

Fundamentos Estatísticos para Ciência dos Dados

Estudos Experimentais x Observacionais

Renato Martins Assunção

DCC, UFMG - 2020



Efeito de X em Y

- Estudos estatísticos: Deseja-se estudar como uma variável aleatória Y responde a mudanças de uma ou mais variáveis independentes X .
- Por exemplo:
 - O webpage design altera a chance de clicar num anúncio patrocinado? Teste A/B: crie duas páginas, tipo A e tipo B, e separe ao acaso parte do tráfego no site para a página A ($X = 0$) ou página B ($X = 1$). Compare a porcentagem Y de usuários que clicam no anúncio em A com a porcentagem que clica no anúncio em B. Alterar X de 0 para 1 altera o valor esperado de Y ?
 - Dois compostos químicos, 1 e 2, combinados nas quantidades X_1 e X_2 . Qual é o efeito dessa combinação no output Y ? Isto é, como podemos representar o valor esperado de Y como uma função $g(x_1, x_2)$?

Efeito de X em Y

- Mais exemplos:
 - Para cada indivíduo, fumar ou não é representada por $X = 0$ ou $X = 1$. E se o indivíduo desenvolver câncer de pulmão ou não é representada por $Y = 1$ e $Y = 0$. Fumar aumenta as chances de que $Y = 1$? Isto é, alterar $X = 0$ para $X = 1$ aumenta as chances de que Y assumo o valor 1?
 - Mais desemprego (X) induz o aumento de crimes (Y)? O valor de Y varia se alteramos o valor de X ?

Experimentais x observacionais

- Dois tipos de análise de dados estatísticos de acordo com o método de coleta dos dados: *Estudos Experimentais* e *Estudos Observacionais*.
- *Experimentais*: As variáveis independentes X são diretamente manipuláveis pelo experimentador.
- *Observacionais*: Coleta-se os pares (X, Y) nas unidades de observação sem manipular os valores de X .
- As consequências disso para a inferência indutiva são muito grandes.

Experimentais x observacionais

- Nos estudos experimentais, podemos chegar a conclusões muito mais firmes e claras sobre os efeitos de X sobre Y .
- Nos estudos observacionais, sempre pairam dúvidas legítimas sobre a validade das conclusões.
- Restante dos slides é “tradução” dos 2 primeiros capítulos de Statistics de Freedman, Pisani e Purves.

Padrão Ouro

- Um novo medicamento é introduzido. Como deveria ser planejado um experimento para testar sua efetividade?
- O método básico é a comparação (ou teste A/B).
- O medicamento é aplicado a indivíduos em um grupo de tratamento, mas outros indivíduos são usados como controle - estes não são tratados.
- Então as respostas dos dois grupos são comparadas.
- Resumo:
 - Pergunta: Alterar o valor de X impacta nos valores de Y ?
 - Crie dois grupos: grupo A tem $X = 0$. Grupo B tem $X = 1$.
 - Em cada grupo, colete a resposta Y em vários itens. Compare a resposta média de Y quando $X = 0$ (no grupo A , ou $\mathbb{E}(Y|X = 0)$) com a resposta média de Y quando $X = 1$ (no grupo B , ou $\mathbb{E}(Y|X = 1)$).

Padrão Ouro

- ESSENCIAL: os indivíduos do controle devem ser escolhidos *aleatoriamente*.
- DESEJÁVEL: experimento deve ser duplamente cego: nem os indivíduos nem os experimentadores que medem as respostas devem saber quem está no grupo de tratamento nem quem está no grupo de controle.

Vacina Salk

- Anos 50: muitas vacinas contra a poliomielite infantil.
- A de Jonas Salk era a mais promissora em labs.
- Em 1954: experimento social gigante.
- Aproximadamente 2 milhões de crianças.
- Idades de 6, 7 e 8 anos (estudantes dos grades 1, 2 e 3)
 - 1/2 milhão: vacinados.
 - 1 milhão: deliberadamente NÃO vacinados (CONTROLE)
 - 1/2 milhão: recusou-se a participar.

Comparação

- O tratamento (vacinação) é aplicado a alguns indivíduos: o grupo de TRATAMENTO ($X = 1$).
- Os outros não recebem o tratamento e são usados como CONTROLE ($X = 0$).
- RESPOSTA Y : taxa (por 100 mil) de casos de pólio no grupo.
- As respostas dos dois grupos são comparadas para ver se o tratamento faz alguma diferença.

Ética

- Se a vacina era promissora, por que não vacinar todas as crianças ?
- Incerteza com novos medicamentos: lab \neq vida real.
- Um experimento de campo não poderia ser vacinar todas as crianças e comparar com as taxas históricas?
- Se a incidência de pólio em 1954 fosse muito menor que em 1953 \rightarrow uma evidência a favor da efetividade da vacina Salk.
- Certo? Na verdade, não.
- Vão pairar muitas dúvidas legítimas.

Dúvidas no ar

- A pólio é uma doença epidêmica, com muita variação de ano para ano.
- Em 1952, ≈ 60000 casos; em 1953, ≈ 30000 casos.
- Em 1954, aplicamos a vacina em todo mundo: sem controles.
- Observamos uma baixa incidência em 1954, menor que os 30 mil casos de 1953.
- Isto ocorreu porque a vacina funcionou ou porque não houve epidemia em 1954?
- Pior ainda: Suponha que ocorra uma ALTA incidência em 1954 (80 mil casos, por exemplo).
- Alguém diz: a vacina é eficaz. Se não tivéssemos aplicado a vacina, o número seria maior ainda!
- Como concluir com segurança?

Deixando crianças sem vacina

- A ÚNICA forma de estar seguro se a vacina iria funcionar era deixar algumas crianças sem vacinar.
- Precisa da aprovação dos pais.
- Ideia “brilhante”: aproveitamos este limão (dificuldade) para fazer uma limonada.
- Criamos o Desenho Amostral 1:
 - crianças cujos pais consentissem: tomam vacina (grupo de tratamento).
 - As demais são as crianças cujos pais não permitem vacinar: OK, essas passam a ser o grupo de controle!
- Problema: Pais com renda mais alta consentem mais facilmente.
- Fonte potencial de viés??
- Sim, CONTRA a vacina (!!)
- Crianças mais ricas eram MAIS vulneráveis à pólio.

Pólio e higiene

- Siga o raciocínio: Pólio é uma doença de higiene.
- As crianças *mais pobres* tem *mais contato* com formas brandas de pólio nos primeiros anos de vida.
- Neste tempo, ainda estão protegidas pelos anticorpos do leite de sua mãe
- → criam anticorpos, que as protegem contra infecções mais sérias mais tarde.
- Assim, os “controles” (com muitos pobres que recusaram a vacina) formaram um grupo diferente daqueles que irão a vacina.
- Antes mesmo de recebe a vacina, os grupos são diferentes: os controles são MENOS susceptíveis á doença!
- Ao comparar os dois grupos, o efeito da vacinação será MENOR que o seu efeito real.
- Poderia até aparecer um aparente efeito DELETÉRIO quando, na verdade, a vacina tivesse um efeito preventivo.

Fatores de confundimento

- Para evitar o viés: grupos de tratamento e controle deveriam ser o mais parecido possível, EXCETO pelo tratamento.
- Assim, QUALQUER diferença substancial entre os grupos seria devido ao tratamento e não a algum outro fator que diferencie os grupos (pois este outro fator não existiria).
- Se os dois grupos diferem em algum fator além do tratamento, os efeitos deste outro fator poderiam ser confundidos (misturados) com os efeitos do tratamento.
- *Fatores de confundimento* são a maior fonte de viés.

Desenho da NFIP

- National Foundation for Infantile Paralysis (NFIP).
- Vacinar todas as crianças do grade 2 cujos pais consentissem (tratamento)
- Todas as crianças dos grades 1 e 3 como controles.
- Implementada por muitos distritos escolares (1 milhão de crianças).
- Dois problemas sérios:
 - 1 A pólio é uma doença contagiosa. A incidência PODERIA ser muito maior no grade 2: viés contra a vacina. Se fosse maior: viés a favor da vacina.
 - 2 Tratamento: grade 2 E CONSENTIMENTO. Controle: grade 1 e 3. Grupo de tratamento teria mais crianças ricas (e portanto, mais vulneráveis) → viés contra a vacina.

Desenho 3: controlado e duplo cego

- Comparação justa: controle deveria ser escolhido da MESMA população que o grupo de tratamento, mesmo que seja uma população mais restrita.
- SUB-POPULAÇÃO onde comparação vai ser feita: crianças cujos pais consentissem a vacinação.
- Como dividir as crianças para o tratamento e para o controle?
- Grupos devem ser o mais similar possível com relação a muitas (infinitas?) variáveis relevantes: renda dos pais, o estado geral de saúde da criança, personalidade, hábitos sociais, etc.

ALEATORIZE

- Como conseguir este equilíbrio? queremos equilíbrio entre TODOS os possíveis fatores de confundimento, aqueles conhecidos e ATÉ MESMO AQUELES DESCONHECIDOS!!!!
- Solução: ALEATORIZE com uma moeda!
- Cara: tratamento. Coroa: controle.

ALEATORIZE

- Cara: tratamento. Coroa: controle.
- Procedimento é imparcial.
- Probabilidade garante que, com um número suficiente de indivíduos, ambos os grupos permaneceriam muito próximos um do outro.
- Isto aconteceria com relação a todas as variáveis importantes.
- MESMO QUE NÃO estejam identificadas !!!!
- Se procedimento aleatório imparcial determina QUEM faz parte do tratamento e do controle: experimento é *controlado com aleatorização*.

Desenho 3: controlado + placebo

- Controle: receberam uma injeção de sal dissolvida em água.
- Crianças não sabiam se estavam em tratamento ou em controle.
- Eram CEGAS sobre seus status como tratamento ou controle.
- Respostas ao tratamento (à vacina) e não à IDÉIA do tratamento.
- Como assim? Protegidos contra a pólio somente pela força do pensamento??
- Estudos com pacientes com forte dor pós-operatória: receberam pílula de farinha MAS descrita como excelente analgésico.
- Aprox 1/3 disseram ter melhorado.

Desenho 3: controlado + DUPLO CEGO

- Mais uma precaução: cegar os MÉDICOS.
- Diagnósticos para decidir se as crianças contraíram pólio durante o tratamento.
- Muitas formas de pólio são difíceis de diagnosticar.
- Em casos duvidosos, o diagnóstico podia ser afetado por saber se a criança tinha sido vacinada.
- Os médicos dos diagnósticos não sabiam se a criança estava em tratamento ou era um controle.
- Chamado duplo cego: nem o indivíduo e nem aquele que avalia a resposta ao tratamento sabem se ele recebeu o tratamento.
- Este é o DESENHO 3: controlado com aleatorização e duplo cego.

Resultados

Grupo	Aleatorizado, Duplo Cego		NFIP	
	No. (em mil)	Taxa	No. (em mil)	Taxa
Trat.	200	28	225	25
Controle	200	71	725	54
Sem Consent	350	46	125	44

Tabela:

- Desenho 3, experimento controlado com aleatorização e duplo cego: $28 - 71 = -43$ crianças por cada 100 mil.
- Desenho 2, NFIP: $25 - 54 = -29$ crianças por cada 100 mil, um efeito MENOR que no Desenho 3.
- NFIP apresentou um viés contra a vacina.
- *Neste exemplo*, o impacto do viés não muda a conclusão. Mas nem sempre é assim.

Vantagem adicional da aleatorização

- Podemos TESTAR estatisticamente se a diferença observada entre os grupos pode acontecer por mero acaso.
- Assuma o papel de advogado do diabo: suponha que a vacina Salk não tem efeito realmente.
- Sob esta hipótese de efeito NULO da vacina, a diferença observada $28 - 71 = -43$ entre as taxas de pólio para os grupos de tratamento (28 por 100 mil) e de controle (71 por 100mil) é devida somente ao acaso.

Vantagem adicional da aleatorização

- Qual a probabilidade de observarmos uma diferença tão grande quanto -43 entre os grupos CASO NÃO EXISTA DIFERENÇA REAL ENTRE ELES?
- $\mathbb{P}(\text{Efeito} \leq -43 | \text{Trat} = \text{Cont}) = 10^{-9}$, ridiculamente pequena.
- SE A VACINA É DE FATO INÓCUA, então um evento MUITO IMPROVÁVEL aconteceu.

Nem sempre...

- Com o desenho 2, da NFIP, muitos fatores de confundimento aleatórios.
- Quais são as famílias voluntárias, quais crianças estão na faixa etária de dois anos, e assim por diante.
- Como estimar a chance de qualquer um destes fatores? Sem chances...
- Não podemos calcular a chance de ter uma diferença entre grupos CAUSADA apenas por estes fatores de confundimento.
- Com o experimento controlado com aleatorização, a chance entra de uma forma simples e planejada ao determinar o tratamento ou o controle.

No caso aleatorizado

- A hipótese de advogado do diabo é que a vacina Salk não faz efeito nenhum.
- Sob esta hipótese, algumas crianças têm mais tendência a contrair pólio, e a maioria não.
- A determinação de tratamento ou controle **não tem nada a ver com isto**.
- Cada criança tem uma chance de 50% de ficar em tratamento ou em controle, somente dependendo do lançamento de uma moeda.

No caso aleatorizado

- Assim, cada criança com maior tendência a contrair pólio tem uma chance meio a meio de estar em tratamento ou em controle.
- Portanto, *sob a hipótese do advogado do diabo*, o número de casos de pólio em ambos os grupos deve ser aproximadamente o mesmo.
- Qualquer diferença é devida à variabilidade do lançamento da moeda.
- Esta variabilidade é bem estudada e conhecida.
- Podemos calcular $\mathbb{P}(\text{Efeito} \leq -43 | \text{Trat} = \text{Cont}) = 10^{-9}$.

Congelamento gástrico

- Em 1961, uma nova técnica para tratar úlcera gástrica (úlcera no estômago):
 - Paciente é anestesiado;
 - um balão é colocado no estômago;
 - um fluido refrigerante é bombeado congelando o estômago por uma hora.
- O procedimento interrompe o processo digestivo e a úlcera cicatriza.
- O criador da técnica testou em 24 pacientes e TODOS apresentaram melhoras imediatas.
- A técnica tornou-se popular.

Testando com controles aleatorizados

- Ruffin (1969): experimento controlado aleatorizado duplo-cego com 82 (trat) e 78 (cont) pacientes.
- Controles: uma imitação do procedimento.
- Havia um desvio no balão: o fluido refrigerante retornava antes de congelar.
- Avaliações de acompanhamento por 2 anos.
- Médicos não sabiam qual o procedimento realizado.

Resultados congelantes

- Maioria do pacientes melhoravam na primeiras 6 semanas nos DOIS GRUPOS.
- Com o tempo, aprox 40% nos dois grupos pioravam ou voltavam à mesma condição anterior.
- Em nenhum momento havia qualquer diferença significativa entre os dois grupos.
- Tratamento ou controle, a resposta média era a mesma.
- Outros experimentos cuidadosos confirmaram as conclusões: o congelamento gástrico não tinha nenhum efeito diferente do controle.

Meta-análise

Ano	Avaliação				
	Favorável	Mod. Fav.	Neutra	Mod. Desfav.	Desfav.
1962	///				
1963	/	////		/	
1964		//// ^a	///	/	/// ^a
1965		^a	//	/	
1966			^b		// ^b
1967				^a	^a
1968					
1969				/	^a

Tabela: Resultados de vários estudos clínicos. Cada barra vertical é um estudo. a: duplo cego, aleatorizado controlado. b: estudo com controles mas não aleatorizado. Fonte: Medical Technology and the Health Care System: a Study of the Diffusion of Equipment-Embodied Technology. By Committee on Technology and Health Care (U.S.), National Academy of Sciences (U.S.), National Science Foundation (U.S.)

Conclusão

- O proponente da técnica era um médico e pesquisador respeitado e criativo.
- Seu experimento inicial era mal planejado: sem controles.
- Ele exagerou o valor de sua técnica.
- Levou mais de 10 anos para aceitar que a técnica não surtia efeito e ainda podia ter efeitos colaterais.
- The same independence and drive that allow the innovator to be a pioneer also enable him to ignore the collective wisdom of his colleagues.

Derivação portacava

- Casos agudos de cirrose do fígado pode levar a hemorragia e morte.
- Um antigo tratamento: redirecionar o fluxo sanguíneo através da derivação portacava.
- Esta cirurgia era longa e arriscada.
- Vale a pena correr o risco da cirurgia?
- Mais de 50 estudos foram realizados.
- Os resultados estão resumidos na tabela seguinte.

Resultados portacava

Desenho	Grau de Entusiasmo			Total
	Marcado	Moderado	Nenhum	
sem controle	24	7	1	32
controles, sem aleatorizar	10	3	2	15
Aleatorizado, controles	0	1	3	4

Tabela: Resultados de vários estudos clínicos. Fonte: G. Muench and J. Chalmers (1966) J. Gastroenterology.

Tratando os melhores

- No caso destes estudos, existe uma explicação.
- Nos casos sem controle nenhum ou sem controle ALEATORIZADO, existe uma tendência a tratar somente os pacientes que estão em relativa boa forma.
- Isto cria um viés a favor do tratamento.
- Nos 3 tipos de desenhos, $\approx 60\%$ dos tratados viveram mais de 3 anos.
- Nos controles aleatorizados, $\approx 60\%$ também viveu mais de 3 anos.
- Mas apenas 45% dos controles não aleatorizados sobreviveu três anos.
- Isto mostra o viés em experimentos sem aleatorização.

Controles e controles

- A palavra controle é usada em estatística com dois sentidos:
 - Um *controle* é um indivíduo que não recebe o tratamento.
 - Um experimento é *controlado* quando os pesquisadores determinam QUAIS indivíduos receberão o tratamento e quais serão controles (por exemplo, lançando uma moeda).
- Podemos ter indivíduos controles num estudo não-controlado!
- Como assim?

Controles e controles

- Em um estudo observacional, os pesquisadores NÃO determinam quais os indivíduos que serão submetidos ao tratamento ou quais serão controles.
- Por exemplo, os estudos sobre os efeitos do cigarro.
- Não é possível jogar uma moeda para decidir quem vai fumar (tratamento) e quem não vai fumar (controle).
- *Mas a terminologia tratamento-controle é ainda usada.*

Controles e controles

- Os pesquisadores comparam um grupo de fumantes (o grupo de tratamento) com um grupo de não-fumantes (o grupo de controle) para determinar os efeitos do cigarro (o “tratamento”).
- Assim, podemos ter:
 - Um estudo observacional com “tratamento” e “controle”
 - Um experimento *controlado* onde (praticamente sempre) os controles são escolhidos por meio de um mecanismo aleatorizado SOB O CONTROLE DO EXPERIMENTADOR.

Qual o busilis?

- Qual é a diferença fundamental entre os dois controles (ou os dois tipos de estudo)?
- Considere o grupo de não-fumantes (controles) versus fumantes (tratamento).
- Os fumantes geralmente se dão mal nesta comparação: doenças cardíacas, câncer de pulmão, acidentes automobilísticos e suicídios são mais comuns entre fumantes que entre não-fumantes.
- De modo que existe uma forte associação entre o hábito de fumar e doenças cardíacas, câncer de pulmão e o resto.
- O cigarro provoca câncer?
- A CAUSA do câncer de pulmão é o ato de fumar?

Um fator de confundimento?

- Pode haver um fator hereditário (genético) que predispõe as pessoas a ambos: fumar e ter câncer de pulmão.
- Observa-se que fumar e câncer de pulmão caminham juntos porque existe um TERCEIRO fator, anterior aos dois, e que induz aos dois.
- Os efeitos do fator hereditário estariam sendo confundidos com os efeitos do cigarro.
- Fazer esquema gráfico: Fumar \leftarrow GENE \rightarrow Câncer de pulmão.
- E nesse caso, não haveria nenhum motivo para deixar de fumar, já que isso não alteraria o fator hereditário e assim não reduziria o risco de câncer.

Outros fatores de confundimento?

- Fumantes possuem uma maior proporção de sedentários?
- A dieta típica é similar a de não-fumantes?
- Estilo de vida?
- Idade?
- Sexo?

Cigarros e experimentos

- Seleccionamos uma grande amostra (muitos milhares)
- Alguns serão mais sedentários, terão genes A ou B, comerão assim ou assado.
- Suponha que pudéssemos decidir quais pessoas vão fumar e quais não vão fumar ao longo de suas vidas.
- Decisão é feita jogando-se uma MOEDA: cara \rightarrow fumante.
- Os dois grupos estarão equilibrados em relação a TODOS OS POTENCIAIS fatores de confundimento.
- A ÚNICA diferença significativa entre os grupos é o ato de fumar ou não fumar.
- Compara-se a incidência de câncer entre os dois grupos.
- A conclusão é muito mais clara, certo?

Um exemplo mais controverso...

- Possuir armas causa violência?
- Nos EUA, venda de armas é livre.
- Slogan da National Rifle Association: *Guns don't kill people; people kill people*
- Quem possui arma: Tratamento.
- Quem não possui arma: controle
- Fatores de confundimento: idade, sexo, escolaridade, personalidade, local de moradia, etc, etc, etc.
- Como fazer uma comparação adequada?
- Como criar dois grupos que sejam diferentes APENAS pela posse da arma?
- Podemos tentar mas sempre haverá espaço para dúvidas.

You name it...

- Quem fez um melhor governo: FHC ou Lula?
- Uma polícia única (civil e militar) é melhor?
- Liberar o consumo de drogas?
- Liberar o consumo, produção e venda de drogas?
- Quem sofre de depressão faz uso mais intenso de bebidas alcoólicas. Existe uma relação de causa e efeito envolvida aqui? Bebida causa depressão ou depressão causa bebida?

O estudo de moradia de Baltimore

- 1950, John Hopkins University: efeito de moradia pública versus favela.
- Vale a pena construir moradias públicas?
- Lafayette Courts em Baltimore: built in 1955 to replace blocks of dilapidated row houses.
- Grupo de Tratamento: 300 famílias de favelas que se mudaram para Lafayette Courts.
- Grupo de Controle: 300 famílias que viviam em favelas.
- Três anos de acompanhamento, de 1955-1958.
- Tratamento aplicado: provisão de moradia pública, que era melhor e mais barata (não era gratuita) que as favelas.

Resultados

- Respostas:
 - Medições de saúde e de atitude social via questionários e exames.
 - Desempenho escolar das crianças.
- Resultado: grupos eram praticamente idênticos em tudo EXCETO dois itens:
 - Grupo de tratamento era mais feliz com seu ambiente físico.
 - Mortalidade no grupo de tratamento era bem menor: 2 mortes versus 10 mortes entre as 300 famílias de cada grupo.
- Estas poucas diferenças podem ser devido ao tratamento?

Discussão

- De onde vieram os controles?
- Milhares de famílias que moravam em favelas próximas candidataram-se para mudar para Lafayette Courts.
- 800 delas foram escolhidas como inquilinas.
- Experimento escolheu 300 dessas 800 famílias.
- Controles foram as 300 famílias que não tiveram sucesso na candidatura.
- Este é um estudo observacional ou experimental?

Discussão

- Os pesquisadores não escolheram (aleatoriamente) quem receberia o tratamento dentre as 800 famílias pré-selecionadas.
- As que tiveram sucesso eram aquelas mais “adequadas” como inquilinos.
- Quais os critérios usados?
- Eles eram associados com aqueles sendo estudados?
- Famílias com longo registro de crimes, de violência doméstica, de desemprego crônico, etc, tinham maior chance de serem escolhidas?
- Se a escolha estava associada com estes fatores então os dois grupos (tratamento e controle) NÃO SÃO comparáveis antes mesmo que qualquer tratamento seja aplicado.

Outros problemas

- Na verdade, no início: tratamento tinha 400 famílias e controle tinha 600.
- Durante os 3 anos, foram “perdidas” 100 famílias do tratamento e 200 do controle.
- Elas cansaram de colaborar com o estudo, mudaram de Baltimore ou saíram de Lafayette Courts (ou da favela).
- Os grupos ficaram desbalanceados em algumas características e várias famílias do controle foram descartadas.
- Qual o efeito e o tamanho dessas mudanças? Impossível determinar.
- Pelo menos um efeito podemos saber: controles que saíram da favela são diferentes dos que ficaram.
- Os que ficaram são menos determinados, menos desejosos de mudanças, etc.

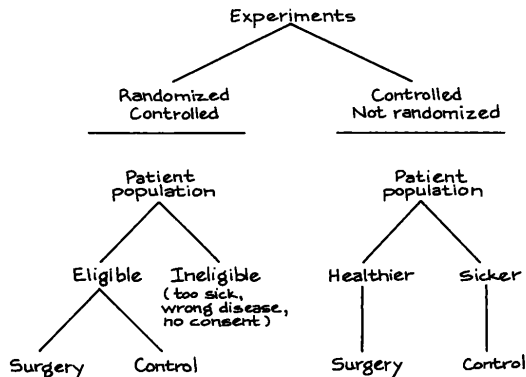
O que poderia ser feito?

- Selecione 1000 famílias que poderiam se mudar para Lafayette Courts.
- Selecione 500 ao acaso para morar lá.
- As outras 500 formam o grupo controle.
- Siga-os no tempo.
- A resposta dos dois grupos é comparável, mesmo que não fiquem onde foram alocados!!

O presente de Lafayette Courts

- Em 1955, Baltimore comemorou quando foi iniciada construção do o conjunto habitacional Lafayette Courts, os “ apartamentos de baixa renda no céu ”.
- Em 1995, existe outro motivo de comemoração: o conjunto habitacional foi cuidadosamente explodido pelas autoridades da cidade.
- Atormentados por anos de crime, drogas e uma batalha sem fim contra a degradação urbana, moradores e autoridades concordaram que o projeto de moradias públicas de Lafayette Courts deveria ser encerrado sem deixar traços.
- Vídeo no YouTube:
<https://www.youtube.com/watch?v=r1FPjiMreTg>

Esquema de desenhos



Critérios de Bradford Hill

- (Sir) Austin Bradford Hill (Londres, 1897 a 1991), epidemiologista e estatístico inglês.
- Pioneiro no uso do raciocínio estatístico em estudos clínicos.
- Com Richard Doll, foi o primeiro a demonstrar a ligação entre o uso do cigarro e o câncer de pulmão.
- Os critérios de Hill é um conjunto de 9 critérios para a determinação de uma associação causal entre X e Y .
- Estes critérios são usados até hoje.



The Bradford Hill causality criteria

- *Força da associação*: quanto mais forte uma associação, mais provável que seja causal.
- *Consistência*: condizente com outros estudos.
- *Especificidade*: exposição causa uma doença específica. O ideal é não ter um agente causando muitas doenças diferentes.
- *Temporalidade*: causa deve ser anterior à doença.
- *Gradiente biológico*: Ter efeito dose-resposta - exposição, mais doença.
- *Plausibilidade biológica*: A associação deve ter uma explicação plausível, concordante com o nível atual de conhecimento do processo patológico.
- *Coerência*: os achados devem seguir o paradigma da ciência atual
- *Evidências experimentais*: Em experimentos de laboratório, mudanças na exposição mudam a incidência da doença
- *Analogia*: com outra doença ou com outra exposição

The Bradford Hill causality criteria

7

Section of Occupational Medicine

295

The Environment and Disease: Association or Causation?

by Sir Austin Bradford Hill CBE DSC FRCP(hon) FRS
(*Professor Emeritus of Medical Statistics,
University of London*)

Amongst the objects of this newly-founded Section of Occupational Medicine are firstly 'to provide a means, not readily afforded elsewhere, whereby physicians and surgeons with a special knowledge of the relationship between sickness and injury and conditions of work may discuss their problems, not only with each other, but also with colleagues in other fields, by holding joint meetings with other Sections of the Society'; and, secondly, 'to make available information about the physical, chemical and psychological hazards of occupation, and in particular about those that are rare or not easily recognized'.

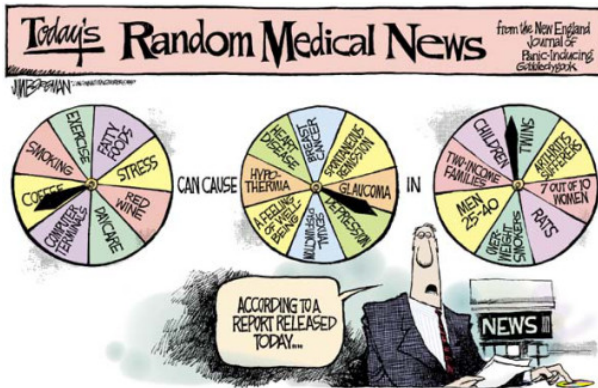
Meeting January 14 1965

President's Address

observed *association* to a verdict of *causation*?
Upon what basis should we proceed to do so?

I have no wish, nor the skill, to embark upon a philosophical discussion of the meaning of 'causation'. The 'cause' of illness may be immediate and direct, it may be remote and indirect underlying the observed association. But with the aims of occupational, and almost synonymously preventive, medicine in mind the decisive question is whether the frequency of the undesirable event B will be influenced by a change in the environmental feature A. *How* such a change exerts that influence may call for a great deal of research. However, before deducing 'causation' and taking action we shall not invariably have to sit around awaiting the results of that research. The whole chain may have to be unravelled or a few links may suffice. It will depend upon circumstances.

Notícias Médicas



Pelagra

- Pelagra é uma doença que estava espalhada em toda a Europa no início do século XX.
- Sintomas: A pele fica escura, seca, áspera e com crostas. Depois, diarreias e demência.
- Condições sanitárias nas casas com doentes eram horríveis, moscas por todo lado.
- Uma mosca que se alimentava de sangue (*Simulium*) tinha a mesma localização geográfica que a pelagra na Europa.
- A mosca era mais ativa na primavera, quando a maior parte dos casos de pelagra aconteciam.
- Médicos concluíram que a doença era infecciosa.
- Como malária e febre amarela e tifo, era transmitida entre as pessoas através dos insetos.

Pelagra

- Uma série de estudos observacionais e experimentais mostraram que pelagra é causada por uma dieta inadequada.
- It can be prevented or cured by certain foods rich in niacina (vitamina B3).
- Niacin occurs naturally in meat, milk, eggs, some vegetables, and certain grains.
- Corn contains relatively little niacin.
- In the pellagra areas, the poor ate corn-and not much else.
- Some villages and some households were poorer than others, and had even more restricted diets.
- The flies were a marker of poverty, not a cause of pellagra.
- Association is not the same as causation.

Cervical cancer and circumcision.

- For many years, cervical cancer was one of the most common cancers among women.
- Many epidemiologists worked on identifying the causes of this disease.
- Cervical cancer was quite rare among Jews.
- The disease was very unusual among Moslems.
- In the 50's, it was thought that circumcision of the males was the protective factor.
- Was this conclusion justified?

Cervical cancer and circumcision.

- There are differences between Jews or Moslems and members of other communities, besides circumcision.
- It turns out that cervical cancer is a sexually transmitted disease, spread by contact (HIV).
- Some women are more active sexually than others, and have more partners;
- they are more likely to be exposed to the viruses causing the disease.
- That seems to be what makes the rate of cervical cancer higher for some groups of women.
- Early studies did not pay attention to this confounding variable, and reached the wrong conclusions.

Ultrasound and low birth weight

- Human babies can now be examined in the womb using ultrasound.
- Several experiments on lab animals have shown that ultrasound examinations can cause low birth weight.
- If this is true for humans, there are grounds for concern.
- Investigators ran an observational study to find out, at the Johns Hopkins hospital in Baltimore.
- Of course, babies exposed to ultrasound differed from unexposed babies in many ways besides exposure; this was an observational study.
- The investigators found a number of confounding variables and adjusted for them.
- Criaram grupos de indivíduos com valores parecidos dos fatores de confundimento.

Ultrasound and low birth weight

- Even so, there was an association.
- Babies exposed to ultrasound in the womb had lower birth- weight, on average, than babies who were not exposed.
- Is this evidence that ultra-sound causes lower birthweight?
-
- Obstetricians suggest ultrasound examinations when something seems to be wrong.
- The investigators concluded that the ultrasound exams and low birthweights had a common cause-problem pregnancies.
- Later, a randomized controlled experiment was done to get more definite evidence.
- If anything, ultrasound was protective.

Controlando fatores de confundimento.

- Indivíduos podem diferir entre si de uma maneira crucial que não seja percebida em um primeiro momento.
- Algumas vezes estas diferenças podem ser encontradas depois.
- Então, ajustamentos podem ser feitos comparando subgrupos menores e mais homogêneos.
- Este processo é chamado *controle do fator de confusão*.
- Este é um TERCEIRO USO da palavra controle...

Discriminação sexual em Berkeley

- Admissão à pós-graduação na Universidade da Califórnia, Berkeley, em 1973.
- 8442 homens e 4321 mulheres se candidataram.
- 44% dos homens e 35% das mulheres foram admitidos.
- Assuma que homens e mulheres estão igualmente qualificados (e não havia nenhuma evidência contrária).
- A diferença na admissão parece indicar que homens e mulheres são tratados de forma diferente.
- A universidade parece preferir homens.

Caçando os culpados

- As admissões são feitas por cada departamento separadamente.
- Deveria ser possível identificar quais deles discriminam as mulheres.
- Olhando cada departamento, não havia nenhum viés contra as mulheres;
- Na verdade, na maioria deles, as mulheres eram favorecidas!!
- Como é possível que cada departamento favoreça a entrada de mulheres mas a universidade como um todo prefira homens??

Simpson's paradox

- Simpson's paradox
- a trend that appears in different groups of data disappears when these groups are combined.
- Mais estranho ainda: The REVERSE trend appear for the aggregate data.
- Exemplo: Década de 70, taxa de mortalidade no Brasil e na Suécia
- Considerando-se as taxas de mortes por 100 mil, a taxa do Brasil era MENOR que a da Suécia.
- No entanto, ao olhar em cada faixa etária separadamente, as taxas da Suécia eram menores que as brasileiras.
- Isto é, a mortalidade parecia ser melhor no Brasil como um todo mas, em cada uma das faixas etárias, ela era pior.

De volta a Berkeley

- Como explicar o caso de Berkeley?
- Algumas pós-graduações eram muito mais fáceis de entrar.
- Outra eram muito mais difíceis.
- Mais de 50% dos homens se candidataram às fáceis.
- Mais de 90% das mulheres se candidataram às difíceis.
- Isto explica tudo.

Imagem das duas redes

- Imagine duas redes no mar.
- Uma é bem fina e quase nenhum peixe passa por ela.
- A outra é larga e todo peixe consegue atravessá-la.
- Existe uma quantidade igual a peixes F e M.
- A rede fina e difícil recebe muitos F e poucos M: do outro lado, vemos alguns F e praticamente nenhum M.
- A rede fácil recebe poucos F e muitos M. Todos passam e do outro lado vemos poucos F e muitos M.
- Resultado global do outro lado: poucos F e muitos M.
- MAS AS REDES NÃO DISCRIMINAM M ou F.

Resumo

- Existia um fator de confundimento neste exemplo.
- Sexo era associado com a dificuldade do curso: mulheres buscavam os cursos mais difíceis de entrar.
- Sexo era (ligeiramente) associado com a aceitação: mulheres eram aceitas mais facilmente.
- O efeito global é resultante da combinação destes dois fatores.
- O resultado que aparece NO TODO é um incerto: pode parecer favorecer as mulheres, pode parecer neutro neutro ou pode parecer estar desfavorecendo as mulheres.

A great and recent example of Simpson's Paradox

- From Revolutions: R Blog, July 2013.
- Since 2000, the median US wage has risen about 1%, adjusted for inflation.
- But over the same period (2000-2013), the median wage for:
 - high school dropouts,
 - high school graduates with no college education,
 - people with some college education, and
 - people with Bachelor or higher degrees
- have all decreased.
- In other words, within every educational subgroup, the median wage is lower now (2013) than it was in 2000.

How can both things be true?

- Overall wages have risen, but wages within every subgroup have fallen?
- This is a great example of Simpson's Paradox.
- The explanation lies in the changing educational profile of the workforce over the past 13 years.
- There are now many more college graduates (who get higher-paying jobs) than there were in 2000.
- However, wages for college graduates collectively have fallen at a much slower rate (down 1.2%) than for those of lower educational attainment (whose wages have fallen precipitously, down 7.9% for high school dropouts).
- The growth in the proportion of college graduates swamps the wage decline for specific groups.