



IBP1263_06

ANÁLISE ESTATÍSTICA DO IMPACTO DA POLUIÇÃO ATMOSFÉRICA NA SAÚDE POPULACIONAL

Yara S. Tadano¹, Cássia M.L. Ugaya², Admilson T. Franco³

Copyright 2006, Instituto Brasileiro de Petróleo e Gás - IBP.

Este Trabalho Técnico foi preparado para apresentação na *Rio Oil & Expo and Conference 2006*, realizada no período de 11 a 14 de setembro de 2006, no Rio de Janeiro. Este Trabalho Técnico foi selecionado para apresentação pelo Comitê Técnico do evento, seguindo as informações contidas na sinopse submetida pelo(s) autor(es). O conteúdo do Trabalho Técnico, como apresentado, não foi revisado pelo IBP. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material, conforme apresentado, não necessariamente reflete as opiniões do Instituto Brasileiro de Petróleo e Gás, seus Associados e Representantes. É de conhecimento e aprovação do(s) autor(es) que este Trabalho Técnico seja publicado nos Anais da *Rio Oil & Expo and Conference 2006*.

Resumo

A previsão do impacto da poluição atmosférica na saúde populacional é bastante complexa, pois cada pessoa responde diferente a uma quantidade de poluentes. Há uma carência de estudos visando analisar os modelos estatísticos capazes de avaliar os efeitos adversos à saúde populacional causados pela poluição do ar. Os modelos estatísticos mais utilizados são as análises de regressão, que são úteis para avaliar a relação entre uma ou mais covariáveis e uma variável resposta. A revisão bibliográfica mostrou que o modelo mais utilizado para o problema apresentado é o de regressão de Poisson, por ser uma modelagem aplicada a dados de contagens. Existem, contudo, outros modelos estatísticos utilizados em estudos epidemiológicos, como o modelo de regressão proporcional de Cox e de regressão logística condicional. No presente estudo foi realizada uma comparação entre o modelo de regressão de Poisson e os demais modelos. Para isto, foi feito um estudo das características dos modelos propostos. Os principais resultados mostraram que o modelo de regressão logística condicional é utilizado para dados binomiais, o proporcional de Cox para dados de sobrevivência, enquanto o modelo de regressão de Poisson é utilizado para dados de contagem, caso mais freqüente em problemas de impacto da poluição na saúde populacional.

Palavras-chave: Impacto Ambiental, Análise de Regressão, Poluição Atmosférica.

Abstract

The impact prevision of atmospheric pollutants in population's health is very complex, because each person responds differently to an amount of pollutants. There is a lack of studies that analyze the statistical models capable to evaluate the adverse effect on population's health caused by air pollution. Regression analyses are the statistical models most used, because they are useful methods to assess the relation between one or more covariates and only one response variable. The bibliographic revision showed the Poisson regression model as the most used, because it is applicable to counts. There exist, however, other statistical models used in epidemiological studies, like the Cox proportional regression and the conditional logistic regression. This study aimed to carry through the comparison between the Poisson regression model and the other models. To do so, a study of the characteristics of each model was done. Main results showed that de conditional logistic regression is used to binomial data, the Cox proportional to survival data, while the Poisson regression model is used for counts, the most common case of pollution impact on population's health.

Key-words: Environmental Impact, Regression Analysis, Atmospheric Pollution.

1. Introdução

Com o surgimento das áreas urbanas, ocorreu um agravamento na qualidade do ar, causando problemas na saúde populacional. Desta forma, houve uma maior preocupação com os efeitos nocivos dos poluentes, que não é simples de ser tratado, por englobar diversos fatores; desde a dispersão, a composição química até seus efeitos na saúde da população.

¹ Mestranda do Programa de Pós-Graduação em Eng. Mecânica e de Materiais-UTFPR

² Doutora, Professora do Programa de Pós-Graduação em Eng. Mecânica e de Materiais- UTFPR

³ Doutor, Professor do Programa de Pós-Graduação em Eng. Mecânica e de Materiais – UTFPR

A dispersão dos poluentes atmosféricos em uma determinada região depende principalmente das condições meteorológicas, da topografia e das características das fontes emissoras. Estudada a dispersão dos poluentes, surge o impacto da poluição atmosférica na saúde, um assunto também complexo, pois a resposta para uma quantidade de poluente é diferente para cada pessoa, dependendo de sua faixa etária; da genética, ou seja, idosos e crianças são mais suscetíveis à poluição e algumas pessoas possuem doenças crônicas, alergias, tornando-as mais sensíveis à poluição do ar.

A quantidade de poluição a qual as pessoas estão sujeitas varia significativamente conforme o local onde elas trabalham e moram; a frequência com que mudam de residência ou cidade, entre outros fatores que também influenciam na quantidade e qualidade do ar a qual cada ser humano está sujeito.

Devido à grande evolução dos métodos estatísticos, atualmente, é possível unir uma vasta quantidade de fatores em um mesmo problema, o que permite uma melhor avaliação dos impactos da poluição atmosférica na saúde populacional.

A partir do desenvolvimento computacional, na década de 70, alguns modelos que exigiam a utilização de processos iterativos na estimativa dos parâmetros começaram a ser utilizados e aperfeiçoados.

Assim, Nelder e Wedderburn (1972), propuseram uma teoria unificadora da modelagem estatística a qual denominaram de modelos lineares generalizados (MLG), como sendo uma extensão dos modelos lineares clássicos. Na realidade, eles mostraram que uma série de técnicas, comumente estudadas separadamente há mais de um século, podem ser reunidas.

Tabela 1 - Histórico dos estudos que levaram à modelagem estatística denominada MLG

Ano	Autor	Estudo realizado	Tipo de distribuição envolvida
1920 a 1935	Fisher	Análise de variância para experimentos planejados	Distribuição binomial com função de ligação logit
1922	Fisher	Procedimento geral para inferência estatística	Qualquer distribuição
1922	Fisher	Modelo complemento log-log para ensaios de diluição	Distribuição binomial
1935	Bliss	Modelo probit para proporções	Distribuição binomial
1952	Dike e Patterson	Modelos logísticos para proporções	Distribuição binomial com função de ligação logit
1960	Rasch	Modelo logístico para análise de itens	Distribuição Bernoulli
1963	Birch	Modelos log-lineares para contagens	Distribuição de Poisson e multinomial
1966	Nelder	Polinômios inversos	Distribuição gama
1965	Feigl e Selen	Modelos para dados de sobrevivência	Distribuição exponencial com função de ligação recíproca ou log
1966	Zippin e Armitage		
1967	Glasser		

Fonte: McCullagh e Nelder (1989).

Tem-se na Tabela 1, um breve histórico dos estudos que levaram a essa visão geral da modelagem estatística denominada modelos lineares generalizados (MLG).

Segundo McCullagh e Nelder (1989), a regressão linear múltipla, envolvendo distribuição normal, iniciou-se no século XIX com os matemáticos Legendre e Gauss.

Além de todos os estudos realizados por Fisher, ele ainda desenvolveu, em 1934, a família exponencial de distribuições, como sendo uma classe de distribuições com propriedades ótimas para a estimação dos parâmetros.

O Trabalho de Dike e Patterson, citado na Tabela 1, foi realizado através de uma análise de dados cruzados relacionando a proporção de pessoas que possuem um bom conhecimento sobre câncer.

O trabalho de Birch é usado freqüentemente nas ciências sociais e o de Nelder, em trabalhos biológicos para expressar a forma das curvas de resposta, o crescimento das curvas e assim por diante. Por sua vez, os trabalhos de Feigl e Selen; Zippin e Armitage; e Glasser, são usados geralmente para tratamentos clínicos e cirúrgicos, por serem modelos para análise de dados de sobrevivência.

O objetivo principal deste trabalho é comparar os modelos estatísticos mais difundidos, capazes de avaliar o impacto da poluição atmosférica na saúde populacional, com o intuito de identificar os casos nos quais cada um pode ser adequadamente empregado, pois apesar de existirem diversos trabalhos que utilizam modelos estatísticos na análise dos dados, geralmente não expõem de forma detalhada o desenvolvimento dos modelos e nem apresentam uma comparação entre eles.

Podem ser citados alguns trabalhos que utilizam os modelos, mas não os abordam detalhadamente, como é o caso dos trabalhos de Neas et al. (1999), Lee e Schwartz (1999), Lee et al. (2000) e Peters et al. (2001), que utilizam o modelo de regressão logística; os trabalhos de Dockery et al. (1993), Abbey et al. (1999), Pope et al. (2002) e Barroso et al. (2003); que utilizam o modelo de regressão proporcional de Cox e os trabalhos de Braga et al. (2001), Martins et al. (2002), Kim et al. (2003), Freitas et al. (2004), Bakonyi et al. (2004), Ha et al. (2005), entre outros; que utilizam o modelo de regressão de Poisson.

Para se entender melhor todas as etapas envolvidas na escolha do modelo adequado aos dados, este trabalho está organizado da seguinte forma. Na Seção 2, são mostradas as características dos modelos lineares generalizados, uma extensão dos modelos lineares clássicos, envolvendo diversas classes de modelos da família exponencial, bem como algumas decisões que devem ser tomadas na escolha do modelo apropriado para a análise dos dados disponíveis. A Seção 3 mostra as diferenças entre os modelos comumente utilizados na análise de problemas que envolvam o impacto da poluição na saúde populacional, destacando a facilidade em se utilizar o modelo de regressão de Poisson devido à sua generalização. Finalmente, na Seção 4, é abordado o modelo de regressão de Poisson, por ser o mais adequado na análise do impacto da poluição atmosférica na saúde populacional, pela “simplicidade” de sua aplicação.

2. Modelos Lineares Generalizados

Nelder e Wedderburn (1972) desenvolveram uma classe de modelos baseados na família exponencial com um parâmetro desconhecido, cujas médias são não-lineares num conjunto de parâmetros lineares. Tais modelos são chamados de modelos lineares generalizados (MLG).

Os MLG são uma união de modelos lineares e não-lineares onde se tem uma distribuição da flexível família exponencial, formada pela distribuição normal, Poisson, binomial, gama, e normal inversa; e incluem modelos lineares tradicionais (erros com distribuição normal), bem como os modelos logísticos, os modelos log-lineares, os modelos de regressão de Poisson, dentre outros (McCullagh e Nelder, 1989).

Os modelos lineares generalizados possibilitam, basicamente, ampliar a distribuição da variável resposta, permitindo que esta variável pertença à família exponencial de distribuições, bem como para dar maior flexibilidade à relação funcional entre a média da variável resposta (μ) e o preditor linear (η). Assim, para dados de contagem, em vez de aplicar-se algum tipo de transformação na variável resposta y com o intuito de normalizá-la, pode-se, por exemplo, usar a distribuição de Poisson e a relação funcional entre a média de y e o preditor linear $\eta = \log(\mu)$. Essa relação garante um valor positivo para quaisquer valores dos parâmetros do preditor linear (Santos et al., 2000).

Muitos outros modelos estatísticos podem ser gerados pela seleção apropriada de uma distribuição de probabilidade para a variável resposta e de uma função de ligação. O modelo linear generalizado é definido por uma distribuição de probabilidade, membro da família exponencial de distribuições, que se constitui pelas seguintes componentes:

- A componente aleatória: admite-se um vetor de observações $y = (y_1, \dots, y_p)^T$ como realização das variáveis aleatórias $Y = (Y_1, \dots, Y_p)^T$, as chamadas variáveis resposta, que seguem alguma distribuição de probabilidade da família exponencial, que tem densidade dada por:

$$f_Y(y; \theta, \varphi) = \exp \left\{ (y\theta - b(\theta)) / a(\varphi) + c(y, \varphi) \right\} \quad (1)$$

onde $a(\cdot)$, $b(\cdot)$ e $c(\cdot)$ são funções conhecidas e θ é o parâmetro natural (ou canônico) e $\varphi > 0$ é chamado de parâmetro de escala e é suposto conhecido para cada observação.

- A componente sistemática: considera-se a estrutura linear de um modelo de regressão $\eta = X^T \beta$, onde $X^T = (x_{i1}, \dots, x_{ip})^T$ são as variáveis explicativas e X é uma matriz modelo;
- A função de ligação: liga as componentes aleatórias e as sistemáticas, ou seja, liga a média da variável resposta e a estrutura linear; é uma função monótona e diferenciável, $g(\cdot)$, definida nos MLG por $g(i) = \eta$, onde
 - $\eta = X^T \beta$ é o preditor linear;
 - $\beta = (\beta_1, \dots, \beta_p)^T$, $p < n$, é um vetor de parâmetros desconhecidos a serem estimados.

Na Tabela 2, são apresentadas as principais regressões e seus respectivos tipos de distribuição da variável resposta, bem como a forma da função de ligação.

Tabela 2 - Característica de algumas regressões comumente utilizadas.

Regressão	Distribuição	Função de ligação
Linear	Normal	$g(\mu) = \mu$
Logística	Binomial	$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$
Poisson	Poisson	$g(\mu) = \log(\mu)$

2.1. Escolha do Modelo

Uma decisão importante na escolha do modelo linear generalizado (MLG) adequado é quanto ao trinômio especificado anteriormente: componente aleatória, componente sistemática e função de ligação.

Outro aspecto importante na escolha do modelo se refere à escolha da escala de análise. Uma escolha comum é entre uma análise da variável resposta Y (escala original), ou sua transformação $\log(Y)$.

Segundo McCullagh e Nelder (1989), a caracterização de uma “boa” escala depende do propósito do estudo, por exemplo, para uma análise de regressão linear clássica, uma boa escala deveria combinar uma variância constante, erros com distribuição aproximadamente normal e adição dos efeitos sistemáticos. Porém, não existe nenhuma razão para acreditar que tal escala exista. Para exemplificar um caso em que isto não ocorre, tem-se a análise de dados discretos, onde os erros são bem aproximados por uma distribuição de Poisson, por exemplo, em que os efeitos sistemáticos, geralmente, são multiplicativos. Neste caso, $Y^{1/2}$ dá uma variância aproximadamente constante, $Y^{2/3}$ tem melhor simetria aproximada ou distribuição normal dos erros, e $\log(Y)$ produz adição dos efeitos sistemáticos. Evidentemente, uma escala sozinha não vai produzir simultaneamente todas as propriedades desejadas. Porém, quando se introduz os MLG, as propriedades, acima citadas, para o erro e a variância não são mais necessárias, embora o modo como a variância depende da média deve ser conhecido.

Após a escolha do modelo que melhor se aproxima dos dados, não se pode esquecer de verificar se o modelo realmente se ajusta aos dados disponíveis, usando, por exemplo, análise de resíduos, ou outras estatísticas.

3. Comparação entre os Modelos

Um dos fatores relevantes na escolha de um modelo é o tipo de variável resposta, comumente chamada de variável dependente. Os modelos a serem comparados aqui possuem variáveis respostas diferentes.

O modelo de regressão logística é utilizado para problemas em que a variável resposta é do tipo binomial, onde se atribui apenas valores 0 ou 1, representando fracasso ou sucesso na análise. Este tipo de observação é conveniente para experiências médicas, onde no final do período de estudo, ou o paciente se recuperou ($Y=1$) ou não se recuperou ($Y=0$). Este tipo de modelo é interessante quando se quer estudar o impacto da poluição em indivíduos específicos, não dando uma visão do impacto geral da poluição em uma determinada região e não é um modelo simples de ser utilizado, pois requer o acompanhamento de pacientes que possuem o diagnóstico no qual se está interessado, acompanhando suas melhoras e piores, porém, eles fornecem resultados praticamente reais da situação, mas apenas para um grupo de pessoas em especial, já a extensão do resultado para outros grupos implicará em erros muito provavelmente maiores que com a aplicação da regressão de Poisson.

Como exemplo de utilização deste modelo pode se citar os trabalhos de Neas et al. (1999), Lee e Schwartz (1999), Lee et al. (2000) e Peters et al. (2001).

Outro tipo de variável que não fornece uma visão geral é a chamada de variável de sobrevivência, correspondendo ao tempo de vida de um componente, ou tempo de sobrevivência de um paciente, também muito utilizada para experiências médicas, ou ainda útil para experimentos de segurança industrial onde a ênfase é dada para o tempo de falha de um equipamento ao invés do tempo de sobrevivência. No caso de experiências médicas, é necessário ter um acompanhamento do paciente durante o período em estudo, o que dificulta o emprego do modelo. O modelo de regressão proporcional de Cox, utilizado nos trabalhos realizados por Dockery et al. (1993), Abbey et al. (1999), Pope et al. (2002) e Barroso et al. (2003), tem como variável resposta dados de sobrevivência.

Já o modelo de regressão de Poisson, muito comum em estudos epidemiológicos, tem por característica a análise de dados contados na forma de proporções ou razões de contagem, ou seja, leva em consideração o total de pessoas que receberam um certo diagnóstico, dados facilmente adquiridos através da secretaria de saúde da região em questão. Quando se aplica este modelo, se está interessado no impacto geral de um certo componente na saúde da população de uma determinada região, para desta forma, iniciar um controle da dispersão deste componente.

Exemplos de trabalhos que utilizaram este modelo são os de Braga et al. (2001), Martins et al. (2002), Kim et al. (2003), Freitas et al. (2004), Bakonyi et al. (2004) e Ha et al. (2005).

Assim, devido à facilidade na utilização deste modelo e na obtenção dos dados necessários, verifica-se que para o problema em questão, o modelo de regressão de Poisson é mais adequado.

4. Modelo de Regressão de Poisson

O modelo de regressão de Poisson é um tipo específico dos modelos lineares generalizados (MLG) e não-lineares, que teve origem por volta de 1970, quando Wedderburn (1974) desenvolveu a teoria da quase-verossimilhança, analisada posteriormente com mais detalhes por McCullagh (1983). Esta teoria especifica apenas uma relação entre a variância e a média da variável resposta, a chamada função de variância.

Conforme Cordeiro (1986), as principais características do modelo de regressão de Poisson são:

- Proporciona, em geral, uma descrição satisfatória de dados experimentais cuja variância é proporcional à média;
- Pode ser deduzido teoricamente de princípios elementares com um mínimo de restrição;
- Se eventos ocorrem independente e aleatoriamente no tempo, com taxa média de ocorrência constante, o modelo determina o número de eventos, num intervalo de tempo especificado.

Na análise de dados categorizados, o modelo de Poisson desempenha, o mesmo papel que o modelo normal na análise de dados contínuos. A principal diferença é que a estrutura multiplicativa para as médias do modelo de Poisson é mais apropriada do que a estrutura aditiva das médias do modelo normal (Schmidt, 2003).

Para o caso do impacto da poluição atmosférica na saúde populacional, a regressão de Poisson tem como trinômio:

- Componente Aleatória: variável resposta como sendo o número de agravos à saúde da população na forma de proporções ou contagens;
- Componente Sistemática: como variável explicativa, pode-se considerar a concentração de um determinado poluente, a temperatura, umidade, entre outros;
- Função de ligação: $\eta = \log(\mu)$.

Como foi dito, após a escolha do modelo a ser utilizado na análise, é preciso verificar a validade deste modelo. Segundo McCullagh e Nelder (1989), no caso do modelo de regressão de Poisson, a estatística utilizada para esta validação é a estatística G^2 , proposta por Bishop, Fienberg e Holland (1975) e outros, que é dada por:

$$2 \sum \{y \log(y/\hat{\mu}) - (y - \hat{\mu})\} \quad (2)$$

onde o segundo termo é freqüentemente omitido. Outra estatística também utilizada para fazer a validação do modelo de Poisson é a chamada estatística de Pearson X^2 , que tem a forma:

$$X^2 = \sum (y - \hat{\mu})^2 / V(\hat{\mu}) \quad (3)$$

onde $V(\hat{\mu})$ é a função variância (McCullagh e Nelder, 1989).

Como o próprio nome já diz, o modelo de regressão de Poisson é um modelo linear generalizado que possuiu a distribuição de Poisson da família exponencial.

4.1. Distribuição de Poisson

A distribuição de Poisson pode ser usada para determinar a probabilidade de um dado número de sucessos quando os eventos ocorrem em um “continuum” de tempo ou espaço. Supõe-se que os eventos são independentes e que o processo é estacionário. A probabilidade de se obter mais um evento num intervalo pequeno é desprezível. Esta distribuição é largamente empregada quando se deseja contar o número de eventos de certo tipo que ocorrem num intervalo de tempo, superfície ou volume. Por exemplo: o número de chamadas telefônicas recebidas durante cinco minutos; número de falhas de um computador num dia de operação, bem como, o número de pessoas com problemas respiratórios durante um ano (Schmidt, 2003).

A distribuição de Poisson é conhecida como distribuição dos pequenos valores, pois o valor da probabilidade (p) é muito pequeno, enquanto o número de observações (n) é muito grande, assim o produto np tenderá para um limite μ . Neste caso tem-se:

$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!}; \quad y = 0, 1, 2, \dots \quad (4)$$

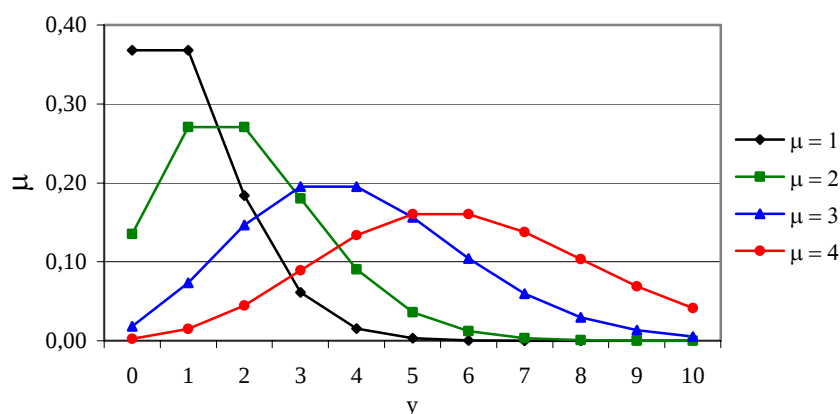
A distribuição que tem como densidade de probabilidade a expressão da Equação 1 é denominada de distribuição de Poisson, com parâmetro $\mu > 0$.

Uma propriedade interessante que uma variável aleatória de Poisson apresenta é: seu valor esperado é igual a sua variância que são iguais a μ , conhecido como parâmetro de Poisson (Meyer, 1983).

Para ilustrar que a distribuição de Poisson é uma distribuição de pequenos valores, tem-se a representação das probabilidades da distribuição de Poisson para $\mu = 1, 2, 4, 6$ na Tabela 3 e na Figura 1.

Tabela 3 - Valores da distribuição de Poisson

y	$\mu = 1$	$\mu = 2$	$\mu = 4$	$\mu = 6$
0	0,367879441	0,135335283	0,018315639	0,002478752
1	0,367879441	0,270670566	0,073262556	0,014872513
2	0,183939721	0,270670566	0,146525111	0,044617539
3	0,06131324	0,180447044	0,195366815	0,089235078
4	0,01532831	0,090223522	0,195366815	0,133852618
5	0,003065662	0,036089409	0,156293452	0,160623141
6	0,000510944	0,012029803	0,104195635	0,160623141
7	0,000072992	0,003437087	0,059540363	0,137676978
8	0,000009123	0,000859272	0,029770181	0,103257734
9	0,000001013	0,000190949	0,013231192	0,068838489
10	0,000000101	0,000038189	0,005292477	0,041303093



Fonte: Schmidt (2003).

Figura 1 – Gráfico da distribuição de Poisson para quatro valores de μ .

Observa-se, pela Figura 1, que com o crescimento do valor de μ , há um achatamento da curva e o seu deslocamento para a direita. Para valores grandes de μ , a distribuição é aproximadamente normal.

Uma distribuição possui igual dispersão, quando a média é igual à variância, porém, quando a variância é maior que a média, há uma superdispersão e quando a variância é menor que a média, há uma subdispersão, sendo este último caso incomum.

Assim, o modelo de Poisson é dado pela Equação 4, onde y é o número de não-conformidades e $\mu > 0$ é o parâmetro da distribuição de Poisson e é, também, a média e a variância. (Schmidt, 2003).

5. Conclusão

Em estudos epidemiológicos, como é o caso do impacto da poluição atmosférica na saúde populacional existem formas diferentes de análise dependendo do tipo de variável resposta que se deseja. Pode ser feita uma avaliação do impacto de indivíduo para indivíduo utilizando tanto variáveis binomiais, onde se utiliza o modelo de regressão logística condicional; como variáveis de sobrevivência, onde se utiliza o modelo de regressão proporcional de Cox. Porém o modelo mais conhecido e utilizado é o modelo de regressão de Poisson, por ser um modelo que leva em consideração dados contados na forma de proporções ou razões de contagem, sendo possível então ter uma visão geral do problema. Por exemplo, saber o impacto da poluição atmosférica em uma região específica, bem como impactos globais, para que então possa se verificar a qualidade do ar em regiões de grande concentração de poluentes, podendo posteriormente ser tomadas medidas de prevenção ou de manutenção da qualidade do ar.

6. Agradecimentos

Este trabalho foi desenvolvido com o apoio financeiro da Agência Nacional do Petróleo – ANP – e da Financiadora de Estudos e Projetos – FINEP – por meio do Programa de Recursos Humanos da ANP para o Setor Petróleo e Gás – PRH-ANP/MCT (PRH10-UTFPR).

7. Referências

- ABBEY, D.E.; NISHINO, N.; McDONNELL, W.F.; BURCHETTE, R.J.; KNYTSEN, S.F.; BEESON, W.L.; YANG, J.X., Long-term inhalable particles and other air pollutants related to mortality in nonsmokers, *American Journal of Respiratory and Critical Care Medicine*, v. 159, p. 373-382, 1999.
- BAKONYI, S.M.C.; DANNI-OLIVEIRA, I.M.; MARTINS, L.C.; BRAGA, A.L.F., Poluição atmosférica e doenças respiratórias em crianças na cidade de Curitiba, PR, *Revista Saúde Pública*, v. 38, n. 5, p. 695-700, 2004.
- BARROSO, E.C.; MOTA, R.M.S.; MORAIS, M.F.M.; CAMPELO, C.L.; BARROSO, J.B.; RODRIGUES, J.L.N., Fatores associados aos tratamentos inadequados em grupo de portadores de tuberculose multiresistente, *Jornal de Pneumologia*, v. 29, n. 6, p. 350-257, 2003.
- BIRCH, M.W., Maximum likelihood in three-way contingency tables, *Journal of the Royal Statistical Society B*, v. 25, p. 220-233, 1963.
- BRAGA, A.L.F.; SALDIVA, P.H.N.; PEREIRA, L.A.A.; MENEZES, J.J.C.; CONCEIÇÃO, G.M.S.; LIN, C.A.; ZANOBETTI, A.; SCHWARTZ, J.; DOCKERY, D.W., Health effects of air pollution exposure on children and adolescents in São Paulo, Brazil, *Pediatric Pulmonology*, v. 31, p. 106-113, 2001.
- CORDEIRO, G. M., Livro Texto do VII SINAPE – Simpósio Nacional de Probabilidade e Estatística. Campinas, SP: UFPE/ABE, 1986.
- DOCKERY, D.W.; POPE III, C.A.; XU, X.; SPENGLER, J.D.; WARE, J.H.; FAY, M.E.; FERRIS, B.G. Jr.; SPEIZER, F.E., An association between air pollution and mortality in six U.S. cities, *New England Journal of Medicine*, v. 29, n. 24, p. 1753-1759, 1993.
- FREITAS, C.; BREMNER, S.A.; GOUVEIA, N.; PEREIRA, L.A.A.; SALDIVA, P.H.N., Internações e óbitos e sua relação com a poluição atmosférica em São Paulo, 1993-1997, *Revista Saúde Pública*, v. 38, n. 6, p. 751-757, 2004.
- HA, E-H.; LEE, J-T.; KIM, H.; HONG, Y-C.; LEE, B-E.; PARK H-S., Infant susceptibility of mortality to air pollution in Seoul, South Korea, *Pediatrics*, v. 111, n. 2, p. 283-290, 2005.
- KIM, H.; KIM, Y.; HONG, Y., The lag-effect pattern in the relationship of particulate air pollution to daily mortality in Seoul, Korea, *International Journal of Biometeorology*, v. 48, p. 25-30, 2003.

- LEE, J.T.; KIM, H.; SCHWARTZ, J., Bidirectional case-crossover studies of air pollution: bias from skewed and incomplete waves, *Environmental Health Perspectives*, v. 108, n. 12, p. 1107-1111, 2000.
- LEE, J.T.; SCHWARTZ, J., Reanalysis of the effects of air pollution on daily mortality in Seoul, Korea: A case-crossover design, *Environmental Health Perspectives*, v. 107, n. 8, p. 633-636, 1999.
- MARTINS, L.C.; LATORRE, M.R.D.O.; CARDOSO, M.R.A.; GONCALVES, F.L.T.; SALDIVA, P.H.N.; BRAGA, A.L.F., Poluição atmosférica e atendimento por pneumonia e gripe em São Paulo, Brasil, *Revista Saúde Pública*, v. 36, n. 1, p. 88-94, 2002.
- McCULLAGH, P.; NELDER, J. A., *Generalized Linear Models*, London: Chapman and Hall, Second Edition, 1989.
- McCULLAGH, P., Quasi-likelihood functions, *Annals of Statistics*, v. 11, p. 59-67, 1983.
- MEYER, P. L., *Probabilidade – Aplicações à Estatística*, 2ª Ed. Rio de Janeiro: Livros Técnicos e Científicos Editora, 1983.
- NEAS, L.M.; SCHWARTZ, J.; DOCKERY, D.A., Case-crossover analysis of air pollution and mortality in Philadelphia, *Environmental Health Perspectives*, v. 107, n. 8, p. 629-631, 1999.
- NELDER, J. A.; WEDDERBURN, R. W. M., Generalized Linear Models, *Journal of the Royal Statistical Society A*, v. 135, n. 2, p. 370-384, 1972.
- POPE III, C.A.; BURNETT, R.T.; THUN, M.J.; CALLE, E.E.; KREWSKI, D.; ITO, K.; THURSTON, G.D., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, *Journal of the American Medical Association*, v. 287, n. i9, p. 1132-1142, 2002.
- SANTOS, C.A.S.T.; FERREIRA, L.D.A.; OLIVEIRA, N.F.; DOURADO, M.I.C.; BARRETO, M.L., Modelagem multinível, *Feira de Santana*, n. 22, p. 89-98, 2000. Disponível em: http://www.uefs.br/sitientibus/exatas-22/modelagem_multinivel.pdf. Acesso em: 20 de Maio de 2006.
- SCHMIDT, C.M.C., Modelo de regressão de Poisson aplicado à área da saúde, Dissertação (Mestrado em Modelagem Matemática), Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí – RS, 2003.
- WEDDERBURN, R.W.M., Quasilikelihood functions, generalized linear models and the Gauss-Newton method, *Biometrika*, v. 61, p. 439-447, 1974.