sum rule $P(X) = \sum_Y P(X, Y) = \sum_y P(X, Y = y)$

$$P_X(x) = \sum_{y \in \mathrm{Dom}(Y)} P_{X,Y}(x, y) \quad \text{for } x \in \mathrm{Dom}(X)$$

sum rule $P(X) = \sum_Y P(X, Y) = \sum_y P(X, Y = y)$

$$P_X(x) = \sum_{y \in \mathrm{Dom}(Y)} P_{X,Y}(x, y) \quad \text{for } x \in \mathrm{Dom}(X)$$

$$P(X = x) = \sum_{y \in \mathrm{Dom}(Y)} P(X = x \wedge Y = y)$$

sum rule $P(X) = \sum_Y P(X,Y) = \sum_y P(X, Y = y)$

$$P_X(x) = \sum_{y \in \mathrm{Dom}(Y)} P_{X,Y}(x,y) \quad \text{for } x \in \mathrm{Dom}(X)$$

$$P(X = x) = \sum_{y \in \mathrm{Dom}(Y)} P(X = x \wedge Y = y)$$

$$\mu(\{\omega \in \Omega \mid \omega \models X = x\}) = \sum_{y \in \mathrm{Dom}(Y)} \mu(\{\omega \in \Omega \mid \omega \models X = x \wedge Y = y\})$$

sum rule $P(X) = \sum_Y P(X, Y) = \sum_y P(X, Y = y)$

$$P_X(x) = \sum_{y \in \mathrm{Dom}(Y)} P_{X,Y}(x, y) \quad \text{for } x \in \mathrm{Dom}(X)$$

$$P(X = x) = \sum_{y \in \mathrm{Dom}(Y)} P(X = x \wedge Y = y)$$

$$\mu(\{\omega \in \Omega \mid \omega \models X = x\}) = \sum_{y \in \mathrm{Dom}(Y)} \mu(\{\omega \in \Omega \mid \omega \models X = x \wedge Y = y\})$$

$$\mu(\{\omega \in \Omega \mid \omega(X) = x\}) = \sum_{y \in \mathrm{Dom}(Y)} \mu(\{\omega \in \Omega \mid \omega(X) = x \text{ and } \omega(Y) = y\})$$

sum rule $P(X) = \sum_Y P(X, Y) = \sum_y P(X, Y = y)$

$$P_X(x) = \sum_{y \in \mathrm{Dom}(Y)} P_{X,Y}(x, y) \quad \text{for } x \in \mathrm{Dom}(X)$$

$$P(X = x) = \sum_{y \in \mathrm{Dom}(Y)} P(X = x \wedge Y = y)$$

$$\mu(\{\omega \in \Omega \mid \omega \models X = x\}) = \sum_{y \in \mathrm{Dom}(Y)} \mu(\{\omega \in \Omega \mid \omega \models X = x \wedge Y = y\})$$

$$\mu(\{\omega \in \Omega \mid \omega(X) = x\}) = \sum_{y \in \mathrm{Dom}(Y)} \mu(\{\omega \in \Omega \mid \omega(X) = x \text{ and } \omega(Y) = y\})$$

From additivity of μ (for finite Ω)

$$\mu(S) = \sum_{\omega \in S} \mu(\{\omega\})$$

Joint probability from a table

	y_1	y_2	$\cdots$	y_c
x_1	$P(x_1, y_1)$	$P(x_1, y_2)$	$\cdots$	$P(x_1, y_c)$
x_2	$P(x_2, y_1)$	$P(x_2, y_2)$	$\cdots$	$P(x_2, y_c)$
$\vdots$				
x_r	$P(x_r, y_1)$	$P(x_r, y_2)$	$\cdots$	$P(x_r, y_c)$

Joint probability from a table

margin

$P(x_i) = \sum_y P(x_i, y)$

	y_1	y_2	$\cdots$	y_c
x_1	$P(x_1, y_1)$	$P(x_1, y_2)$	$\cdots$	$P(x_1, y_c)$
x_2	$P(x_2, y_1)$	$P(x_2, y_2)$	$\cdots$	$P(x_2, y_c)$
$\vdots$				
x_r	$P(x_r, y_1)$	$P(x_r, y_2)$	$\cdots$	$P(x_r, y_c)$

Wikipedia on *Marginal distribution*

> *Marginal variables are those variables in the subset of variables being retained. These concepts are "marginal" because they can be found by summing values in a table along rows or columns, and writing the sum in the margins of the table.*

Sum rule as marginalisation

$$P(X) = \sum_{Y} P(X, Y)$$

Sum rule as marginalisation

$$P(X) = \underbrace{\sum_Y P(X, Y)}^{\text{joint probability } P(X,Y)}$$

Sum rule as marginalisation

joint probability $P(X, Y)$

$$P(X) = \sum_{Y} P(X, Y)$$

marginal probability $P(X)$

Sum rule as marginalisation

joint probability $P(X, Y)$

$$P(X) = \sum_{Y} P(X, Y)$$

marginal probability $P(X)$ marginalising out Y

Sum rule as marginalisation

joint probability $P(X, Y)$

$$P(X) = \sum_{Y} P(X, Y)$$

marginal probability $P(X)$

marginalising out Y

$\approx$ eliminating Y

Sum rule as marginalisation

joint probability $P(X, Y)$

$$P(X) = \sum_{Y} P(X, Y)$$

marginal probability $P(X)$

marginalising out Y

$\approx$ eliminating Y

nuisance variable Y

Sum rule as marginalisation

joint probability $P(X, Y)$

$$P(X) = \sum_Y P(X, Y)$$

marginal probability $P(X)$

marginalising out Y

$\approx$ eliminating Y

nuisance variable Y

We'll define $P(X|Y)$ so that

$$P(X) = \text{expected value of } P(X|Y) \text{ over } Y$$

Sum rule as marginalisation

$$\overbrace{P(X)}^{}\;=\;\sum_{Y} P(X,Y)$$

joint probability $P(X, Y)$

marginal probability $P(X)$

marginalising out Y

$\approx$ eliminating Y

nuisance variable Y

We'll define $P(X|Y)$ so that

$$\begin{aligned} P(X) &= \text{expected value of } P(X|Y) \text{ over } Y \\ &= \sum_{Y} P(X|Y)P(Y) \end{aligned}$$

Sum rule as marginalisation

$$
\begin{array}{rl}
 & \text{joint probability } P(X,Y) \\
P(X) & = \sum_Y P(X,Y) \\
\text{marginal probability } P(X) & \text{marginalising out } Y \\
 & \approx \text{ eliminating } Y \\
 & \text{nuisance variable } Y
\end{array}
$$

We'll define $P(X|Y)$ so that

$$
\begin{aligned}
P(X) &= \text{expected value of } P(X|Y) \text{ over } Y \\
&= \sum_Y P(X|Y)P(Y) \\
P(x) &= \sum_y P(x|y)P(y) \\
&= \mathbb{E}_y[P(x|y)]
\end{aligned}
$$