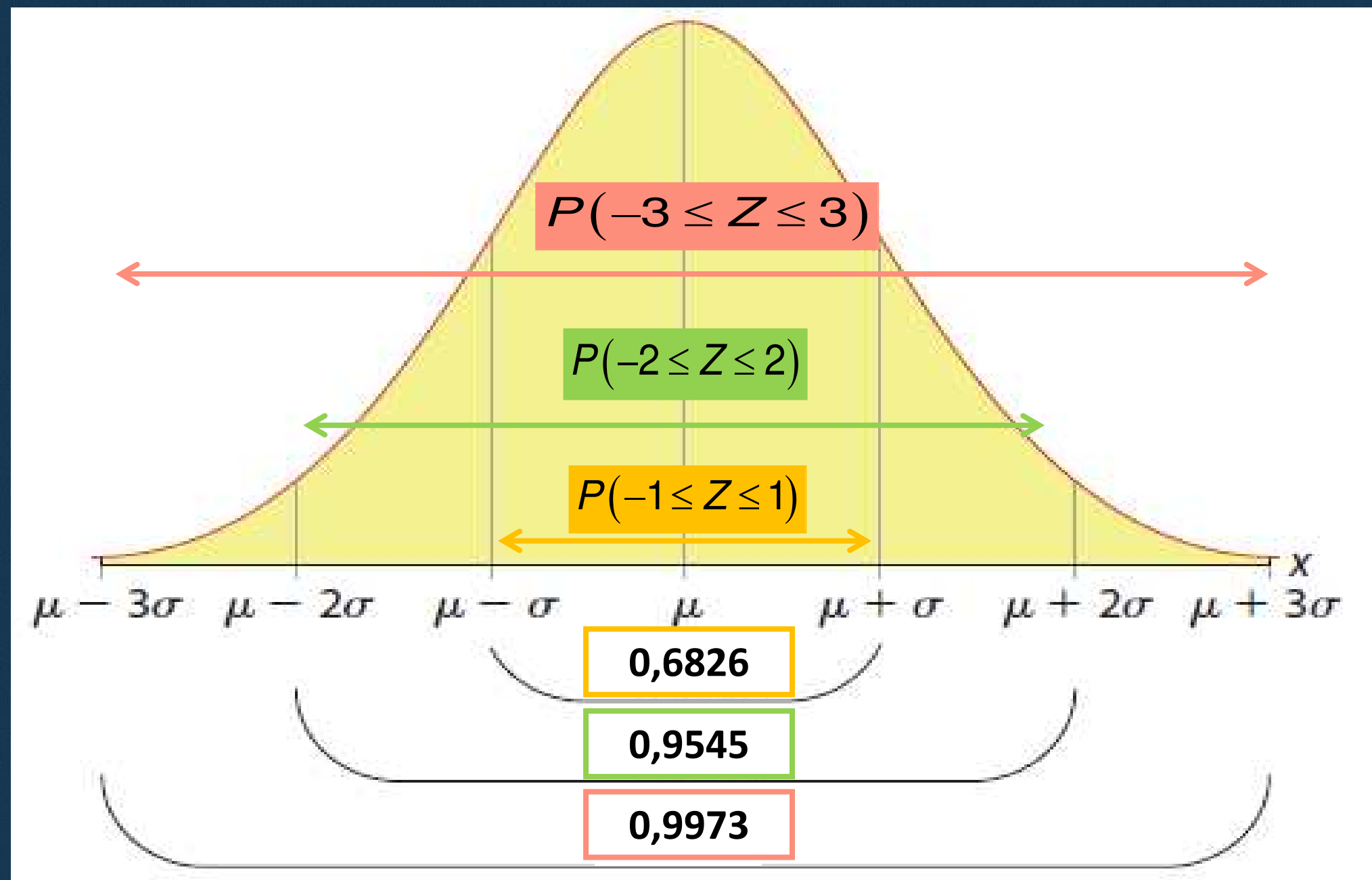
$$z = \frac{x - \mu}{\sigma}$$

Exemplo

As pontuações de QI são normalmente distribuídas com média 100 e desvio padrão 15. A variável aleatória X é a pontuação de QI de um adulto selecionado aleatoriamente. Encontre a probabilidade de que seu QI seja:

- (a) Abaixo de 110
- (b) Entre 100 e 140
- (c) Acima de 80

Regra empírica



z-score & outliers

O z-score é um valor na distribuição normal padrão Z. Ele especifica quantos desvios padrão um valor x cai acima ($z > 0$) ou abaixo ($z < 0$) da média.

Se o z-score for acima de 3, esse valor é considerado um outlier, já que ele tem 0,03% de probabilidade de ocorrência.

O z-score acima de 2 pode também ser considerado um outlier - valores com 5% de probabilidade de ocorrência.

z-score & outliers

