1 Probability Theory

[Probability Model $\langle \Omega, P \rangle$] consists of a countable sample space Ω and a probability function $P : \Omega \to [0;1]$, s.t. $\sum_{\omega \in \Omega} P(\omega) = 1$

[Event] When a random variable X takes on a value x.

[Conditional/Posterior Probability] The probability

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)},$$

i.e. the chance that [def:event]event "a" takes place, given the event "b".

[Conditional Independence] Two [def:event]events a and b are conditionally independent, if $P(a \wedge b \mid c) = P(a \mid c)P(b \mid c)$.

[Probability Distribution] A vector for $\mathbf{P}(X)$ relating each element of the [def:prob-model]sample space to a probability:

$$\langle P(\omega_1), \dots, P(\omega_n) \rangle.$$

Related concepts:

Joint PD Given $Z \subseteq \{X_1, \dots, X_n\}$, results in a array the probabilities of all [def:event]events.

Full joint PD Joint PD for all random variables.

Conditional PD Given X and Y, results in a table for every probability $P(X \mid Y)$.

[Product Rule] $P(a \wedge b) = P(a \mid b)P(b)$

[Chain Rule] Extension of the [def:prod-rule]product rule,

$$P(X_1, \dots, X_n) = P(X_n \mid X_{n-1}, \dots, X_1) \dots P(X_2|X_1)P(X_1)$$

[Marginalisation] $\mathbf{P}(\mathbf{X}) = \sum_{y \in \mathbf{Y}} \mathbf{P}(\mathbf{X}, y)$

[Normalisation] Given $\mathbf{P}(X \mid e)$, and a normalization constant

$$\alpha = \frac{1}{P(x_1 \mid e) + \dots + P(x_n \mid e)},$$

normalization scales each element of the probability distribution s.t. $\sum \alpha \mathbf{P}(X \mid e) = 1$

[Bayes' Rule] Given two propositions a and b,

$$P(a \mid b) = \frac{P(b \mid a)P(a)}{P(b)},$$

where $P(a) \neq 0$ and $P(b) \neq 0$.

[Naive Bayes' Model] In this model, the [def:fjdp]full joint probability distribution is

$$\mathbf{P}(c \mid e_1, \dots, e_n) = \mathbf{P}(c) \prod_i \mathbf{P}(e_i \mid c),$$

i.e. a *single* cause c influences a number of [def:cond-indep]cond. independent effects e_i.