

Análise dos óbitos por causas violentas na cidade de João Pessoa no período de 2000 a 2005 por Regressão de Poisson

Patrícia Silva Nascimento, Ronei Marcos de Moraes

Departamento de Estatística, UFPB,

58051-900, João Pessoa, PB

patriciapb1408@yahoo.com.br, ronei@de.ufpb.br

Resumo: Entre 1971-1991, o Brasil apresentou uma tendência de crescimento nas taxas de mortalidade por causas externas, sendo os homicídios uma das principais causas de morte. Segundo o Mapa da Violência 2006, a capital João Pessoa em 2004 estava em 9º lugar com uma taxa de homicídios de 42,6 por 100 mil habitantes. Objetivando encontrar um modelo para ajustar os óbitos por causas violentas na cidade de João Pessoa no período de 2000 a 2005, utilizou-se a metodologia de Regressão de Poisson. Esta metodologia é apropriada para dados de contagem. Os dados de óbitos foram fornecidos pela Secretaria Municipal de Saúde de João Pessoa e selecionados por causas violentas. O modelo de Poisson obtido para o ajuste dos óbitos por causas violentas com intercepto, utilizou as variáveis raça categoria Parda e estado civil categoria Solteiro, por serem significativas. Os coeficientes estimados foram 1,84027 para o intercepto, 0,02796 para estado civil categoria Solteiro e 0,04658 para raça categoria Parda. Depois de obtido o modelo fez-se uma comparação dos valores observados e ajustados por bairro. Este modelo exibiu bons ajustes para os bairros com contagens maiores, porém para os bairros com contagens menores não há o mesmo padrão de qualidade.

Palavras-chave: Modelos Lineares Generalizados, Dados de contagem, Regressão de Poisson, Óbitos por causas violentas.

Introdução

Entre 1971 e 1991, o Brasil apresentou uma tendência de crescimento nas taxas de mortalidade por causas externas, sendo os acidentes de trânsito e homicídios as principais causas de morte violenta [8]. No Brasil, os homicídios foram responsáveis por cerca de 30% dos óbitos por causas externas, em 1994, ocupando o primeiro lugar dentre elas [3]. Segundo o Mapa da Violência 2006, a Paraíba passou do 20º lugar em 1994 para o 19º lugar em 2004, com uma taxa de 18,6% do total de homicídios do país em 2004. Esta mesma fonte revela que a capital João Pessoa em 2004 estava em 9º lugar com uma taxa de homicídios de 42,6 por 100 mil habitantes [7]. Dentre as capitais do Nordeste João Pessoa foi uma das que

manteve níveis estáveis, sendo 19% em 1977, 19,4% em 1984 e 20,1% em 1994 [3].

Os trabalhos referidos anteriormente para os óbitos por causas violentas em João Pessoa motivaram o estudo destes óbitos. Este trabalho originou-se da carência de estudos para obtenção de um modelo para o ajuste destes óbitos.

O objetivo deste artigo é encontrar um modelo para ajustar os óbitos por causas violentas na cidade de João Pessoa no período de Janeiro de 2000 a Dezembro de 2005. Os dados de óbitos foram fornecidos pela Divisão de Vigilância Epidemiológica da Secretaria Municipal de Saúde da Prefeitura Municipal de João Pessoa.

Metodologia

Os dados mensais de óbitos foram selecionados por causas violentas, depois foram organizados por escolaridade (nenhuma, de 1 a 3 anos, de 4 a 7 anos, de 8 a 11 anos e mais de 12 anos), estado civil (solteiro, casado, viúvo, separado judicialmente e união consensual) e raça ou cor (branca, preta, amarela e parda). Posteriormente separou-se os óbitos por bairro com as mesmas variáveis citadas anteriormente e obteve-se os óbitos ajustados. Para alcançar o objetivo deste artigo, como os dados de óbitos são de contagem, utilizou-se a metodologia de Modelos Lineares Generalizados (MLG) com enfoque na análise de Regressão de Poisson. Utilizou-se o *software R* para obtenção dos resultados da análise.

Modelos lineares generalizados (MLG) são usados quando se quer estimar o valor da variável dependente Y como uma combinação linear de variáveis explicativas X_1, X_2, \dots . Um valor observado y_i é considerado como a soma de uma componente sistemática (μ_i) e uma aleatória (erro) (ϵ_i) [4]:

$$y_i = \mu_i + \epsilon_i \quad (1)$$

onde y_i é o valor observado de Y para i ; μ_i é o valor predito de Y para i ; e ϵ_i é o erro aleatoriamente distribuído.

Os tipos de MLG são diferenciados na suposição da distribuição da probabilidade de Y_i e da função de ligação g que associa a média de $Y_i(\mu_i)$ com o preditor linear [4]:

$$g(\mu_i) = \sum_{j=1}^m \beta_j x_{ij} \quad (2)$$

onde m é o número de parâmetros; β_j é o parâmetro para a j -ésima variável independente; e x_{ij} é o valor da i -ésima observação na j -ésima variável independente. A equação (1) pode ser reescrita [4]:

$$y_i = g^{-1}\left(\sum_{j=1}^m \beta_j x_{ij}\right) + \epsilon_i \quad (3)$$

Uma variedade de modelos lineares pode ser construída especificando diferentes tipos de termos de erro ϵ_i e função de ligação g . Em regressão simples a distribuição do erro é Normal. O valor predito \hat{y}_i é igual a $\sum_{j=1}^m \beta_j x_{ij}$, assim a ligação é simplesmente a função de identidade. Em regressão de *Poisson*, a distribuição do erro é *Poisson*, o preditor linear é o logaritmo natural do valor predito e a função de ligação é o logaritmo. O modelo *loglinear* também usa o modelo de probabilidade de *Poisson* e a função de ligação é o logaritmo. Este modelo é considerado como um caso especial da regressão de *Poisson* quando todas as variáveis independentes são categóricas.

A distribuição de *Poisson* descreve a probabilidade de que um evento ocorre k vezes num período fixo dado que cada ocorrência é independente e tem uma probabilidade constante. A forma de uma distribuição de *Poisson* depende do valor de sua média. Se a média é próximo de zero, então a distribuição é fortemente enviesada; se a média é muito grande, a distribuição de *Poisson* pode ser aproximada pela normal. Na distribuição de *Poisson*, a média e a variância são iguais a λ . A equação da função de densidade de *Poisson* é dada por [5]:

$$P(y) = f(y|\lambda) = \frac{\lambda^y e^{-\lambda}}{y!} \quad (4)$$

onde y é a variável dependente.

O modelo de *Poisson*, ao contrário do modelo normal, supõe que a variância seja proporcional a média e pode ser aplicado para modelar o número de óbitos. O modelo de *Poisson* aplica uma relação exponencial entre λ , X e β , que assegura um número de óbitos sempre não-negativo:

$$\lambda = e^{X\beta} \quad (5)$$

onde X é a matriz de n linhas por p colunas das variáveis independentes; β é o vetor de p linhas por 1 coluna dos parâmetros.

A análise de regressão linear assume que a variável dependente tem uma relação linear com as variáveis independentes. Se existe só uma variável independente X , o valor predito da variável dependente Y para i é dado por [4]:

$$\hat{Y}_i = \beta_0 + \beta_1 X_i \quad (6)$$

onde \hat{Y}_i é o valor predito de Y para i , X_i é o valor de X para i , β_0 é o valor do intercepto de f quando X é zero, e β_1 é a soma da inclinação pela qual \hat{Y}_i muda quando X_i aumenta uma unidade. Em regressão linear, as variáveis Y_i assumem ter distribuição Normal e uma variância constante.

A regressão de *Poisson* é apropriada quando a variável dependente é contagem, tais como o número de óbitos. No modelo de regressão de *Poisson* o valor predito da variável dependente para i é a probabilidade máxima estimada X_i da média de uma variável Y_i com distribuição de *Poisson*. O logaritmo natural deste estimador é igual a uma combinação linear dos valores das variáveis independentes; se existe só uma variável independente a relação tem a forma [1]:

$$\ln(\hat{\lambda}_i) = \beta_0 + \beta_1 x_i \quad (7)$$

Esta equação pode ser reescrita:

$$\hat{\lambda}_i = \exp(\beta_0 + \beta_1 x_i) \quad (8)$$

Diferente da regressão simples, o modelo de *Poisson* não assume que os dados são homocedásticos. A variância de cada caso é igual ao correspondente valor predito. O valor observado de Y_i é uma contagem dos eventos independentes gerados por uma distribuição de *Poisson* com parâmetro X_i .

Os critérios utilizados para verificar a adequação do modelo foram os seguintes: o que obteve o menor *AIC* (Critério de informação de Akaike); análise de variância para verificar a significância das variáveis e análise dos resíduos (comportamento aleatório e distribuição Normal).

Resultados

Primeiramente, ajustou-se um modelo de *Poisson* com todas as variáveis. Entretanto, somente foram significativas as variáveis raça categoria Parda e estado civil categoria Solteiro. Por esta razão ajustou-se os seguintes modelos: um com a variável raça categoria Parda; outro com a variável estado civil categoria Solteiro; e outro com a combinação dessas duas variáveis. O *AIC* do modelo de Regressão de *Poisson* com todas as variáveis foi 365,76, o do modelo com a variável raça categoria Parda foi 348,32, o do modelo com a variável estado civil categoria Solteiro foi 363,21. Verificou-se que o modelo com a combinação das duas variáveis, indicou menor *AIC* (345,68) em relação aos outros modelos testados. Por este motivo escolheu-se este modelo de Regressão de *Poisson* para ajustar os óbitos por causas violentas. Este modelo de Regressão de *Poisson* obteve graus de liberdade total 71 e residual 69. O desvio nulo do modelo foi 87,103, o desvio residual foi 13,178. Com as estatísticas descritivas dos desvios residuais da Regressão de *Poisson* observou-se que os desvios são pequenos, a mediana é 0,0026 indicando um bom ajuste do modelo. Os resultados do ajuste do modelo são apresentados a seguir.

Tabela 1: Estatísticas das variáveis da Regressão de *Poisson* dos dados de óbitos por causas violentas em João Pessoa no período de 2000 a 2005

Variáveis	Coefficientes	Erro Padrão	z valor	P-valor
Intercepto	1,84027	0,1121	16,415	< 2e-16
Parda	0,04658	0,0104	4,466	7,97e-06
Solteiro	0,02796	0,0129	2,159	0,0309

Tabela 2: Análise de variância da Regressão de *Poisson* dos dados de óbitos por causas violentas em João Pessoa no período de Janeiro de 2000 a Dezembro de 2005

Variáveis	gl desvio	Resí duo	gl resíduo	Dev	P(> Chi)
Nulo			71	87,10	
Parda	1	69,29	70	17,82	8,512e-17
Solteiro	1	4,64	69	13,18	0,031

A Tabela 1 mostra os coeficientes estimados na Regressão de *Poisson* dos dados de óbitos por causas violentas na cidade de João Pessoa no período de Janeiro de 2000 a Dezembro de 2005. Observa-se que as variáveis raça categoria Parda (0,04658) e estado civil categoria Solteiro (0,02796) relacionam-se positivamente com o número de óbitos. Mostra também os p-valores das variáveis da Regressão de *Poisson*. Observa-se que as variáveis apresentam p-valores significativos a 0,05, então podem ser incluídas no modelo. Confirma-se a entrada dessas variáveis com a análise de variância da Tabela 2, sendo significativas a 0,05.

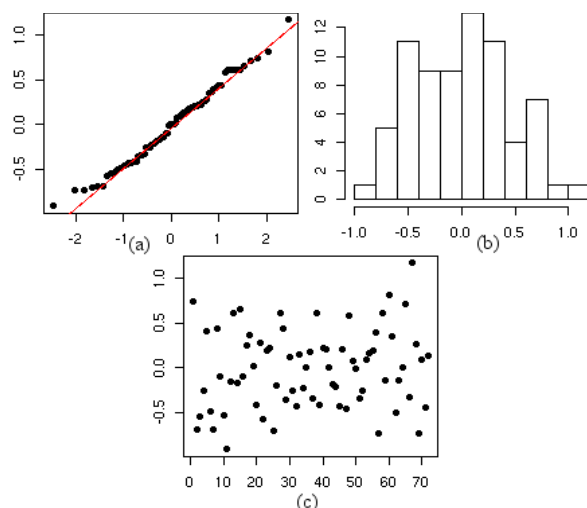


Figura 1: *Qqplot* (a), Histograma (b) e Gráfico dos resíduos (c) do modelo de *Poisson* para verificar a adequação do modelo

A Figura 1 mostra que o modelo é adequado, pois o gráfico dos resíduos (c) se comporta de forma aleatória. Verifica-se também com o *Qqplot* (a) e histograma (b) que os dados apresentam distribuição Normal [2]. O teste de normalidade de *Lilliefors* (*Kolmogorov-Smirnov*) [6] comprova esta normalidade com p-valor de 0,8067. Assim, o modelo de *Poisson* obtido é da seguinte forma:

$$\hat{y} = \exp(1,84027 + 0,02796 x_1 + 0,04658 x_2)$$

onde \hat{y} é o óbito ajustado; x_1 é o valor da variável estado civil categoria Solteiro e x_2 é o valor da variável raça categoria Parda.

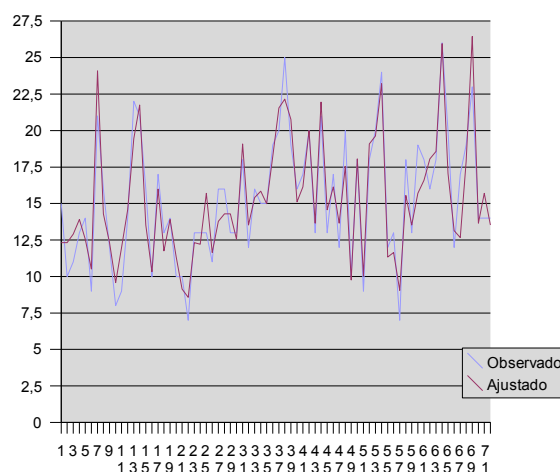


Figura 2: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no período de Janeiro de 2000 a Dezembro de 2005

A Figura 2 apresenta os óbitos observados e ajustados pela regressão de *Poisson*. Observa-se que os óbitos ajustados são próximos dos observados, indicando um bom ajuste do modelo para João Pessoa. Posteriormente, foi feito o ajuste dos óbitos por bairro com o modelo de Regressão de *Poisson* estabelecido.

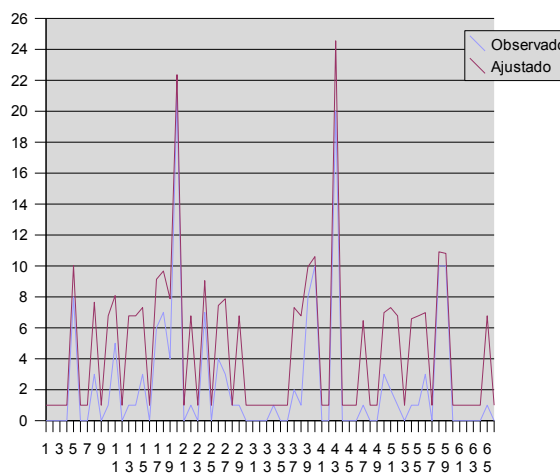


Figura 3: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2000

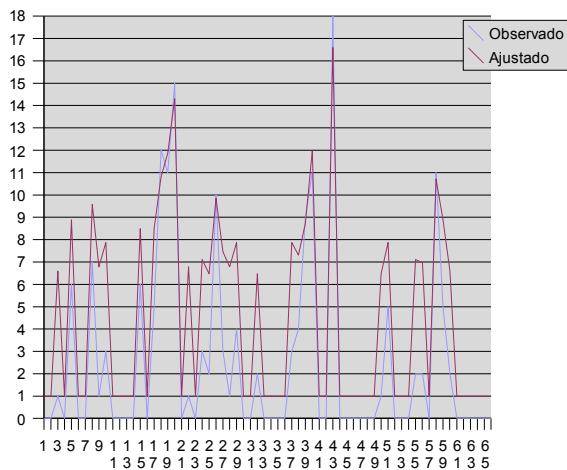


Figura 4: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2001

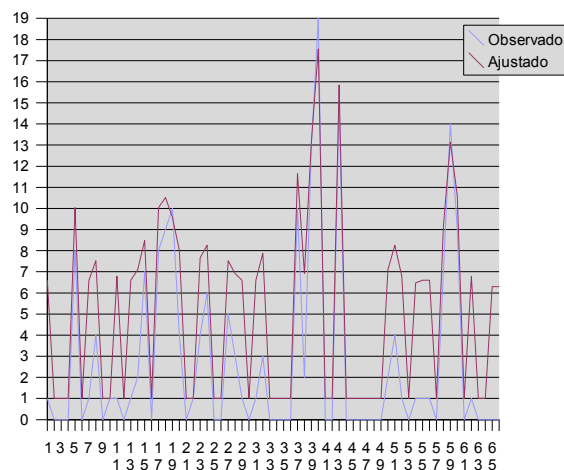


Figura 7: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2004

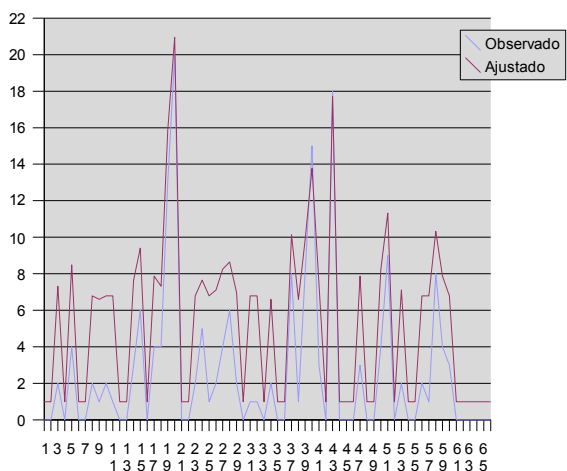


Figura 5: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2002

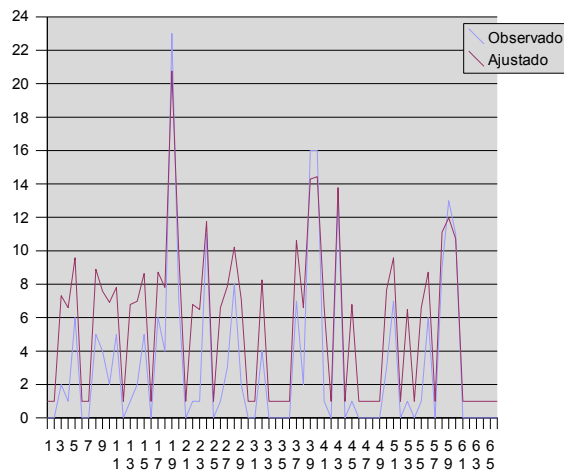


Figura 8: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2005

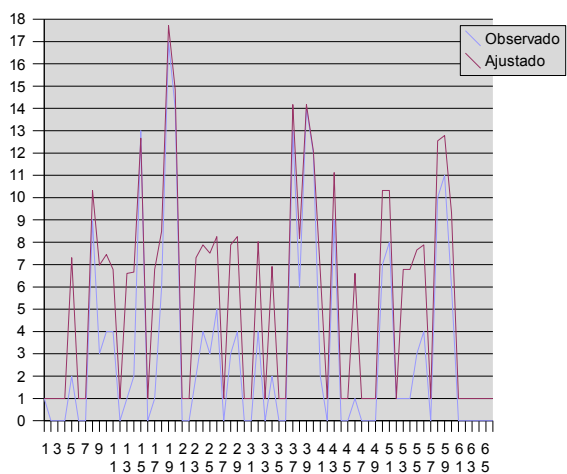


Figura 6: Óbitos por causas violentas observados e ajustados pela regressão de *Poisson* em João Pessoa no ano de 2003

Nas Figuras 3 a 8 tem-se a comparação dos óbitos ajustados pela regressão de *Poisson*, por bairro e para cada ano estudado com respeito aos valores observados. Nota-se que o modelo obteve bons ajustes para os óbitos por causas violentas. Em todos os anos verifica-se uma boa aproximação dos valores ajustados em relação aos bairros com valores de óbitos maiores. Em relação aos bairros com valores de óbitos menores observa-se uma super-estimação dos óbitos. A seguir, fez-se mapas para comparação da distribuição espacial dos valores observados e ajustados.

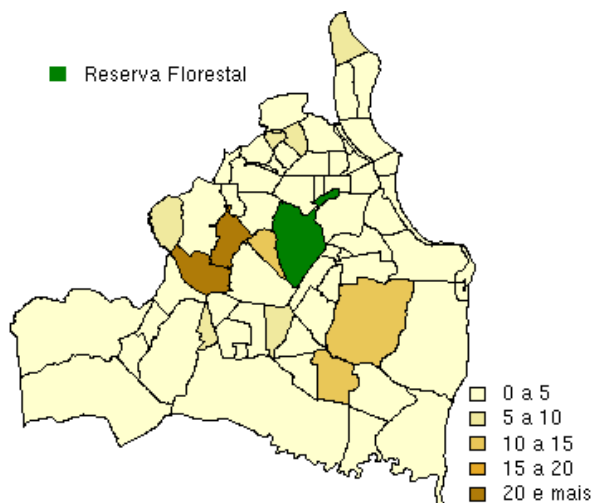


Figura 9: Óbitos por causas violentas observados em João Pessoa no ano de 2000

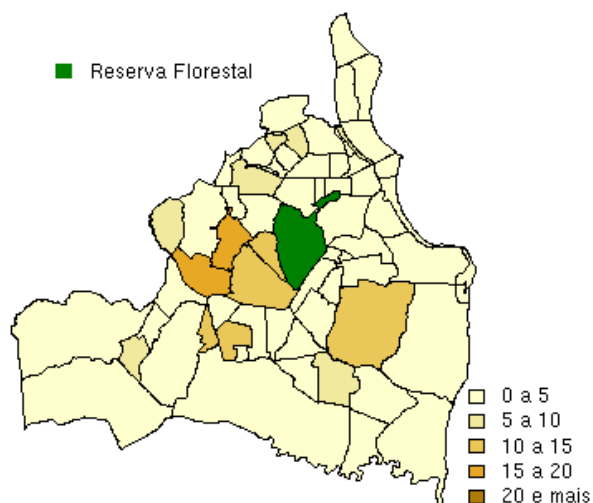


Figura 11: Óbitos por causas violentas observados em João Pessoa no ano de 2001

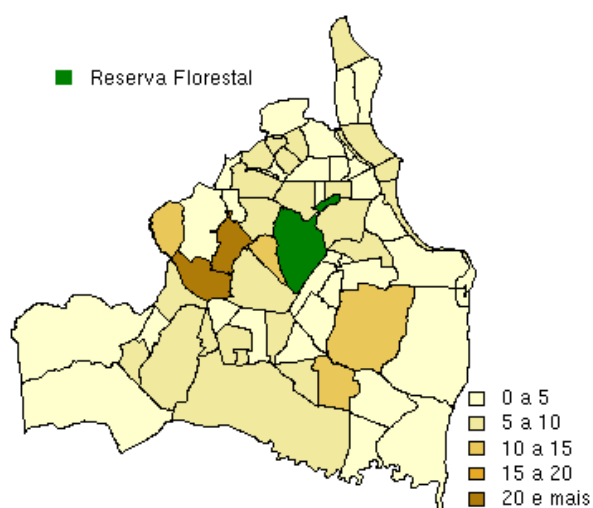


Figura 10: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2000

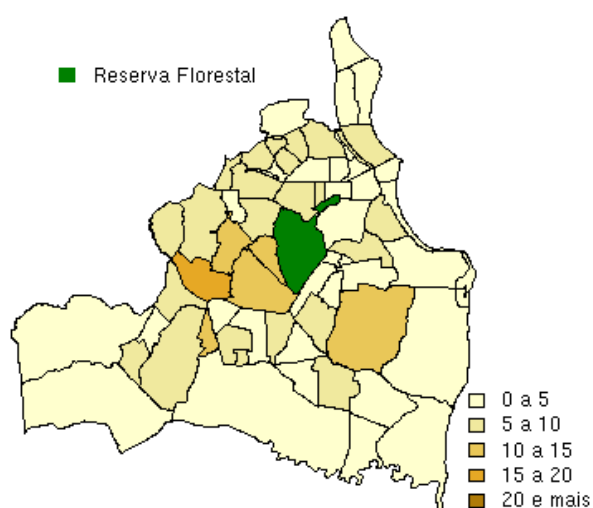


Figura 12: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2001

Pode-se observar na Figura 9 e 10 que em 2000 os óbitos observados e ajustados apresentam a mesma distribuição para os bairros com valores maiores, no caso dos bairros com valores menores a qualidade do ajuste não foi a mesma. Nota-se que os ajustados e observados apresentam mais óbitos nos bairros de Cruz das Armas e Oitizeiro que se localizam na região oeste da cidade. Seguidos por Rangel da região oeste e Mangabeira e Valentina da região sul da cidade.

Nota-se nas Figuras 11 e 12 que em 2001 os valores observados e ajustados se distribuem de forma semelhante para os bairros com valores de óbitos maiores. Observa-se que apresentam mais óbitos no bairro de Oitizeiro. Em seguida, aparecem Cruz das Armas, Cristo Redentor, Costa e Silva e Rangel que se localizam na região oeste e Mangabeira na região sul da cidade.

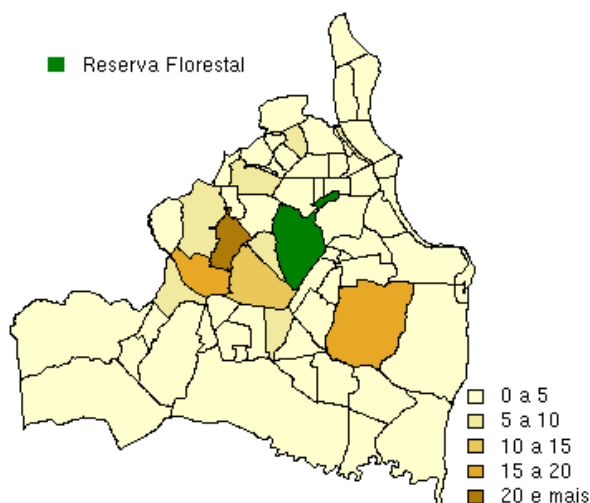


Figura 13: Óbitos por causas violentas observados em João Pessoa no ano de 2002

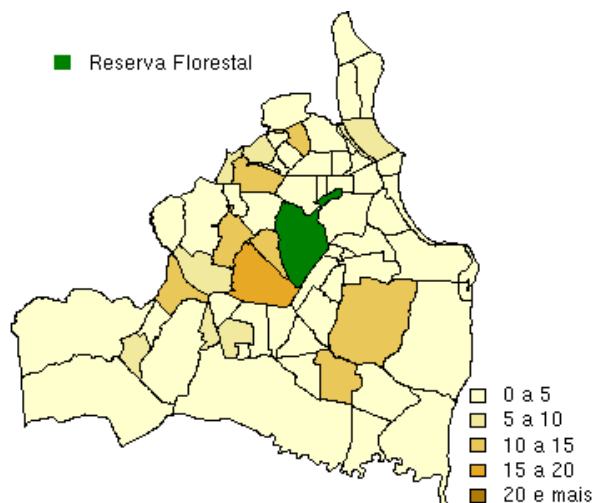


Figura 15: Óbitos por causas violentas observados em João Pessoa no ano de 2003

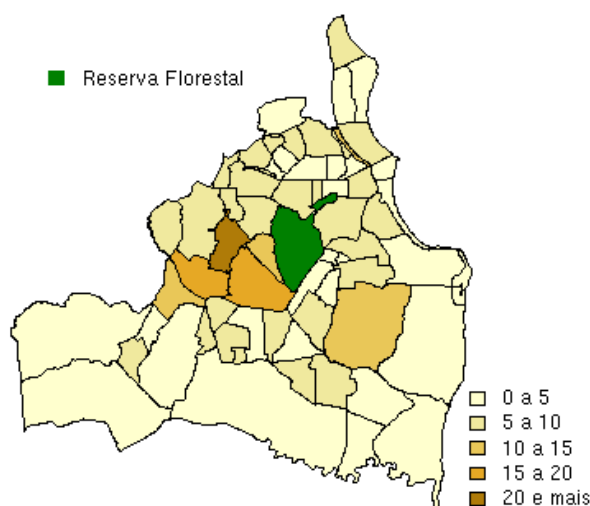


Figura 14: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2002

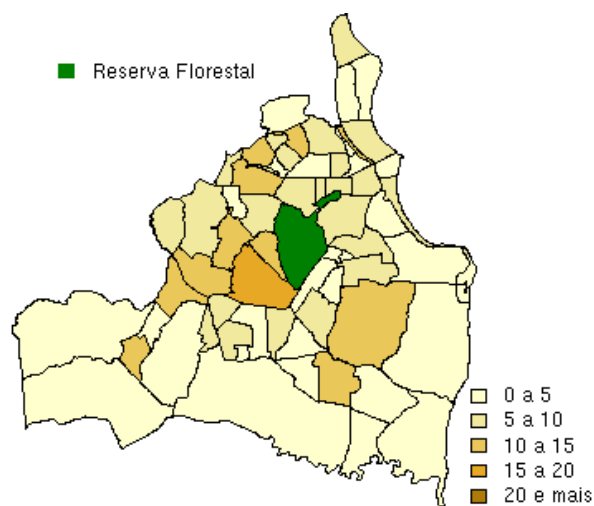


Figura 16: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2003

Nas Figuras 13 e 14 observa-se que os óbitos observados e ajustados apresentam distribuição aproximada para os bairros com valores maiores no ano 2002. Apresentam mais óbitos nos bairros de Cruz das Armas, Oitizeiro e Cristo Redentor que se localizam na região oeste da cidade. Em seguida, aparece Mangabeira que se localiza na região sul da cidade.

Nota-se nas Figuras 15 e 16 que os óbitos observados e ajustados apresentam distribuição semelhante no ano 2003, sendo os ajustes melhores para os bairros com valores de óbitos maiores. Nota-se que apresentam mais óbitos no bairro do Cristo Redentor. Em seguida, aparecem Cruz das Armas, Rangel e Jardim Veneza que se localizam na região oeste, Mangabeira e Valentina da região sul e Centro e Mandacaru da região centro-norte da cidade.

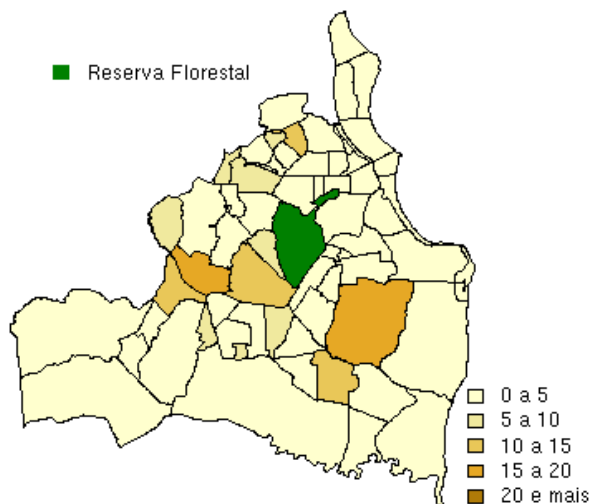


Figura 17: Óbitos por causas violentas observados em João Pessoa no ano de 2004

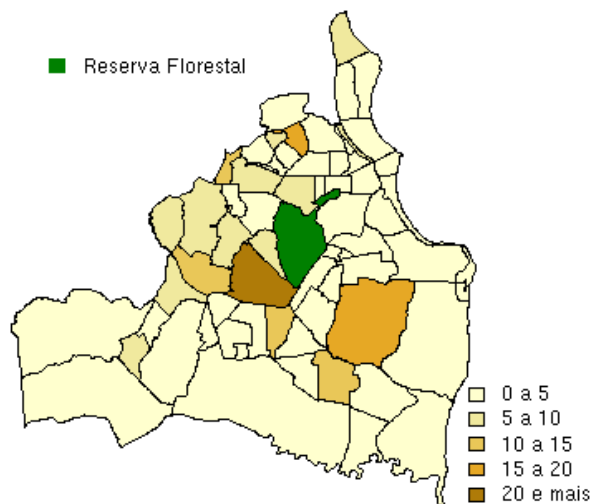


Figura 19: Óbitos por causas violentas observados em João Pessoa no ano de 2005

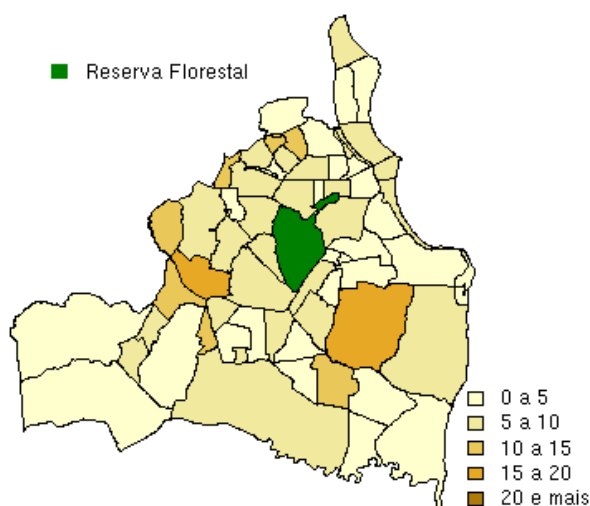


Figura 18: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2004

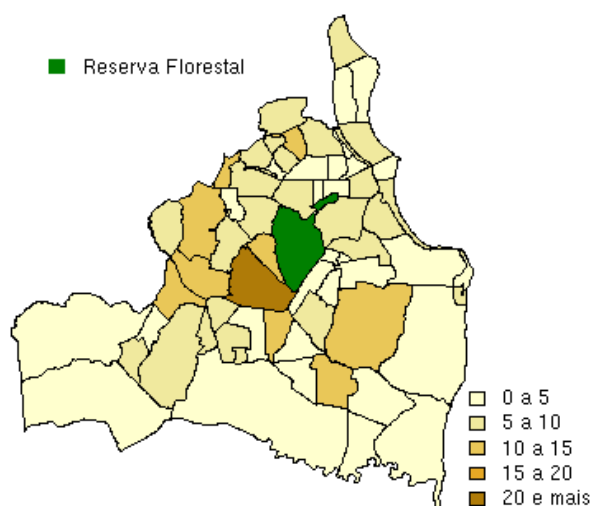


Figura 20: Óbitos por causas violentas ajustados pela Regressão de *Poisson* em João Pessoa no ano de 2005

Observa-se nas Figuras 17 e 18 que em 2004 os óbitos observados e ajustados apresentam a mesma distribuição para os bairros com valores maiores. Sendo os bairros de Oitizeiro que se localiza na região oeste da cidade e Mangabeira que se localiza na região sul os que apresentam mais óbitos neste ano.

Pode-se observar nas Figuras 19 e 20 que os óbitos observados e ajustados se distribuem semelhantemente no ano 2005 para os bairros com valores de óbitos maiores. Nota-se que apresentam mais óbitos no bairro do Cristo Redentor localizado na região oeste. Em seguida, aparecem Oitizeiro que se localizam na região oeste e Mangabeira e Valentina na região sul da cidade. Verifica-se em todos os anos, para os valores observados e ajustados, a predominância de óbitos maiores na região oeste da cidade.

Agradecimentos

Este trabalho tem suporte do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (bolsas), pelo Projeto CNPq-505258/2004-8.

Conclusões

Utilizou-se a metodologia de Regressão de *Poisson*, com o intuito de encontrar um modelo para ajustar os óbitos por causas violentas na cidade de João Pessoa no período de Janeiro de 2000 a Dezembro de 2005. Esta metodologia é apropriada quando os dados são de contagens. Os dados de óbitos foram fornecidos pela Divisão de Vigilância Epidemiológica da Secretaria Municipal de Saúde da Prefeitura Municipal de João Pessoa. Os dados de óbitos mensais foram selecionados por causas violentas, depois foram organizados por escolaridade (nenhuma, de 1 a 3 anos, de 4 a 7 anos, de 8 a 11 anos e mais de 12 anos), estado civil (solteiro, casado,

viúvo, separado judicialmente e união consensual) e raça (branca, preta, amarela e parda).

Inicialmente ajustou-se um modelo com todas as variáveis, mas as variáveis significativas foram apenas raça categoria Parda e estado civil categoria Solteiro. Dessa forma o modelo de *Poisson* obtido para o ajuste dos óbitos por causas violentas com intercepto, utilizou apenas as variáveis raça categoria Parda e estado civil categoria Solteiro. Os coeficientes estimados foram 1,84027 para o intercepto, 0,02796 para a variável estado civil categoria Solteiro e 0,04658 para a variável raça categoria Parda. Depois de obtido o modelo fez-se uma comparação dos valores observados e ajustados por bairro. Observou-se, em todos os anos, que os óbitos por causas violentas predominaram nos bairros da região oeste. Concluiu-se que este modelo exibiu bons ajustes para os bairros com contagens maiores, porém para os bairros com contagens menores os ajustes não seguem a mesma qualidade.

Referências

- [1]-AGRESTE, A., 1996. An introduction to categorical data analysis. A Wiley-Interscience Publication.
- [2]-BUSSAB, W. O.; Morettin, P. A., *Estatística básica*, 4ª edição, São Paulo: Atual, 1987.
- [3] JORGE, M. H. P. M.; GAWRYSZEWSKI, V. P.; LATORRE, M. R. D. O., 1997. I - Análise dos dados de mortalidade. *Rev. Saúde Pública*, Aug. 1997, vol.31, no.4, suplemento, p.05-25.
- [4] LOVETT, A., FLOWERDEW, R., 1989. Analysis of count data using Poisson Regression. University of Lancaster. *Professional Geographer*, vol.41, nº2, p. 190-198.
- [5]-MEYER, P. L., 1984. Probabilidade, Aplicações a Estatística. Rio de Janeiro. LTC.
- [6]-SIEGEL, S., 1975. Estatística não-Paramétrica para as ciências do comportamento, São Paulo: McGraw-Hill do Brasil.
- [7]-WAISELFISZ, J. J., UNESCO, Organização dos Estados Ibero-Americanos (OEI), Mapa da Violência 2006, Os jovens do Brasil, Sumário Executivo. (disponível em <<http://www.oei.org.br/mapa2006.pdf>>). Acesso em: 27 nov. 2006.
- [8]-YUNES, J.; RAJS, D., 1994. Tendencia de la mortalidad por causas violentas en la población general y entre los adolescentes y jóvenes de la región de las Américas. *Cadernos de Saúde Pública*, 10:45-60.