



## Why

### Definition

Consider a joint distribution with  $n$  components. We associate with this joint  $n$  *marginal distributions*.

For  $i = 1, \dots, n$ , the  $i$ th *marginal distribution* of the joint is the distribution over the  $i$ th set in the product which assigns to each element of that set the sum of probabilities of outcomes whose  $i$ th component matches that element.

For  $i, j = 1, \dots, n$  and  $i \neq j$ , the  $i, j$ th *marginal distribution* of the joint is the distribution over the product of the  $i$ th and  $j$ th sets in the original product which assigns to each element in the product the sum of probabilities of outcomes whose  $i$  component matches the first component of the product and whose  $j$ th component matches the  $j$ th component of the product.

### Notation

Let  $A_1, \dots, A_n$  be non-empty finite sets. Define  $A = \prod_{i=1}^n A_i$  and let  $p : A \rightarrow \mathbf{R}$  be a joint distribution.

For  $i = 1, \dots, n$ , define  $p_i : A_i \rightarrow \mathbf{R}$  by

$$p_i(b) = \sum_{a_i=b} p(a).$$

for each  $b \in A_i$ .  $p_i$  is the  $i$ th marginal of  $p$ .

Similarly, for  $i, j = 1, \dots, n$  and  $i \neq j$  define  $p_{ij} : A_i \times A_j \rightarrow$

**R** by

$$p_{ij}(b, c) = \sum_{a_i=b, a_j=c} p(a)$$

for every  $b \in A_i$  and  $c \in A_j$ . Then  $p_{ij}$  is the  $i, j$ th marginal of  $p$ .

