

## Desigualdade social como vetor dos homicídios no Estado de São Paulo, Brasil

Social inequality as a vector of homicides in the States of São Paulo, Brazil

La desigualdade social como vector de homicídios en el Estado de São Paulo, Brasil

Recebido: 04/09/2021 | Revisado: 11/09/2021 | Aceito: 15/09/2021 | Publicado: 17/09/2021

**Calisto Manuel Máquina**

ORCID: <https://orcid.org/0000-0002-6749-3482>

Universidade Zambeze, Moçambique

E-mail: [calisto.calmaqna.maquina@gmail.com](mailto:calisto.calmaqna.maquina@gmail.com)

**Elisa da Conceição José Maria**

ORCID: <https://orcid.org/0000-0003-3134-8061>

Universidade Rovuma, Moçambique

E-mail: [elisajmaria@gmail.com](mailto:elisajmaria@gmail.com)

**Eufrásio João Sozinho Nhongo**

ORCID: <https://orcid.org/0000-0002-5453-7845>

Universidade Zambeze, Moçambique

E-mail: [eufrasionhongo@yahoo.com](mailto:eufrasionhongo@yahoo.com)

**Sérgio Feliciano Come**

ORCID: <https://orcid.org/0000-0002-0866-1423>

Universidade Zambeze, Moçambique

E-mail: [sergiofcome@gmail.com](mailto:sergiofcome@gmail.com)

### Resumo

A violência constitui uma das tristes realidades vividas nas zonas urbanas do Estado de São Paulo. Uma das questões que se tenta responder quando ocorre um ato de violência, por exemplo homicídio, está relacionada com as causas que influenciaram esses atos. Deste modo, o presente trabalho procurou definir um modelo estatístico que permite uma descrição precisa dos casos de homicídios no Estado de São Paulo, tendo em conta alguns indicadores sócio demográficos resultantes do censo de 2000. Foram ajustados três cenários do modelo de distribuição Poisson e quatro cenários de distribuição binomial negativa. Por meio do Critério de Informação de Akaike, foi identificado o modelo de respostas binomiais negativas, como sendo o melhor modelo para modelagem dos casos de homicídios no Estado, tendo-se constatado que a desigualdade sócio econômica contribui significativamente na ocorrência dos casos de homicídios no Estado de São Paulo.

**Palavras-chave:** Violência; Homicídios; Modelo de distribuição binomial negativa; Modelo de Poisson.

### Abstract

Violence is one of the sad realities experienced in urban areas of the State of São Paulo. One of the questions that is tried to answer when there is an act of violence, for example homicide, is related to the causes that influenced these acts. Thus, the present work sought to define a statistical model that allows a precise description of homicide cases in the State of São Paulo, taking into account some socio-demographic indicators resulting from the 2000 census. Three scenarios of the Poisson distribution model and four scenarios of the negative binomial distribution, were adjusted. Through the Akaike Information Criterion, the model of negative binomial responses was identified as the best model for modeling homicide cases in the state, and it was found that socio-economic inequality significantly contributes to the occurrence of homicide cases in the State of São Paulo.

**Keywords:** Violence; Homicides; Negative binomial distribution model; Poisson model.

### Resumen

La violencia es una de las tristes realidades vividas en las zonas urbanas del Estado de São Paulo. Una de las cuestiones que se intenta responder cuando ocurre un acto de violencia, por ejemplo, un homicidio, está relacionada con las causas que influyeron en estos hechos. Así, el presente trabajo buscó definir un modelo estadístico que permita una descripción precisa de los casos de homicidio en el Estado de São Paulo, teniendo en cuenta algunos indicadores sociodemográficos resultantes del censo de 2000. Se han ajustado tres escenarios del modelo de distribución de Poisson y cuatro escenarios de distribución binomial negativa. Mediante el Criterio de Información de Akaike, se identificó que el modelo de respuestas binomiales negativas es el mejor para modelar los casos de homicidio en el Estado, teniéndose constatado que la desigualdad socioeconómica contribuye significativamente en la ocurrencia de casos de homicidio del estado de Sao Paulo.

**Palabras clave:** Violencia; Homicidios; Modelo de distribución binomial negativa; Modelo de Poisson.

## 1. Introdução

O conceito popular de “poucos com muito e muitos com pouco” ainda é considerado como a principal causa da desigualdade social em diversos países do mundo, incluindo o Brasil. A desigualdade social no Brasil, apesar dos avanços alcançados na primeira década dos anos 2000, ainda é considerada uma das mais altas do mundo (Campello, 2018; Barros, Henriques & Mendonça, 2021).

A desigualdade social prejudica cidadãos de todas as faixas etárias, e com maior incidência nos jovens de classe de baixa renda, impossibilitados de ascender socialmente pela falta de uma educação de qualidade, de melhores oportunidades no mercado de trabalho e de uma vida digna, o que propicia a ocorrência da violência nas zonas urbanas brasileiras.

O aumento da violência urbana, constitui um dos fatores que explica o aumento progressivo das taxas de homicídios no Brasil, um cenário considerado grave particularmente nos grandes centros urbanos, como no Estado de São Paulo. De acordo com Gawryszewski e Costa (2005), o Estado de São Paulo conheceu um aumento na ordem de 335% das taxas de homicídios entre 1980 e 2000, de acordo os resultados da análise dos dados do Sistema de Informações de Mortalidade do Ministério da Saúde.

Sob ponto de vista estatístico, *o número de casos de homicídios* faz parte do tipo de dados designado por *dados contagem*, isto é, dados que assumem valores inteiros e não negativos, e para a sua melhor análise requer o uso de modelos estatísticos apropriados. Os dados de contagem são encontrados diariamente, mas para melhor compreensão e extração de informações importantes sobre estes, se requer alguma análise estatística. Pelo exposto, o presente estudo teve como objetivo analisar o impacto da desigualdade social nos homicídios no estado de São Paulo, para tal foi identificado um modelo adequado e parcimonioso, que permita descrever a relação entre a variável aleatória não-negativa (número de homicídios do ano 2000) e um conjunto de variáveis não-aleatórias que descrevem os indicadores sócio demográficos da população de São Paulo, Brasil.

## 2. Metodologia

### 2.1 Caracterização do estudo

Realizou-se um estudo exploratório, de modos a proporcionar maior familiaridade da problemática de ocorrências de homicídios nas zonas urbanas do Estado de São Paulo. A análise realizada sobre os casos de homicídios, foi descritivo, pois consistiu na identificação e desmistificação, dos fatores que contribuem no aumento dos casos de homicídios nos municípios do Estado de São Paulo-Brasil. O estudo foi realizado com base nas fundamentações metodológicas dos autores Koche (2011), Ludke e Andre (2013), Yin (2015), Estrela (2018) e Pereira et al. (2018).

### 2.2 Análise de dados

- Foram ajustados modelos de regressão Poisson e Binomial-Negativa, na tentativa de encontrar um modelo que melhor se adequasse aos dados. Os dados foram obtidos no Instituto Brasileiro de Geografia e Estatística (IBGE). As variáveis analisadas foram as seguintes: Variável dependente (número de homicídios registrados em 621 municípios do Estado de São Paulo no ano de 2000).
- Variáveis independentes: Taxa de mortalidade infantil; Percentagem de jovens de 15 a 17 anos que não frequentam a escola; Percentual de pessoas analfabetas de 25 anos ou mais; Renda *per capita*; Percentual de pessoas que vivem em domicílios com banheiro e água encanada e Percentual de pessoas que vivem em domicílios com energia elétrica e geladeira.

O processamento dos dados foi feito com auxílio do *software* estatístico R.

### 2.3 Modelo Poisson

Segundo Ramalho (1996), Cameron e Trivedi (1998), Winkelmann (2003) e Hilbe (2011a) o modelo de Poisson, ou função de Probabilidade Poisson, é um dos modelos recomendados para a modelagem de um conjunto de observações de uma variável de contagem, a exemplo do número de homicídios que constituem objeto de análise nesta pesquisa. Para além de apresentar uma estrutura simples, o modelo de Poisson permite a realização de inferência sobre o fenômeno em análise.

Seja  $Y_i, i = 1, 2, \dots, n = 620$  uma variável discreta que denota o número de homicídios independentes, registrados num determinado município  $i$ . A função de probabilidade de Poisson de registo de homicídio no município  $i$ , é dada por (Ramalho, 1996, Cameron & Trivedi, 1998)

$$P(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \quad (1)$$

em que  $y_i = 0, 1, 2, \dots$  e  $\lambda_i > 0$ .

O modelo de regressão Poisson, resulta da colocação do parâmetro  $\lambda_i$  em função de um conjunto de variáveis explicativas:  $\lambda_i = f(x_i; \beta)$  em que  $x_i$  é um vector  $1 \times m$ , que contém as observações das  $m$  variáveis exógenas para cada município  $i$  e  $\beta$  um vector  $m \times 1$ , composto por parâmetros desconhecidos. Como  $\lambda_i$  representa o valor esperado condicional de  $Y_i$ , leva em consideração somente os valores não negativos. Sendo assim, a função  $f$  deve conter um contradomínio apenas os números reais não negativos. Por este motivo, a função  $f$  é geralmente expressa na forma exponencial, isto é,

$$\lambda_i = \exp(x_i \beta) \quad (2)$$

Uma vez que a variância depende dos regressores, este modelo é por natureza heterocedástico, que constitui característica comum dos modelos de contagem (Winkelmann, 2003; Hilbe, 2011a). Para avaliar a qualidade de ajuste do modelo Poisson foi utilizada a estatística *deviance* do modelo, por tanto espera-se que o quociente entre a *deviance* e os graus de liberdade seja aproximadamente igual a 1.

### 2.4 Modelo de distribuição binomial negativa

Se a variável  $Y_i, i = 1, 2, \dots, n$  segue uma distribuição binomial negativa (BN), com os parâmetros  $w$  e  $p$ , isto é,  $Y_i \sim BN(w, p)$ , a sua função de probabilidade é determinada da forma (Ross & Preece, 1985)

$$P(Y_i = y_i) = \frac{\Gamma(y_i + p)}{\Gamma(y_i + 1)\Gamma(p)} p^w (1 - p)^{y_i} \quad (3)$$

em que  $0 < p < 1$  e  $w > 0$ .

$$E[Y] = \frac{w(1 - p)}{p} \quad (4)$$

e

$$V[Y] = \frac{E[Y]}{p} \quad (5)$$

em que  $E[Y]$  e  $V[Y]$  denotam os valores esperado e da variância para  $Y$  respectivamente.

A expressão (5) parece impor uma relação linear entre a variância e a média, no entanto, diversas parametrizações para  $p$ , levam a relações distintas entre aqueles dois momentos, uma afirmação sustentada por Johnson, Kotz e Kemp (1992). De acordo com estes autores, diversas parametrizações para  $p$ , ocasionam modelos de regressão distintos, pois conduzem à diferentes hipóteses acerca da forma funcional da heterocedasticidade existente e, em geral, as estimativas de  $\beta$  não são iguais.

Em todos os modelos foi feita a remoção de covariáveis na tentativa de identificar qual é o verdadeiro conjunto das covariáveis, que constituem fatores explicadores do aumento dos casos de homicídios nos municípios do Estado de São Paulo. Esta remoção de covariáveis foi feita tendo em conta o seu peso na modelagem e a significância estatística, e condicionada a minimização do problema da sobredispersão dos dados.

### 3. Resultados e Discussão

#### 3.1 Análise descritiva dos indicadores de demográficos

Foram avaliados seis indicadores demográficos, em 620 municípios do Estado de São Paulo. O indicador demográfico *percentual de jovens de 15 a 17 anos que não frequentam a escola*, apresentou a menor média e o maior valor de coeficiente de variação. A *renda per capita* apresentou maior média entre os indicadores avaliados, de acordo com os resultados da Tabela 1, no entanto, o coeficiente de variação para este indicador é considerado de alta variabilidade de acordo com a classificações de Sampaio (1988) e Gill (1987). Em geral, a maior parte dos indicadores analisados, apresentaram grandes variabilidades, tal como é observado nos valores do coeficiente de variação apresentado para cada um dos indicadores. No entanto, os indicadores *percentual de pessoas que vivem em domicílios com banheiro e água encanada* e *percentual de pessoas que vivem em domicílios com energia elétrica e geladeira*, apresentaram baixa variabilidade. Estes resultados são similares aos resultados obtidos por Gawryszewski e Costa (2005), em um trabalho em que o objetivo foi analisar a relação entre a taxa de homicídios e alguns indicadores socioeconômicos.

**Tabela 1.** Estatísticas descritivas dos indicadores demográficos no Estado de São Paulo, ano de 2000.

Indicador demográfico	N	Média	DP	CV%
Taxa de mortalidade infantil	620	15,52	5,07	32,67
% de jovens de 15 a 17 anos que não frequentam a escola	620	7,88	3,16	40,10
% de pessoas analfabetas de 25 anos ou mais	620	13,73	4,53	32,99
Renda per capita (Em Reais)	620	276,2	90,24	32,67
% de pessoas que vivem em domicílios com banheiro e água encanada	620	94,55	5,9	6,24
% de pessoas que vivem em domicílios com energia elétrica e geladeira	620	94,12	6,89	7,32

N=número de municípios avaliados no Estado de São Paulo; DP= Desvio padrão; CV%=Coeficiente de variação.  
Fonte: Autores.

#### 3.2 Modelagem do número de homicídios em função aos indicadores demográficos

Na Tabela 2, observam-se três cenários de modelos. No primeiro cenário, foi ajustado o modelo de distribuição Poisson ( $ps_1$ ), para os dados de contagem “número de homicídios” considerando todas variáveis independentes, isto é, todos indicadores demográficos. Este modelo, embora tenha apresentado estimativas de todos os coeficientes estatisticamente significativos ao nível de significância de 5%, tal como é observado nos valores de  $Pr(> |z|)$  todos menores que 0,05, apresenta problemas de qualidade de ajuste. Por exemplo, a estatística *deviance* que mensura a qualidade de ajuste para os modelos de contagem (Coxe, West & Aiken, 2009), têm associado o teste qui-quadrado com p-valor igual a 0, daí que, a hipótese nula sobre a qualidade de ajuste foi rejeitada. Resultados similares são observados nos modelos  $ps_2$  e  $ps_3$ .

No modelo  $ps_1$ , observa-se, por exemplo, que o quociente entre o desvio residual da *deviance* e os graus de liberdade é igual a  $118.75$  ( $\frac{\text{desvio residual}}{\text{graus de liberdade}} = \frac{73030}{615} = 118,75 > 1$ ), o que denota a presença de fortes indícios de sobredispersão tal como explicam os autores Hausman, Hall e Griliches (1984) e Hong et al. (2005). No ajuste do modelo  $ps_2$ , foi introduzida na modelagem termo *offset* na sua fórmula, com o objetivo de controlar o problema da sobredispersão, entre tanto, prevaleceu a precariedade do modelo. Estes resultados mostraram que o modelo Poisson, é menos apropriado para ajustar os dados de homicídios nos municípios do Estado de São Paulo, com base nos indicadores demográficos considerados nesta pesquisa.

Na Tabela 3, são apresentados resultados de quatro modelos de respostas Binomiais Negativas, que foram avaliados na tentativa de determinar o modelo que melhor se ajusta aos casos de homicídios ocorridos no ano de 2000, nos municípios do Estado de São Paulo. Para o ajuste do modelo 1 ( $bn_1$ ), foram consideradas quatro indicadores demográficos, nomeadamente: “*Percentual de pessoas analfabetas de 25 anos ou mais*”, “*Renda per capita*”, “*Percentual de pessoas que vivem em domicílios com banheiro e água encanada*” e “*Percentual de pessoas que vivem em domicílios com energia elétrica e geladeira*”. O modelo apresentou uma estatística *deviance* que tem associado o teste de qui-quadrado com p-valor igual a  $0,15904 > 0,05$ , e com este resultado do teste, não se rejeita a hipótese sobre adequabilidade deste modelo. Este modelo não apresenta também indícios de sobredispersão, pois o quociente entre o valor do desvio residual da *deviance* e os graus de liberdade é aproximadamente igual a 1 ( $\frac{\text{desvio residual}}{\text{graus de liberdade}} = \frac{649,99}{615} \approx 1$ ). No entanto, nem todas estimativas dos coeficientes de regressão, associados aos indicadores demográficos, são estatisticamente significativas, como é o caso do coeficiente  $\hat{\beta}_4$  associado ao indicador demográfico *Renda per capita*, que apresentou valor de  $Pr(> |z|) = 0,235992 > 0,05$ . Esta falta de significância observada na estimativa do coeficiente  $\hat{\beta}_4$ , torna o modelo menos adequado para a modelagem do número de casos de homicídios nos municípios de São Paulo.

O modelo 2 ( $bn_2$ ), foi ajustado sem levar em conta a covariável “*Percentual de pessoas que vivem em domicílios com energia elétrica e geladeira*”. Observa-se que esta covariável, é estatisticamente significativa e a sua remoção foi feita pela falta da significância prática com a ocorrência de homicídios. A *deviance está* associada ao teste qui-quadrado a um p-valor igual a  $0,1616782 > 0,05$ , o que leva a não rejeição da hipótese sobre adequabilidade do modelo. No entanto, a presença de variáveis estatisticamente não significativas, torna este modelo menos apropriado para a modelagem dos dados de homicídios.

O modelo 3 ( $bn_3$ ), foi ajustado tomando em consideração duas covariáveis “*Percentual de pessoas de 25 anos ou mais analfabetas*” e “*Renda per capita*”. Foi observada significância estatística das estimativas dos coeficientes de regressão, associados as duas covariáveis consideradas. Uma outra observação satisfatória sobre este modelo, está relacionada com a estatística *deviance* que tem associado o teste qui-quadrado com p-valor igual a  $0,1676803 > 0,05$ , o que leva a não rejeição da hipótese sobre a adequabilidade do modelo (Hausman, Hall, & Griliches, 1984). Para além disso, o modelo não apresenta indícios de sobredispersão, pois o quociente entre a *deviance* e os graus de liberdade é aproximadamente igual a 1 ( $\frac{\text{desvio residual}}{\text{graus de liberdade}} = \frac{650,78}{617} \approx 1$ ). Estas observações satisfatórias que recaem sobre o modelo  $bn_3$ , fazem com que o modelo seja candidato para a modelagem dos casos de homicídios registados nos municípios de São Paulo no ano de 2000.

O modelo 4 ( $bn_4$ ), ainda na tabela 3, foi ajustado tal como no modelo  $bn_1$ , no entanto, para o ajuste do  $bn_4$ , foi eliminada no modelagem a covariável “*Renda per capita*” a qual não foi estatisticamente significativa no modelo  $bn_1$ . Além da presença de coeficientes de regressão estatisticamente significativos em todas covariáveis, observa-se ausência dos indícios de sobredispersão. Com o teste do qui-quadrado não se rejeita a hipótese sobre adequabilidade deste modelo. Estes resultados, tornam o modelo  $bn_4$  também como candidato para a modelagem dos casos de homicídios registados nos municípios do

Estado de São Paulo, no ano de 2000. Desta forma, os modelos  $bn_3$  e  $bn_4$  são os que podem ser tomados em consideração para análise dos casos de homicídios.

A seleção do melhor modelo foi feita recorrendo ao Critério de Informação Akaike (AIC), tal como explicam os autores Akaike (1974), Akaike (1985), Burnham E Anderson (2002), Burnham e Anderson (2004) e Yang (2005). De acordo com estes autores, o melhor modelo, dentro dos modelos candidatos é aquele que apresenta menor valor de AIC. Desta forma, de acordo com os resultados dos valores de AIC dos modelos  $bn_3$  e  $bn_4$ , o modelo  $bn_4$  é mais apropriado para a modelagem dos dados de homicídios registados no ano de 2000, nos municípios de São Paulo, uma vez que foi o modelo que apresentou o menor valor de AIC.

De acordo com os resultados do modelo  $bn_4$ , os casos de homicídios no estado de São Paulo são explicados por três indicadores demográficos: *percentual de pessoas analfabetas de 25 anos ou mais; percentual de pessoas que vivem em domicílios com banheiro e água encanada e percentual de pessoas que vivem em domicílios com energia elétrica e geladeira*". Estes resultados mostram que as desigualdades socioeconômicas, são determinantes para ocorrência de homicídios nos municípios de São Paulo. Por exemplo, os resultados mostram que, a falta de uma habitação condigna, constitui um fator chave para ocorrência da violência. Resultados similares foram encontrados por Gawryszewski e Costa (2005). De acordo com estes autores, a desigualdade na habitação digna, acesso à educação, contribuem significativamente no aumento do número de casos de violência que causam homicídios, nos municípios de São Paulo. Os resultados de Macedo et al, (2001), corroboram com os achados nesta pesquisa. De acordo com estes autores, os altos níveis de pobreza, constituem fatores explicadores da violência, na maioria das cidades brasileiras.

**Tabela 2.** Modelos de distribuição Poisson.

Modelo	Parâmetros	Estimativa	Erro padrão	valor z	Pr(> z )
<i>ps<sub>1</sub></i>	$\beta_0$	-4,3293161	0,3248841	-13,33	<2e-16
	$\beta_1$	0,0821143	0,0021482	38,23	<2e-16
	$\beta_2$	-0,4102577	0,0035537	-115,45	<2e-16
	$\beta_3$	-0,0829820	0,0031475	-26,36	<2e-16
	$\beta_4$	-0,0031548	0,0001511	-20,88	<2e-16
	$\beta_5$	0,1628403	0,0037709	43,18	<2e-16
	$\beta_6$	-0,0508730	0,0026738	-19,03	<2e-16
Desvio Residual=73030 Deviance=0			Graus de liberdade=613 AIC=74424		
<i>ps<sub>2</sub></i>	$\beta_0$	-2,350	0,032610	-72,06	<2e-16
	$\beta_1$	0,0053	0,0019560	18,05	<2e-16
	$\beta_2$	-0,05710	0,0039630	-144,06	<2e-16
	$\beta_3$	0,012960	0,0027570	47,00	<2e-16
	$\beta_4$	-0,010150	0,0001373	-73,98	<2e-16
	$\beta_5$	0,306500	0,0041350	74,13	<2e-16
	$\beta_6$	-0,08975	0,0025320	-35,45	<2e-16
Desvio Residual=92660 Deviance=0			Graus de liberdade=613 AIC=94050		
<i>ps<sub>3</sub></i>	$\beta_0$	-4,129	0,002634	-156,8	<2e-16
	$\beta_4$	-0,0104	0,00008822	-117,7	<2e-16
	Desvio Residual=127500 Deviance=0			Graus de liberdade=618 AIC=128900	

*ps<sub>1</sub>* = Modelo Poisson 1; *ps<sub>2</sub>* = Modelo Poisson 2; *ps<sub>3</sub>* = Modelo Poisson 3;  $\beta_0$  = constate do modelo de regressão;  $\beta_1$  = estimativa do coeficiente de regressão associado a covariável “Taxa de mortalidade infantil”;  $\beta_2$  = estimativa do coeficiente de regressão associado a covariável “% de jovens de 15 a 17 anos que não frequentam a escola”;  $\beta_3$  = estimativa do coeficiente de regressão associado a covariável “% de pessoas analfabetas de 25 anos ou mais”;  $\beta_4$  = estimativa do coeficiente de regressão associado a covariável “Renda per capita”;  $\beta_5$  = estimativa do coeficiente de regressão associado a covariável “% de pessoas que vivem em domicílios com banheiro e água encanada”;  $\beta_6$  = estimativa do coeficiente de regressão associado a covariável % de pessoas que vivem em domicílios com energia elétrica e geladeira”.

Fonte: Autores

**Tabela 3.** Modelos de respostas Binomiais Negativas.

Modelo	Parâmetros	Estimativa	Erro padrão	valor z	Pr(> z )
$bn_1$	$\beta_0$	2,622159	2,008237	1,306	0,191654
	$\beta_3$	-0,090977	0,033145	-2,745	0,006054*
	$\beta_4$	-0,001949	0,001644	-1,185	0,235992
	$\beta_5$	0,116794	0,032633	3,579	0,000345*
	$\beta_6$	-0,093649	0,027553	-3,399	0,000677*
Desvio Residual=649,99 Deviance=0,1590378		Graus de liberdade=615 AIC=3639,8			
$bn_2$	$\beta_0$	2,841683	2,0277	1,401	0,161085
	$\beta_3$	-0,110109	0,033269	-3,31	0,000934*
	$\beta_4$	-0,00311	0,001632	-1,905	0,056798
	$\beta_5$	0,027942	0,018965	1,473	0,140666
Desvio Residual=650,63 Deviance=0,1616782		Graus de liberdade=616 AIC=3648,1			
$bn_3$	$\beta_0$	5,904818	0,837943	7,047	0,0018*
	$\beta_3$	-0,136108	0,03234	-4,209	0,0025*
	$\beta_4$	-0,003287	0,001623	-2,026	0,0428*
Desvio Residual=650,78 Deviance=0,1676803		Graus de liberdade=617 AIC=3649,2			
$bn_4$	$\beta_0$	1,67772	1,93965	0,865	0,38706
	$\beta_3$	-0,05798	0,02424	-2,392	0,016753*
	$\beta_5$	0,12513	0,03259	3,839	0,000123*
	$\beta_6$	-0,10244	0,02714	-3,774	0,000161*
Desvio Residual=650,09 Deviance=0,165307		Graus de liberdade=616 AIC=3639			

$bn_1$  = Modelo Binomial Negativa 1;  $bn_2$  = Modelo Binomial Negativa 2;  $bn_3$  = Modelo Binomial Negativa 3;  $bn_4$  = Modelo Binomial Negativa 4;  $\beta_0$  = constate do modelo de regressão;  $\beta_1$  = estimativa do coeficiente de regressão associado a covariável "Taxa de mortalidade infantil";  $\beta_2$  = estimativa do coeficiente de regressão associado a covariável "% de jovens de 15 a 17 anos que não frequentam a escola";  $\beta_3$  = estimativa do coeficiente de regressão associado a covariável "% de pessoas analfabetas de 25 anos ou mais";  $\beta_4$  = estimativa do coeficiente de regressão associado a covariável "Renda per capita";  $\beta_5$  = estimativa do coeficiente de regressão associado a covariável "% de pessoas que vivem em domicílios com banheiro e água encanada";  $\beta_6$  = estimativa do coeficiente de regressão associado a covariável % de pessoas que vivem em domicílios com energia elétrica e geladeira"; \* = covariável estatisticamente significativa.

Fonte: Autores

#### 4. Conclusão

Verificou-se que o modelo para respostas Binomiais Negativas é o mais apropriado para modelar os dados de homicídios nos municípios de São Paulo.

A falta de acesso à educação tem determinado o aumento da taxa de analfabetismo na maioria dos municípios avaliados, associada com a pobreza extrema que se verifica na maioria das zonas urbanas do Estado de São Paulo são os principais vetores para ocorrência da violência urbana, o que contribui para o aumento dos casos de homicídios.

Desta forma, torna-se imprescindível a realização de estudos futuros que permitem, o mapeamento dos municípios do Estado de São Paulo, com maiores taxas de analfabetismo e pobreza extrema e, auxiliem ao governo local, na tomada de medidas de monitoria dos casos de homicídios.

#### Referências

- Akaike, H. (1974). A new look at the statistical model identification., *IEEE Transactions on Automatic control*, 19 (6):716723, 10.1109/TAC.1974.1100705.
- Akaike, H. (1985). *Prediction and entropy in Atkinson, A. C.; Fienberg, S. E. (eds.). A Celebration of Statistics.* Springer, 1–2.
- Barros, R. P. D., Henriques, R., & Mendonça, R. (2021). Desigualdade e pobreza no Brasil: retrato de uma estabilidade inaceitável. *Revista Brasileira de Ciências Sociais*. 15(42), 123-142. <https://doi.org/10.1590/S0102-69092000000100009>.
- Burnham, K. P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.). Springer-Verlag.

- Burnham, K. P.; Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33: 261–304, doi:10.1177/0049124104268644, S2CID 121861644.
- Cameron, AC, Trivedi, PK (1998). *Regression analysis of count data*. Cambridge University Press.
- Campello, T., Gentili, P., Rodrigues, M., & Hoewell, G. (2018). Faces da desigualdade no Brasil: um olhar sobre os que ficam para trás. *Saúde em Debate*, 42, 54-66. 10.1590/0103-11042018s305.
- Coxe, S., West, S. G. & Aiken, L. S. (2009). The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. *Journal of Personality Assessment*. 91(2), 121–136. <https://doi.org/10.1080/00223890802634175>.
- Estrela, C. (2018). *Metodologia Científica: Ciência, Ensino, Pesquisa*. Editora Artes Médicas.
- Gawryszewski, V. & Costa, L. (2005). Social inequality and homicide rates in Sao Paulo City, Brazil. *Revista de saúde pública*. 39. 191-7. 10.1590/S0034-89102005000200008.
- Gill, J. L. (1987). *Design and analysis of experiments in the animal and medical sciences*. Ames The Iowa State University: Press. 1, 411p.
- Hausman, J., Hall, B. H. & Griliches, Z. (1984). Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica*, 52(4), 909-938.
- Hilbe J M. (2007). *Negative binomial regression*. Cambridge University Press.
- Hilbe, J.M. (2011a). Modeling Count Data. In: Lovric M. (eds) International Encyclopedia of Statistical Science. *Springer*, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-04898-2369>
- Hilbe, JM. (2011b). *Negative binomial regression*, (2nd ed.), Cambridge University Press.
- Hong, D., Kim, J., Kim, W., Lee, Y. & Yang, H. (2005). Development of traffic accident prediction models by traffic and road characteristics urban areas. *Proceedings of the Eastern Asia Society for Transportation Studies*, 5, 2046-2061.
- Johnson, N., Kotz, S. & Kemp, A. (1992). *Univariate discrete distributive*. John Wiley & Sons.
- Koche, J. C. (2011). Fundamentos de metodologia científica. Petrópolis: Vozes. [http://www.brunovivas.com/wp-content/uploads/sites/10/2018/07/K%C3%B6che-Jos%C3%A9-Carlos0D0AFundamentos-de-metodologia-cient%C3%ADfica\\_-teoria-da0D0Aci%C3%Aancia-e-inicia%C3%A7%C3%A3o-%C3%A0-pesquisa.pdfhttps://repositorio.ufsm.br/bitstream/handle/1/15824/Lic\\_Computacao\\_Metodologia-Pesquisa-Cientifica.pdf?sequence=1](http://www.brunovivas.com/wp-content/uploads/sites/10/2018/07/K%C3%B6che-Jos%C3%A9-Carlos0D0AFundamentos-de-metodologia-cient%C3%ADfica_-teoria-da0D0Aci%C3%Aancia-e-inicia%C3%A7%C3%A3o-%C3%A0-pesquisa.pdfhttps://repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-Pesquisa-Cientifica.pdf?sequence=1)
- Ludke, M. & Andre, M. E. D. A. (2013). *Pesquisas em educação: uma abordagem qualitativa*. E.P.U.
- Macedo, A. C. Paim, J. S., Da Silva, L. M. V. & Costa, M. da C. N. (2001). Violência e desigualdade social: mortalidade por homicídios e condições de vida em Salvador, Brasil. *Revista de Saúde Pública*. 35(6), 515-522. <https://doi.org/10.1590/S0034-89102001000600004>. <https://doi.org/10.1590/S0034-89102001000600004>.
- Pekny, A. C. & Ricardo, C. M. (2017). Mapeamento dos principais desafios de violência e criminalidade no Brasil. *Instituto da Paz*. ANÁLISE Nº 36/2017.
- Pereira A. S., Shitsuka, D. M., & Parreira, F. J. & Shitsuka, R (2018). *Metodologia da pesquisa científica*. UFSM.
- Ramalho, J. J. Dos S. (1996). *Modelos de Regressão Para Dados de Contagem*. Universidade Técnica de Lisboa. Portugal.
- Ross, G., & Preece, D. (1985). The Negative Binomial Distribution. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 34(3), 323-335. 10.2307/2987659.
- Ruotti, C., Regina, F. L., De Almeida, J. F., Nasser, M. M. S. & Peres, M. F. T. (2017). A ocorrência de homicídios no município de São Paulo: mutações e tensões a partir das narrativas de moradores e profissionais. *Saúde e Sociedade*. 26(4), 999-1014. <https://doi.org/10.1590/S0104-12902017170254>. ISSN 1984-0470. <https://doi.org/10.1590/S0104-12902017170254>.
- Sampaio, V. R. (1978.) Bananeira: características de desenvolvimento e de produção. In: Congresso Brasileiro De Fruticultura, 4, 1977, Salvador. *Anais. Cruz das almas*: SBF, 424p. 53-57.
- Winkelmann, R. (2003). *Econometric analysis of count data*. (4th ed.). Springer: Heidelberg.
- Yang, Y. (2005). Can the strengths of AIC and BIC be shared? *Biometrika*, 92: 937–950, doi:10.1093/biomet/92.4.937.
- Yin, R. K. (2015). *O estudo de caso*. Bookman.