

# Advanced Econometrics

Mestrado Profissional em Economia

Inspere Instituto de Educação e Pesquisa

Prof. Hedibert Freitas Lopes

## Lista de Exercícios 1

Data de entrega via blackboard: 02 de novembro de 2026 (até 19h30)

Cada grupo de trabalho pode conter até 4 estudantes

## Dados

A tabela abaixo apresenta o número mensal de mortes de condutores de motocicletas no município de São Paulo entre janeiro de 2015 e agosto de 2026, num total de  $n = 140$  meses. Denote por  $y_t$ ,  $t = 1, \dots, n$ , o número de mortes no mês  $t$  (com  $t = 1$  correspondendo a janeiro de 2015). Fonte: INFOSIGA-SP/DETRAN-SP (microdados de vítimas, versão de 16/09/2026; acesso em 20/10/2026).

| Ano  | Jan | Fev | Mar | Abr | Mai | Jun | Jul | Ago | Set | Out | Nov | Dez |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2015 | 21  | 28  | 33  | 38  | 35  | 39  | 27  | 24  | 24  | 26  | 22  | 38  |
| 2016 | 23  | 26  | 30  | 36  | 22  | 21  | 16  | 28  | 23  | 20  | 22  | 39  |
| 2017 | 16  | 21  | 27  | 31  | 25  | 31  | 31  | 16  | 20  | 24  | 22  | 32  |
| 2018 | 26  | 23  | 22  | 40  | 23  | 29  | 32  | 24  | 28  | 24  | 25  | 51  |
| 2019 | 32  | 19  | 19  | 16  | 26  | 22  | 24  | 35  | 29  | 34  | 14  | 27  |
| 2020 | 19  | 19  | 38  | 17  | 40  | 13  | 21  | 26  | 32  | 16  | 24  | 27  |
| 2021 | 33  | 23  | 19  | 18  | 25  | 28  | 30  | 28  | 34  | 31  | 27  | 17  |
| 2022 | 29  | 19  | 38  | 32  | 40  | 38  | 30  | 35  | 40  | 46  | 32  | 37  |
| 2023 | 27  | 19  | 40  | 25  | 26  | 28  | 42  | 31  | 34  | 38  | 39  | 54  |
| 2024 | 38  | 28  | 43  | 42  | 38  | 50  | 44  | 38  | 41  | 44  | 28  | 56  |
| 2025 | 31  | 33  | 39  | 31  | 49  | 39  | 46  | 46  | 50  | 36  | 25  | 54  |
| 2026 | 31  | 34  | 41  | 40  | 43  | 50  | 48  | 47  |     |     |     |     |

Para facilitar, os dados em formato R:

```
y <- c(21,28,33,38,35,39,27,24,24,26,22,38, 23,26,30,36,22,21,16,28,23,20,22,39,
16,21,27,31,25,31,31,16,20,24,22,32, 26,23,22,40,23,29,32,24,28,24,25,51,
32,19,19,16,26,22,24,35,29,34,14,27, 19,19,38,17,40,13,21,26,32,16,24,27,
33,23,19,18,25,28,30,28,34,31,27,17, 29,19,38,32,40,38,30,35,40,46,32,37,
27,19,40,25,26,28,42,31,34,38,39,54, 38,28,43,42,38,50,44,38,41,44,28,56,
31,33,39,31,49,39,46,46,50,36,25,54, 31,34,41,40,43,50,48,47)
y <- ts(y, start = c(2015, 1), frequency = 12)
n <- length(y) # 140
```

## Exercícios

1. **Histograma.** Faça um histograma dos dados  $y_1, \dots, y_n$  e comente: forma da distribuição (simetria, assimetria, caudas), valores típicos, dispersão e possíveis valores atípicos. Reporte também média, mediana, variância e desvio-padrão amostrais.

2. **Modelo Gaussiano.** Assuma que  $y_1, \dots, y_n$  são i.i.d. com

$$y_t | \mu, \sigma^2 \sim N(\mu, \sigma^2), \quad t = 1, \dots, n.$$

- (a) Escreva a função de verossimilhança  $L(\mu, \sigma^2; y)$  e obtenha os estimadores de máxima verossimilhança  $\hat{\mu}$  e  $\hat{\sigma}^2$ .
- (b) Calcule as estimativas com os dados e um intervalo de confiança de 95% para  $\mu$ .
- (c) Sobreponha a densidade  $N(\hat{\mu}, \hat{\sigma}^2)$  ao histograma do item 1. Comente a qualidade do ajuste. O que o modelo Gaussiano “ignora” sobre a natureza destes dados?

**3. Modelo Poisson.** Como os dados são contagens, assuma agora que

$$y_t | \lambda \sim \text{Poisson}(\lambda), \quad p(y_t | \lambda) = \frac{e^{-\lambda} \lambda^{y_t}}{y_t!}, \quad t = 1, \dots, n,$$

independentes.

- (a) Mostre que o estimador de máxima verossimilhança de  $\lambda$  é  $\hat{\lambda} = \bar{y}$  e obtenha seu erro-padrão assintótico.
  - (b) Calcule  $\hat{\lambda}$  e um intervalo de confiança de 95% para  $\lambda$ .
  - (c) Compare as frequências observadas com as probabilidades  $p(y | \hat{\lambda})$  (por exemplo, sobrepondo-as ao histograma).
  - (d) Lembrando que, sob o modelo Poisson,  $E(y_t) = V(y_t) = \lambda$ , compare a média e a variância amostrais. Há evidência de *sobredispersão*? O que isso sugere sobre a adequação do modelo?
  - (e) Compare os ajustes Gaussiano e Poisson (por exemplo, via log-verossimilhança maximizada ou AIC).
- 4. Série temporal.** Faça o gráfico de  $y_t$  contra  $t$  (ou contra a data) e comente: nível, tendência, sazonalidade, mudanças de regime e meses atípicos. À luz deste gráfico, discuta a hipótese de que  $y_1, \dots, y_n$  são i.i.d., implícita nos itens 2 e 3.
- 5. Quebra estrutural.** O gráfico do item 4 sugere uma mudança na incidência de mortes mais ou menos no meio da série. Escolha (arbitrariamente, mas justificando a escolha) um ponto de corte  $\tau$  e divida a amostra em

$$\text{Parte 1: } y_1, \dots, y_\tau \quad \text{e} \quad \text{Parte 2: } y_{\tau+1}, \dots, y_n.$$

- (a) *Modelos normais.* Assuma  $y_t \sim N(\mu_1, \sigma_1^2)$  para  $t \leq \tau$  e  $y_t \sim N(\mu_2, \sigma_2^2)$  para  $t > \tau$ . Estime  $(\mu_1, \sigma_1^2)$  e  $(\mu_2, \sigma_2^2)$ .
  - (b) *Modelos Poisson.* Assuma  $y_t \sim \text{Poisson}(\lambda_1)$  para  $t \leq \tau$  e  $y_t \sim \text{Poisson}(\lambda_2)$  para  $t > \tau$ . Estime  $\lambda_1$  e  $\lambda_2$ .
  - (c) Construa intervalos de confiança de 95% e/ou teste as hipóteses  $H_0 : \mu_1 = \mu_2$  e  $H_0 : \lambda_1 = \lambda_2$  (por exemplo, via teste da razão de verossimilhanças comparando o modelo com quebra ao modelo sem quebra dos itens 2 e 3).
  - (d) Os resultados confirmam (ou não) a existência da quebra? Quão sensíveis são as conclusões à escolha de  $\tau$ ? Experimente pelo menos dois outros valores de  $\tau$ .
  - (e) Repita a comparação média versus variância do item 3(d) dentro de cada parte. A sobre-dispersão diminui? Interprete.
- 6. (Para os mais energizados) Taxa variando no tempo.** Mantenha o modelo Poisson com taxa constante na Parte 1,

$$y_t \sim \text{Poisson}(\lambda), \quad t = 1, \dots, \tau,$$

mas, para a Parte 2, assuma uma taxa linear no tempo:

$$y_{\tau+s} \sim \text{Poisson}(\theta_0 + \theta_1 s), \quad s = 1, 2, \dots, n - \tau,$$

onde  $s$  conta os meses após o ponto de corte. O parâmetro  $\theta_1$  mede a importância do tempo: se  $\theta_1 > 0$ , a taxa de mortes é crescente.

- (a) Escreva a log-verossimilhança de  $(\theta_0, \theta_1)$  para a Parte 2. Que restrição os parâmetros devem satisfazer para que o modelo esteja bem definido?
- (b) Obtenha numericamente os estimadores de máxima verossimilhança  $(\hat{\theta}_0, \hat{\theta}_1)$  e seus erros-padrão (via a matriz de informação observada). *Dica:* use `optim()` com `hessian=TRUE`, ou `glm(y ~ s, family=poisson(link="identity"))`.
- (c) Teste  $H_0 : \theta_1 = 0$  (taxa constante na Parte 2) contra  $H_1 : \theta_1 > 0$ . Compare o modelo com o do item 5(b) via razão de verossimilhanças e/ou AIC.
- (d) Ajuste também  $y_t \sim \text{Poisson}(\lambda_0 + \lambda_1 t)$  apenas para a Parte 1. A hipótese de taxa constante antes do corte é razoável?
- (e) Sobreponha ao gráfico do item 4 as taxas estimadas ( $\hat{\lambda}$  na Parte 1 e  $\hat{\theta}_0 + \hat{\theta}_1 s$  na Parte 2) e comente.
- (f) (*Opcional*) Discuta vantagens e desvantagens da especificação log-linear  $\log \lambda_s = \beta_0 + \beta_1 s$  em relação à especificação linear acima.

---

**Observações.** Entregue um relatório com gráficos, tabelas de estimativas e comentários, acompanhado do código utilizado (R, Python ou outra linguagem de sua preferência).