



UNIVERSIDADE FEDERAL DO PIAUÍ  
CENTRO DE CIÊNCIAS DA NATUREZA  
CURSO DE GRADUAÇÃO EM ESTATÍSTICA

**Análise das infrações de trânsito de excesso de velocidade  
através da Teoria de Valores Extremos**

**Filipe Mateus de Sousa Costa**

**TERESINA - 2022**

**Filipe Mateus de Sousa Costa**

**Monografia:**

**Análise das infrações de trânsito de excesso de velocidade através da Teoria  
de Valores Extremos**

Monografia submetida à Coordenação do  
curso de graduação em Estatística, da Uni-  
versidade Federal do Piauí, como requisito  
parcial para obtenção do grau de Bacharel  
em Estatística.

**Orientador:**

Profa. Dra. Valmária Rocha da Silva Ferraz

**Co-orientador:**

Prof. Dr. Fernando Ferraz do Nascimento

**TERESINA - 2022**

# Contents

|  |           |
|--|-----------|
| <b>List of Figures</b>   | <b>4</b>  |
| <b>List of Tables</b>  | <b>5</b>  |
| <b>1 Agradecimentos</b>  | <b>6</b>  |
| <b>2 Resumo</b>  | <b>7</b>  |
| <b>3 Abstract</b>  | <b>8</b>  |
| <b>4 Introdução</b>  | <b>9</b>  |
| 4.1 Apresentando os dados . . . . .                              | 9         |
| 4.1.1 Pontos de medição de excesso de velocidade . . . . .       | 11        |
| 4.1.2 Detalhando os dados . . . . .                              | 12        |
| 4.1.3 Análise descritiva e da distribuição dos excessos. . . . . | 13        |
| 4.1.4 Alguns testes . . . . .                                    | 15        |
| 4.1.5 Arrecadação possível . . . . .                             | 17        |
| <b>5 Distribuição Pareto Generalizada (GPD)</b>                  | <b>19</b> |
| 5.1 Domínio da atração . . . . .                                 | 20        |
| 5.2 Determinação do Limiar . . . . .                             | 21        |
| 5.3 Estimação da GPD . . . . .                                   | 22        |
| <b>6 Inferência Bayesiana</b>                                    | <b>23</b> |
| <b>7 Modelos dinâmicos para Valores Extremos</b>                 | <b>24</b> |
| 7.1 Alameda Parnaíba . . . . .                                   | 24        |
| 7.2 Av. Maranhão . . . . .                                       | 24        |
| 7.3 Av. Raul Lopes - Shopping . . . . .                          | 26        |
| 7.4 Av. dos Ipês . . . . .                                       | 27        |
| 7.5 Av. Av. Jose Francisco de Almeida Neto . . . . .             | 28        |
| 7.6 Av. Barão de Castelo Branco . . . . .                        | 29        |
| <b>8 Conclusão</b>   | <b>31</b> |
| <b>9 Referências Bibliográficas</b>                              | <b>32</b> |

## List of Figures

|    |   |    |
|----|---|----|
| 1  | Tabela com relação Velocidade Medida (VM) e Velocidade Considerada (VC) . . . . . | 10 |
| 2  | Distribuição de excesso na Alameda Parnaíba . . . . .                             | 13 |
| 3  | Distribuição de excesso na Av. Raul Lopes . . . . .                               | 14 |
| 4  | Distribuição de excesso na Av. Maranhão . . . . .                                 | 14 |
| 5  | Distribuição de excesso na Av. B. Castelo Branco . . . . .                        | 15 |
| 6  | Boxplot comparativo das distribuições dos excessos . . . . .                      | 17 |
| 7  | Alameda Parnaíba - Função fggpd . . . . .   | 24 |
| 8  | Alameda Parnaíba - Função fmgpd . . . . .   | 24 |
| 9  | Av. Maranhão - Função fggpd . . . . .   | 25 |
| 10 | Av. Maranhão - Função fmgpd . . . . .   | 25 |
| 11 | Av. Raul Lopes - Função fggpd . . . . .   | 26 |
| 12 | Av. Raul Lopes - Função fmgpd . . . . .   | 26 |
| 13 | Av. dos Ipês - Função fggpd . . . . .   | 27 |
| 14 | Av. dos Ipês - Função fmgpd . . . . .   | 27 |
| 15 | Av. Jose - Função fggpd . . . . .   | 28 |
| 16 | Av. Jose - Função fmgpd . . . . .   | 28 |
| 17 | Av. Barão - Função fggpd . . . . .  | 29 |
| 18 | Av. Barão - Função fmgpd . . . . .  | 29 |

## List of Tables

# 1 Agradecimentos

Primeiramente, eu preciso agradecer a Deus por ter permitido a possibilidade de realizar uma segunda graduação e ter me sustentado durante todo o curso, foram vários os desafios, mas Ele permitiu que eu conseguisse superar tudo. Soli Deo Glória.

Preciso agradecer a minha esposa que ficou do meu lado durante este período, sacrificando-se para permitir que eu estudasse, motivando-me. Superando o ciúme por ter sido trocada, muitas vezes, por um livro cheio de fórmulas esquisitas. Incluo meu filho Mateus, que chegou no durante o curso para dar mais emoção à vida. Mesmo pequeno, me ensinou a ser mais produtivo. E claro, meu pai, minha mãe e minha irmã, mesmo a distância, me apoiavam.

Agradeço a Professora Valmaria por ter aceitado o desafio de me orientar nesse TCC, juntamente com o professor Fernando que ofereceu o apoio teórico para a realização desse projeto. Incluo a STRANS, na pessoa do superintendente Claudio Pessoa, que gentilmente, disponibilizaram os dados para este trabalho.

Finalizo os agradecimentos aos amigos que foram criados durante o curso. Edvaldo que, como eu, chefe de família. Compartilhamos muito do desafio de trabalhar, cuidar do lar e estudar. Eva, cheia de energia, se esforçava para tirar as melhores notas nas disciplinas. Envelheceu alguns anos durante as madrugadas acordadas, mas ok. Obrigado a todos, de coração.

## 2 Resumo

### 3 Abstract



## 4 Introdução

Eventos extremos são situações ou comportamentos que não ocorrem com tanta frequência. No desenho da distribuição normal, são aqueles que ficam próximos as caudas, distantes do pico onde está localizado a média, a situação de normalidade. Entretanto, é possível que estes eventos ocorram, na verdade, é esperado que eles ocorram. Diante dessa situação, é importante que possamos prever seus acontecimentos, pois quando ocorrem, as consequências podem ser trágicas.

Exemplos que demonstram com facilidade os efeitos de eventos extremos estão na área da climatologia. Para os portais de notícias, o aquecimento global tem trazido mais manchetes de catástrofes climáticas que causaram prejuízos substanciais e, infelizmente, mortes. Uma busca rápida em sites de pesquisa nos direciona facilmente para estes eventos.

O surgimento da teoria de valores extremos surge da necessidade de pesquisas sobre eventos extremos de forma mais eficiente. As distribuições mais comuns trabalham melhor na análise dos eventos centrais, mais frequentes. Dessa forma, os acontecimentos mais raros possuem ficam mais difíceis de prever.

### 4.1 Apresentando os dados

Os dados que iremos trabalhar neste projeto têm como origem as medições de excesso de velocidade registradas por radares no município de Teresina. Baseado no Código de Trânsito Brasileiros, ruas, avenidas e estradas no Brasil, possuem limites de velocidade especificados, entretanto, para a realização do registro infração, é preciso que a via esteja sinalizada informando a velocidade máxima.

Um ponto importante sobre as infrações de velocidade é que existe uma margem de erro para medição da velocidade registrada. Para fim de registro de infração, não se registra a velocidade obtida por instrumentos de medição, mas subtrai-se um valor na velocidade medida (para vias de velocidade máxima até 100km/h, o valor é 7), o que passar deste valor acima da velocidade permitida, será considerado para infração.

Por exemplo, para vias de com velocidade máxima de 60 km/h, as infrações serão registradas somente quando o veículo passar pelo ponto de medição quando a velocidade calculada for de 68 km/h, registrando um excesso de 1km/h. O resultado dessa distinção é o uso de dois termos: Velocidade medida, que consiste na velocidade registrada e Velocidade Considerada, que é a velocidade medida menos o fator de correção. Na tabela abaixo, temos a relação entre velocidade medida e velocidade considerada. A infração é feita quando a considerada é maior que permitida na via.

| VM<br>(Km/h) | VC<br>(Km/h) | VM<br>(Km/h) | VC<br>(Km/h) | VM<br>(Km/h) | VC<br>(Km/h) | VM<br>(Km/h) | VC<br>(Km/h) |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 27           | 20           | 69           | 62           | 111          | 103          | 153          | 142          |
| 28           | 21           | 70           | 63           | 112          | 104          | 154          | 143          |
| 29           | 22           | 71           | 64           | 113          | 105          | 155          | 144          |
| 30           | 23           | 72           | 65           | 114          | 106          | 156          | 145          |
| 31           | 24           | 73           | 66           | 115          | 107          | 157          | 146          |
| 32           | 25           | 74           | 67           | 116          | 108          | 158          | 147          |
| 33           | 26           | 75           | 68           | 117          | 109          | 159          | 148          |
| 34           | 27           | 76           | 69           | 118          | 110          | 160          | 149          |
| 35           | 28           | 77           | 70           | 119          | 111          | 161          | 150          |
| 36           | 29           | 78           | 71           | 120          | 112          | 162          | 151          |
| 37           | 30           | 79           | 72           | 121          | 113          | 163          | 152          |
| 38           | 31           | 80           | 73           | 122          | 113          | 164          | 153          |
| 39           | 32           | 81           | 74           | 123          | 114          | 165          | 153          |
| 40           | 33           | 82           | 75           | 124          | 115          | 166          | 154          |
| 41           | 34           | 83           | 76           | 125          | 116          | 167          | 155          |
| 42           | 35           | 84           | 77           | 126          | 117          | 168          | 156          |
| 43           | 36           | 85           | 78           | 127          | 118          | 169          | 157          |
| 44           | 37           | 86           | 79           | 128          | 119          | 170          | 158          |
| 45           | 38           | 87           | 80           | 129          | 120          | 171          | 159          |
| 46           | 39           | 88           | 81           | 130          | 121          | 172          | 160          |
| 47           | 40           | 89           | 82           | 131          | 122          | 173          | 161          |
| 48           | 41           | 90           | 83           | 132          | 123          | 174          | 162          |
| 49           | 42           | 91           | 84           | 133          | 124          | 175          | 163          |
| 50           | 43           | 92           | 85           | 134          | 125          | 176          | 164          |
| 51           | 44           | 93           | 86           | 135          | 126          | 177          | 165          |
| 52           | 45           | 94           | 87           | 136          | 126          | 178          | 166          |
| 53           | 46           | 95           | 88           | 137          | 127          | 179          | 166          |
| 54           | 47           | 96           | 89           | 138          | 128          | 180          | 167          |
| 55           | 48           | 97           | 90           | 139          | 129          | 181          | 168          |
| 56           | 49           | 98           | 91           | 140          | 130          | 182          | 169          |
| 57           | 50           | 99           | 92           | 141          | 131          | 183          | 170          |
| 58           | 51           | 100          | 93           | 142          | 132          | 184          | 171          |
| 59           | 52           | 101          | 94           | 143          | 133          | 185          | 172          |
| 60           | 53           | 102          | 95           | 144          | 134          | 186          | 173          |
| 61           | 54           | 103          | 96           | 145          | 135          | 187          | 174          |
| 62           | 55           | 104          | 97           | 146          | 136          | 188          | 175          |
| 63           | 56           | 105          | 98           | 147          | 137          | 189          | 176          |
| 64           | 57           | 106          | 99           | 148          | 138          | 190          | 177          |
| 65           | 58           | 107          | 100          | 149          | 139          | 191          | 178          |
| 66           | 59           | 108          | 100          | 150          | 140          | 192          | 179          |
| 67           | 60           | 109          | 101          | 151          | 140          | 193          | 179          |
| 68           | 61           | 110          | 102          | 152          | 141          | 194          | 180          |

Figure 1: Tabela com relação Velocidade Medida (VM) e Velocidade Considerada (VC)

Na tabela acima, temos os valores entre as velocidades medidas e consideradas, note-se que na medida 107 e 108 km/h, a considerada se repete, neste ponto se inicia um aumento da margem de erro. Para este trabalho, iremos utilizar a medição calculada com a Velocidade considerada, pois o cálculo da infração é realizado com ela.

Em relação as infrações, o CTB normatiza três tipos de infrações para o excesso de velocidade baseado no percentual registado acima. Para medições até 20% acima da permitida, a infração é média, entre 20% e 50% é considerada grave e, superior a 50%, é gravíssima. O aumento da gravidade resulta no aumento do valor pago da multa. Nas tabelas abaixo, temos um resumo da divisão de tipificação, o intervalo de excesso e o valor pago. A primeira tabela refere-se para vias com velocidade máxima até 60 km/h e a segunda tabela considera a máxima 40km/h

- **Radares de 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%   | $\geq 31$           | R\$ 880,41     |

- **Radares de 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%   | $\geq 21$          | R\$ 880,41     |

#### 4.1.1 Pontos de medição de excesso de velocidade

Para a realização deste trabalho, escolhemos 4 endereços, são eles:

- Alameda Parnaíba, próximo Ponte Estaiada João Isidoro França – Zona Norte;
- Av. Raul Lopes, em frente ao Teresina Shopping – Zona Leste;
- Av. Maranhão, trecho entre o centro Administrativo e ponte da Amizade – Zona Sul;
- Av. Barão de Castelo Branco, próximo Igreja Católica do Cristo Rei. – Zona Sul.

Estes radares são considerados famosos com velocidade máxima de 60 km/h, exceção do radar da Av. Barão de Castelo Branco, onde o excesso é de 40 km/h. Seus bancos de dados são extensos e iremos apresentar algumas medidas descritivas para termos uma noção melhor da distribuição dos excessos.

É importante destacar que não iremos realizar uma distinção do sentido que o veículo está indo, para análises futuras, pode-se aprofundar nesta análise. Destaco, também, que utilizaremos todas os registros realizados, independente se auto não foi expedido por motivos técnicos ou administrativos.

#### 4.1.2 Detalhando os dados

Os dados utilizados apresentam algumas diferenças de período de registro por via, entretanto, todos os dados têm a data limite de 31/12/2021, sendo seus inícios variados. Isto ocorre devido a fatores administrativos e técnicos. Um fator que não podemos mensurar facilmente são os dias zerados para autuações, pois não podemos identificar se foi problema técnico ou se realmente não houve nenhum registro naquele específico.

Abaixo, apresentamos uma tabela especificando a data de início e fim dos registros, a quantidade dias entre as datas e a quantidade de dias zerados.

| 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          |

A próxima tabela apresenta a quantidade de dias com registros de autuação, a quantidade de autos registrados em todo o período analisando e a divisão percentual por tipo de infração apresentada anteriormente, média, grave e gravíssima.

| 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                 |

| Endereço                 | Dias | Nº de Autos | Percentual até 20% | Percentual entre 20% e 50% | Percentual acima 50% |
|--------------------------|------|-------------|--------------------|----------------------------|----------------------|
| Av. B. de Castelo Branco | 1496 | 30954       | 79.18              | 17.66                      | 3.16                 |

#### 4.1.3 Análise descritiva e da distribuição dos excessos.

O objetivo deste tópico é apresentar algumas medidas básicas dos radares que estamos analisando. Primeiramente, temos um tabela com as informações descritivas dos radares. Num segundo momento, construímos algumas imagens com `ggplot` com as distribuições dos excessos dos radares.

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

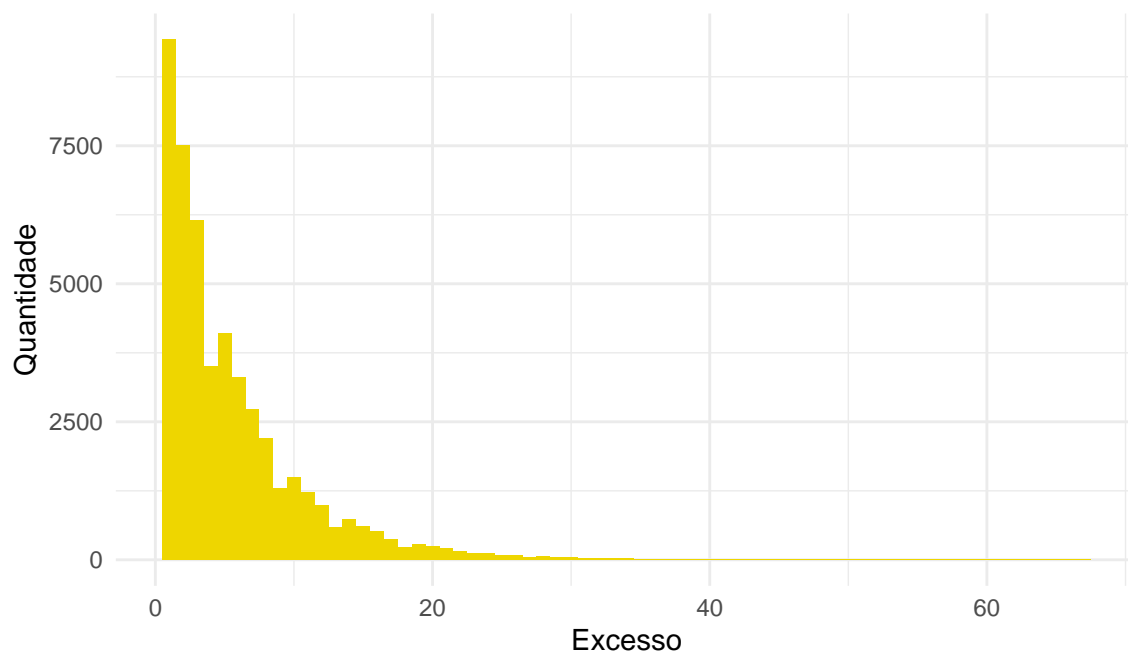


Figure 2: Distribuição de excesso na Alameda Parnaíba

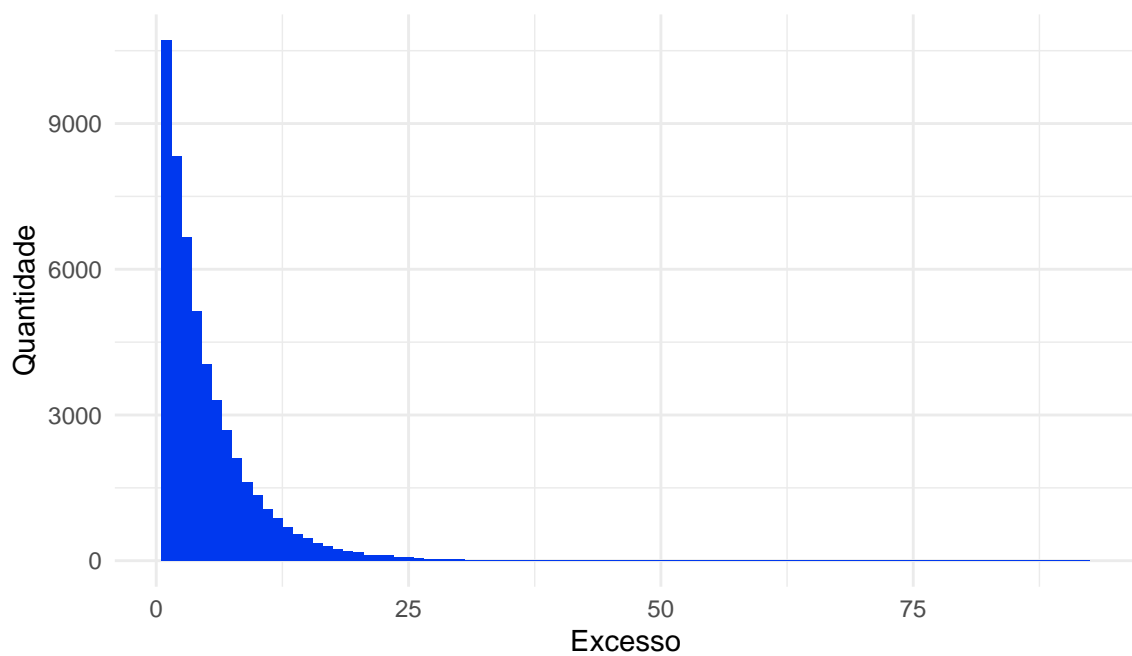


Figure 3: Distribuição de excesso na Av. Raul Lopes

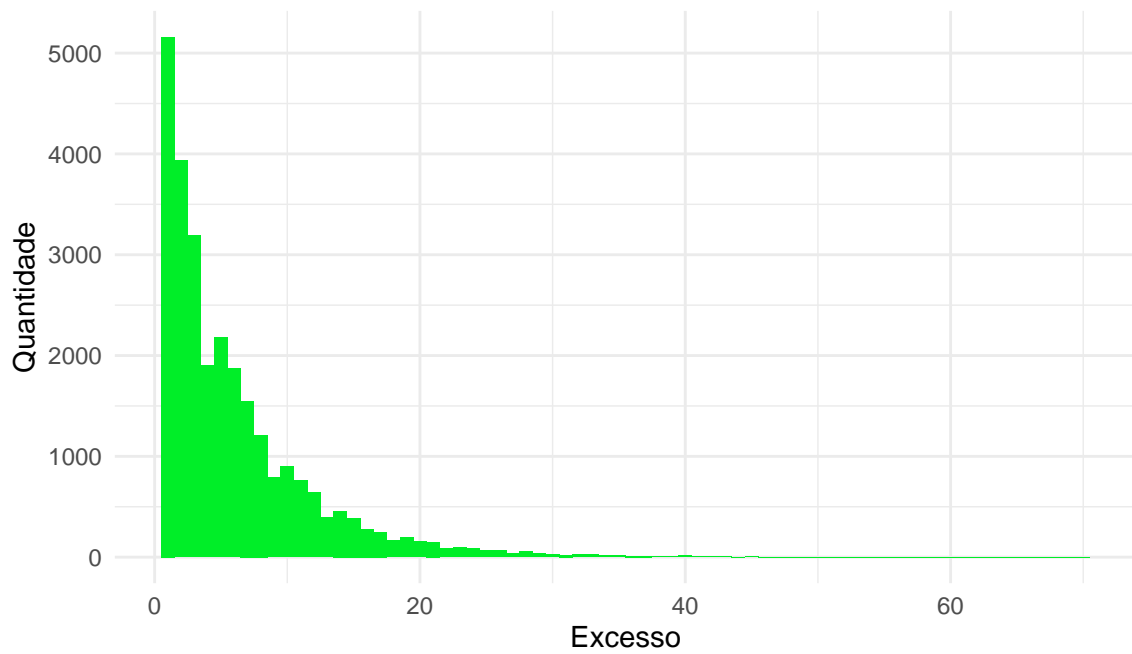


Figure 4: Distribuição de excesso na Av. Maranhão

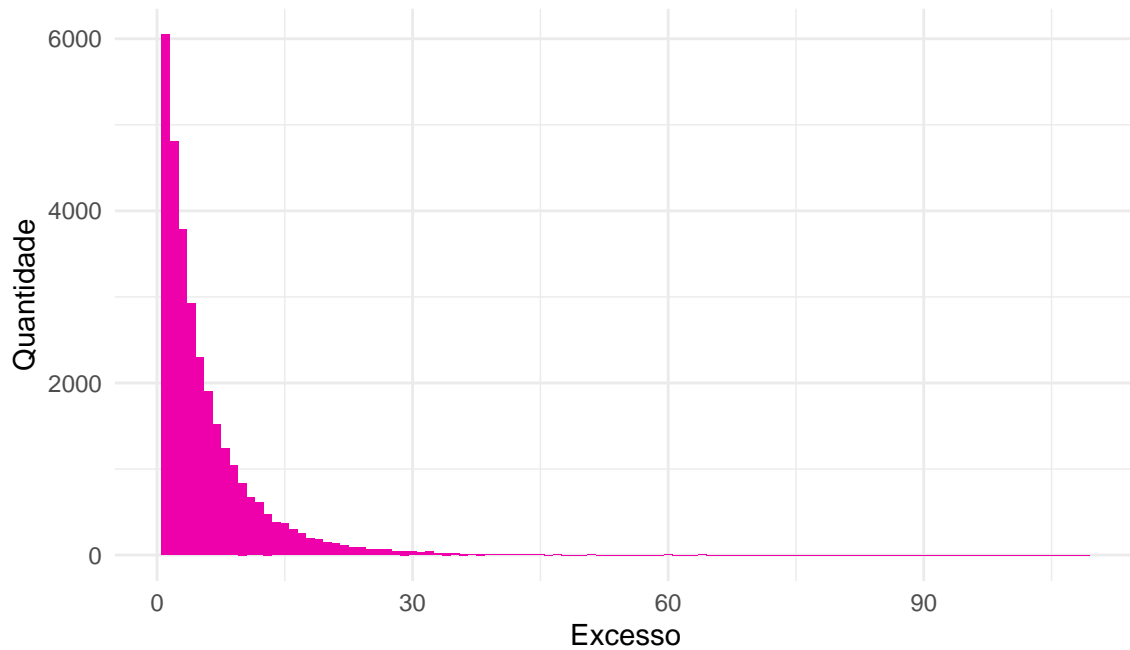


Figure 5: Distribuição de excesso na Av. B. Castelo Branco

#### 4.1.4 Alguns testes

##### 4.1.4.1 Teste de Corridas

- **Alameda Parnaíba**

```
##
## Runs Test for Randomness
##
## data: Alameda$Excesso
## z = -2.8499, runs = 23806, m = 26564, n = 22083, p-value = 0.004373
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##          4
```

- **Av. Raul Lopes**

```
##
## Runs Test for Randomness
##
## data: Shopping$Excesso
```

```
## z = -2.6169, runs = 24539, m = 30826, n = 20779, p-value = 0.008872
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##          4
```

- **Av. Maranhão**

```
##
## Runs Test for Randomness
##
## data: Maranhao$Excesso
## z = -3.1474, runs = 13355, m = 14191, n = 13082, p-value = 0.001647
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##          4
```

- **Av. B. de Castelo Branco**

```
##
## Runs Test for Randomness
##
## data: Barao$Excesso
## z = -2.1968, runs = 15009, m = 17554, n = 13400, p-value = 0.02803
## alternative hypothesis: true number of runs is not equal the expected number
## sample estimates:
## median(x)
##          4
```

#### 4.1.4.2 Teste de Kruskal White

```
## # A tibble: 1 x 4
##   statistic p.value parameter method
##   <dbl>     <dbl>     <int> <chr>
## 1      435. 5.45e-94         3 Kruskal-Wallis rank sum test
```



#### 4.1.4.3 Pós-Teste

| .y.      | group1 | group2 | n1    | n2    | statistic  | p         | p.adj | p.adj.signif |
|----------|--------|--------|-------|-------|------------|-----------|-------|--------------|
| excessos | 1      | 2      | 48647 | 51605 | -13.916882 | 0.0000000 | 0     | ****         |
| excessos | 1      | 3      | 48647 | 27273 | 7.843274   | 0.0000000 | 0     | ****         |
| excessos | 1      | 4      | 48647 | 30954 | -0.867789  | 0.3855098 | 1     | ns           |
| excessos | 2      | 3      | 51605 | 27273 | 19.672889  | 0.0000000 | 0     | ****         |
| excessos | 2      | 4      | 51605 | 30954 | 11.355477  | 0.0000000 | 0     | ****         |
| excessos | 3      | 4      | 27273 | 30954 | -7.903749  | 0.0000000 | 0     | ****         |

#### • Gráfico comparativo

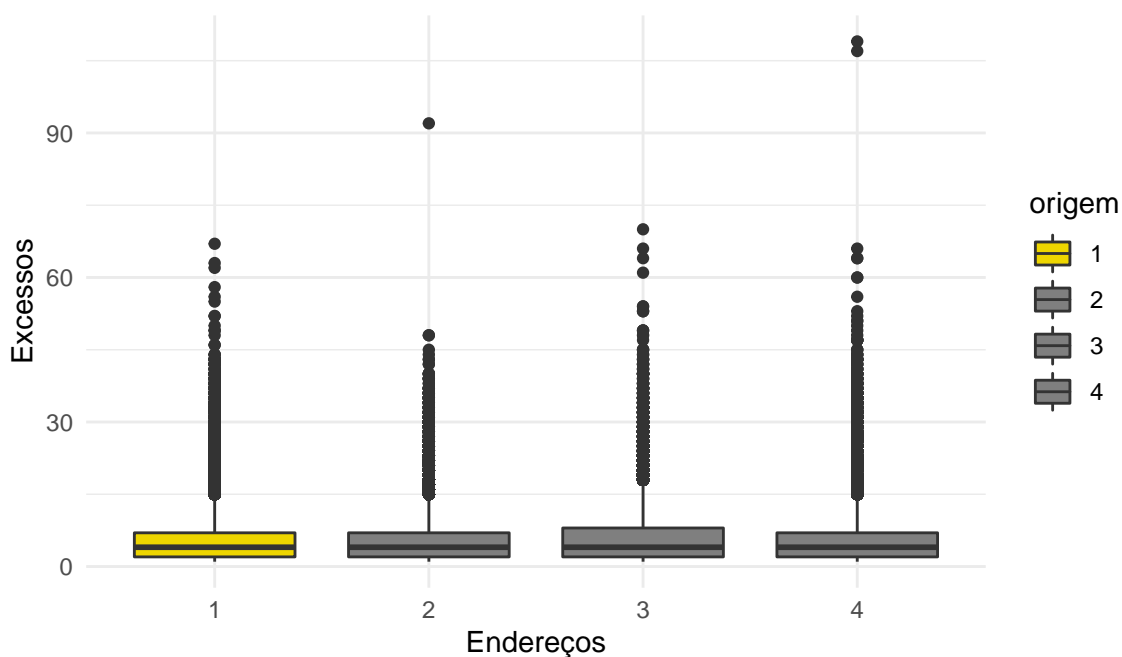


Figure 6: Boxplot comparativo das distribuições dos excessos

#### 4.1.5 Arrecadação possível

Finalizando esta análise no ponto que consiste o objetivo final do trabalho, a arrecadação. A idéia do trabalho é utilizar as informações detalhadas acima para que seja possível realizar uma previsão de arrecadação e, desta forma, aperfeiçoar a forma de administração das receitas e despesas públicas. Abaixo, temos a arrecadação possível se todas as multas fossem pagas com o valor integral.

| Local                    | Arrecadação Possível |
|--------------------------|----------------------|
| 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        |

## 5 Distribuição Pareto Generalizada (GPD)

A distribuição Pareto Generalizada (GPD) analisa a distribuição dos excessos de acordo com um limiar determinado. Esse formato de análise é mais eficaz pois evita a perda de informações que uma análise em períodos (ou blocos) pode gerar, afetando, principalmente, pesquisas com um grande volume de dados.

A distribuição Pareto Generalizada foi desenvolvida por Pickands [1975] é baseada no seguinte teorema:

*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 parâmetros são  $\xi$ ,  $\sigma$  e  $\mu$ , que representam, respectivamente, a forma, a escala e o limiar da distribuição.

A função de densidade da distribuição GPD é dada por:

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

onde  $x - \mu > 0$  para  $\xi \geq 0$  e  $0 \leq x - \mu < -\frac{\mu}{\xi}$  para  $\xi > 0$ .

A distribuição GPD possui as seguintes características em relação aos parâmetros:

$$E(X) = \frac{\sigma}{1-\xi}; \xi < 1; Md(x) = \frac{\sigma(2\xi - 1)}{\xi}; V(X) = \frac{\sigma^2}{(1-\xi)^2(1-2\xi)}$$

Justificando o uso da GPD Pickands (1975) e Davidson e Smith (1990) demonstram as propriedades e provam que GPD é única que satisfaz estas propriedades. “Por exemplo, estabilidade do limiar, ou seja, se  $Y$  possui distribuição GPD, e se  $\mu > 0$ , então a distribuição de  $P(Y - \mu | Y > \mu)$  também possui distribuição GPD”. *prof fernando*

Em valores extremos, além de encontrar a estimativa dos parâmetros do modelo, também é

muito importante encontrar uma forma para determinar os quantis altos, acima do limiar, de tal forma que se  $X$  possui distribuição GPD, é importante saber com qual probabilidade ocorre um evento maior ou igual a  $q$ , ou seja,  $P(X > q) = 1 - q$ .

Com os cálculos destes quantis, podemos realizar previsões com os dados de autos de excesso velocidade de trânsito de Teresina nos próximos anos, considerando uma manutenção da estrutura dos radares e, incluindo novos endereços de medição, uma previsão de quantos autos poderão ser registrados. Outra variável que pode ser analisada consiste no excessos de velocidade, determinando possíveis valores máximos e a previsão da quantidade de infrações.

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

## 5.1 Domínio da atração

As distribuições de valores extremos são obtidas como distribuição limite ( $n \rightarrow \infty$ ) do máximo de um conjunto de variáveis aleatórias (v.a.s) independente e identicamente distribuídas (i.i.d) e são unicamente determinadas, a menos de transformações afins. O teorema de Fisher-Tippett implica que se  $F_x^n(C_n x + d_n)$  converge para uma distribuição degenerada quando ( $n \rightarrow \infty$ ), para certas constantes  $C_n > 0$ ,  $d_n \in \mathbb{R}$ , então

$$|F_x^n(x) - H(\frac{x - d_n}{c_n})| \rightarrow 0, n \rightarrow \infty$$

para alguma distribuição  $H$ . A coleção das distribuições  $F_x$  tais que os respectivos máximos possuem a mesma distribuição limite é chamada de domínio de atração.

*Definição 2.1.2* Se (função) se verifica dizemos que  $F_x$  pertence ao domínio de atração do máximo da distribuição de valores extremos  $H$ . Notação:  $F_x \in MDA(H)$

Existem 3 casos possíveis para as distribuições limites das excedências de um limiar. Para domínio do tipo I ( $\gamma = 0$ ), a distribuição se torna

$$H(y) = 1 - e^{-\frac{y}{\sigma}}, y > 0$$

sendo assim, o domínio, uma distribuição Exponencial com parâmetro  $\frac{1}{\sigma}$ . Para o domínio tipo II ( $\gamma > 0$ ), a distribuição limite será a distribuição de Pareto. Já para o domínio tipo III ( $\gamma < 0$ ), quando  $\sigma = -\frac{1}{\gamma}$ , a distribuição limite será uma Beta e quando  $\sigma \neq -\frac{1}{\gamma}$ , a distribuição limite será uma Beta reescalada com suporte em  $(0, \frac{\sigma}{\gamma})$ .

## 5.2 Determinação do Limiar

A análise via GPD exige um cuidado inicial pois é preciso determinar um limiar para os dados. O valor escolhido pode ser determinado pelo pesquisador, entretanto, correm-se riscos que podem influenciar os cálculos, resultando em análises incorretas.

A escolha de um limiar " $\mu$ " muito alto implica em um número muito pequeno de observações resultando em estimadores com grande variabilidade. Um limiar muito pequeno resulta na violação do Teorema de Pickands (1945), modelando de forma errada os valores com limiar baixo, dessa forma, não se garante a convergência dos excessos  $Y$  para a família da GPD, levando a um vício alto.

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 (MRL < *Mena 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  $\mu$ , e  $x_{max}$  é o valor mais elevado das observações.

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

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

Se o modelo é adequado a partir de  $\mu_0$  o gráfico apresentará um comportamento linear a partir de  $u$ . Um problema recorrente com a utilização desse gráfico é que o limiar pode limitar o número de excessos devido a escolha de limiar muito alto.

Outra técnica gráfica utilizada é Dipersion Index Plot (DIP), baseado em Cunnane [1979] (Citado por Nascimento [2012]), que diz que, o número de excessos sobre um limiar alto em um determinado

período (geralmente meses ou anos), pode ser distribuído através de um processo de Poisson. Assim, a razão entre a variância e a média é igual a 1. Assim, pode-se fazer um gráfico

$$\left\{ \left( \mu, \frac{Var(Y)}{E(Y)} \right) : \mu < x_{max} \right\}$$

### 5.3 Estimação da GPD

Após determinar o limiar, a estimação dos parâmetros da distribuição GPD podem ser estimados por vários métodos, entre eles, tem-se o da máxima verossimilhança, que iremos trabalhar. Existem outros métodos como de momentos proposto por Smith [1987] (citado por Mendes [2004]) e o métodos dos momentos ponderados (Singh e Guo [1995], (citado por Mendes [2004])), em que a eficiência de cada método depende da situação estudada.

Os estimadores de máxima verossimilhança (EMV) que maximizam a função de log-verossimilhança, quando  $\xi \neq 0$ , é dado por

$$l(\mu, \xi) = -n_u \log(\sigma) - \left(1 + \frac{1}{\xi} \sum_{i=1}^{n_u} \log\left(1 + \xi \frac{y_i}{\sigma}\right)\right)$$

definida em  $(1 + \xi \frac{y_i}{\sigma}) > 0$ , para todo  $i = 1, 2, \dots, N_u$ . No caso particular onde  $\xi = 0$ , a a log-verossimilhança é dada por

$$l(\sigma) = -n_u \log(\sigma) - \sum_{i=1}^{n_u} \left(\frac{y_i}{\sigma}\right)$$

No caso de  $\xi = 0$ , a maximização dos parâmetros não pode ser obtida analiticamente, sendo necessárias técnicas numéricas de maximização.

## 6 Inferência Bayesiana

## 7 Modelos dinâmicos para Valores Extremos

### 7.1 Alameda Parnaíba

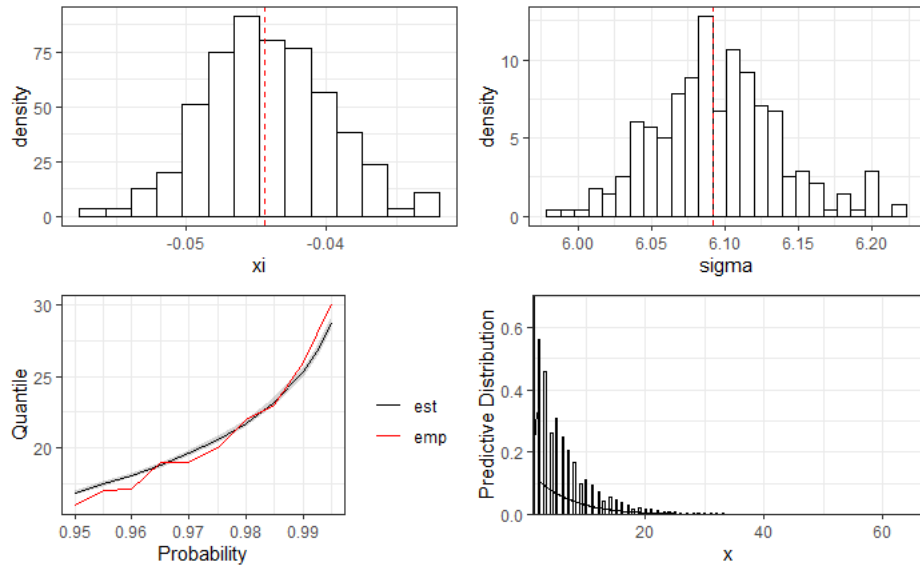


Figure 7: Alameda Parnaíba - Função  $fgcpd$

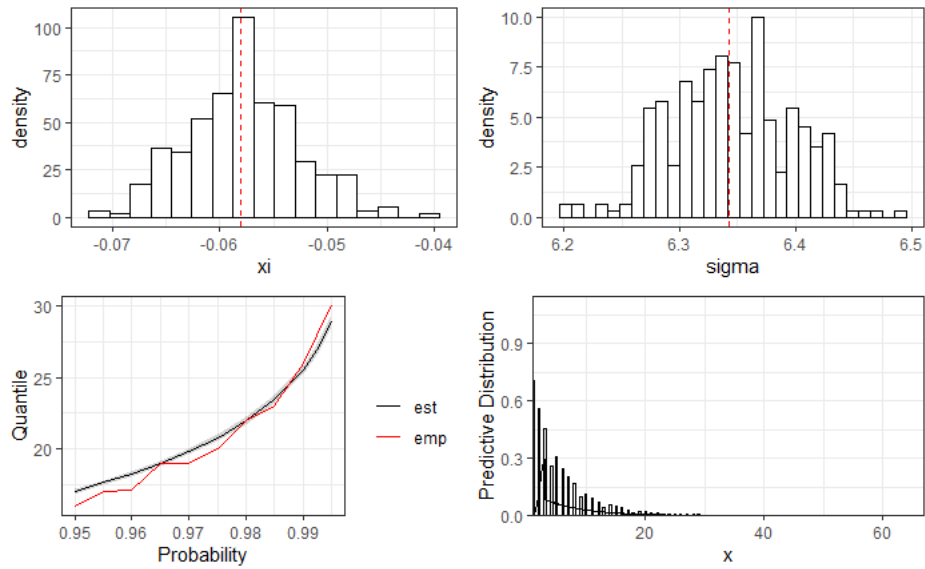


Figure 8: Alameda Parnaíba - Função  $fmcpd$

| Ajuste              | DIC      | BIC      |
|---------------------|----------|----------|
| Alameda com $fgcpd$ | 247310.8 | 247351.3 |
| Alameda com $fmcpd$ | 215577.4 | 215641.7 |

### 7.2 Av. Maranhão



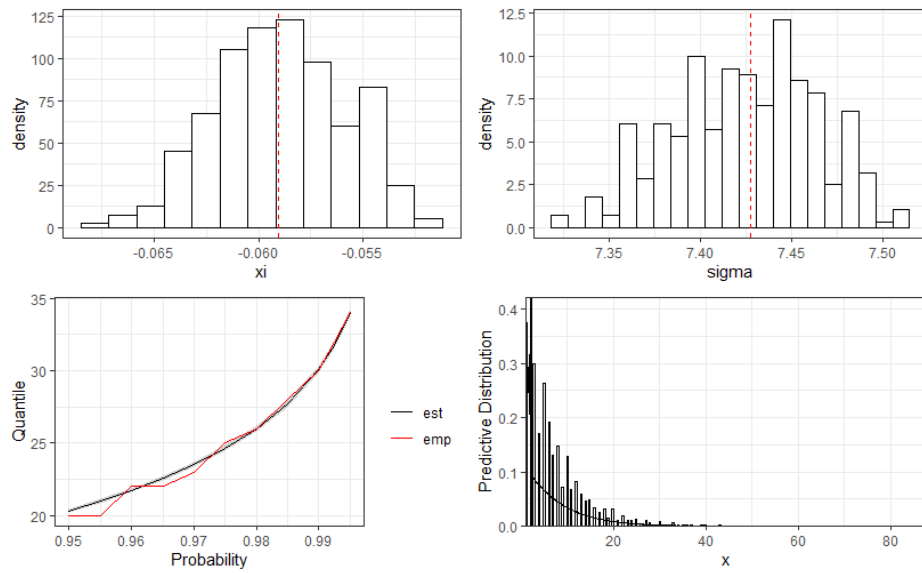


Figure 9: Av. Maranhão - Função fggpd

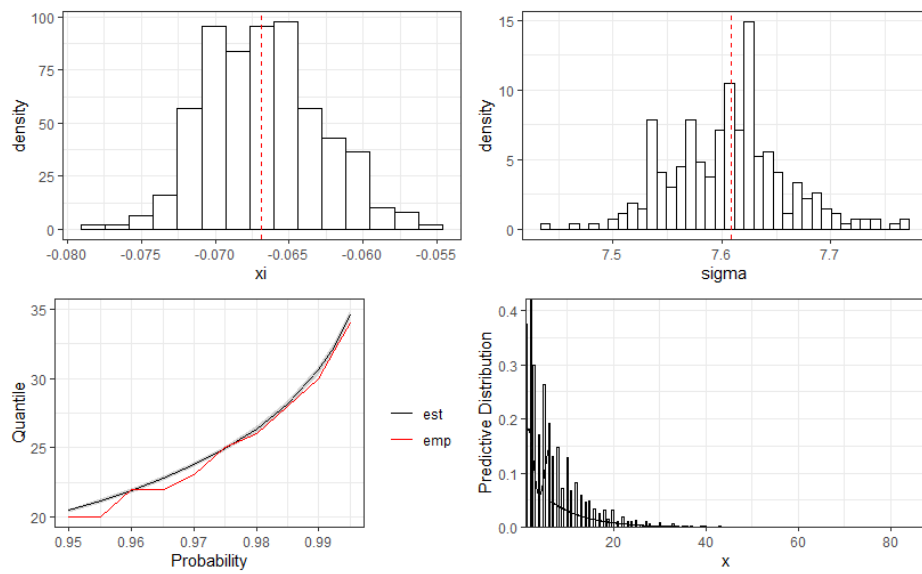


Figure 10: Av. Maranhão - Função fmgpd

| Ajuste                 | DIC      | BIC      |
|------------------------|----------|----------|
| Av. Maranhao com fgdpd | 480153.4 | 480201.7 |
| Av. Maranhão com fmgpd | 485451.3 | 485518.7 |

### 7.3 Av. Raul Lopes - Shopping

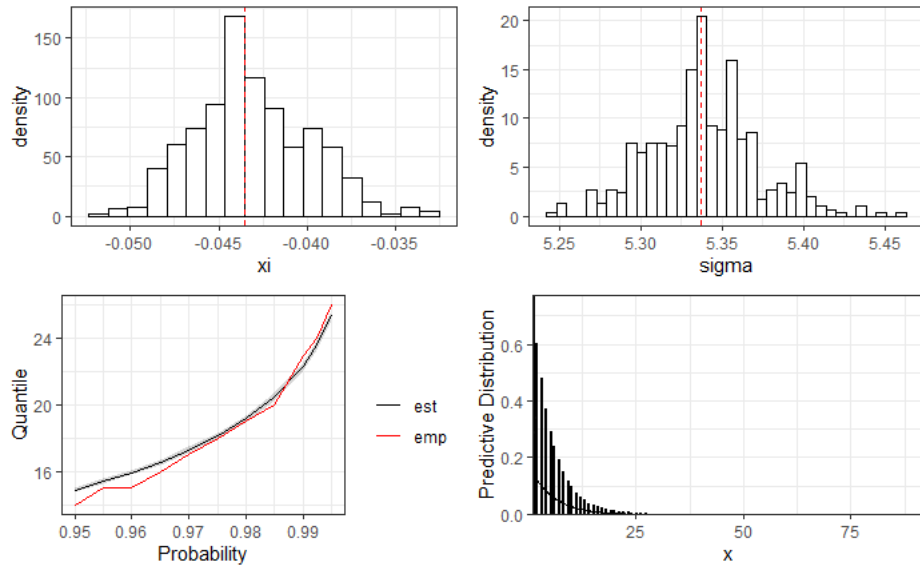


Figure 11: Av. Raul Lopes - Função fgdpd

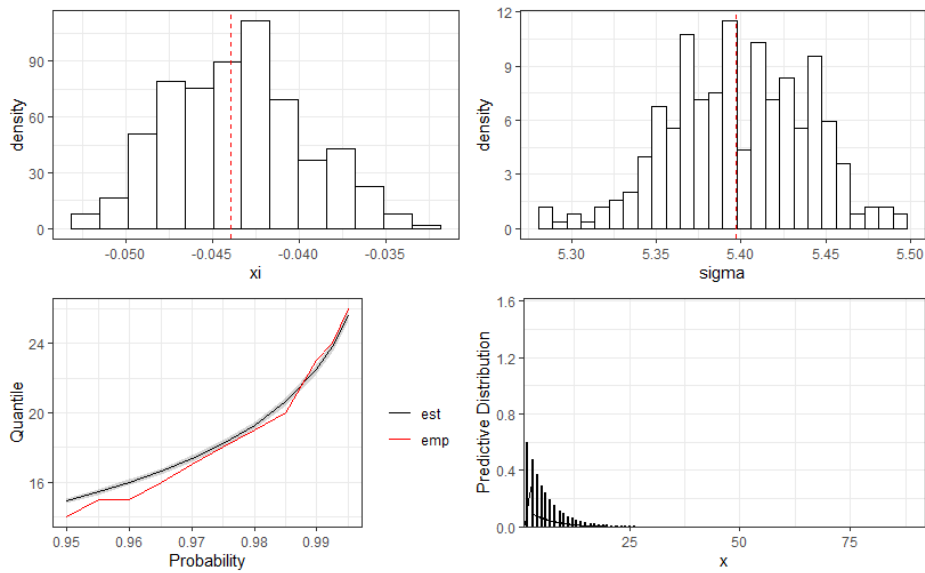


Figure 12: Av. Raul Lopes - Função fmgpd

| Ajuste                   | DIC      | BIC      |
|--------------------------|----------|----------|
| Av. Raul Lopes com fggpd | 253692.7 | 253754.8 |
| Av. Rul Lopes com fmgpd  | 211497.7 | 211567.5 |

## 7.4 Av. dos Ipês

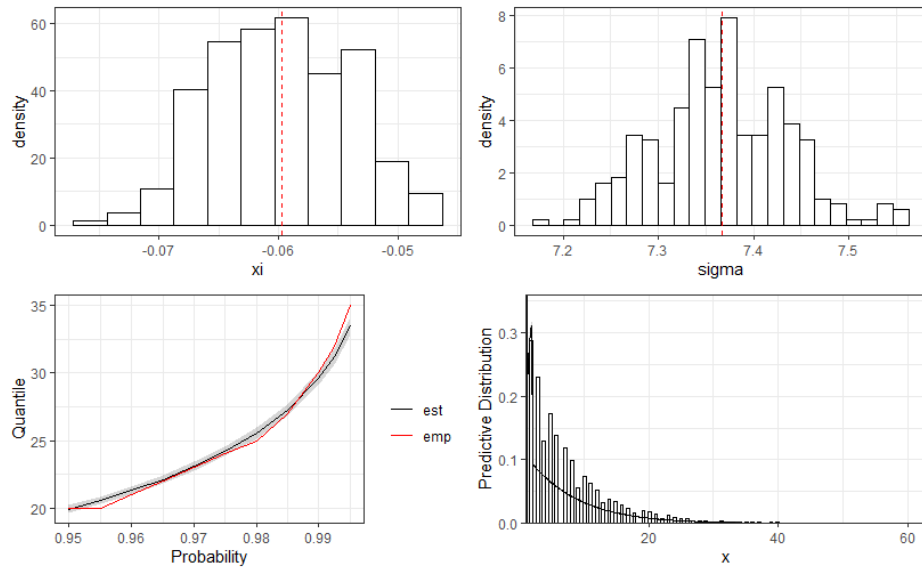


Figure 13: Av. dos Ipês - Função fggpd

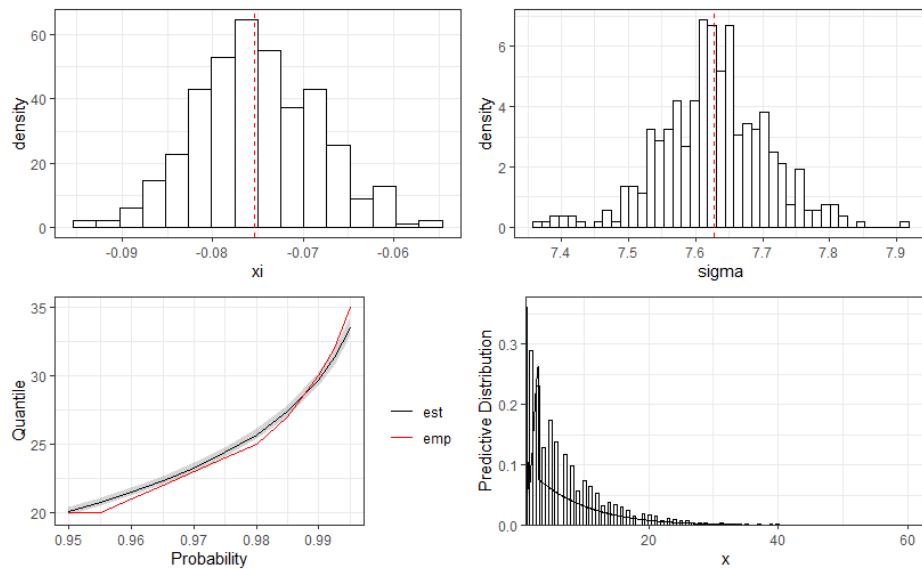


Figure 14: Av. dos Ipês - Função fmgpd

| Ajuste                      | DIC      | BIC      |
|-----------------------------|----------|----------|
| Av. Av. dos Ipês com fggpdp | 136611.9 | 136654.4 |
| Av. dos Ipês com fmgpdp     | 131887.8 | 131672.3 |

## 7.5 Av. Av. Jose Francisco de Almeida Neto

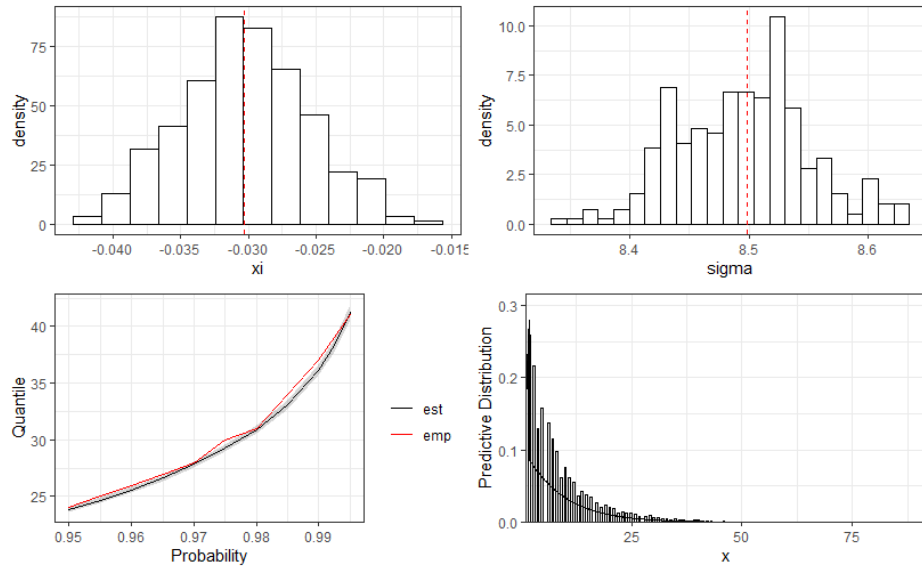


Figure 15: Av. Jose - Função fggpdp

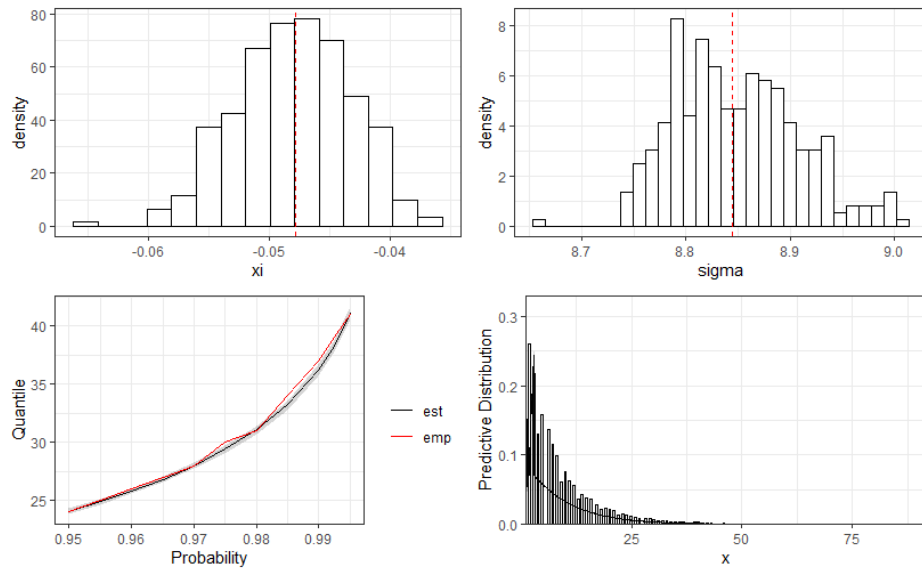


Figure 16: Av. Jose - Função fmgpdp

| Ajuste                 | DIC      | BIC      |
|------------------------|----------|----------|
| Av. Av. Jose com fgcpd | 313760.1 | 303626.5 |
| Av. Jose com fmgcpd    | 313806.8 | 303289.6 |

## 7.6 Av. Barão de Castelo Branco

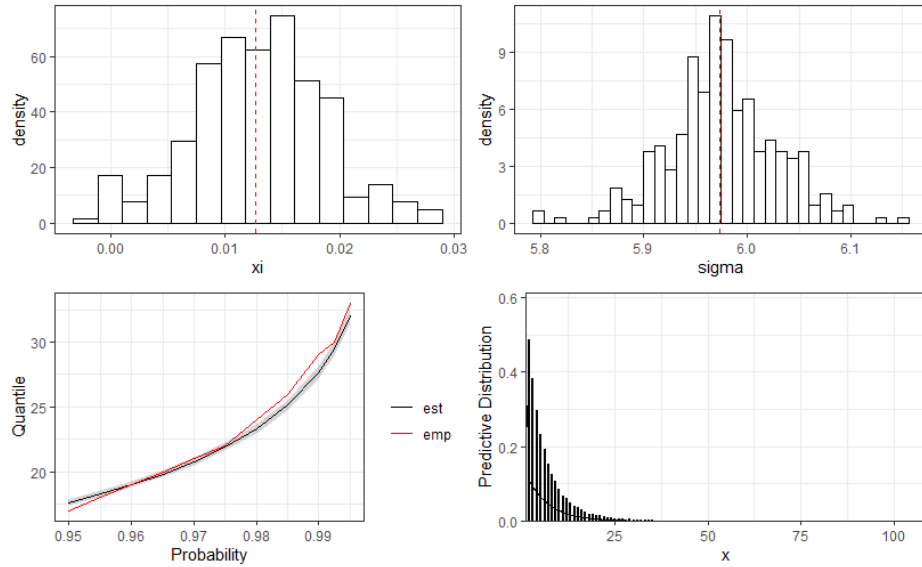


Figure 17: Av. Barão - Função fgcpd

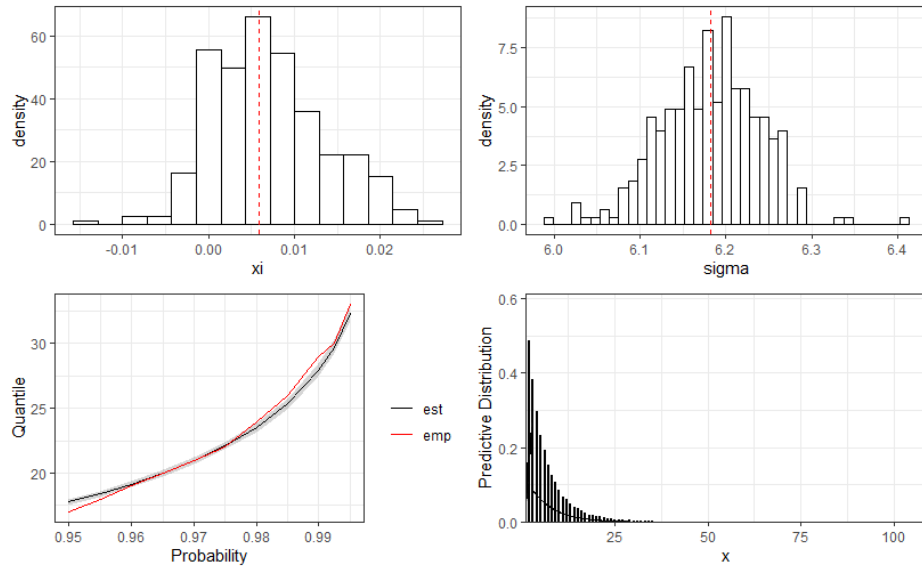


Figure 18: Av. Barão - Função fmgcpd

| Ajuste              | DIC      | BIC      |
|---------------------|----------|----------|
| Av. Barão com fgdpd | 160774.3 | 160813.9 |
| Av. Barão com fmgpd | 153049.2 | 152921.9 |

## 8 Conclusão

## 9 Referências Bibliográficas

@book{AREL:03, author={Nascimento, F. F.}, title={Modelos probabilísticos para dados extremos: teoria e aplicações}, publisher={EDUFPI}, year={2012.} }