

ANÁLISE DE CORRELAÇÃO CANÔNICA APLICADA AO FLUXO DE TRÁFEGO VEICULAR - ESTUDO DE CASO DA RODOVIA FEDERAL BR-116

BOGO R. L.^c, GRAMANI L. M.^a, NETO A. C.^b e BALBO F. A. N.^d

^a*Departamento de Matemática, Universidade Federal do Paraná, Centro Politécnico, Curitiba, Paraná, Brasil. <http://www.mat.ufpr.br/>*

^b*Departamento de Estatística, Universidade Federal do Paraná, Centro Politécnico, Curitiba, Paraná, Brasil. <http://www.est.ufpr.br/>*

^c*Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Centro Politécnico, Curitiba, Paraná, Brasil. <http://www.ppgmne.ufpr.br/>*

Palavras Chave: Tráfego de Veículos, Análise de Correlação Canônica, BR-116.

Resumo. O tráfego de veículos nas rodovias federais representa um fenômeno de grande importância sócio-econômica, principalmente, por ser o principal meio de transporte utilizado pela sociedade para se deslocar de uma cidade para outra. Também representa a principal alternativa para o transporte de produtos comercializados pela indústria. Com o crescente aumento do número de veículos nas estradas e a deteriorização da malha asfáltica em grande parte das rodovias federais, a falta de segurança tornou-se um problema evidente, elevando assim, o número de acidentes ocorridos. A partir de um banco de dados fornecido pela Polícia Rodoviária Federal do Paraná referente aos acidentes ocorridos no ano de 2009 ao longo da rodovia federal BR-116, trecho compreendido no estado do Paraná (aproximadamente 211km), realizou-se um estudo para desenvolver métodos que possam identificar o perfil dos acidentes nesse trecho da rodovia. A técnica estatística usada neste trabalho é denominada análise de correlação canônica, onde verificou-se as inter-relações entre dois grupos de variáveis (tipos de acidentes x outras características). As correlações canônicas obtidas entre dois grupos indicaram uma forte correlação entre os mesmos. Desta forma, foi possível determinar o perfil dos acidentes em alguns trechos da rodovia com possíveis problemas, para que possa ser feito um planejamento para ser resolver esse problemas na tentativa de reduzir o grande número de acidentes.

1 INTRODUÇÃO

O tráfego de veículos nas rodovias federais representa um fenômeno de grande importância sócio-econômica, principalmente, por ser o principal meio de transporte utilizado pela sociedade para se deslocar de uma cidade para outra. Também representa a principal alternativa para o transporte de produtos comercializados pela indústria. O rápido crescimento do número de veículos nas rodovias federais do Brasil vem causando diversos problemas à sociedade, e um aumento considerável no número de acidentes. Dentre estas rodovias destaca-se a rodovia BR-116, umas das principais país.

A BR-116 (ver Figura 1) é uma rodovia longitudinal que tem início em Fortaleza, estado do Ceará, e término na cidade de Jaguarão, estado do Rio Grande do Sul. A extensão total é de aproximadamente 4.385 quilômetros passando por diversos estados brasileiros, entre eles, o estado do Paraná. No Paraná, a rodovia é conhecida também como rodovia Régis Bittencourt¹. Neste estado, a rodovia tem início na divisa com o estado de São Paulo atravessando toda a cidade Curitiba (capital paranaense) e terminando na cidade de Rio Negro. A extensão total no estado do Paraná é de aproximadamente 211 quilômetros. Por ser um dos meios de ligação norte-sul mais importante do país, o aumento do número de veículos e, como consequência, o aumento dos acidentes atinge números preocupantes.



Figura 1: BR-116

O presente trabalho teve como objetivo principal determinar o perfil dos acidentes ocorridos nesta rodovia a partir de um banco de dados fornecido pela Polícia Rodoviária Federal do Paraná (PRF), referente os acidentes ocorridos na BR-116 no ano de 2009. Para isso, utilizamos a técnica estatística multivariada denominada análise de correlação canônica (Johnson e Wichern, 2002).

A análise de correlação canônica tem por objetivo principal verificar a associação entre dois grupos de variáveis, podendo assim, relacionar os tipos de acidentes com determinadas características da pista. É a técnica que apresentou melhores resultados, uma vez que a análise de agrupamentos (Mingoti, 2007) não apresentou resultados positivos e na análise fatorial Hair e Anderson (2002) o conjunto de dados não foi adequado para aplicação da mesma.

¹Edmundo Régis Bittencourt foi um engenheiro civil brasileiro, empreiteiro e ex-presidente do Departamento Nacional de Estradas de Rodagem (DNER). Em sua homenagem, a rodovia BR-116 no trecho entre São Paulo e Curitiba, foi designada Régis Bittencourt.

Na literatura existente, não foram encontrados trabalhos referentes à acidentes que trânsito que utilizam a análise de correlação canônica. Os trabalhos de alguns autores que utilizaram esta técnica na área da química e administração, podem ser encontrados em [Trugilho et al. \(2003\)](#) e [Costa \(2008\)](#), respectivamente.

A estrutura deste trabalho está dividido em mais quatro seções envolvendo a descrição matemática da análise de correlação canônica, a apresentação dos materiais e métodos usados no problema, a discussão dos resultados obtidos e a conclusão. Em detalhes:

Na Seção 2 será feito o desenvolvimento matemático da análise de correlação canônica onde explica-se os testes que verificam se a aplicação do método é adequada ao conjunto de dados.

Na Seção 3 será descrito todo o processo de análise de dados, tais como, a composição da amostra utilizada, os tipos de variáveis envolvidas, a maneira como as técnicas e testes multivariados foram aplicados.

Na Seção 4 será discutido os resultados obtidos com a análise de correlação canônica, onde apresentaremos as correlações canônicas obtidas entre os conjuntos de variáveis, os testes de validação da análise de correlação canônica que irão auxiliar na interpretação e conclusão na Seção 5.

2 DESENVOLVIMENTO DO TEMA

Nesta seção será apresentado a teoria matemática da análise de correlação canônica, bem como o teste de hipótese para verificação da estrutura da correlação da matriz de dados.

2.1 Análise de Correlação Canônica

A análise de correlação canônica é uma técnica para a identificação e quantificação da associação entre dois grupos de variáveis. Conforme [Johnson e Wichern \(2002\)](#), o objetivo desta técnica é determinar os pares de combinações lineares $U = a'X$ e $V = b'Y$ tais que tenham a maior correlação possível. Estas correlações podem estabelecer diferenças convenientes entre os dois conjuntos de variáveis.

Segundo [Lira \(2004\)](#), a análise de correlação canônica é considerada uma extensão da análise de regressão múltipla. Na análise de regressão múltipla, as variáveis formam o conjunto das covariáveis X ($p - 1$ variáveis independentes) e a variável resposta Y (variável única dependente). Assim, o problema na análise de regressão consiste em buscar a combinação linear $\beta'X$ que é altamente correlacionada com a variável Y . Por outro lado, na análise de correlação canônica tem-se o conjunto X com $p \geq 1$ variáveis, portanto, devendo-se achar os vetores a' e b' para os quais a correlação entre $U = a'X$ e $V = b'Y$ seja máxima.

2.2 Variáveis Canônicas e Correlações Canônicas

Conforme [Neto \(2002b\)](#), tem-se interesse em medir a associação entre 2 (dois) grupos de variáveis. Seja o primeiro grupo, de p variáveis, representado pelo vetor aleatório $X(p \times 1)$. O segundo grupo, de q variáveis, é representado pelo vetor aleatório $Y(q \times 1)$ considerando $p \leq q$.

Para os vetores aleatórios X e Y , tem-se:

$$\begin{aligned} E(X) &= \mu_1 & \text{Cov}(X) &= \Sigma_{11} \\ E(Y) &= \mu_2 & \text{Cov}(Y) &= \Sigma_{22} \\ \text{Cov}(X, Y) &= \Sigma_{12} = \Sigma'_{21} \end{aligned} \quad (1)$$

Considerando X e Y conjuntamente, encontramos o vetor aleatório

$$\underbrace{X}_{((p+q) \times 1)} = \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \\ Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{bmatrix} \quad (2)$$

com o vetor de médias

$$\underbrace{\mu}_{((p+q) \times 1)} = E(X) = \begin{bmatrix} E(X) \\ E(Y) \end{bmatrix} = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix} \quad (3)$$

e matriz de covariância

$$\begin{aligned} \underbrace{\Sigma}_{(p+q) \times (p+q)} &= E(X - \mu)(X - \mu)' = \\ &= \begin{bmatrix} E(X - \mu^{(1)})(X - \mu^{(1)})' & E(X - \mu^{(1)})(Y - \mu^{(2)})' \\ E(Y - \mu^{(2)})(X - \mu^{(1)})' & E(Y - \mu^{(2)})(Y - \mu^{(2)})' \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \end{aligned} \quad (4)$$

Assim, as covariâncias entre os pares de variáveis dos diferentes conjuntos - uma variável de X e outra de Y - estão contidas em Σ_{12} ou, equivalentemente, em Σ_{21} . Isto é, os pq elementos medem a associação entre os dois conjuntos de variáveis. Quando p e q são relativamente grandes, a interpretação dos elementos de Σ_{12} torna-se impraticável. Em consequência disso, as combinações lineares são introduzidas para tornar possível a interpretação desejada.

Seja as combinações lineares:

$$\begin{aligned} U &= a'X \\ V &= b'Y \end{aligned} \quad (5)$$

então

$$\text{Corr}(U, V) = \frac{\text{Cov}(U, V)}{\sqrt{V(U)V(V)}} = \frac{E[(U - \bar{U})(V - \bar{V})]}{\sqrt{V(U)V(V)}} \quad (6)$$

onde:

$$\begin{aligned} \text{Var}(U) &= a'\text{Cov}(X)a = a'\Sigma_{11}a \\ \text{Var}(V) &= b'\text{Cov}(Y)b = b'\Sigma_{22}b \\ \text{Cov}(U, V) &= a'\text{Cov}(X, Y)b = a'\Sigma_{12}b \end{aligned} \quad (7)$$

Portanto, devemos encontrar os vetores coeficientes a e b tais que:

$$\text{Corr}(U, V) = \frac{a'\Sigma_{12}b}{\sqrt{a'\Sigma_{11}a}\sqrt{b'\Sigma_{22}b}} \quad (8)$$

seja a maior possível.

O primeiro par de variáveis canônicas são as combinações lineares U_1 e V_1 , com variância unitária, e que maximiza a correlação 8.

O segundo par de variáveis canônicas são as combinações lineares U_2 e V_2 , com variância unitária, e que maximiza a correlação 8 dentre aquelas que não são correlacionadas com o primeiro par de variáveis canônicas.

Na k -ésima etapa, o k -ésimo par de variáveis de canônicas são as combinações lineares U_k e V_k , com variância unitária, e que maximiza a correlação 8 entre aquelas que não são correlacionadas com os $k-1$ primeiros pares de variáveis canônicas definidos.

A correlação entre o k -ésimo par de variáveis canônicas é denominada por k -ésima correlação canônica.

O resultado a seguir fornece detalhes de como as variáveis canônicas e suas correlações são obtidas.

Resultado 1. Suponhamos que $p \leq q$ e seja os vetores aleatórios X e Y com $Cov(X) = \Sigma_{11}$, $Cov(Y) = \Sigma_{22}$ e $Cov(X, Y) = \Sigma_{12}$. Sejam as combinações lineares $U = a'X$ e $V = b'Y$. Então

$$\max_{a,b} \text{Corr}(U, V) = \rho_1^*$$

é obtida pelas combinações lineares (primeiro par de variável canônica)

$$U_1 = e_1' \Sigma_{11}^{-\frac{1}{2}} X \text{ e } V_1 = f_1' \Sigma_{22}^{-\frac{1}{2}} Y$$

O k -ésimo par de variáveis canônicas, $k = 1, 2, \dots, p$,

$$U_k = e_k' \Sigma_{11}^{-\frac{1}{2}} X \text{ e } V_k = f_k' \Sigma_{22}^{-\frac{1}{2}} Y$$

maximiza

$$\text{Corr}(U_k, V_k) = \rho_k^*$$

entre as combinações lineares não correlacionadas com as precedentes $1, 2, \dots, k-1$ variáveis canônicas.

Consideramos aqui $\rho_1^{*2} \geq \rho_2^{*2} \geq \dots \geq \rho_p^{*2}$ como sendo os autovalores de $\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}}$, e e_1, e_2, \dots, e_p são os (px1) autovetores associados. Os valores de $\rho_1^{*2}, \rho_2^{*2}, \dots, \rho_p^{*2}$ são também os p autovalores da matriz $\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}}$ com os (qx1) autovetores correspondentes f_1, f_2, \dots, f_p . Cada vetor f_i é proporcional a matriz $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-\frac{1}{2}} e_i$. Logo, a correlação entre U_k e V_k é dada por:

$$\text{Corr}(U_k, V_k) = \rho_k^* = \sqrt{\rho_k^{*2}}$$

As variáveis canônicas têm as seguintes propriedades:

- $\text{Var}(U_k) = \text{Var}(V_k) = 1$;
- $\text{Cov}(U_k, U_l) = \text{Corr}(U_k, U_l) = 0, k \neq l$;
- $\text{Cov}(V_k, V_l) = \text{Corr}(V_k, V_l) = 0, k \neq l$;
- $\text{Cov}(U_k, V_l) = \text{Corr}(U_k, V_l) = 0, k \neq l$ para todo $k, l = 1, 2, \dots, p$.

Quando utilizado apenas uma amostra da população, as matrizes de covariância Σ podem ser substituídas pelas matrizes de correlação ρ . A matriz de covariância e de correlação serão R e $\hat{\rho}$, respectivamente. Assim, as correlações e variáveis canônicas serão obtidas da mesma forma, a partir da matriz de covariância ou de correlação.

2.3 Interpretação das Variáveis Canônicas Populacionais

As variáveis canônicas em gerais são artificiais, isto é, não têm significado. A interpretação de \hat{U}_k, \hat{V}_k pode ser feita calculando as correlações amostrais entre as variáveis canônicas e as variáveis nos conjuntos X e Y . Definimos as matrizes

$$A_{p \times p} = [\hat{a}_1, \hat{a}_2, \dots, \hat{a}_p] \quad B_{q \times q} = [\hat{b}_1, \hat{b}_2, \dots, \hat{b}_q] \quad (9)$$

onde as linhas são os vetores coeficientes das variáveis canônicas amostrais, então:

$$U_{p \times 1} = \hat{A}x^{(1)} \quad V_{q \times 1} = \hat{B}y^{(2)} \quad (10)$$

e definimos

$$R_{\hat{U}, x^{(1)}} = \text{matriz de correlações amostrais de } \hat{U} \text{ com } x^{(1)}$$

$$R_{\hat{V}, y^{(2)}} = \text{matriz de correlações amostrais de } \hat{V} \text{ com } y^{(2)}$$

$$R_{\hat{U}, y^{(2)}} = \text{matriz de correlações amostrais de } \hat{U} \text{ com } y^{(2)}$$

$$R_{\hat{V}, x^{(1)}} = \text{matriz de correlações amostrais de } \hat{V} \text{ com } x^{(1)}$$

sendo

$$\begin{aligned} R_{\hat{U}, x^{(1)}} &= \hat{A}S_{11}D_{11}^{-\frac{1}{2}} \\ R_{\hat{V}, y^{(2)}} &= \hat{B}S_{22}D_{22}^{-\frac{1}{2}} \\ R_{\hat{U}, y^{(2)}} &= \hat{A}S_{12}D_{22}^{-\frac{1}{2}} \\ R_{\hat{V}, x^{(1)}} &= \hat{B}S_{21}D_{11}^{-\frac{1}{2}} \end{aligned} \quad (11)$$

onde $D_{11}^{-\frac{1}{2}}$ é uma matriz diagonal da ordem $p \times p$ com o i -ésimo elemento diagonal amostral igual a $(\text{var}(X_i))^{-\frac{1}{2}}$ e $D_{22}^{-\frac{1}{2}}$ é uma matriz diagonal da ordem $q \times q$ com o i -ésimo elemento diagonal correspondente a $(\text{var}(Y_i))^{-\frac{1}{2}}$.

2.4 Validação da Análise de Correlação Canônica

Para verificar a aplicabilidade da análise de correlação canônica ao conjuntos de dados do problema, deve-se verificar se os vetores X e Y são independentes entre si, ou não correlacionados (Mingoti, 2007). Caso isso aconteça, a aplicação do método será inútil, pois $a'X$ e $b'Y$ terão correlação zero para qualquer escolha de a' e b' . Portanto, para validar a análise de correlação canônica se faz necessário uma análise da matriz de covariância ou de correlações a fim de determinar se elas são próximas ou não da matriz nula.

Assim vamos testar as seguintes hipóteses:

$$H_0 : \Sigma_{XY} = 0_{p \times q}$$

contra

$$H_A : \Sigma_{XY} \neq 0_{p \times q}$$

ou, equivalentemente, para $\Sigma = R$ e a estatística do teste é definida por:

$$-2 \ln(\Lambda) = - \left(n - 1 - \frac{1}{2}(p + q + 1) \right) \ln \left(\prod_{i=1}^p (1 - \lambda_i) \right) \quad (12)$$

onde n é o tamanho da amostra. A estatística do teste está associada a uma distribuição Qui-quadrado χ_v^2 com pq graus de liberdade. Se o valor obtido em 12 for maior ou igual ao respectivo valor da distribuição qui-quadrado, com o nível de significância igual a 5%, rejeitamos a hipótese nula.

3 MATERIAIS E MÉTODOS

Com base em um banco de dados fornecido pela Polícia Rodoviária Federal do Paraná referente aos 2338 acidentes ocorridos no período de janeiro à dezembro de 2009 na BR-116 (Figura 1), analisamos o relacionamento de um conjunto variáveis para traçar o perfil dos acidentes nesta rodovia.

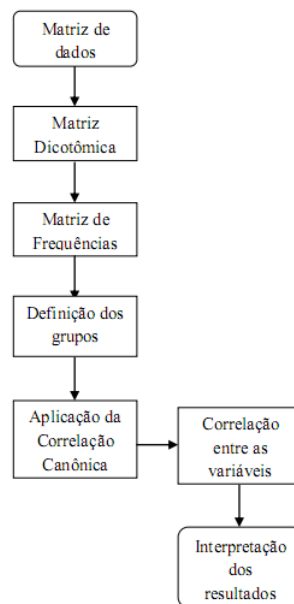


Figura 2: Fluxograma da Aplicação da Análise de Correlação Canônica

O Banco de dados contém informações sobre o tipo de acidente (capotamento, colisão frontal, colisão traseira, tombamento, saída de pista, entre outros), causas do acidente (desobediência a sinalização, velocidade incompatível, falta de atenção, não guardar distância de segurança), fases do dia (dia, noite, amanhecer, entre outros), condição metereológica (sol, chuva, nevoeiro, neblina, entre outros), quilômetro da rodovia onde ocorreu cada acidente e algumas características da pista. No total, foram consideradas 33 variáveis qualitativas selecionadas do banco de dados para aplicação da análise de correlação canônica. Todas essas variáveis foram divididos em dois grupos (ver Tabela 1) de modo a verificar as relações existentes entre os mesmos.

| | Tipos de Acidentes (Y) | | Outras características (X) |
|-------|----------------------------|----------|------------------------------------|
| Y_1 | Saída de Pista | X_1 | Desobediência à sinalização |
| Y_2 | Colisão lateral | X_2 | Velocidade incompatível |
| Y_3 | Tombamento | X_3 | Falta de Atenção |
| Y_4 | Colisão com objeto fixo | X_4 | Não guardar distância de segurança |
| Y_5 | Colisão traseira | X_5 | Acostamento com erosão |
| Y_6 | Capotamento | X_6 | Sem canteiro |
| Y_7 | Colisão transversal | X_7 | Faixa ruim |
| | | X_8 | Pista simples |
| | | X_9 | Pista dupla |
| | | X_{10} | Reta |
| | | X_{11} | Curva |
| | | X_{12} | Pista com erosão |
| | | X_{13} | Pista molhada |
| | | X_{14} | sol |
| | | X_{15} | chuva |
| | | X_{16} | nevoeiro/neblina |
| | | X_{17} | 0-30km |
| | | X_{18} | 30-60km |
| | | X_{19} | 60-90km |
| | | X_{20} | 90-120km |
| | | X_{21} | 120-150km |
| | | X_{22} | 150-180km |
| | | X_{23} | 180-210km |
| | | X_{24} | Noite |
| | | X_{25} | Dia |
| | | X_{26} | Anoitecer |

Tabela 1: Variáveis divididas em dois grupos

O trabalho foi desenvolvido conforme o fluxograma da Figura 2, em que, inicialmente, observamos o banco de dados do qual foram extraídos as variáveis qualitativas, para em seguida obter a matriz de dados dicômicos (2338 acidentes x 33 variáveis), onde cada elemento dessa nova matriz representa a ocorrência (1) ou não ocorrência (0) de uma variável em um determinado acidente. A matriz de correlação foi utilizada na análise de correlação canônica para obter as correlações entre os dois grupos. O cálculo das correlações entre as variáveis canônicas e originais foram feitas no sentido de verificar quais variáveis são mais representativas no seu respectivo grupo.

Todos os procedimentos efetuados neste trabalho foram feitos usando os programas Statgraphics 5.1 e Matlab R2009b.

4 RESULTADOS

Nesta seção apresentamos os resultados obtidos com a aplicação da análise de correlação canônica.

A primeira etapa do trabalho foi encontrar os quilômetros da rodovia onde se concentram o maior número de acidentes. Conforme a Figura 3, podemos observar que os acidentes se

| Número | Autovalores | Correlações Canônicas | Qui-Quadrado | Graus de Liberdade | p-value |
|--------|-------------|-----------------------|--------------|--------------------|---------|
| 1 | 0,9917 | 0,9958 | 2096,17 | 238 | 0,00 |
| 2 | 0,9498 | 0,9746 | 1233,49 | 198 | 0,00 |
| 3 | 0,8026 | 0,8959 | 694,885 | 160 | 0,00 |
| 4 | 0,6227 | 0,7891 | 402,759 | 124 | 0,00 |
| 5 | 0,4190 | 0,6473 | 227,287 | 90 | 0,00 |
| 6 | 0,3644 | 0,6036 | 129,528 | 58 | 0,00 |
| 7 | 0,2338 | 0,4835 | 47,9446 | 28 | 0,01 |

Tabela 2: Resultado da aplicação da análise de correlação canônica

concentram nos primeiros 60 quilômetros da rodovia e entre os quilômetros 100 a 150.

Na aplicação da análise de correlação canônica para verificar as relações existentes entre os tipos de acidentes e as características dos acidentes (ver Tabela 2), observando o *p-value*, vemos que todas as correlações canônicas são estatisticamente importantes considerando um nível de significância igual a 5% pelo teste do qui-quadrado. Analisando ainda os autovalores, podemos concluir que os dois primeiros autovalores bem como as correlações canônicas (raiz quadrada de cada autovalor) são mais significantes que os demais. Assim, obtemos dois pares de variáveis canônicas, U_1 e V_1 e U_2 e V_2 , com correlações 0,9958 e 0,9746, respectivamente. Não detalharemos aqui os pares de variáveis canônicas e seus respectivos autovetores.

A Tabela 3 apresenta as correlações entre as variáveis originais e variáveis canônicas. Podemos observar que as associações intergrupos são estabelecidas, principalmente, pela influência de:

1. O primeiro par canônico, com correlação canônica 0,9958, associa os tipos de acidentes: saída de pista, colisão com objeto fixo e capotamento, com velocidade incompatível, pista molhada, chuva, pleno dia, pista dupla, curva e o trecho entre os KM30 e KM60;
2. O segundo par canônico, com correlação canônica 0,9746, associa os tipos de acidentes: colisão traseira e transversal com desobediência, falta de atenção, acostamento com erosão, pista simples e com erosão, sem canteiro, céu claro, plena noite, anoitecer e o trecho entre os KM90 e KM120.

Estes resultados nos fornecem um perfil dos acidentes nos trechos determinados pela análise feita, como pode ser observado pelo histograma (Figura 3) onde há uma concentração grande de acidentes.

5 CONCLUSÕES

O principal do objetivo da análise de correlação canônica aplicada à matriz de frequência obtida por meio do banco de dados, permite associar os diversos tipos de acidentes, com as causas, condições da pista e tempo, entre outras variáveis consideradas neste trabalho, podendo assim traçar um perfil dos acidentes que ocorreram na rodovia considerada para o ano de 2009. Os resultados da análise efetuada permitiram concluir que:

1. As correlações canônicas foram elevadas considerando um nível de significância igual a 5%

| $U_1 e x^{(1)}$ | $V_1 e x^{(2)}$ | $U_1 e x^{(2)}$ | $V_1 e x^{(1)}$ | $U_2 e x^{(1)}$ | $V_2 e x^{(2)}$ | $U_2 e x^{(2)}$ | $V_2 e x^{(1)}$ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.8626 | 0.7023 | -0.0045 | -0.0580 | 0.0629 | 0.0030 | 0.4808 | 0.8485 |
| 0.0532 | -0.0677 | 0.8468 | 0.5147 | 0.7908 | 0.9677 | 0.0852 | 0.0430 |
| 0.3759 | 0.6271 | 0.1401 | -0.0635 | 0.0151 | 0.1124 | 0.7722 | 0.5194 |
| 0.6221 | 0.5187 | -0.0005 | -0.0220 | 0.1397 | -0.0480 | 0.6723 | 0.7058 |
| 0.0574 | 0.1358 | 0.8205 | 0.5935 | 0.7357 | 0.8964 | 0.2922 | 0.0296 |
| 0.6166 | 0.5125 | 0.4855 | -0.0476 | 0.0448 | 0.4936 | 0.5964 | 0.5518 |
| -0.1051 | 0.4463 | 0.5656 | 0.7487 | 0.4747 | 0.5997 | 0.6427 | -0.0851 |
| | 0.6277 | 0.0979 | | | 0.0753 | 0.5881 | |
| | 0.0399 | 0.4424 | | | 0.5010 | 0.2550 | |
| | 0.5812 | 0.0710 | | | 0.0199 | 0.6336 | |
| | 0.1164 | 0.8106 | | | 0.9062 | 0.2525 | |
| | 0.6614 | 0.3418 | | | 0.4035 | 0.5524 | |
| | 0.5743 | 0.2581 | | | 0.3741 | 0.4039 | |
| | 0.3956 | -0.0277 | | | -0.0351 | 0.1993 | |
| | 0.1915 | 0.6332 | | | 0.6424 | 0.5429 | |
| | 0.6896 | -0.0749 | | | -0.0480 | 0.3003 | |
| | 0.6561 | 0.1466 | | | 0.0878 | 0.7594 | |
| | -0.0904 | 0.8128 | | | 0.9490 | 0.0663 | |
| | 0.5812 | 0.3952 | | | 0.4464 | 0.5647 | |
| | 0.6286 | 0.0227 | | | -0.0590 | 0.5570 | |
| | 0.2554 | -0.1103 | | | -0.1466 | 0.3069 | |
| | -0.0343 | 0.4855 | | | 0.5742 | 0.1212 | |
| | -0.0873 | 0.5110 | | | 0.6032 | 0.0515 | |
| | -0.0575 | -0.1326 | | | -0.1638 | -0.0929 | |
| | -0.0949 | -0.1013 | | | -0.1189 | -0.0608 | |
| | 0.0305 | -0.1434 | | | -0.1532 | -0.0608 | |

Tabela 3: Correlações entre as variáveis canônicas e originais

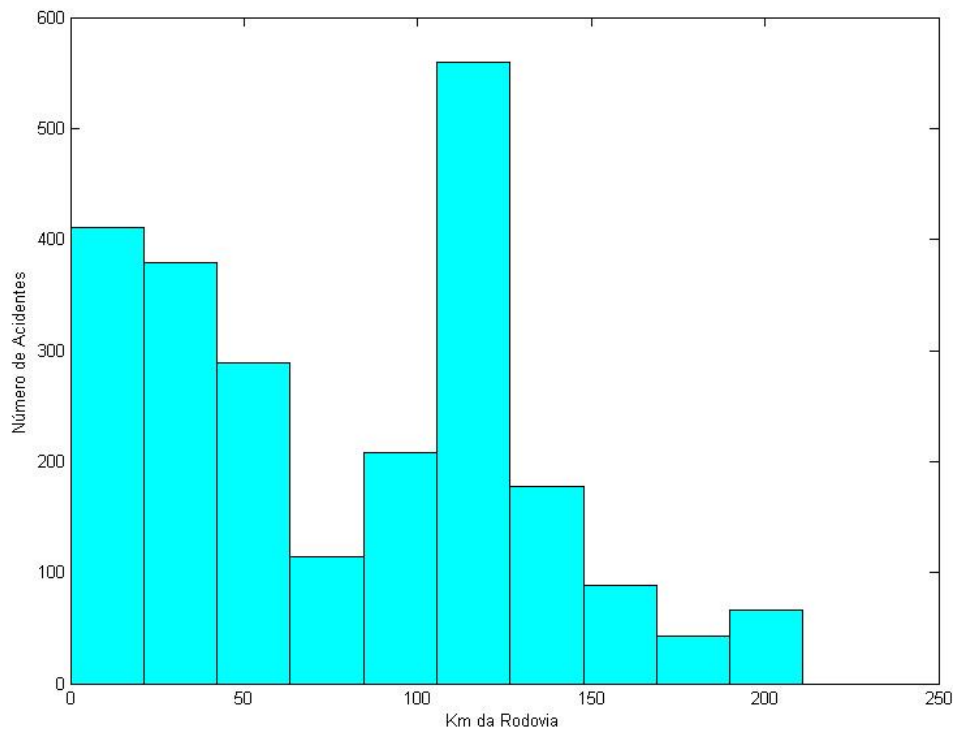


Figura 3: Distribuição dos Acidentes na BR-116 (Jan/09 à Dez/09)

2. As análises de correlação canônica permitem observar que os grupos considerados não são independentes, isto é os grupos considerados estão fortemente correlacionados;
3. Estes resultados demonstram que da extensão total de 211km no estado do Paraná da BR-116, nos trechos do KM30 ao KM60 e do KM90 ao KM120 ocorrem mais acidentes sendo as suas causas determinadas pela análise entre os pares de variáveis canônicas.

Desta forma, estes resultados poderão ser utilizados pelos engenheiros responsáveis pela rodovia BR-116 (trecho estado do Paraná) a fim de obter um melhor planejamento das mudanças que devem ser realizadas na mesma, com o principal objetivo de se reduzir o número de acidentes.

REFERÊNCIAS

- Costa G.G.O. Análise de correlação canônica entre conjuntos de princípios da qualidade total. *Cadernos de Administração*, 1, 2008.
- Hair J.F. e Anderson R.E. *Applied Multivariate Statistical Analysis*. New York, 2002.
- Johnson R.A. e Wichern D.W. *Applied Multivariate Statistical Analysis*. New York, 2002.
- Lira S.A. *Análise de Correlação: Abordagem Teórica e de Construção dos Coeficientes com Aplicações*. Tesis de Mestrado, UFPR, 2004.
- Mingoti S.A. *Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada*. Belo Horizonte, 2007.
- Neto A.C. *Análise Estatística Multivariada Aplicada à Pesquisa. Notas de Aula*. UFPR, Curitiba, 2002b.
- Trugilho P.F., Lima J.T., e Mori F.A. Correlação canônica das características químicas e físicas da madeira de clones de eucalyptus grandis e eucalyptus saligna. *CERNE*, 9:66–80, 2003.