

PALESTRA 1

17 de outubro de 2011

Jonathan D. Mahnken, Ph.D., PStat®

Parte I - Esboço

- Tipos de estudos clínicos
 - Estudos Observacionais
 - Estudos Experimentais
 - Meta-análise

Tipos de Estudos

□ Observacional

- Diferentes grupos são examinados por diferenças
 - Resultados (exemplo: taxas de mortalidade, tempo de sobrevivência)
 - Exposições (exemplo: idade, raça, comportamento de saúde)

□ Experimental

- Envolve uma intervenção
- Os grupos são similares **w.r.t.** tudo exceto a intervenção (exposição)
- Seleccionados vs. não-seleccionados

Comentários Gerais

- Em ambos estudos observacionais e experimentais
 - Melhor inferência se as populações forem idênticas em todos aspectos exceto pela exposição
 - Um grupo é “substituto” viável para o outro exceto para o status de exposição
 - Ambos os grupos: amostras aleatórias da população-alvo
 - Única diferença entre os grupos sendo o status de exposição
 - Melhor resultado seria se os sujeitos pudessem voltar no tempo sob exatamente as mesmas condições *exceto status de exposição w.r.t*
 - Rothman e Greenland (1998) consideram isso uma experiência “contra factual”

Estudos Observacionais

❑ Estudo de série de casos

- Descreve um número menor de observações
- Leva à uma hipótese testável
- Útil para se desenvolver um estudo mais eficaz
 - Pode fornecer estimativas iniciais – muito brutas – de associações tais como razão de prevalência, proporção de diferenças — mas frequentemente pequena demais mesmo para este.
- Exemplo: uma observação clínica de quatro pacientes com uma doença comum.

Estudos Observacionais

❑ Estudo de Caso-controle

- Resultado conhecido – olhe para trás para determinar se foi exposto
- Estudo Retrospectivo
- Delineamentos pareados e não-pareados
- Razão de chances
- Exemplo: Pessoas com câncer no pulmão (casos) e controles hospitalares pareados inscritos
 - Olhe para trás para ver quem fumava?

Estudos Observacionais

- ❑ Estudo Transversal (cross-sectional)
 - Ponto isolado no tempo
 - Estudos de prevalência (pesquisas)
 - Nenhuma inferência causal
 - A causou B ou B causou A?
 - Razão de prevalência, proporção de diferenças
 - Exemplo, pesquisa que recolha:
 - Status de Tabagismo atual
 - Status de doença atual

Estudos Observacionais

□ Estudos de coorte

- Exposição conhecida – acompanhar para ver o resultado
- Estudo prospectivo
- Pode fornecer evidência de causalidade
- Coorte histórica
- Taxa de risco, diferença de risco
- Exemplo, identificar grupo de pessoas
 - Discernir seu status de tabagismo na linha de base
 - Follow up através do tempo para ver quem desenvolverá câncer

Estudos Observacionais

❑ Estudos de Coorte

- Direção da pesquisa correta para determinar a causalidade
- \$\$\$

❑ Estudos de caso-controle

- Mais rápidos, baratos, mais viável para resultados raros.
- Mais suscetível a viés especialmente viés de seleção.

Estudos Observacionais

- ❑ Estudos prospectivos
 - Determinam associações
 - Causalidade não pode ser provada
 - Sem intervenção
 - Entretanto fornece evidências
- ❑ Estudos de Caso
 - Exame aprofundado dos sujeitos
 - Não podem ser generalizáveis
- ❑ Vieses são uma grande preocupação em estudos observacionais porque auto-seleção em grupo de exposição pode confundir *gravemente e inconscientemente* os resultados

Estudos Observacionais

❑ Ensaios Clínicos Aleatórios

- Possui grupo de controle (comparação)
- Intervenção (exposição) distribuídas aleatoriamente
 - Randomização deve equilibrar as variáveis de confusão
- Prospectivo
- Cego
 - Ensaio Cego
 - Ensaio Duplo-cego
- Taxa de risco, diferença de risco
- Exemplo: inscrever fumantes no estudo
 - Alocar aleatoriamente para condições de tratamento ou controle
 - Follow-up para verificar que proporção deixa de fumar em cada grupo

Estudos Experimentais

❑ Ensaios Clínicos não-randomizados

- Possui grupo de controle (comparação)
- Intervenção (exposição) distribuída – mas não aleatoriamente
 - Variáveis de confusão não podem ser balanceadas
- Prospectivo
- Taxa de risco, diferença de risco
- Exemplo: identificar grupo de pessoas
 - Discernir seu status de tabagismo na linha de base
 - Follow up através do tempo para ver quem desenvolverá câncer
 - Outros fatores, como idade, equilibrados entre fumantes e não-fumantes?

Estudos Experimentais

□ Ensaios com autocontroles

- Cada sujeito atua como seu próprio controle
 - Recebem intervenção por um período
 - Não recebem nenhuma intervenção por um período
- Estudo Cruzado
- Medidas de diferença
 - Exemplo: peso numa dieta especial vs peso numa dieta normal
- Exemplo: inscrever fumantes no estudo
 - Em primeiro lugar, alocar aleatoriamente para tratamento ou controle.
 - Follow-up para verificar qual a proporção de quem para de fumar em cada grupo
 - Alternar para outra condição de tratamento
 - Follow-up para verificar qual a proporção de quem para de fumar em cada grupo

Estudos Experimentais

❑ Ensaios com controles externos

- Grupo de controle de outro estudo
- Controles históricos
- Outros fatores relacionados mudaram com o tempo?
- Razão de risco, diferença de risco
- Exemplo: identificar um grupo de fumantes e dar a todos um novo tratamento
 - Comparar o taxa de quem parou de fumar à taxa do estudo anterior

Estudos Experimentais

❑ Estudos sem controle

- Sem controle ou grupo de comparação
- Supõe que a intervenção (tratamento ou exposição) é melhor disponível
- Similar a estudo de caso da intervenção
- Razão de prevalência, proporção de diferenças
- Similar a ensaios com controles externos — mas não faz quaisquer comparações
 - Apenas descritivo

Estudos Experimentais

□ Ensaios Clínicos Randomizados

- Confundidores equilibrados através da randomização
- Evidência mais forte de causalidade
- \$\$\$

□ Ensaios Clínicos Não Randomizados

- Desprotegidas contra vieses na atribuição de tratamento

Estudos Experimentais

❑ Ensaios com autocontroles

- Observações bem equilibradas **w.r.t.** a maior parte era confundidores
- Deve haver um período apropriado de **falha/fiasco**

❑ Ensaios com controles externos

- Úteis quando curas não existem (AIDS)
- Mudanças temporais em casos contribuintes podem ter ocorrido

❑ Estudos sem controle

- Não há uma forma de determinar se a população amostrada era “normal”.

Meta-Análise

- ❑ Combina estudos
- ❑ Revisão de literatura quantitativa
- ❑ Útil quando a informação disponível for inconclusiva
 - Todos os estudos são de baixa potência
 - Resultados conflitantes
 - Pode apontar em direção à fatores relacionados a conflitos
- ❑ Estudos de tipos diferentes devem ser analisados separadamente.
 - Medida do resultado depende dos tipos de estudos
- ❑ *Use com cuidado!*

Resumo

- ❑ Quando feito corretamente, ensaios clínicos aleatórios criam a evidência mais forte para uma conexão causal
- ❑ Inferência Causal requer:
 - Substituto válido para o grupo de tratamento
 - Controle de vieses
 - Randomização é uma forma otimizada
 - Direção apropriada de averiguação

Resumo

- ❑ Como que o resultado no grupo exposto se compara com o resultado que se teria obtido se estes mesmos sujeitos pudessem reviver tudo identicamente com exceção de seu status de exposição?
 - Um substituto válido, imparcial responderá esta questão
 - Na realidade este substituto não existe
 - O valor da comparação feita entre indivíduos que se expuseram e os que não se expuseram dependerá da intensidade que o substituído chegará da experiência “contra factual”

Parte II – Esboço

- ❑ Introdução à probabilidade e inferência estatística
 - Probabilidade
 - Definições
 - Exemplos
 - Notação
 - Regras de Probabilidade
 - Tipos de probabilidades
 - Probabilidade conjunta
 - Probabilidade condicional
 - Independência estatística
- ❑ Populações e amostras
 - Tipos de amostras
- ❑ Parâmetros e estatísticas
- ❑ Distribuições
- ❑ Inferência Estatística
 - Estatística versus significância clínica

Inferência Estatística

□ Propósito da estatística:

- Para “generalizar as descobertas pelo conjunto de observações de um grupo de sujeitos com outros que são similares àqueles sujeitos” (p 64)
- Inferência
- População
- Amostra

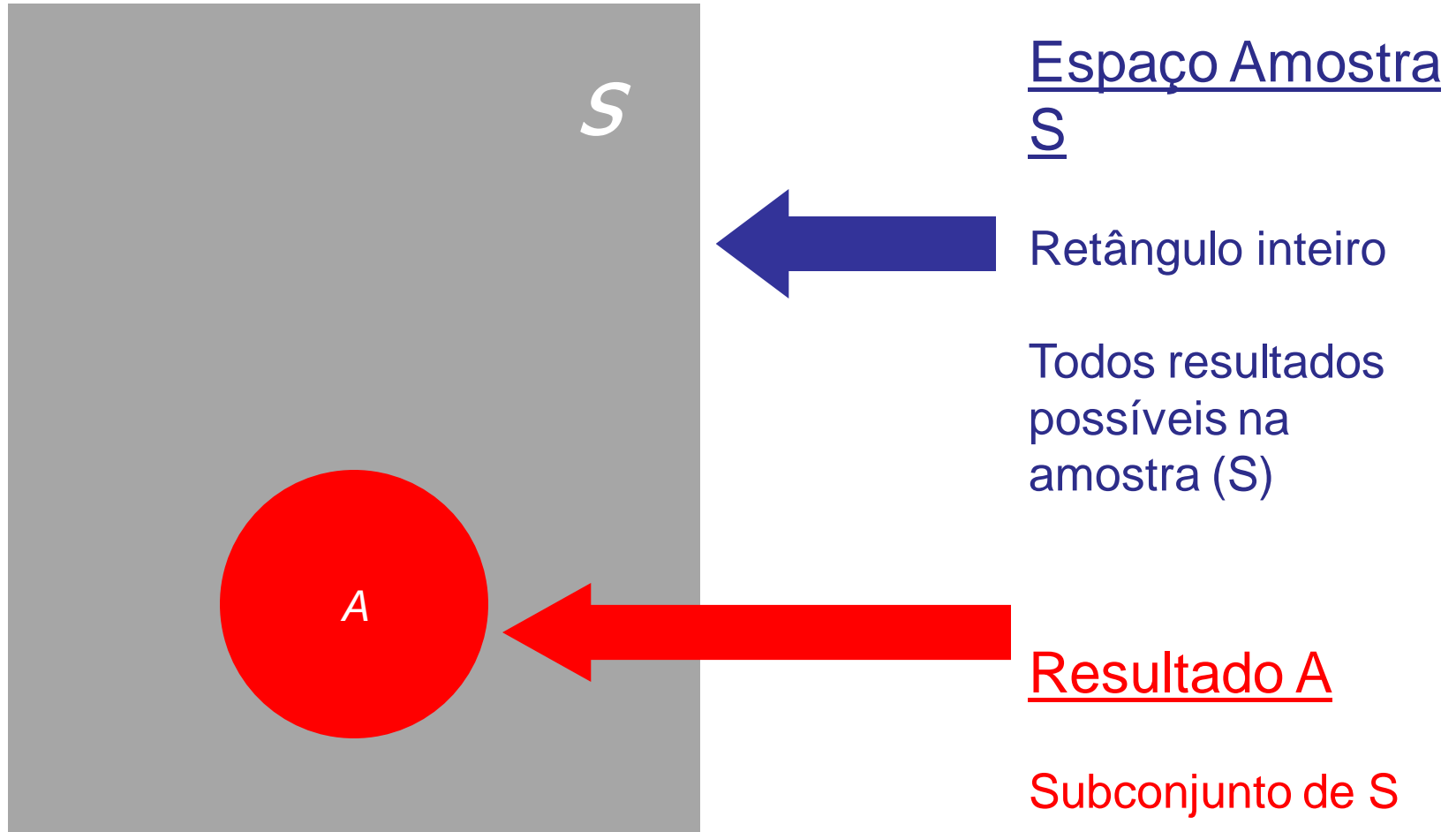
Probabilidade

$$\text{Probabilidade} = \frac{\text{Número de vezes que um resultado ocorre}}{\text{Número total de ensaios}}$$

- Ensaio, experimento, replicação
- $\Pr\{A\} \in [0, 1]$
- Exemplos
 - Ensaio: jogar uma moeda para o alto
 - $\Pr\{\text{Cara}\} = 0.5$
 - Ensaio: Jogar um par de dados
 - $\Pr\{\text{Soma} = 3\} = 2/36 \approx 0.056$

$$\Pr\{\text{Soma} = 3\} = \frac{\{(1,2), (2,1)\}}{\{(1,1), (1,2), (1,3) \dots (6,5), (6,6)\}}$$

Probabilidade



Exemplo: Jogar Um Par De Dados

Resultado A

Subconjuntos
onde a soma
das faces é
igual a 3



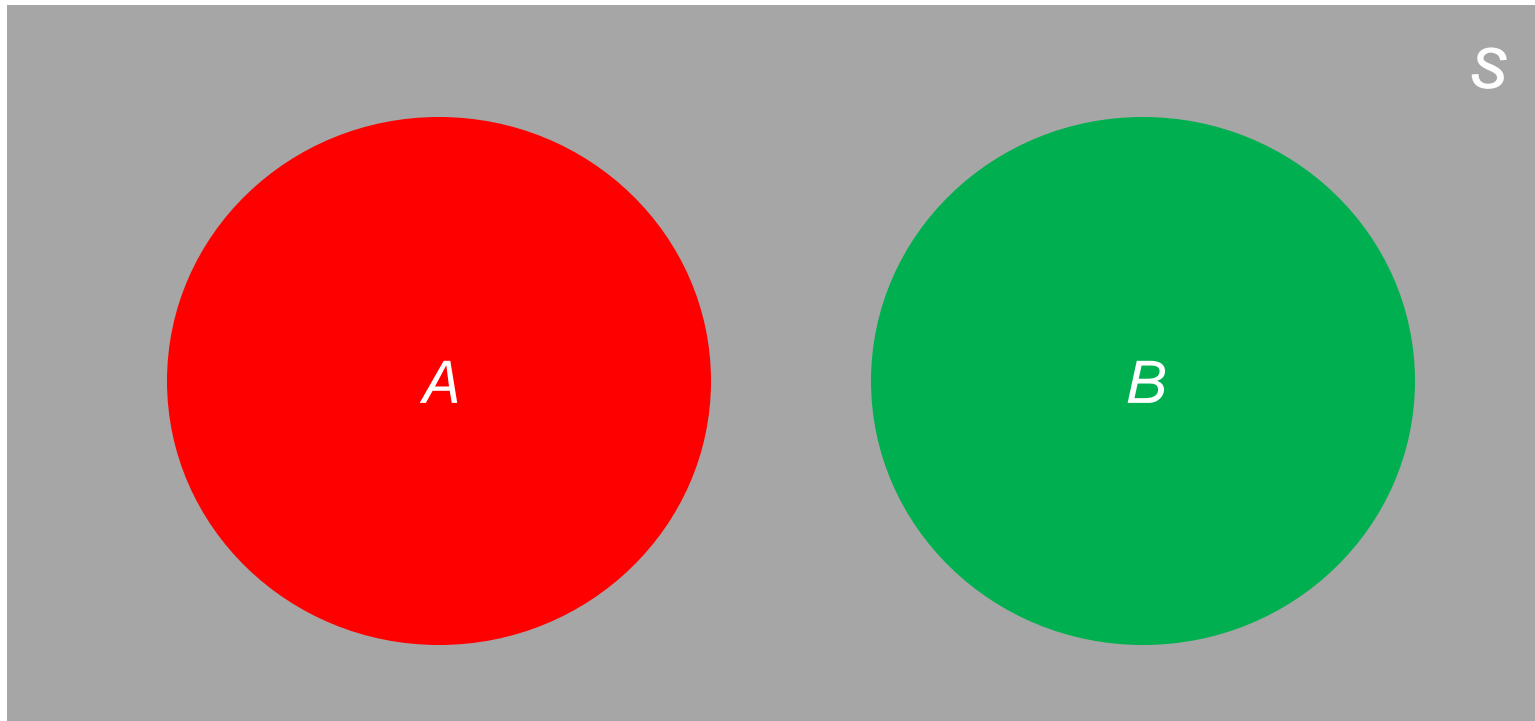
S

$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(1,6)$
$(2,1)$	$(2,2)$	$(2,3)$	$(2,4)$	$(2,5)$	$(2,6)$
$(3,1)$	$(3,2)$	$(3,3)$	$(3,4)$	$(3,5)$	$(3,6)$
$(4,1)$	$(4,2)$	$(4,3)$	$(4,4)$	$(4,5)$	$(4,6)$
$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$
$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$

Espaço Amostra
 S

Todas as possíveis
combinações das
duas faces.

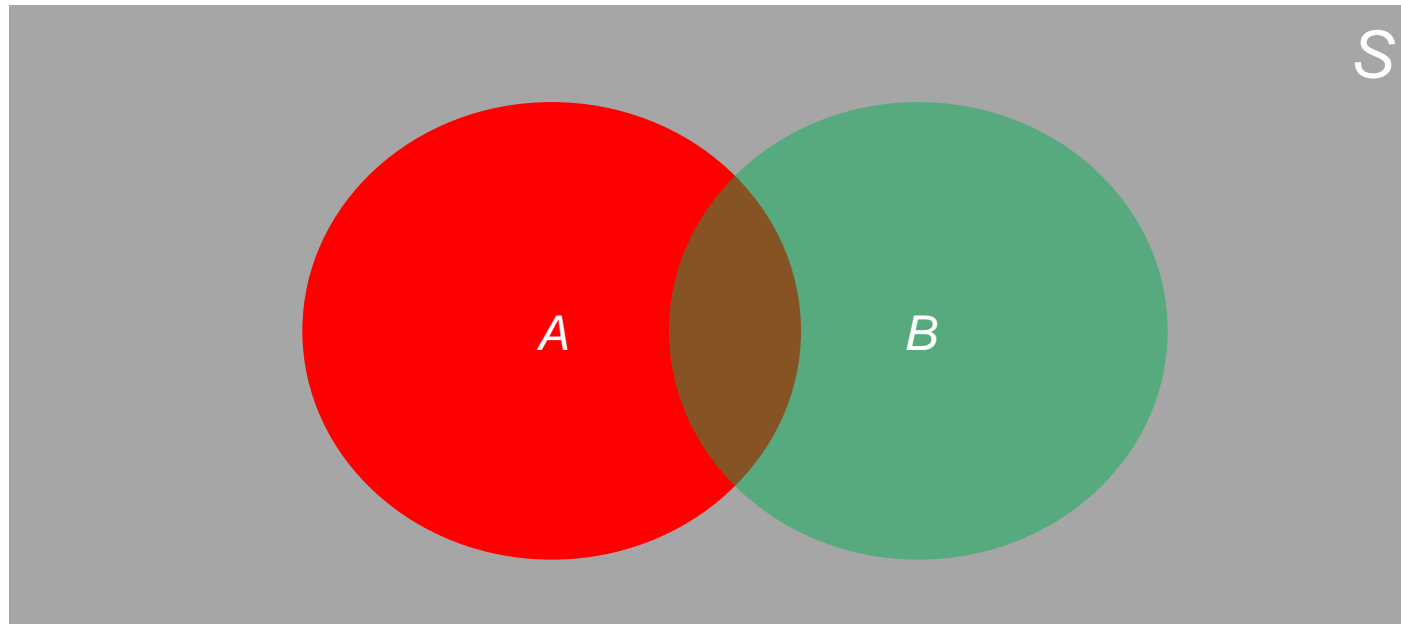
Probabilidade – Dois Eventos



□ Mutuamente Exclusivos

- A e B **não** se sobrepõe (sem resultados comuns)
- $P(A \text{ ou } B) = P(A \cup B) = P(A) + P(B)$
- $P(A \text{ e } B) = P(A \cap B) = 0$

Probabilidade – Dois Eventos

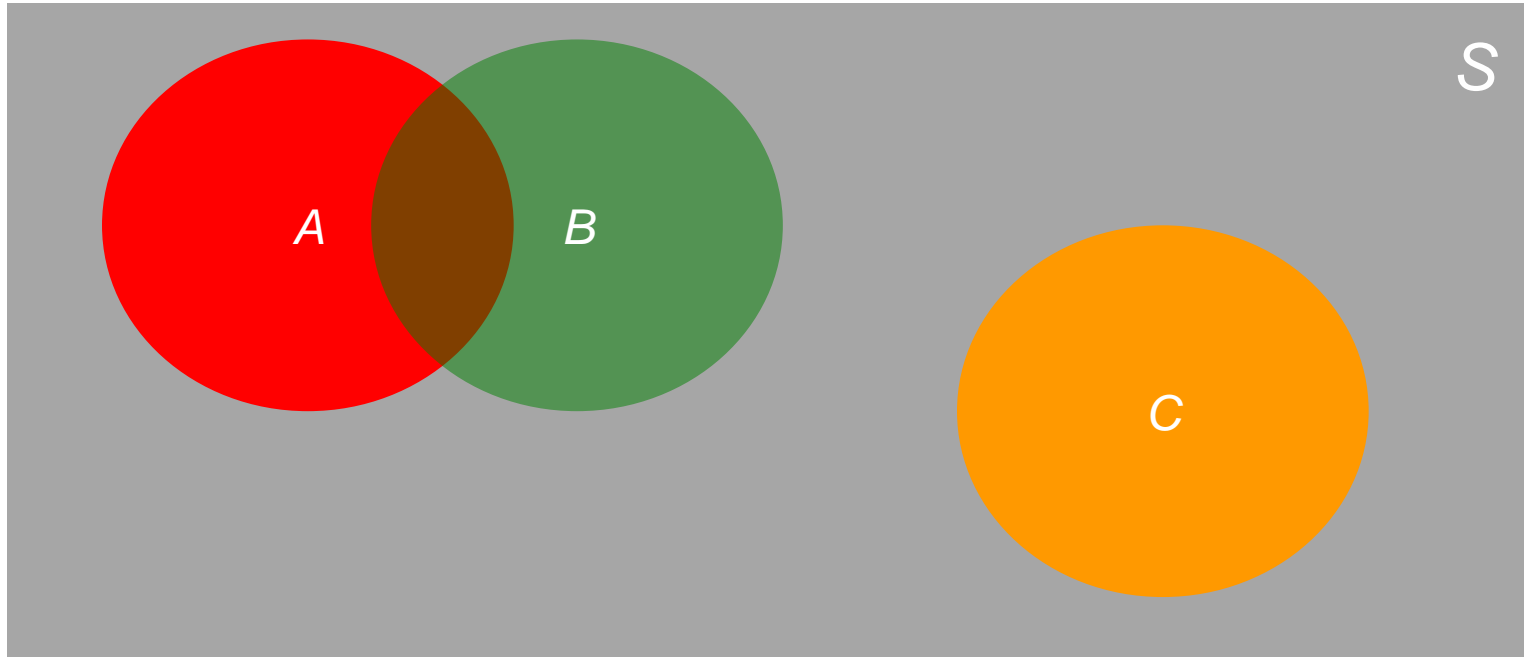


□ A e B *não* são mutuamente exclusivos

- A e B se sobrepõe (há resultados comuns)
- $P(A \text{ ou } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - $P(A)$ e $P(B)$ *ambos possuem* $P(A \cap B)$, então somá-los inclui a área de sobreposição duas vezes!
- $P(A \text{ e } B) = P(A \cap B) \neq 0$

Probabilidade – Mais de Dois

Eventos



- ❑ A e B **não** são mutuamente exclusivos
- ❑ A e C são mutuamente exclusivos
- ❑ B e C são mutuamente exclusivos

Regras de Probabilidade

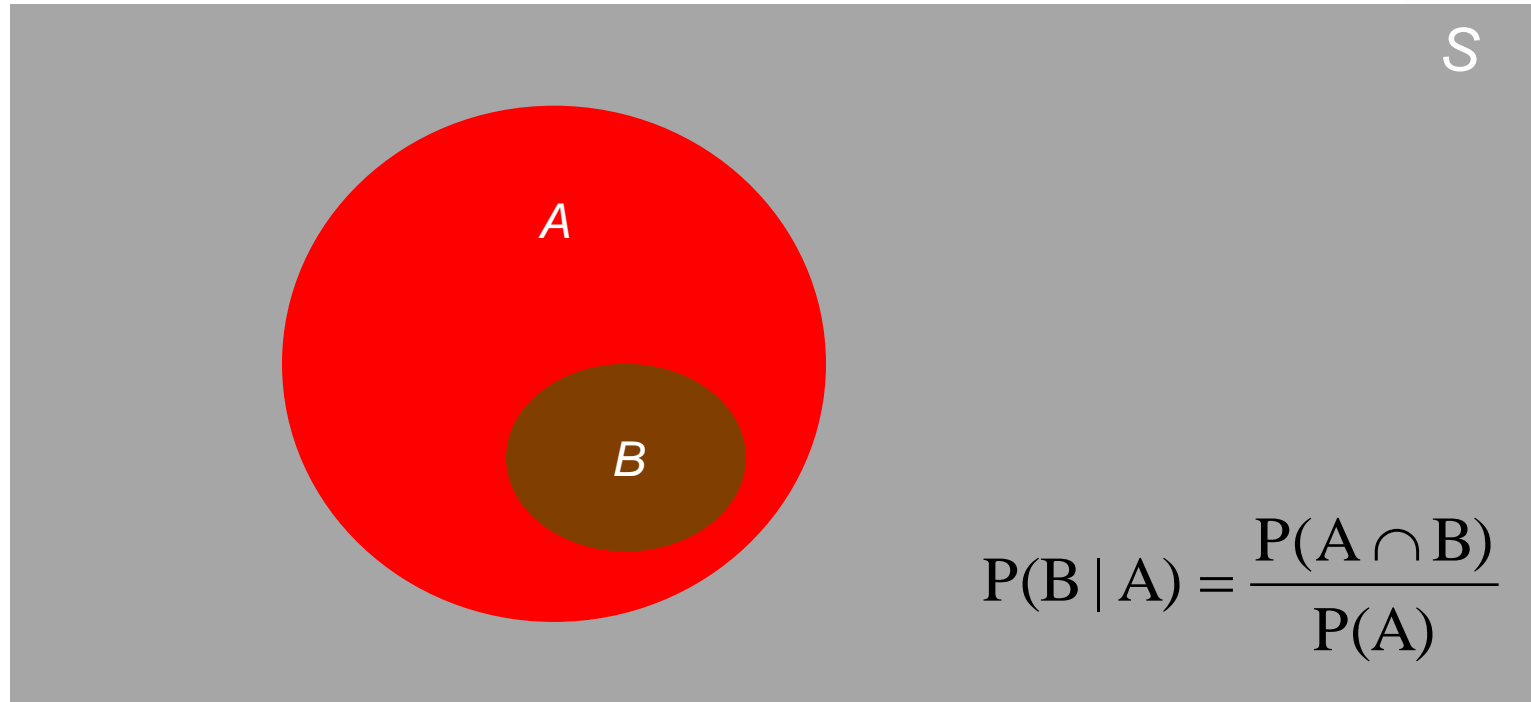
- $P(\text{conjunto vazio}) = P(\emptyset) = 0$
 - Probabilidade de evento que não ocorre no conjunto amostra S é zero
- $P(\text{não } A) = P(A^c) = 1 - P(A)$
 - Pois a soma das probabilidades no espaço da amostra S é igual a 1
- $P(A \text{ ou } B) = P(A \cup B)$
$$= P(A) + P(B) - P(A \cap B)$$
 - Se A, B são eventos mutuamente exclusivos, então $P(A \cap B) = 0$, assim $P(A \text{ ou } B) = P(A) + P(B)$

Probabilidade Conjunta

- $P(A \text{ e } B) = P(A \cap B)$ é chamada de probabilidade “conjunta” de A e B
 - \cap significa “e” (\cup significa “ou”)
- Exemplo: jogando um dado vermelho e um dado azul
 - $P(\text{dado vermelho} = 4 \cap \text{dado azul} = 3)$
- $P(A)$ e $P(B)$ são referidas como probabilidades “marginais”
- Qual a probabilidade conjunta de que alguém seja do sexo feminino e tenha tipo sanguíneo B? (Veja tabela no próximo slide.)

Tipo Sanguíneo	Probabilidades		
	Femenino	Masculino	Total
O	0,21	0,21	0,42
A	0,215	0,215	0,43
B	0,055	0,055	0,11
AB	0,02	0,02	0,04
Total	0,5	0,5	1,00

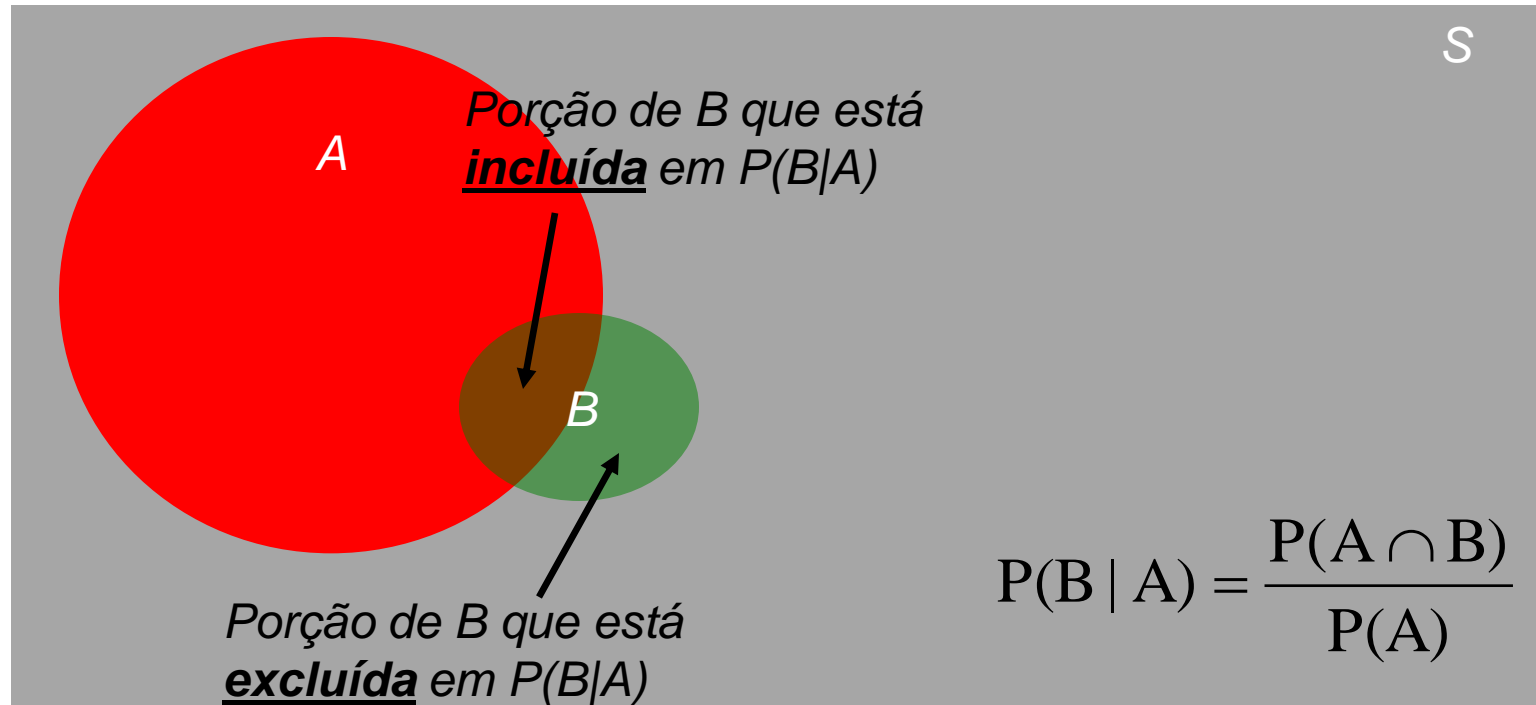
Probabilidade Condicional



□ $P(B|A)$ trata A como se fosse o espaço amostral (em vez de S)

$$P(B | A) = \frac{\text{Número de vezes que B ocorre dentro de A}}{\text{Número total de ensaios em A}}$$

Probabilidade Condicional



- Nem todos de B tem que estar contidos dentro de A
 - B não tem que ser um subconjunto de A, mas ele pode ser (como exibido no slide anterior)

Probabilidade Conjunta

□ Utilizando a relação anterior, temos

$$P(B | A) = \frac{P(A \cap B)}{P(A)}$$

Portanto

$$P(A \cap B) = P(B | A)P(A)$$

E da mesma forma

$$P(A \cap B) = P(A | B)P(B)$$

Independência

Definição:

A é independente de B se $P(A|B) = P(A)$

- Em palavras, se A é independente de B, então a probabilidade de $A = a_1$ (por exemplo) é a mesma independente do nível de B (tal como $B = b_1, b_2$ or b_3)
- Implicações de independência
 - $P(A \cap B) = P(A)P(B)$
- Veja um exemplo de variáveis independentes no próximo slide

Tipo Sanguíneo	Probabilidades		
	Feminino	Masculino	Total
O	0,21	0,21	0,42
A	0,215	0,215	0,43
B	0,055	0,055	0,11
AB	0,02	0,02	0,04
Total	0,5	0,5	1,00

Regras de Probabilidade

$$\square P(A \text{ ou } B) = P(A \cup B)$$

$$= P(A) + P(B) - P(A \cap B)$$

- $P(A \cap B) = 0 \Rightarrow A, B$ mutuamente exclusivos

$$\square P(A \text{ e } B) = P(A \cap B)$$

$$= P(A|B)P(B)$$

- $P(A|B) = P(A) \Rightarrow A, B$ independentes
- $P(A \cap B) = P(A)P(B) \Rightarrow A, B$ independentes

Teorema de Bayes

$$P(B | A) = \frac{P(A \cap B)}{P(A)}$$

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)}$$

$$P(B | A) = \frac{P(A | B)P(B)}{P(A \cap B) + P(A \cap B')}$$

$$P(B | A) = \frac{P(A | B)P(B)}{P(A | B)P(B) + P(A | B')P(B')}$$

- O Teorema de Bayes é útil pois “investigadores frequentemente só sabem uma das probabilidades pertinentes e devem determinar a outra.” (p 68)

Teorema de Bayes

□ Probabilidade

■ Exemplo

- Paciente apresenta dor no peito
- Causas possíveis:
 - MI (infarto do miocárdio), pneumonia, esofagite de refluxo
- Conhecidas:
 - $P(\text{dor no peito} \mid \text{Infarto do miocárdio})$
 - $P(\text{dor no peito} \mid \text{pneumonia})$
 - $P(\text{dor no peito} \mid \text{esofagite de refluxo})$
- Problema
 - Qual é a $P(\text{Infarto do miocárdio} \mid \text{dor no peito})$?
 - Pode-se resolver com o Teorema de Bayes se $P(\text{dor no peito})$ e $P(\text{Infarto do miocárdio})$ forem também conhecidas

Populações e Amostras

□ População

- Conjunto ou coleção de itens com algo em comum
- População possui sua própria distribuição

□ Amostra

- Subconjunto da população da qual é retirada
- Seleccionada para ser “representativa” da população
- Descrições (estatísticas) de uma amostra representativa podem ser generalizadas à população-fonte da qual a amostra foi retirada

Razões para Amostragem

- ❑ Censo não pode ser possível
- ❑ Se um censo for possível, a amostragem é
 - Mais rápida
 - Mais barata
 - Mais “precisa”
 - Maiores informações podem ser obtidas a partir de cada observação
 - Útil para a inferência

Métodos de Amostragem

□ Amostragem probabilística

■ Amostra aleatória simples

- Todos na população possui a mesma probabilidade de serem selecionados na amostra

■ Amostra Sistemática

- Cada sujeito k th é selecionado
- $k = [\text{número na população}] / [\text{tamanho estimado da amostra}]$
- A primeira observação selecionada deve ser feita em aleatório.

Métodos de Amostragem

❑ Amostragem probabilística

■ Amostragem estratificada

- População dividida em estratos (subgrupos)
 - Geralmente os subgrupos são níveis diferentes de variáveis de confusão
- Amostra aleatória retirada de dentro de cada estrato
- Ex. – estratificar por idade e sexo

■ Amostragem por Clusters

- População dividida em grupos de temas como
 - Famílias, blocos, escolas, etc.
- Clusters selecionados aleatoriamente
- Requer um tamanho maior de amostra pois as observações são correlacionadas dentro do cluster

Métodos de Amostragem

□ Amostragem Não Probabilística

- A probabilidade de que um sujeito da população-fonte fosse selecionado na amostragem é desconhecida
- Nenhuma atribuição aleatória
 - Vieses
 - Randomização é usada para equilibrar *tudo* exceto a intervenção (exposição)
 - Remove todas as vieses
 - A amostra é representativa da população
 - » Resultados podem ser generalizados à população

Amostragem

- ❑ População-alvo
- ❑ População amostrada
 - Base de amostragem
- ❑ Amostra
- ❑ Os resultados só podem ser generalizados à população-alvo se a amostra for representativa
 - Propriedades de distribuição serão similares

Parâmetros versus Estatísticas

- Parâmetros são características que descrevem uma *população*
 - Letras Gregas

- Estatísticas são características que descrevem uma *amostra*
 - Letras Romanas

Variáveis Aleatórias e Distribuições de Probabilidade

□ Variável Aleatória

- Uma série de observações com uma estrutura de probabilidade correspondente a cada resultado
 - Observações podem ser numéricas ou categóricas
 - Exemplo – altura dos pacientes na UTMB

□ Distribuição de Probabilidade

- Probabilidades que correspondem a cada resultado de uma variável aleatória
- pdf (Função de distribuição de probabilidade)
 - Para variáveis aleatórias contínuas
- pmf (função massa de probabilidade)
 - Para variáveis aleatórias categóricas

Distribuição Binomial

- Evento dicotômico (sucesso ou fracasso)
 - Ensaio de Bernoulli
 - π = probabilidade de sucesso
- $B(n, \pi)$
 - n = número de ensaios
 - π = probabilidade de sucesso
 - $X \sim B(n, \pi)$ então X é a soma de n ensaios *independentes* de Bernoulli
 - $Pmf = P(X=x) = \{n!/[x!(n-x)!]\} \pi^x (1-\pi)^{n-x}$
 - Exemplo: $X \sim B(10, 0,8)$
 - $P(X=0) = 1 \cdot (0,8)^0 (1-0,8)^{10-0}$
 - $P(X=1) = 10 \cdot (0,8)^1 (1-0,8)^{10-1}$

Distribuição de Poisson

□ Dados de “Contagem”

□ Se $X \sim B(n, \pi)$ e

- $n \rightarrow \infty$ (n é grande)
- $\pi \rightarrow 0$ (π é pequeno)
- $n\pi \rightarrow \lambda$ (λ é a média)

então X pode ser aproximado pela
Distribuição de Poisson – $\text{Poi}(\lambda)$

□ Média = variância

- Variância = λ

□ Pmf = $P(X = x) = \lambda^x e^{-\lambda} / x!$

- Exemplo: $X \sim \text{Poi}(3,22)$
 - $P(X = 0) = 3,22^0 e^{-3,22} / 0!$
 - $P(X = 1) = 3,22^1 e^{-3,22} / 1!$

Distribuição Normal

□ Variável aleatória contínua

- Gaussiana
- Curva em formato de sino
- Simétrica
- $(-\infty, \infty)$

□ $N(\mu, \sigma^2)$

- μ = média
- σ^2 = variância
- Normal padronizada $N(0,1)$
 - Distribuição z
 - Escore z (1,645, 1,96)

□ Pdf

- $f(x) = [1/(2\pi\sigma^2)^{1/2}] \exp[-(x-\mu)^2/\sigma^2]$

Distribuição Normal

- ❑ "Na verdade, poucas características são normalmente distribuídas." (p 80)
- ❑ "Inferência estatística geralmente envolve valores médios de uma população, e não os valores relacionados aos indivíduos." (p 80)
- ❑ "...se tivermos que fazer declarações de probabilidade sobre indivíduos utilizando as regras de desvio médio e de desvio padrão, a distribuição da característica de interesse deve ser aproximadamente normalmente distribuída." (p 80)

Distribuições de Amostragem

- ❑ Distribuição de amostragem é a distribuição das médias
 - Menos variação que distribuição de indivíduos
- ❑ Teorema do Limite Central

Teoria do Limite Central

\bar{X} = *média amostral*

$$\text{Média}(\bar{X}) = E[\bar{X}] = \mu$$

$$\text{Variância}(\bar{X}) = V[\bar{X}] = \sigma^2 / \sqrt{n}$$

Para amostra grande (grande n),

$$\bar{X} \overset{\text{approx}}{\sim} N(\mu, \sigma^2 / \sqrt{n})$$

Desvio Padrão versus Erro-padrão

- ❑ Desvio padrão é a medida de variabilidade entre as *observações* numa *população*
- ❑ Erro-padrão é a medida da variabilidade entre as *estatísticas* (médias) numa *amostra*

Inferência Estatística

- ❑ Descreve os parâmetros populacionais baseado nas estatísticas da amostra
 - Estimativa
 - Peso médio entre pessoas de 65 anos de idade é de 170 lbs. (95% CI de 162-178 lbs.)
 - Intervalos de Confiança
 - Teste de hipótese
 - H_0 : peso médio entre pessoas de 65 anos de idade é 170 lbs.
 - H_1 : peso médio entre pessoas de 65 anos de idade é maior que 170 lbs.
 - Valor p
- ❑ Ambos possuem a mesma conclusão em qualquer nível de "confiança" dado

Inferência Estatística

- ❑ “Boa” estimativa da amostra terá as seguintes propriedades
 - Imparcial
 - Variância mínima
- ❑ O objetivo é ser capaz de descrever a população baseado nas informações obtidas a partir da amostra
 - Combine observações similares e mantenha observações diferentes separadas.

Significância

- ❑ Significância estatística é uma função de
 - Suposições do modelo
 - Tamanho da Amostra
 - Tamanho da diferença
 - Precisão
- ❑ Resultados estatisticamente significantes nem sempre são clinicamente significantes
- ❑ Significância Clínica é uma diferença biológica ou cientificamente importante
- ❑ Resultados clinicamente significativos nem sempre são estatisticamente significantes

Referência

- ❑ Dawson B and Trapp RG (2001). *Basic & Clinical Biostatistics*, 3rd ed., McGraw Hill: New York