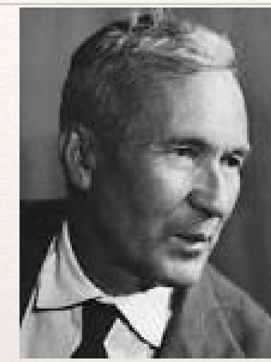
LS1 Probability Theory

Probability Elements



- Andrey Nikolaevich Kolmogorov (1903-1987)
- Russian (Soviet) mathematician made major contribution to theory of probability
- Probability tuple (Ω,Ƒ,P)
  - Ω – Sample space, space of all possible $\underbrace{\text{outcomes,}}_{}$ of that experiment e.g. Ω = { 1,2,3,4,5,6}
    - Event – subset of sample space, e.g. {1}, {2,3}
  - event space Ƒ -- Set of desired events built on sample space, e.g. primes {2,3,5} Also Omega, null set - subsets of Ω
  - P -- Probability mapping of given Ƒ to real number line in interval [0,1]
  - Σ -- [redacted], all sets of outcomes, all subsets of the sample space; for countable space, this is the power set or all subsets of Ω, including Ω and Ø and |Σ| or cardinality of Σ = $2^{|\Omega|}$ The largest possible event space for Ω
  - The power set has a relationship to the binomial theorem (which we will learn about) and provides us with an important counting rule
- Later we will definite for more complex events a function called a Random Variable (RV) which maps set of events to the real number line

Sample And Event Space Examples

Working with a Sample Space, Part I

1. **You roll two six-sided dice**:
    1. How would you define an appropriate sample space, Ω?
    2. How many elements exist in Ω?
    3. What is an appropriate event space, and how many elements does it have?
    4. Give an example of an event.

Working with a Sample Space, Part II

2. **For a random sample of 1,000 Berkeley students**:
    1. How would you define an appropriate sample space, Ω?
    2. How big is Ω? How many elements does it contain?
    3. What is an example of an event for this scenario?
    4. Can a single person be represented in the space twice? Why or why not?

Working with a Sample Sapce, Part III

3. **Suppose that you're sitting in a surf lineup, and you have to pick a wave that is the right height. Too small, and you won't get anywhere, too large and you'll get crushed.**
    1. What sample space is appropriate to represent the height of a single wave, Ω?
    2. How big is Ω? How many elements does it contain?
    3. What is an example of an event that could be part of the event space?
    4. What sample space is appropriate to represent the height of the next 10 waves? How large is this sample space?

To represent 10 waves, you should use $\mathbb{R}^{10}$. It is an interesting mathematical fact that $\mathbb{R}$ and $\mathbb{R}^{10}$ actually have the same cardinality. There exists a 1-to-1 function between these sets.

Sample And Event Space Examples-Solns

Working with a Sample Space, Part I

1. **You roll two six-sided dice:**
   1. How would you define an appropriate sample space, Ω? *Let's assume that you can't tell the dice apart. Use unordered pairs of integers:* $\Omega = \{\{a, b\} : a, b \in \{1, 2, 3, 4, 5, 6\}\}$.
   2. How many elements exist in Ω? *This is a little tricky because, for example,* $\{2, 3\}$ *and* $\{3, 2\}$ *are the same outcome. There are* $6 \cdot 5/2$ *outcomes in which the numbers don't match, and* 6 *more in which the numbers match, for a total of* 21.
   3. What is an appropriate event space, and how many elements does it have? *For countable sample spaces, you will want to use the power set:* $2^\Omega = \{S : S \subseteq \Omega\}$. = 2^21 = 2,097,152
   4. Give an example of an event. *Rolling two numbers that are the same:* $e = \{(a, b) \in \Omega : a = b\}$

Working with a Sample Space, Part II

2. **For a random sample of 1,000 Berkeley students:**
   1. How would you define an appropriate sample space, Ω? *First, define* B *to be the set of all Berkeley students. Then let* $\Omega = \{S : S \subseteq B, |S| = 1000\}$.
   2. How big is Ω? How many elements does it contain? $\binom{|B|}{1000}$
   3. What is an example of an event for this scenario? *The event that student Anika Patel is in the sample:* $E = \{\omega \in \Omega : \text{"Anika Patel"} \in \omega\}$
   4. Can a single person be represented in the space twice? Why or why not?

No, independence and uniqueness are required

Working with a Sample Sapce, Part III

3. **Suppose that you're sitting in a surf lineup, and you have to pick a wave that is the right height. Too small, and you won't get anywhere, too large and you'll get crushed.**
   1. What sample space is appropriate to represent the height of a single wave, Ω? *The real numbers* $\Omega = \mathbb{R}$. *If there is a clear minimum value (such as the bottom of the ocean), the nonnegative real numbers would work.*
   2. How big is Ω? How many elements does it contain? *There are uncountably many real numbers. Mathematicians use the symbol* c *(for continuum) to represent the cardinality of the reals.*
   3. What is an example of an event that could be part of the event space? *The event that the wave is over 5 foot:* $e = \{\omega \in \Omega : \omega > 5\}$
   4. What sample space is appropriate to represent the height of the next 10 waves? How large is this sample space? *You should use 10 real numbers to represent 10 waves:* $\Omega = \mathbb{R}^{10}$. *You may be interested to know that even though this set seems much larger than* $\mathbb{R}$, *it actually has the same cardinality. That is, there is a one-to-one map between the sets.*

Counting rules

- Permutations (ordered events) – Types: Repetition allowed or repetition not allowed
    - Repetition allowed: n^k (no reduction of choice) – three number lock combo
    - Repetition not allowed (reduction of choice) – orderings of 3 billiard balls out of 16

$$P_{k,n} = \frac{n!}{(n-k)!}$$

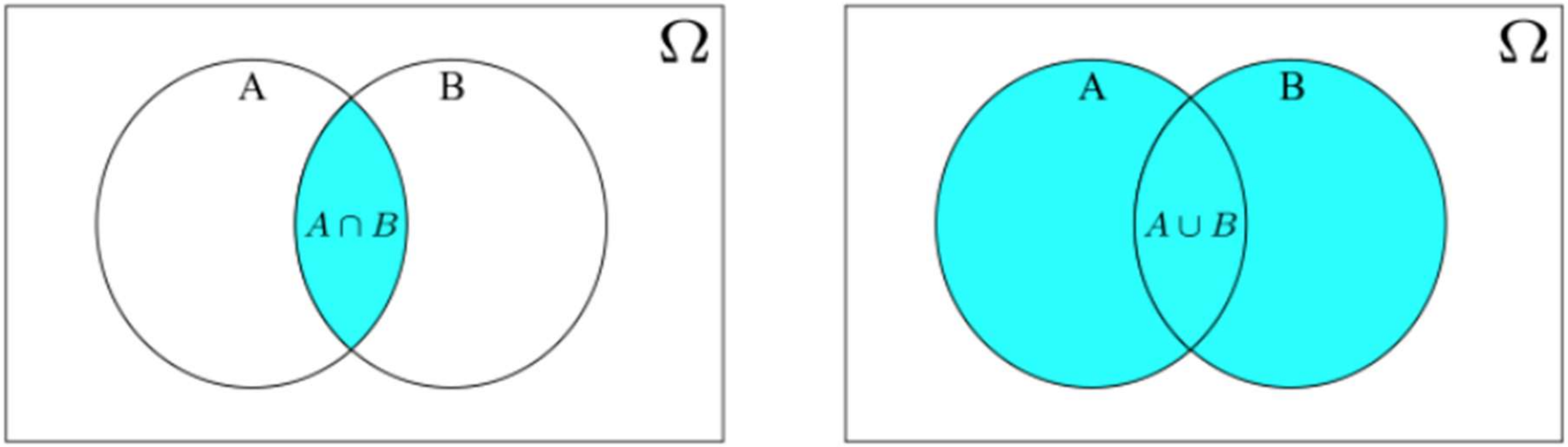
- Combinations ("n choose k" - unordered events) Types: Repetition allowed or repetition not allowed
    - Repetition not allowed – assume order matters, then ignore, |Perm(1,2,3)| = 6, but one combination

$$\binom{n}{k} = \frac{P_{k,n}}{k!} = \frac{n!}{k!(n-k)!}$$

Also known as binomial coef

n = group size , *k*= subset size

Events



** 1.2 Event $A \subset \Omega$ **: Is a collection of outcomes

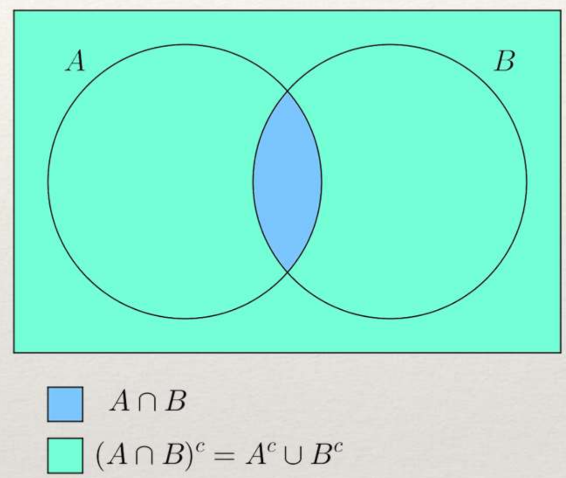
- $A = \{2, 4, 6\}$ is the event of rolling a 6 sided die and getting an even number
- $B = \{(H, H, T), (H, T, H), (T, H, H), (H, H, H)\}$ is the event that you fill a coin 3 times and get a least 2 heads.

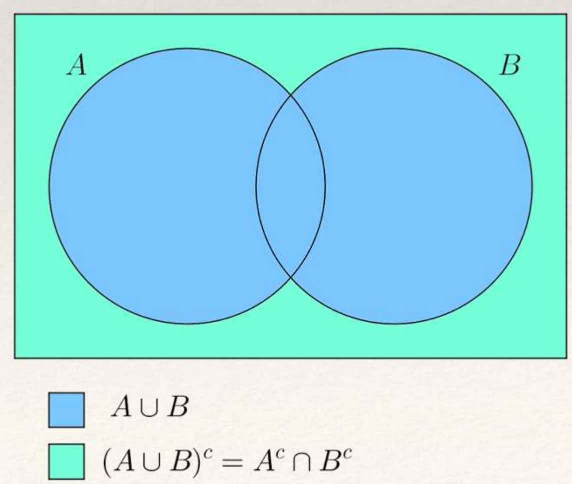
Union and Intersection

- For $A \cup B$, say Event in A or B
- For $A \cap B$, say Event in A and B

De Morgan’s Laws (https://brilliant.org/wiki/de-morgans-laws/)

- **De Morgan's Laws** relate the intersection and union of sets through complements.
- A, B are finite sets
- $(A \cap B)^c = A^c \cup B^c$
- $(A \cup B)^c = A^c \cap B^c$
- Generalization
- $(\bigcap_{k=1}^{n} A_k)^c = \bigcup_{k=1}^{n} A_k^c$
- $(\bigcup_{k=1}^{n} A_k)^c = \bigcap_{k=1}^{n} A_k^c$





Mutually Exclusive

- Events A and B are said to be mutually exclusive events if their intersection is the null space that is $A \cap B = \varnothing$ which is equivalent to $P(A \cap B) = 0$
- Events A and A^c are mutually exclusive

Axioms of Probability

- $P(A) >= 0$, A in Ω
- $P(\Omega) = 1$
- $P(\cup A(i)) = \sum P(A(i))$, $\{i = 1, \ldots, \infty\}$, For a countable sequence of disjoint events or $P(A(i) \cap A(j)) = \emptyset$, for all $i \neq j$

Important Laws And Results

- Law of Additive Probability
- Law of Total Probability
- Law of Multiplicative Probability
- Conditional Probability
- Bayes Rule
- Law of Complementary Probability

Events act like areas

- Addition rule for mutually exhaustive events

$$P(X \cup Y) = P(X) + P(Y)$$

- Addition rule (general form)

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$

- What is the difference can you draw a picture to illustrate this?

Proof of Law of Addition

- Prove $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B) = P(A) + P(B \setminus A)$
  - By countable infinite or finite and disjoint
- $= P(A) + P(B \setminus (A \cap B))$ (1)
- $P(B) = P(B \setminus (A \cap B)) + P(A \cap B)$
- ➔ $P(B) - P(A \cap B) = P(B \setminus (A \cap B))$ (2)
- Substituting (2) into (1) we have what we want

Conditional probability

$$\text{conditional probability, } P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$$

Multiplication rule

- Multiplication Rule

  $P(X \cap Y) = P(Y)P(X|Y)$

- if X and Y are independent (one event does not control the happening of another)

  $P(X \cap Y) = P(Y)P(X)$ (definition)

- Extended to three or more events:

  $P(X \cap Y \cap Z) = P(Y)P(X|Y)P(Z|X \cap Y)$

Mutual Exclusivity And Independence

- If events A and B are mutually exclusive events then A and B are independent iff (if and only if) $P(A) = 0$ or the $P(B) = 0$ or both
- However if $P(A) > 0$ and $P(B) > 0$ ➔ $P(A)*P(B) > 0$ and if event A and B are mutually exclusive ➔ $P(A \cap B) = 0$
    - So if we have mutually exclusive events A and B and $P(A \cap B) = 0 \neq P(A)*P(B)$ and A and B have non-zero probability of occurring ➔ A and B are not independent

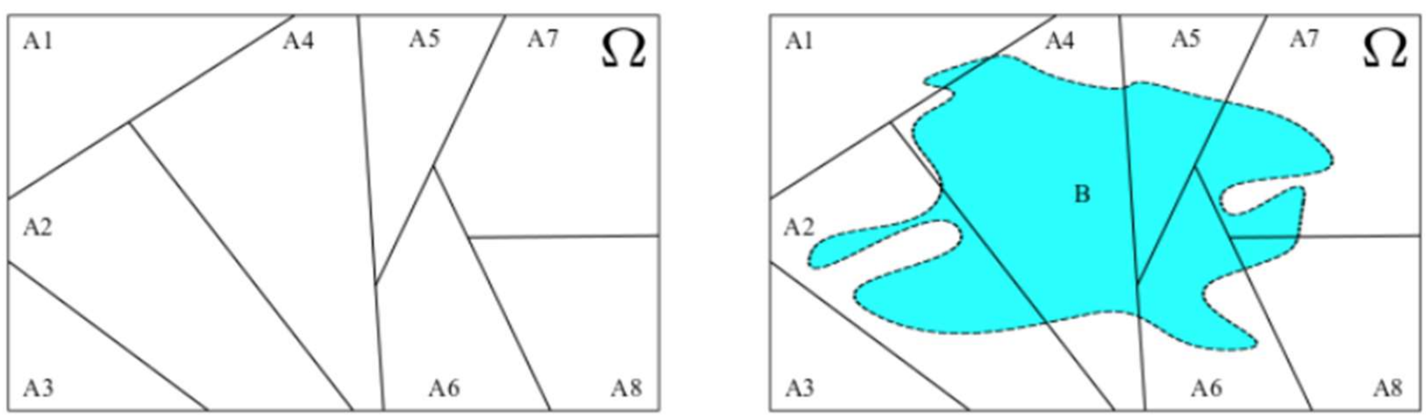
Partition complex events

we can characterize a complex event B if we understand its relationship to “partition” A

** 1.3 Partition of Ω **: Is a collection of events $A_1, A_2, \cdots, A_N$ with 3 properties

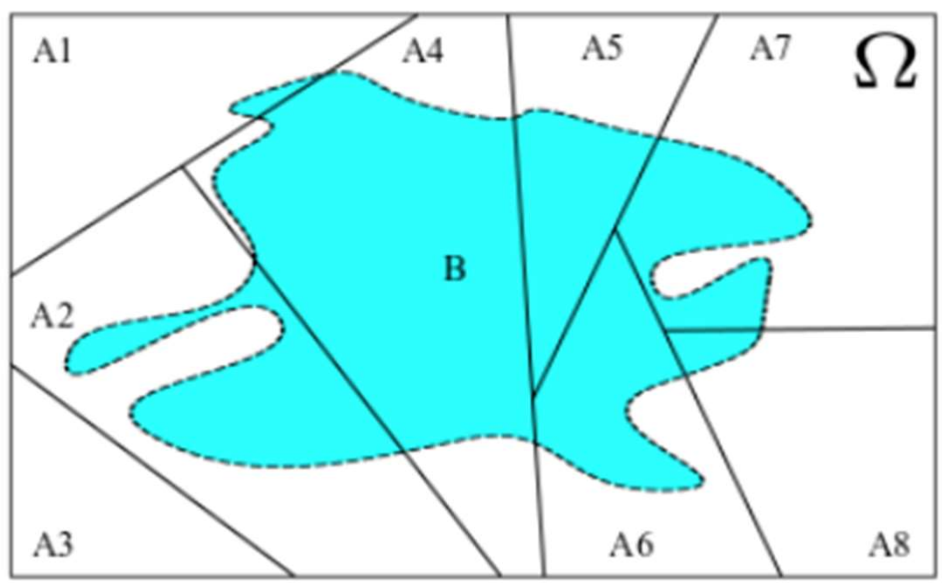
1. Relevant: $A_j \subset \Omega$ for all $j \in \{1, 2, 3, \cdots, N\}$
2. Disjoint: $A_i \cap A_j = \varnothing$ whenever $i \neq j$
3. Exhaustive: $A_1 \cup A_2 \cup \cdots \cup A_N = \Omega$

In other words, each event A_j is made up of outcomes in Ω, No two A_j's overlap/share the same outcome, together all A_j's combine to equal Ω



The law of total probability

- break down complex events



$$P(B) = \sum_{i=1}^{N} P(A_i \cap B) = \sum_{i=1}^{N} P(B|A_i)P(A_i)$$

Bayes' Rule

$$\text{Bayes' rule: } \mathrm{P}(I|E) = \frac{\mathrm{P}(E|I)\mathrm{P}(I)}{\mathrm{P}(E)}$$

Law of Complementary Probability

- Want to find P(!A) given you know P(A)
- ➔P(!A) = Prob of all events in Ω which are not in A or Ω/A
- Given we known P(Ω) = 1 and A and Ω/A are disjoint
- From Law of Add Prob we have
- P(Ω) = P(A) + P(!A) ➔ P(!A) = 1-P(A)

Sample Problem 1

- Suppose you're taking a statistics class and that in each week you are either caught up or behind on the readings. It is difficult to incorporate a new class in to your schedule so the probability that you will be caught up in Week 1, having viewed all async and completed the pre class exercises before the live session, is 0.7. From then on, if you are caught up in a given week, the probability that you will be caught up in the next week is 0.7. If you are behind in a given week, the probability that you will be caught up in the next week is 0.4. What is the probability that you are caught up in week 3?
- TIP: write down all givens and implications

Set-up Problem 1

- Let C(t) be event caught up in week t.
- Given P(C(1)) = 0.7 ➔ by law of complementary probability (LCP) P(!C(1)) = 0.3
- Also know:
  - P(C(t)|C(t-1)) = 0.7 ➔ P(!C(t)|C(t-1)) = 0.3
  - P(C(t)|!C(t-1)) = 0.4 ➔ P(!C(t)|!C(t-1)) = 0.6
  - P(C(t)) = P(C(t)|C(t-1))*P(C(t-1)) + P(C(t)|!C(t-1))*P(!C(t-1)) for t>1 by law of total probability and definition of conditional probability
- Want P(C(3))

Sample Problem 2

- A test for certain disease is assumed to be correct 95% of the time: if a person has a disease the test will give a positive result with probability 0.95. If a person does not have disease the test will give a negative result with probability 0.95. A random person drawn from a certain population has a probability 0.001 of having the disease. Given that a person drawn at random just tested positive, what is the probability that they have the disease?

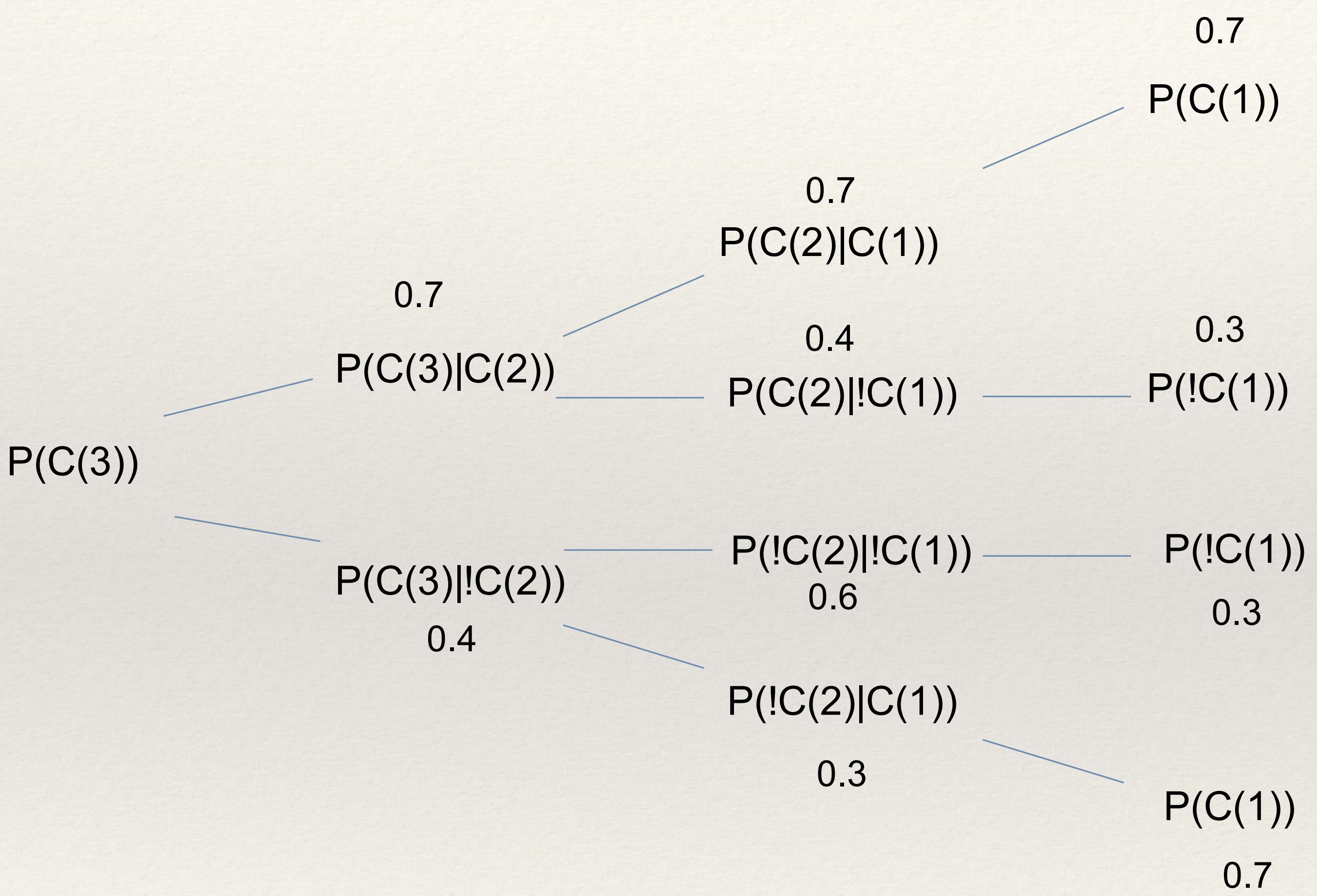
Set-up Problem 2

- Let T be event positive test result for disease D
- Let !T be event negative test result for disease D
- Givens
  - P(T|D) = 0.95 ➔ LCP P(!T|D) = 0.05
  - P(!T|!D) = 0.95 ➔ P(T|!D) =0.05
  - P(D) = 0.001 ➔ P(!D) = 0.999
- Want P(D|T) or the reverse condition
- We know from condition probability and Bayes Rule that
- P(D ∩ T)/P(T) = P(T|D)*P(D)/P(T) and
- By Law of Total Probability we can compute P(T) based upon the partition of D and !D

Algebraic Solution Problem 1

- P(C(2)) = P(C(2)|C(1))*P(C(1)) + P(C(2)|!C(1))*P(!C(1)) = 0.61
- P(!C(2)) = P(!C(2)|!C(1))*P(!C(1)) + P(!C(2)|C(1))*P(C(1)) = 0.39
- P(C(3)) = P(C(3)|C(2))*P(C(2)) + P(C(3)|!C(2))*P(!C(2)) = 0.7* 0.61 + 0.4*0.39 = 0.583

Graphical Solution Problem 1



Algebraic Solution Problem 2

- $P(T) = P(T|D)*P(D) + P(T|!D)*P(!D) = 0.95*0.001 + 0.05*0.999 = 0.0509$
- $P(D \cap T)/P(T) = P(T|D)*P(D)/P(T) = (0.95*0.001)/0.0509 = 0.0187$