Mathematics for Artificial Intelligence
Compilation: Yogesh Kulkarni

Basics

Land of Mathematics

High School Mathematics

- Which Mathematics related words do you remember?
- Which Mathematics related branches do you know?
- Mathematics is the Language of Science.
- It provides grammar, words, structure.
- Our first encounter: Equations ...

Numbers

Numbers

- $x = 3 \Rightarrow \mathbb{N}$: Natural numbers: 1,2,3,...
- $x + 5 = 3 \Rightarrow \mathbb{Z}$: Integers: -1,0,1,2,3,...
- $2x = 3 \Rightarrow \mathbb{Q}$: Rational numbers, ratios: $\frac{3}{2}$, $\frac{1}{3}$,...
- $x^2 = 2 \Rightarrow \mathbb{P}$: Irrational numbers, cannot be expressed as ratios: $\sqrt{2}$, $\sqrt[3]{2}$,....
- $\pi, e \Rightarrow \mathbb{R}$: Real numbers, these cannot be represented by polynomials, cannot be roots, etc.
- $x^2 + 1 = 0 \Rightarrow \mathbb{C}$: Complex numbers: $i, 2 + 3i$,....All polynomials have roots in $\mathbb{C}$

$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{P} \subseteq \mathbb{R} \subseteq \mathbb{C}$
Note: all the coefficients in the equation are $\mathbb{N}$ but the resultant x is of different types.

Equations

Algebra

- Equation 'equates' two expressions. E.g. $3x + 4 = 10$
- We solve equations for unknowns such as x
- x in the equation above have coefficient 3.
- Others like 4 and 10 are constants.

Equations

$3x + 4 = 10$

- To Solve, we isolate x on one side by operations.
- $3x + 4 - 4 = 10 - 4$
- Although a short hand way is just to transfer 4 from one side to other and changing the sign.
- $3x = 6$
- In case of coefficients, you divide after crossing the $=$
- $3x/3 = 6/3$
- $x = 2$
- To test, plug $x = 2$, back in the original equation and see if both sides 'equate'.

The Distributive Property

- $3(x + 2) = 18$ is same as $3x + 6 = 18$
- Intuitively we multiply everything from outside, as if its a coefficient.
- Now instead of just 3 I have another expression to multiply $(3x + 3)(x + 2) = 18$
- Same distributive property works?

Two Variables

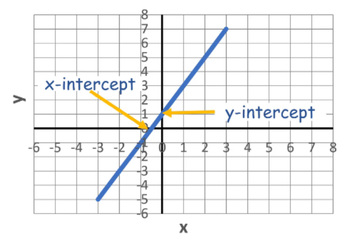
- $2y - 4x = 2$
- How to solve?
- Can we solve?

Two Variables

- $2y - 4x = 2$
- Isolate y first
- $2y = 2 + 4x$
- $y = 1 + 2x$
- Thats y's definition in terms of x.
- For any x, y can be calculated.
- x, y if plotted form a line, thus the original equation is called as Linear Equation.

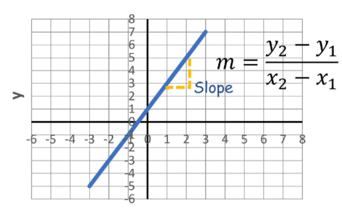
Line Plot

- X intercept is where line crosses x axis, where y is 0 (put in the eqn)
- Y intercept is where line crosses y axis, where x is 0 (put in the eqn)



Line Plot

- Slope is change in y per change in x
- Calculated using any two points.
- Slope-Intercept form is $y = mx + b$



Exercise

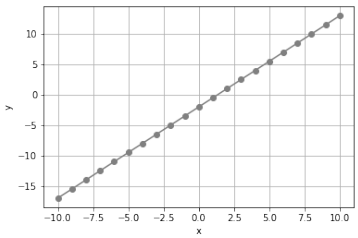
Find slope and both intercepts of $2y + 3 = 3x - 1$

Exercise

```
import pandas as pd

# Create a dataframe with an x column containing values from
    -10 to 10
df = pd.DataFrame ({'x': range(-10, 11)})

# Add a y column by applying the solved equation to x
df['y'] = (3*df['x'] - 4) / 2

print(df)
```

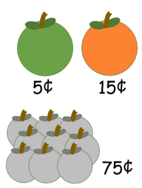
Exercise

```python
from matplotlib import pyplot as plt

plt.plot(df.x, df.y, color="grey", marker = "o")
plt.xlabel('x')
plt.ylabel('y')
plt.grid()
plt.show()
```



System of Equations

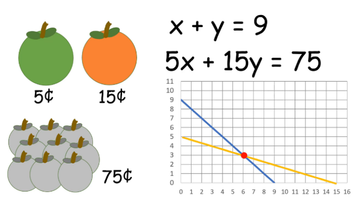
9 items cost 75c, how many are apples and how many oranges?



(Ref: Essentials of Mathematics - DAT 256 EdX)

System of Equations

Let number of apples be x and oranges be y.

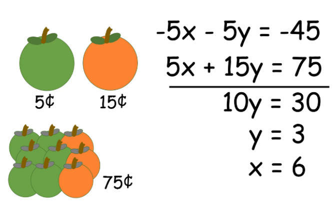


(Ref: Essentials of Mathematics - DAT 256 EdX)

Blue line represents $x + y = 9$ and yellow line $5x + 15y = 75$. Their intersection, the common point shared by two equations, is the solution.

System of Equations

System of Linear equations can be solved, apart from line intersection method, by elimination method.



(Ref: Essentials of Mathematics - DAT 256 EdX)

Make one of the equations (say, first) as NEGATIVE of the other in one variable, so that their addition cancels it out, leaving only one variable. Easy to find solution. Put that value back in the original equation, to get the other variable value.

Exercise

Solve $x + y = 16$ and $10x + 25y = 250$

Exponentials

Power

- $3 \times 3 = 3^2 = 9$
- $3 \times 3 \times 3 = 3^3 = 27$
- $4^7 = 16384$ here, 4 is the base, and 7 is the power or exponent in this expression.
- The 'raised to' can be seen as whole number.
- Can it be a fraction?
- Can it be negative?

Radicals (Roots)

- Sometimes you'll need to calculate one or other of the elements themselves.
- $?^2 = 9$
- This expression is asking, given a number (9) and an exponent (2), what's the base?
- In other words, which number multiplied by itself results in 9?
- This type of operation is referred to as calculating the root, and in this particular case it's the square root $\sqrt{9} = 3$
- $36^{1/2} = ?$
- $81^{1/3} = ?$

Radicals (Roots)

```python
import math

# Calculate square root of 25
x = math.sqrt(25)
print (x)

# Calculate cube root of 64
cr = round(64 ** (1. / 3))
print(cr)
```

Logarithms

- To determine the exponent for a given number and base.
- In other words, how many times do I need to multiply a base number by itself to get the given result.
- $4^x = 16$, what is x?
- Such solutions are found by logarithms
- $x = \log_4(16) = 2$
- Log finds the POWER to given base, of a number.
- Whats Log of 1000 to base 10?

Logarithms

```python
import math

# Natural log of 29
print (math.log(29))

# Common log of 100
print(math.log10(100))
```

Logarithms

- We know of log to base 10 and those log tables. But is there a function/formula?
- What is a log? Is it defined for all real values?
- $10^{2.3} = 10^{\frac{23}{10}}$ Thats 10th root of 10^{23}
- $10^{-0.07} = \frac{1}{10^{0.07}}$
- $10^{\sqrt{2}} = 10^{(2^{1/2})}$. Need to approximate $\sqrt{2}$ to some decimal places and then evaluate.
- $10^x = y$ then $\log y = x$
- x can be any real number and y will be corresponding unique number.

Sets

Sets

Example 1. The set consisting of all positive integers less than 10 can be denoted by $\{1,2,3,4,5,6,7,8,9\}$. $\{1,2,3,\ldots,9\}$ is also used since the general pattern is obvious.

Sets

Example 2. Members of a set need not be numeric. The set of all possible outcomes when tossing a coin is $\{H,T\}$.

Set Builder Notation

For example, $O=\{1,3,5,7,9\}$ can also be written as

$$O=\{x : x \text{ is a positive odd integer less than } 10\}.$$

Example 3.
- $A=\{x : x \text{ is a positive even number less than } 15\}=\{2,4,6,8,10,12,14\}$.
- $B=\{n : n \text{ is a positive integer}\}=\{1,2,3,4,5,\ldots\}$.
- $C=\{2n : n \text{ is a whole number and } 1 \le n \le 4\}=\{2,4,6,8\}$.

Some Basic Notations

- $\emptyset$, The empty set (a set with *no* elements). $\{\quad\}$ is also used.
- $\mathbb{N}=\{1,2,3,\ldots\}$, the set of natural numbers.
- $\mathbb{Z}=\{\ldots,-2,-1,0,1,2,\ldots\}$, the set of integers.
- $\mathbb{R}$ or $(-\infty,\infty)$, the set of real numbers.
- $[a,b]$, the set of all real numbers x such that $a \le x \le b$.
- (a,b), the set of all real numbers x such that $a < x < b$.
- $(a,b]$, the set of all real numbers x such that $a < x \le b$.
- $[a,b)$, the set of all real numbers x such that $a \le x < b$.

Subsets and Equality

Definition 4. A set A is said to be a subset of another set B if every element of A is also an element of B. We use the notation $A \subseteq B$. Two sets A and B are said to be equal (notation $A=B$) if they have the same elements. In other words, A and B are equal if $A \subseteq B$ and $B \subseteq A$.

Subsets and Equality

Example 5. $\{2,3,5\} \subseteq \{-1,2,3,5,7\}$. However, $\{1,2,3,4\} \not\subseteq \{1,3,4,5,6,7\}$, because $2 \notin \{1,3,4,5,6,7\}$.

Example 6. By definition, $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{R}$.

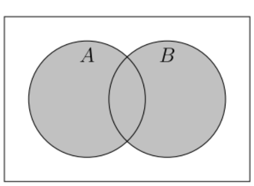
Set Operations

Let A and B be sets.

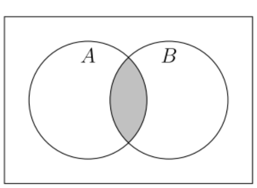
- The union of A and B, denoted by $A \cup B$, is the set that contains those elements that are either in A or B, or in both. In other words, $A \cup B=\{x : x \in A \text{ or } x \in B\}$.
- The intersection of A and B, denoted by $A \cap B$, is the set that contains those elements in both A and B. In other words, $A \cap B=\{x : x \in A \text{ and } x \in B\}$.
- Two sets are said to be disjoint if their intersection is the empty set.
- Let S be the universal set (the set of all elements under consideration. It depends on the context). Then the complement A^c of A is defined to be the set of all elements in S that are not in A.

Set Operations, continued

Example 7. Let $A=\{1,3,4,5\}$ and $B=\{1,2,3\}$, then $A \cup B=\{1,2,3,4,5\}$ and $A \cap B=\{1,3\}$.



Union



Intersection

Set Operations, continued

Example 8. If $P=(-\infty,2)$ and $Q=[-1,\infty)$, then $P \cap Q=[-1,2)$ and $P \cup Q=(-\infty,\infty)=\mathbb{R}$.

Set Operations, continued

Example 9. Let $O=\{2k+1 : k \in \mathbb{Z}\}$ and $E=\{2k : k \in \mathbb{Z}\}$, then O and E are disjoint. If one takes $\mathbb{Z}$ as the universal set, then $O^c=E$ and $E^c=O$.

Set algebra laws

Commutative law:
$A \cap B = B \cap A$
$A \cup (B \cup C) = (A \cup B) \cup C$
Associative law:
$(A \cap B) \cap C = A \cap (B \cap C)$
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Distributive law:
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Cardinality of the Union of Sets

Let $|S|$ denote the number of elements of a set S. Obviously, $A \cup B$ contains both A and B, so $|A \cup B| \ge |A|$ and $|A \cup B| \ge |B|$. But exactly how many elements are there in $A \cup B$?

Theorem 10. *Let A,B be sets with finitely many elements, then*

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

In particular, if A and B are disjoint, then $|A \cup B| = |A| + |B|$.

Cardinality of the Union of Sets

Theorem 11. *Let A,B,C be sets with finitely many elements, then*

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|.$$

Probability

Why Probability?

- Real life scenarios are inherently random or too complex to be completely known.
- If you knew every "Physics" aspect of motion of a dice, you could predict the outcome, deterministically.
- However this can never be done in practice, so there is a CHANCE of something happening.

Probability: the Likeliness

- How likely something is to happen.
- Many events can't be predicted with total certainty.
- The best we can say is how likely they are to happen, using the idea of probability.
- Probability is the fraction of time that outcome would occur with repeated experiments.

Probability: Example

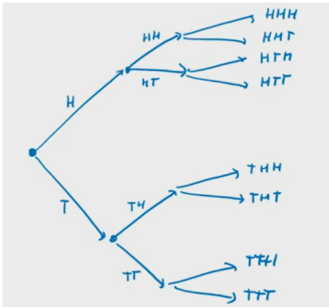
- Probability (event) = Number of ways it can happen / Total number of outcomes
- Example: the chances of rolling a "4" with a die
- Number of ways it can happen: 1 (there is only 1 face with a "4" on it)
- Total number of outcomes: 6 (there are 6 faces altogether)
- Probability of this event = 1/ 6

Probability: Example

- There are 5 marbles in a bag: 4 are blue, and 1 is red.
- What is the probability that a blue marble gets picked?
- Number of ways it can happen: 4 (there are 4 blues)
- Total number of outcomes: 5 (there are 5 marbles in total)
- Probability of this event = 4 / 5

Probability: Example

- Flip 3 coins. What is the probability that we get exactly two heads?
- Flip 1 coin, you get either H or T. Then flip 2nd coin, draw further branches (like below)
- At the end you will have 8 possible outcomes



(Ref: Math for Machine Learning - AWS Brent Werness)

- What we want are HHT, HTH, THH. So 3 out of 8.
- Probability of this event = 3 / 8

Probability

Definition 12. Given an experiment, let S be the set of all possible outcomes (S is often called the sample space) and A be the event of some particular outcomes (so $A \subseteq S$). The probability of A, denoted by $P(A)$, is the relative proportion of A in S.

Probability

Example 13. Suppose we are rolling a fair die, then $S = \{1, 2, 3, 4, 5, 6\}$. Let A be the event that the outcome is an even number, that is, $A = \{2, 4, 6\}$. Then $P(A) = 1/2$.

Example 14. Suppose we are rolling two fair dice, and let A be the event that the sum of outcomes is either 6 or 7. Then $P(A) = 11/35$.

Some terminologies

- Experiment or Trial: an action where the result is uncertain. Example: Tossing a coin, throwing dice are examples of experiments.
- Outcome: A single possibility from the experiment. E.g. $\Omega = HHT, THT, \ldots$, each of these are outcomes.

Event

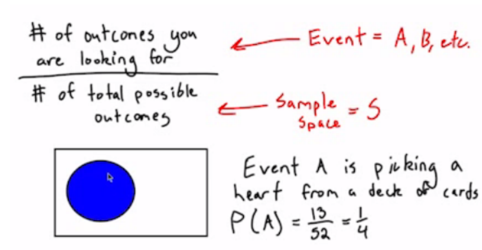
- Event: a set of outcomes/results of an experiment
- An event is what we are looking for ue $E = ExactlyTwoHeads$. E is any subset of the sample space. The set of all events associated with a given experiment is called the event space.
- Example Events:
  - Getting a Tail when tossing a coin is an event
  - Rolling a "5" is an event.
- An event can include one or more possible outcomes:
  - Choosing a "King" from a deck of cards (any of the 4 Kings) is an event
  - Rolling an "even number" (2, 4 or 6) is also an event

Sample space

- Sample Space: all the possible outcomes of an experiment. Example: choosing a card from a deck
  - There are 52 cards in a deck (not including Jokers)
  - So the Sample Space is all 52 possible cards: Ace of Hearts, 2 of Hearts, etc.
- A Sample Point is just one possible outcome.
- And an Event can be one or more of the possible outcomes.
- Flipping the coin: the sample space is $\{H, T\}$.
- Whats the sample space of number of heads if we flip a coin 10 times?
- $\{0, 1, \ldots, 10\}$

Probability of an event?

- Both 'event' and 'probability' are intuitively understood by most people, but we need to establish certain rules.
- 'Rain tomorrow', '3 or fewer cyclones next year', 'Crop yield will exceed a given threshold' are all examples of 'events' whose 'probabilities' might be of interest.



Axioms* of Probability

- Probability is between 0 and 1. ie $P(E) \in [0, 1]$
- Outcome is always from the sample space.
- Summation of probabilities of all possible outcomes is 1

* Axioms are statements considered to be true without any proof.

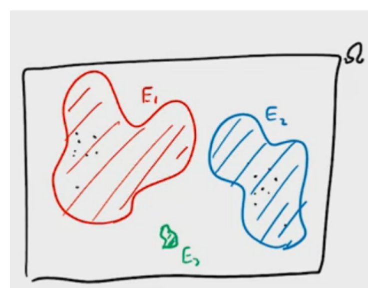
Properties of Probability

Let S be the sample space and A, B be events. Then

- $P(S) = 1, P(\emptyset) = 0$.
- If $A \subseteq B$, then $P(A) \leq P(B)$. In particular, $0 \leq P(A) \leq 1$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. In particular, if A and B are disjoint, then $P(A \cup B) = P(A) + P(B)$.
- $P(A^c) = 1 - P(A)$.

Visualizing Probability

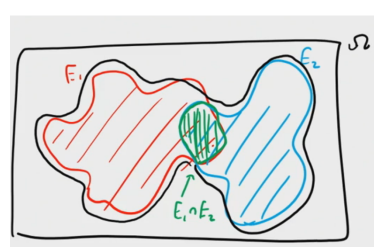
- Let rectangular region being of entire sample space. Area is "1".
- Points are outcomes
- Events are collection of outcomes, so sub-regions.
- Probability is shown as area of the sub-region.



(Ref: Math for Machine Learning - AWS Brent Werness)

Visualizing Probability

- Union of Events is sum of probability when they are Exclusive.
- If they are not, meaning there is some intersection between events, then total probability is sum of both minus the intersection (as it got included twice)
- For 3 events summation, add all 3, minus 3 double overlaps, plus add single triple overlap.



(Ref: Math for Machine Learning - AWS Brent Werness)

Probability of multiple variables

Joint Probability

- To denote the probability of multiple variables AT THE SAME TIME.
- Say, two variables, gender and hair-length. $P(male, short)$ is the probability that a person is male and has short hair.

Conditional Probability

- To denote the probability of one event AFTER THE OTHER HAS OCCURED.
- $P(longhair|male)$ is the probability of having long hair given that a person is male.

Marginal Probability

- To denote the probability of one event IRRESPECTIVE OF OTHER EVENTS.
- Probability of person being male is given by $P(male)$. It doesn't matter whether or not the person has short hair or long.

Relation between joint, conditional and marginal probabilities

- $P(A, B) = P(B|A)P(A)$
- $P(male, longhair) = P(longhair|male)P(male)$
- That is, probability of a person being male and having long hair = probability of person being male times probability of having long hair given person is male.
- Bayes theorem is a relation between conditional and marginal probabilities of two variables $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

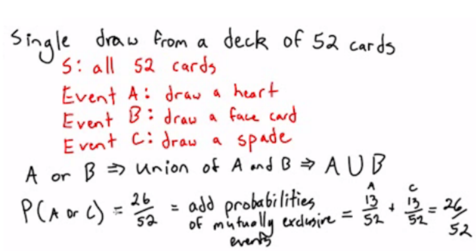
Probability Formulation

Unions and intersections

- The union of two events A and B consists of everything included in A or B or both. Let
  - A = rain tomorrow
  - B = rain the day after tomorrow
  - C = 3 or fewer cyclones
  - D = 4 or 5 cyclones
- Then
  - $A \cup B$= rain in the next 2 days
  - $C \cup D$ = 5 or fewer cyclones
- $P(C \cup D) = P(C) + P(D)$, because C and D are mutually exclusive (they don't overlap).

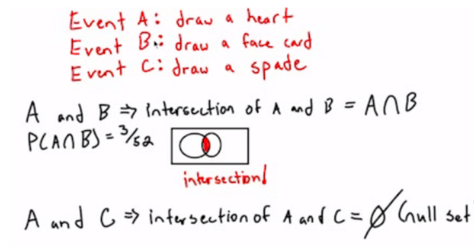
Unions and intersections

- $P(A \cup B) \neq P(A) + P(B)$ because A and B do overlap.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- $P(A \cap B)$ is the intersection of A and B; it includes everything that is in both A and B, and is counted twice if we add P(A) and P(B).



Unions and intersections

- In our example
  - $A \cap B$ = (rains tomorrow and the day after tomorrow).
  - $C \cap D$ is empty - it is impossible for C and D to occur simultaneously, so $P(C \cap D) = 0$.



Joint Probability

- Multiple events and we are trying to find chance of happening all of them together
- So? Intersection set of all events
- If A and C are independent, $P(A \cup C) = P(A)P(C)$.

Independence

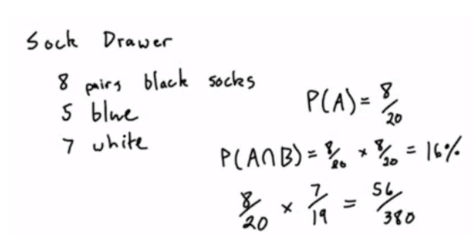
- Two events are independent, if the occurrence of one does not affect the probability of occurrence of the other.
- Similarly, two random variables are independent if the realization of one does not affect the probability distribution of the other.
- Example: rolling dice. First outcome, does not influence second outcome at all.
- In pack of cards: pick a card and then pick a card again, are independent if the first card is put back.

Independence

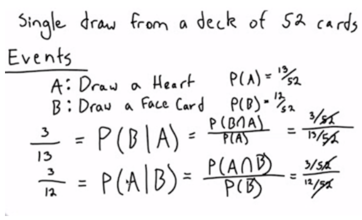
- In dice example
  - A = getting 3
  - B = getting 6
  - C = getting even number
- Total outcome of 2 events, 6x6 pairs of possible outcomes, e.g. 1-1, 1-2,1-3 . . . , 6-6.
- $P(A \cap B) = P(A)P(B) = 1/6 \times 1/6 = 1/36$.
- $P(A \cap C) = P(A)P(C) = 1/6 \times 3/6 = 3/36$.
- **NOTE: A and C are not outcome of same event, which would have been NULL. They are two separate events, one after another, thus there is non-zero probability**

Without replacement?

- First pair pulled is black
- Second pair getting black without putting the first one back.
- These are Dependent events. So answer is not pure multiplication, but with some adjustment.
- Conditional Probability is for Dependent probabilities



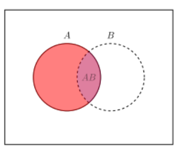
Conditional probability



- Calculate probabilities of A and B, independently.
- but then if I say, that the first card is Heart, then whats the probability of Face Card.
- There are 13 hearts. Among hearts there are 3 face cards. So, 3 /13. 3 face-heart cards, total 13 hearts.
- reverse, first face, then whats prbability of getting heart. So, 12 face cards and 3 out of them are hearts. So, 3/12.

Conditional probability

- If we know that one event has occurred it may change our view of the probability of another event. Let
- A = rain today, B = rain tomorrow.
- It is likely that knowledge that A has occurred will change your view of the probability that B will occur.
- We write $P(B|A) \neq P(B), P(B|A) = P(A \cap B)/P(A)$ denotes the conditional probability of B, given A.



Conditional Probability using Venn Diagram

	A	B	C	Total
Male	8	18	13	39
Female	10	4	12	26
Total	18	22	25	65

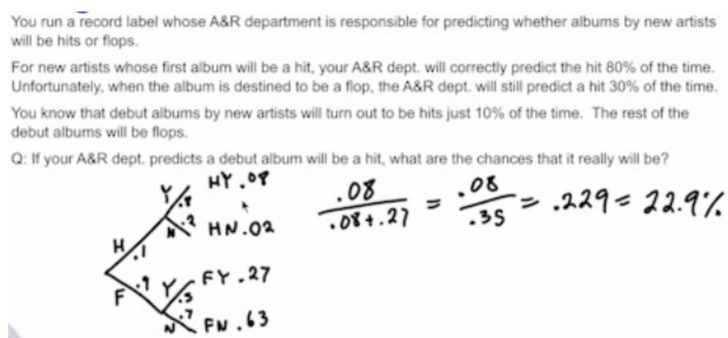
If one student was chosen at random, find the probability that the student got an A given that he is male.

$P(\text{A}|\text{male}) = \frac{8}{39}$ Find the probability that a student is male given that he got an A. $P(\text{male}|\text{A}) = \frac{8}{18}$

Conditional Probability Formula

If events A and B are not independent, then $P(A \text{ and } B) = P(A) \cdot P(B|A)$. If you pull 2 cards out of a deck, what is the probability that both are twos? The probability that the first card is a two is $\frac{4}{52}$. The probability that the second card is a two, given that the first was a two, is $\frac{3}{51}$. The probability that both cards are twos is $\frac{4}{52} \cdot \frac{3}{51} = \frac{12}{2652} \simeq 0.00452$.

Graphical Conditional Probability



Copyleft © Send suggestions to yogeshkulkarni@yahoo.com