



Introdução à bioestatística utilizando Python

Porque mais um curso de estatística???

PhD Flávio Lichtenstein

Bioinformatics, Systems Biology, and Biostatistics

Instituto Butantan – CENTD - Bioinformática

Janeiro 2026



**Porque a maioria dos cursos de estatística
dissociam Estatística e Ciência.**

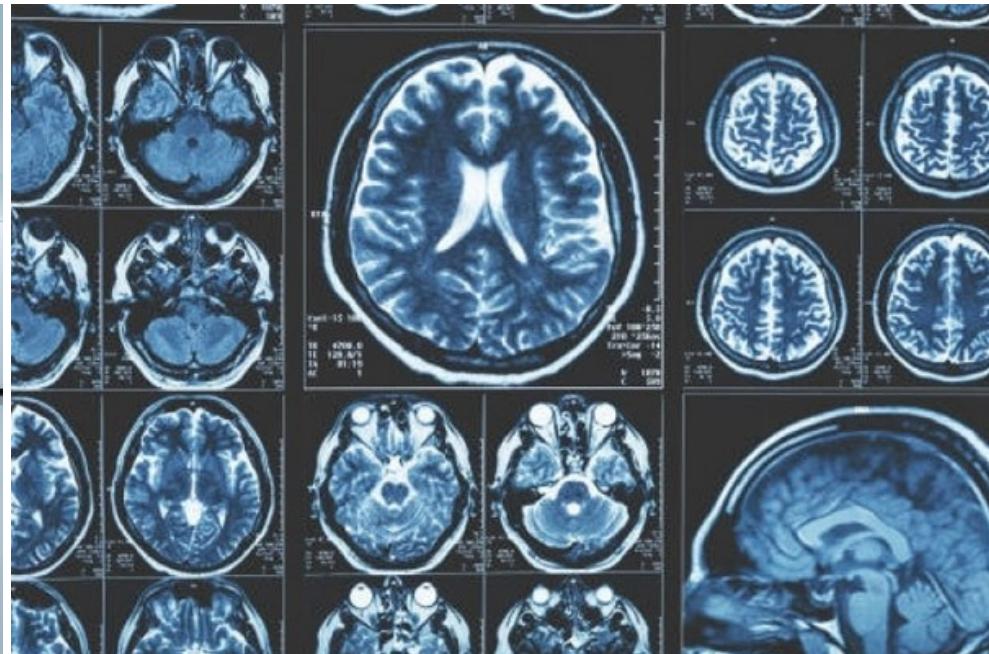
**Aqui vamos fazer uma forte associação
entre Estatística e Ciência Biológica**

**Ciência tem que ser quantitativa (Galileo),
logo não existiria sem matemática,
geometria e estatística.**

Porque a Biologia tem que ser quantitativa apesar de sua complexidade + biologia descritiva



Porque praticamente não existem aparelhos que medem efeitos físico-químicos sem Matemática, Geometria e Estatística



Quantificar tudo!
E conhecer as incertezas
mean(SSD)

mean: media amostral

SSD: desvio padrão amostral

Todas as mensurações têm incertezas

O mundo exato não existe

**Se você esqueceu de colocar o desvio
padrão amostral: você errou!**

O que aprendemos a quantificar?

Quantificação é o processo de transformar dados qualitativos em dados quantitativos, permitindo a análise estatística e a tomada de decisões baseadas em evidências.

Os dados qualitativos são informações descritivas, que podem ser classificadas em categorias ou tipos. Por exemplo, a cor de um carro, o gênero de uma pessoa, ou a opinião de alguém sobre algo.

Os dados quantitativos, por outro lado, são informações numéricas que podem ser contadas ou medidas. Por exemplo, o número de pessoas que compraram um carro, a idade média de uma população, ou a taxa de sucesso de uma campanha publicitária.

A quantificação é importante porque permite que os dados sejam analisados estatisticamente, o que significa que podemos identificar tendências, padrões e relações entre variáveis. Isso pode ser útil para muitos tipos de decisões, como a estratégia de marketing de uma empresa, a política pública de um governo, ou a tomada de decisões médicas.

Além disso, a quantificação também pode ajudar a identificar problemas ou desafios que podem ser resolvidos com a aplicação de soluções baseadas em evidências. Por exemplo, se uma empresa descobre que uma grande proporção de seus clientes não está satisfeita com seu produto, ela pode usar esse dado para desenvolver uma estratégia de melhoria contínua.

No entanto, é importante lembrar que a quantificação não é a única forma de analisar dados. A qualificação também é importante, pois nos permite entender melhor as características e as nuances das informações que estamos analisando. Por exemplo, se uma empresa descobre que uma grande proporção de seus clientes não está satisfeita com seu produto, ela também deve considerar as razões por trás desse sentimento, como a qualidade do produto ou o atendimento ao cliente.

Portanto, a quantificação é uma ferramenta poderosa para a tomada de decisões, mas é importante sempre lembrar que ela deve ser complementada por outras formas de análise e entendimento dos dados.



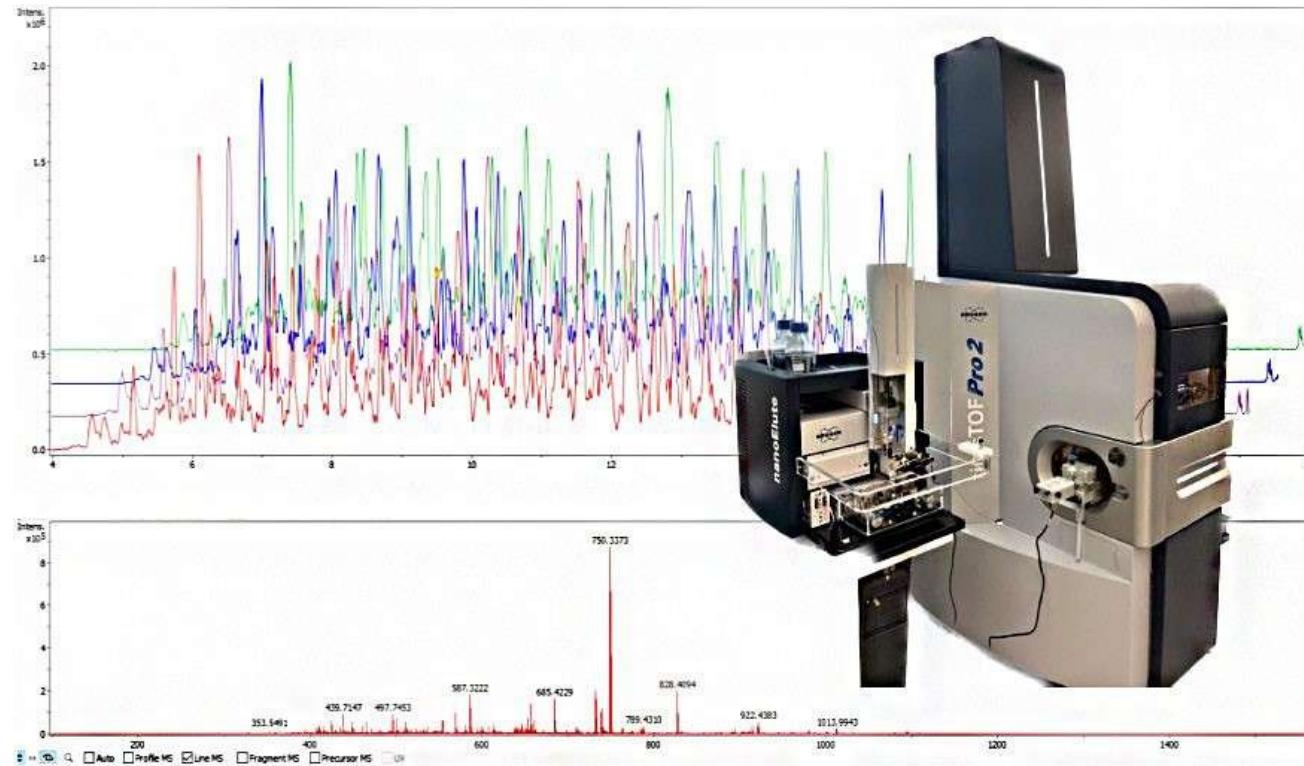
Pressão Sanguínea

Stephen Hales medindo a pressão sanguínea em um cavalo (1705)



Batimentos cardíacos
Eletrocardiograma (ECG)

Espectômetro de Massas Proteômica

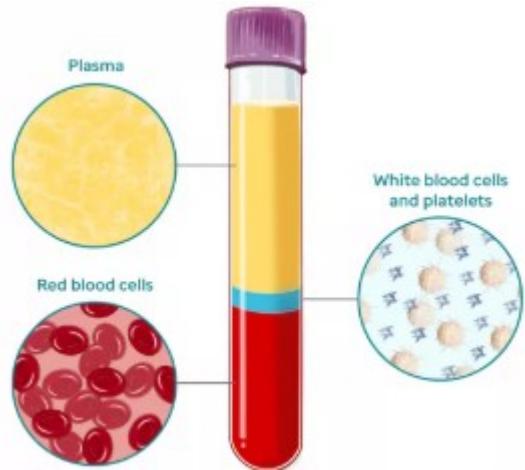


Fusion Orbitrap ETD mass spectrometer.



Contagem de Hemácias

Proporções: RBC + WBC + Plasma



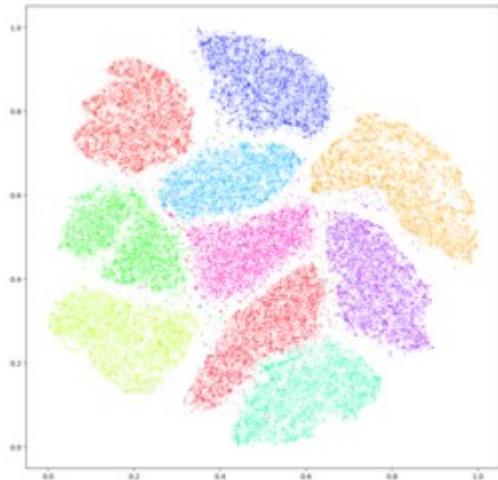
Contagem de todos os componentes da série vermelha (eritrograma) e branca (leucograma)

Ressonância Nuclear Magnética (NMR)



Núcleos como os de H e C apresentam spin (propriedade quântica relativa a momento angular - prótons e nêutrons podem girar no seu próprio eixo), logo funcionam como magnetos. Uma vez alinhado a um campo magnético externo (B) eles se alinham e no relaxamento emitem radiação que pode ser interpretada de acordo com a matéria existente nos núcleos vizinhos.

<https://www.youtube.com/watch?v=cqOewFZPgOM>



Clusterização por t-SNE (ou UMAP)

Quantificação de tipos celulares e seus principais genes diferencialmente expressos (LFC)

Tranciptômica (transcrição de mRNA)

LFC: \log_2 fold-change - \log_2 da variação transcrecional relativa a um controle.

Como medir?

Como quantificar?

- Vírions, Bactérias e Virulência
- Eficácia de uma vacina
- miRNA (micro RNA)
- Biodiversidade na floresta Amazônica

Para cada uma destas perguntas os cientistas criaram métodos, os validaram e publicaram, gerando conhecimento, patentes, ciência (conhecimento oficial) e tecnologia reproduutíveis.

Como funciona?

Como funciona?

- Multiplex (analizador de proteínas)
- Citômetro de Fluxo
- Microscópio confocal
- Microscópio eletrônico de varredura
- High-Content Screening
- Tomógrafo
- Ressonância Magnética

**A ciência tem que ser quantitativa (Galileo),
logo não existiria sem matemática e
estatística.**

Palavras chaves

**A ciência tem que ser quantitativa (Galileo),
logo não existiria sem matemática e
estatística.**

Palavras chaves

**Quantificar e Medir
Reprodutibilidade
Significância e Tamanho do Efeito**

Problemas e Soluções

Falta de reproduzibilidade experimental:

- Declaração explícita de todos materiais e recursos utilizados
- Declaração explícita de todos os métodos utilizados
- Materiais e Métodos validados

Falta de reproduzibilidade resultados:

- Todos os dados medidos e inferidos têm que ser apresentados (supplemental material)
- Caso os dados sejam de pacientes, estes têm que ser anonimizados
- Métodos estatísticos tem que ser declarados
- Resultados têm que ser apresentados com:
 - Media e desvio padrão: mean(SSD) - nunca use mean(\pm SSD), NUNCA!
 - Número amostral
 - Teste de hipótese
 - p-valor (só p-valor NÃO PODE; é insuficiente)
 - intervalo de confiança
- Inferências só podem ser aceitas se seguirem os testes de hipóteses

Problemas e Soluções

Falta de reproduzibilidade computacional:

- Códigos fontes têm que ser de livre acesso (depositados no GitHub, Kaggle ou Zenodo)
- Códigos fontes e respectivos dados têm que ser reproduzíveis
 - Python, R, C, etc
 - Docker
- Dados similares de outros laboratórios têm que ser reproduzíveis com os códigos fontes apresentados

Utilização de Inteligência Artificial (IA):

- IA não pode ser um referência científica
- Nada impede do Pesquisador usar IA + PubMed ou Web of Science:
 - Caso o pesquisador utilize IA, tem que mostrar que todas as anotações apresentadas pela IA são reais, válidas e constam em referências validadas por pares
- IA não pode fazer hipóteses, e sim o pesquisador
- IA não pode fazer inferência, e sim o pesquisador

Porque mais um curso de estatística???

Em caso de dúvidas escreva no chat para o moderador

Algumas perguntas serão respondidas e depositadas no Q&A
Possivelmente, devido ao grande número de alunos não conseguiremos responder a todas as perguntas, mas tentaremos.

PhD Flávio Lichtenstein

Bioinformatics, Systems Biology, and Biostatistics

Instituto Butantan – CENTD - Bioinformática

Janeiro 2026

