10. Counting and Probability 2

Aaron Tan

17-21 October 2016

9.5 Counting Subsets of a Set: Combinations

Counting Subsets of a Set: Combinations

- Given a set S with n elements, how many subsets of size r can be chosen from S?
- Each subset of size r is called an r-combination of the set.

Definition: r-combination

Let n and r be non-negative integers with $r \leq n$.

An ***r*-combination** of a set of n elements is a subset of r of the n elements.

$\binom{n}{r}$, read "n choose r", denotes the number of subsets of size r (r-combinations) that can be chosen from a set of n elements. Other symbols used are $C(n, r)$, ${}_nC_r$, $C_{n,r}$, or ${}^nC_r$.

Example 1 – 3-Combinations

Let S = {Ann, Bob, Cyd, Dan}. Each committee consisting of three of the four people in S is a 3-combination of S.

a. List all such 3-combinations of S.

b. What is $\binom{4}{3}$?

a. The 3-combinations are:

{Bob, Cyd, Dan}, {Ann, Cyd, Dan},
{Ann, Bob, Dan}, {Ann, Bob, Cyd}

b. $\binom{4}{3}$ = **4**

Example 2 – Ordered and Unordered Selection

Two distinct methods that can be used to select r objects from a set of n elements:

Ordered selection	**Unordered selection**
Also called r-permutation	Also called r-combination

Example: $S = \{ 1, 2, 3 \}$

2-permutations of S		2-combinations of S
$\{1, 2\}$	$\{2, 1\}$	$\{1, 2\}$
$\{1, 3\}$	$\{3, 1\}$	$\{1, 3\}$
$\{2, 3\}$	$\{3, 2\}$	$\{2, 3\}$

Example 3 – Relationship between Permutations and Combinations

Write all 2-permutations of the set {0, 1, 2, 3}. Find an equation relating the number of 2-permutations, $P(4, 2)$, and the number of 2-combinations, $\binom{4}{2}$, and solve this equation for $\binom{4}{2}$.

{0, 1}, {1, 0},
{0, 2}, {2, 0},
{0, 3}, {3, 0},
{1, 2}, {2, 1},
{1, 3}. {3, 1}.
{2, 3}, {3, 2}

According to Theorem 9.2.3,

$$P(4, 2) = 4!/(4-2)! = 4!/2! = 12$$

The construction of a 2-permutation of {0, 1, 2, 3} can be thought of comprising two steps:

Step 1: Choose a subset of 2 elements from {0, 1, 2, 3}.
Step 2: Choose an ordering for the 2-element subset.

Example 3 – Relationship between Permutations and Combinations

This can be illustrated by the following possibility tree:

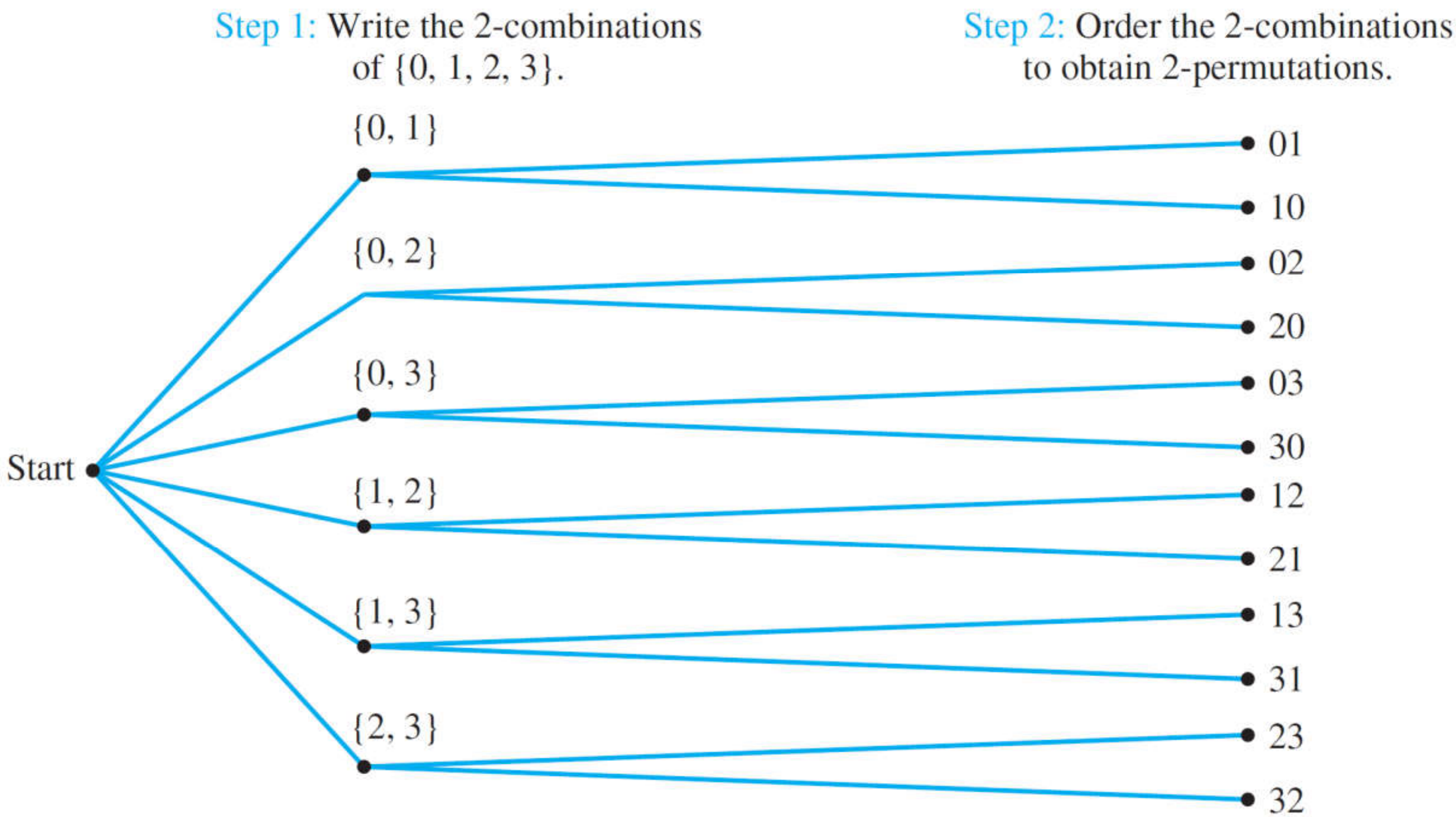


Figure 9.5.1 Relationship between Permutations and Combinations

Example 3 – Relationship between Permutations and Combinations

The number of ways to perform step 1 is $\binom{4}{2}$.

The number of ways to perform step 2 is 2!

Hence,

$$P(4, 2) = \binom{4}{2} \cdot 2!$$

$$\binom{4}{2} = P(4, 2) / 2! = 12 / 2 = 6$$

Theorem 9.5.1 Formula for $\binom{n}{r}$

The number of subsets of size *r* (or *r*-combinations) that can be chosen from a set of *n* elements, $\binom{n}{r}$, is given by the formula

$$\binom{n}{r} = \frac{P(n,r)}{r!}$$

or, equivalently,

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

where *n* and *r* are non-negative integers with $r \leq n$.

Example 4 – Teams with Members of Two Types

Suppose the group of 12 consists of 5 men and 7 women.

a. How many 5-person teams can be chosen that consist of 3 men and 2 women?

Hint: Think of it as a two-step process:

Step 1: Choose the men.

Step 2: Choose the women.

Example 4 – Teams with Members of Two Types

Suppose the group of 12 consists of 5 men and 7 women.

b. How many 5-person teams contain at least one man?

Hint: May use **difference rule** or **addition rule**.
The former is shorter.

Let *A* be the set of all 5-person teams,
and *B* be the set of 5-person teams without any men.

Then $N(A)$ = , and $N(B)$ =

Therefore number of 5-person teams that contain at least one man =

Example 4 – Teams with Members of Two Types

Suppose the group of 12 consists of 5 men and 7 women.

c. How many 5-person teams contain at most one man?

Number of teams without any man =

Number of teams with one man =

Therefore number of 5-person teams that contain at most one man =

Example 5 – Permutations of a Set with Repeated Elements

Order the letters in the word

M I S S I S S I P P I

How many distinguishable orderings are there?

Four-step process:

- **Step 1:** Choose a subset of 4 positions for the S's. $\binom{11}{4}$
- **Step 2:** Choose a subset of 4 positions for the I's. $\binom{7}{4}$
- **Step 3:** Choose a subset of 2 positions for the P's. $\binom{3}{2}$
- **Step 4:** Choose a subset of 1 position for the M. $\binom{1}{1}$

$$\binom{11}{4} \times \binom{7}{4} \times \binom{3}{2} \times \binom{1}{1} = \mathbf{34650}$$

Theorem 9.5.2 Permutations with sets of indistinguishable objects

Suppose a collection consists of *n* objects of which

n_1 are of type 1 and are indistinguishable from each other
n_2 are of type 2 and are indistinguishable from each other
:
n_k are of type *k* and are indistinguishable from each other

and suppose that $n_1 + n_2 + \ldots + n_k = n$. Then the number of distinguishable permutations of the *n* objects is

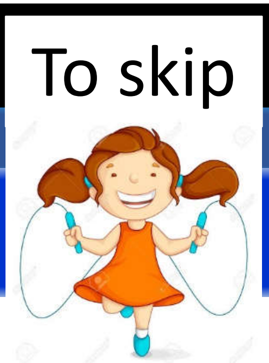
$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\cdots-n_{k-1}}{n_k}$$

$$= \frac{\boldsymbol{n}!}{\boldsymbol{n_1}!\boldsymbol{n_2}!\boldsymbol{n_3}!\cdots\boldsymbol{n_k}!}$$

The Number of Partitions of a Set into r Subsets

- In an ordinary (or *singly indexed*) sequence, integers n are associated to numbers a_n. In a *doubly indexed* sequence, ordered pairs of integers (m, n) are associated to numbers $a_{m,n}$.
- For example, combinations can be thought of as terms of the doubly indexed sequence defined by $C_{n,r} = \binom{n}{r}$ for all integers n and r with $0 \leq r \leq n$.
- An important example of a doubly indexed sequence is the sequence of *Stirling numbers of the second kind*.

The Number of Partitions of a Set into *r* Subsets



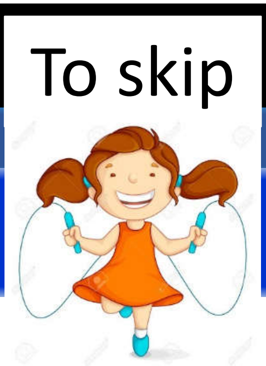
- Observe that if a set of three elements $\{x_1, x_2, x_3\}$ is partitioned into two subsets, then one of the subsets has one element and the other has two elements. Therefore, there are three ways the set can be partitioned.

$\{x_1, x_2\}\ \{x_3\}$ put x_3 by itself

$\{x_1, x_3\}\ \{x_2\}$ put x_2 by itself

$\{x_2, x_3\}\ \{x_1\}$ put x_1 by itself

The Number of Partitions of a Set into r Subsets

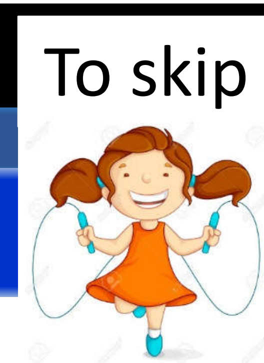


In general, let

$S_{n,r}$ = number of ways a set of size n can be partitioned into r subsets.

- Then, by the above, $S_{3,2} = 3$.
- The numbers $S_{n,r}$ are called Stirling numbers of the second kind.

Example 6 – Values of Stirling Numbers



Find $S_{4,1}$, $S_{4,2}$, $S_{4,3}$, and $S_{4,4}$.

Given a set with four elements, denote it by $\{x_1, x_2, x_3, x_4\}$.

Stirling number $S_{4,1} = 1$.

$\{x_1, x_2, x_3, x_4\}$

Stirling number $S_{4,4} = 1$.

$\{x_1\}\ \{x_2\}\ \{x_3\}\ \{x_4\}$

Example 6 – Values of Stirling Numbers

To skip

Stirling number $S_{4,2} = 7$.

Any partition of $\{x_1, x_2, x_3, x_4\}$ into 2 subsets must consist either of (a) 2 subsets of size two, *or*, of (b) 1 subset of size three and 1 subset of size one.

(a)

$\{x_1, x_2\}\ \{x_3, x_4\}$

$\{x_1, x_3\}\ \{x_2, x_4\}$

$\{x_1, x_4\}\ \{x_2, x_3\}$

(b)

$\{x_1\}\ \{x_2, x_3, x_4\}$

$\{x_2\}\ \{x_1, x_3, x_4\}$

$\{x_3\}\ \{x_1, x_2, x_4\}$

$\{x_4\}\ \{x_1, x_2, x_3\}$

Example 6 – Values of Stirling Numbers

Stirling number $S_{4,3} = 6$.

Any partition of $\{x_1, x_2, x_3, x_4\}$ into 3 subsets must have 2 elements in one subset and the other two elements in subsets by themselves.

There are $\binom{4}{2} = 6$ ways to choose the 2 elements to put together, which gives the following 6 possible partitions:

$\{x_1, x_2\}\ \{x_3\}\ \{x_4\}$
$\{x_1, x_3\}\ \{x_2\}\ \{x_4\}$
$\{x_1, x_4\}\ \{x_2\}\ \{x_3\}$

$\{x_2, x_3\}\ \{x_1\}\ \{x_4\}$
$\{x_2, x_4\}\ \{x_1\}\ \{x_3\}$
$\{x_3, x_4\}\ \{x_1\}\ \{x_2\}$

9.6 *r*-Combinations with Repetition Allowed

r-Combinations with Repetition Allowed

How many ways are there to choose *r* elements without regard to order from a set of *n* elements if *repetition is allowed*?

Definition: Multiset

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

If *X* = $\{x_1, x_2, \cdots, x_n\}$, we write an *r*-combination with repetition allowed as $[x_{i_1}, x_{i_2}, \cdots, x_{i_r}]$ where each x_{i_j} is in *X* and some of the x_{i_j} may equal each other.

Example 7 – *r*-Combinations with Repetition Allowed

Write a complete list to find the multisets of size 3, that can be selected from {1, 2, 3, 4}.

All combinations with 1, 1:	[1, 1, 1]; [1, 1, 2]; [1, 1, 3]; [1, 1, 4]
All additional combinations with 1, 2:	[1, 2, 2]; [1, 2, 3]; [1, 2, 4]
All additional combinations with 1, 3:	[1, 3, 3]; [1, 3, 4]
All additional combinations with 1, 4:	[1, 4, 4]
All additional combinations with 2, 2:	[2, 2, 2]; [2, 2, 3]; [2, 2, 4]
All additional combinations with 2, 3:	[2, 3, 3]; [2, 3, 4]
All additional combinations with 2, 4:	[2, 4, 4]
All additional combinations with 3, 3:	[3, 3, 3]; [3, 3, 4]
All additional combinations with 3, 4:	[3, 4, 4]
All additional combinations with 4, 4:	[4, 4, 4]

20 3-combinations with repetition allowed.

Example 7 – *r*-Combinations with Repetition Allowed

Consider the numbers 1, 2, 3, 4 in {1, 2, 3, 4} as categories and imagine choosing a total of 3 numbers from the categories with multiple selections from any category allowed.

	Category 1		Category 2		Category 3		Category 4
[1, 1, 1]:	xxx	|		|		|	
[1, 3, 4]:	x	|		|	x	|	x
[2, 4, 4]:		|	x	|		|	xx

Hence, we may write [1, 1, 1] as "xxx|||", [1,3,4] as "x||x|x" and [2,4,4] as "|x||xx".

This is the same as $\binom{6}{3}$ or **20**.

Theorem 9.6.1 Number of *r*-combinations with Repetition Allowed

The number of *r*-combination with repetition allowed (multisets of size *r*) that can be selected from a set of *n* elements is:

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

This equals the number of ways *r* objects can be selected from *n* categories of objects with repetitions allowed.

Example 8 – Counting Triples (i, j, k) with $1 \le i \le j \le n$

Consider the numbers 1, 2, 3, 4 in {1, 2, 3, 4} as categories and imagine choosing a total of 3 numbers from the categories with multiple selections from any category allowed.

Any triple of integers (i, j, k) with $1 \le i \le j \le k \le n$ can be represented as a string of $n - 1$ |'s and three x's, with the positions of the x's indicating which three integers from 1 to n are included in the triple.

Example:

For $n = 5$, we can represent [3,3,5] by "||xx||x" and [1,2,4] by "x|x||x|".

Which Formula to Use?

Earlier we have discussed four different ways of choosing *k* elements from *n*. The order in which the choices are made may or may not matter, and repetition may or may not be allowed. The following table summarizes which formula to use in which situation.

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

9.7 Pascal's Formula and the Binomial Theorem

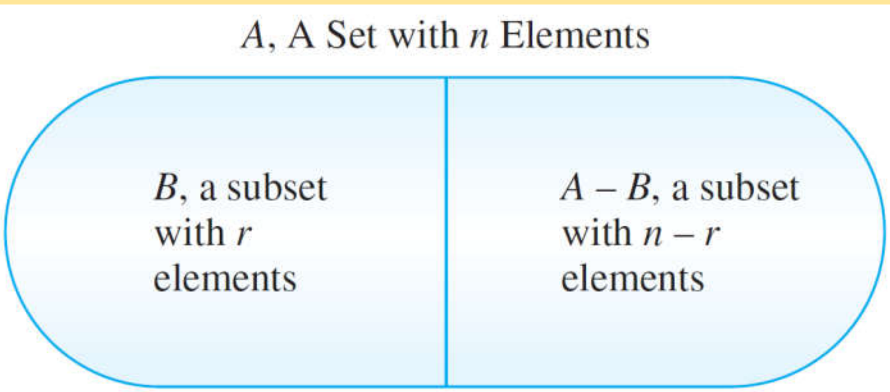
Example 9 – Deduce $\binom{n}{r} = \binom{n}{n-r}$

Deduce the formula

$$\binom{n}{r} = \binom{n}{n-r}$$

for all non-negative integers n and r with $r \leq n$, by interpreting it as saying that a set A with n elements has exactly as many subsets of size r as it has subsets of size $n - r$.

Observe that any subset of size r can be specified either by saying which r elements lie in the subset or by saying which $n - r$ elements lie outside the subset.



Any subset B with r elements completely determines a subset, $A - B$, with $n - r$ elements.

Example 9 – Deduce $\binom{n}{r} = \binom{n}{n-r}$

Suppose A has k subsets of size r : $B_1, B_2, \ldots, B_k$.

Then each B_i can be paired up with exactly one set of size $n - r$, namely its complement $A - B_i$ as shown below.

Subsets of Size r		Subsets of Size $n - r$
B_1	$\longleftrightarrow$	$A - B_1$
B_2	$\longleftrightarrow$	$A - B_2$
$\vdots$		$\vdots$
B_k	$\longleftrightarrow$	$A - B_k$

All subsets of size r are listed in the left-hand column, and all subsets of size $n - r$ are listed in the right-hand column.

Number of subsets of size r = number of subsets of size $n - r$.

The type of reasoning used in this example is called *combinatorial*, because it is obtained by counting things that are combined in different ways.

Pascal's Formula

Suppose *n* and *r* are positive integers with $r \leq n$. Then

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

Pascal's Formula

Pascal's triangle is a geometric version of Pascal's formula.

$n \backslash r$	0	1	2	3	4	5	...	$r-1$	r	...
0	1							·	·	...
1	1	1						·	·	...
2	1	2	1					·	·	...
3	1	3	3	1				·	·	...
4	1	4	6 +	4	1			·	·	...
5	1	5	10	= 10	5	1		·	·	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮
n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	...	$\binom{n}{r-1}$ +	$\binom{n}{r}$	...
$n+1$	$\binom{n+1}{0}$	$\binom{n+1}{1}$	$\binom{n+1}{2}$	$\binom{n+1}{3}$	$\binom{n+1}{4}$	$\binom{n+1}{5}$	...		= $\binom{n+1}{r}$	...
·		·	·	·	·	·		·	·	...
·		·	·	·	·	·		·	·	...
·		·	·	·	·	·		·	·	...

Table 9.7.1 Pascal's Triangle

Pascal's formula can be derived by two entirely different arguments. One is algebraic; it uses the formula for the number of *r*-combinations obtained in Theorem 9.5.1

The other is combinatorial; it uses the definition of the number of *r*-combinations as the number of subsets of size *r* taken from a set with a certain number of elements.

Theorem 9.7.1 Pascal's Formula

Let *n* and *r* be positive integers, $r \leq n$. Then

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

Example 10 – Deriving New Formulas from Pascal's Formula

Use Pascal's formula to derive a formula for $\binom{n+2}{r}$in terms of values of $\binom{n}{r}$, $\binom{n}{r-1}$ and $\binom{n}{r-2}$.
Assume *n* and *r* are non-negative and $2 \le r \le n$.

By Pascal's formulas, $\binom{n+2}{r} = \binom{n+1}{r-1} + \binom{n+1}{r}$.

Now apply Pascal's formula to $\binom{n+1}{r-1}$ and $\binom{n+1}{r}$ and substitute into the above to obtain:

$$\binom{n+2}{r} = \left[\binom{n}{r-2} + \binom{n}{r-1}\right] + \left[\binom{n}{r-1} + \binom{n}{r}\right]$$

$$\binom{n+2}{r} = \binom{n}{r-2} + 2\binom{n}{r-1} + \binom{n}{r}$$

The Binomial Theorem

In algebra a sum of two terms, such as *a* + *b*, is called binomial.

The *binomial theorem* gives an expression for the powers of a binomial $(a + b)^n$, for each positive integer *n* and all real numbers *a* and *b*.

$\binom{n}{r}$ is called a binomial coefficient

Theorem 9.7.2 Binomial Theorem

Given any real numbers *a* and *b* and any non-negative integer *n*,

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

$$= a^n + \binom{n}{1} a^{n-1} b^1 + \binom{n}{2} a^{n-2} b^2 + \cdots + \binom{n}{n-1} a^1 b^{n-1} + b^n$$

Example 11 – Substituting into the Binomial Theorem

Expand the following using the binomial theorem:

a. $(a+b)^5$ b. $(x-4y)^4$

a.

$$(a+b)^5 = \sum_{k=0}^{5} \binom{5}{k} a^{5-k} b^k$$

$$= a^5 + \binom{5}{1} a^{5-1} b^1 + \binom{5}{2} a^{5-2} b^2 + \binom{5}{3} a^{5-3} b^3 + \binom{5}{4} a^{5-4} b^4 + b^5$$

$$= a^5 + 5a^4 b + 10a^3 b^2 + 10a^2 b^3 + 5ab^4 + b^5$$

Example 11 – Substituting into the Binomial Theorem

Expand the following using the binomial theorem:

a. $(a+b)^5$ b. $(x-4y)^4$

b.

$$(x-4y)^4 = \sum_{k=0}^{4} \binom{4}{k} x^{4-k}(-4y)^k$$

$$= x^4 + \binom{4}{1}x^{4-1}(-4y)^1 + \binom{4}{2}x^{4-2}(-4y)^2 + \binom{4}{3}x^{4-3}(-4y)^3 + (-4y)^4$$

$$=$$

Example 12 – Using Combinatorial Argument to Derive some Identity

According to Theorem 6.3.1, a set with n elements has 2^n subsets.

Theorem 6.3.1 Number of elements in a Power Set

For all integers $n \geq 0$, if a set X has n elements, then $\wp(X)$ has 2^n elements.

Apply this fact to give a combinatorial argument to justify the identity

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$$

Example 12 – Using Combinatorial Argument to Derive some Identity

Suppose S is a set with n elements. Then every subset of S has some number k of elements, where $0 \le k \le n$.

It follows that the total number of subsets of S, $N(\wp(S))$, can be expressed as follows:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$$

By Theorem 6.3.1, $N(\wp(S)) = 2^n$. Hence

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$$

Example 13 – Using the Binomial Theorem to Simplify a Sum

Express the following sum in **closed form** (without using a summation symbol and without using an ellipsis …):

$$\sum_{k=0}^{n} \binom{n}{k} 9^k$$

$$\sum_{k=0}^{n} \binom{n}{k} 9^k = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} 9^k = (1+9)^n = 10^n$$

9.8 Probability Axioms and Expected Value

Probability Axioms

Recall: a sample space is a set of all outcomes of a random process or experiment and that an event is a subset of a sample space.

Probability Axioms

Let S be a sample space. A probability function P from the set of all events in S to the set of real numbers satisfies the following axioms: For all events A and B in S,

1. $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$ and $P(S) = 1$
3. If A and B are disjoint ($A \cap B = \emptyset$), then
   $P(A \cup B) = P(A) + P(B)$

Probability of the Complement of an Event

Suppose that A is an event in a sample space S. Deduce that $P(A^c) = 1 - P(A)$.

By Theorem 6.2.2 (5) Complement Laws:

$$\text{(a) } A \cup A^c = U \qquad \text{and} \qquad \text{(b) } A \cap A^c = \varnothing$$

With S playing the role of the universal set U,

$$A \cup A^c = S \qquad \text{and} \qquad A \cap A^c = \varnothing$$

Thus S is the disjoint union of A and A^c, and so

$$P(A \cup A^c) = P(A) + P(A^c) = P(S) = 1$$

Subtracting $P(A)$ from both sides:

$$P(A^c) = 1 - P(A)$$

Probability of the Complement of an Event

Probability of the Complement of an Event

If A is any event in a sample space S, then

$$P(A^c) = 1 - P(A)$$

Probability of a General Union of Two Events

Probability of a General Union of Two Events

If A and B are any events in a sample space S, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Proof using two steps:

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

b. Prove that for any events U and V in a sample space S, if $U \subseteq V$ then $P(V - U) = P(V) - P(U)$.

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

Part 1: To show $A \cup B \subseteq (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$

Given any element $x \in A \cup B$, x satisfies exactly one of the following three conditions:

1) $x \in A$ and $x \in B$

2) $x \in A$ and $x \notin B$

3) $x \in B$ and $x \notin A$

1) $x \in A \cap B$, and so
$x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$
by definition of union.

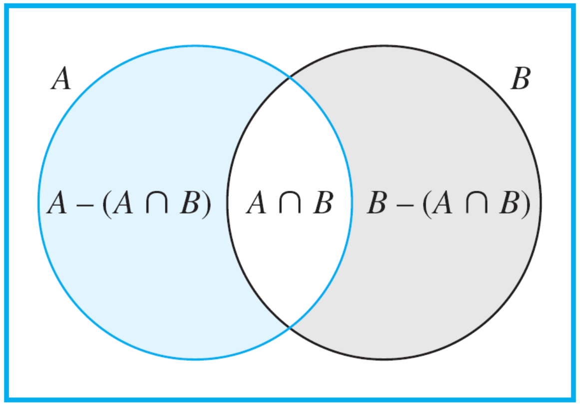


Figure 9.8.1

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

Part 1: To show $A \cup B \subseteq (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$

Given any element $x \in A \cup B$, x satisfies exactly one of the following three conditions:

1) $x \in A$ and $x \in B$

2) $x \in A$ and $x \notin B$

3) $x \in B$ and $x \notin A$

2) $x \notin A \cap B$ (because $x \notin B$) and so $x \in A - (A \cap B)$. Therefore, $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$ by definition of union.

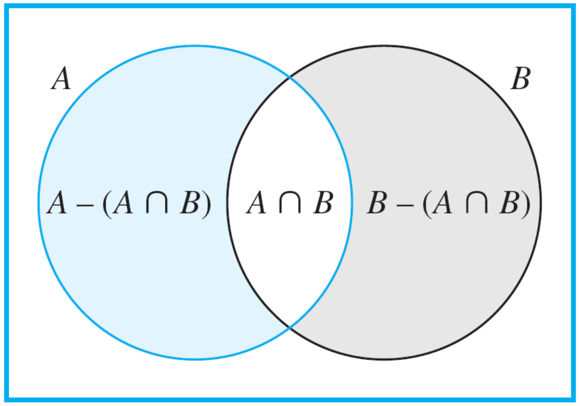


Figure 9.8.1

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

Part 1: To show $A \cup B \subseteq (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$

Given any element $x \in A \cup B$, x satisfies exactly one of the following three conditions:

1) $x \in A$ and $x \in B$

2) $x \in A$ and $x \notin B$

3) $x \in B$ and $x \notin A$

3) $x \notin A \cap B$ (because $x \notin A$) and so $x \in B - (A \cap B)$. Therefore, $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$ by definition of union.

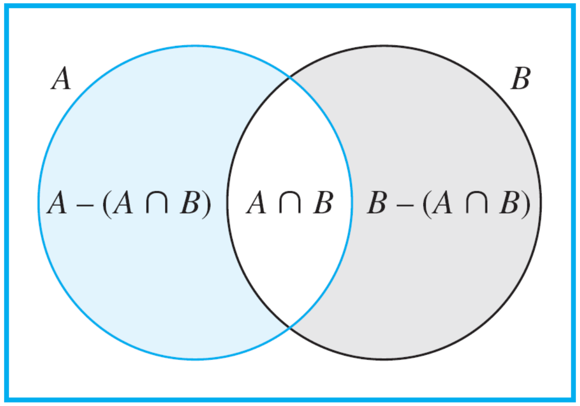


Figure 9.8.1

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

Part 1: To show $A \cup B \subseteq (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$

Given any element $x \in A \cup B$, x satisfies exactly one of the following three conditions:

1) $x \in A$ and $x \in B$

2) $x \in A$ and $x \notin B$

3) $x \in B$ and $x \notin A$

Hence, in all three cases,

$x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$.

Moreover, since the three conditions are mutually exclusive, the three sets $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$ are mutually disjoint.

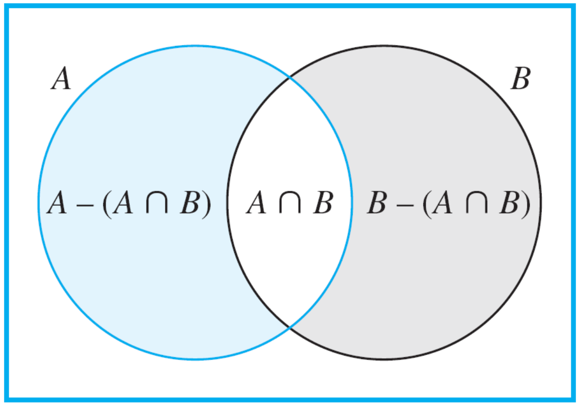


Figure 9.8.1

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

Part 2: To show $(A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B) \subseteq A \cup B$

Suppose x is any element in $(A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$.

By definition of union, $x \in A - (A \cap B)$ or $x \in B - (A \cap B)$ or $x \in A \cap B$.

In case $x \in A - (A \cap B)$, then $x \in A$ and $x \notin A \cap B$ by definition of set difference. In particular, $x \in A$ and so $x \in A \cup B$.

Similarly for the case $x \in B - (A \cap B)$, and the case $x \in A \cap B$.

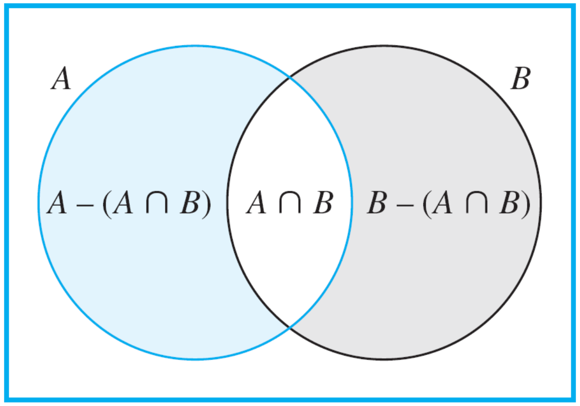


Figure 9.8.1

Probability of a General Union of Two Events

b. Prove that for any events U and V in a sample space S, if $U \subseteq V$ then $P(V - U) = P(V) - P(U)$.

For all sets U and V, $U \cup (V - U) = U \cup V$.

If $U \subseteq V$ then $U \cap (V - U) = \emptyset$ and $U \cup V = V$.

Therefore, $P(U) + P(V - U) = P(V)$

$\Rightarrow P(V - U) = P(V) - P(U)$.

Probability of a General Union of Two Events

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$. ✓

b. Prove that for any events U and V in a sample space S, if $U \subseteq V$ then $P(V - U) = P(V) - P(U)$. ✓

$$P(A \cup B) = P((A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)) \text{ by (a)}$$

$$= P(A - (A \cap B)) + P(B - (A \cap B)) + P(A \cap B)$$

$$= P(A) - P(A \cap B) + P(B) - P(A \cap B)) + P(A \cap B) \text{ by (b)}$$

$$= P(A) + P(B) - P(A \cap B)$$

Expected Value

People who buy lottery tickets regularly often justify the practice by saying that, even though they know that on average they will lose money, they are hoping for one significant gain, after which they believe they will quit playing.

Unfortunately, when people who have lost money on a string of losing lottery tickets win some or all of it back, they generally decide to keep trying their luck instead of quitting.

Expected Value

The technical way to say that on average a person will lose money on the lottery is to say that the *expected value* of playing the lottery is negative.

Definition: Expected Value

Suppose the possible outcomes of an experiment, or random process, are real numbers $a_1, a_2, a_3, \cdots, a_n$ which occur with probabilities $p_1, p_2, p_3, \cdots, p_n$. The **expected value** of the process is

$$\sum_{k=1}^{n} a_k p_k = a_1 p_1 + a_2 p_2 + a_3 p_3 + \cdots + a_n p_n$$

Example 14 – Expected Value of a Lottery

Suppose that 500,000 people pay \$5 each to play a lottery game with the following prizes:

- *A grand prize of \$1,000,000,*
- *10 second prizes of \$1,000 each,*
- *1,000 third prizes of \$500 each, and*
- *10,000 fourth prizes of \$10 each.*

What is the expected value of a ticket?

Each of the 500,000 lottery tickets has the same chance as any other of containing a winning lottery number, and so $p_k = \frac{1}{500000}$ for all k = 1, 2, 3, …, 500000.

Example 14 – Expected Value of a Lottery

Let $a_1, a_2, a_3, \ldots, a_{500000}$ be the net gain for an individual ticket, where $a_1 = 999995$.

- *A grand prize of $1,000,000,*
- *10 second prizes of $1,000 each,*
- *1,000 third prizes of $500 each, and*
- *10,000 fourth prizes of $10 each.*

$a_2 = a_3 = \ldots = a_{11} = 995$ (the net gain for each of the 10 second prize tickets)

$a_{12} = a_{13} = \ldots = a_{1011} = 495$ (the net gain for each of the 1000 third prize tickets)

$a_{1012} = a_{1013} = \ldots = a_{11011} = 5$ (the net gain for each of the 10000 fourth prize tickets)

$a_{11012} = a_{11013} = \ldots = a_{500000} = -5$ (the remaining 488989 tickets lose $5)

Example 14 – Expected Value of a Lottery

- *A grand prize of $1,000,000,*
- *10 second prizes of $1,000 each,*
- *1,000 third prizes of $500 each, and*
- *10,000 fourth prizes of $10 each.*

The expected value of a ticket is therefore:

$$\sum_{k=1}^{500000} a_k p_k = \sum_{k=1}^{500000} \left(a_k \cdot \frac{1}{500000} \right)$$

$$= \frac{1}{500000} \sum_{k=1}^{500000} a_k$$

$$=$$

$$=$$

In other words, a person who continues to play this lottery for a very long time will probably win some money occasionally but on average will lose $1.78 per ticket.

9.9 Conditional Probability, Bayes' Formula, and Independent Events

Conditional Probability

Imagine a couple with two children, each of whom is equally likely to be a boy or a girl. Now suppose you are given the information that one is a boy. What is the probability that the other child is a boy?

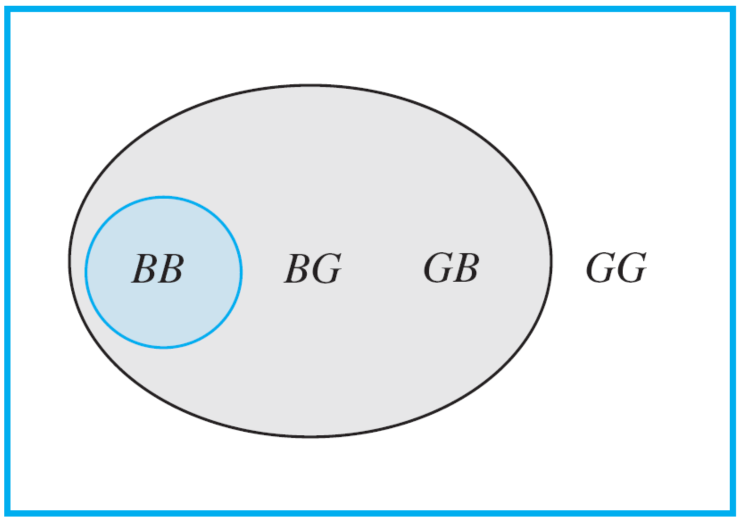
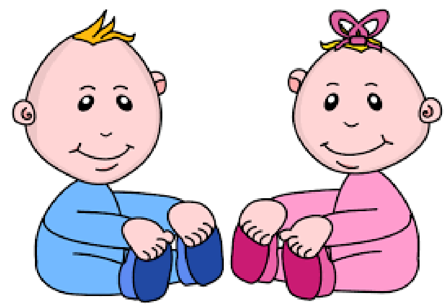


Figure 9.9.1



New sample space = gray region.

Within the new sample space, there is one combination where the other child is a boy (blue-gray region).

Hence, the likelihood that the other child is a boy given that at least one is a boy = 1/3.

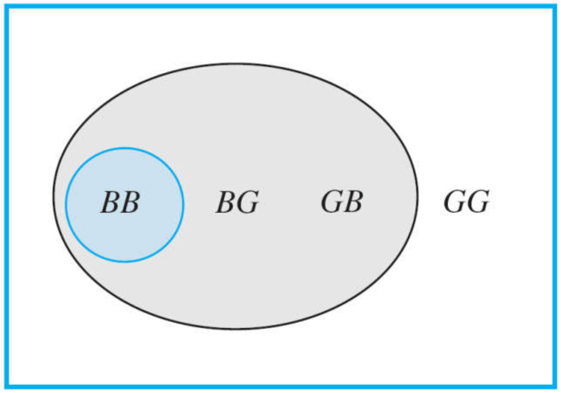


Figure 9.9.1

Note also

$$\frac{P(\text{at least one child is a boy and the other child is also a boy})}{P(\text{at least one child is a boy})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

A generalization of this observation forms the basis for the following definition.

Definition: Conditional Probability

Let A and B be events in a sample space S. If $P(A) \neq 0$, then the **conditional probability of B given A**, denoted $P(B|A)$, is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad 9.9.1$$

Multiplying both sides of formula 9.9.1 by $P(A)$, we get

$$P(A \cap B) = P(B|A) \cdot P(A) \qquad 9.9.2$$

Dividing both sides of formula 9.9.2 by $P(B|A)$, we get

$$P(A) = \frac{P(A \cap B)}{P(B|A)} \qquad 9.9.3$$

Example 15 – Representing Conditional Probabilities in a Tree Diagram

An urn contains 5 blue and 7 gray balls. Let us say that 2 are chosen at random, one after the other, without replacement

a. Find the following probabilities and illustrate them with a tree diagram: the probability that both balls are blue, the probability that the first ball is blue and the second is not blue, the probability that the first ball is not blue and the second ball is blue, and the probability that neither ball is blue.

b. What is the probability that the second ball is blue?

c. What is the probability that at least one of the balls is blue?

d. If the experiment of choosing two balls from the urn were repeated many times over, what would be the expected value of the number of blue balls?

Example 15 – Representing Conditional Probabilities in a Tree Diagram

An urn contains 5 blue and 7 gray balls. Let us say that 2 are chosen at random, one after the other, without replacement

Let

- S denote the sample space of all possible choices of two balls from the urn,
- B_1 be the event that the first ball is blue (then B_1^c is the event that the first ball is not blue),
- B_2 be the event that the second ball is blue (then B_2^c is the event that the second ball is not blue).

Example 15 – Representing Conditional Probabilities in a Tree Diagram

An urn contains 5 blue and 7 gray balls. Let us say that 2 are chosen at random, one after the other, without replacement.

a. Find the following probabilities and illustrate them with a tree diagram: the probability that both balls are blue, the probability that the first ball is blue and the second is not blue, the probability that the first ball is not blue and the second ball is blue, and the probability that neither ball is blue.

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray. a. Find $P(B_1 \cap B_2)$, $P(B_1 \cap B_2{}^c)$, $P(B_1{}^c \cap B_2)$, $P(B_1{}^c \cap B_2{}^c)$.

$$P(B_1) = \frac{5}{12}$$

$$P(B_1^c) = \frac{7}{12}$$

$$P(B_2|B_1) = \frac{4}{11}$$

$$P(B_2^c|B_1) = \frac{7}{11}$$

$$P(B_2|B_1^c) = \frac{5}{11}$$

$$P(B_2^c|B_1^c) = \frac{6}{11}$$

By formula 9.9.2

$$P(A \cap B) = P(B|A) \cdot P(A)$$

$$P(B_1 \cap B_2) = P(B_2|B_1) \cdot P(B_1) = \frac{4}{11} \cdot \frac{5}{12} = \frac{20}{132}$$

$$P(B_1 \cap B_2^c) = P(B_2^c|B_1) \cdot P(B_1) = \frac{7}{11} \cdot \frac{5}{12} = \frac{35}{132}$$

$$P(B_1^c \cap B_2) = P(B_2|B_1^c) \cdot P(B_1^c) = \frac{5}{11} \cdot \frac{7}{12} = \frac{35}{132}$$

$$P(B_1^c \cap B_2^c) = P(B_2^c|B_1^c) \cdot P(B_1^c) = \frac{6}{11} \cdot \frac{7}{12} = \frac{42}{132}$$

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray. a. Find $P(B_1 \cap B_2)$, $P(B_1 \cap B_2^c)$, $P(B_1^c \cap B_2)$, $P(B_1^c \cap B_2^c)$.

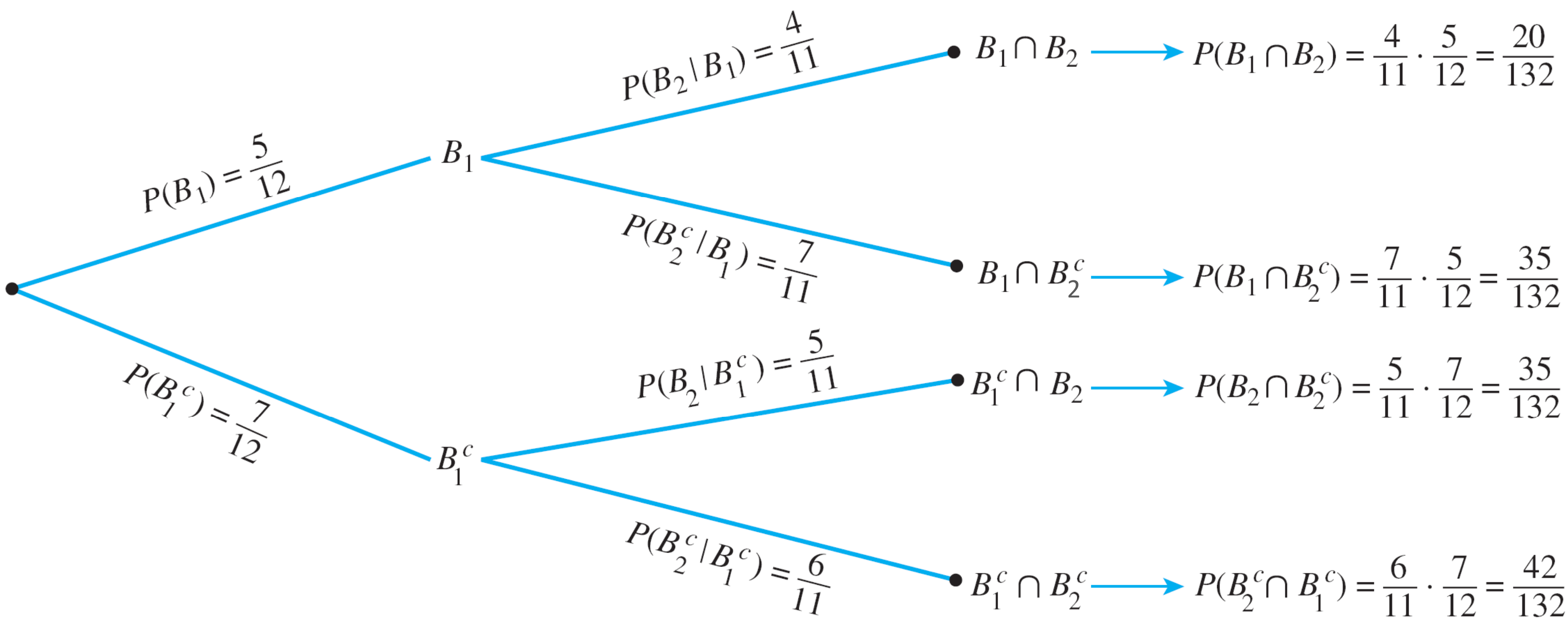


Figure 9.9.2

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray.

b. What is the probability that the 2nd ball is blue?

First ball is blue and second is also blue, OR
first ball is gray and second is blue.
They are mutually exclusive.

$$P(B_2) = P((B_2 \cap B_1) \cup (B_2 \cap B_1^c))$$

$$= P(B_2 \cap B_1) + P(B_2 \cap B_1^c)$$ by probability axiom 3

$$= \frac{20}{132} + \frac{35}{132} = \frac{55}{132} = \mathbf{\frac{5}{12}}$$

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray. c. What is the probability that at least one ball is blue?

Probability of a General Union of Two Events

If A and B are any events in a sample space S, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

$$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 \cap B_2)$$

$$=$$

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray. d. The expected value of the number of blue balls?

P(no blue balls) = 1 – *P*(at least one blue ball) $= 1 - \frac{15}{22} = \frac{7}{22}$

The event that one ball is blue → first ball is blue but second ball is not; or second ball is blue but first ball is not.

From part (a), $P(B_1 \cap B_2^c) = \frac{35}{132}$ and $P(B_1^c \cap B_2) = \frac{35}{132}$

Hence, *P*(1 blue ball) $= \frac{35}{132} + \frac{35}{132} = \frac{70}{132}$

From part (a), *P*(2 blue balls) $= \frac{20}{132}$

Example 15 – Representing Conditional Probabilities in a Tree Diagram

5 blue, 7 gray. d. The expected value of the number of blue balls?

P(no blue balls) = $\frac{7}{22}$ P(1 blue ball) $= \frac{70}{132}$ P(2 blue balls) $= \frac{20}{132}$

Expected value of the number of blue balls

= 0·P(no blue balls) + 1·P(1 blue ball) + 2·P(2 blue balls)

$=$

$=$

Bayes' Theorem

Suppose that one urn contains 3 blue and 4 gray balls and a second urn contains 5 blue and 3 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**?

This problem can be solved by carefully interpreting all the information that is known and putting it together in just the right way.

Let

- A be the event that the chosen ball is blue,
- B_1 the event that the ball came from the first urn, and
- B_2 the event that the ball came from the second urn.

Bayes' Theorem

3 of the 7 balls in the first urn are blue, 5 of the 8 balls in the second urn are blue: $P(A|B_1) = \frac{3}{7}$ and $P(A|B_2) = \frac{5}{8}$

The urns are equally like to be chosen: $P(B_1) = P(B_2) = \frac{1}{2}$

By formula 9.9.2 $P(A \cap B) = P(B|A) \cdot P(A)$

$$P(A \cap B_1) = (A|B_1) \cdot P(B_1) = \frac{3}{7} \cdot \frac{1}{2} = \frac{3}{14}$$

$$P(A \cap B_2) = (A|B_2) \cdot P(B_2) = \frac{5}{8} \cdot \frac{1}{2} = \frac{5}{16}$$

A is a disjoint union of $(A \cap B_1)$ and $(A \cap B_2)$, so by probability axiom 3,

$$P(A) = P\big((A \cap B_1) \cup (A \cap B_2)\big)$$
$$= P(A \cap B_1) + P(A \cap B_2) = \frac{3}{14} + \frac{5}{16} = \frac{59}{112}$$

Bayes' Theorem

3 of the 7 balls in the first urn are blue, 5 of the 8 balls in the second urn are blue.

From previous slide, $P(A \cap B_1) = \frac{3}{14}$ and $P(A) = \frac{59}{112}$

By definition of conditional probability,

$$P(B_1|A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{\frac{3}{14}}{\frac{59}{112}} = \frac{336}{826} \cong 40.7\%$$

Thus, if the chosen ball is blue, the probability that it came from the first urn is approximately 40.7%.

Theorem 9.9.1 Bayes' Theorem

Suppose that a sample space S is a union of mutually disjoint events $B_1, B_2, B_3, \ldots, B_n$.

Suppose A is an event in S, and suppose A and all the B_i have non-zero probabilities.

If k is an integer with $1 \leq k \leq n$, then

$$P(B_k|A) = \frac{P(A|B_k) \cdot P(B_k)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \cdots + P(A|B_n) \cdot P(B_n)}$$

Example 16 – Applying Bayes' Theorem

Most medical tests occasionally produce incorrect results, called false positives and false negatives.

When a test is designed to determine whether a patient has a certain disease, a **false positive** result indicates that a patient has the disease when the patient does not have it.

A **false negative** result indicates that a patient does not have the disease when the patient does have it.

Example 16 – Applying Bayes' Theorem

Consider a medical test that screens for a disease found in 5 people in 1,000. Suppose that the false positive rate is 3% and the false negative rate is 1%.

Then 99% of the time a person who has the condition tests positive for it, and 97% of the time a person who does not have the condition tests negative for it.

a. What is the probability that a randomly chosen person who tests positive for the disease actually has the disease?

b. What is the probability that a randomly chosen person who tests negative for the disease does not indeed have the disease?

Example 16 – Applying Bayes' Theorem

Consider a person chosen at random from among those screened. Let

- A be the event that the person tests positive for the disease,
- B_1 the event that the person actually has the disease, and
- B_2 the event that the person does not have the disease.

Then

$$P(A|B_1) = 0.99 \qquad P(A^c|B_2) = 0.97$$

$$P(A^c|B_1) = 0.01 \qquad P(A|B_2) = 0.03$$

Also, because 5 people in 1000 have the disease,

$$P(B_1) = 0.005 \qquad P(B_2) = 0.995$$

Example 16 – Applying Bayes' Theorem

a. What is the probability that a randomly chosen person who tests positive for the disease actually has the disease?

By Bayes' Theorem,

$$P(B_1|A) = \frac{P(A|B_1)\cdot P(B_1)}{P(A|B_1)\cdot P(B_1) + P(A|B_2)\cdot P(B_2)}$$

$$= \frac{(0.99)\cdot(0.005)}{(0.99)\cdot(0.005) + (0.03)\cdot(0.995)}$$

$$\cong 0.1422 \cong \mathbf{14.2\%}$$

Thus the probability that a person with a positive test result actually has the disease is approximately 14.2%.

Example 16 – Applying Bayes' Theorem

b. What is the probability that a randomly chosen person who tests negative for the disease does not indeed have the disease?

By Bayes' Theorem,

$$P(B_2|A^c) = \frac{P(A^c|B_2)\cdot P(B_2)}{P(A^c|B_1)\cdot P(B_1) + P(A^c|B_2)\cdot P(B_2)}$$

$$= \frac{(0.97)\cdot(0.995)}{(0.01)\cdot(0.005) + (0.97)\cdot(0.995)}$$

$$\cong 0.999948 \cong \mathbf{9\ 99\ 9\ 5\%}$$

Thus the probability that a person with a negative test result does not have the disease is approximately 99.995%.

Independent Events

A is the event that a head is obtained on the first toss and B is the event that a head is obtained on the second toss, then if the coin is tossed randomly both times, events A and B should be ***independent*** in the sense that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

To obtain the final form for definition of independence, observe that

If $P(B) \neq 0$ and $P(A|B) = P(A)$,

then $P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B)$

By the same argument,

If $P(A) \neq 0$ and $P(B|A) = P(B)$,
then $P(A \cap B) = P(B|A) \cdot P(A) = P(A) \cdot P(B)$

Conversely,

If $P(A \cap B) = P(A) \cdot P(B)$ and $P(A) \neq 0$, then $P(B|A) = P(B)$.

If $P(A \cap B) = P(A) \cdot P(B)$ and $P(B) \neq 0$, then $P(A|B) = P(A)$.

Thus, for convenience and to eliminate the requirement that the probabilities be nonzero, we use the following product formula to define independent events.

Definition: Independent Events

If A and B are events in a sample space S, then A and B are **independent**, if and only if,

$$P(A \cap B) = P(A) \cdot P(B)$$

Example 17 – Computing Probabilities of Intersections of Two Independent Events

A coin is loaded so that the probability of heads is 0.6. Suppose the coin is tossed twice. Although the probability of heads is greater than the probability of tails, there is no reason to believe that whether the coin lands heads or tails on one toss will affect whether it lands heads or tails on the other toss. Thus it is reasonable to assume that the results of the tosses are independent.

a. What is the probability of obtaining two heads?

b. What is the probability of obtaining one head?

c. What is the probability of obtaining no heads?

d. What is the probability of obtaining at least one head?

Example 17 – Computing Probabilities of Intersections of Two Independent Events

Sample space S consists of the 4 outcomes {HH, HT, TH, TT} which are not equally likely.

Let

- E be the event that a head is obtained on the first toss
- F be the event that a head is obtained on the second toss

$P(E) = P(F) = 0.6$.

Example 17 – Computing Probabilities of Intersections of Two Independent Events

a. What is the probability of obtaining two heads?

$P(\text{two heads}) = P(E \cap F) = P(E) \cdot P(F) = (0.6)(0.6) = 0.36 =$ **36%**

b. What is the probability of obtaining one head?

$P(\text{one head}) = P((E \cap F^c) \cup (E^c \cap F)) = P(E) \cdot P(F^c) + P(E^c) \cdot P(F)$

$= (0.6)(0.4) + (0.4)(0.6) = 0.48 =$ **48%**

c. What is the probability of obtaining no heads?

$P(\text{no heads}) =$

Example 17 – Computing Probabilities of Intersections of Two Independent Events

d. What is the probability of obtaining at least one head?

Method 1:

P(at least one head) = P(one head) + P(two heads) = 0.48 + 0.36
= 0.84 = **84%**

Method 2:

P(at least one head) = 1 – P(no heads) = 1 – 0.16 = 0.84 = **84%**

Pairwise Independent/Mutually Independent

We say three events A, B, and C are *pairwise independent* if, and only if,

$P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$ and

$P(B \cap C) = P(B) \cdot P(C)$

Events can be pairwise independent without satisfying the condition
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$.

Conversely, they can satisfy the condition
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ without being pairwise independent.

Four conditions must be included in the definition of independence for three events

Definition: Pairwise Independent and Mutually Independent

Let A, B and C be events in a sample space S. A , B and C are **pairwise independent**, if and only if, they satisfy conditions 1 – 3 below. They are **mutually independent** if, and only if, they satisfy all four conditions below.

1. $P(A \cap B) = P(A) \cdot P(B)$
2. $P(A \cap C) = P(A) \cdot P(C)$
3. $P(B \cap C) = P(B) \cdot P(C)$
4. $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

The definition of mutual independence for any collection of n events with $n \geq 2$ generalizes the two definitions given previously.

Definition: Mutually Independent

Events $A_1, A_2, \ldots, A_n$ in a sample space S are **mutually independent** if, and only if, the probability of the intersection of any subset of the events is the product of the probabilities of the events in the subset.

Example 18 – Tossing a Loaded Coin Ten Times

A coin is loaded so that the probability of heads is 0.6 (and thus the probability of tails is 0.4). Suppose the coin is tossed ten times. As in Example 17, it is reasonable to assume that the results of the tosses are mutually independent.

a. What is the probability of obtaining eight heads?

b. What is the probability of obtaining at least eight head?

For each $i = 1, 2, \ldots, 10$, let H_i be the event that a head is obtained on the ith toss, and let T_i be the event that a tail is obtained on the ith toss.

Example 18 – Tossing a Loaded Coin Ten Times

a. What is the probability of obtaining eight heads?

Suppose that the eight heads occur on the first eight tosses and that the remaining two tosses are tails. This is the event

$$H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5 \cap H_6 \cap H_7 \cap H_8 \cap T_9 \cap T_{10}.$$

For simplicity, we denote it as *HHHHHHHHTT*.

By definition of mutually independent events,

$$P(HHHHHHHHTT) = (0.6)^8(0.4)^2$$

By commutative law for multiplication, if the eight heads occur on any other of the ten tosses, the same number is obtained. Eg:

$$P(HHTHHHHHTH) = (0.6)^2(0.4)(0.6)^5(0.4)(0.6) = (0.6)^8(0.4)^2$$

Example 18 – Tossing a Loaded Coin Ten Times

a. What is the probability of obtaining eight heads?

Now there are as many different ways to obtain eight heads in ten tosses as there are subsets of eight elements (the toss numbers on which heads are obtained) that can be chosen from a set of ten elements. This number is $\binom{10}{8}$.

Hence

$$P(\text{eight heads}) = \binom{10}{8}(0.6)^8(0.4)^2$$

Example 18 – Tossing a Loaded Coin Ten Times

b. What is the probability of obtaining at least eight heads?

By similar reasoning,

$$P(\text{nine heads}) = \binom{10}{9}(0.6)^9(0.4)$$

and

$$P(\text{ten heads}) = \binom{10}{10}(0.6)^{10}$$

Therefore,

$$P(\text{at least 8 heads}) = P(\text{8 heads}) + P(\text{9 heads}) + P(\text{10 heads})$$

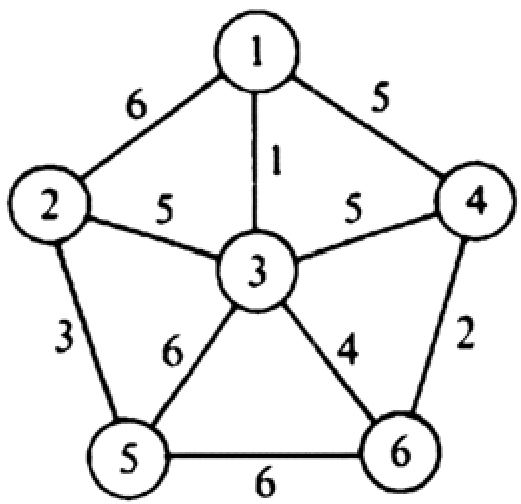
$$= \binom{10}{8}(0.6)^8(0.4)^2 + \binom{10}{9}(0.6)^9(0.4) + \binom{10}{10}(0.6)^{10}$$

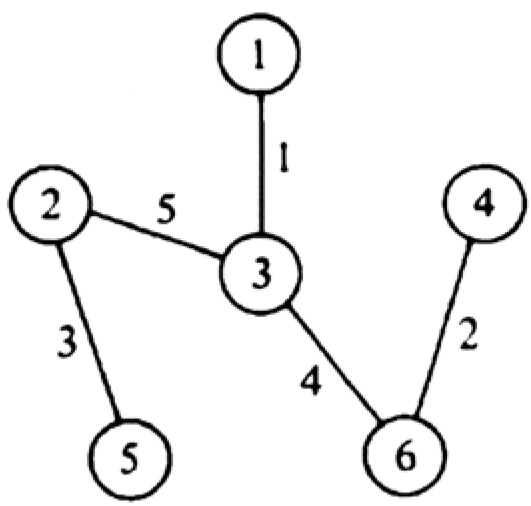
$\cong 0.167 =$ **16.7%**

Probabilities of the form $\binom{n}{k}p^{n-k}(1-p)^k$, where $0 \leq p \leq 1$, are called **binomial probabilities**.

Next week's lectures

Graphs and Trees





END OF FILE