Review of Lecture #2

- Experiments with a finite number of outcomes were discussed
- An event is a subset of the sample space
- An event occurs on a trial if the observed outcome on the trial belongs to the event
- On any trial, only one outcome occurs
  but many events occur
- On a trial, exactly one of A and A^c occurs

Review of Lecture #2 (continued)

- Probabilities were assigned to events using the classical equally-likely outcomes approach
- More general assignments of probabilities (when the outcomes are not equally likely) were discussed
- Disjoint unions were defined
- The axioms of probability were stated and some consequences were derived

Probability Axioms for finite spaces

- Probabilities are numbers assigned to events, subject to the following rules:
- Axiom I: $P(A) \geq 0$ for all events A
- Axiom II: $P(\Omega) = 1$
- Axiom III: If events A and B are disjoint, then P(A B) = P(A) + P(B)
- Consequences: P( ) = 0
  $P(A^c) = 1 - P(A)$; $P(A) = 1 - P(A^c)$
  $0 \leq P(A) \leq 1$ for all events A

YOU MUST REMEMBER THIS!!!

$$0 \leq P(A) \leq 1$$

for all events A

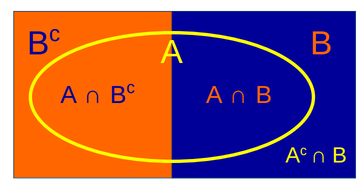
In this course, answers such as
$P(A) = -0.1$ or $P(A) = 1.57$
lead to course grades that resemble the sixth letter of the roman alphabet

Partitions

- For any given event A, Ω is the disjoint union of A and A^c : Ω = A A^c
- A and A^c are said to be a partition of Ω
- Sets {A, B, C, …, G} are a partition of set H if H is the disjoint union of A, B,…,G
- Axiom III straightforwardly generalizes to give P(H) = P(A) + P(B) + … + P(G)
- $A \cap B$ and $A \cap B^c$ are a partition of A

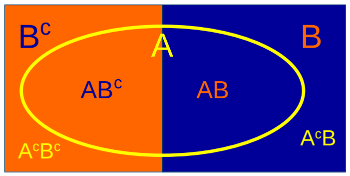
An important partition

- $A \cap B$ and $A \cap B^c$ are a partition of A
- $A \cap B$ and $A^c \cap B$ are a partition of B



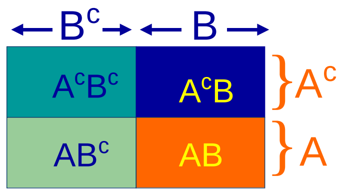
Another partition of Ω

- $A \cap B$ is often abbreviated to just AB
- AB, A^cB, AB^c, and A^cB^c are a partition of Ω



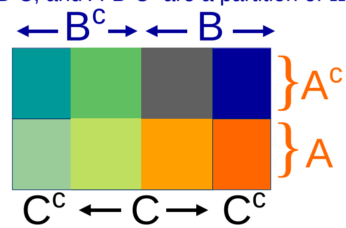
Venn diagrams vs. Karnaugh maps

- Venn diagrams are terrible for computation
- Always use Karnaugh maps!



A finer partition of Ω

- ABC, ABC^c, AB^cC, AB^cC^c, A^cBC, A^cBC^c, A^cB^cC, and $A^cB^cC^c$ are a partition of Ω



MinSets

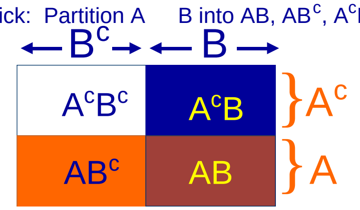
- ABC, ABC^c, AB^cC, AB^cC^c, A^cBC, A^cBC^c, A^cB^cC, and $A^cB^cC^c$ are called the minsets induced by A, B, and C
- It is possible for a minset to be empty
- Example: Ω = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
  A = {x Ω: x is a multiple of 2}
  B = {x Ω: x is a multiple of 3}
  C = {x Ω: x is a multiple of 5}
- ABC = . Are any other minsets empty?

A basic technique

- Sets {A, B, C, …, G} are a partition of set H if H is the disjoint union of A, B,…,G
- If so, then P(H) = P(A) + P(B) + … + P(G)
- A basic and very useful technique, that will be applied repeatedly in this course, is to find a clever partition A, B,…,G of an event H, and then calculate P(H) as
  P(H) = P(A) + P(B) + … + P(G)

Application of the basic technique

- What is P(A B) when A and B are not disjoint events?
- Trick: Partition A B into AB, AB^c, A^cB

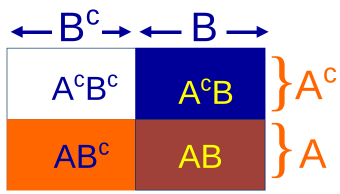


P(A B): the general case

- What is P(A B) when A and B are not disjoint events?
- Trick: Partition A B into AB, AB^c, A^cB
- P(A B) = P(AB) + P(AB^c) + P(A^cB)
  = P(AB) + P(AB^c)
  + P(A^cB) + P(AB) – P(AB)
  = P(A) + P(B) – P(AB)

P(A B): the general case

- P(A B) = P(A) + P(B) – P(AB)
- We included P(AB) twice; in P(A) and P(B)



Very useful formulas

- P(A B) = P(A) + P(B) – P(AB)
- P(A B) = P(A) + P(B) for disjoint unions
- If A and B are disjoint, P(AB) = P( ) = 0
- P(A B) = P(A) + P(A^cB)
- P(A B) = P(B) + P(B^cA)
- P(A B) = P{only one of A and B occur}
  = P(A) + P(B) – 2P(AB)

Inclusion/Exclusion Principle

- P(A B) = P(A) + P(B) – P(AB) illustrates the principle of inclusion/exclusion
- P(A B C) = P(A) + P(B) + P(C)
  – P(AB) – P(BC) – P(AC)
  + P(ABC)
- Include the probability of the events, exclude the probabilities of the pairwise intersections, include the probabilities of the triplewise intersections, …

Inclusion/Exclusion Principle

- The formula for the probability of the union of many events is given in the textbook (Ross, p. 35, Proposition 4.4)
- The inclusion/exclusion principle is so general that it hardly ever can be used to solve any particular problem
  — you need to know the probabilities of far too many events

Example of use

- Example: A and B are events defined on a sample space Ω
  If P(A) = 0.3, P(B) = 0.4, and P(A^cB^c) = 0.5, what is P(AB)?
- First step is to

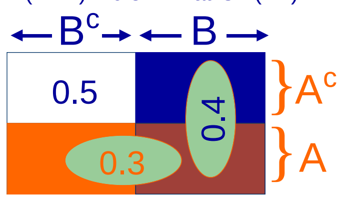
PANIC!

Method of attack

- Example: A and B are events defined on a sample space Ω
  If P(A) = 0.3, P(B) = 0.4, and $P(A^cB^c) = 0.5$, what is P(AB)?
- First step is to draw a Karnaugh map and mark the given information on it
- Then use the various results described to obtain the answer

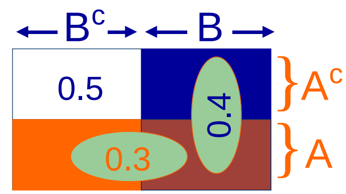
How to do it

- Example: A and B are events defined on a sample space Ω. P(A) = 0.3, P(B) = 0.4, and $P(A^cB^c) = 0.5$. What is P(AB)?



Completion of example

- $P(A^cB^c) = 0.5$ P(A B) = 0.5
- P(AB) = P(A) + P(B) – P(A B) = 0.2



Critique of solution of the problem

- What exactly was the sample space in the example just worked?
- How many outcomes belonged to A? to B? and what were their probabilities?
- The solution methodology that we used does not require the underlying details
- The same solution would be obtained for different sample spaces (as long as the assigned probabilities were the same)

DeMorgan's Laws

- DeMorgan's Laws are useful in simplifying the calculations in many problems
- $(A \quad B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \quad B^c$
- More generally, the complement of the union (intersection) of countably many sets is the intersection (union) of the complements of the sets

Equally likely outcomes

- Vast literature on sample spaces with equally likely outcomes
- Many of the problems basically are problems in counting
- Combinatorial problems
- Sometimes, interpreting the statement of the problem and deciphering what is being asked is the hardest part

Equally likely outcomes

- Example: An urn contains 6 identical red and 3 identical green balls. A trial of the experiment consists of drawing a ball at random from the urn
- Is the sample space Ω = {R, G}?
- Or is the sample space Ω = {R_1, R_2, R_3, R_4, R_5, R_6, G_1, G_2, G_3} where we assume the balls are numbered and thus distinguishable?

Identical things are distinguishable!

- We always assume that objects said to be identical actually are distinguishable to the very perceptive observer
- Physically, the assumption is that two "identical" objects differ in minute ways that do not affect the probabilistic behavior
- Example: Two absolutely identical coins are tossed. The sample space is {(T,T), (T,H), (H,T), (H,H)}

P{red ball is drawn}

- Example: An urn contains 6 identical red and 3 identical green balls. A trial of the experiment consists of drawing a ball at random from the urn
- The phrase at random means that the outcomes are be equally likely
- Ω = {R_1, R_2, R_3, R_5, R_5, R_6, G_1, G_2, G_3}
- P{red ball is drawn} = 6/9 = 2/3

Drawing a random sample

- Example: An urn contains 6 identical red and 3 identical green balls. A trial of the experiment consists of simultaneously drawing two balls at random from the urn
- Ω = {R_1, R_2, R_3, R_4, R_5, R_6, G_1, G_2, G_3}?
- No. The outcomes of this experiment are of the form {R_2, R_5} or {R_3, G_2} or {G_1, G_3}
- The outcomes of this experiment are the subsets of size 2 from the set of 9 balls

More on drawing a random sample

- The subset drawn is called a random sample of size 2 from the urn
- Random means that all subsets of size 2 are equally likely to be chosen
- But, how many such subsets are there?
- 36 subsets of size 2 from a set of size 9
- More generally, given a set of size n, how many subsets of size k are there?

A combinatorial digression

- The number of subsets of size k from a set of size n is denoted by

$$\binom{n}{k}$$

  Read this as n-choose-k. It is also called a binomial coefficient
- Many calculators have built-in programs

More combinatorial digression

- Some obvious values are
  $$\binom{n}{0}=1;\quad \binom{n}{1}=n;\quad \binom{n}{n}=1$$
  For a sample space of size n, these values correspond to the empty set (there is only one!), the elementary events (there are n of these) and the entire sample space
  $$\binom{n}{k}=\binom{n}{n-k}$$

Combinatorial digression continues

- A subset of size k–1 can be enlarged into a subset of size k by including one of the elements not already in the subset
- {a,b,c}, a subset of size k–1 = 3 from the n = 6 elements {a,b,c,d,e,f} can be enlarged to {a,b,c,d} or to {a,b,c,e} or to {a,b,c,f}
- A subset of size k–1 can be enlarged into n–(k–1) different subsets of size k

Combinatorial digression continues

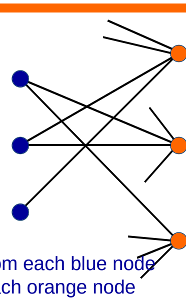
- However, {a,b,c,d} could have resulted from enlarging the set {a,b,c} or the set {a,b,d} or the set {a,c,d} or the set {b,c,d}
- Removing an element from a subset of size k reduces it to a subset of size k–1
- A subset of size k can be reduced to k different subsets of size k–1 by removing one of the elements

Combinatorial digression continues

- Consider the graph shown
- Each blue node is a subset of size k–1
- Each orange node is a subset of size k
- A line joins a blue node to an orange node if the subset of size k–1 denoted by the blue node can be enlarged to the subset of size k denoted by the orange node
- We now count the number of lines

Graph with blue and orange nodes

Blue nodes are subsets of size k–1



Orange nodes are subsets of size k

n–k+1 lines from each blue node
k lines from each orange node

Combinatorial digression continues

- Each blue node (subset of size k–1) has n–(k–1) lines coming out of it because each can be enlarged in n–(k–1) different ways
- Each orange node (subset of size k) has k lines coming out of it because each can be reduced in k different ways
- Parts is parts
- Lines is lines

Putting it all together

$$k\binom{n}{k} = (n-k+1)\binom{n}{k-1}$$

$$\binom{n}{k} = \frac{n-k+1}{k}\binom{n}{k-1}$$

$$= \frac{n-k+1}{k} \times \frac{n-k+2}{k-1}\binom{n}{k-2}$$

$$= \frac{n-k+1}{k} \times \frac{n-k+2}{k-1} \times \cdots \times \frac{n}{1}$$

What does n-choose-k count?

- n-choose-k is the number of subsets of size k from a set of size n
- n-choose-k is the number of events, defined on a sample space with n outcomes, that contain exactly k outcomes
- n-choose-k is the number of binary vectors, of length n that have Hamming weight k, that is, have k ones and n–k zeroes

The example continued (finally!)

- There are 9-choose-2 = 36 subsets of size 2 from a set of 9 balls, and the collection of these subsets is the sample space
- There are 6-choose-2 = 15 outcomes that contain two red balls
- There are 3-choose-2 = 3 outcomes that contain two green balls
- There are 6×3 = 18 outcomes that contain one red ball and one green ball

Summary

- We explored the consequences of the axioms of probability and their use in solving problems
- We studied binomial coefficients
- We looked at some simple combinatorial problems
- We will next consider relative frequencies and sample spaces with infinitely many outcomes