Prova 3 - [redacted]

1. $P(y_i = 1 \mid x_i) = S(x_i \theta)$ p/ $S(x_i) = 1 / (1 + \exp(-x_i))$

$$P(y_i = 0 \mid x_i) = 1 - S(x_i \theta)$$

$$P(Y = y \mid \theta) = \theta^k (1-\theta)^{N-k} = L(\theta)$$

$$L(\theta, y_i, x_i) = [S(x_i \theta)]^{y_i} [1 - S(x_i \theta)]^{1-y_i}$$

$$L(\theta, y, X) = \prod_{i=1}^{N} [S(x_i \theta)]^{y_i} [1 - S(x_i \theta)]^{1-y_i}$$

$$\log L(\theta, y, X) = \ln\left(\prod [S(x_i \theta)]^{y_i} [1 - S(x_i \theta)]^{1-y_i}\right)$$

$$= \sum_{i=1}^{N} \left[ y_i \ln\left(\overbrace{\frac{1}{1+e^{-x_i\theta}}}^{P}\right) + (1-y_i) \ln\left(\frac{\cancel{1} + e^{-x_i\theta} \cancel{-1}}{1+e^{-x\theta}}\right)\right]$$

$$= \sum_{i=1}^{N} \left[ \ln\left(\frac{\exp(-x_i\theta)}{1+e^{-x\theta}}\right) + y_i \left(\ln\left(\frac{1}{1+e^{-x\theta}}\right) - \ln\left(\frac{e^{-x\theta}}{1+e^{-x\theta}}\right)\right)\right]$$

$$= \sum_{i=1}^{N} \left[ \ln\left(\frac{1}{1+e^{x_i\theta}}\right) + y_i \left(\ln\left(\frac{1}{e^{-x_i\theta}}\right)\right)\right]$$

$$= \sum_{i=1}^{N} \left[ \ln(1) - \ln(1+e^{x\theta}) + y_i (\ln(1) - \ln(e^{-x\theta}))\right]$$

$$= \sum_{i=1}^{N} \left[ -\ln(1+\exp(x_i^t \theta)) + y_i x_i \theta \right]$$

p/ maximizar $\ell(\theta) \rightarrow$ derivada

$$0 = \frac{\partial \ell(\theta)}{\partial \beta_j} = \frac{\partial \sum^{N} [-\ln(1+e^{x_i \beta_j}) + y_i x_i \beta_j}{\partial \beta_j}$$

$$= \sum_{i=1}^{N} \left( \frac{-e^{x_i \beta_j}}{1+e^{x_i \beta_j}} x_i + y_i x_i \right)$$

$$= \sum_{i=1}^{N} \left( y_i - \frac{e^{x_i \beta_j}}{1+e^{x_i \beta_j}} \right) x_i$$

$$= \sum_{i=1}^{N} \left( y_i - \frac{1}{1+e^{-x_i \beta_j}} \right) x_i$$

$$= \sum_{i=1}^{N} x_i (y_i - p_i)$$

$$= X^t (y - p)$$

2. A primeira derivada, obtida na questão 1 é:

$$\frac{\partial l(\theta)}{\partial \beta_j} = \sum_{i=1}^{N} x_i (y_i - p_i) \quad (a)$$

$$\hookrightarrow \sum x_{ij} y_i - \sum x_{ij} p_i$$

$$\frac{\partial^2 l(\theta)}{\partial \beta_n \partial \beta_j} = \nabla_\theta \ (a)$$

$$= -\sum_{i=1}^{N} \left( \frac{(1+e^{x_i\theta}) x_{in} x_{ij} e^{x_i^t\theta} - x_{in} x_{ij} (e^{x_i\theta})^2}{(1+e^{x_i\theta})^2} \right)$$

$$= -\sum_{i=1}^{N} \left( \frac{x_{in} x_{ij} e^{x_i\theta}}{1+e^{x_i\theta}} \cdot \frac{(1+e^{x_i\theta})}{1+e^{x_i\theta}} - \frac{x_{in} x_{ij} (e^{x_i\theta})^2}{(1+e^{x_i\theta})^2} \right)$$

$$= -\sum_{i=1}^{N} \left( \frac{x_{in} x_{ij} e^{x_i\theta}}{1+e^{x_i\theta}} - x_{in} x_{ij} \left( \frac{e^{x_i\theta}}{1+e^{x_i\theta}} \right)^2 \right)$$

$$= -\sum_{i=1}^{N} \left( \frac{x_{in} x_{ij} \cancel{e^{x_i\theta}}}{1+e^{x_i\theta}} \cdot \frac{\cancel{e^{-x_i\theta}}}{e^{-x_i\theta}} - // \right)$$

$$= -\sum_{i=1}^{N} \left( x_{in} x_{ij} \cdot \frac{1}{1+e^{-x_i\theta}} - x_{in} x_{ij} \left( \frac{e^{x_i\theta}}{1+e^{x_i\theta}} \cdot \frac{e^{-x_i\theta}}{e^{-x_i\theta}} \right)^2 \right)$$

$$= -\sum_{i=1}^{N} \left( x_{in} x_{ij} p_i - x_{in} x_{ij} \overbrace{\left( \frac{1}{1+e^{-x_i\theta}} \right)^2}^{p_i^2} \right)$$

$$= -\sum_{i=1}^{N} x_{in} x_{ij} p_i - x_{in} x_{ij} p^2$$

$$= -\sum_{i=1}^{N} x_{in} x_{ij} p_i (1 - p_i)$$

Na forma matricial, temos,

$$\frac{\partial^2 l(\theta)}{\partial\theta\,\partial\theta^t} = -\sum_{i=1}^{n} x_i x_i^t p_i (1-p_i)$$

$$= -X^t W X$$

$W$ é matriz $n \times n$ diagonal com $i$-ésimo termo dado por $p_i(1-p_i)$

$$\hookrightarrow W = \begin{bmatrix} p_1(1-p_1) & 0 & \cdots & 0 \\ 0 & p_2(1-p_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & p_n(1-p_n) \end{bmatrix}$$

3. $N(\mu_0, \sigma^2)$

$$f(Y_i, \mu_0, \sigma^2) = \frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} \exp\left[-\frac{(y_i-\mu_0)^2}{2\sigma^2}\right]$$

$$\log L(\mu_0, \sigma^2) = \frac{n}{2}\log \sigma^2 - \frac{n}{2}\log(2\pi) - \frac{\sum (Y_i - \mu_0)^2}{2\sigma^2}$$

~ considerando $\mu_0$ conhecido

$$\frac{\partial \log L(\mu_0, \sigma^2)}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{\sum(Y_i - \mu_0)^2}{2\sigma^2} \equiv 0$$

$$= -n\sigma^2 + \sum(Y_i - \mu_0)^2 = 0$$

$$\sigma^2 = \sum\frac{(Y_i - \mu_0)^2}{n}$$

~ Checar pto de máxima

$$\hat{\sigma}^2 = \sum\frac{(Y_i - \bar{Y})^2}{n}$$

$$\frac{\partial^2 \log L(\mu_0, \sigma^2)}{(\partial \sigma^2)^2} = \frac{n}{\sigma^4} - \frac{\sum(Y_i - \mu_0)^2}{\sigma^6}$$

$$\frac{n}{\sigma^4} - \frac{\sum(Y_i - \mu_0)^2}{\sigma^6} = 0 \rightarrow n\sigma^2 - \sum(Y_i - \mu_0)^2 = 0$$

$$\hat{\sigma}^2 = \frac{\sum(Y_i - \mu_0)^2}{n}$$

$\rightarrow \sum(Y_i - \mu_0)^2$ é sempre $>0$ e $n>0$ $\therefore$ $\hat{\sigma}^2 > 0$ e, como a 2ª derivada é negativa temos que esta é o ponto de máxima.

Assim, $\hat{\sigma}^2 = \frac{\sum(Y_i - \mu_0)^2}{n}$

4. $E(Y) = \mu$

$$E(Y-m)^2 = E(Y-\mu+\mu-m)^2$$
$$= E[(Y-\mu)^2 + 2(Y-\mu)(\mu-m) + (\mu-m)^2]$$
$$= E[(Y-\mu)^2] + 2E[(Y-\mu)(\mu-m)] + E[(\mu-\theta)^2]^2$$
$$= Var(Y) + 2(\mu-m)E(Y-\mu) + (\mu-m)^2$$
$$E(Y-m)^2 = Var(Y) + 2(\mu-m).0. + (\mu-m)^2$$
$$= Var(Y) + 0 + (\mu-m)^2$$
$$= Var(Y) + \underbrace{(\mu-m)^2}_{\text{bias}}$$

→ derivando em relação a $m$

↳ $-2\mu + 2m$

↳ $m = \mu$ p/ derivada $= 0$

↳ $m = \mu$ minimiza

5. Seja $H_1$ a matriz de projeção ortogonal em $\mathcal{M}(X)$, ou seja, $H_1 = X(X'X)^{-1}X'$. Então temos:

- Para $Y \in \mathbb{R}^n$, $HY \in \mathcal{M}(X)$ e $HY$ é projeção ortogonal de $Y$ em $\mathcal{M}(X)$. $H_1 = 1(1^t 1)^{-1}1'$

- A solução de mínimos quadrados $\hat{\beta} = \arg\min_b \|Y - Xb\|^2$
  - Assim $Xb$ é $HY = X(X'X)^{-1}X'Y$
  - A projeção é $\hat{Y} = X\hat{\beta} \Rightarrow \hat{Y} = H_1 Y$
- Em uma boa predição $Y \approx \hat{Y} = X.\hat{\beta}$
  - Se $\bar{y} = \sum y_i/N$ e $\bar{Y} = \bar{y}\,1'$ então,
    - $Y = \hat{Y} + (Y - \hat{Y}) = \bar{Y} + \hat{Y} - \bar{Y} + (Y - \hat{Y})$
    - $Y - \bar{y}1 = (\hat{Y} - \bar{y}1) + (Y - \hat{Y})$

$Y = X\hat{\beta} + (Y - X\hat{\beta})$

$Y = X(X'X)^{-1}X'Y + (Y - X(X'X)^{-1}X'Y)$

$Y = H_1 Y + (Y - H_1 Y)$

$Y = H_1 Y + (I - H_1)Y = \bar{Y} + \hat{Y} - \bar{Y} + (Y - \hat{Y})$

$Y = H_1 Y + (I - H_1)Y = \bar{y}1' + H_1 Y - \bar{y}1' + (I - H_1)Y$

$Y = H_1 Y + (I - H_1)Y = \bar{y}1' + H_1 Y - \bar{y}1' + (I - H_1)Y$

~> Aqui devemos notar que $H_1$ é simétrica e idempotente

$Y = H_1 Y + (I - H_1)Y = \bar{y}1' + \overset{Y}{\widetilde{H_1 Y}} - \bar{y}1' + Y - \overset{Y}{\widetilde{H_1 Y}}$

$Y = H_1 Y + (I - H_1)Y = \bar{y}1' + (Y - \bar{y}1') + \cancel{Y - Y}$

~> próxima pg ~>

· $\langle u, v \rangle$ são ortogonais se o seu produto interno é 0

$$\rightarrow \langle \bar{y}1, (Y - \bar{y}1) \rangle \rightarrow \langle \bar{y}1, (y_i - \bar{y})1 \rangle$$

$$\rightarrow \bar{y}(y_1 - \bar{y}) + \bar{y}(y_2 - \bar{y}) + \dots + \bar{y}(y_n - \bar{y}) = \bar{y} \sum_i^n (y_i - \bar{y})$$

$$\rightarrow \bar{y} = \sum y_i/n \therefore \sum^n \bar{y} = \sum^n y_i \therefore \sum^n (y_i - \bar{y}) = 0 \therefore$$

$$\langle \bar{y}1, (Y - \bar{y}1) \rangle = 0$$

FORONI