ARTIFICIAL INTELLIGENCE & CYBER SECURITY

REASONING USING A FULL JOINT PROBABILITY DISTRIBUTION

Nacir Bouali

n.bouali@utwente.nl

UNIVERSITY OF TWENTE.

FULL JOINT PROBABILITY DISTRIBUTION

- A full joint probability distribution assigns a value for each random variable.

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

JOINT PROBABILITY

- The joint probability is the probability of assigning a value to each random variable.

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

- P(Headache=yes, Fever=no, Covid=Yes)=0.05
- P(Headache=yes, Fever=no, Covid=No)=0.05

MARGINAL PROBABILITY

- The marginal probability is a probability over a subset of the random variables.
- P(Headache=yes, Fever=no)=?

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

$$\sum_{x\in\{yes,no\}} P(Headache = yes, Fever = no, Covid = x)$$

- P(Headache=yes, Fever=no, Covid=Yes)=0.05
- P(Headache=yes, Fever=no, Covid=No)=0.05

MARGINAL PROBABILITY

- The marginal probability is a probability over a subset of the random variables.
- P(Headache=yes)=0.47

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

$$\sum_{x \in \{yes,no\}} \sum_{y \in \{yes,no\}} p(Headache = yes, Fever = y, Covid = x)$$

- P(Headache=yes, Fever=yes, Covid=yes)=0.25
- P(Headache=yes, Fever=yes, Covid=no)=0.12
- P(Headache=yes, Fever=no, Covid=yes)=0.05
- P(Headache=yes, Fever=no, Covid=no)=0.05

DISJUNCTION

- The disjunction rule states that $P(A\ or\ B) = P(A) + P(B) - P(A, B)$

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

- P(Headache=yes or Covid=no)=0.89

DISJUNCTION

- The disjunction rule states that $P(A \text{ } or \text{ } B) = P(A) + P(B) - P(A, B)$

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

- P(Headache=yes or Covid=no)=0.89
- It can also be calculated using the following formula:

$$P(headache \text{ } or \text{ } \neg covid) = 1 - P(\neg headache, covid)$$

DISJUNCTION

- I can also estimate more complex events

$$P(headache, \neg covid) \vee (fever, covid)$$

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

- =0.25+0.12+0.05+0.10=0.52

CONDITIONAL PROBABILITY

- $P(headache|covid) = P(headache, covid)/p(covid)$

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

Only if P(covid)>0

$P(A|B) = P(A, B)/P(B)$

= 0.30/0.41=0.73

PROBABILITY DISTRIBUTION

- Depicts the possible outcomes for all the given values of a random variable.
- $\boldsymbol{P}(Headache) = [P(headache), P(\neg headache)] = [0.47, 0.53]$
- $\boldsymbol{P}(Headache, Covid) = [[P(headache, covid), P(headache, \neg covid)],$ $[P(\neg headache, covid), P(\neg headache, \neg covid)]]$
- =[[0.3,0.17],[0.11,0.42]]
- The sum is always 1

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

CONDITIONAL DISTRIBUTION

- Depicts the possible outcomes for all the given values of a random variable.
- $\boldsymbol{P}(Headache|covid) = [P(headache|covid), P(\neg headache|covid)]$ = [0.73,0.27]
- $\boldsymbol{P}(Headache|Covid) = [[P(headache|covid), P(\neg headache|covid)],$
  $[P(headache|\neg covid), P(\neg headache|\neg covid)]]$

  =[[0.73,0.27],
  [0.29,0.71]]

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

CONDITIONAL DISTRIBUTION

- Depicts the possible outcomes for all the given values of a random variable.
- $\boldsymbol{P}(Headache|covid) = [P(headache|covid), P(\neg headache|covid)]$ = [0.73,0.27]
- $\boldsymbol{P}(Headache|Covid) = [[P(headache|covid), P(\neg headache|covid)],$
  $[P(headache|\neg covid), P(\neg headache|\neg covid)]]$
  =[[0.73,0.27],
  [0.29,0.71]]
- A conditional distribtuion can be seen as a distribution normalized so that it sums to 1.
- $\boldsymbol{P}(Headache|covid) = \boldsymbol{P}(Headache, covid)/\alpha$ = [0.3,0.11]/α
  $\alpha = 0.3 + 0.11$
  =[0.73,0.27]

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

CONDITIONAL DISTRIBUTION

- A conditional distribtuion can be seen as a distribution normalized so that it sums to 1.
- $\boldsymbol{P}(Headache|Covid) = [[P(headache, covid), P(\neg headache, covid)]/\alpha_{covid},$
  $[P(headache, \neg covid), P(\neg headache, \neg covid)]/\alpha_{\neg covid}]$

=[[0.3,0.11]/ α_{covid},
[0.17,0.42]/ $\alpha_{\neg covid}$]

α_{covid}=0.3+0.11=0.41

$\alpha_{\neg covid}$= 0.17+0.42=0.59

=[[0.73,0.27],
[0.29,0.71]]

Headache	Fever	Covid	Probability
Yes	Yes	Yes	0.25
Yes	Yes	No	0.12
Yes	No	Yes	0.05
Yes	No	No	0.05
No	Yes	Yes	0.10
No	Yes	No	0.10
No	No	Yes	0.01
No	No	No	0.32

OTHER TYPES OF RANDOM VARIABLES

- We work with random variables which are boolean or binary.
- This also applies to random variables which are discrete
  - $weather \in \{sunny, cloudy, rainy, snowy\}$
  - When you marginalize a variable you need to sum over all the values of the random variable.
- This also applies to random variables which are continuous.
  - posX=4.2
  - Marginalizing a variable requires integrals instead of summations.