Statistical details

Once the set of non-genetic variables to be evaluated is established, each combination of values of the variables defines a bracket. First, to simplify the notation, we number the brackets from 1 to m, being m the total number of brackets. Let us call N_j the total number of possible victims belonging to the j bracket.

We consider a Dirichlet distribution as the prior probabilities for obtaining every brackets.

$$Priori\left((\theta_1, \theta_2, ..., \theta_j, ...)|(\alpha_1, \alpha_2, ..., \alpha_j, ...)\right) = \frac{\Gamma(\alpha_1 + \alpha_2 + ... + \alpha_j + ...)}{\Gamma(\alpha_1)\Gamma(\alpha_2)...\Gamma(\alpha_j)...}\theta_1^{\alpha_1}\theta_2^{\alpha_2}...\theta_j^{\alpha_j}... \qquad (1)$$

The hyperameters are calibrated in order to be consistent with the prior knowledge, thus $E(\theta j) = \frac{N_j}{N}$ for all $j = 1, 2..., m$ (where $E(\theta_j)$ means expected value or "average" value of θ_j). Therefore, in the instance of the knowledge prior of the data of the already-solved cases, all the individuals have the same chance of corresponding with the skeletal remains S. As for the Dirichlet distribution, the expected value of the variable θ_j is α_j/α_0 being $\alpha_0 = \alpha_1 + \alpha_2 + ... + \alpha_m$, then what must be satisfied are$m$ conditions, one for each bracket. Assuming an extra condition for the k bracket such as N_k is maximum, such as $\frac{\sqrt{Var(\theta_k)}}{E(\theta_k} = 1$, the solution for the α_0 hyperparameter is $\alpha_0 = \frac{N}{N_k} - 2$, and the expression for each α_i hyperparameter corresponding to the i bracket is:

$$\alpha_i = \frac{N_i}{N}\left(\frac{N}{N_k} - 2\right) \quad \forall\ i$$

Finally, using that Dirichlet and Multinomial distributions are conjugate distributions, then if the likelihood is a multinomial and the a prior probability of the parameters is a Dirichlet, then the Posterior distribution of the parameters $Post\left((\theta_1, \theta_2...\theta_m)|Data, \mathcal{H}\right)$ is also a Dirichlet distribution but with other parameters, which are going to be a function of the parameters of the prior and the $Data$, such as

$$\alpha_j^{'} = \alpha_j + n_j$$

being n_j the total number of already-solved cases of the massive event associated to j bracket. This means that in the ideal problem the probability that the outcome result of j bracket will be

$$\frac{\alpha_j + n_j}{\alpha_0 + n} = \frac{\frac{N_j}{N}\left(\frac{N}{N_k} - 2\right) + n_j}{\frac{N}{N_k} - 2 + n} \equiv \theta_j^{post}$$

being n the total number of already-solved cases.

In order to construct a probabilitisc ranking for each possible victim we assume that whithin a particular bracket, all of the people which are unidentified have the same chance of corresponding with the skeletal remain of the same event. Thus the probability of a particular individual whose non-genetic variables values are associated to bracket j is

$$P = \frac{\theta_j^{post}}{N_j - n_j}$$

Odds probability are calculated as $ODDS = P/(1 - P)$.

References

A. O'Hagan, C.E. Buck, A. Daneshkhah,J.R. Eiser, P.H. Garthwaite,D.J. Jenkinson,J.E. Oakleyy T. Rakow, Uncertain Judgements: Eliciting Experts' Probabilities (Statistics in Practice). Wiley,Nueva York, EE.UU.(2006).