

## Uso do modelo logito generalizado na análise de dados da criminalidade em Santarém, PA

Edilan S. Quaresma<sup>1</sup> †, Carlos T. S. Dias<sup>2</sup>, Renata Alcarde<sup>2</sup>, Ezequiel A. L. Bautista<sup>3</sup>, Jarsen Luis C. Guimarães<sup>1</sup>.

<sup>1</sup> Universidade Federal do Oeste do Pará, Instituto de Ciências da Educação (UFOPA-PA).

<sup>2</sup> Universidade de São Paulo, Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP).

<sup>3</sup> Universidad de San Carlos de Guatemala - USAC, Facultad de Agronomía, Guatemala.

**Resumo:** *A criminalidade é um fenômeno social amplamente disseminado nas mais diferentes civilizações modernas e objeto de estudo cada vez mais presente e defendido por pesquisadores de distintas áreas do conhecimento. Compreender a relação existente entre diferentes variáveis relacionadas ao crime pode ser uma ferramenta de grande importância no planejamento e execução de políticas públicas concretas, direcionadas a minimização do problema. Uma das ferramentas disponíveis para essa compreensão pode ser realizada por meio do uso da análise de dados categóricos e suas ramificações. O trabalho aqui apresentado objetiva utilizar o modelo logito generalizado para analisar a relação da renda e do uso de drogas no tipo de crime (contra a vida, contra o patrimônio, contra costumes e tráfico de entorpecentes) cometido por detentos da penitenciária agrícola Silvio Hall de Moura, em Santarém-Pa. Foram entrevistados 344 detentos nos anos de 2004 e 2005, estudando o agente ativo do crime e não a vítima. Considerando o tipo de crime como variável resposta, o modelo logito generalizado foi o melhor ajustado aos dados, a partir do qual probabilidades de ocorrências de cada cruzamento das categorias foram preditas e, finalmente, chances e razões de chances foram obtidas. O uso do modelo logito generalizado mostrou-se eficiente no processo de análise de dados categóricos identificando, dentre outros fatores, que detentos com menores rendas apresentam maiores chances de cometerem crimes (patrimônio, costumes e tráfico de entorpecentes) quando comparados com crimes contra a vida, sendo este o tipo de crime mais frequente no processo de condenação daqueles detentos.*

**Palavras-chave:** *Criminalidade; Dados Categóricos; Modelo Logito Generalizado.*

**Abstract:** *Crime is a social phenomenon spread in different modern civilizations and studied by researchers in different fields of knowledge. Understanding the relation between different variables related to crime can be a tool of great importance in planning and implementation of specific public policies. One of the tools available for this understanding can be accomplished by through the use of the analysis of categorical data and its branches. The work presented here objective using the generalized logit model to analyze the relationship between income and the use of drugs on the type of crime (against life, against property, against customs and trafficking narcotics) committed by inmates from agricultural Silvio Hall de Moura prison, in Santarém-Pará-Brazil. 344 detainee interviewed in 2004 and 2005, studying the agent asset of the crime and not the victim. Considering the type of crime as the dependent variable, the logit generalized model was the best fit to the data, from which probabilities of occurrences of each intersection of the categories were predicted, and finally odds and odds ratios were obtained. The use of the logit generalized model proved efficient in the analysis of categorical data process identifying, among other factors, which inmates with lower incomes are more likely to commit crimes (equity, customs and narcotics trafficking) when compared with crimes against life, and this is the most frequent type of crime in the sentencing process those detainees.*

**Keywords:** *Crime; Categorical data; generalized model logito.*

---

† Autor correspondente: [edilan@usp.br](mailto:edilan@usp.br).

## Introdução

A criminalidade é um fenômeno social amplamente disseminado nas mais diferentes civilizações modernas e objeto de estudo cada vez mais presente e defendido por pesquisadores de distintas áreas do conhecimento. Compreender a relação existente entre diferentes variáveis relacionadas ao crime pode ser uma ferramenta de grande importância no planejamento e execução de políticas públicas concretas, direcionadas a minimização do problema.

No Brasil, vive-se um momento de grande preocupação com este fenômeno, justificado dentre outros, pelos grandes eventos internacionais, como por exemplo, a copa do mundo, programados para serem realizados em diferentes estados da nação. O interior da Amazônia brasileira, como parte integrante da nação, também apresenta problemas relacionados à criminalidade. Dados da Polícia Civil do Pará apontam crescimento de 134,52% no índice de criminalidade no município de Santarém entre os anos 1999 e 2004. Embora as estatísticas oficiais do estado do Pará apontem redução na maioria dos tipos de crimes em 2011, quando comparado com 2010, merece destaque o aumento de 44,55% na incidência do tráfico de entorpecentes, defendido por muitos pesquisadores como a causa de outros tipos de delito.

O trabalho aqui apresentado objetiva estudar a influência das variáveis renda e o fato do delinquente usar ou não drogas no tipo de crime cometido por detentos da penitenciária agrícola Silvio Hall de Moura, localizada no município de Santarém, Oeste do Pará, por meio da utilização do modelo logito generalizado.

## Material e métodos

Para a realização deste trabalho, foram coletadas informações, por meio de aplicação de questionários a 344 detentos da Penitenciária Agrícola Silvio Hall de Moura localizada em Santarém-PA, nos meses de novembro de 2004 a abril de 2005, estudando-se o agente ativo do crime e não a vítima (GUIMARÃES, 2009). Os crimes foram divididos em quatro categorias: crimes contra a vida (homicídio e tentativa de homicídio), contra o patrimônio (roubo, furto e extorsão), contra os costumes (estupro e atentado violento ao pudor) e tráfico de entorpecentes. Além do tipo de crime como variável resposta, foram também analisadas as variáveis “Uso de drogas” e “Renda” do detento, no momento em que o crime foi cometido.

Em estudos dessa natureza, parece razoável compreender que um elenco de variáveis econômicas e sociais, apresentam algum tipo de relação com o fato do sujeito cometer um delito. Uma das alternativas que viabilizam tal compreensão diz respeito ao uso da análise de dados categóricos, mais especificamente para o caso em que se pretende relacionar variáveis explicativas com uma variável resposta do tipo politômica, o uso do modelo logito generalizado.

O modelo logito generalizado sugerido por Stukel (1988) é proposto para se trabalhar com a situação em que o interesse é estudar a existência de associação entre um conjunto de variáveis explicativas  $X = (X_1, X_2, \dots, X_r)$  com uma variável resposta politômica  $Y$ , que pode ser nominal ou ordinal, situação em que o uso de modelos de regressão podem ser úteis.

Para o caso em que se tem uma única variável resposta e, por exemplo, duas variáveis explicativas, pode-se denotar por  $\pi_{hik}$  a probabilidade de uma categoria da variável explicativa  $h$  ( $h = 1, 2, \dots, s$ ) e variável explicativa  $i$  ( $i = 1, 2, \dots, I$ ) preferir uma das  $k$  categorias da variável resposta  $k = 1, 2, \dots, r$ , em

$$\text{que } \sum_{k=1}^r \pi_{hik} = 1$$

Segundo Giolo (2013), este modelo considera  $(r - 1)$  logitos generalizados, tal que:

$$\text{logit}_{hik} = \ln \left( \frac{\pi_{hik}}{\pi_{hir}} \right) = \beta_0 + \beta'_k x_{hi}, k = 1, \dots, r - 1 \quad (1)$$

Cada logito  $k$  ( $k = 1, 2, \dots, r - 1$ ) é definido a partir da probabilidade associada à  $k$ -ésima categoria de resposta sobre a da última categoria ( $k = r$ ) e há diferentes interceptos e diferentes vetores de parâmetros  $\beta_k$  associados aos  $(r - 1)$  logitos.

A ideia inicial é propor modelos que possam se ajustar aos dados de tal forma que seja viável identificar a influência das variáveis Droga e Renda no tipo de crime cometido pelo detento. Um dos critérios utilizados para avaliar a qualidade do modelo ajustado consiste em observar a diferença entre os valores das probabilidades observadas e ajustadas pelo modelo.

## Resultados e discussões

Os dados mostrados na Tabela 1 foram fornecidos por detentos da penitenciária agrícola Silvio Hall de Moura, em Santarém-Pa e contemplam as variáveis Droga (usuário ou não usuário), Renda (até dois salários mínimos e mais de dois salários mínimos) e o tipo de crime cometido pelo detento (contra a vida, contra o patrimônio, de costumes e tráfico de entorpecentes).

Tabela 1: Tipos de crimes por renda e uso de drogas cometidos pelos detentos.

Droga	Renda	Vida	Patrimônio	Costumes	Tráfico de entorpecentes
Sim	≤ 2 SM	89	54	14	9
Sim	> 2 SM	10	1	1	3
Não	≤ 2 SM	33	52	21	30
Não	> 2 SM	13	7	2	5

Quatro modelos foram ajustados, conforme mostrado na Tabela 2.

Tabela 2: Diferença de *Deviances* entre os modelos propostos.

Modelo	gl	deviance	≠ de deviance	≠ de gl	valor p
Nulo	9	46,55983	-	-	-
D	6	14,08215	32,47768	3	$4 \times 10^{-7}$ ii
R D	3	5,316181	8,76597	3	0,03257
D*R D,R	0	0	5,31618	3	0,15006

As informações referentes aos quatro modelos propostos, apresentadas na Tabela 2 mostraram que o efeito da interação *Droga \* Renda* é nulo. Ao incluir a variável Droga no modelo, a *deviance* reduziu de 46,55983 para 32,47768 e ainda, ao incluir a variável “Renda” no modelo, considerando que a variável “Droga” estava presente, a *deviance* foi reduzida a 8,76597, com valor *p* igual a aproximadamente 0,0326, indicando que estas variáveis têm suas contribuições significativas no tipo de crime cometido pelo detento

Tabela 3: Diferença entre valores das probabilidades observados e preditas a partir do modelo

Vida	Patrimônio	Costumes	Tráfico
-0,08089	-0,07698	0,025564	0,132306
0,007309	0,006956	-0,00231	-0,01196
0,044939	0,042766	-0,0142	-0,0735
-0,00892	-0,00849	0,00282	0,014593

Para o modelo aqui proposto, que considera como variável resposta o tipo de crime e como variáveis explicativas o uso ou não de drogas e a classe de renda do detento, as diferenças entre valores observados e preditos apresentadas na Tabela 3 são pequenas, indicando evidências favoráveis ao modelo ajustado. Outro fato que justifica a aceitação do modelo proposto é apresentado na Figura 1, onde se observa a aparente ausência de valores atípicos.

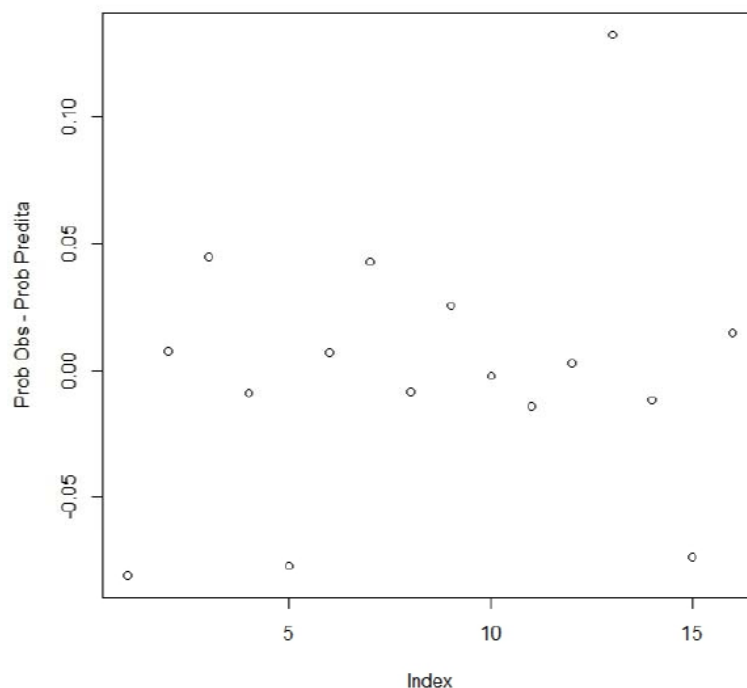


Figura 1: Resíduos da diferença entre probabilidades observada e predita pelo modelo

As estimativas dos parâmetros dos 3 modelos logito generalizados, comparando os tipos de crime com o crime contra a vida, são apresentadas na Tabela 4 que podem ser utilizadas para prever probabilidades (ver Tabela 5) associadas às categorias das variáveis aqui estudadas e, dessa forma, viabilizar a obtenção das chances e razões de chances de interesse.

Tabela 4: Estimativas dos parâmetros dos 3 modelos logitos (referência: crimes contra a vida).

Parâmetros	Logito 1	Logito 2	Logito 3
	Patrimônio:Vida	Costumes:Vida	Tráfico:Vida
Intercepto	-0,70133	-1,59842	-0,52326
Droga (S:N)	-0,94813	-1,30231	-1,87855
Renda( $\leq$ 2SM: $>$ 2SM)	1,14192	1,09189	0,32340

Na Tabela 5 observa-se, dentre outras coisas, que a probabilidade de crimes contra a vida são cometidos, em sua maioria, por usuários de drogas e, dentre estes, por aqueles sujeitos com as menores rendas. Este quadro se inverte quando se trata de crimes contra o patrimônio. Outra informação interessante diz respeito ao fato das probabilidades de crimes de tráfico de entorpecentes serem maiores naqueles classificados como não usuários, com aparente indiferença entre as classes de renda.

Tabela 5: Probabilidades estimadas pelo modelo “Crime = Droga + Renda”

Vida	Patrimônio	Costumes	Tráfico de entorpecentes
0,747557	0,143646	0,041103	0,067694
0,528835	0,318345	0,086647	0,066172
0,436542	0,216493	0,088276	0,258689
0,251569	0,390843	0,151592	0,205996

Na Tabela 6 é mostrado a chance de um sujeito cometer crimes contra o patrimônio, costumes ou tráfico de entorpecentes, em relação ao crime contra a vida. Nesta tabela observa-se, por exemplo, que a chance de um detento ter cometido crime contra o patrimônio, em relação ao crime contra a vida, é

maior naqueles sujeitos não usuários de drogas e que, além disso, apresentam maiores rendas (maior que 2 salários mínimos), fato que se repete para os outros tipos de crime (costumes e tráfico) quando comparados com o crime contra a vida.

Tabela 6: Chances dos diferentes tipos de crime em relação ao crime contra a vida

Logito 1	Logito 2	Logito 43
Patrimônio:Vida	Costumes:Vida	Tráfico:Vida
0,192154	0,054983	0,090553
0,601974	0,163846	0,125128
0,495927	0,202216	0,592585
1,553624	0,602588	0,818844

Finalmente, a Tabela 7 mostra as razões de chance de crimes contra o patrimônio, contra costumes e tráfico de entorpecentes, comparados em relação ao crime contra a vida. Observa-se que a maior relação do tipo de crime quando comparado com o crime contra a vida, está relacionado ao tráfico de entorpecentes

Tabela 7: Razão de chances considerando as diferentes rendas

Logito 2	Logito 3	Logito 4
Patrimônio:Vida	Costumes:Vida	Tráfico:Vida
0,319206	0,33558	0,723685

Na Tabela 7 observa-se que a chance de detentos terem cometido crimes contra o patrimônio em relação a crimes contra a vida é 3 vezes menor para sujeitos com menores rendas (até 2 salários mínimos) do que para sujeitos com rendas maiores que 2 salários mínimos, fato que se repete ao comparar crimes contra costumes em relação a crimes contra a vida. Por outro lado, ao se fazer a mesma comparação entre crimes de tráfico de entorpecentes em relação a crimes contra a vida, a relação é de 0,7234.

## Conclusões

O modelo logito generalizado mostrou-se interessante no processo de análise de dados categóricos, provenientes de detentos da penitenciária agrícola de Santarém, identificando, dentre outros fatores, que detentos mais pobres apresentam menores chances de cometer crimes (patrimônio, costumes e tráfico de entorpecentes) quando comparados com crimes contra a vida, sendo este o tipo de crime mais frequente no processo de condenação daqueles detentos.

## Agradecimentos

À CAPES e CNPq pelo apoio financeiro.

## Referências

GIOLO, S.R. *Introdução à análise de dados categóricos*. UFPR,(Minicurso), 2013. 225p.

GUIMARÃES, J. L. C. Criminalidade econômica: análise de fatores econômicos e sociais que influenciam as categorias de crimes no município de Santarém-Pa. *Revista Jus Navigandi*, Porto Alegre, 2009.

STUKEL, T. A. Generalized logistic models. *Journal of the American Statistical Association*, 83, 1988. p.426-431.