I started writing this note by saying "I do not remember exactly when and where I first met Professor Caussinus, but it could very well be in Cambridge in 1985." As I investigate my chronological history more systematically, I find my conjecture was indeed correct. The first European Meeting of the Psychometric Society was held in Queen's College at Cambridge University in 1985, and there was a satellite meeting on Multidimensional Data Analysis in Pembroke College just prior to the Psychometric Society Meeting. There Professor Caussinus and I each gave an invited talk on probabilistic and statistical aspects of multidimensional data analysis. The topic of his talk at the satellite meeting was, as I recall, principal component analysis which had little to do with either quasi-symmetry or quasi-independence, while my talk was on comparison of models for stimulus identification data [9], and one of the models compared was closely related to the model of quasi-symmetry for square contingency tables (more on this shortly). I met Professor Caussinus at least one more time. A few years later he visited McGill to give a colloquium talk. I do not remember what he talked about then, but I remember clearly the fine moment we shared over lunch at the Faculty Club.

In a stimulus identification (or as sometimes called, stimulus recognition) experiment, one stimulus is presented at a time from a set of $n$ stimuli. The subject is asked to identify which one of the $n$ possible stimuli is actually presented. Stimuli are usually presented under degraded conditions to deliberately create confusions among the stimuli. The number of times stimulus $i$ is identified (or misidentified) as stimulus $j$ is counted out of $N_i$ repeated presentations of stimulus $i$. It is a very popular experimental paradigm used in psychology for investigating the structure of similarities between stimuli (see references in [9, 10]). A variety of models have been proposed for this kind of data, and one of them due to Luce [5], called the biased choice model, posits that the conditional probability of response $j$ when stimulus $i$ is presented ($p_{ji}$) is proportional to the similarity between stimuli $i$ and $j$ ($s_{ij}$) and the response bias for stimulus $j$ ($w_j$). That is, $p_{ji} \propto w_j s_{ij}$, or $p_{ji} = c_j w_j s_{ij}$, where $c_j = [\sum_{k=1}^{n} w_k s_{ik}]^{-1}$. It is assumed that $s_{ij} = s_{ji} \leq s_{ii} = s_{jj} = 1$ for $i, j = 1, \ldots, n$, and that $\sum_{j=1}^{n} w_j = 1$. It can be easily verified that this model satisfies the cycle condition ($p_{ij}p_{jk}p_{ki} = p_{ji}p_{kj}p_{ik}$), characteristic of the quasi-symmetry model. The biased choice model is thus a special case of the quasi-symmetry model. I
learned this from Smith [7] and Townsend and Landon [11], when I was working on [9]. Both articles refer to Professor Caussinus’ original article [2] and Bishop, Fienberg, and Holland’s [1] book on log-linear models for contingency tables. I suppose that the latter has had enormous impact on disseminating the idea of quasi-symmetry.

A number of models have been proposed in psychometrics which specialize Luce’s biased choice model. Typically, they further impose restrictions on $s_{ij}$. For example, $s_{ij} = \exp(-d_{ij})$ (Shepard [6]; Euclidean distance choice model) or $s_{ij} = \exp(-d_{ij}^2)$ (Squared Euclidean distance choice model), where $d_{ij}$ is the distance between stimuli $i$ and $j$ represented as points in a multidimensional Euclidean space. This is one point of contact between multidimensional scaling (MDS) and the analysis of square contingency tables. MDS is a popular technique developed in psychometrics for analysis of proximity data in general by a distance model. Keren and Baggen’s [3] model for stimulus identification data follows from assuming a special kind of distance function for $d_{ij}$ in the exponent of $s_{ij}$. Shepard’s [7] model (see above) can also be derived from Krumhansl’s [4] distance-density model which postulates $\hat{d}_{ij} = d_{ij} + a_i + b_j$, where $a_i$ and $b_j$ indicate the stimulus density near stimuli $i$ and $j$, respectively, and by assuming $p_{ij} \propto \exp(-\hat{d}_{ij})$ with no explicit bias parameter. It can be readily seen that $\exp(-a_i)$ falls out (being cancelled out in the numerator and the denominator), and that $\exp(-b_j)$ comes out as a kind of bias parameter (having a similar role to $w_j$).

All these are special cases of the quasi-symmetry model as far as the symmetry of $s_{ij}$ is maintained. A few years later I used a similar idea to develop a model [8] for general contingency tables (usually rectangular). In this model, row and column categories of contingency tables are represented as (separate) points in a joint MD Euclidean space in such a way that the probability of column $j$ in a given row (say, $i$) is obtained by a decreasing function of the squared Euclidean distance between row $i$ and column $j$. One obvious candidate for this monotonically decreasing function is the negative exponential function, $\exp(-d_{ij}^2)$. The form of the model is essentially the same as the one for square contingency tables (Squared Euclidean distance choice model), only what $i$ and $j$ refer to are different.

References


