Supplementary material

May 23, 2016

Suppose $\mathbf{p}$ is the vector of observations for a random variable $y \sim \mathcal{C}(\hat{\mu}_1, \ldots, \hat{\mu}_k)$. $\hat{\boldsymbol{\mu}} = \mathcal{L}(\mathbf{p})$ is the estimated probability from $\mathbf{p}$. Let $d_{ij} = \frac{\partial \hat{\mu}_i}{\partial p_j}$. $\theta_c(\hat{\mu}) = \ln \hat{\mu}_c - \ln \hat{\mu}_k$. Equation (1) can be written as:

$$
\begin{aligned}
\sum_{i=1}^{k-1}\sum_{j=1}^{k-1} \mathrm{cov}(p_i, p_j) \frac{\partial \theta_i(\mathcal{L}(\mathbf{v}))}{\partial v_j}|_{\mathbf{v}=\hat{\boldsymbol{\mu}}} &= \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} \mathrm{cov}(p_i, p_j) \frac{\partial [\ln \hat{\mu}_i - \ln \hat{\mu}_k]}{\partial p_j} \\
&= \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} \mathrm{cov}(p_i, p_j) [\frac{d_{ij}}{\hat{\mu}_i} - \frac{d_{kj}}{\hat{\mu}_k}] \\
&= \sum_{i=1}^{k-1} \hat{\mu}_i (1 - \hat{\mu}_i) [\frac{d_{ii}}{\hat{\mu}_i} - \frac{d_{ki}}{\hat{\mu}_k}] - \sum_{i=1}^{k-1}\sum_{j \neq i}^{k-1} \hat{\mu}_i \hat{\mu}_j [\frac{d_{ij}}{\hat{\mu}_i} - \frac{d_{kj}}{\hat{\mu}_k}] \\
&= \sum_{i=1}^{k-1} d_{ii} - \sum_{i=1}^{k} \hat{\mu}_i d_{ii} - \sum_{i=1}^{k-1} \hat{\mu}_i (\hat{\mu}_k + \sum_{m \neq i}^{k-1} \hat{\mu}_m) \frac{d_{ki}}{\hat{\mu}_k} - \sum_{i=1}^{k-1}\sum_{j \neq i}^{k-1} [\hat{\mu}_j d_{ij} - \hat{\mu}_i \hat{\mu}_j \frac{d_{kj}}{\hat{\mu}_k}] \\
&= \sum_{i=1}^{k-1} d_{ii} - \sum_{i=1}^{k} \hat{\mu}_i d_{ii} - \sum_{i=1}^{k-1} \hat{\mu}_i d_{ki} + \sum_{i=1}^{k}\sum_{m \neq i}^{k-1} \hat{\mu}_i \hat{\mu}_m \frac{d_{ki}}{\hat{\mu}_k} - \sum_{i=1}^{k-1}\sum_{j \neq i}^{k-1} [\hat{\mu}_j d_{ij} - \hat{\mu}_i \hat{\mu}_j \\
&= \sum_{i=1}^{k-1} d_{ii} - \sum_{i=1}^{k} \hat{\mu}_i d_{ii} + \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} \hat{\mu}_i d_{ji} - \sum_{i=1}^{k-1}\sum_{j \neq i}^{k-1} \hat{\mu}_j d_{ij} \\
&= \sum_{i=1}^{k-1} d_{ii}
\end{aligned}
$$