

Pesquisa Experimental

Profa. M. Cristina
cristina@icmc.usp.br

Baseado em material preparado e cedido por:

Nemesio Freitas Duarte Filho

Kleberson Junio do Amaral Serique

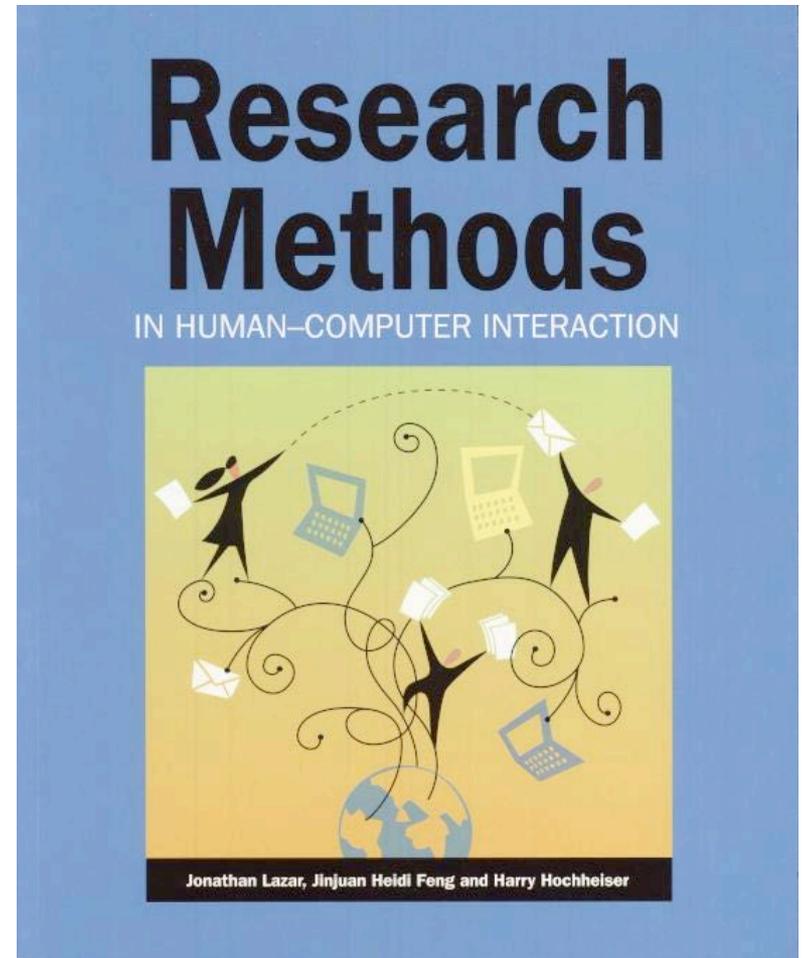
Prof. Dra. Renata Pontin

Prof. José Fernando Rodrigues Jr.

Leitura recomendada 1

- Capítulo 2:
Experimental research

(Jonathan Lazar, Jinjuan Heidi Feng, & Harry Hochheiser - Research Methods in Human-Computer Interaction, Wiley, 2010. ISBN 0-470-72337-8, 978-0-470-72337-1)



Roteiro

- Conceitos básicos da pesquisa experimental
- Passos de um experimento
- Variável dependente & independente
- Hipóteses nulas e hipóteses alternativas
- Tipos de erros
- Limitações da pesquisa experimental

Introdução

- A investigação científica e os seus fundamentos são elementos importantes para a Ciência, proporcionando a descoberta de novos conhecimentos em diversas áreas
- Entretanto, existem diferenças significativas nos modos de condução de investigação científica, devido à diversidade de perspectivas em relação ao foco, objetivos, contextos e aspectos a serem pesquisados
- Uma das abordagens possíveis é a empírica
 - pesquisa experimental

Introdução

- A **pesquisa experimental** busca dar elementos para um pesquisador determinar o efeito causal entre dois fatores
 - P.ex. observou-se que um grupo de adolescentes que jogam um certo jogo digitam mais rápido do que outro grupo de adolescentes que não jogam
 - Existe uma relação causal, i.e., jogar esse jogo faz com que os adolescentes digitem mais rápido?
- Para isso, é preciso testar hipóteses que expressam a convicção do pesquisador
 - Uma hipótese é uma afirmação precisa relativa um problema, que pode ser verificada (testada) diretamente por meio de uma investigação empírica

Conceitos básicos da Pesquisa Experimental

A pesquisa experimental consiste, essencialmente, em determinar um objeto de estudo, selecionar as variáveis capazes de influenciá-lo e definir as formas de controle dessas variáveis e de observação dos efeitos que elas produzem no objeto de estudo

Trata-se, portanto, de uma pesquisa em que o pesquisador é um agente ativo, e não um observador passivo.

Conceitos básicos da Pesquisa Experimental

Princípios da pesquisa experimental:

Temos uma idéia de relação causa-efeito

Acreditamos que existe uma relação entre a construção da causa e a construção do efeito

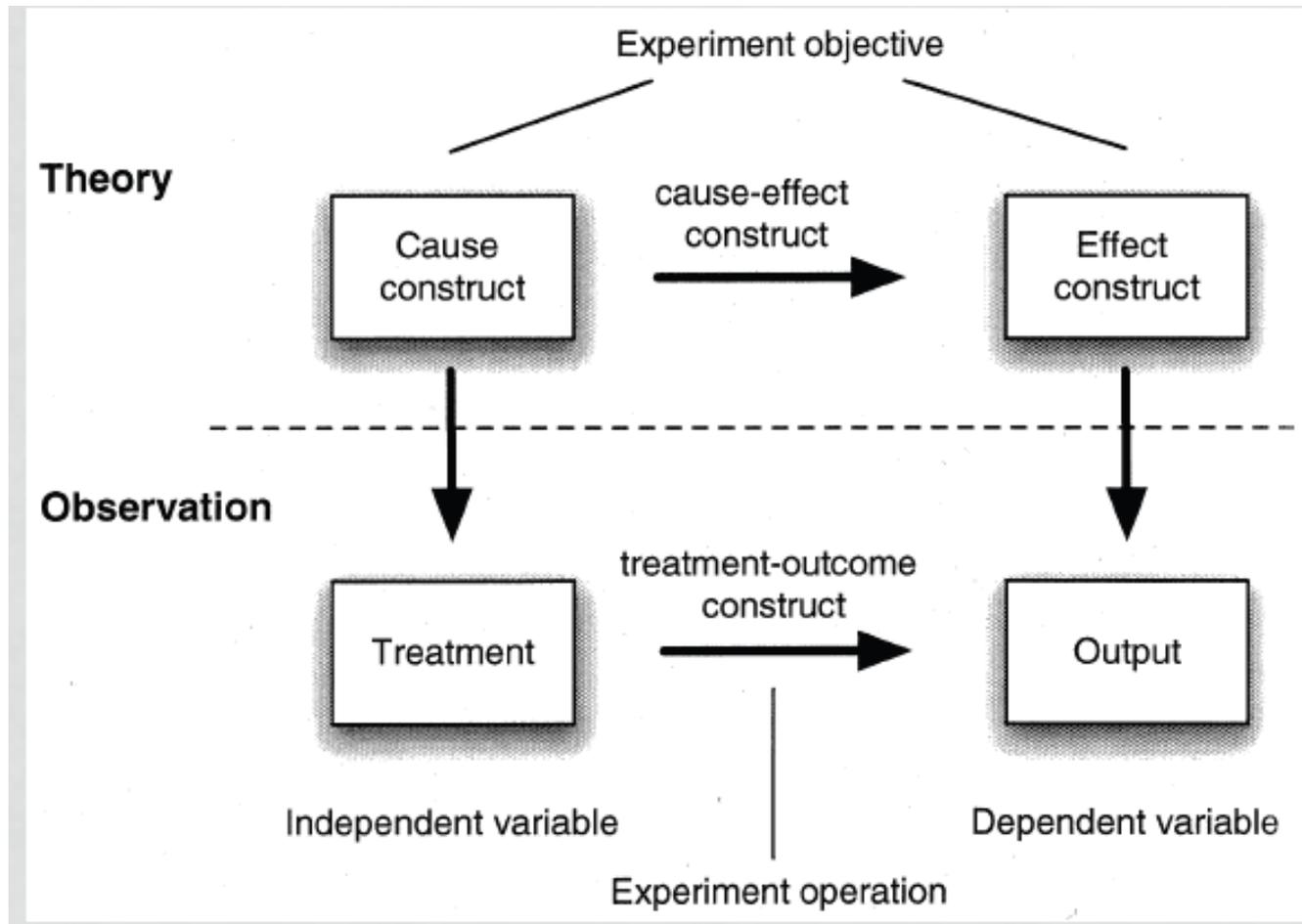
Temos formulações de hipóteses a serem testadas

Temos vários tratamentos (variáveis independentes)

Executamos o experimento e observamos a saída (variáveis dependentes)

Se o experimento foi corretamente elaborado, podemos formular conclusões a respeito da relação causa-efeito para a hipótese estabelecida

Conceitos básicos da Pesquisa Experimental



Passos de um Experimento

1-Definição do Contexto

O escopo do experimento é estabelecido em termos do problema existente, objetivos e metas

2-Planejamento

Faz-se o projeto do experimento, a instrumentação é preparada e as ameaças à validade do experimento são avaliadas

3-Execução

Faz-se as medições => dados

4-Análise e Interpretação

Dados coletados são analisados com suporte estatístico

5-Apresentação e Empacotamento

Os resultados são apresentados e empacotados

Hipóteses de Pesquisa

A definição de um experimento é formalizada por meio de hipóteses

Uma hipótese deve ser declarada formalmente os dados coletados durante a execução experimental permitirão verificar a hipótese (rejeitar/aceitar)

Se a hipótese pode ser rejeitada/aceita então conclusões podem ser obtidas que são generalizáveis para outras situações semelhantes, levando em consideração alguns riscos

Hipótese nula vs Hipóteses alternativas

Um experimento normalmente tem ao menos uma hipótese nula e uma hipótese alternativa

Hipótese nula (H_0) – declara que não existe diferença entre os tratamentos experimentais, i.e., que não será possível observar uma padrão de causa-efeito no experimento

Hipótese alternativa (H_1) – declara o oposto, i.e., que será possível observar uma padrão de causa-efeito no experimento

As hipóteses devem ser mutuamente exclusivas, i.e., se uma é falsa a outra é verdadeira, e vice-versa

Hipótese nula vs Hipóteses alternativas

H_0 é a hipótese que queremos **REJEITAR** com a maior significância (certeza) possível

Exemplo: “Uma nova técnica de inspeção de software encontra, na média (μ), o mesmo número de falhas ($\#F$) do que uma técnica existente”

$$H_0: \mu_{\#F_antiga} = \mu_{\#F_nova}$$

Hipótese nula vs Hipóteses alternativas

Hipótese Alternativa (H_1)

Exemplo: “A nova técnica de inspeção encontra, na média (μ), mais falhas ($\#F$) do que a técnica existente”.

$$H_1: \mu_{\#F_antiga} < \mu_{\#F_nova}$$

Hipótese nula vs Hipóteses alternativas

O objetivo de um experimento é achar evidência estatística que permita **rejeitar a hipótese nula**, para fortalecer a hipótese alternativa

Alguns experimentos podem investigar múltiplos pares de hipóteses nulas/alternativas

mas quanto mais hipóteses serão testadas, mais fatores precisarão ser controlados e mais variáveis precisarão ser medidas – aumentando os riscos

Exemplos

- H_0 : there is no difference between user satisfaction rating between the pull-down menu and the pop-up menu
- H_1 : there is a difference between user satisfaction rating between the pull-down menu and the pop-up menu

Hipótese nula vs Hipóteses alternativas

A hipótese nula H_0 representa *a circunstância que está sendo testada, e o objetivo dos testes de hipóteses é sempre tentar rejeitar a hipótese nula.*

A hipótese alternativa H_1 representa o que se deseja provar ou estabelecer, sendo formulada para contradizer a hipótese nula.

Uma hipótese bem formulada coloca claramente quais são as **variáveis dependentes** e as **variáveis independentes** do estudo.

Variáveis dependentes e independentes

Ao conduzir um experimento, definimos um **tratamento experimental** que nos permite observar (ou medir) **variáveis de saída**, manipulando (ou controlando) algumas **variáveis de entrada** do processo.

Variáveis independentes: aquelas que são manipuladas ou controladas (ou seja, as variáveis de entrada, que podem 'causar' efeitos na variável de saída)

Variáveis dependentes: aquelas que queremos observar para verificar os efeitos das manipulações aplicadas às variáveis de entrada (ou seja, as variáveis de saída observadas)

Variáveis dependentes e independentes

Independentes

Todas aquelas que são manipuladas ou controladas

Identificar as v.i. não é fácil e, normalmente, exige *conhecimento do domínio*

Dependentes

As que são observadas (medidas) para ver os efeitos das manipulações nas variáveis independentes

Normalmente temos apenas uma variável dependente, que na maioria das vezes *não é diretamente mensurável*

Variáveis dependentes e independentes

- Ex.
 - H_0 : *There is no difference between the pull-down menu and the pop-up menu in the time spent locating pages*
 - Independent variable: type of menu (pull-down ou pop-up)
 - + individual factors?
 - Dependent variable: time to locate a page
- H_0 : Uma técnica nova de inspeção não encontra, na média (μ), mais falhas (#F) do que a técnica antiga
 - Variável independente: número de falhas encontradas
 - Variável dependente: a técnica de inspeção

Exemplos (HCI)

- V.I. típicas:
 - Relacionadas a tecnologias ou dispositivos, p. ex. mouse vs joystick, ...
 - Relacionadas às pessoas, p.ex. idade, gênero, nível educacional, profissão, experiência com computação, limitações físicas, ...
 - Relacionadas ao ambiente, p.ex. nível de ruído, nível de iluminação, ...
- V.D. típicas:
 - Medidas de eficiência, de precisão, de satisfação, de esforço cognitivo, ...

Componentes do experimento

- Identificada a hipótese de pesquisa, o design do experimento envolve 3 componentes: **tratamento, unidades, método de atribuição**
- **Tratamentos:** as diferentes técnicas, dispositivos, procedimentos a serem comparados
- **Unidades:** os objetos aos quais os tratamentos serão aplicados (participantes)
- **Método de atribuição:** definição de como unidades serão alocadas aos diferentes tratamentos (deve garantir que seja aleatório) (*randomization*)

Aleatorização

- A alocação dos diferentes tratamentos experimentais às unidades deve ser totalmente aleatória
 - Requisito essencial para a validade do estudo experimental
- É a atribuição aleatória dos participantes aos diferentes tratamentos experimentais que potencialmente anula a influência de variáveis ocultas (não consideradas)
- Pode ser necessário 'randomizar' outros fatores ou condições, p.ex., a ordem de execução de tarefas...

Teste de hipótese

- Testes estatísticos da hipótese (teste de significância). Porque são necessários?
- Ex. H_0 : *There is no difference between the pull-down menu and the pop-up menu in the time spent locating pages*
- Porque não podemos simplesmente comparar, p.ex., duas médias da medida de tempo (ou de uma medida de erro) e concluir que o tratamento que produziu o menor valor tem melhor desempenho do que o outro?

Teste de hipótese

- Testes estatísticos da hipótese (teste de significância). Porque são necessários?
- Porque o experimento observou o comportamento de uma “amostra” da população alvo
 - Até que ponto podemos confiar no resultado obtido observando o comportamento dessa amostra?
 - Que certeza temos de que esse resultado generaliza para a população como um todo?
 - Até que ponto podemos confiar que o efeito observado não foi causado por outros fatores não considerados?

Testes de significância

- Em termos técnicos, um teste de significância é um processo em que uma hipótese nula (H_0) é contrastada com uma hipótese alternativa (H_1) para determinar se a hipótese nula pode ser rejeitada, e com que confiança, em função dos dados observados
 - resultado **positivo**: hipótese nula é **rejeitada**
 - resultado **negativo**: hipótese nula **não pode ser rejeitada**
- O resultado tem uma certa probabilidade de estar correto, e tem uma certa probabilidade de estar errado.

Testes de significância

- Ex.

“On average, participants performed significantly better ($F(1,25)=20,83$, $p < 0.01$)... in condition A than in condition B ... “

“A t test showed that there was a significant difference in the number of lines of text entered ($t(11) = 6.28$, $p < 0.001$) with more entered in the tactile condition. ... ”

Erros: tipo I e tipo II

Ao testarmos uma hipótese nula, chegamos a uma conclusão:
rejeitá-la (positivo), ou não rejeitá-la (negativo)

Entretanto, devemos lembrar que as conclusões ora são corretas, ora são incorretas (mesmo quando fazemos tudo corretamente!)

Este é o preço a ser pago por estarmos trabalhando em uma situação em que a variabilidade é inerente !!!

Testes de significância

- O teste da hipótese resulta em 4 cenários possíveis:
 - Hipótese nula rejeitada (corretamente) (positivo)
 - Hipótese nula não rejeitada (corretamente) (negativo)
 - Hipótese nula rejeitada (incorretamente) (falso positivo)
 - Hipótese nula não rejeitada (incorretamente) (falso negativo)

Erros: tipo I e tipo II

Erro Tipo-I

Ocorre quando um teste estatístico rejeita a hipótese nula, sendo que ela é verdadeira e não deveria ser rejeitada (**falso positivo**)

A probabilidade de cometer um erro desse tipo pode ser expressa como:

$$P(\text{type-I-error}) = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira})$$

No exemplo da técnica de inspeção, um erro tipo-I é dado pela probabilidade de **rejeitar H_0** mesmo que as duas técnicas, na média (μ), encontrem o mesmo número de falhas ($\#F$)

Erros: tipo I e tipo II

Erro Tipo-II

Ocorre quando um teste estatístico não rejeita a hipótese nula quando ela é falsa, e deveria ter sido rejeitada (**falso negativo**)

A probabilidade de cometer um erro desse tipo pode ser expressa como:

$$P(\text{type-II-error}) = P(\text{n\~{o} rejeitar } H_0 \mid H_0 \text{ \textit{\'e} falsa})$$

No exemplo apresentado, o erro tipo-II é dado pela probabilidade de **n\~{o} rejeitar H_0** mesmo que as duas t\~{e}cnicas, na m\~{e}dia, possuam m\~{e}dias (μ) do n\~{u}mero de falhas ($\#F$) encontradas diferentes (e, portanto, H_0 deveria ter sido rejeitada)

Erros: tipo I e tipo II

		Jury decision	
		Not guilty	Guilty
Reality	Not guilty	✓	Type I error
	Guilty	Type II error	✓

Table 2.3 Type I and Type II errors in the judicial case.

Erros: tipo I e tipo II

		Study conclusion	
		No difference	Touchscreen ATM is easier to use
Reality	No difference	✓	Type I error
	Touchscreen ATM is easier to use	Type II error	✓

Table 2.4 Type I and Type II errors in a hypothetical HCI experiment.

Erros: tipo I e tipo II

		O Verdadeiro Estado da Natureza	
		A hipótese nula é verdadeira.	A hipótese nula é falsa.
Decisão	Decidimos rejeitar a hipótese nula.	Erro tipo I (rejeição de uma hipótese nula verdadeira)	Decisão correta
	Não rejeitamos a hipótese nula.	Decisão correta	Erro tipo II (Não rejeição de uma hipótese nula falsa)

Controle de riscos: erros tipo I e tipo II

- O grau de risco de incorrer em erros do tipo I ou II depende de diferentes fatores
 - Probabilidade de incorrer em um erro do tipo I (falso positivo, i.e., erroneamente rejeitar a hipótese nula): *alpha*, ou nível de significância do teste estatístico, ou *p-value*
 - Probabilidade de incorrer em um erro do tipo II (falso negativo, i.e., erroneamente aceitar a hipótese nula): *beta*

Controle de riscos: erros tipo I e tipo II

- *alpha, ou p-value*: probabilidade de erro do tipo I (falso positivo)
- *beta*: probabilidade de erro do tipo II (falso negativo)
- **Poder do Teste Estatístico (P)** é dado por $(1 - beta)$: dá a probabilidade do teste tomar a decisão correta, i.e., rejeitar a hipótese nula, quando ela é de fato falsa.

Erros: tipo I e tipo II

		O Verdadeiro Estado da Natureza	
		A hipótese nula é verdadeira.	A hipótese nula é falsa.
Decisão	Decidimos rejeitar a hipótese nula.	Erro tipo I (rejeição de uma hipótese nula verdadeira) alpha	Decisão correta 1 - beta
	Não rejeitamos a hipótese nula.	Decisão correta 1 - alpha	Erro tipo II (Não rejeição de uma hipótese nula falsa) beta

Controle de riscos: erros tipo I e tipo II

Para tanto, ao realizar um experimento devemos escolher um teste com o maior P possível

$$P = (\text{rejeitar } H_0 \mid H_0 \text{ é falsa}) = 1 - P(\text{type-II-error})$$

Entretanto, α e β são relacionados: um valor muito baixo para α reduz a chance de incorrer em erros do tipo I, mas aumenta a chance de incorrer em erros do Tipo II...

Controle de riscos: erros tipo I e tipo II

Exemplo:

A eficácia de certa vacina após um ano é de 25% (isto é, o efeito imunológico se prolonga por mais de um ano em apenas 25% das pessoas que a tomam). Desenvolve-se uma nova vacina, mais cara, e deseja-se saber se esta é, de fato, melhor (mais eficaz)

Sendo “r” a proporção de imunizados por mais de um ano com a nova vacina...

Quais hipóteses devem ser formuladas?

Que erros pode-se cometer?

Controle de riscos: erros tipo I e tipo II

Exemplo:

Hipótese nula $H_0: r \leq 0,25$

Hipótese alternativa $H_1: r > 0,25$

Erro tipo I (FP): rejeitar H_0 , i.e., aprovar a nova vacina quando, na realidade, ela não tem um efeito superior ao da vacina em uso

Erro tipo II (FN): aceitar H_0 , i.e., não aprovar a nova vacina quando ela é, de fato, mais eficaz do que a vacina em uso

Seleção dos participantes

A seleção dos participantes afeta diretamente a capacidade de generalização dos resultados de um experimento

Para tanto, a seleção deve ser representativa da população alvo

Seleção de participantes = amostra de uma população

A amostragem pode ser probabilística ou não-probabilística

Amostragem probabilística: *a probabilidade da seleção de cada participante é conhecida*

Amostragem não-probabilística: *a probabilidade da seleção de cada participante não é conhecida*

Seleção dos participantes

O tamanho da amostra tem impacto sobre a generalização dos resultados de um experimento

Quanto maior a amostra, menor é a chance de errar ao generalizar os resultados

Princípios gerais para escolher o tamanho da amostra:

Se existir uma ampla variabilidade na população, é necessário uma amostra de tamanho maior;

A análise dos dados pode influenciar a escolha do tamanho da amostra.

Limitações da Pesquisa Experimental

As pesquisas experimentais constituem um valioso procedimento disponível aos cientistas para testar hipóteses que estabelecem relações de causa e efeito entre as variáveis

Em virtude de suas possibilidades de controle, os experimentos oferecem garantia muito maior do que qualquer outro delineamento de que a variável independente causa efeitos na variável dependente.

A despeito de suas vantagens, a pesquisa experimental apresenta várias limitações.

Primeiramente, existem muitas variáveis, cuja manipulação experimental pode ser muito difícil ou mesmo impossível

Limitações da Pesquisa Experimental

Uma série de características humanas, tais como idade, sexo ou histórico familiar, não podem ser conferidas às pessoas de forma aleatória.

Outra limitação consiste no fato de que muitas variáveis que poderiam ser tecnicamente manipuladas estão sujeitas as considerações de ordem ética que proíbem sua manipulação.

Não se pode, por exemplo, submeter pessoas a atividades estressantes com vistas a verificar alterações em sua saúde física ou mental.

Task (next class)

- Bring **two** examples of papers (in HCI, Imaging or Visualization) that verify research hypothesis
 - Describe the null and alternative hypotheses in each case
 - Identify the independent variable(s) and the dependent variable
 - Identify the experimental design adopted
 - Identify and explain the choice of statistical significance test applied