Exercícios: revisão A2

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

Junho/2025

Questões

1. **(GLM com Weibull)** Seja $\boldsymbol{Y} = \{Y_1, \dots, Y_n\}$ observações do tempo de falha de um certo equipamento associadas à matriz de covariáveis $\boldsymbol{X} \in \mathbb{R}^{n \times p}$. Considere o seguinte modelo:

$$Y_i | X_i \sim \text{Weibull}(\alpha, \theta_i),$$
$$\mathbb{E}[Y_i^{\alpha} | X_i] = g^{-1}(X_i \boldsymbol{\beta}),$$

   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 familia exponencial.

   (b) Mostre como escolher g e calcule a log-verossimilhança do modelo.

   (c) Para calcular $\hat{\boldsymbol{\beta}}$, precisamos resolver a equação $\nabla l(\boldsymbol{\beta}) = 0$. 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

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

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

2. **(LMM):** Seja $\boldsymbol{Y}_i = (Y_{i1}, \dots, Y_{1d})^{\top}$ vetores independentes correspondentes a observações para cada grupo $i = 1, \dots, n$. Considere matrizes de covariáveis $\boldsymbol{X}_i$ e $\boldsymbol{Z}_i$. Assuma que

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{U}_i + \boldsymbol{\varepsilon}_i,$$

onde $\boldsymbol{\beta}$ é um vetor p de efeitos fixos, $\boldsymbol{U}_i \sim \text{Normal}(0, \boldsymbol{\Sigma})$ é um vetor $q \times 1$ de efeitos aleatórios, e $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{id})^\top \sim \text{Normal}(0, \boldsymbol{R})$.

(a) Podemos escrever $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{U} + \boldsymbol{\varepsilon}$. Escreva explicitamente $\boldsymbol{X}$ e $\boldsymbol{Z}$.

(b) É possível mostrar que $\boldsymbol{Y} \sim \text{Normal}(\boldsymbol{\mu}, \boldsymbol{V})$. Determine $\boldsymbol{\mu}$ e $\boldsymbol{V}$.

(c) Determine o estimador de máxima verossimilhança, $\hat{\boldsymbol{\beta}}_{\text{EMV}}$.

Dica: Para $\boldsymbol{W}$ simétrica, $\dfrac{d}{d\boldsymbol{s}}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s})^\top \boldsymbol{W}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s}) = -2\boldsymbol{A}^\top \boldsymbol{W}(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s})$.

(d) Encontre a distribuição conjunta de $(\boldsymbol{Y}, \boldsymbol{U})^\top$.

(e) Mostre que $\mathbb{E}[\boldsymbol{U} \mid \boldsymbol{Y} = \boldsymbol{y}] = \boldsymbol{\Sigma}\boldsymbol{Z}^\top \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$.

Dica: Use que se $\begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^\top & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right)$, então,

$\boldsymbol{X}_1 \mid \boldsymbol{X}_2 = \boldsymbol{x}_2 \sim \text{Normal}\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)$.

3. **(Modelo Poisson Multinível)** Considere Y_{ij} contagens para $i = 1, 2, \ldots, n$ no grupo $j = 1, 2, \ldots, m$. Suponha que observamos uma covariável x_{ij} e postulamos que as observações são Poisson condicionalmente independentes com média

$$\mathbb{E}[Y_{ij}, \mid u_j, x_{ij}] = \exp(u_j + \beta_0 + \beta_1 x_{ij}),$$

onde u_j é um efeito aleatório por grupo.

(a) Uma propriedade fundamental em um modelo multinível é que os efeitos aleatórios introduzem dependência nas observações. Vamos mostrar isso quando $u_j \sim \text{Normal}(0, \tau^2)$.

i. Temos que $\mathbb{E}[Y_{ij}] = \mathbb{E}[\exp(u_j)] \exp(\beta_0 + \beta_1 x_{ij})$. Determine $\mathbb{E}[\exp(u_j)]$.

ii. Determine a variância marginal de Y_{ij}.

iii. Determine $\text{Cov}(Y_{kj}, Y_{lj})$ para $k \neq l$.

(b) Para estimar $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$ por máxima verossimilhança precisamos marginalizar $\boldsymbol{u} = (u_1, \ldots, u_m)^\top$. Exiba a verossimilhança marginal conjunta $f_{\boldsymbol{Y}}(\boldsymbol{y} \mid \boldsymbol{X}, \boldsymbol{\beta}, \tau^2)$, onde $\boldsymbol{Y} = (Y_{11}, Y_{12}, \ldots, Y_{nm})^\top$ e $\boldsymbol{X}$ é a matriz de desenho.

4. **(Estimação do ATE)** Suponha que realizamos um ensaio clínico aleatorizado (*randomized controlled trial*, RCT) com n participantes, onde atribuímos um tratamento binário $W_i \in \{0, 1\}$ e medimos uma variável resposta Y_i para cada participante $i = 1, \ldots, n$. Para cada $w \in \{0, 1\}$, definimos respostas potenciais $Y_i(w)$ tal que $Y_i = Y_i(W_i)$. Uma quantidade

de interesse, então, é o efeito do tratamento nos indivíduos, $\Delta_i = Y_i(1) - Y_i(0)$. No entanto, um problema fundamental na inferência causal é que só podemos atribuir um tratamento a cada indivíduo, isto é, só observamos $Y_i(1)$ ou $Y_i(0)$. Em amostras finitas, sem hipóteses sobre como os participantes foram gerados, é possível obter estimadores não viesados e consistentes do efeito médio amostral do tratamento, $\overline{\Delta} := \frac{1}{n}\sum_{i=1}^{n} Y_i(1) - Y_i(0)$. Se assumirmos que os participantes são selecionados de uma população P, também é possível obter estimadores não-viesados e consistentes do efeito médio do tratamento na população, $\tau = \mathbb{E}_P[Y_i(1) - Y_i(0)]$. Vamos definir e verificar algumas propriedades desses métodos de estimação.

(a) A maneira mais intuitiva de proceder é através do estimador de diferença de médias,

$$\hat{\tau}_{\mathrm{DM}} = \frac{1}{n_1}\sum_{i=1}^{n} Y_i(1)W_i - \frac{1}{n_0}\sum_{i=1}^{n} Y_i(0)(1 - W_i),$$

onde $n_w = \sum_{i=1}^{n} \mathbb{1}\{W_i = w\}$. Suponha que o tratamento é aleatorizado de fato, ou seja,

$$\Pr(W_i = 1 \mid \{Y_i(1), Y_i(0)\}_{i=1}^{n}, n_1) = \frac{n_1}{n}.$$

Mostre que $\hat{\tau}_{\mathrm{DM}}$ é um estimador não viesado de $\overline{\Delta}$, dado $\{Y_i(1), Y_i(0)\}_{i=1}^{n}$.

(b) Uma limitação do resultado anterior é que não caracteriza o erro padrão de $\hat{\tau}_{\mathrm{DM}} - \overline{\Delta}$. Sendo assim, suponha que produzimos um ensaio Bernoulli, com

$$W_i \mid \{Y_i(0), Y_i(1)\} \sim \mathrm{Bernoulli}(p),\ 0 < p < 1.$$

Supondo que $\{Y_i(0), Y_i(1)\} \sim P$, mostre que

$$\sqrt{n}(\hat{\tau}_{\mathrm{DM}} - \tau) \xrightarrow{d} \mathrm{Normal}(0, V_{\mathrm{DM}}),$$

onde $V_{\mathrm{DM}} = \frac{\mathrm{Var}(Y_i(1))}{p} + \frac{\mathrm{Var}(Y_i(0))}{1-p}$.

(c) Uma outra maneira de estimar o ATE é via regressão linear. Para isso, suponha que

$$Y_i = \alpha + W_i\tau + \varepsilon_i,$$

com $\varepsilon_i \sim \mathrm{Normal}(0, \sigma^2)$. Assim, podemos obter um estimador de τ por máxima verossimilhança ou mínimos quadrados, $\hat{\tau}_{\mathrm{MLE}}$.

i. Mostre que $\hat{\tau}_{\mathrm{MLE}} = \hat{\tau}_{\mathrm{DM}}$.

ii. Usando as propriedades do estimador de máxima verossimilhança, mostre como construir um invertalo de confiança γ para τ.