Exercícios: revisão A1

Disciplina: Modelagem Estatística
Instrutor: Luiz Max Carvalho
Monitor: Ezequiel Braga

Abril/2025

Questões

1. **(Weighted regression)** Em muitas situações faz sentido ponderar alguns pontos de dados mais do que outros ao ajustar um modelo de regressão, o que é feito usando mínimos quadrados ponderados. Para isso, considere que estamos interessados em minimizar $\sum_{i=1}^{n} w_i(y_i - \boldsymbol{X}_i\boldsymbol{\beta})^2$, para $\boldsymbol{X} \in \mathbb{R}^{n\times(p+1)}$, com a primeira coluna de $1's$; $\boldsymbol{Y} \in \mathbb{R}^n$; $\boldsymbol{\beta} \in \mathbb{R}^{p+1}$; e $w_i > 0,\ i = 1, \dots, n$.

   (a) Escreva o problema de minimização na forma matricial, usando $\boldsymbol{X}$, $\boldsymbol{Y}$ e $\boldsymbol{W} = \text{diag}(\boldsymbol{w})$, com $\boldsymbol{w} = (w_1, w_2, \dots, w_n)^T$.

   (b) Determine o estimador de mínimos quadrados para $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}_{\text{wls}}$.

   (c) O modelo acima pode ser usado quando a premissa de homocedasticidade é violada, isto é, os erros não possuem a mesma variância. Assim, suponha que $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, com $\boldsymbol{\epsilon} \sim \text{MVN}(0, \boldsymbol{V})$ e $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ tal que $\boldsymbol{V}_{ij} = \begin{cases} \sigma_i^2, & \text{se } i = j; \\ 0, & \text{c.c.} \end{cases}$.

      i. Calcule a log-verossimilhança do modelo.

      ii. Determine o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ e deduza a matriz de pesos $\boldsymbol{W}$.

2. O modelo linear (de regressão) é um dos cavalos de batalha da Estatística, sendo aplicado em problemas de Finanças, Medicina e Engenharia. Vamos agora estudar como utilizar as propriedades deste modelo para desenhar experimentos com garantias matemáticas de desempenho e obter estimadores de quantidades de interesse.

   a) Uma prática comum em regressão é a de **centrar** a variável independente (covariável), isto é subtrair a média; isto facilita a interpretação do intercepto e também simplifica alguns cálculos importantes. Mostre que no caso com a covariável centrada, $\hat{\beta}_0$ e $\hat{\beta}_1$ são independentes;

   b) Mais uma vez considerando o caso centrado, mostre como obter o número de observações n que faz com que a variância do estimador de máxima verossimilhança do intercepto seja menor que $v > 0$;

c) Mostre como obter um estimador não-viesado da quantidade $\theta = a\beta_0 + b\beta_1 + c$, com $a, b, c \neq 0$, e encontre o seu erro quadrático médio.

d) Quando $x_{\text{pred}} = \bar{x}$, mostre como obter o número de observações n necessário para que o intervalo de predição de $100(1 - \alpha_0)\%$ para a variável-resposta (Y) tenha largura menor ou igual a $l > 0$ com probabilidade pelo menos γ.

   Dicas:(i) A expressão dependerá *também* da variância dos resíduos, σ^2 e (ii) Você não precisa calcular n, apenas mostrar o procedimento para obtê-lo.

3. **(Regressão polinomial)** Para $i = 1, 2, \ldots, n$, considere o modelo

$$Y_i = aX_i + bX_i^2 + cX_i^3 + d + \varepsilon_i, \tag{1}$$

onde ε_i são variáveis aleatórias com média 0 e variância constante e X_i são constantes observadas.

a) O modelo em questão é linear?

b) Como você estimaria os parâmetros desse modelo? Escreva a função de perda.

c) É possível obter um estimador não-viesado de $\theta = (a, b, c, d)$?

d) Qual a interpretação de d? E a de b?

e) (Desafio) Mostre que o estimador de mínimos quadrados está bem definido se e somente se todos os X_i são distintos.

4. **(Família exponencial)** Considere as distribuições de probabilidade abaixo. Usando as propriedades da família exponencial, para cada uma, mostre:

i) A estatística suficiente e sua esperança

ii) O estimador de máxima verossimilhança.

(a) `Pareto`. Para $y > 1$ e $\theta > 0$ a densidade vale

$$f(y; \theta) = \theta y^{-\theta-1}.$$

(b) `Binomial negativa`. Para $y \in \mathbb{N} \cup \{0\}$ e $\theta, \phi > 0$ a função de massa vale

$$f(y; \theta) = \binom{y + \phi - 1}{y} \left(\frac{\theta}{\theta + \phi}\right)^y \left(\frac{\phi}{\theta + \phi}\right)^\phi,$$

com ϕ fixo e conhecido.