

## Tabelas Bidimensionais

### Unidade I Parte 2

## RAZÃO DE CHANCES ( *ODDS RATIO* )

- Medida de associação para tabelas de contingência 2x2

Grupo	Y		Total
	Sucesso	Insucesso	
1	$n_{11}$	$n_{12}$	$n_{1+}$
2	$n_{21}$	$n_{22}$	$n_{2+}$
Total	$n_{+1}$	$n_{+2}$	$N$

$\pi_1$  - probabilidade de sucesso na linha 1 (grupo1)

$\pi_2$  - probabilidade de sucesso na linha 2 (grupo2)

## "Odds"

Define-se para:

■ linha 1:  $odds_1 = \frac{\pi_1}{(1 - \pi_1)}$

■ linha 2:  $odds_2 = \frac{\pi_2}{(1 - \pi_2)}$

- Em qualquer das linhas, a probabilidade de sucesso é uma função da *odds*.

$$odds = \frac{\pi}{1 - \pi} \Leftrightarrow \pi = \frac{odds}{odds + 1}$$

- Quando  $\pi_1 = \pi_2$ , as odds satisfazem  $odds_1 = odds_2$ . As variáveis são, portanto, **independentes**.

## "Odds Ratio"

- É a razão das "odds" de dois grupos (linhas):

$$\theta = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)}$$

- A **odds ratio** pode ser qualquer número não negativo.
- Quando X e Y são independentes ( $\pi_1 = \pi_2$ ) então  $\theta = 1$ .
  - Quando  $1 < \theta < +\infty$  a **odds** de sucesso é maior na linha 1 do que na linha 2, consequentemente,  $\pi_1 > \pi_2$ .
  - Quando  $0 < \theta < 1$  a **odds** de sucesso é menor na linha 1 do que na linha 2, consequentemente,  $\pi_1 < \pi_2$ .

## "Odds Ratio"

- Valores de  $\theta$  longe de 1 em uma dada direção representam os níveis de intensidade da associação.
  - Dois valores para  $\theta$  representam o mesmo nível de associação, mas em direções opostas, quando um valor é o inverso do outro.

Exemplo:

$\theta = 0,25 \rightarrow$  odds de sucesso da linha 1 é 0,25 vezes a odds de sucesso da linha 2;

$\theta' = 1/0,25 = 4 \rightarrow$  odds de sucesso da linha 2 é 4 vezes maior do que a da linha 1.

## "Odds Ratio" - Propriedades

- Quando a ordem das linhas ou das colunas é trocada, o novo valor de  $\theta$  é o inverso do valor original.
- A odds ratio trata as variáveis simetricamente, isto é, a odds ratio não muda de valor quando a orientação da tabela muda de sentido - linhas tornam-se colunas e colunas tornam-se linhas.
  - O valor da odds ratio independe da identificação da variável resposta.
  - Já o risco relativo depende da identificação de qual é a variável resposta e, qual dos dois níveis será considerado como sucesso.

## "Odds Ratio" - Propriedades

- Quando ambas as variáveis são variáveis respostas, a odds ratio pode ser definida usando probabilidades conjuntas como:

$$\theta = \frac{\pi_{11} \cdot \pi_{22}}{\pi_{12} \cdot \pi_{21}}$$

A odds ratio é chamada também de razão de produtos cruzados.

## Inferência para Odds Ratio

- Estimador da odds ratio:

$\Rightarrow$  Odds ratio amostral

$$\hat{\theta} = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} = \frac{n_{11} / n_{12}}{n_{21} / n_{22}} = \frac{n_{11} \cdot n_{22}}{n_{12} \cdot n_{21}}$$

- estimador de máxima verossimilhança da odds ratio, para os esquemas de amostragem padrão
- Distribuição amostral: - altamente assimétrica para amostras de tamanho pequeno a moderado.

$\Rightarrow$  usar  $\ln(\hat{\theta})$ .

## Inferência para Odds Ratio

### ■ Distribuição amostral de $\ln(\hat{\theta})$ .

Para amostras grandes,

$$\ln(\hat{\theta}) \approx N(\ln \theta, ASE(\ln \theta))$$

onde:

$$ASE(\ln \hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}.$$

( Erro padrão  
assintótico )

### ■ Intervalo de Confiança para de $\ln(\theta)$

$$\ln \hat{\theta} \pm z_{\alpha/2} ASE(\ln \hat{\theta}).$$

- Tomando-se a exponencial em ambos os limites finais deste intervalo, se constrói um para  $\theta$ .

## Inferência para Odds Ratio

### ■ Observação:

Se qualquer célula  $n_{ij} = 0 \rightarrow \hat{\theta}$  é igual a 0 ou  $\infty$ .

ou

se ambas as células de uma linha ou de uma coluna são nulas

$\rightarrow \theta$  é indefinido.

Um estimador levemente corrigido de  $\theta$  é dado por:

$$\hat{\theta} = \frac{(n_{11} + 0,5)(n_{22} + 0,5)}{(n_{12} + 0,5)(n_{21} + 0,5)}$$

não tem este problema.

- É preferido quando as frequências das células são pequenas ou existe alguma frequência nula.

Neste caso na fórmula do ASE usa-se  $\{n_{ij} + 0,5\}$  em vez de  $\{n_{ij}\}$ .