

Utilização de estratificação e modelo de regressão logística na análise de dados de estudos de caso-controle.

Suely Godoy Agostinho Gimeno

Departamento de Epidemiologia da Faculdade de Saúde Pública, SP - Brasil

José Maria Pacheco de Souza

Universidade de São Paulo, SP - Brasil

Aluna

Fabiana Grifante

Índice

- 1 Descrição do problema e metodologia
- 2 Análise Bruta
- 3 Análise Estratificada
- 4 Regressão Logística
- 5 Conclusões
 - Comparações gráficas
 - Tipo de análise

Objetivo e descrição do problema

Objetivo => Câncer de esôfago
Tipo de Estudo: Caso - Controle

85 CASOS → 292 Controles

Variáveis:

- ✓ Idade
- ✓ Sexo
- ✓ Hábito de Beber
- ✓ Hábito de Fumar
- **"Odds Ratio"**

Análise Bruta dos Dados

Tabela - Resultados da análise bruta

Variável	<i>Odds ratio</i>	X^2_{MH}	ρ
Hábito de Beber	6,91	35,88	0,00
Hábito de fumar	8,93	33,09	0,00
Sexo	3,41	17,83	0,00
Idade	0.80	0,82	0,37

(Idade excluída das demais análises)

Análise Estratificada dos Dados

Tabela - Análise estratificada para hábito de beber, controlando-se hábito de fumar.

Estrato		Caso	Controle	Total	Odds ratio	Intervalo com 95% de confiança
Não fumante:	Bebe	1	32	33	0,54	0,06 - 4,78
	Não Bebe	5	86	91		
Fumante:	Bebe	74	120	194	6,66	2,55 - 17,41
	Não Bebe	5	54	59		
Total		85	292	377	4,50*	2,11 - 9,61

*Estimativa do *odds ratio* ponderado de Mantel-Haenszel (OR_{MH})

Teste de homogeneidade entre os estratos: $\chi^2_{1gl} = 4,27$ ($p = 0,04$)

Análise Estratificada dos Dados

Tabela - Análise estratificada para hábito de beber, controlando-se sexo.

Estrato		Caso	Controle	Total	Odds ratio	Intervalo com 95% de confiança
Feminino:	Bebe	8	35	43	2,69	0,94 - 7,71
	Não Bebe	8	94	102		
Masculino:	Bebe	67	117	184	13,17	3,10 - 55,99
	Não Bebe	2	46	48		
Total		85	292	377	6,28	2,91 - 13,55

*Estimativa do *odds ratio* ponderado de Mante-Haenszel (OR_{MH})

Teste de homogeneidade entre os estratos: $\chi^2_{1gl} = 3,03$ ($p = 0,08$)

Análise Estratificada dos Dados

Tabela - Análise estratificada para hábito de beber, controlando-se sexo e hábito de fumar.

Sexo	Hábito de fumar		Caso	Controle	Total	Odds ratio	Intervalo com 95% de confiança
Feminino	Não fumante:	Bebe	-	20	20	0	
		Não Bebe	5	71	76		
	Fumante:	Bebe	8	15	23	4,09	0,93-17,92
		Não Bebe	3	23	26		
Masculino	Não fumante:	Bebe	1	12	13	indefinido	
		Não Bebe	-	15	15		
	Fumante:	Bebe	66	105	171	9,74	2,26-42,07
		Não Bebe	2	31	33		
Total		85	292	377	4,79*	2,09 - 10,99	

*Estimativa do odds ratio ponderado de Mantel-Haenszel (OR_{MH})

Teste de homogeneidade entre os estratos: não foi possível (caselas vazias)

Análise com regressão logística

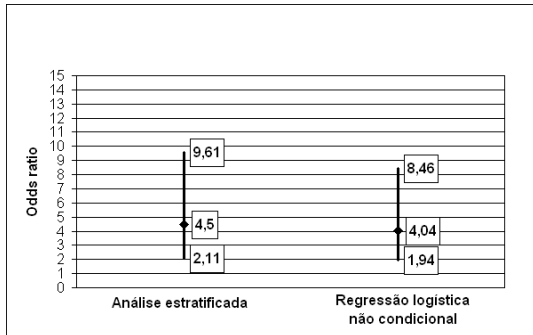
Tabela - Análise com regressão logística não condicionada para a variável hábito de beber, controlando-se (separadamente) hábito de fumar e sexo.

Modelo	Variável de controle	Hábito de beber <i>Odds ratio</i>	Intervalo com 95% de confiança
1	Hábito de fumar	4,04	1,94 - 8,46
2	Hábito de fumar	0,54	0,06 - 4,78
	Hábito de fumar e hábito de beber	12,39	2,55 - 17,40
3	Sexo	5,42	2,56 - 11,50
4	Sexo	2,69	0,94 - 7,71
	Sexo e hábito de beber	4,90	0,82 - 29,38

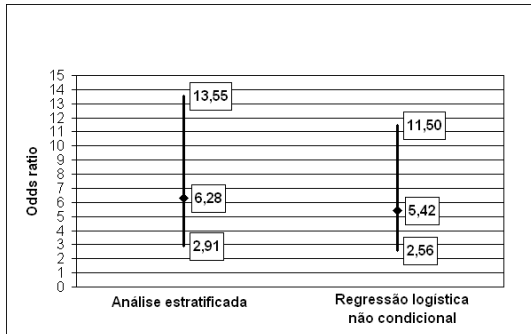
Análise com regressão logística

Tabela - Análise com regressão logística não condicionada para a variável hábito de beber, hábito de fumar e sexo (simultaneamente).

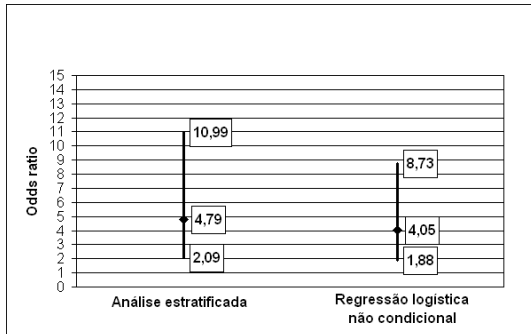
Variáveis	<i>Odds ratio</i>	Intervalo com 95% de confiança
Hábito de beber (variável principal)	4,05	1,88 - 8,73
Hábito de fumar	5,01	1,92 - 13,12
Sexo	1,00	0,49 - 2,05



Odds ratio e intervalo com 95% de confiança para hábito de beber, controlando-se o hábito de fumar, segundo tipo de análise.



Odds ratio e intervalo com 95% de confiança para hábito de beber, controlando-se sexo, segundo tipo de análise.



Odds ratio e intervalo com 95% de confiança para hábito de beber, controlando-se sexo e hábito de fumar, segundo tipo de análise.

Regressão Logística

- + Controla grande número de variáveis simultaneamente.
- + É a técnica estatística apropriada quando a variável dependente é categórica.
- Uso de recursos computacionais.
- Barreira entre o pesquisador e os dados.

Análise Estratificada

- + Simplicidade de execução.
- + Proximidade dos dados.
- + Facilidade de entedimento.
- Dificuldade quando cresce o número de variáveis.
- Menor poder de exploração das variáveis.

Conclusões

- Os intervalos de confiança construídos para estimativa da *Odds ratio* para regressão logística não condicional foram mais estreitos que os construídos para análise estratificada.
- Observou-se consistência entre os resultados obtidos com a aplicação da análise estratificada e regressão logística.