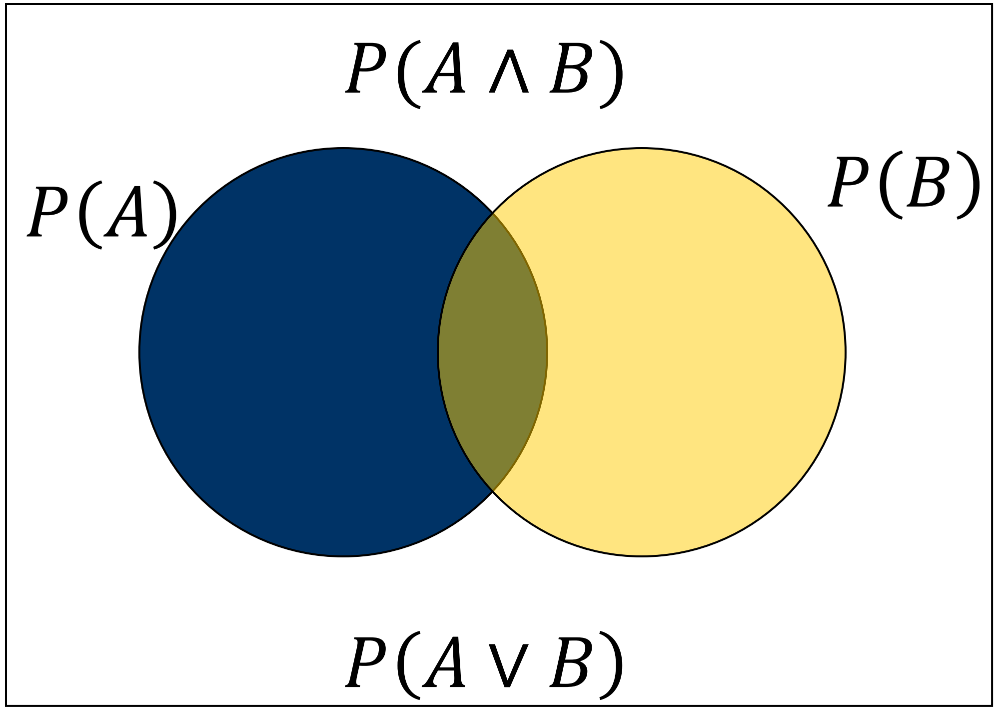
Axioms of Probability Theory

P(A) denotes probability that proposition A is true.

- $0 \leq P(A) \leq 1$
- $P(True) = 1$ $\qquad\qquad$ $P(False) = 0$
- $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

A Closer Look at Axiom 3

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$



Using the Axioms

$$\begin{aligned}
P(A \vee \neg A) &= P(A) + P(\neg A) - P(A \wedge \neg A) \\
P(True) &= P(A) + P(\neg A) - P(False) \\
1 &= P(A) + P(\neg A) - P(False) \\
P(\neg A) &= 1 - P(A)
\end{aligned}$$

Discrete Random Variables

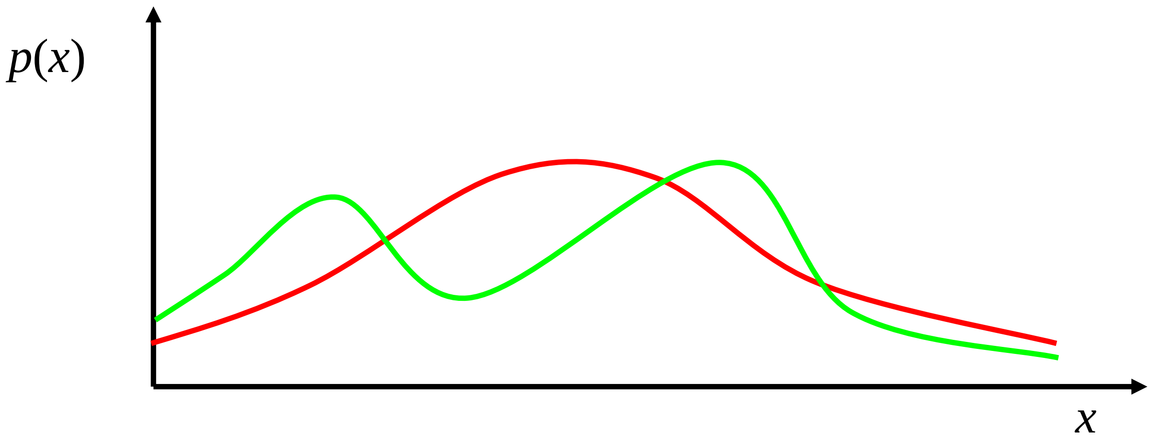
- X denotes a random variable
- X can take on a countable number of values in $\{x_1, x_2, \ldots, x_n\}$
- $P(X=x_i)$ or $P(x_i)$ is the probability that the random variable X takes on value x_i
- $P(\ )$ is called probability mass function
- E.g. $P(Room) =< 0.7, 0.2, 0.08, 0.02 >$

Continuous Random Variables

- X takes on values in the continuum.
- $p(X=x)$ or $p(x)$ is a probability density function

$$P(x \in [a,b]) = \int_a^b p(x)dx$$

- E.g.



“Probability Sums up to One”

Discrete case

$$\sum_{x} P(x) = 1$$

Continuous case

$$\int_{X} P(x)dx = 1$$

Joint and Conditional Probability

- $P(X{=}x \text{ and } Y{=}y) = P(x,y)$

- If X and Y are independent then

$$P(x,y) = P(x)\,P(y)$$

- $P(x \mid y)$ is the probability of x given y

$$P(x \mid y) = P(x,y) \,/\, P(y)$$

$$P(x,y) \quad = P(x \mid y)\,P(y)$$

- If X and Y are independent then

$$P(x \mid y) = P(x)$$

Law of Total Probability

Discrete case

$$P(x) = \sum_y P(x \mid y) P(y)$$

Continuous case

$$p(x) = \int p(x \mid y) p(y) dy$$

Marginalization

Discrete case

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

Continuous case

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

Bayes Formula

$$P(x, y) = P(x \mid y)P(y) = P(y \mid x)P(x)$$

$$P(x \mid y) = \frac{P(y \mid x)P(x)}{P(y)}$$

$$\text{Posterior} = \frac{\text{Likelihood} * \text{Prior}}{\text{Evidence}}$$

Normalization

$$P(x \mid y) = \frac{P(y \mid x)P(x)}{P(y)}$$

- At the same time: $P(y) = \sum_x P(y \mid x)P(x)$

$$P(x \mid y) = \frac{P(y \mid x)P(x)}{\sum_x P(y \mid x)P(x)}$$

- $P(y)$ is independent of x and thus constant for all x

$$P(x \mid y) = \eta P(y \mid x)P(x)$$

Bayes Rule with Background Knowledge

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

Conditional Independence

- $P(x, y \mid z) = P(x \mid z)P(y \mid z)$
- Equivalent to $P(x \mid z) = P(x \mid z, y)$ and $P(y \mid z) = P(y \mid z, x)$
- But this does not necessarily mean $P(x, y) = P(x)P(y)$
- Marginal independence does not mean independence