Defining conditional probability

In the discrete case

$$P\left(A_2|A_1\right)=\frac{P\left(A_1\cap A_2\right)}{P\left(A_1\right)}$$

In general, the definition is not so simple because we can't divide by zero. Instead, the definition is inspired by partitioning the outcome space:

$$\cup_{i=1}^{n}A_2^i=\Omega$$

$$A_2^i\cap A_2^j=0 \text{ for } i\neq j$$

$$i=1,...,n$$

$$P\left(A_1\right)=\sum_{i=1}^{n}P\left(A_1\cap A_2^i\right)$$

$$P\left(A_1\right)=\sum_{i=1}^{n}P\left(A_1|A_2^i\right)P\left(A_2^i\right)$$

.

Definition of conditional probability for arbitrary distributions

.

random variable X with distribution Q_X

collection of probability measures P_x – conditional distribution given X

requirements:

(1) $P_x\left(\{X=x\}\right)=1$except (...)

(2) (Borel measurability requirement)

(3) For all events $A \in \mathcal{A}$ ($\mathcal{A}$ is the collection of events in the probability space $(\Omega, \mathcal{A}, P)$):

$$P(A) = \int_{\mathbb{R}} P_x(A) \, dQ_X(x)$$

$P_x(A)$ is the same as $P(A|X = x)$