

Exercícios: revisão A1

Disciplina: Modelagem Estatística

Instrutor: Luiz Max Carvalho

Monitores: Eduardo Adame & Ezequiel Braga

Abril/2024

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 $\sum_{i=1}^n w_i (y_i - \mathbf{X}_i \boldsymbol{\beta})^2$, para $\mathbf{X} \in \mathbb{R}^{n \times (p+1)}$, com a primeira coluna de 1's; $\mathbf{Y} \in \mathbb{R}^n$; $\boldsymbol{\beta} \in \mathbb{R}^{p+1}$; e $w_i > 0$, $i = 1, \dots, n$.
 - (a) Escreva o problema de minimização na forma matricial, usando \mathbf{X} , \mathbf{Y} e $\mathbf{W} = \text{diag}(\mathbf{w})$, com $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$.
 - (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 $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, com $\boldsymbol{\epsilon} \sim \text{MVN}(0, \mathbf{V})$ e $\mathbf{V} \in \mathbb{R}^{n \times n}$ tal que $\mathbf{V}_{ij} = \begin{cases} \sigma_i^2, & \text{se } i = j; \\ 0, & \text{c.c.} \end{cases}$.
 - i. Calcule a log-verossimilhança do modelo.
 - ii. Determine o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ e deduza a matriz de pesos \mathbf{W} .
2. Seja $\mathbf{Y} = \{Y_1, \dots, Y_n\}$ observações do tempo de falha de um certo equipamento associadas à matriz de covariáveis $\mathbf{X} \in \mathbb{R}^{n \times p}$. Considere o seguinte modelo:

$$Y_i | X_i \sim \text{Weibull}(\alpha, \theta_i),$$

com α conhecido.

Dica: Se $Y \sim \text{Weibull}(\alpha, \theta)$, então sua p.d.f. é dada por

$$f_Y(y; \alpha, \theta) = \frac{\alpha y^{\alpha-1}}{\theta^\alpha} \exp \left[- \left(\frac{y}{\theta} \right)^\alpha \right].$$

Além disso $E[Y^r] = \theta^r \Gamma(1 + r/\alpha)$.

- (a) Para α conhecido, mostre que a distribuição de Weibull pertence à família exponencial.
- (b) Calcule a log-verossimilhança do modelo.
- (c) Para calcular $\hat{\beta}$, precisamos resolver a equação $\nabla l(\beta) = 0$. Porém, muitas vezes não existe solução analítica e precisamos usar algum método numérico. Neste caso, podemos utilizar Newton Raphson, que consiste em fazer atualizações da forma

$$\beta^{(m+1)} = \beta^{(m)} - \mathcal{H}^{-1}(\beta^{(m)}) \nabla l(\beta^{(m)}),$$

para algum chute inicial $\beta^{(0)}$. (\mathcal{H} denota a matriz hessiana). Calcule $\nabla l(\beta^{(m)})$ e $\mathcal{H}(\beta^{(m)})$.

3. 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.