$\hat{y}_i \neq 0$ for all $i \in I''$, $I' \cup I'' = I$. I' and I'' specify the standard event observation. Standard conditioning requires that

$$\hat{y}_i = \frac{y_i}{\sum_{k \in I''} y_k}. \tag{20}$$

To solve this problem using PME, we want to minimize the cross-entropy with the constraint that the non-zero $\hat{y}_i$ sum to 1. The Lagrange function is (writing in vector form $\hat{y} = (\hat{y}_i)_{i \in I''}$)

$$\Lambda(\hat{y}, \lambda) = \sum_{i \in I''} \hat{y}_i \ln \frac{\hat{y}_i}{y_i} + \lambda \left(1 - \sum_{i \in I''} \hat{y}_i \right). \tag{21}$$

Differentiating the Lagrange function with respect to $\hat{y}_i$ and setting the result to zero gives us

$$\hat{y}_i = y_i e^{\lambda - 1} \tag{22}$$

with λ normalized to

$$\lambda = -1 + \ln \sum_{i \in I''} y_i. \tag{23}$$

(20) follows immediately. PME generalizes standard conditioning.

A.2. Jeffrey Conditioning

Let $\theta_i, i = 1, \ldots, n$ and $\omega_j, j = 1, \ldots, m$ be finite partitions of the event space with the joint prior probability matrix (y_{ij}) (all $y_{ij} \neq 0$). Let κ be defined as in Section 3, with (1) true (remember that in Section 5, (1) is no longer required). Let P be the type II prior probability distribution and $\hat{P}$ the posterior probability distribution.

Let $\hat{y}_{ij}$ be the posterior probability distribution derived from Jeffrey conditioning with

$$\sum_{i=1}^{n} \hat{y}_{ij} = \hat{P}(\omega_j) \text{ for all } j = 1, \ldots, m \tag{24}$$

Jeffrey conditioning requires that for all $i = 1, \ldots, n$

$$\hat{P}(\theta_i) = \sum_{j=1}^{m} P(\theta_i | \omega_j) \hat{P}(\omega_j) = \sum_{j=1}^{m} \frac{y_{ij}}{P(\omega_j)} \hat{P}(\omega_j) \tag{25}$$

Using PME to get the posterior distribution $(\hat{y}_{ij})$, the Lagrange function is (writing in vector form $\hat{y} = (x_{11}, \ldots, x_{n1}, \ldots, x_{nm})^\top$ and $\lambda = (\lambda_1, \ldots, \lambda_m)^\top$)

$$\Lambda(\hat{y}, \lambda) = \sum_{i=1}^{n} \sum_{j=1}^{m} \hat{y}_{ij} \ln \frac{\hat{y}_{ij}}{y_{ij}} + \sum_{j=1}^{m} \lambda_j \left(\hat{P}(\omega_j) - \sum_{i=1}^{n} \hat{y}_{ij} \right). \tag{26}$$

Consequently,

$$\hat{y}_{ij} = y_{ij} e^{\lambda_j - 1} \tag{27}$$

with the Lagrangian parameters λ_j normalized by

$$\sum_{i=1}^{n} y_{ij} e^{\lambda_j - 1} = \hat{P}(\omega_j) \tag{28}$$