HW7-Q2

Tuesday, November 19, 2024 11:18 AM

Derive: $D^2 = -2 \sum_{i=1}^{g} [S_i \ln \hat{p}_i + (n_i - S_i) \ln(1-\hat{p}_i)\}$

$- \{S_i \ln f_i + (n_i - S_i) \ln(1-f_i)\}]$

$= -2 \sum_{i=1}^{g} \left[S_i \ln\left(\frac{\hat{p}_i}{f_i}\right) + (n_i - S_i) \ln\left(\frac{1-\hat{p}_i}{1-f_i}\right)\right]$ $g \geq 2$

---

likelihood for logistic regression

① $L_i = \binom{n_i}{S_i} p_i^{S_i} (1-p_i)^{n_i - S_i}$

grouped $L = \prod_{i=1}^{g}$ ↓

② deviance is based on diff in log-likelihood

$l = \ln L \rightarrow \sum_{i=1}^{g} \ln \binom{n_i}{S_i} + \sum_{i=1}^{g} [S_i \ln(p_i) + (n_i - S_i) \ln(1-p_i)]$

③ $f_i = \frac{S_i}{n_i}$ saturated $\rightarrow$ $\sum_{i=1}^{g} [S_i \ln(f_i) + (n_i - S_i) \ln(1-f_i)]$ ← l(f)

④ $D^2 = -2[l(\hat{p}) - l(f)]$

$D^2 = -2 \sum_{i=1}^{g} \{S_i \ln(\hat{p}_i) + (n_i - S_i) \ln(1-\hat{p}_i) - [S_i \ln(f_i) + (n_i - S_i) \ln(1-f_i)]\}$

$$D^2 = -2 \sum_{i=1}^{g} \left[S_i \ln\left(\frac{\hat{p}_i}{f_i}\right) + (n_i - S_i) \ln\left(\frac{1-\hat{p}_i}{1-f_i}\right)\right]$$

---

degree of freedom: $g = 6$, # of par $= 2$

↳ $6 - 2 = 4$

Toxicity Dataset

$f_i = \frac{S_i}{n_i} \rightarrow [0.112, 0.212, 0.372, 0.504, 0.688, 0.788]$

$D^2 = -2 \sum_{i=1}^{g} \left[S_i \ln\left(\frac{\hat{p}_i}{f_i}\right) + (n_i - S_i) \ln\left(\frac{1-\hat{p}_i}{1-f_i}\right)\right]$

$\hat{p}_i = [0.136, 0.207, 0.318, 0.475, 0.671, 0.818]$

$S_1 \rightarrow 28, n_1 = 250, f_1 = 0.112, \hat{p}_1 = 0.136$

$28 \ln\left(\frac{.136}{.112}\right) = 28 \times 0.199 = 5.572$

$222 \ln\left(\frac{1-0.136}{1-0.112}\right) = 222 \times -0.027 = -6.011$ } -0.439

$S_2 \rightarrow 53, n_2 = 250, f_2 = 0.212, \hat{p}_2 \rightarrow 0.207$

$53 \ln(0.978) \rightarrow -1.166$

$197 \ln(1.006) \rightarrow 1.182$ $\quad -1.166 + 1.182 = 0.016$

$S_3 \rightarrow 93, n_3 \rightarrow 250, f_3 \rightarrow 0.372, \hat{p}_3 \rightarrow 0.317$

$93 \ln(0.852) \rightarrow -14.97$ $\quad \rightarrow -1.942$

$157 \ln\left(\frac{0.683}{0.628}\right) = 157(1.087) \rightarrow 13.031$

$S_4 \rightarrow 126, n_4 \rightarrow 250, f_4 \rightarrow 0.504, \hat{p}_4 = 0.475$

$126 \ln(0.943) \rightarrow -7.434$

$124 \ln(1.058) \rightarrow 6.444$ } -0.490

$S_5 \rightarrow 172, n_5 = 250, f_5 \rightarrow 0.688, \hat{p}_5 = 0671$

$172 \ln(0.975) \rightarrow -4.300$

$78 \ln(1.054) \rightarrow 4.134$

$-4.3 + 4.13 \rightarrow -0.166$

$S_6 \rightarrow 197, n_6 \rightarrow 250, f_6 \rightarrow 0.788, \hat{p}_6 = 0.818$

$197 \ln(1.038) \rightarrow 7.289$

$53 \ln(0.859) \rightarrow -8.056$ } -0.767

$-2 \times (-0.590 + 0.016 - 1.942 - 0.490 - 0.166 - 0.767) \rightarrow \boxed{1.45}$

deviance ↑