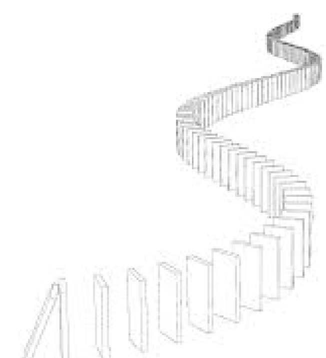
COMP 1002

Logic for Computer Scientists

Lecture 29



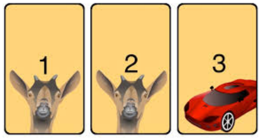
- **Conditional probability** of an event A given event B, denoted $\Pr(A|B)$, is the probability of A if we know that B occurred.
  – Probability of car a behind door 2 if we chose door 1, and door 3 had a goat behind it.
- So it is probability of both A and B, given that we know B happened for sure:
$$\Pr(A|B) = \frac{\Pr(A \cap B\ )}{\Pr(B)}$$
– Assume that $\Pr(B) >\ 0$: after all, B did happen.

Conditional probabilities

- **Conditional probability** of an event A given event B, denoted $\Pr(A|B)$, is the probability of A if we know that B occurred.
  – Probability of car a behind door 2 if we chose door 1, and door 3 had a goat behind it.
- So it is probability of both A and B, given that we know B happened for sure:

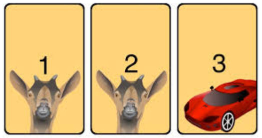
$$\Pr(A|B) = \frac{\Pr(A \cap B\ )}{\Pr(B)}$$

  – Assume that $\Pr(B) >\ 0$: after all, B did happen.



Bayes theorem

- Consider a medical test that
  - Has false positive rate of 3% (healthy labeled as sick)
    - Specificity 97%
  - Has false negative rate of 1% (sick labeled as healthy).
    - Sensitivity 99%
- What is the probability that a person has the disease given that the test came positive?
  - Let A: person tested positive, B: person is sick. $\Pr(B|A)$?
  - Pr(A|B) = 0.99, $\Pr(\bar{A}|B) = 0.01\ldots$
- Not enough information!

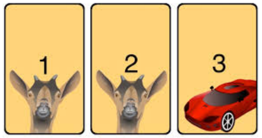


Bayes theorem

- **Bayes theorem** allows us to get $\Pr(B|A)$ from $\Pr(A|B)$, if we know probabilities of A and B:

$$\Pr(B|A) = \frac{\Pr(\mathrm{A}|\mathrm{B})\cdot\Pr(B)}{\Pr(A)} = \frac{\Pr(\mathrm{A}|\mathrm{B})\cdot\Pr(B)}{\Pr(A|B)\,\Pr(B)+\Pr(A|\bar{B})\Pr(\bar{B})}$$

- Proof:
  - $\Pr(A|B) = \frac{\Pr(A\cap B)}{\Pr(B)}$. $\Pr(B|A) = \frac{\Pr(A\cap B)}{\Pr(A)}$.
  - $\Pr(\mathrm{A} \cap B) = \Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A)$
  - So $\Pr(B|A) = \frac{\Pr(A|B)\,\Pr(B)}{\Pr(A)}$



Bayes theorem

- Consider a medical test that
  - Has false positive rate of 3% (healthy labeled as sick).
  - Has false negative rate of 1% (sick labeled as healthy).
  - Tests for a disease that occurs in 5 in 1000 people.
- What is the probability that a person has the disease given that the test came positive?
  - Let A: person tested positive, B: person is sick. $\Pr(B|A)$?
  - Pr(A|B) = 0.99, Pr($\bar{A}|B$) $= 0.01$. Pr($\bar{A}|\bar{B}$) = 0.97 , Pr($A|\bar{B}$) $= 0.03$
  - $\Pr(B) = 0.005.$
  - $\Pr(A) = \Pr(A|B)\Pr(B) + \Pr(A|\bar{B})\Pr(\bar{B})$ = 0.0348
  - By Bayes theorem, Pr(B|A) = $\frac{\Pr(A|B)\Pr(B)}{\Pr(A)}$ = 0.1422
- So the probability that a person who tested positive has the disease is just 0.1422, that is, 14.22%.
  - By a similar argument, probability that a person who tested negative does not have a disease is whopping 0.99995 = 99.995%.

Hat-check problem

- Suppose n men came to an event, and checked in their hats at the door.
  - On the way out, in a hurry, they each picked up a random hat.
  - On average, how many men picked their own hat?

Expectations

- Often we are interested in what outcome we would see "on average".
  - How fast does this program run "on average"?

- Suppose that possible outcomes of an experiment are numbers $a_1, \ldots, a_n$
  - E.g., time a program takes to sort n elements

- Its **expected value (mean)** is $\Sigma_{k=1}^{n} a_k \mathrm{Pr}(a_k)$
  - Often phrased in terms of a "random variable" X, where X is a *function* from outcomes to numbers.
  - Write E(X) to mean the expectation of X.

Random variables

- Random variable X: a function from outcomes to numbers. $X: S \to \mathbb{R}$
  - Number of heads in 3 coin tosses
    - X(HHH)=3,
    - X(HHT)=X(HTH)= H(THH)=2,
    - X(THT)=X(TTH)=X(HTT)=1,
    - X(TTT)=0
  - Number of steps until a program crashes
  - Number of items in a hash table until collision

- Distribution of X: set of pairs $(r,\ Pr(X = r))$
  - Number of heads in 3 coin tosses: distribution of X is
    - (0,1/8), (1,3/8), (2,3/8), (3,1/8)

- Expectation of X: $E(X) = \sum_{r \in range(X)} r \cdot \Pr(X = r)$
  - X: number of heads in 3 coin tosses. Range(X) ={0,1,2,3}
  - $E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2}$

Expected win in a lottery

- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of $5,000,000 or more.
    - There are also smaller prizes; let's ignore them for simplicity.
  - A ticket costs $3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

Expected win in a lottery

- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of $5,000,000 or more.
    - There are also smaller prizes; let's ignore them for simplicity.
  - A ticket costs $3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

- What is the expected amount a player would win if the jackpot is 5,000,000?
  - Pr(win) = 1/14,000,000. Pr(loss)=1-Pr(win) = 13,999,999/14,000,000.
  - Let the random variable X encode the amount a player wins.
    - For all but one player, that amount is -3. So Pr(X=-3)=Pr(loss)
    - For the lucky one, the amount is the jackpot minus ticket price. Pr(X=4,999,997)=Pr(win)
  - Expected amount to win is E(X) = Pr(loss)*(-3) +Pr(win)*(5,000,000-3) = *-2.64*
    - If counting smaller prizes , just add their amount*odds to the sum, and adjust Pr(loss)
    - E(X)=Pr(loss)*(-3)+Pr(jackpot)*(4,999,997)+Pr(5/6+bonus)*374,997+Pr(5/6)*312,497…

Expected win in a lottery

- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of \$5,000,000 or more.
  - A ticket costs \$3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

- How large should be the jackpot so that the players expect at least to break even?
  - Let's call the jackpot amount J.
  - Expected amount to win is E(X) = Pr(loss)*(-3) +Pr(win)*(J-3).
    - To break even, want E(X)=0.
  - J= 3+(E(X)-Pr(loss)*(-3))/Pr(win) = *42,000,000*

Saturday, March 25, 2017
MAIN DRAW
05-09-14-24-30-35
Bonus: 21
GUARANTEED PRIZE DRAW
45768958-02

Expected number until…

- Suppose that Alan insists on buying lottery tickets until he wins. If probability of winning is p, how many tickets in expectation (“on average”) would he have to buy?
  - Let X be a random variable for how many tickets he has to buy.
  - The probability of winning on exactly i^{th} ticket is $p(1-p)^{i-1}$
  - $E(X) = \Sigma_{i \in \mathbb{N}}\, i * \Pr(X = i) = \frac{1}{p}$
    - So for Lotto 6/49 he’d have to buy 14,000,000 tickets (and spend \$42,000,000 -- that’s jackpot that would let him break even!)

- Same reasoning applies to other processes, where there is a fixed probability of something happening at each experiment or time step.
  - A system has a 1% probability of hanging in any given hour. How long, on average, will it stay up?
    - 100 hours: a little over 4 days.

Linearity of expectation

- Expectation is a very well-behaved operation:
  - $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$
  - $E(aX + b) = a\,E(X) + b$
    - Where $X_1 \dots X_n$ are random variables on some sample space S, and $a, b \in \mathbb{R}$

- Proof:
  - $E(X_1 + X_2) = \Sigma_{s \in S}\, p(s)(X_1(s) + X_2(s))$
    $= \Sigma_{s \in S}\, p(s)X_1(s) + \Sigma_{s \in S}\, p(s)X_2(s)$
    $= E(X_1) + E(X_2)$
  - Similar for $E(aX + b) = a\,E(X) + b$
    - Using the fact that $\Sigma_{s \in S}\, p(s)$ =1

Hat-check problem

- Suppose n men came to an event, and checked in their hats at the door.
  - On the way out, in a hurry, they each picked up a random hat.
  - How many men are expected to have picked their own hat?
- For each man, introduce a random variable X_i, where $X_i = 1$ iff he picked his own hat
  - Such random variables are called **indicator variables**.
  - The quantity we want is $E(X_1 + \cdots + X_n)$
  - Now, for each $X_i, E(X_i) = 1 \cdot \Pr(X_i = 1) = \frac{1}{n}$
  - By linearity of expectation, $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = \frac{1}{n} + \cdots + \frac{1}{n} = n \cdot \frac{1}{n} = 1$

- So on average, just one man will go home with his own hat!