

## MULTIWAY CLASSIFICATION MODELS

## Why

1

## Definition

The two-sample model (or the two-sample problem) has input space  $\{1,2\}$  and regression function  $\phi: \{1,2\} \to \mathbf{R}^2$  where  $\phi(1) = (1,0)^{\top}$  and  $\phi(2) = (0,1)^{\top}$ . In this case, we model two populations (corresponding to 1 and 2) with different means but a common variance. The regression range is the set  $\{(1,0),(0,1)\}$ .

Generalizing from two to a populations gives the one-way classification model. In that case the input space is  $\{1, ..., a\}$  and the regression function is  $\phi : \{1, ..., a\} \to \mathbf{R}^a$  defined by  $\phi(i) = e_i$  where  $e_i$  is the standard unit vector in  $\mathbf{R}^a$ .<sup>2</sup> The regression range is  $\{e_1, ..., e_a\}$ . In this case we say that the factor population takes levels 1, ..., a.

If there are more than one factors, then we have a multiway classification model. For example, the two-way classification model with no interaction has input domain  $X = \{1, ..., a\} \times \{1, ..., b\}$  and the regression function  $\phi: X \to \mathbf{R}^{a+b}$  is defined by  $\phi(i, j) = (e_i, e_j)^{\top}$ .

<sup>&</sup>lt;sup>1</sup>Future editions will include.

<sup>&</sup>lt;sup>2</sup>Future editions will define the standard unit vector.

