Week 4: Introduction to Probability

Unit 3: Understanding Bayes' Theorem

Introduction

- In the 18th century, Thomas Bayes came up with an approach to how statistical probabilities should change *in response to new evidence*.
- Bayes' theorem provides us with the ability to update our beliefs based on prior knowledge of factors that might be related to the event.
- In this lesson, you should remember that P(A|B) means the probability of outcome 'A' given evidence 'B'.
  - For example, the probability of getting lung cancer, given that the person is a smoker:

    P(Lung Cancer | Smoker)

Where can we use Bayes' theorem?

What is the likelihood that a CEO of a stock market listed company will be fired, given that the company's share price underperforms its competitors' by more than 10% over a year?

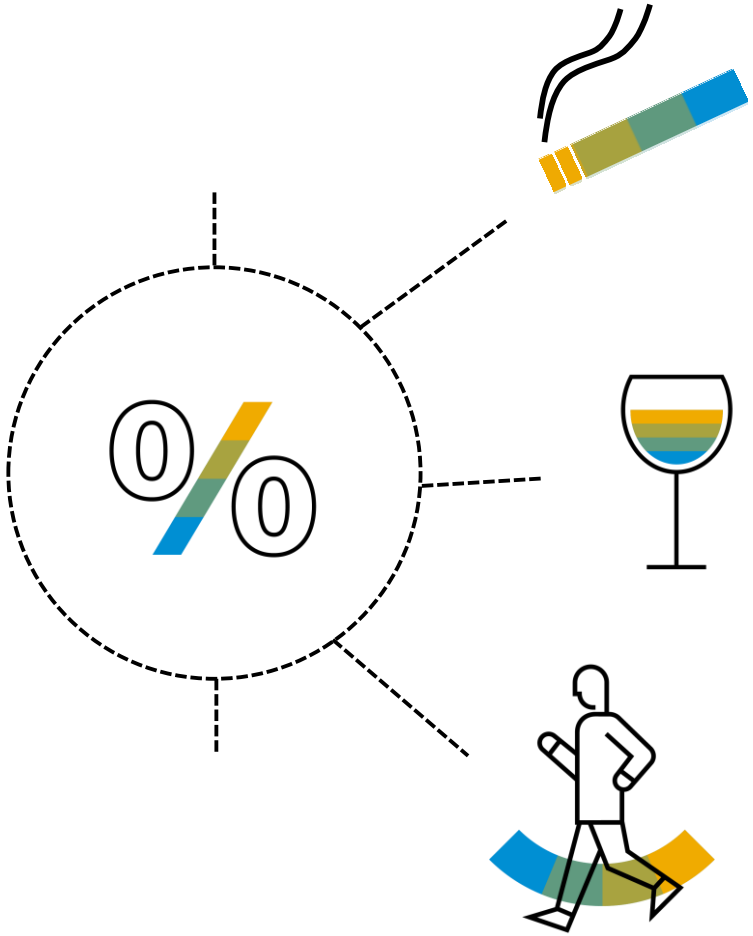
You have developed an app to review travel expense claims, after an audit found that 3% of them contain errors.

- The app finds errors in 94% of the examples picked up in the audit and in 8% of the claims with no errors.
- What is the likelihood of there actually being an error, given the expense claim has been marked as such by your app?

What is the likelihood of people who work outdoors without a hat getting skin cancer, given they work in the UK?

Medical test example

- We know that a certain % of the population will get lung cancer, but we want to know how that is affected by prior conditions like drinking, smoking, or lack of exercise.
- Bayes' theorem allows us to calculate the likelihood (probability) of someone getting cancer given that they smoke.



Bayes' theorem – The formula

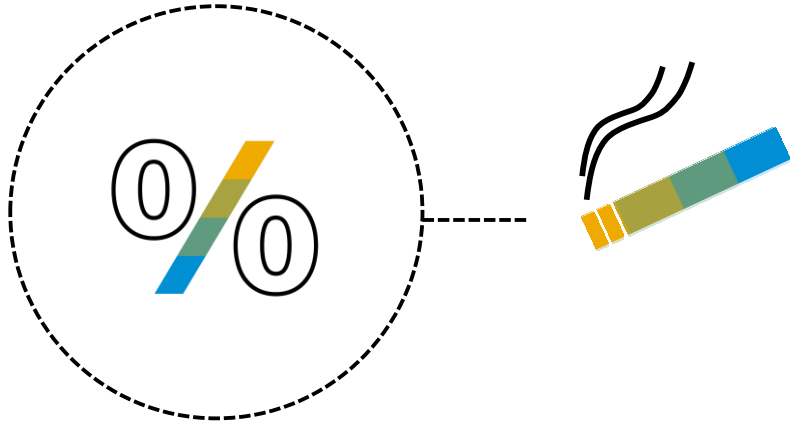
$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

- P(A|B) is the probability of event A (such as 'cancer') given B (if we know the person is a smoker).
- P(B|A) is the probability of B given A.
- Therefore, the formula calculates the probability of event A, given the prior information we have about event B.

What's the probability of getting cancer from smoking?

Lets assume that the prior probabilities are:

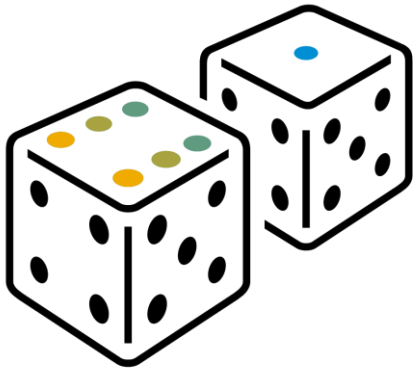
- 4% of the population have some form of cancer
- 20% of the population smoke
- 60% of people with cancer smoke,
  i.e. P(Smoker|Cancer)



What's the probability of getting cancer from smoking?

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

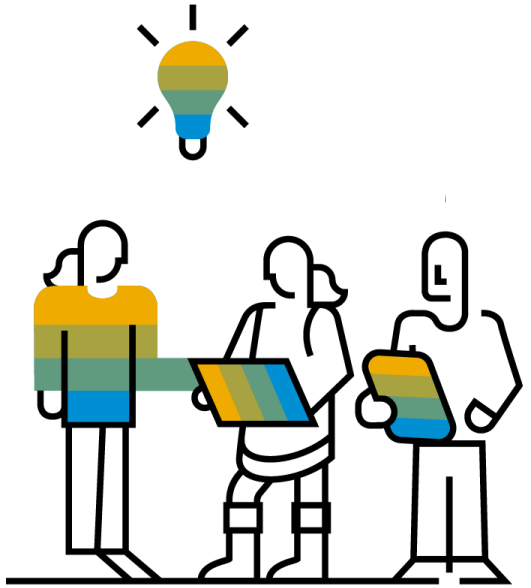
- The probability of the evidence conditional on the hypothesis – probability of being a smoker given you have cancer:
  $P(B|A) = 0.6$
- The prior probability of the hypothesis – having cancer:
  $P(A) = 0.04$
- The prior probability of the evidence – simple prior probability of being a smoker:
  $P(B) = 0.2$



What's the probability of getting cancer from smoking?

We will use S for smoker and C for cancer:

$$P(C|S) = \frac{P(S|C) \times P(C)}{P(S)} = \frac{(0.6 \times 0.04)}{0.2} = 0.12\ (12\%)$$

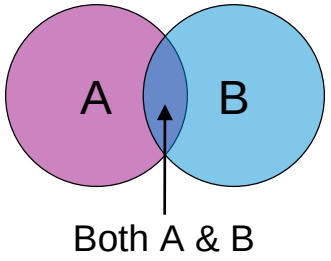


Summary

- You can use Bayes' theorem to allow probabilities to be re-evaluated based on new evidence.
- Bayes' theorem allows you to evaluate the likelihood of an event based on the occurrence of a prior event.
- The formula for Bayes' theorem:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

VENN DIAGRAM!



- The probability of A occurring is calculated based on the relevant event B, which has occurred.

Thank you.

Contact information:

open@sap.com

Follow all of SAP

www.sap.com/contactsap

THE BEST RUN SAP®