Discrete Probability

- Counting
  - Permutations – Combinations
  - r- Combinations
  - r- Combinations with repetition Allowed
  - Pascal’s Formula
  - Binomial Theorem
  - Conditional Probability
  - Baye’s Formula
  - Independent Events

Reading (Epp’s textbook)
9.5-9.9

Permutations

A **permutation** of a set of distinct objects is an ordered arrangement of these objects

- ✓ No object can be selected more than once.
- ✓ Order of arrangement matters.

How many Permutations?

In general, given a set of *n* objects, how many permutations does the set have?

Imagine forming a permutation as an *n*-step operation:

Step 1: Choose an element to write first. (n choices)

Step 2: Choose an element to write second. (n – 1 choices)

.
.
.

Step *n*: Choose an element to write *n*th. (1 choice)

Multiplication rule: There are $n \times (n-1) \times \cdots \times 2 \times 1 = n!$ permutations.

r-Permutations

- An ***r*-permutation** of a set of *n* elements is an ordered selection of *r* elements taken from the set of *n* elements.
- The number of *r*-permutations of a set of *n* elements is denoted ***P(n, r).***

If *n* and *r* are integers and 1 ≤ *r* ≤ *n*, then the number of *r*-permutations of a set of *n* elements is given by the formula:

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

Example:

How many different ways can three of the letters of the word *EXAMPLE* be chosen and written in a row?

r-Combinations

- An **r-combination** of elements of a set is an unordered selection of r elements from the set.
- Thus, an r-combination is simply a subset of the set with r elements.

Example: Let S = {1, 2, 3, 4}.

Then {1, 3, 4} is a 3-combination from S.

- The number of r-combinations of a set with n distinct elements is denoted by C(n, r).

Example Cont.:

C(4, 2) = 6, since, for example, the 2-combinations of a set {1, 2, 3, 4} are {1, 2}, {1, 3}, {1, 4}, {2, 3}, {2, 4}, {3, 4}.

Number of r-Combinations

- C (n, r) is often also written as $\binom{n}{r}$, read "n choose r"
- $\binom{n}{r}$ is also called the binomial coefficient.
- Theorem:

$$C\ (n, r\ ) = \binom{n}{r} = \frac{n!}{r! \times (n-r)!}$$

Proof of Theorem

- What is the relationship between $P(n, r)$ and $C(n, r)$?

- Let's decompose $P(n, r)$ using the multiplication rule:

  First choose r elements

  Then, order these r elements

- How many ways to choose r elements from n? … $C(n, r)$
- How many ways to order r elements? … *r!*
- Thus, $P(n, r) = C(n, r) * r!$

Therefore, $C(n, r) = \frac{P(n,r)}{r!} = \frac{n!}{r! \times (n-r)!}$

Combinations – Examples

- We need to create a team with 5 members out of 10 candidates. How many different teams are possible?

- ✓ When creating a team, we don't care about order in which players were picked. Thus, we want C (n, r), not P(n, r)

- How many hands of 5 cards can be dealt from a standard deck of 52 cards?

Combinations

Corollary:

Let n and r be nonnegative integers with r ≤ n.

Then $C(n, r) = C(n, n - r)$.

✓ Note that **"picking a group of r people from a group of n people"** is the same as **"splitting a group of n people into a group of r people and another group of (n – r) people".**

Combinations Examples – in class

- How many bit strings of length 8 contain at least 6 ones?

C(8,6) + C(8,7) + C(8,8) = 28 + 8 + 1 = 37

- How many bit strings of length 8 contain at least 3 ones and at least 3 zeros?

C(8,3) + C(8,4) + C(8,5) = 56 + (35/12) + 56 ≈ 115

- If 3 persons are chosen at random from a set of 5 men and 6 women, what is the probability that 3 women are chosen?

$$P(E) = \frac{C(6,3)}{C(11,3)} = \frac{4}{33}$$

r-Combinations with repetition

- An ***r*-combination with repetition allowed**, or **multi-set of size *r***, chosen from a set *X* of *n* elements is an unordered selection of elements taken from *X* with repetition allowed.

- The number of *r*-combinations with repetition allowed (multi-sets of size *r*) that can be selected from a set of *n* elements is

$$\binom{r+n-1}{r}$$

A Motivating Example

- How many ways can I select 15 cans of soda from a cooler containing large quantities of Coke, Pepsi, Diet Coke, Root Beer and Sprite?
- We have to model this problem using the chart:

	Coke	Pepsi	Diet Coke	Root Beer	Sprite	
A:	111	111	111	111	111	=15
B:	11	111111	111111	1		=15
C:			1111	1111111	1111	=15

- Here, we set an order of the categories and just count how many from each category are chosen.

A Motivating Example

- Now, each event will contain fifteen 1’s, but we need to indicate where we transition from one category to the next. If we use 0 to mark our transitions, then the events become:

A: 1110111011101110111

B: 1100111111011111101

C: 0011110111111101111

- Thus, associated with each event is a binary string with #1’s = #things to be chosen and #0’s = #transitions between categories.

A Motivating Example

- From this example we see that the number of ways to select 15 sodas from a collection of 5 types of soda is:

$$C(15 + 4,15) = C(19,15) = C(19,4).$$

- Note that #zeros = #transitions = #categories - 1.

Theorem: The number of ways to fill r slots from n categories with repetition allowed is: $C(r + n - 1, r) = C(r + n - 1, n - 1)$.

- In words, the counts are: C(#slots + #transitions, #slots) or C(#slots + #transitions, #transitions).

Which Formula to Use?

	Order Matters	Order Does Not Matter
Repetition Is Allowed	n^k	$\binom{k+n-1}{k}$
Repetition Is Not Allowed	$P(n,k)$	$\binom{n}{k}$

Examples – in class

- How many ways to assign 3 jobs to 6 employees if every employee can be given more than one job?
  - Perm with repetition; each of three jobs can be done by 6 choices: $6^3 = 216$.
- How many different four-digit numbers can be formed from the digits 1, 2, 3, 4 (without repetition)?
  - Perm without repetition; P(4,4) = 4! / 0! = 24.
- How many nonnegative integer solutions exist for the equation x1 + x2 + x3 + x4 = 17?
  - Comb with repetition; 17 similar things of four types:

    C(17+4–1, 17) = C(20, 3) = 1140.

Pascal's Identity

Pascal's Identity:

- Let n and k be positive integers with $n \geq k$.
- Then $C(n + 1, k) = C(n, k – 1) + C(n, k)$.

How can this be explained?

What is it good for?

Pascal's Identity

Imagine a set S containing n elements and a set T containing (n + 1) elements, namely all elements in S plus a new element **a**.

✓ Calculating C(n + 1, k) is equivalent to answering the question:
How many subsets of T containing k items are there?

Case I: The subset contains (k – 1) elements of S
plus the element **a**: C(n, k – 1) choices.

Case II: The subset contains k elements of S and
does not contain **a**: C(n, k) choices.

Addition Rule: C(n + 1, k) = C(n, k – 1) + C(n, k).

Pascal's Triangle

In **Pascal's triangle**, each number is the sum of the numbers to its upper left and upper right:

					1					
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
…		…		…		…		…		…

Pascal's Triangle

Since we have $C(n + 1, k) = C(n, k - 1) + C(n, k)$ and $C(0, 0) = 1$, we can use Pascal's triangle to simplify the computation of $C(n, k)$:

k

n

$C(0, 0) =$ **1**

$C(1, 0) =$ **1** $C(1, 1) =$ **1**

$C(2, 0) =$ **1** $C(2, 1) =$ **2** $C(2, 2) =$ **1**

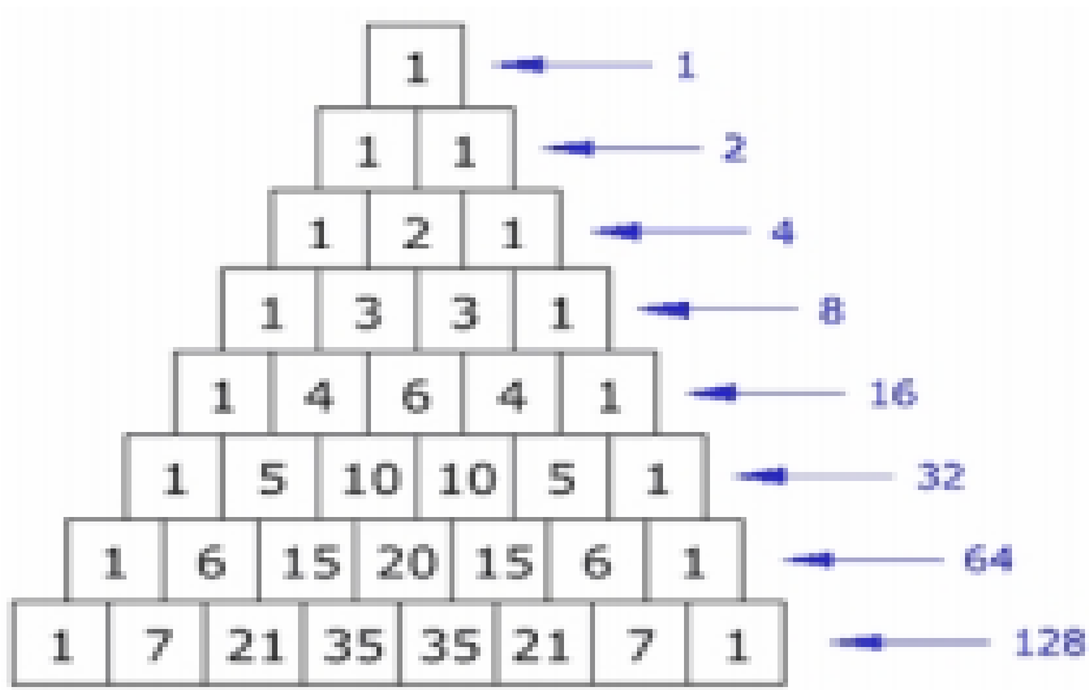
$C(3, 0) =$ **1** $C(3, 1) =$ **3** $C(3, 2) =$ **3** $C(3, 3) =$ **1**

$C(4, 0) =$ **1** $C(4, 1) =$ **4** $C(4, 2) =$ **6** $C(4, 3) =$ **4** $C(4, 4) =$ **1**

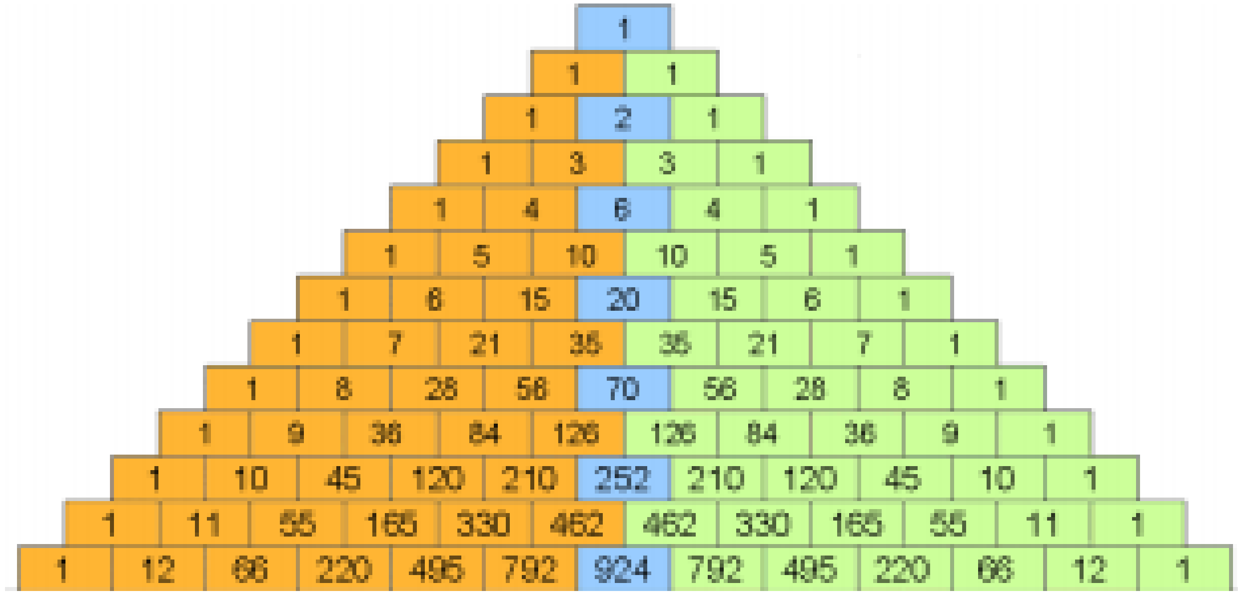
Interesting facts about Pascal's Triangle

What is the sum of numbers in n'th row in Pascal's triangle (starting at n = 0)?

- Observe: $\sum_{k=0}^{n} \binom{n}{k} = 2^n$



Interesting facts about Pascal's Triangle



- ✓ Pascal's triangle is perfectly symmetric.
- ✓ Numbers on left are mirror image of numbers on right.
- ✓ Which of the theorems we saw today explains this fact?

Binomial Coefficients

➢ Expressions of the form C(n, k) are also called **binomial coefficients**.

How come?

- A **binomial expression** is the sum of two terms, such as (a + b).

Now consider $(a + b)^2 = (a + b)(a + b)$.

- When expanding such expressions, we have to form all possible products of a term in the first factor and a term in the second factor:

$$(a + b)^2 = a \cdot a + a \cdot b + b \cdot a + b \cdot b$$

Then we can sum identical terms:

$$(a + b)^2 = a^2 + 2ab + b^2$$

Binomial Coefficients

For $(a + b)^3 = (a + b)(a + b)(a + b)$ we have:

$$(a + b)^3 = aaa + aab + aba + abb + baa + bab + bba + bbb$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

- There is only one term a^3, because there is only one possibility to form it: Choose **a** from all three factors: C(3, 3) = 1.
- There is three times the term a^2b, because there are three possibilities to choose **a** from two out of the three factors: C(3, 2) = 3.
- Similarly, there is three times the term ab^2 (C(3, 1) = 3)
- And once the term b^3 (C(3, 0) = 1).

Binomial Coefficients

This leads us to the following formula:

$$(a+b)^n = \sum_{j=0}^{n} C(n,j) \cdot a^{n-j} b^j$$

(Binomial Theorem)

- With the help of Pascal's triangle, this formula can considerably simplify the process of expanding powers of binomial expressions.

- For example, the fifth row of Pascal's triangle (1 – 4 – 6 – 4 – 1) helps us to compute $(a + b)^4$:

  - ✓ $(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$

Complimentary Events

- Let E be an event in a sample space S. The probability of an event –E, the **complimentary event** of E, is given by:

$$p(-E) = 1 - p(E).$$

- This can easily be shown:

$$p(-E) = (|S| - |E|)/|S| = 1 - |E|/|S| = 1 - p(E).$$

- This rule is useful if it is easier to determine the probability of the complimentary event than the probability of the event itself.

Complimentary Events - Examples

Example I: A sequence of 10 bits is randomly generated. What is the probability that at least one of these bits is zero?

Solution:

- ✓ There are $2^{10} = 1024$ possible outcomes of generating such a sequence.
- ✓ The event –E, **"none of the bits is zero"** , includes only one of these outcomes, namely the sequence 1111111111.
- ✓ Therefore, p(-E) = 1/1024.
- ✓ Now p(E) can easily be computed as:

$$p(E) = 1 – p(-E) = 1 – 1/1024 = 1023/1024.$$

Complim. Events Example - in class

What is the probability that at least two out of 36 people have the same birthday?

Solution:

- ✓ The sample space S encompasses all possibilities for the birthdays of the 36 people, so $|S| = 365^{36}$.
- ✓ Let us consider the event –E (“no two people out of 36 have the same birthday”).
- ✓ –E includes P(365, 36) outcomes (365 possibilities for the first person’s birthday, 364 for the second, and so on).
- ✓ Then $p(-E) = P(365, 36)/365^{36} = 0.168$, so $p(E) = 0.832$ or 83.2%

Probability Axioms

Let E_1 and E_2 be events in the sample space S.
Then, for all events E_1 and E_2 in S we have:

1. $0 \leq P(E_1) \leq 1$
2. $P(\emptyset) = 0 \text{ and } P(S) = 1$
3. $If\ E_1 \cap E_2 = \emptyset\ then\ \mathrm{P}(E_1 \cup E_2) = P(E_1) + P(E_2)$

Does this remind you of something?

Of course, the principle of **inclusion-exclusion**.

Discrete Probability Example

What is the probability of a positive integer selected at random from the set of positive integers not exceeding 100 to be divisible by 2 or 5?

Solution:

E_2: "integer is divisible by 2"

E_5: "integer is divisible by 5"

$E_2 = \{2, 4, 6, \ldots, 100\}$

$|E_2| = 50$

$p(E_2) = 0.5$

Discrete Probability Example

$E_5 = \{5, 10, 15, \ldots, 100\}$

$|E_5| = 20$

$p(E_5) = 0.2$

$E_2 \cap E_5 = \{10, 20, 30, \ldots, 100\}$

$|E_2 \cap E_5| = 10$

$p(E_2 \cap E_5) = 0.1$

$p(E_2 \cup E_5) = p(E_2) + p(E_5) - p(E_2 \cap E_5)$

$p(E_2 \cup E_5) = 0.5 + 0.2 - 0.1 = 0.6$

Discrete Probability

What happens if the outcomes of an experiment are **not** equally likely?

- In that case, we assign a probability $P(s)$ to each outcome $s \in S$, where S is the sample space.
- Two conditions have to be met:
  (1): $0 \leq P(s) \leq 1$ for each $s \in S$, and
  (2): $\sum_{s \in S} P(s) = 1$

This means, as we already know, that (1) each probability must be a value between 0 and 1, and (2) the probabilities must add up to 1, because one of the outcomes is **guaranteed** to occur.

Discrete Probability

How can we obtain these probabilities P(s) ?

The probability P(**s**) assigned to an outcome **s** equals the limit of the number of times **s** occurs divided by the number of times the experiment is performed.

Once we know the probabilities p(s), we can compute the **probability of an event E** as follows:

$$P(E) = \sum_{s \in E} P(s)$$

Discrete Probability Example

A die is biased so that the number 3 appears twice as often as each other number. What are the probabilities of all possible outcomes?

Solution:

There are 6 possible outcomes $s_1, \ldots, s_6$.

$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6)$

$p(s_3) = 2p(s_1)$

Since the probabilities must add up to 1, we have:

$5p(s_1) + 2p(s_1) = 1$

$7p(s_1) = 1$

$p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = 1/7$, $p(s_3) = 2/7$

Discrete Probability Example – in class

For the biased die from the previous Example, what is the probability that an odd number appears when we roll the die?

Solution:

$E_{odd} = \{s_1, s_3, s_5\}$

Remember the formula $p(E) = \sum_{s \in E} p(s)$.

$p(E_{odd}) = \sum_{s \in E_{odd}} p(s) = p(s_1) + p(s_3) + p(s_5)$

$p(E_{odd}) = 1/7 + 2/7 + 1/7 = 4/7 = 57.14\%$

Conditional Probability

If we toss a coin three times, what is the probability that an odd number of tails appears **(event E)**, if the first toss is a tail **(event F)** ?

- If the first toss is a tail, the possible sequences are TTT, TTH, THT, and THH.
- In two out of these four cases, there is an odd number of tails.
- Therefore, the probability of E, under the condition that F occurs, is 0.5.

- We call this **conditional probability**.

Conditional Probability

If we want to compute the **conditional probability** of E given F, we use F as the sample space.

✓ For any outcome of E to occur under the condition that F also occurs, this outcome must also be in $E \cap F$.

Definition: Let E and F be events with $P(F) > 0$.
The conditional probability of E given F, denoted by $P(E \mid F)$, is defined as:

$$P(E \mid F) = P(E \cap F)/P(F)$$

$P(F) = P(E \cap F) / P(E \mid F)$

$P(E \cap F) = P(E \mid F) \times P(F)$

Conditional Probability Example

What is the probability of a random bit string of length four contains at least two consecutive 0s, given that its first bit is a 0 ?

Solution:

E: "bit string contains at least two consecutive 0s"

F: "first bit of the string is a 0"

We know the formula $p(E \mid F) = p(E \cap F)/p(F)$.

$E \cap F = \{0000, 0001, 0010, 0011, 0100\}$

$p(E \cap F) = 5/16$

$p(F) = 8/16 = 1/2$

$p(E \mid F) = (5/16)/(1/2) = 10/16 = 5/8 = 0.625$

Baye’s Theorem - Example

- Suppose that one urn contains 2 blue and 3 gray balls and a second urn contains 6 blue and 4 gray balls. A ball is selected by choosing one of the urns at random and then picking a ball at random from that urn. If the chosen ball is blue, what is the probability that it came from the first urn?

Solution:

Let A be the event that the chosen ball is blue, $B1$ the event that the ball came from the first urn, and $B2$ the event that the ball came from the second urn.

$$P(A|B1) = 2/5 \text{ and } P(A|B2) = 6/10,$$

$$P(B1) = P(B2) = 1/2.$$

Baye's Theorem – Example Cont.

- Suppose that one urn contains 2 blue and 3 gray balls and a second urn contains 6 blue and 4 gray balls. A ball is selected by choosing one of the urns at random and then picking a ball at random from that urn. If the chosen ball is blue, what is the probability that it came from the first urn?

<u>Solution:</u>

$P(A \cap B1) = P(B1) \times P(A|B1)$ and $P(A \cap B2) = P(B2) \times P(A|B2)$

$P(A \cap B1) = (1/2) \times (2/5)$ and $P(A \cap B2) = (1/2) \times (6/10)$

$P(A \cap B1) = 1/5$ and $P(A \cap B2) = 3/10$

But *A* is the disjoint union of $(A \cap B1)$ and $(A \cap B2)$, so by probability axiom 3:
$P(A) = P(A \cap B1) + P(A \cap B2) = 1/2$.

$$P(B1 \mid A) = P(B1 \cap A) / P(A) = (1/5) / (1/2) = 2/5 = 40\%.$$

Baye's Theorem

Suppose that a sample space *S* is a union of **mutually disjoint** events *B*1, *B*2, *B*3, . . . ,*Bn*, suppose *A* is an event in *S*, and suppose *A* and all the *B*1 have nonzero probabilities.

If *k* is an integer with 1 ≤ *k* ≤ *n*, then:

$$\mathrm{P}(B_k|\mathrm{A}) = \frac{P(A|B_k)\,P(B_k)}{P(A|B_1)P(B_1)\ +\ P(A|B_2)P(B_2) + \ \dots\ +\ P(A|B_n)P(B_n)}$$

Independence

Let us return to the example of tossing a coin three times.

- Does the probability of event E (odd number of tails) **depend** on the occurrence of event F (first toss is a tail) ?

In other words, is it the case that $p(E \mid F) \neq p(E)$?

We actually find that $p(E \mid F) = 0.5$ and $p(E) = 0.5$,
so we say that E and F are **independent events**.

Independence

- We have $p(E \mid F) = p(E \cap F)/p(F)$,

- $p(E \mid F) = p(E)$ if and only if $p(E \cap F) = p(E)p(F)$.

Definition: The events E and F are said to be **independent** if and only if $p(E \cap F) = p(E)p(F)$.

Obviously, this definition is **symmetrical** for E and F. If we have $p(E \cap F) = p(E)p(F)$, then it is also true that $p(F \mid E) = p(F)$.

Independence Example – in class

Suppose E is the event that a randomly generated bit string of length four begins with a 1, and F is the event that a randomly generated bit string contains an even number of 0s. Are E and F independent?

Solution:

Obviously, $p(E) = p(F) = 0.5$.

$E \cap F = \{1111, 1001, 1010, 1100\}$

$p(E \cap F) = 0.25$

$p(E \cap F) = p(E)p(F)$

Conclusion: E And F are **independent**.