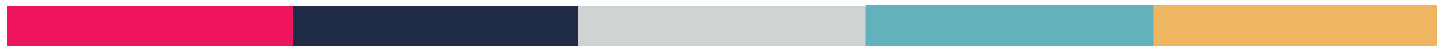


SPTECH

Distribuições

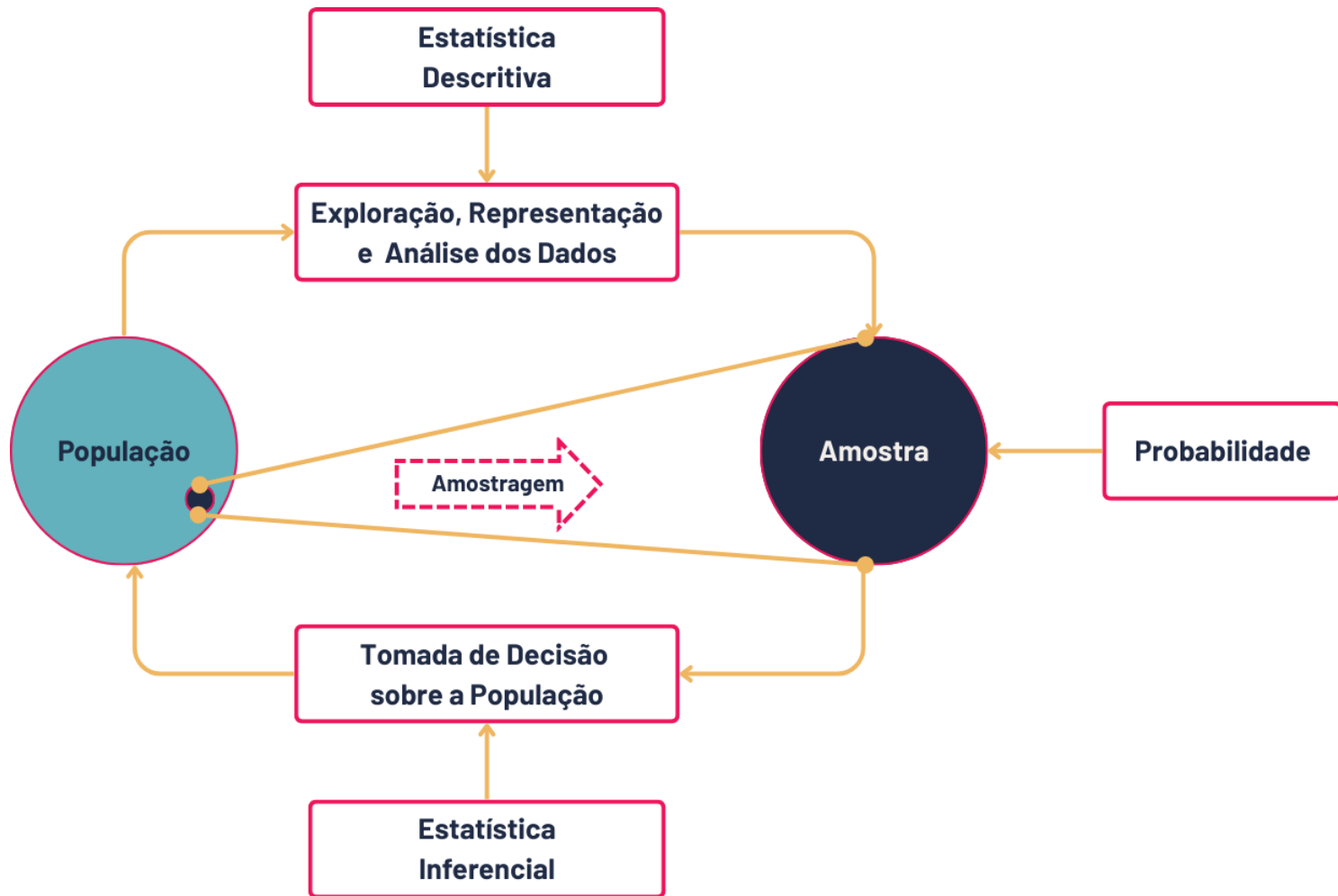
Cálculo Computacional e Estatística





Teoria da Probabilidade

- Ramo da matemática que estuda a incerteza e o acaso.
- Nos ajuda a quantificar a chance de um evento acontecer.
- A base para a estatística, ciência de dados e muitas outras áreas.



Tipos de Distribuições

Contínuas

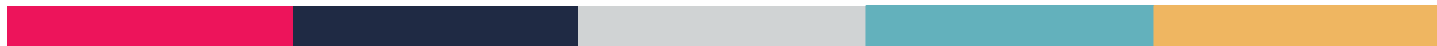
Para eventos que podem ter qualquer valor dentro de um intervalo (ex: altura de uma pessoa, temperatura).

Exemplos: Distribuição Normal, Distribuição Exponencial.

Discretas

Para eventos que só podem ter valores inteiros (ex: número de caras em 10 lançamentos de moeda).

Exemplos: Distribuição Binomial, Distribuição de Poisson.





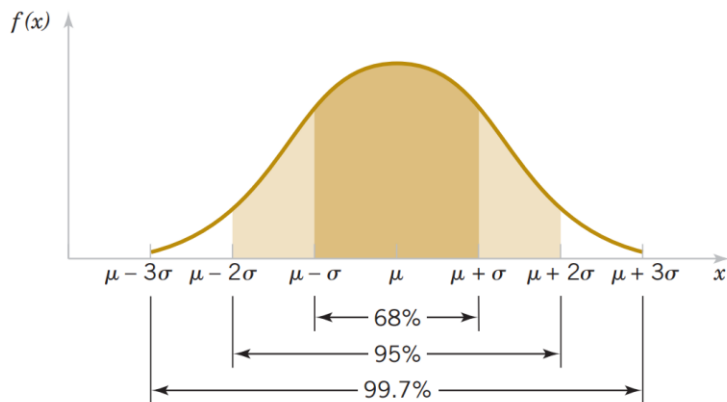
Distribuição Normal (Gaussiana)

A Estrela da Estatística

É a distribuição contínua mais utilizada e conhecida


Se ajusta bem a fenômenos naturais e sociais

Base para muitos testes estatísticos.



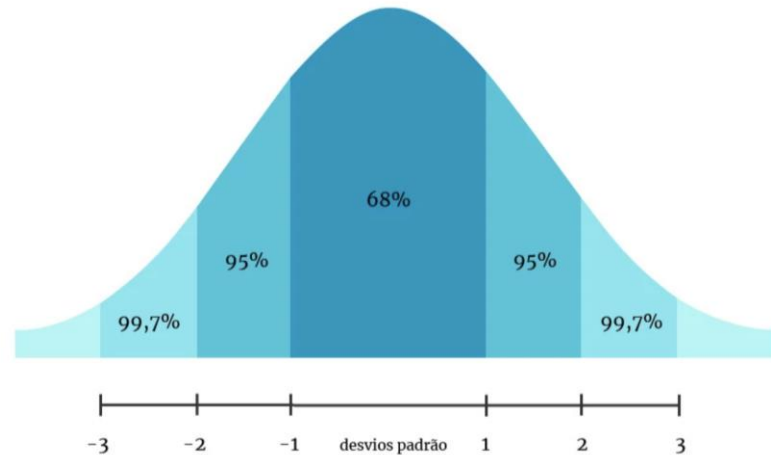
CARACTERÍSTICAS PRINCIPAIS

- **Curva em Forma de Sino:**
Perfeitamente simétrica
- **Média, Mediana e Moda:** Estão todas no centro da curva.
- **Desvio Padrão:** Controla a "largura" do sino. Um desvio padrão pequeno indica dados mais próximos da média; um desvio padrão grande indica dados mais espalhados.


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

Regra Empírica

Aproximadamente **68%** dos dados estão a uma distância de um desvio padrão da média, **95%** a uma distância de dois desvios padrão e **99.7%** a uma distância de três desvios padrão.



⚠ Quando a curva não é perfeita: mais de 68% para curvas "pontudas", menos de 68% para curvas "achatadas"

Quando a distribuição normal é usada?

- ✓ onde a maioria das medições se **agrupa** em torno de um **valor central**.
- **Biologia e Medicina:** Para analisar a altura, peso, pressão arterial e outras características físicas de uma população.
- **Economia e Finanças:** Para modelar retornos de ações ou flutuações de preços.
- **Controle de Qualidade:** Para verificar se um produto (por exemplo, o diâmetro de um parafuso ou o peso de um pacote de biscoitos) está dentro das especificações aceitáveis.
- **Psicologia e Educação:** Para analisar resultados de testes de QI ou pontuações de exames.



Além da Curva de Sino

POR QUE OUTRAS DISTRIBUIÇÕES IMPORTAM?

- Nem todos os fenômenos do mundo real seguem a **Distribuição Normal**.
- Entender outras distribuições nos ajuda a escolher a ferramenta estatística correta e a **tomar melhores decisões**.

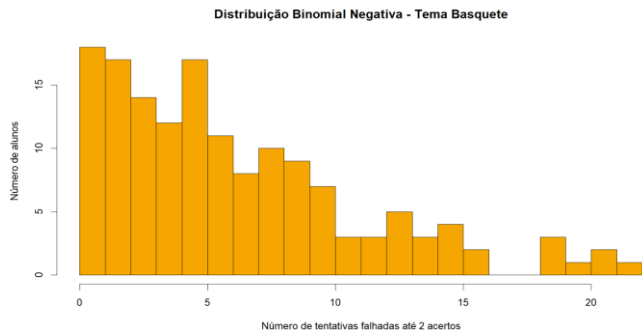
Distribuições que vamos explorar:

Binomial: Para eventos de "sucesso ou fracasso".

Poisson: Para contagem de eventos raros.

Distribuição Binomial

Probabilidade de um evento acontecer um certo número de vezes, quando há apenas dois resultados possíveis: **sucesso** ou **fracasso**



Binomial: Lançar uma moeda 10 vezes, e contar quantas vezes sai cara.

Negativa: Quantas tentativas são necessárias até obter 3 caras?

CARACTERÍSTICAS PRINCIPAIS

- **Número fixo de tentativas:** Você define quantas vezes o evento será repetido. (ex: 10 lançamentos de moeda ou 500 visitas ao site.)
- **Resultados binários:** Cada tentativa tem apenas dois resultados.
- **Probabilidade constante:** A probabilidade de sucesso é a mesma em cada tentativa.
- **Tentativas independentes:** O resultado de uma tentativa não afeta a próxima.

Quando a distribuição binomial é usada?

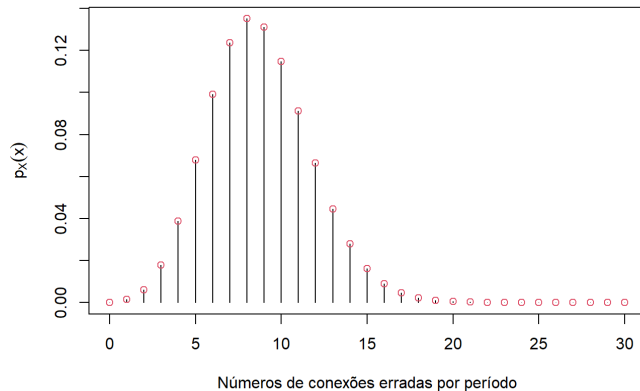


Número **fixo de tentativas** e cada tentativa só puder ter **dois resultados**

- **Taxa de conversão de um anúncio:** Você envia um e-mail de marketing para 500 clientes. A probabilidade de cada cliente abrir o e-mail é de 15%. A binomial pode prever a probabilidade de exatamente 80 clientes abrirem o e-mail.
- **Controle de qualidade:** Uma fábrica produz 1000 peças de um produto. A probabilidade de cada peça ter um defeito é de 0.5%. Você pode usar a binomial para calcular a probabilidade de 3 ou menos peças serem defeituosas.
- **Resultados de pesquisas:** Em uma pesquisa com 200 pessoas, a probabilidade de uma pessoa responder "sim" a uma pergunta é de 70%. A binomial te ajuda a descobrir qual a probabilidade de pelo menos 150 pessoas responderem "sim".

Distribuição Poisson

Conta a **ocorrência de um evento** em um **intervalo** contínuo de **tempo** ou **espaço** (ideal para eventos que são considerados "raros" e acontecem aleatoriamente).



CARACTERÍSTICAS PRINCIPAIS

- **Taxa média constante:** A taxa de ocorrência do evento é a mesma ao longo do intervalo.
- **Eventos independentes:** A ocorrência de um evento não afeta a probabilidade de outro ocorrer.
- **Contagem:** O foco é contar o número de eventos em um período, e não o número de tentativas.

Quando a distribuição poisson é usada?

- ✓ Contando **eventos** em um **intervalo de tempo**/espaço (sem número fixo de "tentativas")
- **Atendimento ao cliente:** O número médio de chamadas que um call center recebe por hora é de 12. A Poisson pode ser usada para estimar a probabilidade de o call center receber exatamente 20 chamadas na próxima hora, ou menos de 5.
- **Análise de tráfego:** Em uma rodovia, a média de acidentes por mês é de 3. Você pode usar a Poisson para calcular a probabilidade de não haver acidentes no próximo mês, ou de haver mais de 5.
- **Monitoramento de servidores:** Um servidor web tem uma média de 2 falhas por semana. A Poisson pode prever a probabilidade de o servidor falhar 4 vezes na próxima semana.

Lembrando...

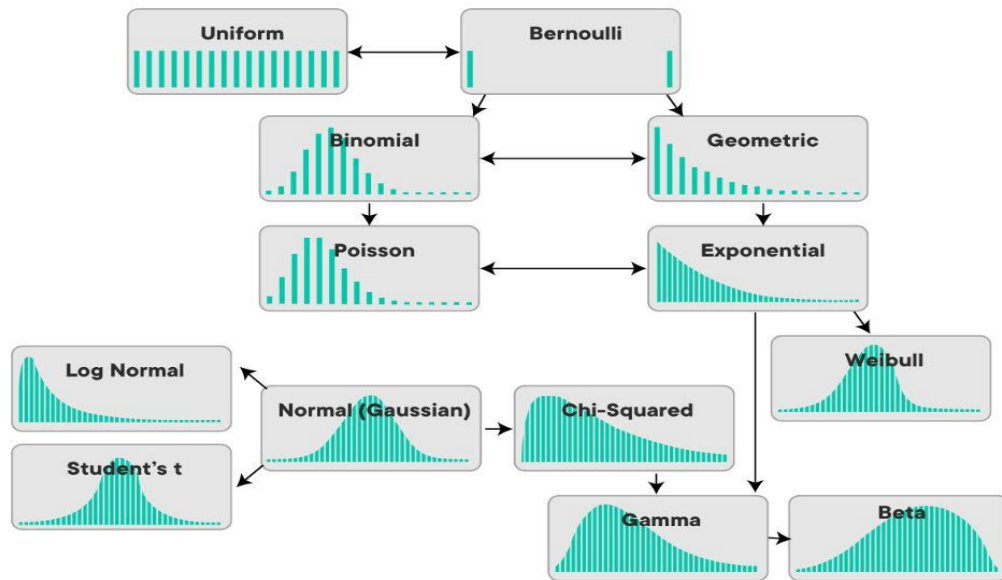
Tipo de Distribuição	Cenário Comum	Aplicação em Negócios
Normal (Gaussiana)	Dados se agrupam em torno da média	Altura, peso, notas de testes, tempo de vida útil de um produto.
Binomial	Contagem de sucessos em um número fixo de tentativas	Taxa de conversão de um anúncio, número de clientes que usam um cupom.
Poisson	Contagem de eventos em um intervalo de tempo	Número de vendas por dia, número de visitas a uma página web por hora.

A ponta do iceberg...

A mensagem principal é: Não se preocupe em dominar todas as distribuições.

Concentre-se em saber

identificar o padrão nos seus dados e, a partir daí, pesquisar e aplicar a distribuição correta com as bibliotecas e ferramentas que você já usa (com Python ou R).





Obrigada!

giuliana.franca@sptech.school