

## O USO DE TABACO POR ESTUDANTES DE PRIMEIRO E SEGUNDO GRAUS EM DEZ CAPITALS BRASILEIRAS: POSSÍVEIS CONTRIBUIÇÕES DA ESTATÍSTICA MULTIVARIADA PARA A COMPREENSÃO DO FENÔMENO\*

Maria Tereza Serrano Barbosa\*\*  
Beatriz Carlini-Cotrim\*\*  
Armando Ramos Silva-Filho\*\*

---

BARBOSA, M.T.S. et al. O uso do tabaco por estudantes de primeiro e segundo graus em dez capitais brasileiras: possíveis contribuições da estatística multivariada para a compreensão do fenômeno. *Rev. Saúde públ., S. Paulo*, 23:401-9, 1989.

**RESUMO:** Analisou-se o uso de tabaco por estudantes de primeiro e segundo graus de dez capitais brasileiras, utilizando técnicas estatísticas bi e multivariadas (modelo logit). Pela análise bi-variada detectou-se que o uso de tabaco está positivamente associado ao fato de o estudante trabalhar, freqüentar o curso noturno, estar atrasado nos estudos e ter pais fumantes. Pela aplicação de técnicas multivariadas e do modelo logit, verificou-se que as variáveis de maior peso são a defasagem escolar (no curso noturno) e o fato de os pais fumarem (diurno). Discutiu-se a importância da utilização de técnicas multivariadas para a compreensão de fenômenos complexos como o do uso de tabaco e outros psicotrópicos.

**DESCRIPTORIOS:** Fumo. Análise multivariada. Risco.

---

### INTRODUÇÃO

Pretende-se analisar, a partir de técnicas estatísticas bi e multivariada as informações sobre o uso de tabaco por estudantes de primeiro e segundo graus da rede estadual de dez capitais brasileiras (Belém, Brasília, Belo Horizonte, Curitiba, Fortaleza, Porto Alegre, Recife, Rio de Janeiro, Salvador e São Paulo), pesquisadas recentemente por Carlini e col.<sup>5</sup> (1989).

Trabalhos anteriores<sup>3,11,12</sup> apontam para um uso de tabaco (uso na vida), entre estudantes de primeiro e segundo graus, que varia de 21,5% em escolas paulistas<sup>11</sup>, a 33,7% em escolas do Distrito Federal<sup>3</sup> passando por um índice de 30,51% em Ribeirão Preto, interior do Estado de São Paulo<sup>11</sup>. No entanto, estes estudos tratam de realidades setoriais, específicas, não permitindo uma visão global das várias regiões do país.

Por considerar-se o uso de tabaco por jovens e adolescentes um fenômeno de inúmeras causas, procurou-se fugir da análise mais freqüentemente utilizada no Brasil, nesta área<sup>1, 3,4</sup>, onde se busca detectar separadamente a existência de relações entre as diversas variáveis (por exemplo: sexo,

idade, situação escolar ou familiar) e a variável de interesse (uso de tabaco, álcool ou outros produtos psicotrópicos). Estas relações são encontradas geralmente a partir do cálculo da estatística qui-quadrado em tabelas com duas variáveis, sob o controle de uma terceira. Este tipo de análise bivariada permite identificar, por exemplo, que o uso de tabaco está relacionado com a idade, ou com o fato de o estudante trabalhar, mas não permite compor de que forma estes e outros fatores se complementam ou interagem na influência que exercem sobre as chances de o estudante fazer ou não parte do grupo dos fumantes. Além disso, não há, de antemão, a determinação de uma variável resposta que seria explicada por um conjunto de variáveis auxiliares. No teste do qui-quadrado, a interpretação das associações significantes fica a cargo do autor ou leitor dos dados.

O presente trabalho objetiva superar estas limitações, utilizando a análise multivariada. Com este tipo de análise, procurou-se identificar quais combinações de categorias das variáveis medidas caracterizam uma situação de risco de o estudante ser fumante (variável resposta). Além disso, cada tabela foi ajustada a um modelo Logit para quan-

---

\* Trabalho realizado com o apoio financeiro do Ministério da Saúde e da Associação Fundo de Incentivo à Psicofarmacologia (AFIP).

\*\* Departamento de Psicobiologia da Escola Paulista de Medicina e da Associação Fundo de Incentivo à Psicofarmacologia – Rua Botucatu, 862 – 1º andar – 04023 – São Paulo, SP – Brasil.

tificar os efeitos individuais ou conjuntos das diversas variáveis na formação dessas chances.

## MATERIAL E MÉTODO

Os dados foram obtidos a partir da aplicação de um questionário de auto-preenchimento a respeito do uso de drogas psicotrópicas, respondidos por 16.149 estudantes da rede estadual, de primeiro e segundo graus, de dez capitais brasileiras<sup>5</sup>, em 1987.

A amostra de cada capital estudada foi obtida em dois estágios, após uma estratificação realizada de acordo com a localização e o tamanho das escolas. No primeiro estágio sortearam-se as escolas e, no segundo, as turmas (classes) a serem pesquisadas.

Para a presente análise foram considerados conjuntamente os estudantes das dez capitais. Esta decisão deveu-se principalmente ao comportamento semelhante das diversas variáveis nas cidades em relação ao hábito de fumar e mesmo de usar outras drogas<sup>5</sup>, mas, também por desejar-se uma visão global do uso de tabaco no país.

O questionário continha cinco perguntas a respeito de cigarro (uso na vida, uso no ano, uso no mês, quantidade de cigarros fumados na semana e idade que fumou pela primeira vez). Elegueu-se o uso de tabaco no último mês como a variável resposta a ser analisada, uma vez que o interesse do estudo era o uso *atual* de cigarros. O uso semanal não foi escolhido por praticamente coincidir com o uso mensal. O aluno que consumiu tabaco no último mês foi denominado *fumante* no presente trabalho.

## RESULTADOS

Como pode ser visto na Figura, o uso de tabaco na vida situou-se em 19,5%, o uso no último ano em 15,9% e o no último mês, assim como o semanal, atingiu 10,5%.

Os resultados do último mês, principal objeto deste trabalho, são apresentados de duas formas (bi e multivariada), visando ilustrar as diferenças entre estas duas formas de tratamento.

### Resultados Bivariados

A Tabela 1 apresenta os números e as percentagens de fumantes segundo as diversas caracte-

terísticas sociais e demográficas dos estudantes. Ao analisar a associação entre cada variável e o

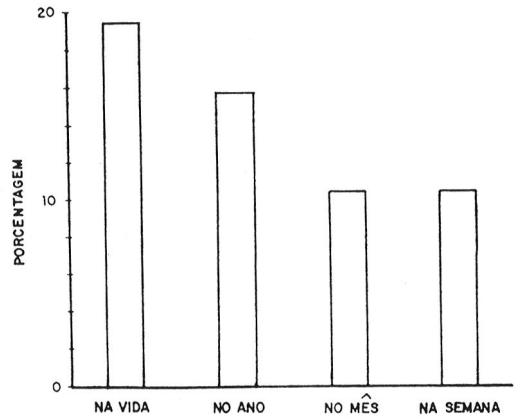


Figura — Percentagem do uso de tabaco em estudantes de dez capitais brasileiras, em 1987.

TABELA 1

Percentagens e chances dos fumantes em uma amostra de 16.149 estudantes, segundo características sociais.

		Total (N)	Fumantes (%)	Razão de chances
Turno	diurno	11.768	6,5	0.07
	noturno	3.817	21,7*	0.28
Sexo	masculino	6.883	10,4	0.12
	feminino	8.993	8,8*	0.10
Idade (anos)	10-12	3.260	2,2	0.02
	13-15	7.273	7,9	0.09
	16-18	3.991	15,0	0.18
	> 18	1.350	24,6*	0.33
Trabalho	não	11.821	7,9	0.09
	sim	3.985	16,8*	0.20
Defasagem (nos estudos)	não tem	1.619	2,5	0.03
	1-2 anos	5.926	5,5	0.06
	> 2 anos	8.410	14,3*	0.17
Pai fuma?	não	7.840	8,2	0.09
	sim	6.737	11,0	0.12
	não tem	1.397	14,0*	0.16
Mãe fuma?	não	11.269	8,9	0.10
	sim	4.332	12,0	0.14
	não tem	386	16,6*	0.20
Total		16.149	9,8	0.11

Nota — As somas dos Ns para cada variável não coincidem, pois estão excluídos os brancos e nulos.

\*  $p \leq 0,01$  em relação aos não fumantes (teste do qui-quadrado)

hábito de fumar vê-se que a idade acima de 18 anos, o sexo masculino, o trabalho, a freqüência ao curso noturno, a defasagem nos estudos e o fato de os pais fumarem estão positivamente associados ao fato de o estudante ser fumante ( $p < 0,01$ , pelo teste do qui-quadrado). Ainda na Tabela 1 é apresentada outra possibilidade de leitura de resultados bivariados, a razão de chances. A diferença entre a leitura das percentagens e destas razões pode ser explicada da seguinte maneira: as percentagens relacionam o número de fumantes com o total (10% significa que 10 pessoas em 100 fumam), e as razões de chances relacionam o número de fumantes com o de não fumantes (0,10 significa que existem 10 fumantes para 100 não fumantes). Assim, 10% de fumantes equivale a uma chance de fumar de 10/90 ou uma razão de 0,11.

Como já foi dito anteriormente, estes resultados, embora forneçam dados importantes, são limitados no sentido de detectar-se a complementação ou interação das associações feitas. Vejamos então o que um tratamento multivariado dos dados e a aplicação do Modelo Logit podem revelar.

### Análise Multivariada

Para a realização do estudo multivariado, procurou-se adaptar as variáveis em estudo às características do modelo estatístico a ser utilizado (Logit). Este modelo recomenda, visando não prejudicar as estimativas que realiza, que se evite o aparecimento de celas vazias ou com freqüências muito baixas nas tabelas a serem montadas. Por isso, optou-se, nesta fase do trabalho, por eliminar ou agrupar variáveis ou mesmo categorias de variáveis.

Assim sendo, eliminou-se a variável sexo desta análise o que se justifica pela pequena diferença de percentagem entre os sexos e agruparam-se as categorias "pai fuma" e "mãe fuma" na variável "pai e/ou mãe fumam" e as categorias "sem defasagem escolar" e "um ano de defasagem" em "um ano ou menos de defasagem" (Tabela 2). Do mesmo modo, excluiu-se da análise os estudantes de 10 e 11 anos, uma vez que o número de fumantes nessas idades foi praticamente zero.

Finalmente, optou-se por tratar em separado os dados referentes aos alunos do curso diurno e noturno, não somente por verificar-se que o turno escolar era a variável que mais discriminava o grupo dos fumantes, mas, também, pelo fato de as clientela dos dois turnos serem bastante distin-

TABELA 2

Agrupamento de categorias para o tratamento multivariado dos dados

Classificação	Nome	Código	Categorias
Variável dependente	Fumante	(F)	1- Não fumou no mês 2- Fumou no mês
Variáveis Independentes	Pai e/ou mãe fuma	(P)	1- Nenhum dos dois fumam 2- Pelo menos um fuma
	Defasagem	(D)	1- Apresenta um ano ou menos de defasagem 2- Apresenta dois anos ou mais de defasagem
	Idade	(I)	1- 12-15 anos 2- > 15 anos
	Trabalho	(T)	1- Não trabalha 2- Trabalha

tas em termos de sua composição etária e sócio-econômica.

Cabe esclarecer, ainda, que o modelo Logit estima as razões de chances e não as percentagens.

### Estudantes do Diurno

Os 10.015 estudantes do diurno que têm mais do que 11 anos estão distribuídos de acordo com as características apresentadas nas quatro variáveis, nas 32 celas da Tabela 3. A variável a ser explicada será a razão de chances de ser fumante, que foi calculada dividindo-se em cada linha o número de não-fumantes pelo de fumantes (ver penúltima coluna, Tabela 3). Pode-se, numa análise inicial, identificar o conjunto de categorias que fornecem as maiores e menores razões.

Observa-se que a maior razão de chances acontece na seguinte situação:

— O estudante, cujos pais são fumantes, está defasado nos estudos, tem mais do que 15 anos e trabalha (0,234).

Ao alterar-se apenas a categoria da variável PAIS, percebe-se que esta razão cai para 0.076. Esta queda parece indicar grande importância desta variável.

Observe-se, ainda, o que aconteceria com as razões ao alterar-se cada uma das variáveis:

— O estudante, cujos pais não são fumantes, está defasado nos estudos, tem mais do que 15 anos e trabalha (0,076)

TABELA 3

Frequências e razões observadas e esperadas dos estudantes do diurno

Pais (P) Fumam	Defasagem (D) Escolar	Idade (I)	Trabalho (T)	Fumante (F)		Total	Razão de Chances	Razão Esperada
				Não	Sim			
Não	Não	12-15	Não	1.735	57	1.792	0.033	0.031
			Sim	229	11	240	0.048	0.045
		> 15	Não	188	11	199	0.058	0.049
			Sim	45	2	47	0.044	0.071
Não	Sim	12-15	Não	869	45	914	0.052	0.052
			Sim	204	17	221	0.083	0.077
		> 15	Não	570	44	614	0.077	0.084
			Sim	132	10	142	0.076	0.122
Sim	Não	12-15	Não	2.222	111	2.333	0.050	0.055
			Sim	342	24	366	0.070	0.080
		> 15	Não	187	21	208	0.112	0.088
			Sim	36	6	42	0.167	0.129
Não	Não	12-15	Não	1.303	127	1.430	0.097	0.093
			Sim	313	43	356	0.137	0.136
		> 15	Não	740	107	847	0.145	0.149
			Sim	214	50	264	0.234	0.217
Total				9.329	686	10.015	0.073	

— O estudante, cujos pais são fumantes, não está defasado, tem mais do que 15 anos e trabalha (0.167).

— O estudante cujos pais são fumantes, está defasado nos estudos, tem menos do que 15 anos e trabalha (0.137).

— O estudante cujos pais são fumantes, está defasado nos estudos, tem mais do que 15 anos e não trabalha (0.145).

Percebe-se que o fato de os pais serem fumantes exerce um grande efeito no aumento da razão de chances.

Identifica-se, também, apenas pela observação desta Tabela, que as menores razões de chances ocorrem quando:

— Os pais não são fumantes, o estudante não tem defasagem, tem idade entre 12-15 anos e não trabalha (0.033). Para este grupo percebe-se pouca alteração nas razões com qualquer mudança de variável.

Esta análise exploratória ajuda a perceber os efeitos do comportamento conjunto das variáveis independentes na razão de chances da variável resposta. O ajuste do modelo Logit permitirá recompor estas razões, identificando em cada situação quais as variáveis que mais interferiram. Para o estudante cujos pais são fumantes, está defasado nos estudos, têm mais do que 15 anos e trabalha,

o modelo estimará o quanto cada variável contribuiu na formação da razão observada (0.234).

### Estudantes do Noturno

Os 3.419 estudantes do noturno estão distribuídos de acordo com as características apresentadas nas quatro variáveis (Tabela 4).

Percebe-se que as razões de chances são mais elevadas do que para os estudantes do diurno em todas as combinações de categorias das variáveis. Só este fato já caracteriza o turno noturno como um fator que discrimina quanto ao hábito de fumar. As maiores razões aconteceram também para os estudantes com pais fumantes, que estão defasados nos estudos, possuem mais do que 15 anos e trabalham (0.416). Mas já podemos observar que ao alterarmos o fator trabalho esta razão praticamente não se modifica (0.37). Isto indica que o fator trabalho tem uma influência menor para os estudantes do noturno. O fato de os pais serem fumantes, ao contrário do encontrado nos estudantes do diurno, não parece ter também influência, pois ao alterarmos esta variável, a razão cai apenas para 0.332.

TABELA 4

Frequências observadas e razões observadas e estimadas dos estudantes do noturno

País (P) Fumam	Defasagem (D) Escolar	Idade (I)	Trabalho (T)	Fumante (F)		Total	Razão de Chances	Razão Esperada
				Não	Sim			
Não	Não	12-15	Não	105	11	116	0.105	0.098
			Sim	62	7	69	0.113	0.122
		> 15	Não	37	3	40	0.081	0.123
			Sim	43	8	51	0.186	0.150
Não	Sim	12-15	Não	104	11	115	0.106	0.150
			Sim	76	9	85	0.118	0.184
		> 15	Não	314	77	391	0.245	0.250
			Sim	361	120	481	0.332	0.308
Sim	Não	12-15	Não	157	20	177	0.127	0.142
			Sim	84	15	99	0.179	0.174
		> 15	Não	50	6	56	0.120	0.175
			Sim	77	18	95	0.234	0.217
Não	Não	12-15	Não	131	33	164	0.252	0.214
			Sim	112	33	145	0.295	0.264
		> 15	Não	408	151	559	0.370	0.359
			Sim	548	228	776	0.416	0.442
<b>Total</b>				<b>2.669</b>	<b>750</b>	<b>3.419</b>	<b>0.281</b>	

Ainda pode-se perceber que a maior alteração se dará ao passarmos para o grupo dos não defasados nos estudos (0.234).

O ajuste do modelo Logit permitirá confirmar, negar e quantificar as observações acima. Antes de se iniciar a procura do modelo, será feita uma apresentação teórica do modelo Log Linear e sua comparação com o modelo Logit.

### Modelo Logit

Nesta etapa do trabalho deseja-se encontrar um modelo que descreva como as chances de ser fumante estão relacionadas com as variáveis escolhidas, ou seja, construir um modelo para prever a probabilidade de o estudante ser fumante a partir do conhecimento das categorias em que este se situa nas diversas variáveis em estudo. Este é um problema análogo ao de regressão múltipla, mas como a variável resposta bem como as variáveis independentes não são contínuas e sim categóricas, será utilizado o Modelo Logit<sup>2,7,8,9</sup>. O Modelo Logit ajustará o logaritmo das razões entre o número observado de fumantes e o número observa-

do de não fumantes em relação aos efeitos de cada variável e de cada interação significantes. Partindo do modelo com todas as variáveis e todas as interações possíveis (modelo saturado), a escolha será feita a partir da análise de quais efeitos seriam decorrentes apenas de variações casuais. Considerando-se que o valor padronizado de cada parâmetro pode ser considerado como tendo distribuição normal, escolhe-se os efeitos que apresentam em alguma categoria um valor maior do que 1.96. Isto porque numa distribuição  $N(0,1)$ , só 5% das observações têm valores maiores do que + 1.96 ou menores do que - 1.96, o que equivale a rejeitar erradamente que os efeitos são nulos com uma probabilidade de no máximo 0.05.

### Modelo Log-Linear

Seja uma tabela de contingência  $I \times J \times K$  com as variáveis denotadas respectivamente por A, B, C e  $P_{ijk}$  a probabilidade de uma observação cair na cela  $(i, j, k)$ , onde  $i = 1, \dots, I, j = 1, \dots, J$  e  $K = 1, 2$  são as respectivas categorias das variáveis A, B e C

B	A	C			
		1	2		
1	1	n <sub>111</sub>	n <sub>112</sub>	n <sub>11+</sub>	
	2	n <sub>211</sub>	n <sub>212</sub>	n <sub>21+</sub>	
	.				
	.				
	I	n <sub>I11</sub>	n <sub>I12</sub>	n <sub>I1+</sub>	
	1	n <sub>121</sub>	n <sub>122</sub>	n <sub>12+</sub>	
	2	2	n <sub>221</sub>	n <sub>222</sub>	n <sub>22+</sub>
		.			
		.			
		.			
I		n <sub>I21</sub>	n <sub>I22</sub>	n <sub>I2+</sub>	
.	.	.	.		
.	.	.	.		
.	.	.	.		
.	.	.	.		
J	1	n <sub>1j1</sub>	n <sub>1j2</sub>	n <sub>1j+</sub>	
	2	n <sub>2j1</sub>	n <sub>2j2</sub>	n <sub>2j+</sub>	
	.				
.	.	.	.		
.	.	.	.		
I	n <sub>Ij1</sub>	n <sub>Ij2</sub>	n <sub>Ij+</sub>		

N<sub>ijk</sub> é a freqüência observada na cela (i, j, k)  
 N<sub>ij+</sub> é a freqüência total observada na linha (i, j)  
 F<sub>ijk</sub> a freqüência esperada na cela (i, j, k) obtida a partir de um modelo especificado.

Supondo-se que as N observações desta Tabela foram retiradas de uma distribuição Multinomial com vetor de parâmetros P<sub>ijk</sub>.

Então tem-se que F<sub>ijk</sub> = E (f<sub>ijk</sub>) = Np<sub>ijk</sub> onde f<sub>ijk</sub> ~ M(N, P<sub>ijk</sub>).

Analogamente à Análise de Variância, o modelo log linear<sup>(2)</sup> decompõe o valor esperado das freqüências nas celas como segue:

$$\ln F_{ijk} = \alpha + \alpha_i^A + \alpha_j^B + \alpha_k^C + \alpha_{ij}^{AB} + \alpha_{ik}^{AC} + \alpha_{jk}^{BC} + \alpha_{ijk}^{ABC}$$

$$\text{onde } \sum_i \alpha_i^A = \sum_j \alpha_j^B = \sum_k \alpha_k^C = 0$$

$$\sum_{ij} \alpha_{ij}^{AB} = \sum_{jk} \alpha_{jk}^{BC} = \sum_{ik} \alpha_{ik}^{AC} = 0$$

$$\sum_{ijk} \alpha_{ijk}^{ABC} = 0$$

O α é o efeito geral e os α<sub>is</sub> representam os efeitos das 3 variáveis no logaritmo do valor esperado das freqüências nas celas.

O BMDP: 4F pacote estatístico utilizado neste trabalho, ajusta modelos log lineares e calcula os estimadores de Máxima Verosimilhança dos valores esperados através de um método iterativo adaptado por Haberman<sup>10</sup> baseado no ajuste proporcional das marginais que se deseja ajustar. Todo modelo logit é equivalente a algum modelo log-linear, desde que as marginais conjuntas das variáveis independentes sejam consideradas fixas<sup>7,8,9</sup>.

No caso da Tabela e do modelo acima descritos o modelo logit equivalente será:

$$(\ln F_{ij1} - \ln F_{ij2}) = (\alpha_1^C - \alpha_2^C) + (\alpha_{i1}^{AC} - \alpha_{i2}^{AC}) + (\alpha_{j1}^{BC} - \alpha_{j2}^{BC}) + (\alpha_{ij1}^{ABC} - \alpha_{ij2}^{ABC}) = \lambda^C + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

onde λ<sup>C</sup> é um efeito geral, λ<sub>i</sub><sup>A</sup> e λ<sub>j</sub><sup>B</sup> são os efeitos das variáveis A e B na variável C e λ<sub>ij</sub><sup>AB</sup> é o efeito conjunto das variáveis A e B na variável C.

O modelo log-linear ao ajustar as freqüências das celas de uma tabela de contingência, considera que as variáveis são simétricas, enquanto o modelo Logit, explica a razão das chances em favor de uma das categorias da variável considerada dependente.

**Modelo Logit - Diurno**

O modelo escolhido para ajustar as razões de chances observadas na Tabela 3 considerou como significativo todos os efeitos individuais das variáveis independentes (exemplo: idade, trabalho, defasagem) na variável resposta, mas nenhum efeito de interação (exemplo: idade e defasagem, país fumantes e trabalho, etc.). As estimativas dos parâmetros do Modelo Logit foram calculadas a partir da equivalência existente entre este e o modelo log-linear. Na Tabela 5 encontra-se a descrição destas estimativas.

## Estimativas dos Parâmetros

TABELA 5

Categoria	Período Diurno	
	1	2
Efeito		
Geral (G)		- 2,502
Trabalho (T)	- 0,188	0,188
Idade (I)	- 0,234	0,234
Pais fumantes (P)	- 0,288	0,288
Defasagem (D)	- 0,266	0,266

A partir dos valores dos parâmetros encontrados pode-se construir a tabela ajustada e identificar-se quais efeitos estão tendo uma maior influência nas razões de chances (última coluna da Tabela 3). Para o caso da maior razão de chance observada (0.234), tem-se que o valor estimado foi de (0.217) e

$$\ln(0.217) = - 2.502 + 0.288 + 0.266 + 0.234 + 0.188$$

(G) (P) (D) (I) (T)

Verifica-se que esta razão de chances está influenciada, principalmente, pelo fato de os pais fumarem (efeito de + 0.288) e, em segundo lugar, pelo fato de o aluno estar defasado nos estudos (efeito de + 0.266).

Importante observar que, nos casos em que existem poucas observações na cela da tabela, as estimativas são pouco precisas, e as interpretações devem ser mais cuidadosas. Percebe-se, por exemplo, que, se mantivermos todas as variáveis acima constantes e mudarmos para o grupo dos que tem pais não fumantes, a razão de chances observada cai para 0.076, enquanto a estimativa é de 0.129 (apenas 10 observações). Se mudamos para o grupo de 12 a 15 anos, o valor observado é de 0.137 e o estimado de 0.136. O efeito do trabalho para o estudante do diurno é o menor entre todos os outros.

## Modelo Logit - Noturno

Na escolha do modelo verificou-se a significância dos efeitos de todas as variáveis individuais e ainda a possibilidade de inclusão da interação da defasagem e idade ( $p < 0.08$ ). Nas Tabelas 6 e 7 encontra-se a descrição dos parâmetros do modelo:

## Estimativas dos Parâmetros

TABELA 6

Efeito	Período Noturno	
	1	2
Geral (G)		- 1,640
Trabalho (T)	- 0,104	0,104
Idade (I)	- 0,184	0,184
Pais fumantes (P)	- 0,180	0,180
Defasagem (D)	- 0,282	0,282

TABELA 7

D1xI1 e D1xI2

Defasagem (D)	Idade (I)	
	1	2
1	0,074	- 0,074
2	- 0,074	0,074

A Tabela 4 (última coluna) apresenta as razões ajustadas pelo modelo escolhido. As estimativas das maiores razões observadas serão assim descritas:

$$\ln(0.442) = - 1.64 + 0.104 + 0.184 + 0.180 + 0.282 + 0.074$$

(G) (T) (I) (P) (D) (DXI)

$$\ln(0.217) = - 1.64 + 0.104 + 0.184 + 0.180 - 0.282 - 0.074$$

(G) (T) (I) (P) (D) (DXI)

Tal resultado demonstra que as maiores razões de chances são explicadas principalmente pelo fato de o aluno estar defasado nos estudos. Verificou-se que o fato de o aluno mais velho não estar defasado diminui a sua razão de chances. Confirmou-se, também, que o menor efeito é o do trabalho.

## DISCUSSÃO

Os dados do presente trabalho apontaram para um consumo de tabaco de proporções significativas (cerca de 20%) entre estudantes, embora ligeiramente menor do que os índices encontrados em estudos anteriores<sup>3,11,12</sup>. Seria precipitado atribuir tal diferença a uma real diminuição do uso de cigarros na população estudantil. Possivelmente, tal diferença deve-se a não homogeneização dos procedimentos metodológicos adotados nos

vários estudos e ao fato de o presente trabalho apresentar dados que abrangem estudantes de dez capitais brasileiras.

Em relação ao consumo nos últimos trinta dias, objeto de nossa análise, nota-se que este hábito não se distribui aleatoriamente entre o estudantado, mas parece obedecer a uma lógica onde características demográficas e sociais conferem maiores ou menores chances ao aluno de vir a ser fumante.

Deste modo, é possível verificar, pelo tratamento multivariado dos dados, que os alunos que freqüentam o curso noturno, têm pais que fumam, estão defasados nos estudos, trabalham e são maiores de 15 anos apresentam uma chance muito maior de serem fumantes (0.416) do que seus outros colegas. Da mesma forma, os alunos que estudam no diurno, os pais não fumam, não estão defasados nos estudos, não trabalham e têm 15 anos ou menos apresentam a menor chance de virem a ser fumantes de toda a amostra estudada (0.033).

Entre estes dois extremos, é possível traçar uma série de situações de maior ou menor risco de ser fumante, combinando-se as variáveis em questão.

Assim, os presentes resultados indicam a necessidade de superarem-se visões unicasais e admitirem-se formas mais complexas de determinação do fenômeno do consumo de tabaco, ou mesmo de outros psicotrópicos, onde o embricamento, a interação, a complementação de fatores é que na verdade vão fornecer pistas reais para a discussão da questão.

Isto não quer dizer que alguns fatores não tenham maior peso que outros, em determinadas circunstâncias. Com a aplicação do modelo Logit foi possível detectar, por exemplo, que a defasagem escolar é o fator que mais contribui no caso dos alunos do curso noturno em situação de risco (trabalham, estão defasados nos estudos, os pais fumam e são maiores de 15 anos), na determinação do comportamento de ser fumante. Já entre alunos do curso diurno, com as mesmas características, a variável de maior peso é o fato de os pais fumarem.

Tal achado nos remete à discussão da influência do modelo familiar na configuração de hábitos como fumar, beber e abusar de medicamentos. Se aceitarmos como hipótese que os alu-

nos do curso noturno são, quando comparados com os do diurno, diferentes não só em termos de idade e inserção no mercado de trabalho (variáveis controladas nesta comparação) mas, também, do ponto de vista de sua maior autonomia em relação à família, a interpretação dos resultados sugere pistas bastante interessantes de análise.

É possível que, em setores de adolescentes de maior dependência dos pais, o papel destes seja realmente decisivo no comportamento de fumar, enquanto que para adolescentes já mais autônomos a influência dos pais, embora existente, não seja tão grande. Nestes casos, a trajetória escolar (ou seja, um fator relacionado com seu desempenho social mais geral) tem peso decisivo.

Esta relatividade do papel dos pais é discutida com bastante propriedade por alguns autores. Swadi<sup>13</sup> (1988), por exemplo, ao analisar o papel da família e dos amigos em uma amostra de 3.333 estudantes londrinos identificou tendência cinco vezes maior de usar drogas habitualmente em jovens que afirmam ter pais usuários de algum psicotrópico; mas detectou uma influência externa ainda maior, determinada pelo fato de que o estudante que convive com amigos que usam drogas têm uma possibilidade treze vezes maior de usar drogas habitualmente do que outros jovens. Do mesmo modo, Dishion<sup>6</sup> (1985) defende a hipótese de que o uso de drogas em adolescentes está associado a processos familiares, mas que a inserção mais geral em outros grupos sociais é tão ou mais importante que a família.

Estar atento a este papel parcial da família, assim como à multideterminação do comportamento estudado (uso de tabaco), é fato de grande importância na elaboração de políticas preventivas adequadas. A utilização de técnicas multivariadas permitiram, não só definir as situações de risco para os estudantes se tornarem fumantes, mas também detectar que os diversos fatores agem e interagem com intensidades diferentes para os estudantes do diurno e noturno. Estes resultados parecem indicar que os procedimentos utilizados merecem ser seguidos em estudos futuros.

#### AGRADECIMENTOS

Ao Prof. Dr. E. A. Carlini pela valiosa revisão do manuscrito.



BARBOSA, M.T.S. et al. [The use of tobacco by school children in ten Brazilian state capitals: possible contributions of multivariate statistical analysis to an understanding]. *Rev. Saúde públ.*, S.Paulo, 23:401-9, 1989.

**ABSTRACT:** The use of tobacco among school children in ten Brazilian cities, by means of statistical techniques such as bivariate and multivariate analysis (Logit Model), was analyzed. Bivariate analysis revealed a positive association between the use of tobacco and the school child's non-academic occupation, evening course attendance, poor school performance, and parents' smoking habits. Multivariate techniques and the Logit Model showed that poor school performance (on the evening course) and parents' smoking habits (on the day-time course) were most significant variables. The importance of the utilization of multivariate technics for the real understanding of the various factors involved in complex behavior patterns such as the use of tobacco, alcohol and drugs, was discussed.

**KEYWORDS:** Smoking. Multivariate analysis. Risk.

#### REFERÊNCIAS BIBLIOGRÁFICAS

1. ALMEIDA FILHO, N. & SANTANA, V.S. Consumo de drogas entre escolares de Salvador. [Relatório de pesquisa apresentado à DINSAM/Ministério da Saúde, 1988]
2. BISHOP, Y.M.M.; FIENBERG, S.E.; HOLLAND, P.W. *Discrete multivariate analysis: theory and practice*. Cambridge, MIT Press, 1975.
3. BUCHER, R. & TOTUGUI, M.L. Conhecimento e uso de drogas entre alunos de Brasília. *Psicol. Teor. e Pesq.*, Brasília, 3:178-94, 1987.
4. CARLINI-COTRIM, B.H. & CARLINI, E.A. O consumo de solventes e outras drogas em crianças e adolescentes de baixa renda na cidade de São Paulo. Parte I: Estudantes de primeiro e segundo graus da rede estadual. *Rev. ABP-APAL*, 9(2):49-58, 1987.
5. CARLINI, E.A.; CARLINI-COTRIM, B.H.; SILVA-FILHO, A.R.; BARBOSA, M.T.S. *Consumo de drogas psicotrópicas no Brasil, em 1987*. Brasília, Centro de Documentação do Ministério da Saúde/Ministério da Justiça, 1989.
6. DISHION, T.J. Adolescent marijuana and alcohol use: the role of parents and peers revisited. *Amer. J. Drug Alcohol Abuse*, 11:11-25, 1985.
7. GOODMAN, A.L. The multivariate analysis of qualitative data: interactions among multiple classifications. *J. Amer. statist. Ass.*, 65:227-56, 1970.
8. GOODMAN, A.L. The analysis of multidimensional contingency tables: a stepwise procedures and direct estimation methods for building models for multiple classifications. *Technometrics*, 13(1):33, 1971.
9. GOODMAN, A.L. A modified multiple regression approach to the analysis of dichotomous variables. *Amer. sociol. Rev.*, 37:28-46, 1972.
10. HABERMAN, S.J. *Analysis of qualitative data*. New York, Academic Press, 1978/1979.
11. SIMÕES, B.J.G. & SIMÕES, M.J.S. Levantamento sobre o uso de fumo entre os estudantes de nível secundário de Ribeirão Preto, SP, 1975. *Medicina*, Ribeirão Preto, 9/10:37-46, 1976, 1977.
12. STEWIEN, G.T.M. & MARCONDES, R.S. Uso de cigarros entre adolescentes de uma escola de primeiro grau da cidade de São Paulo, Brasil. *Rev. Saúde públ.*, S.Paulo, 13:235-43, 1979.
13. SWADI, H.S. Adolescent drug taking: role of family and peers. *Drug Alcohol Depend.*, 21:157-60, 1988.

Recebido para publicação em 5/4/1989

Reapresentado em 3/7/1989

Aprovado para publicação em 22/8/1989