

ME623 - Planejamento e Pesquisa

Parte 01

Tatiana Benaglia - 2024S1

Teoria Estatística de Experimentação

Ciência é um processo de organização do conhecimento através de explicações testáveis sobre o universo.

Testabilidade é uma propriedade de uma teoria: se é possível conceber um experimento replicável em que a teoria pode ser negada, então ela é testável, isto é, uma teoria científica.

Karl Popper: testabilidade é o que distingue as ciências de outras áreas do conhecimento, como matemática e filosofia (isto é, a metodologia de descoberta).

Richard Feynman e a essência da ciência em 60 segundos:

<https://www.youtube.com/embed/LlxvQMhttq4>

Estatística: Ciência voltada a tirar conclusões válidas a partir de inferências de dados coletados em estudos observacionais e experimentais.



Estudo Experimental x Estudo Observacional

- **Estudo Observacional:** o pesquisador é um observador passivo que anota variáveis/fatores de interesse. Esse tipo de estudo mostra apenas associação (não causa) entre fatores de interesse e a variável resposta.

Exemplo: Estudos Epidemiológicos nos quais podemos verificar quais fatores de risco estão associados a uma certa doença. Um exemplo seria: a partir de prontuários de um hospital, verificar se existe associação entre fumo e câncer de pulmão.

- **Estudo Experimental:** o pesquisador manipula as variáveis/fatores e avalia seus efeitos na variável resposta. Esse tipo de estudo pode mostrar causa entre fatores e a variável resposta.

Exemplo: Ensaios Clínicos nos quais podemos testar se um medicamento é efetivo no tratamento de uma certa doença.

Nesse curso focaremos nos **estudos experimentais**. A palavra *experimento* será usada neste curso para designar uma investigação onde o sistema em estudo está sob o controle do investigador.

Planejamento/Delineamento de Experimentos

Na Estatística, é uma área de estudos que desenvolve técnicas de planejamento e análise de experimentos.

Em inglês chama-se *Design of Experiments* (DOE). Usaremos esta sigla ao longo do curso em algumas situações.

Há atualmente todo um arsenal de técnicas, com vários níveis de sofisticação e uma quantidade enorme de livros sobre o assunto.

Nesse curso, iremos aprender os tipos de experimentos mais conhecidos e utilizados nas mais diversas áreas.



Planejamento de Experimentos

É o processo de planejar e conduzir um ensaio/experimento, incluindo sua implementação, de modo que seja possível coletar dados que possam ser analisados, usando metodologias estatísticas apropriadas, e que conduzam a conclusões válidas e objetivas.

Algumas aplicações em diversas áreas

- **Agricultura:** comparar vários tipos de fertilizantes e tipos de solo na produção de laranja.
- **Indústria:** reduzir variabilidade, tempo e custo na produção de um componente eletrônico.
- **Biologia:** comparar a resistência de diferentes tipos bactérias à temperaturas extremas.
- **Saúde:** Estabelecer a dosagem ótima de um antibiótico no tratamento de uma infecção.
- **Economia:** comparar o desempenho de carteiras de investimento.
- **Física:** estudar o impacto de fontes radioativas em certos ambientes.
- **Psicologia:** Estudar se métodos de ensino diferentes produzem benefícios aos alunos.

Objetivos do Curso

- Identificar sob quais condições podemos falar de causalidade (experimento planejado) contra apenas associação (experimento observacional).
- Identificar as fontes de variação na resposta de um experimento (fatores/tratamentos, efeitos ambientais, confundimentos, dependências, erro) e planejar a melhor maneira de controlá-las (**antes** de fazer o experimento). Saber como essas complicações afetam a resposta.
 - Ferramentas essenciais: aleatorização e replicação (e blocagem/estratificação, quando não temos saída).
- Estudar metodologias experimentais mais avançadas para experimentos complexos (quadrados latinos, frações de experimentos fatoriais, estudos longitudinais etc.). Em geral escolhemos experimentos simples, e usamos experimentos complexos só quando há necessidade!

Experimentos Planejados

Lembre-se que:

“Todos os experimentos são planejados. A diferença é que alguns são mal planejados e outros são bem planejados.”

A análise estatística dos resultados de um experimento é uma consequência natural do planejamento, mas objetivo primário do curso é **planejar bons experimentos**.



Experimento

Experimento: é um teste (ou série de testes) no qual mudanças propositalmente são feitas nos inputs de um processo, de modo que se possa observar e identificar as razões das possíveis mudanças observadas na variável resposta.

Objetivos de um Experimento

- Identificar as principais causas de variação na variável resposta.
- Encontrar as condições para que seja atingida uma resposta máxima/mínima.
- Comparar as respostas obtidas em níveis diferentes configurações das variáveis controláveis.
- Obter um modelo que possa ser usado para prever observações futuras.



Checklist de um Experimento Planejado

Vamos definir alguns desses conceitos ainda (Dean e Voss, Página 8).

- 1 Defina os objetivos do experimento.
- 2 Identifique todas as fontes de variação incluindo:
 - fatores e seus níveis (tratamentos).
 - unidades experimentais.
 - efeitos de blocos, efeitos de erro, e covariáveis.
- 3 Escolha uma regra para alocar unidades experimentais aos tratamentos.
- 4 Escolha como serão feitas suas medições (isto é, a **variável resposta**).
- 5 Execute um experimento piloto (testar se experimento funciona, obter informação sobre a variância do erro).
- 6 Especifique um modelo.
- 7 Escolha a ferramenta de análise: quais testes fazer, que procedimentos pós-teste (falaremos sobre descobertas falsas!).
- 8 Calcule o número de observações ou replicações (“amostras” no jargão aplicado).
- 9 Revisar e executar.

Checklist de um Experimento Planejado

Vamos definir alguns desses conceitos ainda (Dean e Voss, Página 8).

- 1 Defina os objetivos do experimento.
- 2 Identifique todas as fontes de variação incluindo:
 - **fatores** e seus níveis (**tratamentos**).
 - **unidades experimentais**.
 - efeitos de **blocos**, efeitos de **erro**, e **covariáveis**.
- 3 Escolha uma regra para alocar unidades experimentais aos tratamentos.
- 4 Escolha como serão feitas suas medições (isto é, a **variável resposta**).
- 5 Execute um experimento piloto (testar se experimento funciona, obter informação sobre a **variância do erro**).
- 6 Especifique um modelo.
- 7 Escolha a ferramenta de análise: quais testes fazer, que procedimentos pós-teste (falaremos sobre descobertas falsas!).
- 8 Calcule o número de observações ou **replicações** (“amostras” no jargão aplicado).
- 9 Revisar e executar.

Unidade Experimental (1 de 2)

Unidade experimental (UE): é a unidade física ou biológica que recebe cada um dos tratamentos, e na qual se irá quantificar o efeito dos fatores. É o menor nível do experimento, em que um tratamento pode ser aplicado.

Nos exemplos a seguir:

- Num ensaio clínico para determinar a imunização por uma vacina: o **indivíduo** é a UE.
- Na agricultura, para comparar o efeito de fertilizantes, o **pedaço de terra** deve ser a UE.
- Na avicultura, com objetivo de determinar o crescimento de galinhas a partir de um aditivo de vitaminas na ração:
 - Se você tem uma galinha em cada gaiola, com sua própria ração, cada galinha é uma unidade experimental.
 - Se as galinhas estão soltas, mas separadas em pequenos grupos com acesso a uma única tigela de ração: a unidade experimental é o grupo.

Veja que no último caso, a maneira como o tratamento é aplicado (gaiola vs. tigela) é o que determina a unidade experimental.

Unidade Experimental (2 de 2)

- A unidade experimental deve ser representativa da população. Ter 500 pacientes em um ensaio clínico é inútil se todos eles são homens de 20 a 30 anos. No máximo, seu experimento identificará o efeito no remédio em homens de 20 a 30 anos (se tudo for feito corretamente!).
 - Planejamento de experimentos e amostragem estão intrinsecamente ligadas.
- Agricultura: se eu tenho um pedaço de terra, mas meus tratamentos (fertilizante ou sem fertilizante) precisam ser aplicados através do sistema de irrigação, então a região coberta por cada mangueira de irrigação corresponde a uma unidade experimental (limitação física: mesmo que tenha centenas de plantas).

Fator e Tratamento (1 de 4)

Fator: variável independente controlada pelo experimentador que pode influenciar os valores observados na variável resposta. São os efeitos que temos interesse em comparar nas unidades experimentais.

Em geral definimos **níveis** de um fator e fazemos a comparação entre níveis. Quando temos mais de um fator de interesse, cada combinação entre os níveis dos fatores é chamada de **tratamento**.

As unidades experimentais devem ser alocadas aos tratamentos **aleatoriamente**.

- Se meu experimento busca comparar a resistência à pressão de tijolos, todos feitos da mesma argila, todos moldados do mesmo tamanho, todos cozidos por 6 horas e resfriados naturalmente, um fator pode ser a temperatura de cocção: digamos, 250C, 280C, 310C.
 - Notem que eu listei algumas fontes de variabilidade (tipo de material, molde etc.) mas como estão fixos, não considero como fator de interesse.
 - Estou presumindo que tenho quantos fornos precisar, com espaço suficiente, temperatura controlada etc.
 - Se tudo estiver controlado, e eu escolher **aleatoriamente** qual a temperatura de cocção de cada um dos tijolos, tenho um experimento planejado para comparar o efeito do fator **temperatura**.

Fator e Tratamento (2 de 4)

Se eu tenho um ensaio clínico para comparar uma droga experimental para controle de colesterol, 12 pacientes, 6 pílulas com a droga (em dose fixa) e 6 pílulas com placebo; e 8 pacientes são mulheres e 4 são homens.

- Se eu aleatorizar as pílulas (de preferência com aleatorização duplamente cega), então o fator **droga** tem dois níveis (**tratamentos**): dose e controle (placebo).
- O sexo dos participantes **não** é um fator, porque **não pode ser aleatorizado**. Note que o sexo pode ser uma fonte de variação! Neste caso, é uma fonte de variação descontrolada.

É importante observar que a resposta de um indivíduo a um tratamento não é “aleatória”, mas sim complexa.

Fator e Tratamento (3 de 4)

Se eu estou testando dose versus controle, meu fator droga tem dois níveis/tratamentos. E se quisermos comparar a resposta dos pacientes a duas pílulas diferentes?

- Fatores podem ser numéricos contínuos (e.g. temperatura do forno, volume de água na receita, log-dose de veneno), discretos (número de sementes por hectare) ou simplesmente categóricos (tinta com verniz vs. tinta sem verniz).
- Temos que decidir quais níveis vamos investigar.
 - Limitações naturais/práticas.
 - Amostras piloto podem sugerir níveis interessantes.
 - Experimentos planejados são iterativos: se temperaturas de 250C e 280C mostraram diferenças na amostra piloto, mas 280C e 310C ficaram iguais, talvez seja o caso de olhar 250C, 265C e 280C num segundo experimento.
 - Procedimento sistemático de execução e otimização da resposta em função de níveis do tratamento: metodologia de superfície de resposta (veremos mais adiante).
- Tipos de Fatores: **Fixos** (número fixo de níveis) ou **Aleatórios** (amostra aleatória dos possíveis níveis).

Fator e Tratamento (4 de 4)

Podemos usar a flexibilidade para determinarmos os fatores a nosso favor.

- Em aulas futuras vamos considerar experimentos com dois ou mais fatores. Por exemplo, nos tijolos, podemos estar interessados em temperatura (250C, 280C e 310C) e porcentagem de areia misturada com a argila (0%, 5%).
- Suponha que seja possível aleatorizar todas as combinações de níveis dos fatores (tratamentos), na forma:

| | $T = 250C$ | $T = 280C$ | $T = 310C$ |
|-----------|---------------|---------------|---------------|
| $A = 0\%$ | $(0\%, 250C)$ | $(0\%, 280C)$ | $(0\%, 310C)$ |
| $A = 5\%$ | $(5\%, 250C)$ | $(5\%, 280C)$ | $(5\%, 310C)$ |

Então podemos tratar o experimento como um fator (Areia \times Temperatura) de seis níveis.

- Quando todos os tratamentos, ou seja, todas as combinações de níveis de diferentes fatores estão em um experimento, isso é chamado de **Experimento Fatorial**.

Replicação (1 de 2)

Replicação: é uma **repetição independente** de um nível do tratamento, ou seja, a atribuição do mesmo tratamento a várias UEs.

A palavra **independente** é chave. Repetições independentes são necessárias pra gente conseguir estimar a **variabilidade do erro**.

Duas propriedades importantes:

- Permite estimar o erro experimental, que será usado nos testes de hipóteses.
- Se a média amostral é usada para estimar o efeito de um fator, a replicação produz estimativas mais precisas desse efeito.

Replicação, medida repetida e repetição

- Em um ensaio clínico, se eu tenho um indivíduo tomando a droga A, e avalio seu colesterol hoje e daqui 10 dias, eu tenho uma **medida repetida**, não uma replicação. A variância de medidas repetidas é uma medida da variação do processo ao longo do tempo.
- Se eu meço o colesterol do indivíduo (com amostra de sangue), e imediatamente meço o colesterol, eu tenho uma **repetição**, não uma replicação. A repetição pode ser útil quando estamos examinando a variabilidade do instrumento (incerteza do diagnóstico), mas definitivamente não é estimativa para a variabilidade do erro.

Portanto, medida repetida ou repetição não é replicação!

Replicação (2 de 2)

Pra ter uma estimativa da variância do erro, eu preciso de **replicações**, observações independentes. Dois indivíduos (com perfil absolutamente parecido mas nenhuma relação entre si) podem ser aproximadamente tratados como independentes – mas independência verdadeira é difícil de garantir.

- Replicações são repetições independentes. Se eu aplico um remédio na raiz de uma planta (tratamento), e decidir medir o efeito do remédio após alguns dias nas folhas de um galho, e medir depois noutro galho da mesma planta, eu não tenho resultados independentes: de fato, como o nível em que variamos o tratamento é a raiz, eu só tenho medidas repetidas de uma mesma unidade experimental.
- Conceitualmente, o livro do Box, Hunter & Hunter (2005) distingue entre replicações do experimento inteiro e replicações das unidades experimentais. É interessante ler a respeito. Na prática, podemos ficar satisfeitos com a garantia de independência.

Aleatorização

Se temos um experimento em que tudo o que podemos controlar está controlado (i.e. fixo, inalterado), e decidimos quais níveis do fator iremos investigar, precisamos de uma última ferramenta: a **aleatorização**.

Aleatorização: É o processo pelo qual a alocação dos tratamentos às UEs e/ou a ordem dos ensaios do experimento são determinados de forma completamente aleatória.

- A aleatorização garante que não estamos introduzindo nenhum **viés de seleção**, ou seja, elimina o viés sistemático causado por fatores externos .
- Também faz com que a suposição de erros independentes seja válida.
- A **aleatorização** deve ser repetida para as **replicações**. Em conjunto, replicação e aleatorização garantem uma relação **causal** (e não meramente associativa) entre o tratamento e a resposta. É possível provar que isso é verdade, mas dá certo trabalho.
- Outra maneira de pensar em causalidade é que a combinação de **aleatorização** e **replicação** num experimento onde tudo o mais está **controlado**, garante que efeitos que nós não conhecemos (e consequentemente não controlamos) não vai interferir no resultado.

Blocos (1 de 2)

Blocos são o último conceito chave da aula de hoje, mas como só vamos considerar experimentos com blocos mais adiante, vamos falar rapidamente e revisamos no futuro.

Blocos: são agrupamentos das unidades experimentais, tais que todas as UEs dentro de um nível do bloco encontram-se em condições similares entre si, ou seja, são as mais homogêneas possíveis, e distintas das UEs de outro nível do bloco.

Blocagem: técnica que permite reduzir ou eliminar a variabilidade transmitida por fatores perturbadores, reduzindo assim a variância do erro. Como não queremos determinar o efeito dos blocos, normalmente não aleatorizamos os blocos.

Blocos (2 de 2)

Exemplo 1: Se tenho 7 ratos, mas 3 vieram de uma mesma ninhada (fêmea A), e os outros 4 de uma mesma ninhada (fêmea B), tenho um bloco com 2 níveis (fêmea A, B) de tamanhos 3 e 4.

Exemplo 2: Se eu tenho uma fazenda com 6 plots, mas 2 deles encontram-se na parte superior do terreno (pouca acúmulo de matéria orgânica, muito sol), 2 deles encontram-se em colinas (boa drenagem de água) e 2 deles em terreno de baixa elevação (acúmulo de água e matéria orgânica), então os dois plots superiores encontram-se em condições similares, os dois plots na colina encontram-se em condições similares, e os dois plots no terreno baixo encontram-se em condições similares. Temos um bloco (elevação) com três níveis (alto, inclinado, baixo) de tamanho dois.

Exemplo 3: Se estou comparando porcentagem de areia em tijolos mas meu forno só consegue armazenar 24 tijolos por fornada, tenho um bloco (forno) com b níveis (quantas fornadas eu for coletar) e tamanho 24.

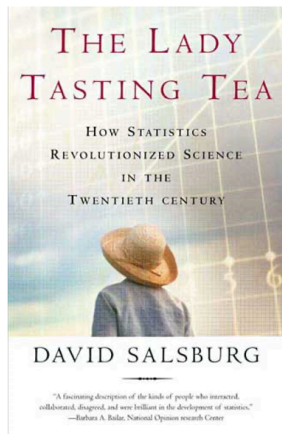
Princípios Básicos em Delineamento de Experimentos



Princípios Básicos

Fisher (1935) ilustrou os princípios da experimentação com o seguinte experimento.

Vocês conhecem o livro **The Lady Tasting Tea** ou **Uma Senhora Toma Chá** (em português)?



The Lady Tasting Tea

Dr. Bristol alega que ela consegue distinguir, experimentando uma xícara de chá, se o leite ou o chá foi colocado na xícara primeiro.

Decidiram então testá-la!

O resultado do experimento nunca foi divulgado!

A significância deste experimento está nas seguintes questões:

- Quantas xícaras de chá deveríamos testar?
- Oferecer xícaras A e B aos pares com a ordem aleatória dentro de cada par, melhora a precisão?
- As condições para fazer e experimentar os chás devem ser as mais semelhantes possíveis para que a diferença no sabor seja apenas atribuída ao método de preparação?
- Quais fatores podem influenciar no gosto do chá?



Quatro Eras na História de DOE

- **1919 – 1940: Era da Agricultura:**

- Sir Ronald A. Fisher com seu trabalho no Rothamsted Agricultural Experimental Station, UK.
- Princípios Básicos, ANOVA, Experimentos Fatoriais.
- Em 1935 publicou o livro: The Design of Experiments.

- **1951 – 1970: Primeira Era Industrial**

- Box & Wilson (1951)
- Superfície de resposta: usar uma sequência de experimentos para encontrar a resposta ótima.
- Aplicações na indústria química.

- **1970 – 1990: Segunda Era Industrial**

- Japão utiliza técnicas de controle de qualidade para melhorar seus produtos depois da WWII.
- Taguchi: Experimentos Fatoriais Fracionados, Planejamento Robusto.

- **A partir de 1990 – Era Moderna**

- Competitividade econômica e globalização levando todos os setores da economia a serem mais competitivos.
- Six Sigma torna-se popular e é hoje usada em diferentes setores da indústria.

- **Por volta de 1960 - Ensaios Clínicos**

- Avanços médicos eram feitos sem rigor científico até então.
- Ensaios clínicos duplamente cegos tornaram-se padrão na aprovação de qualquer produto ou intervenção.

- Montgomery, DC. Design and Analysis of Experiments. Capítulo 1.
- Dean, A. e Voss, D. - Design and Analysis of Experiments. Capítulo 1.

Agradecimentos

Parte deste material foi criado pelo prof. Guilherme Ludwig - IMECC/UNICAMP