

Suppose that the outcome of game i is multinomial with single trial; that is $\mathbf{y}_i \sim MN(1; \mathbf{p}_i)$, where $\mathbf{y}_i = (y_{i1}, y_{i2}, y_{i3})$ and $\mathbf{p}_i = (p_i^H, p_i^D, p_i^A)$. Suppose that the probabilities are given by

$$\begin{aligned}\text{logit}(p_i^H) &= \log(o_i^H) + \delta^H = lo_i^H + \delta^H \\ \text{logit}(p_i^D) &= \log(o_i^D) + \delta^D = lo_i^D + \delta^D,\end{aligned}$$

where o_i^H are the odds of a home win based on the opta probabilities. Then the likelihood function based on n games is

$$L(\boldsymbol{\delta}) = \prod_{i=1}^n (p_i^H)^{y_{i1}} (p_i^D)^{y_{i2}} (p_i^A)^{y_{i3}}$$

and hence the log-likelihood is

$$\log L(\boldsymbol{\delta}) = \sum_{i=1}^n \left[y_{i1} \log p_i^H + y_{i2} \log p_i^D + y_{i3} \log p_i^A \right]. \quad (1)$$

Note that the model implies that the probabilities are given by

$$\begin{aligned}p_i^H &= \frac{\exp(lo_i^H + \delta^H)}{\exp(lo_i^H + \delta^H) + \exp(lo_i^D + \delta^D) + 1} \\ p_i^D &= \frac{\exp(lo_i^D + \delta^D)}{\exp(lo_i^H + \delta^H) + \exp(lo_i^D + \delta^D) + 1} \\ p_i^A &= \frac{1}{\exp(lo_i^H + \delta^H) + \exp(lo_i^D + \delta^D) + 1}.\end{aligned}$$

Hence (1) can be written as

$$\log L(\boldsymbol{\delta}) = \sum_{i=1}^n \left\{ y_{i1}(lo_i^H + \delta^H) + y_{i2}(lo_i^D + \delta^D) - \log \left[\exp(lo_i^H + \delta^H) + \exp(lo_i^D + \delta^D) + 1 \right] \right\}$$

This formula can be used to determine the gradient vector and Hessian of the log-likelihood (with respect to $\boldsymbol{\delta}$). Specifically,

$$\begin{aligned}\frac{\partial}{\partial \delta^H} \log L &= \sum_{i=1}^n (y_{i1} - p_i^H) \\ \frac{\partial}{\partial \delta^D} \log L &= \sum_{i=1}^n (y_{i2} - p_i^D)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial \delta^{H2}} &= -\sum_{i=1}^n p_i^H (1 - p_i^H) \\ \frac{\partial^2}{\partial \delta^{D2}} &= -\sum_{i=1}^n p_i^D (1 - p_i^D) \\ \frac{\partial^2}{\partial \delta^H \partial \delta^D} &= -\sum_{i=1}^n p_i^H p_i^D .\end{aligned}$$