$$M^2(N_J(I)) = \sum_{i \in I} f_i \| f_J^i - f_J \|_{f_J}^2 = \sum_{i \in I} f_i \rho^2(i) \qquad (2.3)$$

In the latter term, ρ is the Euclidean distance from the cloud center, and f_i is the mass of element i. The mass is the marginal distribution of the input data table. Let us take a step back: the given contingency table data is denoted $k_{IJ} = \{k_{IJ}(i,j) = k(i,j); i \in I, j \in J\}$. We have $k(i) = \sum_{j \in J} k(i,j)$. Analogously $k(j)$ is defined, and $k = \sum_{i \in I, j \in J} k(i,j)$. Next, $f_{IJ} = \{f_{ij} = k(i,j)/k; i \in I, j \in J\} \subset \mathbb{R}_{I \times J}$, similarly f_I is defined as $\{f_i = k(i)/k; i \in I, j \in J\} \subset \mathbb{R}_I$, and f_J analogously.

Next back to the first right hand side term in equation 2.3: the conditional distribution of f_J knowing $i \in I$, also termed the jth profile with coordinates indexed by the elements of I, is

$$f_J^i = \{f_j^i = f_{ij}/f_i = (k_{ij}/k)/(k_i/k); f_i \neq 0; j \in J\}$$

and likewise for f_I^j.

The cloud of points consists of the couple: profile coordinate and mass. We have $N_J(I) = \{(f_J^i, f_i); i \in I\} \subset \mathbb{R}_J$, and again similarly for $N_I(J)$.

From equation 2.3, it can be shown that

$$M^2(N_J(I)) = M^2(N_I(J)) = \| f_{IJ} - f_I f_J \|_{f_I f_J}^2 = \sum_{i \in I, j \in J} (f_{ij} - f_i f_j)^2 / f_i f_j$$

$$(2.4)$$

The term $\| f_{IJ} - f_I f_J \|_{f_I f_J}^2$ is the χ^2 metric between the probability distribution f_{IJ} and the product of marginal distributions $f_I f_J$, with as center of the metric the product $f_I f_J$.

In correspondence analysis, the choice of χ^2 metric of center f_J is linked to the *principle of distributional equivalence*, explained as follows. Consider two elements j_1 and j_2 of J with identical profiles: i.e., $f_I^{j_1} = f_I^{j_2}$. Consider now that elements (or columns) j_1 and j_2 are replaced with a new element j_s such that the new coordinates are aggregated profiles, $f_{ij_s} = f_{ij_1} + f_{ij_2}$, and the new masses are similarly aggregated: $f_{ij_s} = f_{ij_1} + f_{ij_2}$. Then there is *no effect* on the distribution of distances between elements of I. The distance between elements of J, other than j_1 and j_2, is naturally not modified. This description has followed closely [47] (chapter 2).

The principle of distributional equivalence leads to representational self-similarity: aggregation of rows or columns, as defined above, leads to the same analysis. Therefore it is very appropriate to analyze a contingency table with fine granularity, and seek in the analysis to merge rows or columns, through aggregation.

2.2.3 Notation for Factors

Correspondence analysis produces an ordered sequence of pairs, called factors, (F_α, G_α) associated with real numbers called eigenvalues $0 \leq \lambda_\alpha \leq 1$. The