Probability Basics

Frederik Mallmann-Trenn
6CCS3AIN

Probability basics

- Begin with a set Ω—the sample space.
- This is all the possible things that could happen.
  - 6 possible rolls of a die.
  - How many if I have two dice?
- $\omega \in \Omega$ is a sample point, atomic event.

(pngimg.com)

Probability basics

- A probability space or probability model is a sample space with an assignment $P(\omega)$ for every $\omega \in \Omega$ such that:

$$0 \leqslant P(\omega) \leqslant 1$$
$$\sum_{\omega} P(\omega) = 1$$

- For a typical die:
  $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6.$

(pngimg.com)

(hartsport.com.au)

Probability basics

- An event A is any subset of Ω

$$P(A) = \sum_{\{\omega \in A\}} P(\omega)$$

- Again, for a regular die:
  $P(\text{die roll} < 4) = P(1) + P(2) + P(3) = 1/6 + 1/6 + 1/6 = 1/2$

(pngimg.com)

(hartsport.com.au)

Random variables

- A random variable is a function from sample points to some range.
  - $raining(London) \in \{true, false\}$.
  - $temperature(lecture\ room) \in \{0, 1, \ldots, 30\}$.
- P induces a probability distribution for any r.v. X:

$$P(X = x_i) = \sum_{\{\omega : X(\omega) = x_i\}} P(\omega)$$

- In our dice example, we could set $\omega =$ die shows an odd number:

$$\begin{aligned} P(Odd = true) &= P(1) + P(3) + P(5) \\ &= 1/6 + 1/6 + 1/6 \\ &= 1/2 \end{aligned}$$

Propositions

- We describe the world in terms of propositions, which are mathematical statements such as “the die shows an odd number’ or “it is raining”.
- Think of a proposition as the event (set of sample points) where the proposition is true
- Given Boolean random variables A and B:

  event a = set of sample ω points where $A(\omega) = true$
  event $\neg a$ = set of sample points ω where $A(\omega) = false$
  event $a \wedge b$ = points ω where $A(\omega) = true$ and $B(\omega) = true$

- Example:

  The set of sample points is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

  $A(\omega) = true$ or simply a is the event that the number is odd, i.e., $\{1, 3, 5\}$.

  $B(\omega) = true$ or simply b is the event that the number is < 4, i.e., $\{1, 2, 3\}$.

  $a \wedge b$ is given by the sample points in $\{1, 3\}$.

Propositions

- A state can be defined by a set of Boolean variables.

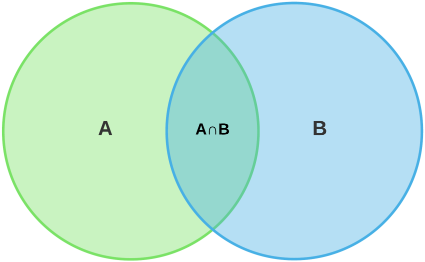
$$a \wedge b \vee \neg c \qquad A = true, B = true, C = false$$

This is then just a sample point.

Union and Intersection

- The definitions imply that certain logically related events must have related probabilities

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$



(lucidchart.com)

Example: Let a denote the event that the die shows an odd number. Let b denote the event that the die shows a number < 4. Hence, the probability that we get an odd number or a number < 4 is
$P(a \vee b) = P(a) + P(b) - P(a \wedge b) = \frac{1}{2} + \frac{1}{2} - P(\mathit{roll}\ 1\ \mathit{or}\ 3) = 2/3.$

Prior and posterior probability

- Prior or unconditional probabilities of propositions

  $P(Cavity = true) = 0.1$ and $P(Weather = sunny) = 0.72$ correspond to belief before (prior) to arrival of any (new) evidence.

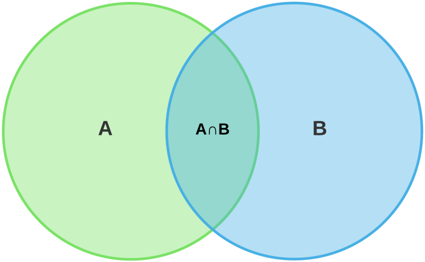
- In contrast, $P(Cavity = true \mid toothache) = 0.2$ is the posterior or conditional probability. Here, we have additional evidence.

Conditional Probabilities

- Conditional Probabilities

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

assuming $P(b) > 0$.



(lucidchart.com)

Example: Let a denote the event that the die shows an odd number. Let b denote the event that the die shows a number < 4. Hence, the probability that we get an odd number given that the is number < 4 is $P(a|b) = \frac{P(a \wedge b)}{P(b)} = \frac{1/3}{1/2} = 2/3$.

Syntax for propositions

- Propositional or Boolean random variables
  - $Cavity$ (do I have a cavity?)
  - $Cavity = true$ is a proposition, also written $cavity$
  - $Cavity = false$ is a proposition, also written $\neg cavity$ or $\overline{cavity}$
- Discrete random variables (finite or infinite)
  - $Weather$ is one of $\{sunny, rain, cloudy, snow\}$
  - $Weather = rain$ is a proposition
- Continuous random variables (bounded or unbounded)
  - $Temp = 21.6$; also allow, e.g., $Temp < 22.0$.

Syntax for propositions

- We allow arbitrary combination of logical operators (AND, OR, NOT) and comparison operators ($<, \leqslant, =, \neq, \ldots$).
- E.g., $Temp < 22.0$ AND $Cavity = true$

Notation

- Probability distribution gives values for all possible assignments (assumes a fixed ordering):

$$\mathbf{P}(Weather) = \begin{pmatrix} 0.72 \\ 0.1 \\ 0.08 \\ 0.1 \end{pmatrix} \tag{1}$$

  means
  - $P(Weather = sunny) = 0.72$,
  - $P(Weather = rain) = 0.1$,
  - $P(Weather = cloudy) = 0.08$ and
  - $P(Weather = snow) = 0.1$
- Values must be exhaustive (everything covered) and mutually exclusive (no overlap)
- Values have to sums to 1
- Note that the book uses the notation $\langle 0.72, 0.1, 0.08, 0.1 \rangle$

Quiz

There is a KEATS quiz to see if you understood. Have a look!