AI - Probability

Three types of Probability :

- Marginal | Unconditional Probability : Addition of probability of all values of Variable X, summing out all other values.
- Joint Probability : Probabilityies of any 2 or more occuring simultaneously.
- Conditional Probability : Probability of one event occuring provided another event has already occured.

Eg. In a deck of 52 cards.

Probability of choosing a red card is $\frac{26}{52} = \frac{1}{2} = 0.5$ $P(R)$

Probability of choosing a card '6' is $\frac{4}{52} = \frac{1}{13}$ $P(6)$

∴ Joint probability of choosing a red card with is also '6' $= P('R') \cdot P'6' = \frac{1}{13} \times \frac{1}{2} = \frac{1}{26}$.

written as $P(A \wedge B)$.

∴ Conditional Probability = $\left(\frac{1}{13} \times \frac{1}{2}\right) / \frac{1}{2} = \frac{1}{13}$

written as $P(A|B) = P(A \wedge B) / P(B)$

Formulae

- $P(a \wedge b) = P(a) \cdot P(b)$
- $P(a|b) = P(a \wedge b) / P(b)$
- $P(b|a) = P(a \wedge b) / P(a)$ $\qquad A \wedge B = B \wedge A$.
- $P(a|b) = \dfrac{P(b|a)\, P(a)}{P(b)}$

Same formulae for Probability Distributions $\underline{P}$:

- $\underline{P}(X|Y) = \underline{P}(X \wedge Y) / \underline{P}(Y)$

Probabilty Distribution P is used when dealing with a variable that can have multiple values, or set of variables.

- Marginal Probability

$\underline{P}(\underline{y}) = \sum_{z \in \underline{Z}} \underline{P}(\underline{y}, z)$

Here $\underline{y}, \underline{Z}$ are variable sets
$\underline{z}$ is a set of values.
Formula for JPD

- If conditional probability then then the formula is

$\underline{P}(\underline{y}) = \sum_{z \in \underline{Z}} \underline{P}(\underline{y} \mid z) P(\underline{z})$ → Here P is not underlined, so not a P.D. as $\underline{z}$ is set of values & not a variable/variable set

- $\underline{P}(X, y) = \underline{P}(X \mid y)\underline{P}(y)$. - Probabilities for all x & y values
- $\underline{P}(y \mid X) = \alpha \underline{P}(X \mid y)\underline{P}(y)$ - General Bayes with normalization

Conditional Independence with 3rd variable

- $\underline{P}(X, y \mid z) = \underline{P}(X \mid z) \underline{P}(y \mid z)$

Absolute Independence

- $\underline{P}(X \mid y, z) = \underline{P}(X \mid z)$ & $\underline{P}(y \mid X, z) = \underline{P}(y \mid z)$.

Naive Bayes

$P(\text{Cause}, \text{Effect}_1, \ldots, \text{Effect}_n) = \underline{P}(\text{Cause}) \prod_{i=1}^{n} \underline{P}(\text{Effect}_i \mid \text{Cause})$

- $P(x_1, \ldots x_n) = \prod_{i=1}^{n} P(x_i \mid \text{parents}(X_i))$.

- $P(x_1, \ldots, x_n) = P(x_n \mid x_{n-1}, \ldots x_1) P(x_{n-1} \mid x_{n-2}, \ldots, x_1) \ldots P(x_2 \mid x_1) P(x_1)$

$= P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i-1}, \ldots x_1)$.

$\therefore \underline{P}(X_i \mid X_{i-1}, \ldots, X_1) = P(X_i \mid \text{Parents}(X_i))$.

Markov Chain First-order ↓

$P(X_t \mid X_{0:t-1}) = P(X_t \mid X_{t-1})$

- $\underline{P}(A \mid B, \underline{e}) = (\underline{P}(B \mid A, \underline{e}) \times \underline{P}(A \mid \underline{e})) / \underline{P}(B \mid \underline{e})$

JPD in temporal models

$P(X_0, X_1, \ldots X_t, E_1, \ldots, E_t) = P(X_0) \prod_{i=1}^{t} P(X_i \mid X_{i-1}) P(E_i \mid X_i)$

- Ashish R. Gavande