COMP 233/2
Probability and Statistics for Computer Science
Week 1

Contributions by: T. Fevens, H. Mellah,
S. Mudur, and NETA powerpoint from Nelson

Introduction
Axioms of probability
Sample spaces having equally likely outcomes

Reading: Chap 1, 4

Course Description

This course introduces students to the fundamentals of probability theory and the basics of statistical analysis.

For a detailed course description and the list of learning objectives/topics see the Course Outline on the course webpage in Moodle.

Instructor: Hakim Mellah (H961-15)

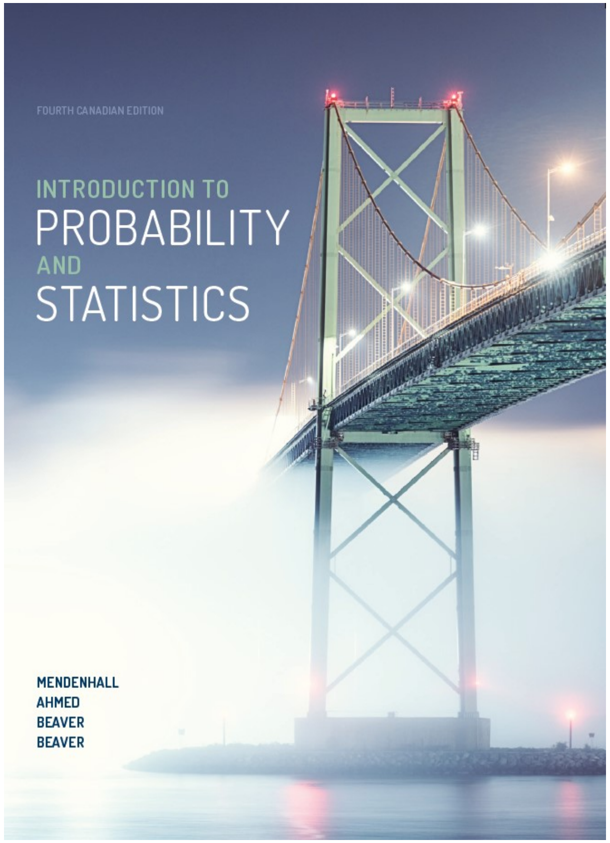
Email: hmellah@encs.concordia.ca

Office Hours: Wednesdays, 15h00-16h00;
or by appointment

Course Information

- See Course Outline.
- Specific Comments:
  - Textbook
  - i>Clicker participation
  - Number of assignments (4)
  - Assignment submission through Moodle (keep electronic copies of your submission)
  - Mark breakdown

Class Participation Points	10%
Assignments	20%
Midterm (in class)	20%
Comprehensive final exam	50%



Course Structure

1. Fundamentals of Probability Theory: Conditional Probability and Independent Events. Random Variables. (6 Weeks – Chapters 4,5,6,10,11,14)
2. Basic Notions of Statistics. Descriptive Statistics: Organizing, Presenting and Summarizing Data. Central Limit Theorem. (2 Weeks – Chapters 2,3)
3. Inferential Statistics: Parameter Estimation, Hypothesis Testing, Regression. (3 Weeks – Chapters 7,8,9,12,13)
4. Random Number Generation. (1 Week – Chapter 6)

How to do well in this course

- Do the assignments!!!
- Practice using WebAssign.
- Come to the classes. Answer questions (clicker). Participate in the follow-up discussion.
- Attend the tutorials. Many more examples (similar to the questions in the midterm and the final) will be done here.
- Write-hand rule: write out by hand each example done in class/tutorial.

Questions We Seek Answers To: Example 1

On average, one request in 1000 causes a server to crash.

What are the chances that one server can process 2000 requests or more without crashing?

Questions We Seek Answers To: Example 2

Suppose that, on average, 3600 messages are received within one hour.

1. What are the chances that one message per second is received?
2. What are the chances that the time between two consecutive messages is less than one second?

Questions We Seek Answers To: Example 3

Ten (10) measurements of the current at a circuit junction resulted in an average value of 123mA. You expected, however, a value not greater than 120mA.

1. So… was your assumption wrong? Could it be correct?
2. How would the answer be affected, if twenty (20) measurements were carried out?
3. What other factors are there to consider?

Introduction to Probability Theory

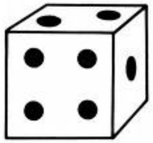
Mainly theory. Real data comes with statistics later in this course.

Probability Experiments

A *probability (random) experiment* is an experiment that can result in different *outcomes*, even though it is repeated in the same manner every time.

An experiment:

- Flipping a coin once.


- Rolling a die once.


- Pulling a card from a deck.


Further Examples

- More relevant (to us CSE/ECE) examples are
  - An experiment:
    - Transmitting bits until the first error.
    - Measuring the time between two consecutive requests to a server.
    - Measuring the thickness of a wafer used in semiconductor manufacturing.
    - Counting the number of bugs per module.

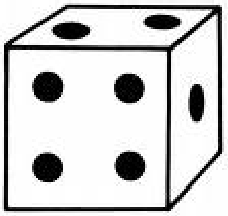
Sample Spaces

- The set of all possible *outcomes* of an experiment is called a ***sample space***.
- You may think of a sample space as the set of all values that a variable (e.g., X) representing the outcome may assume.
- We are going to denote the sample space by S.

Examples

- Experiment: Tossing a coin once.
  - $S = \{H, T\}$

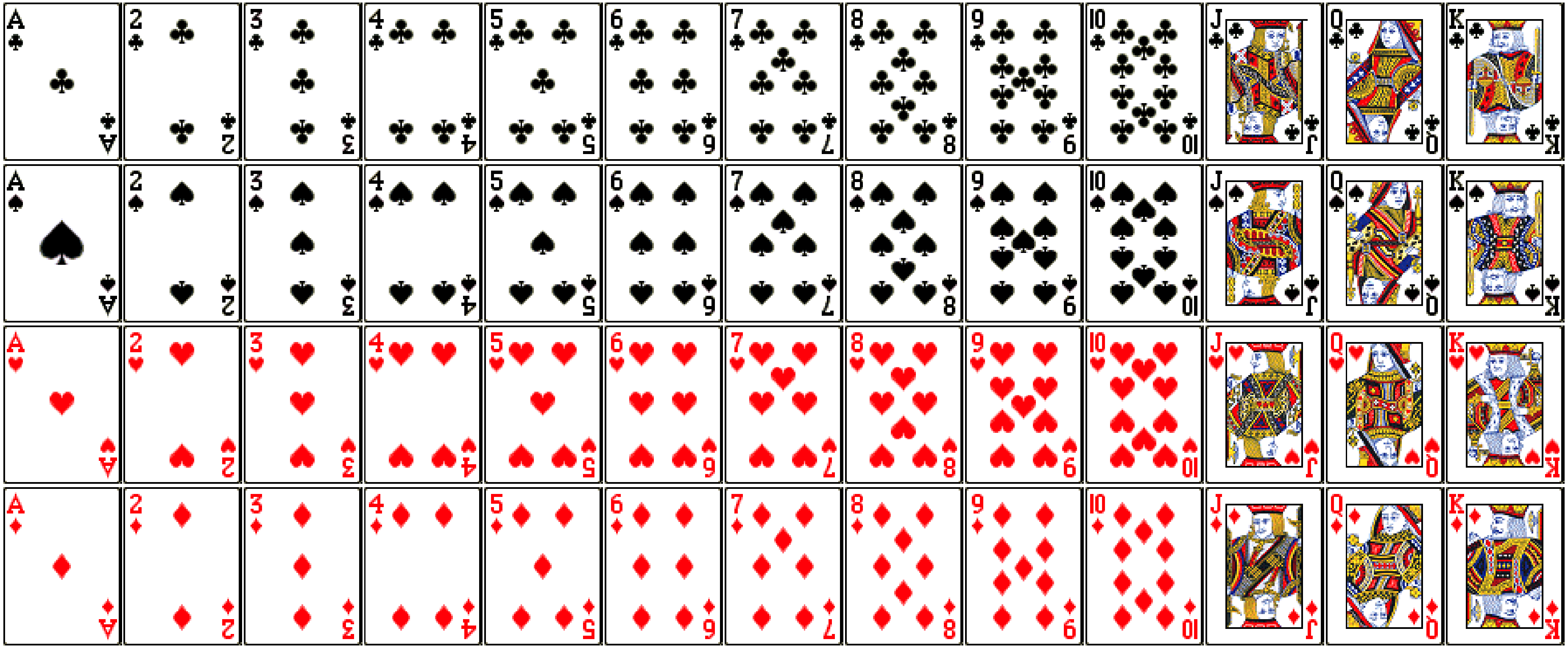
- Experiment: Rolling a cubic die once.
  - $S = \{1, 2, 3, 4, 5, 6\}$



Examples

- Experiment: Drawing a card from 52 card deck.

S =



Examples

- If the experiment consists of the running of a race among the seven horses having post positions 1, 2, 3, 4, 5, 6, 7, then

  S = {all orderings of (1, 2, 3, 4, 5, 6, 7)}

  The outcome (2, 3, 1, 6, 5, 4, 7) means, for instance, that the number 2 horse is first, followed by the number 3 horse, then the number 1 horse, and so on.
- What is the size of S? Large?

Examples

- Experiment: Checking a bit.
  - $S = \{o, e\}$

    (o means the bit is OK, and e means an error)
- Experiment: Transmitting until the first error.
  - $S = \{e, oe, ooe, oooe, ooooe, \ldots\}$
- Experiment: Measuring the time.
  - $S = [0, \infty]$.

Classification of Sample Spaces

- We distinguish between *discrete* and *continuous* sample spaces.
- The outcomes in *discrete* sample spaces can be counted. The number of outcomes can be finite or infinite.
- In the case of *continuous* sample spaces, the outcomes fill an entire region in the space (e.g., an interval of real numbers).

Events

- An *event*, E, is a set of outcomes of a probability experiment, i.e. a subset in the sample space.
  - $E \subseteq S$.
- If an event E contains no outcomes from S, then E is an *impossible* event.
- The *union* and the *intersection* of events are defined in a natural way (set operations), giving possibly new events.

Examples

- In the tossing-a-coin-once experiment, if $E = \{H\}$, then E is the event that the coin comes up heads. Similarly, if $F = \{T\}$, then F is the event that it comes up tails.
- In the race among 7 horses experiment, if $E = \{$all outcomes in S starting with a 3$\}$, then E is the event that the number 3 horse wins the race.

Mutually Exclusive Events

- Two events, *E* and *F*, are called ***mutually exclusive***, if

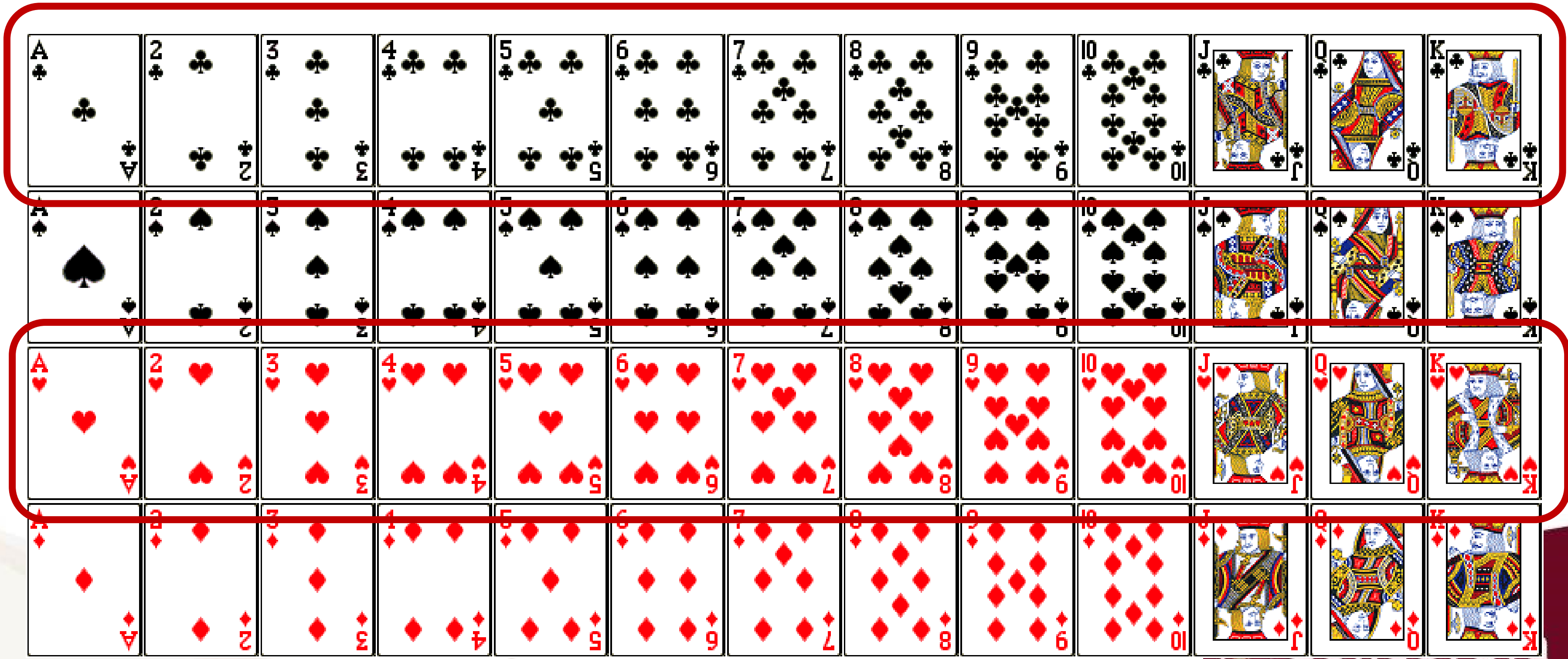
$$EF = \phi$$

  (EF is the intersection of sets E and F, and ϕ denotes an empty set)
- That is, it is impossible for an outcome to be an occurrence of both events.
  - E.g., for the cast die experiment, events E = {2,4,6} and F = {1,5} are mutually exclusive.

Examples

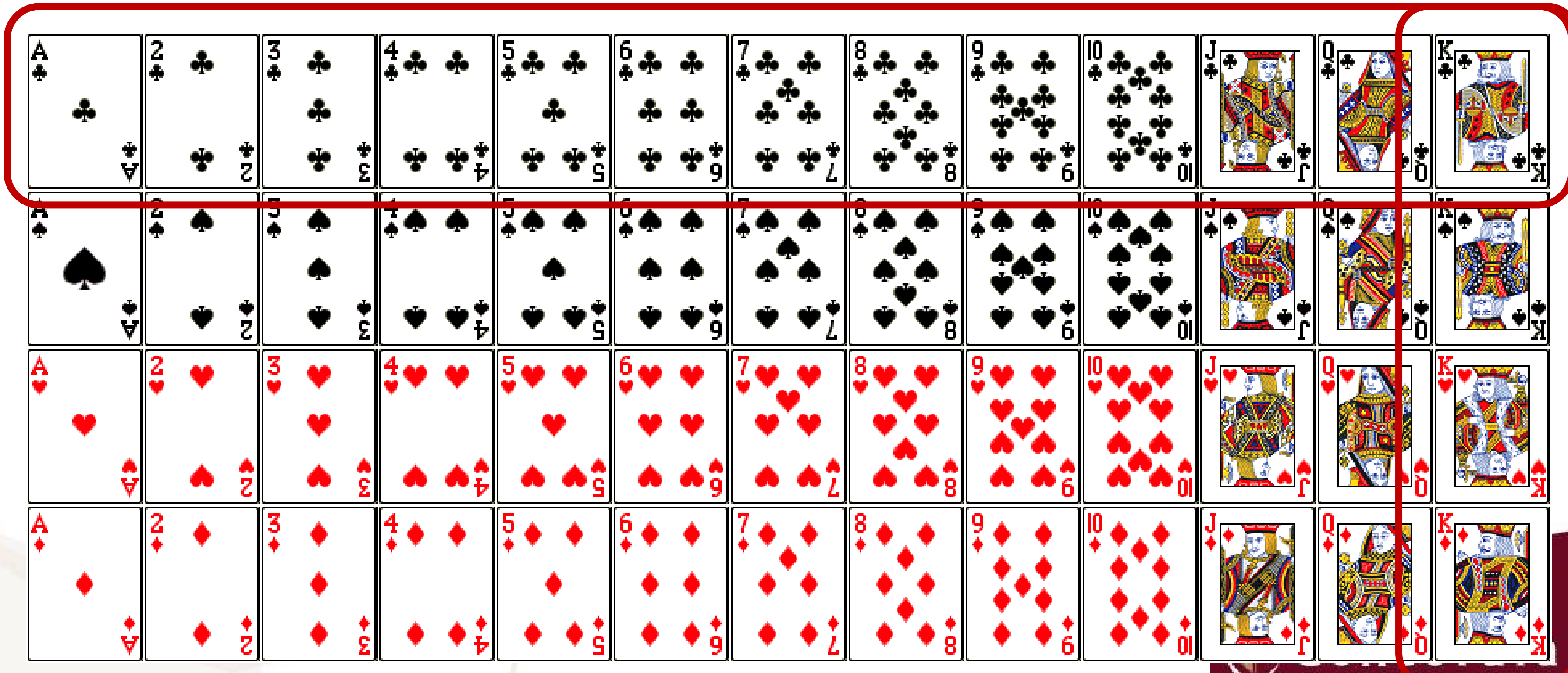
When drawing a card from a deck:

- Drawing a club (E) and drawing a heart (F) are mutually exclusive events ($EF=\phi$).



Examples

- Drawing a club (C) and drawing a king (K) are not mutually exclusive ($CK \neq \phi$).
  - Drawing the king of clubs is an outcome of both events.



Example

- Experiment: Rolling two dice.
  - $S = \{ (1,1), (1,2), \ldots , (1,6),$
    $(2,1), (2,2), \ldots , (2,6),$
    ……………………………
    $(6,1), (6,2), \ldots , (6,6) \}$

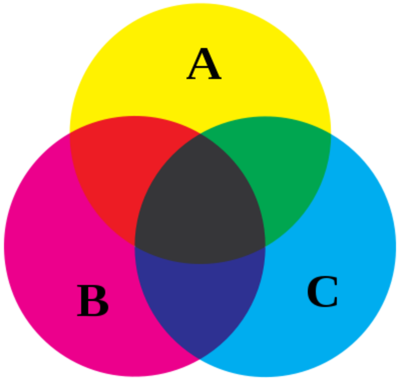
Example

- E is the event of getting two numbers with different parities (i.e., even/odd).
- F is the event of getting two numbers whose sum is 6.
- E = {(1,2), (1,4), (1,6),
  (2,1), (2,3), (2,5),
  ...,
  (6,1), (6,3), (6,5)}
- F = {(1,5), (2,4), (3,3), (4,2), (5,1)}
- Thus, E and F are mutually exclusive.
  - For the sum of two numbers to be even, they have to have the same parity.

Complementary Events

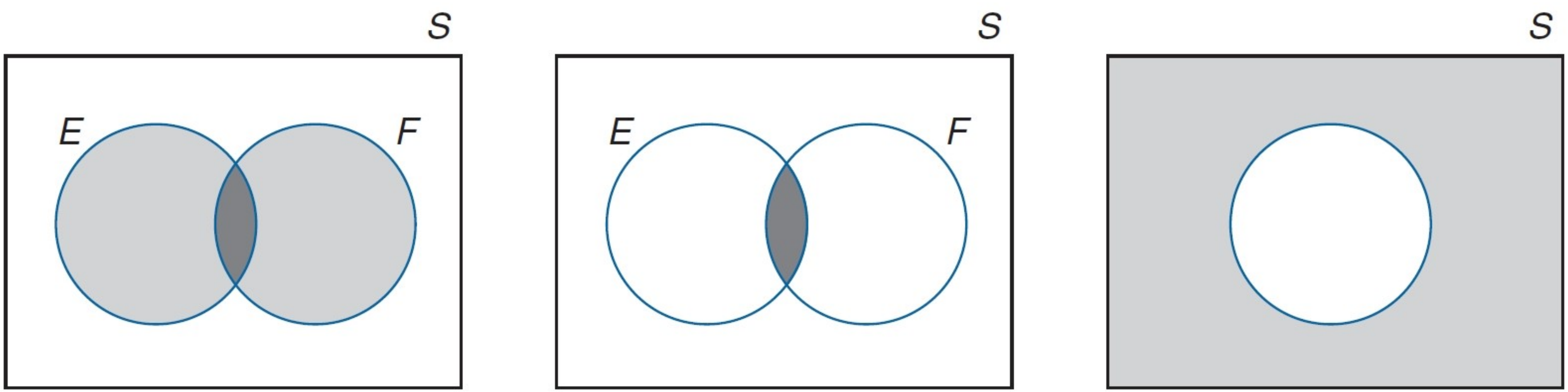
- The *complement* of an event E is defined by E^c
  $$E^c = S \backslash \mathrm{E}$$
  where "\" denotes set difference. That is, the complement contains all outcomes that are not in E.
- It is clear that E and E^c are *mutually exclusive.*
- E.g., for the cast die experiment,
  if event E = {2,4,6}

  then E^c = {1,2,3,4,5,6}\{2,4,6} = {1,3,5}.



Venn Diagrams

Sometimes it is convenient to represent the sample space by a rectangular region, with events being circles within the region. Such a representation is called a ***Venn diagram***.



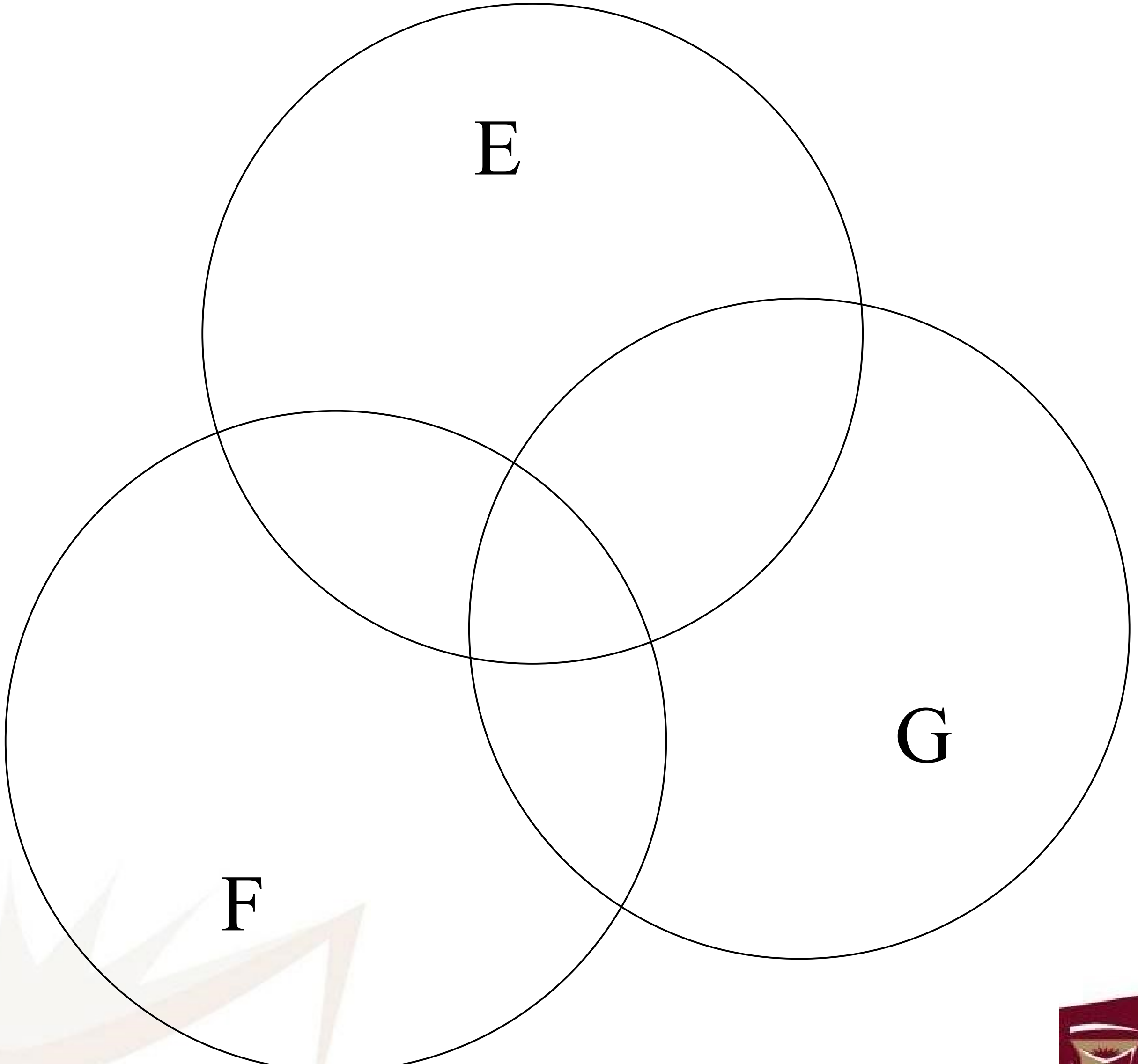
(a) Shaded region: $E \cup F$ (b) Shaded region: EF (c) Shaded region: E^C

Set Operations

- The operations of forming unions, intersections, and complements of events obey certain rules not dissimilar to the rules of algebra. We list a few of these.

Commutative law	$E \cup F = F \cup E$	$EF = FE$
Associative law	$(E \cup F) \cup G = E \cup (F \cup G)$	$(EF)G = E(FG)$
Distributive law	$(E \cup F)G = EG \cup FG$	$EF \cup G = (E \cup G)(F \cup G)$

- These operations are easily shown using Venn diagrams.


E
G
F

De Morgan Laws

- Using the Venn diagrams one can also easily show that

$$(1^{st}\ \text{Law})(E \cup F)^c = E^c F^c$$

- The 2nd Law can be obtained from the first one by first replacing E and F by E^c and F^c, respectively, on both sides of the equation, and then applying the complement to both sides and simplifying:

$$(2^{nd}\ \text{Law})(EF)^c = E^c \cup F^c$$

Definition of Probability

- The probability of an event *E* is a number, *P(E)*, such that

$$0 \le P(E) \le 1$$

$$P(S) = 1$$

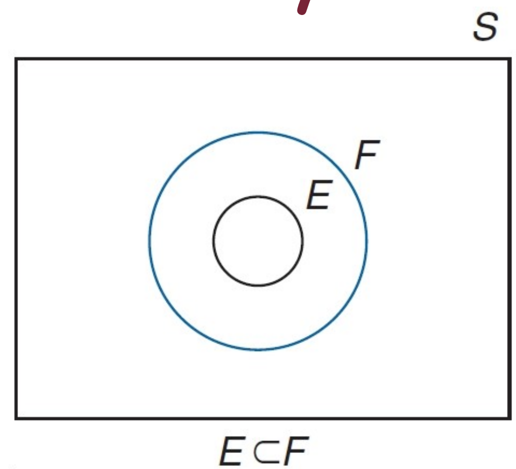
$$P(E \bigcup F) = P(E) + P(F),$$

if the events *E* and *F* are *mutually exclusive.*

- The latter implies that, in general,

$$P(E \bigcup F) = P(E) + P(F) - P(EF)$$

Further Properties of Probability



- Probability of sub-event:

$$\because E \subseteq F$$

$$\therefore P(E) \leq P(F)$$

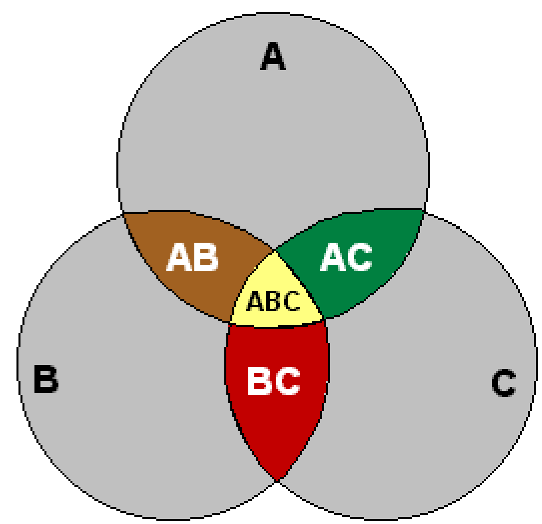
- Probability of the union of three events:

$$P(A \bigcup B \bigcup C)$$

$$= P(A) + P(B) + P(C)$$

$$- P(AB) - P(BC) - P(AC)$$

$$+ P(ABC)$$



Further Properties of Probability

- *Odds* of an event A:

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1-P(A)}$$

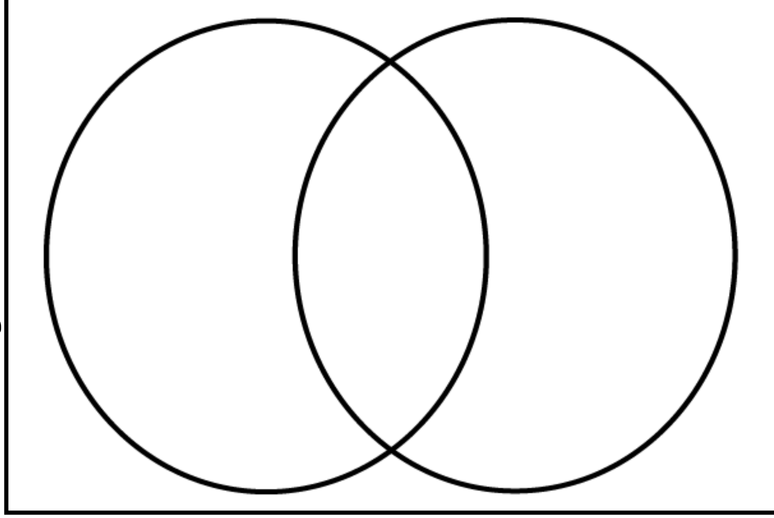
  The odds tell us how much more likely it is for A to occur than that it does not occur.

- E.g., P(A) = 3/4 means that

$$\frac{P(A)}{1-P(A)} = 3$$

  or that A is three times more likely to occur than not.

Example



A total of 37 percent of Computer Science students play computer games daily, 73 percent check their email daily, and 24 percent do both daily. What percentage of Computer Science students neither play computer games daily nor check their email daily?

Classical Probability

- The (*classical*) probability, $P(E)$, of an event E is given by

$$P(E) = \frac{N(E)}{N(S)}$$

where N() stands for the number of elements (outcomes) in a set.

- It is assumed here that the sample space is finite and all outcomes are <u>equally likely</u> to occur.

Example, revisited

- Experiment: Rolling two dice.
  - $S = \{ (1,1), (1,2), \ldots, (1,6),$
    $(2,1), (2,2), \ldots, (2,6),$
    ………………………………
    $(6,1), (6,2), \ldots, (6,6) \}$

N(S) =

Example, revisited

- E is the event of getting two numbers with different parities
- E = {(1,2), (1,4), (1,6),
  (2,1), (2,3), (2,5),
  …,
  (6,1), (6,3), (6,5)}

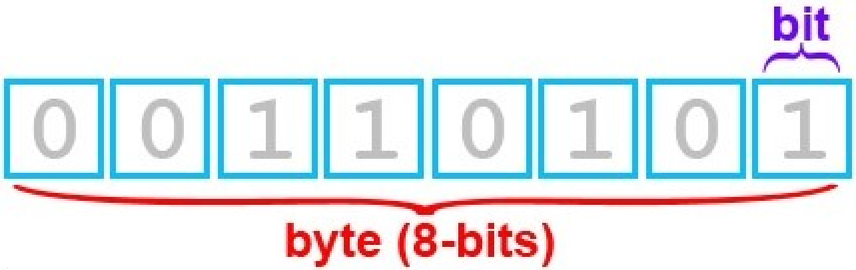
P(E) =

- F is the event of getting two numbers whose sum is 6.
- F = {(1,5), (2,4), (3,3), (4,2), (5,1)}

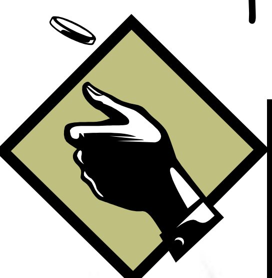
P(F) =

Computing Probabilities

- Find the probability that a 1-byte long message contains exactly 4 zeros.



- If a coin is flipped 8 times, find the probability that heads appear exactly four times.



1	2	3	4	5	6	7	8
0	1	1	0	1	0	0	1

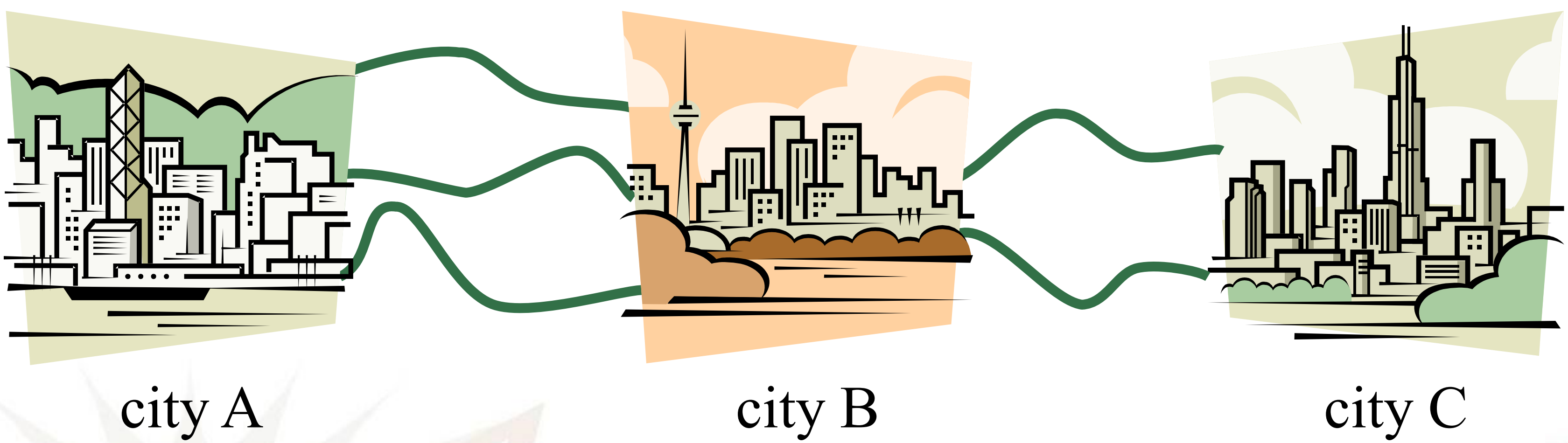
Computing Probabilities

- It is clear that an efficient way to count outcomes in large sample spaces is required.
- So next topic is

COUNTING RULES

Example

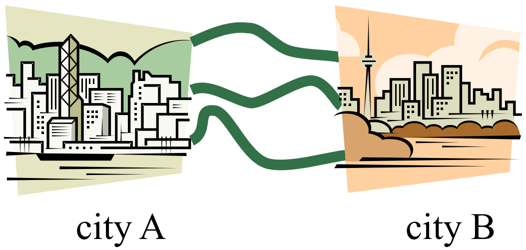
A salesperson needs to get from city A to city B and then to city C. Three roads lead from A to B, and two lead from B to C. How many different routes can the salesperson take to accomplish their task?



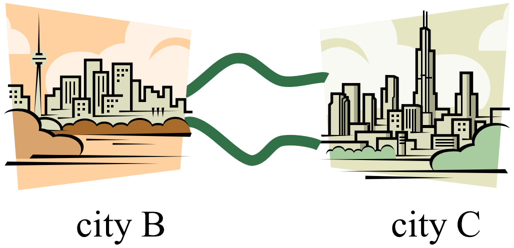
Solution

The event of getting from A to C is a sequence of two events.

- The first (from A to B) can occur in 3 different ways.



- The second (from B to C) can occur in 2 different ways.



- Thus, getting from A to C can occur in

$$3 \times 2 = 6 \text{ different ways}$$

Multiplication Rule

- In a sequence of two experiments, if the first experiment can occur in m different ways and the second one can occur in n different ways, then the sequence of outcomes can occur in mn different ways.

- The rule is sometimes called the

  Basic Principle of Counting.

Multiplication Rule, Generalized

Generalized Basic Principle of Counting:

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if, …, **then** there are a total of $n_1 n_2 \cdots n_r$ possible outcomes of the r experiments.

Example

Consider the manufacturing of license plates consisting of two letters followed by four digits.

a. How many plates are possible?

b. How many plates are possible, if no letter or digit can be repeated?

c. If repetitions are allowed, on how many plates are the two letters vowels and all digits even?

Solution

a. 26 letters and 10 digits

$26 \times 26 \times 10 \times 10 \times 10 \times 10 = 6{,}760{,}000.$

b. Same, no letter or digit repeated

$26 \times 25 \times 10 \times 9 \times 8 \times 7 = 3{,}276{,}000.$

c. 5 vowels and 5 even digits

$5 \times 5 \times 5 \times 5 \times 5 \times 5 = 15{,}625.$

Example

- In a class of ten (10) students, six (6) are to be chosen and seated in a row for a picture. How many different pictures are possible? If the same students are seated in a different order, we consider the picture to be different.

Solution

- Two pictures are different in two respects:
  - which six students are in the picture, and
  - how they are linearly arranged.
- Thus, there are

$$10 \times 9 \times 8 \times 7 \times 6 \times 5$$

different pictures.

A Note on Factorials

- For an integer $n \geq 0$, **n factorial** (denoted by $n!$) is defined by

$$0! = 1$$

$$\begin{aligned} n! &= n(n-1)! \\ &= n(n-1)(n-2)\cdots(2)(1), \text{ for } n > 0 \end{aligned}$$

A Note on Factorials

- In the last example, note that

$$
\begin{aligned}
& 10\times 9\times 8\times 7\times 6\times 5 \\
&= 10\times 9\times 8\times 7\times 6\times 5\times \frac{4\times 3\times 2\times 1}{4\times 3\times 2\times 1} \\
&= \frac{10!}{4!}
\end{aligned}
$$

Permutations

- The number of repetition-free **permutations** (linear arrangements) **of size *r* from a set of *n* distinct objects** is given by

$$^{n}P_{r} = \frac{n!}{(n-r)!}; \qquad 0 \le r \le n$$

Note

- If $r = n$ $\qquad {}^{n}P_{n} = \dfrac{n!}{(n-n)!} = n!$

- If $r = 0$ $\qquad {}^{n}P_{0} = \dfrac{n!}{n!} = 1$

- If $r > n$, then the rule does not work.
  - permutations of size > n would have to involve *repetitions*.

Example

- What is the number of (possibly meaningless) words that are made up of <u>all</u> the letters in the word "computing"?

Solution

- Two words are distinct in two respects
  - which symbols they are made up of, and
  - how these symbols are arranged.
- Words made up of the letters in "computing" are all possible permutations of the letters.
- Thus, the number of such words is (n = 9; r = 9):

$$^{n}P_{r} = {}^{9}P_{9} = \frac{9!}{(9-9)!} = \frac{9!}{0!} = 9!$$

Example

- What is the number of permutations of the letters in the word "ball"?
  - Note that the answer is not 4!.
  - We do not have 4 distinct objects.
  - We have a set of only 3 distinct objects.
  - But note that the answer is not 3!.
    - The permutations here involve repetitions (r (=4) > n (=3)).

Solution

- Suppose that there are no repetitions:

  ball

- In this case, there are 4! permutations.
- But for each of these permutations, there is exactly one other permutation where l and l switch positions.
- These are, really, the same permutations.
- Thus, the number of permutations of the letters in "ball" is

$$4!/2 = 4!/2! = 12$$

Now, consider this

- What about the word “pepper”?

Food for thought

- What about the word "pepper"?

$$\frac{6!}{3! \cdot 2!}$$

- Why?

Example

In a class of 10 students, three are to be chosen to represent the class in a contest.

How many selections are possible?

Solution

- Note that, here, the students are not arranged in any order.
  - As if we only care about who is in the group photo, and not how they are seated.
- The number of selections is

$$\frac{{}^{10}P_3}{3!} = \frac{10!}{7! \cdot 3!}$$

- The denominator accounts for the permutations of the 3 selected students.

Combinations

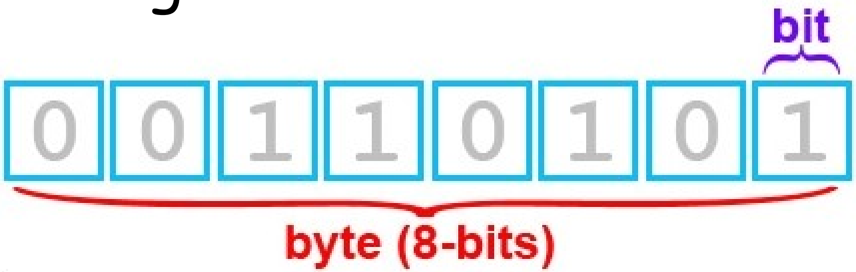
- In general, the number of repetition-free ***combinations*** (selections) ***of size r from a set of n distinct objects*** is given by

$$\binom{n}{r} = \frac{{}^{n}P_{r}}{r!} = \frac{n!}{(n-r)!r!}; \quad 0 \le r \le n$$

- Note that $\binom{n}{0} = \binom{n}{n} = 1$

Example

- A fair coin is tossed 8 times, what is the probability of getting
  - exactly 4 heads.
  - at least 4 heads.
- In a 1-byte long message, what is the probability of having
  - exactly 4 zeros.
  - at least 4 zeros.



Solution

- By the multiplication rule,

$$N(S) = 2\times2\times\ldots\times2 = 2^8 = 256.$$

- For the event of 4 HEADS, we need to select 4 tosses out of 8. The other 4 will be TAILS. Hence,

$$N(E) = \binom{8}{4} = 70.$$

1st	2nd	3rd	4th	5th	6th	7th	8th
0	1	1	0	1	0	0	1

Computing Probabilities

- Find the probability that a 1-byte long message contains exactly 4 zeros.
- If a coin is flipped 8 times, find the probability that heads appear exactly four times.

$$P(E) = \frac{N(E)}{N(S)} = \frac{\binom{8}{4}}{2^8} = \frac{8!}{4! \cdot 4!} \cdot \frac{1}{2^8} = \frac{70}{256} \approx 0.273$$

Computing Probabilities II

- If a coin is flipped 8 times, find the probability that heads appear at least four times.

Note on Number of Combinations

- The numbers $\binom{n}{r}$ (or ${}^{n}C_{r}$) have a special name in algebra – the ***binomial coefficients.***
- They are the coefficients of the expansion $(a+b)^n$:

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$$

Pascal triangle

- The *binomial coefficients* can be conveniently visualized by means of the so-called ***Pascal Triangle***.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

$\binom{2}{0}$

$\binom{5}{2}$

Blaise Pascal
(1623 - 1662)

Pascal triangle

- Here the *Pascal Triangle* with 17 rows depicted.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1
1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1
1 15 105 455 1365 3003 5005 6435 6435 5005 3003 1365 455 105 15 1
1 16 120 560 1820 4368 8008 11440 12870 11440 8008 4368 1820 560 120 16 1

Food for thought

- Show that

$$2^n = \sum_{r=0}^{n} \binom{n}{r}$$

- Can you give a probabilistic reasoning that proves the above property?

Food for thought

- Show that $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$

Food for thought

- Ten students (A, B, . . ., J) are to be divided into five teams of two to work on a term project.
  a) Find the number of possible divisions into teams.
  b) Find the probability that A and B form a team.

Solution to “10 students divided into 5 teams”

a) Here the sample space is the set of all possible pairings into teams.
   - Each element in the sample space is a set of five pairs.
   - $N(S) = (10)!/((2!)(2!)(2!)(2!)(2!)(5!))$
     $= (10)!/((2^5)(5!))$

b) E is the set of pairings in which A and B form a pair.
   - $N(E) = (8!)/((2^4)(4!))$
   - Hence, get $P(E) = N(E)/N(S)$.

Food for thought

- A student is taking a Data Structures test in which 7 questions out of 10 must be answered. In how many ways can the student answer the exam if
  1. Any 7 questions may be selected.
  2. The first two questions must be selected.
  3. The student must choose three questions from the first five and four questions from the last five.

“7 questions out of 10” Solution – 1).

Any 7 questions may be chosen.

- Thus, the student needs to select 7 objects out of a set of 10 distinct objects.
- Hence, the student can answer the exam in $\binom{10}{7}$ different ways.
- So the answer is 10!/(3!7!) = 120.

Solution - 2).

The first two questions must be chosen.

- We know that 2 of the selected 7 are the first two.
- Thus, different selections correspond to different selections of the remaining 5 out of only 8 questions.
- Hence, the student can answer the exam in $\binom{8}{5}$ = 8!/(3!5!) = 84 different ways.

Solution – 3).

3 selected from the first 5 and 4 selected from the last 5.

- The selection is actually a sequence of two selections: 3 from 5 and 4 from 5.
- I.e., $\binom{5}{3}$ possible selections followed by

  $\binom{5}{4}$ possible selections
- Thus, by the rule of products, the student can answer the exam in $10\times5 = 50$ different ways.

More on Probability – next class

- Classical probability works in a lot of simple situations, but frequently if an event has occurred, it will change the probability of another event occurring.
- Next week we are going to expand the types of probabilities we consider to include conditional probability.
- So next class' topic is

CONDITIONAL PROBABILITY

References/Resources Used

- Lecture Slides for MATH 401 of Dr. Oleksiy Us, Department of Mathematics, German University of Cairo. [PPT]