CSE422: Artificial intelligence

Probability

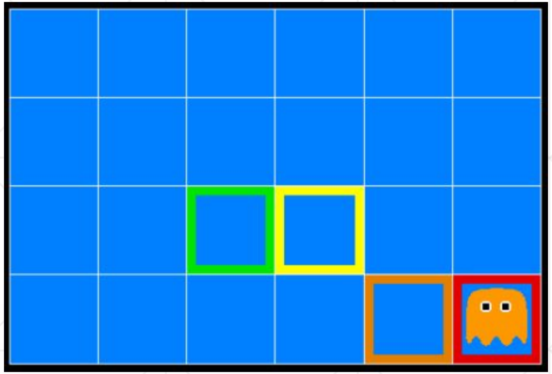
Asif Shahriar
Lecturer, CSE, BRACU

You PROBABLY Know These

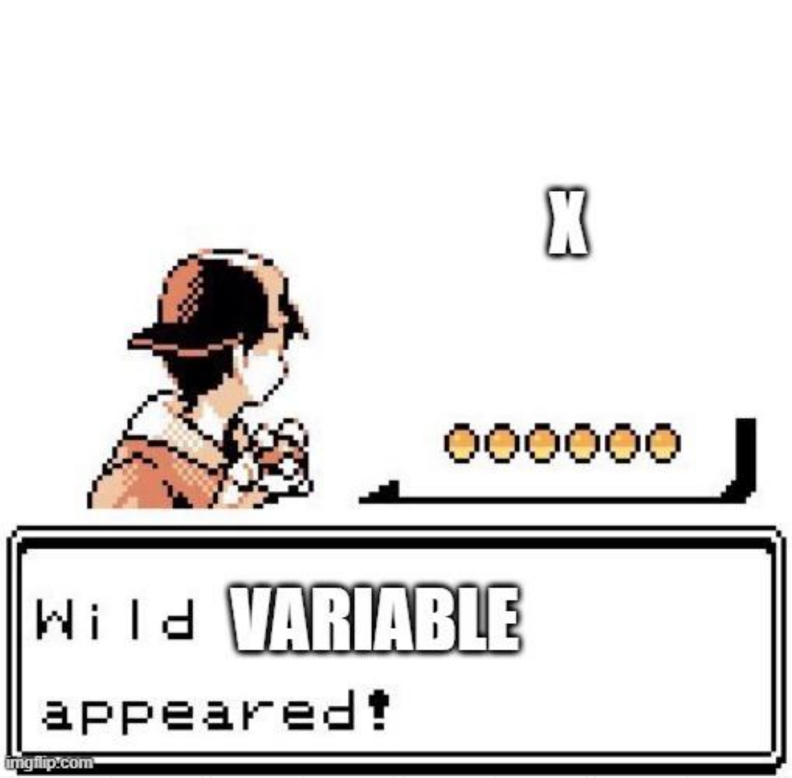
- Probability of an Event is defined by how LIKELY the event is to occur in an experiment
- If we repeat an experiment many times and record how often an event E happens, then probability of event E : P(E) = Number of times E occurs / Total number of trials
- In a coin toss, probability of heads occurring: P(head) = 1 / 2 = 0.5
- In a fair dice rolling (where all faces of the dice are equal)
  - Probability of getting 4: P(4) = 1 / 6
  - Probability of NOT getting 4: P(not 4) = 1 - P(4) = 1 - 1/6 = 5/6
  - Probability of getting an even number = P({2, 4, 6}) = P(2) + P(4) + P(6) = 3/6 = 0.5

WHY Probability

- There is always some uncertainty in environment
- Consider a game where there is a ghost in a grid, we need to avoid it
- Sensor readings tell how CLOSE a square is to a ghost
    - On the ghost: RED
    - 1 squares away: ORANGE
    - 2-3 squares away: YELLOW
    - 4+ squares away: GREEN
- Problem: Sensors CAN have noise, so UNRELIABLE
- Solution: Use the probability P(color | distance)



Random Variables



Random Variables

- A **Random Variable (RV)** is an aspect of the environment for which we have some uncertainty
- A Random variable can have ONE of a set of values from its **domain**, each with an associated probability
  - Example: Weather is a RV with domain = {sunny, rainy, cloudy, snowy}
- Domain values must be exhaustive and mutually exclusive
- Random variable can be **DISCRETE** or **CONTINUOUS**
  - Discrete RV can take a discrete value from a finite range. Example: weather, mark you will get in the MID
  - Continuous RV can take ANY value from an infinite / continuous range. Example: Temperature of a day
- Let X be a RV defined as the SUM of two fair dice, what are the domain values?

Probability Distribution

- A probability distribution assigns a probability to each value of a Random Variable
- Let X be a random variable that represents YOUR CURRENT MODE, with the domain being {bored, sleepy, annoyed, excited, anxious about mid marks}
  - P(X = bored) = 0.25
  - P(X = sleepy) = 0.20
  - P(X = annoyed) = 0.10
  - P(X = excited) = 0.05
  - P(X = anxious about mid marks) = 0.4
- We write this as: P(X) = <0.25, 0.20, 0.10, 0.05, 0.4> [called Vector Notation]
- The sum of all probabilities in a distribution must be 1 : TRUE or FALSE?

Some Notations

- Suppose **weather** is a discrete RV with the domain {sunny, rainy, cloudy, snowy}
  - *"Weather = sunny"* is written as *"sunny"*
  - *"P(Weather = sunny) = 0.6"* is written as *"P(sunny) = 0.6"*
- Suppose cavity is a discrete RV with the domain {TRUE, FALSE}
  - *"Cavity = TRUE"* is written as *"cavity"*
  - *"Cavity = False"* is written as *"¬cavity"*
  - *"P(Cavity = False) = 0.6"* is written as *"P(¬cavity) = 0.6"*
- Probability distribution is often presented in a table
  - P(Weather) = <0.6, 0.15, 0.24, 0.01> is the same as:

Weather	Probability
sunny	0.60
rainy	0.15
cloudy	0.24
snowy	0.01

Joint Probability Distribution

- Probability distribution over a SET of Random variables (i.e. two or more random variables are being considered)
- Let’s consider two boolean RVs: **Toothache** and **Cavity**

	toothache	¬toothache
cavity	0.22	0.15
¬cavity	0.08	0.55

- Sum of ALL entries in the joint probability distribution table must be 1
- Every question about a domain can be answered by looking at the table
- Quiz: P(cavity, toothache) = ? P(¬cavity, toothache) = ?
- Quiz: If I have cavity, probability of toothache is 0.22 -> TRUE or FALSE?

Marginal Probability Distribution

	toothache	¬toothache
cavity	0.22	0.15
¬cavity	0.08	0.55

- Probability of a proposition is the sum of probabilities of all events where that proposition holds TRUE
- P(cavity) = sum of all events where (Cavity=cavity) holds = 0.22 + 0.15 = 0.37
- P(¬toothache) = sum of all events where (Toothache=¬toothache) holds = 0.15 + 0.55 = 0.70
- Quiz: P(¬cavity) = ? P(toothache) = ?

AND OR

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cap B) = P(A) \times P(B)$ [ONLY IF A and B are independent, else not]
- $P(A \cup B)$ is also written as $P(A \lor B)$
- $P(A \cap B)$ is also written as $P(A \land B)$

Quiz on Probability Distribution

Let’s consider a Joint probability distribution with THREE boolean RV: H, B, S

Quiz on Probability Distribution

Let's consider a Joint probability distribution with THREE boolean RV: H, B, S

H	*B*	*S*	Prob.
+h	+b	+s	0.44
+h	+b	-s	0.15
+h	-b	+s	0.14
+h	-b	-s	0.07
-h	+b	+s	0.10
-h	+b	-s	0.04
-h	-b	+s	0.05
-h	-b	-s	0.01

Quiz on Probability Distribution

Let's consider a Joint probability distribution with THREE boolean RV: H, B, S

H	*B*	*S*	**Prob.**
+h	+b	+s	0.44
+h	+b	-s	0.15
+h	-b	+s	0.14
+h	-b	-s	0.07
-h	+b	+s	0.10
-h	+b	-s	0.04
-h	-b	+s	0.05
-h	-b	-s	0.01

- P(+h, -b, +s) = 0.14
- P(+b) = 0.44 + 0.15 + 0.10 + 0.04 = 0.73
- P((+h, -b) OR +s)

  = P(+h, -b) + P(+s) - P(+h, -b, +s)

  = (0.14 + 0.07) + (0.44 + 0.14 + 0.10 + 0.05) - 0.14

  = 0.80

Conditional Probability

- Conditional probability is the probability of an event A occurring **given** that another event B has already occurred
- Written as P(a|b) [read as Probability of a given b] and defined by:
  - $P(a \mid b) = P(a \cap b) / P(b)$
  - Also written as $P(a \mid b) = P(a, b) / P(b)$
- Here b is **evidence** (we know its probability) and a is **query** (we WANT to know its probability)

	toothache	¬toothache
cavity	0.22	0.15
¬cavity	0.08	0.55

Probability of toothache IF i have cavity

= P(toothache | cavity)

= P(toothache, cavity) / P(cavity)

= 0.22 / (0.22 + 0.15)

= 22 / 37 = 0.595

Conditional Probability

	Male	Female
GoT	0.16	0.24
BB	0.2	0.05
TBBT	0.1	0.25

- P(GoT | Male) = ?
- P(Male | GoT) = ?
- P(TBBT | Female) = ?

Conditional Probability

	Male	Female
GoT	0.16	0.24
BB	0.2	0.05
TBBT	0.1	0.25

- P(GoT | Male) = P(GoT, Male) / P(Male) = 0.16 / 0.46 = 0.3478
- P(Male | GoT) = P(Male, GoT) / P(GoT) = 0.16 / 0.4 = 0.4
- P(TBBT | Female) = ? Do it yourself

Marginal Probability

	toothache		¬toothache	
	shows	¬shows	shows	¬shows
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- Consider three Boolean RV Toothache, Cavity, and ShowsOnXRay
- P(cavity) = ?
- P(toothache) = ?
- P(cavity ∨ toothache) = ?

Marginal Probability via Conditioning

	toothache		¬toothache	
	shows	¬shows	shows	¬shows
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- If A and B are two boolean RVs,
  - $P(a) = P(a \land b) + P(a \land \neg b)$ [easier to find if Joint distribution table given]
  - In other words, $P(a) = P(a \mid b) \times P(b) + P(a \mid \neg b) \times P(\neg b)$ [if table not given]
- $P(cavity) = P(cavity \land shows \land toothache) + P(cavity \land \neg shows \land toothache) + P(cavity \land shows \land \neg toothache) + P(cavity \land \neg shows \land \neg toothache)$

BAYES' Rule

- From the formula of conditional probability $P(x \mid y) = P(x, y) / P(y)$, we can infer: $P(x, y) = P(x \mid y) \times P(y)$ [called the product rule]
- Similarly, $P(x, y) = P(y \mid x) \times P(x)$
- Dividing we get:

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

- In other words:

$$P(\text{cause}|\text{effect}) = \frac{P(\text{effect}|\text{cause})P(\text{cause})}{P(\text{effect})}$$

- Why is this SO important?
  - It is hard to know the cause, but it is easy to see the effect
  - Using this formula, we can infer the cause by analyzing the effect

BAYES' Rule

	Right handed	Left handed
Male	0.41	0.08
Female	0.45	0.06

- From the joint probability distribution given above, find the probability that a randomly selected person is:
  - Male given right handed
  - Right handed given male
  - Female given left handed
  - Are the events "being female" and "being left handed" INDEPENDENT?

BAYES' Rule

	Right handed	Left handed
Male	0.41	0.08
Female	0.45	0.06

- From the joint probability distribution given above, find the probability that a randomly selected person is:
    - Male given right handed [0.477]
    - Right handed given male [0.837]
    - Female given left handed [0.429]
    - Are the events "being female" and "being left handed" INDEPENDENT? [No]

BAYES' Rule

Example 1.14 A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply he has the disease.) If 0.5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive?

BAYES' Rule

Example 1.14 A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply he has the disease.) If 0.5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive?

Solution: Let D be the event that the tested person has the disease, and E the event that his test result is positive. The desired probability $P(D|E)$ is obtained by

$$
\begin{aligned}
P(D|E) &= \frac{P(DE)}{P(E)} = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\
&= \frac{(0.95)(0.005)}{(0.95)(0.005) + (0.01)(0.995)} \\
&= \frac{95}{294} \approx 0.323
\end{aligned}
$$

Note: If A and B are boolean RV, P(a | b) + P (¬a | b) = 1

BAYES' Rule

A person is brought in front of a jury. The jury finds the defendant guilty in 98% of the cases in which they indeed committed the crime; and it finds the defendant not guilty 97% of the cases when the defendant did not commit the crime. Only 0.8% of the population has committed a crime.

If a person is found guilty by the jury, what is the probability that person has truly committed the crime?

Answer: 49 / 235 = 0.2085

Chain Rule of Conditional Probability

- We know, $P(a, b) = P(a \mid b) \times P(b)$ [product rule]
- We can extend this using chain rule of multiplication
- $P(a, b \mid c) = P(a, b, c) / P(c)$
  - Hence, $P(a, b, c) = P(a, b \mid c) \times P(c)$
- $P(a, b, c, d, e) = P(a \mid b, c, d, e) \times P(b, c, d, e)$

$$= P(a \mid b, c, d, e) \times P(b \mid c, d, e) \times P(c, d, e)$$

$$= P(a \mid b, c, d, e) \times P(b \mid c, d, e) \times P(c \mid d, e) \times P(d, e)$$

$$= P(a \mid b, c, d, e) \times P(b \mid c, d, e) \times P(c \mid d, e) \times P(d \mid e) \times P(e)$$

Probabilistic Independence

- If two events A and B are independent, then the probability of A occurring does not depend on B
- So, P(A | B) = P(A)
- Hence, P(A ∧ B) = P(A | B) × P(B) = P(A) × P(B) [proved]
- This is also called **absolute independence**
- Consider the joint distribution of the RVs Toothache, Cavity, and ShowsOnXRay
  - Suppose we bring in another RV Weather = {sunny, cloudy, rainy, snowy}
  - Dental problems have no effect on weather
  - So, P(Weather = cloudy | toothache, cavity, shows) = P(Weather = cloudy)
  - P(Weather = cloudy, toothache, cavity, shows) = P(Weather) × P(toothache, cavity, shows)
  - **Caution!** P(toothache, clavity) ≠ P(toothache) × P(cavity) [toothache and cavity are not independent]

Probabilistic Independence

	Male	Female
GoT	0.16	0.24
BB	0.2	0.05
TBBT	0.1	0.25

- Are Male viewers and GoT independent?
- P(Male ∧ GoT) = ? P(Male)= ? P(GoT) = ?
- Is P(Male ∧ GoT) = P(Male) × P(GoT)?

Conditional Independence

- A and B are **conditionally independent** given C if $\mathbf{P(A \wedge B \mid C) = P(A \mid C) \times P(B \mid C)}$
- In that case we can decompose the joint distribution:
  - $P(A \wedge B \wedge C) = P(A \mid C)\, P(B \mid C)\, P(C)$ [Note: $P(A \wedge B \wedge C)$ is the same as $P(A, B, C)$]
- Consider the OG boolean RVs Toothache, Cavity, Shows
- If I have cavity, probability of X-Ray showing it does not depend on whether I have toothache or not
  - P(+shows | +toothache, +cavity) = P(+shows | +cavity)
- Same independence holds if I do not have cavity
  - P(+shows | +toothache, –cavity) = P(shows | –cavity)
- Hence, the random variable *Shows* is conditionally independent of *Toothache* given *Cavity*
  - P(Shows ∧ Toothache | Cavity) = P(Shows | Cavity) × P(Toothache | Cavity)

Conditional Independence

	toothache		¬toothache	
	shows	¬shows	shows	¬shows
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- From this join distribution, find out whether the random variables Shows and Toothache are independent given Cavity
  - Is P(Shows ∧ Toothache | Cavity) = P(Shows | Cavity) × P(Toothache | Cavity)?
    - P(Shows ∧ Toothache | Cavity) = ?
    - P(Shows | Cavity) = ?
    - P(Toothache | Cavity) = ?

Conditional Independence

	toothache		¬toothache	
	shows	¬shows	shows	¬shows
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- From this join distribution, find out whether the random variables Shows and Toothache are independent given Cavity
  - Is P(Shows ∧ Toothache | Cavity) = P(Shows | Cavity) × P(Toothache | Cavity)?
    - P(Shows ∧ Toothache | Cavity) = 0.54
    - P(Shows | Cavity) = 0.9
    - P(Toothache | Cavity) = 0.6

Exercise

	smart		¬smart	
	study	¬study	study	¬study
prepared	.432	.16	.084	.008
¬prepared	.048	.16	.036	.072

Queries:

- What is the prior probability of smart?
- What is the prior probability of study?
- What is the conditional probability of prepared, given study and smart?
- Save these answers for later! ☺️

Exercise

	smart		¬smart	
	study	¬study	study	¬study
prepared	.432	.16	.084	.008
¬prepared	.048	.16	.036	.072

Queries:

- Is smart independent of study?
- Is prepared independent of study?

Exercise

	smart		¬smart	
	study	¬study	study	¬study
prepared	.432	.16	.084	.008
¬prepared	.048	.16	.036	.072

Queries:

- Is smart conditionally independent of prepared, given study?
- Is study conditionally independent of prepared, given smart?