The probability distribution over the subset is known as the **marginal probability** distribution.

For example, suppose we have discrete random variables x and y, and we know P (x, y). We can find P (x) with the sum rule:

$$\forall x \in \mathrm{x}, P(\mathrm{x} = x) = \sum_{y} P(\mathrm{x} = x, \mathrm{y} = y).$$

The name “marginal probability” comes from the process of computing marginal probabilities on paper. When the values of P (x, y) are written in a grid with different values of x in rows and different values of y in columns, it is natural to sum across a row of the grid, then write P(x) in the margin of the paper just to the right of the row.

For continuous variables, we need to use integration instead of

$$p(x) = \int p(x, y)dy.$$