

O uso da distribuição Poisson composta na teoria de valores extremos para previsão de arrecadação de multas por excesso de velocidade.

Filipe Costa

10/10/2022



UNIVERSIDADE
FEDERAL DO PIAUÍ

Objetivos

- Realizar uma previsão de arrecadação em nível diário, mensal e anual, utilizando distribuição composta de Poisson e Teoria de Valores de Extremos.

Metodologia:

- Distribuição de Poisson Composta e Teoria de Valores Extremos

Radares em Análise:

- Alameda Parnaíba, próximo a Ponte Estaiada João Isidoro França – Zona Norte;
- Av. Raul Lopes, em frente ao Teresina Shopping – Zona Leste;
- Av. Maranhão, no trecho entre o centro Administrativo e ponte da Amizade – Zona Sul;
- Av. Barão de Castelo Branco, próximo a Igreja Católica do Cristo Rei. – Zona Sul.

Medida e Considerada

- Por determinação do CTB - Código de Trânsito Brasileiro, existe uma margem de erro para o registro da velocidade.

Table: Velocidade medida e Velocidade Considerada

| Medida | Considerada |
|--------|-------------|
| 40 | 33 |
| 48 | 41 |
| 60 | 53 |
| 68 | 61 |
| 80 | 73 |
| 100 | 93 |
| 107 | 100 |

Tipificação da infração

- **Radares de 60 km/h**

Table: Divisão do tipo de infração por excesso de velocidade a 60 km/h

| Tipo | Faixa percentual | Faixa de Excesso | Valor da Multa |
|------------|------------------|---------------------|----------------|
| Média | até 20% | $1 \leq v \leq 12$ | R\$ 130,16 |
| Grave | entre 20% e 50% | $13 \leq v \leq 30$ | R\$ 195,23 |
| Gravíssima | superior a 50% | ≥ 31 | R\$ 880,41 |

- **Radares de 40 km/h**

Table: Divisão do tipo de infração por excesso de velocidade a 40 km/h

| Tipo | Faixa percentual | Faixa de Excesso | Valor da Multa |
|------------|------------------|--------------------|----------------|
| Média | até 20% | $1 \leq v \leq 8$ | R\$ 130,16 |
| Grave | entre 20% e 50% | $9 \leq v \leq 20$ | R\$ 195,23 |
| Gravíssima | superior a 50% | ≥ 21 | R\$ 880,41 |

Análise descritiva do dados - 1

Table: Informações sobre os dados registrados por dia

| Endereço | Data de Inicio | Data Final | Dias | Dias Zerados |
|------------------|----------------|------------|------|--------------|
| Alameda Parnaíba | 05/05/2016 | 31/12/2021 | 2089 | 282 |
| Av. Raul Lopes | 12/05/2017 | 19/10/2020 | 1696 | 595 |
| Av. Maranhão | 05/05/2017 | 31/12/2021 | 1702 | 292 |
| Av. B. C. Branco | 26/09/2016 | 31/12/2020 | 1830 | 334 |

Table: Probabilidade de termos 0 multas no dia

| Endereço | $p(0; \lambda)$ |
|---------------------|-----------------|
| Alameda Parnaiba | 1/12987052869 |
| Av. Raul Lopes | 1/1.638646e+13 |
| Av. Maranhão | 1/9102770 |
| Av. Barão C. Branco | 1/22181166 |

Análise descritiva do dados - 2

- Informações sobre quantidade de autos por radar e o tipo de infração:

Table: Divisão por tipo de infração para os radares

| Endereço | Dias | Nº de Autos | Percentual até 20% | Percentual entre 20% e 50% | Percentual acima 50% |
|--------------------------|------|-------------|--------------------|----------------------------|----------------------|
| Alameda Parnaíba | 1807 | 48647 | 90.23 | 9.30 | 0.45 |
| Av. Raul Lopes | 1101 | 51605 | 92.75 | 7.03 | 0.22 |
| Av. Maranhão | 1410 | 27273 | 88.34 | 10.94 | 0.72 |
| Av. B. de Castelo Branco | 1496 | 30954 | 79.18 | 17.66 | 3.16 |

Análise descritiva do dados - 3

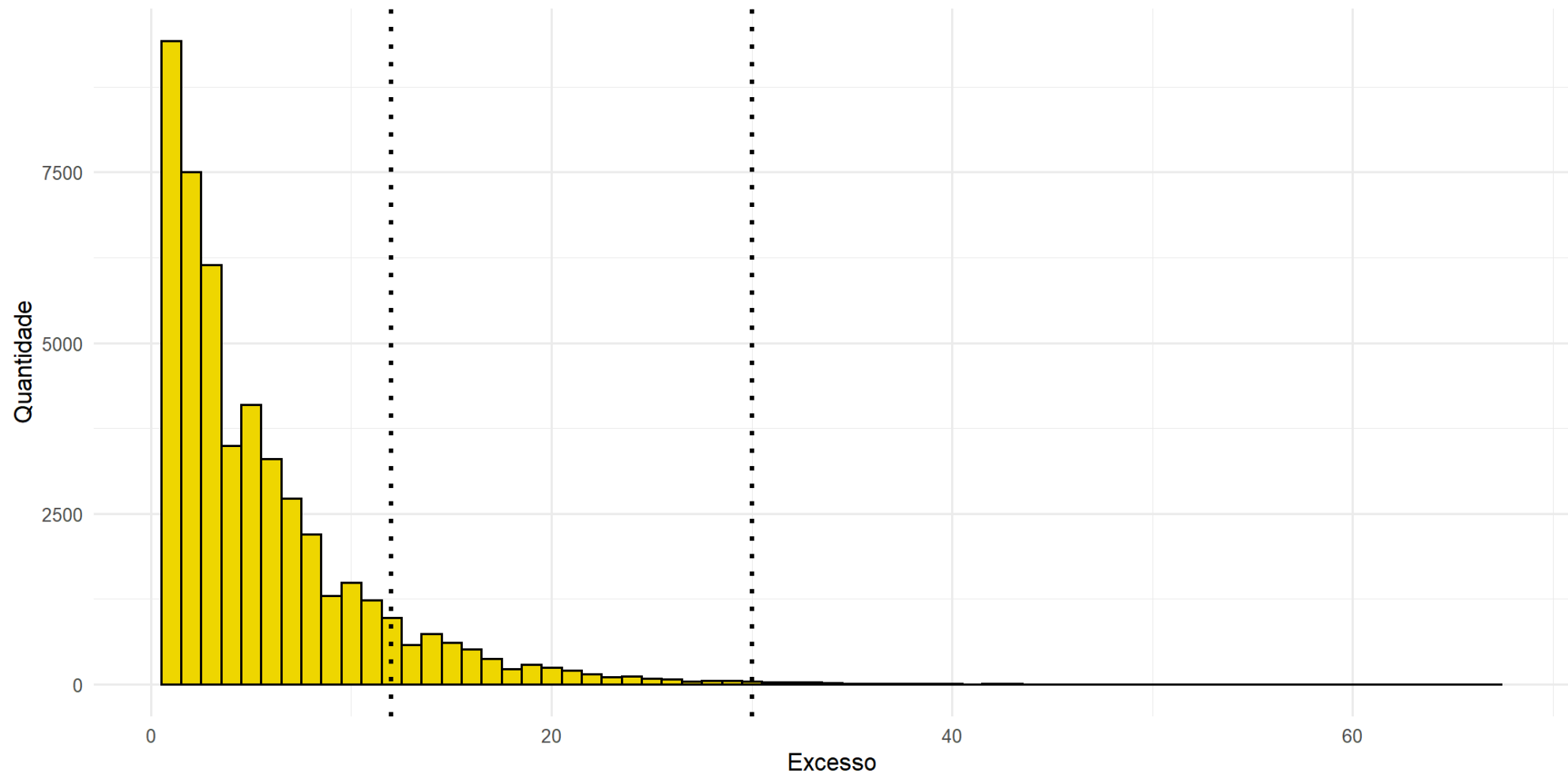
Table: Média e desvios-padrão diários por radar

| Endereços | Média diária | Desvio padrão |
|--------------------------|--------------|---------------|
| Alameda Parnaíba | 23.29 | 21.93 |
| Av. Raul Lopes | 30.43 | 25.91 |
| Av. Maranhão | 16.02 | 13.77 |
| Av. B. de Castelo Branco | 16.91 | 14.98 |

Table: Medidas descritivas das infrações de excesso por radar

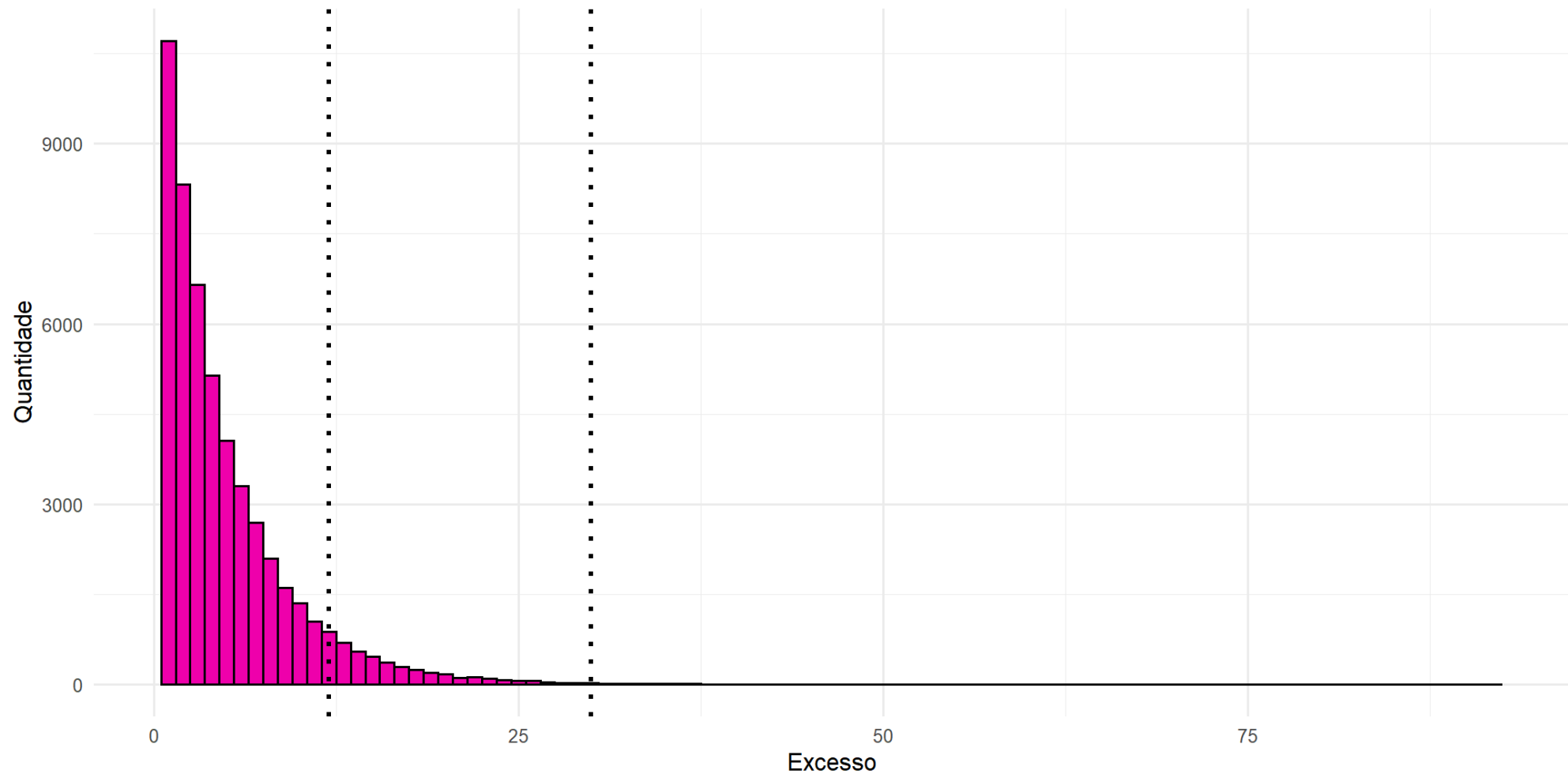
| Endereço | Média | Desvio Padrão | Mediana | Mínimo | Máximo |
|--------------------------|-------|---------------|---------|--------|--------|
| Alameda Parnaíba | 5.61 | 5.39 | 4 | 1 | 67 |
| Av. Raul Lopes | 5.02 | 4.70 | 4 | 1 | 92 |
| Av. Maranhão | 6.05 | 5.94 | 4 | 1 | 70 |
| Av. B. de Castelo Branco | 5.74 | 5.89 | 4 | 1 | 109 |

Análise descritiva do dados - 4



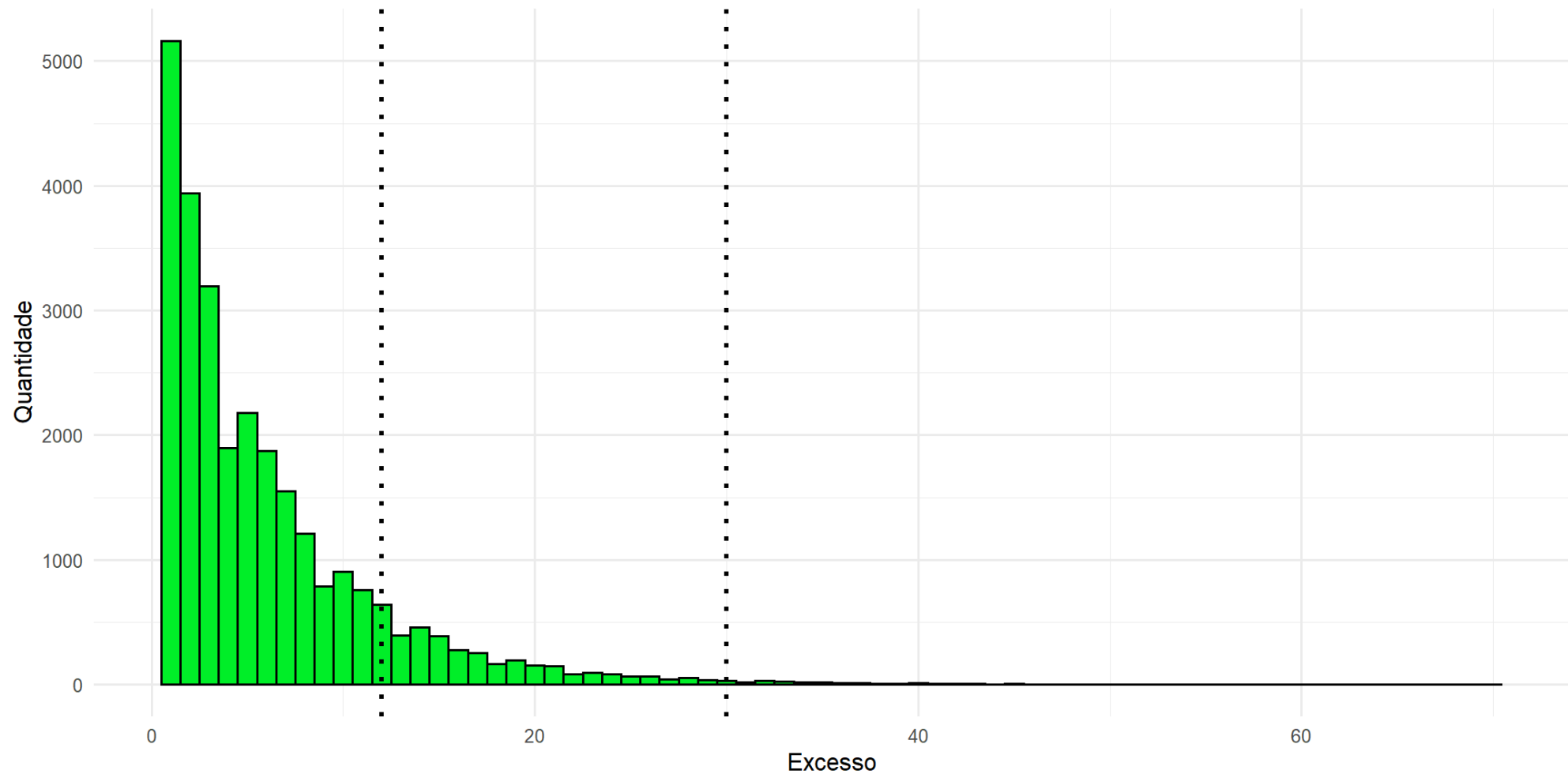
Distribuição de excesso na Alameda Parnaíba

Análise descritiva do dados - 6



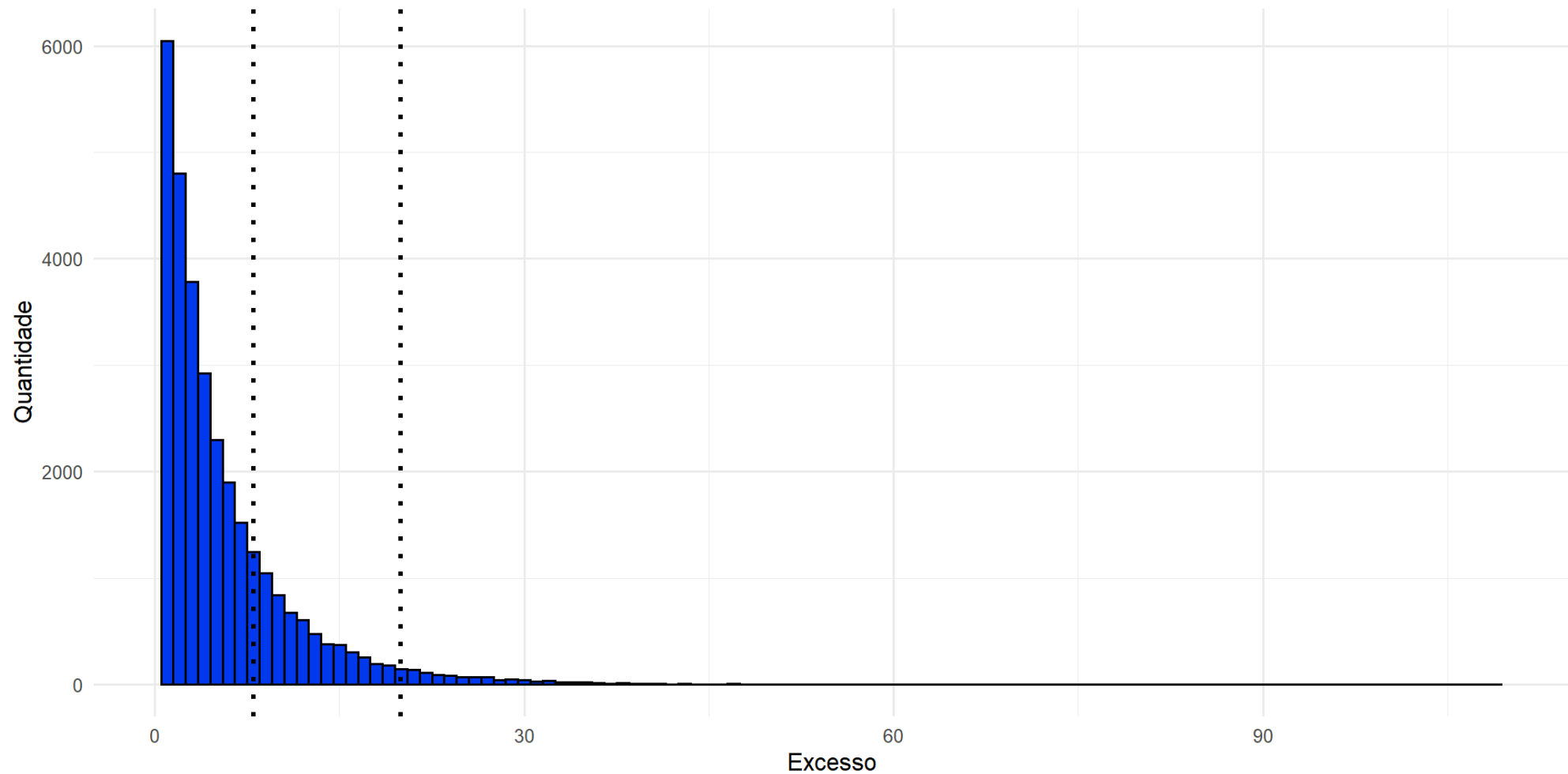
Distribuição de excesso na Av. Raul Lopes

Análise descritiva do dados - 7



Distribuição de excesso na Av. Maranhão

Análise descritiva do dados - 8



Distribuição de excesso na Av. B. Castelo Branco

Testes Hipoteses

Teste de corrida:

- H_0 : A sequência dos dados é aleatória
- H_1 : A sequência dos dados não é aleatória

Table: Valor-p do Teste de Corridas

| Endereço | p-valor |
|------------------------|---------|
| Alameda Parnaíba | 0.004 |
| Av. Raul Lopes | 0.009 |
| Av. Maranhão | 0.002 |
| Av. Barão de C. Branco | 0.028 |

Testes Hipoteses

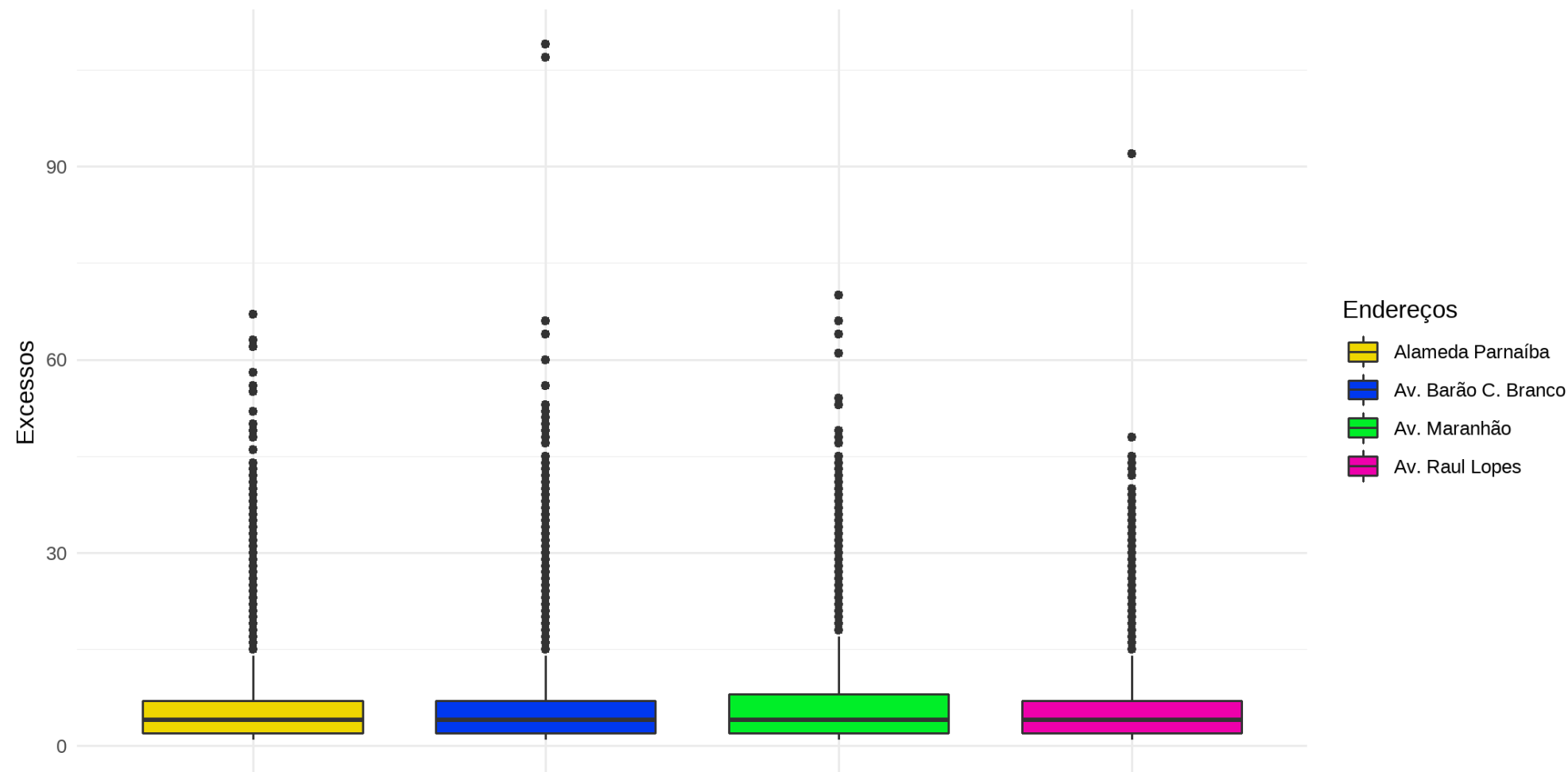
Teste de Kruskal-Wallis

- H_0 o comportamento semelhante para todos as distribuições em análise
- H_1 , o comportamento diferente para, pelo menos, duas das distribuições em análise
- $p - \text{valor}: < 2.2e - 16$

Table: Teste de Dunn

| group1 | group2 | n1 | n2 | statistic | p |
|---------------------|---------------------|-------|-------|------------|-----------|
| Alameda Parnaíba | Av. Barão C. Branco | 48647 | 30954 | -0.867789 | 0.3855098 |
| Alameda Parnaíba | Av. Maranhão | 48647 | 27273 | 7.843274 | 0.0000000 |
| Alameda Parnaíba | Av. Raul Lopes | 48647 | 51605 | -13.916882 | 0.0000000 |
| Av. Barão C. Branco | Av. Maranhão | 30954 | 27273 | 7.903749 | 0.0000000 |
| Av. Barão C. Branco | Av. Raul Lopes | 30954 | 51605 | -11.355477 | 0.0000000 |
| Av. Maranhão | Av. Raul Lopes | 27273 | 51605 | -19.672889 | 0.0000000 |

Comparação da distribuição



Boxplot comparativo das distribuições dos excessos

Arrecadação possível

Tabela com os valores possíveis de arrecadação durante o período de análise dos radares em análise:

Table: Arrecadação Possível

| Local | Valores |
|--------------------------|---------------|
| Alameda Parnaíba | R\$ 6.794.927 |
| Av. Raul Lopes | R\$ 7.036.259 |
| Av. Maranhão | R\$ 3.891.757 |
| Av. B. de Castelo Branco | R\$ 5.119.140 |

Distribuição Poisson

Modelando o número diário de Infrações

A distribuição Poisson possui um único parâmetro, λ , que corresponde a um parametro de intensidade. Sua função de distribuição é:

$$P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, \dots$$

Para esta distribuição, temos que $E(X) = \lambda$ e $VAR(X) = \lambda$

Distribuição Poisson composta

Supondo que $N \sim \text{Poisson}(\lambda)$ e N seja um conjunto de variáveis baseada em *Poisson* com valor esperado de λ e X_1, X_2, X_3, \dots sejam identicamente distribuídas e independentes, então a distribuição de probabilidade da soma de N

$$Y = \sum_{n=1}^N x_n$$

- Tabela da média diária:

| Endereços | a_1 | b_1 | λ | Limite Inferior | Limite Superior |
|---------------------|-------|----------|-----------|-----------------|-----------------|
| Alameda Parnaíba | 48647 | 1807.001 | 26.921 | 26.683 | 27.161 |
| Av. Raul Lopes | 51605 | 1101.001 | 46.871 | 46.467 | 47.276 |
| Av. Maranhão | 27273 | 1410.001 | 19.343 | 19.114 | 19.573 |
| Av. Barão c. Branco | 30954 | 1496.001 | 20.691 | 20.461 | 20.922 |

Função GPD

Teorema 1: Se x for uma variável aleatória (v.a.) com função distribuição (f.d.) F_x , que pertence ao domínio da de atração de uma distribuição GEV, então, quando $\mu \rightarrow \infty$, $F(x|u) = \Pr X > u + x | X > u$, possui distribuição GPD, com a seguinte função de distribuição:

$$P(x|\xi, \sigma, \mu) = \begin{cases} 1 - (1 + \xi \frac{(x-\mu)}{\sigma})^{-\frac{1}{\xi}}, & \text{se } \xi \neq 0 \\ 1 - \exp\left\{-\frac{(x-\mu)}{\sigma}\right\}, & \text{se } \xi = 0 \end{cases}$$

onde $\mu > 0$, $x - \mu \geq 0$, se $\xi > 0$, e $0 \leq x - \mu \leq -\frac{\sigma}{\xi}$, se $\xi < 0$. O caso $\xi = 0$ é interpretado como sendo o limite quando $\xi \rightarrow 0$, e tem como caso particular a distribuição exponencial de parâmetro $\frac{1}{\sigma}$. Os parametros são ξ , σ e μ , que representam, respectivamente, a forma, a escala e o limiar da distribuição.

- Densidade:

$$p(x|\xi, \sigma, \mu) = \begin{cases} \frac{1}{\sigma} (1 + \xi \frac{x-\mu}{\sigma})^{-\frac{1}{\xi}}, & \text{se } \xi \neq 0 \\ \frac{1}{\sigma} \exp\left\{-\frac{(x-\mu)}{\sigma}\right\}, & \text{se } \xi = 0 \end{cases}$$

Determinação do limiar

Métodos mais convencionais de determinação do limiar utilizam-se de análises gráficas da linearidade de N_u . Um método muito utilizado é o gráfico de médias de excessos (MLP\< *Mean Residual Life Plot*), baseado na espera da GPD, NASCIMENTO (2012). Sua construção segue o seguinte formato:

$$\left\{ \left(\mu, \frac{1}{n_u} \sum_{i=1}^{n_u} \right) : \mu < x_{max} \right\}$$

onde $x_1 \leq x_2 \leq \dots x_n$ consistem nas N_u observações que excedem μ , e x_{max} é o valor mais elevado das observações.

Considerando a distribuição GPD válida para os excesso, esta também é válida para os excesso acima de todos os limiares $\mu > \mu_0$, sujeito a mudanças no parâmetro de escala $\sigma_\mu = \sigma_{\mu_0} + \xi_\mu$. Então, para $\mu > \mu_0$

$$E(X - \mu | X > \mu) = \frac{\sigma_\mu}{1 - \xi} = \frac{\sigma_{\mu_0} + \xi_\mu}{1 - \xi}$$

Mistura de Gammas

Modelo para dados extremos que utiliza aproximação não-paramétrica baseado em mistura de distribuição Gama para valores extremos maiores que um limiar e GPD para cauda dos dados, a função densidade é definida como:

$$f(x|\theta, p, \Psi) = \begin{cases} h(x|\mu, \eta, p), & \text{se } x \leq \mu \\ (1 - H(x|\mu, \eta, p))p(x|\Psi), & \text{se } x > \mu \end{cases}$$

onde H é a função de distribuição acumulada da mistura de Gammas, $\Psi = (\xi, \sigma, \mu)$, $\sigma > 0$, $(x - \mu) \leq -\sigma/\xi$ e $x > \mu$.

- Encontrando os quantis altos

Na distribuição GPD, pode-se encontrar um quantil q com probabilidade $P(X < q)$ em função dos parâmetros. Invertendo a função acumulada, obtém-se a seguinte função dos quantis da cauda:

$$q_x p = \frac{((1 - p^*)^{-\xi} - 1)}{\xi},$$

onde $p^* = 1 - (1 - p)N/N_u$.

Parâmetros Estimados - Alameda Parnaíba

-Alameda Parnaíba

| Parâmetro | Estimado | Limite Superior | Limite Inferior |
|-----------|----------|-----------------|-----------------|
| ξ | -0,06 | -0,07 | -0,05 |
| σ | 6,33 | 6,18 | 6,45 |
| u | 3,01 | 3,01 | 3,01 |
| μ_1 | 1,00 | 1,00 | 1,00 |
| μ_2 | 3,55 | 3,52 | 3,57 |
| η_1 | 219,16 | 219,05 | 219,18 |
| η_2 | 10,07 | 9,85 | 10,40 |
| w_1 | 0,19 | 0,19 | 0,19 |
| w_2 | 0,81 | 0,81 | 0,81 |

• Dados:

| Parâmetro | Medidas |
|-------------------|------------|
| Acima do Limiar | 52,28% |
| Velocidade Máxima | 55,33 km/h |
| Nível médio | 87,94% |
| Nível grave | 11,67% |
| Nível gravíssima | 0,39% |

Parâmetros Estimados - Av. Raul Lopes

- Av. Raul Lopes

| Parâmetro | Estimado | Limite Superior | Limite Inferior |
|-----------|----------|-----------------|-----------------|
| ξ | -0,04 | -0,05 | -0,04 |
| σ | 5,40 | 5,31 | 5,46 |
| u | 3,00 | 3,00 | 3,00 |
| μ_1 | 1,00 | 1,00 | 1,00 |
| μ_2 | 3,47 | 3,46 | 3,49 |
| η_1 | 392,15 | 392,07 | 392,24 |
| η_2 | 10,35 | 10,12 | 10,47 |
| w_1 | 0,21 | 0,20 | 0,21 |
| w_2 | 0,79 | 0,79 | 0,80 |

- Dados:

| Parâmetro | Medidas |
|-------------------|---------|
| Acima do Limiar | 52,28% |
| Velocidade Máxima | 60 km/h |
| Nível médio | 91,01% |
| Nível grave | 8,81 % |
| Nível gravíssima | 0,18% |

Parâmetros Estimados - Av. Maranhão

- Av. Maranhão

| Parâmetro | Estimado | Limite Superior | Limite Inferior |
|-----------|----------|-----------------|-----------------|
| ξ | 0,19 | 0,16 | 0,21 |
| σ | 4,25 | 4,17 | 4,35 |
| u | 2,00 | 1,99 | 2,00 |
| μ_1 | 1,00 | 1,00 | 1,00 |
| μ_2 | 10,36 | 8,30 | 11,93 |
| η_1 | 353,15 | 353,04 | 353,17 |
| η_2 | 10,40 | 8,52 | 11,90 |
| w_1 | 0,19 | 0,19 | 0,19 |
| w_2 | 0,81 | 0,81 | 0,81 |

- Dados:

| Parâmetro | Medidas |
|-------------------|----------|
| Acima do Limiar | 66,64% |
| Velocidade Máxima | ∞ |
| Nível médio | 88,29% |
| Nível grave | 10,51 % |
| Nível gravissima | 1,16% |

Parâmetros Estimados - Av. Barão C. Branco

- Av. Barão C. Branco

| Parâmetro | Estimado | Limite Superior | Limite Inferior |
|-----------|----------|-----------------|-----------------|
| ξ | 0,00 | -0,01 | 0,01 |
| σ | 6,20 | 6,09 | 6,31 |
| u | 3,01 | 3,01 | 3,02 |
| μ_1 | 1,02 | 1,00 | 1,09 |
| μ_2 | 3,62 | 3,60 | 3,67 |
| η_1 | 13,00 | 9,98 | 14,20 |
| η_2 | 8,85 | 8,58 | 9,14 |
| w_1 | 0,20 | 0,20 | 0,20 |
| w_2 | 0,80 | 0,80 | 0,80 |

- Dados:

| Parâmetro | Medidas |
|-------------------|----------|
| Acima do Limiar | 52,74% |
| Velocidade Máxima | ∞ |
| Nível médio | 87,46% |
| Nível grave | 11,83% |
| Nível gravissima | 0,70% |

Resultados

Considerando a estimação do número de autuções diárias pela distribuição de Poisson e a estimação do valor do excesso pela distribuição de mistura por GPD, obtemos a estimação do valor esperado de arrecadação diária, mensal e anual, considerando os valores de cada atuação e de cada faixa de multa. Será utilizado o valor esperado da distribuição poisson composta.

- **Arrecadação diária:**

| Endereços | IC- | IC+ | Mediana |
|---------------------|-------------|------------|-------------|
| Alameda Parnaíba | R\$3704.35 | R\$3764.83 | R\$3727.48 |
| Av. Raul Lopes | R\$6336.13 | R\$6427.17 | R\$6383.88 |
| Av. Maranhão | R\$ 2657.42 | R\$2728.49 | R\$ 2693.00 |
| Av. Barão c. Branco | R\$2841.20 | R\$2900.33 | R\$2875.46 |

Resultados

- Arrecadação mensal

| Endereços | IC- | IC+ | Mediana |
|---------------------|---------------|---------------|--------------|
| Alameda Parnaíba | R\$ 111130.57 | R\$ 112944.98 | R\$ 111824.4 |
| Av. Raul Lopes | R\$190083.90 | R\$ 192815.17 | R\$191516.4 |
| Av. Maranhão | R\$ 79722.68 | R\$ 81854.93 | R\$ 80790.0 |
| Av. Barão c. Branco | R\$85236.23 | R\$ 87010.12 | R\$ 86263.8 |

- Arrecadação anual

| Endereços | IC- | IC+ | Mediana |
|---------------------|---------------|---------------|-------------|
| Alameda Parnaíba | R\$ 1352088.7 | R\$1374163.9 | R\$1360530 |
| Av. Raul Lopes | R\$ 2312687.5 | R\$ 2345918.0 | R\$ 2330116 |
| Av. Maranhão | R\$969959.2 | R\$ 995901.6 | R\$ 982945 |
| Av. Barão c. Branco | R\$1037040.7 | R\$ 1058623.2 | R\$ 1049543 |

Conclusão

O trabalho foi dividido em duas partes. Num primeiro momento, realizamos uma análise através da Distribuição Poisson analisando a frequência diária das infrações de trânsito. A segunda parte consistiu em parametrizar a distribuição dos excessos através da teoria dos valores extremos utilizando a função GPD.

Acreditamos que os resultados obtidos são satisfatórios e que podem auxiliar o poder público na sua organização. Ressaltamos, também, que a análise feita pode ser replicada em outras pesquisas, independente da área, ou seja, não fica restrita a análise de radares, mas, qualquer área que possa ser analisada através da Teoria de Valores Extremos relacionando com a Poisson.

Referências

- [1] Nascimento FF, Gamerman D, Lopes HF (2011) \textbf{Regression models for exceedance data via the full likelihood}. Environ Ecol Stat 18:495-512.
- [2] CTB - \textbf{Código de Trânsito Brasileiro}. Disponível em *[http : //www.planalto.gov.br/ccivil_03/leis/l9503compilado.htm](http://www.planalto.gov.br/ccivil_03/leis/l9503compilado.htm)*
- [3] Beijo, Luiz Alberto e Avelar, Fabricio Goecking (2011). \textbf{Distribuição generalizada de valores extremos no estudo de dados climáticos : Uma breve revisão e aplicação.}, Revista da Estatística da UFOP, Minas Gerais, v. 1, p.10-15, jan. 2011.
- [4] Bussab, W. O. e Morettin, P. A. (2010). \textbf{Estatística Básica}, Editora Saraiva, 6a. Edição.
- [5] Nascimento, F. F. (2012). \textbf{Modelos Probabilísticos Para Dados Extremos: Teoria e Aplicações.}In: II COLOQUIO DE MATEMÁTICA DA REGIÃO NORDESTE, 2012, Teresina, Piauí. Universidade Federal do Piauí, Edufpi.
- [6] Gibbons, JD e Chakraborti. \textbf{Nonparametric Statistical Inference}. Fourth Edition, Revised and Expanded. New York.
- [7] CASELLA, G. e BERGER, R.L,. \textbf{Inferência Estatística}. Tradução: Solange Aparecida Visconte. Cengage Learning. São Paulo, 2020.

Obrigado!!!