Counting and Probability

Permutations, Addition Rule, Inclusion/Exclusion

Lecture Notes #11

Peter Blomgren
Department of Mathematics and Statistics
San Diego State University
San Diego, CA 92182-7720
blomgren@terminus.SDSU.EDU
http://terminus.SDSU.EDU

$Id: lecture.tex,v 1.2 2006/11/13 21:31:57 blomgren Exp $

Quick Recap

Last lecture we started talking about ***Counting and Probability***.

We introduced the concepts: **random process**, **sample space (S)** (all the possible outcomes of a random process), **event (E)** (a subset of the sample space), and **probability** (the relative size of the event vs. the sample space):

$$P(E) = \frac{n(E)}{n(S)} = \frac{\textbf{\# element is the event}}{\textbf{\# elements in the sample space}}$$

this formula is valid *if and only if* all outcomes are ***equally likely***.

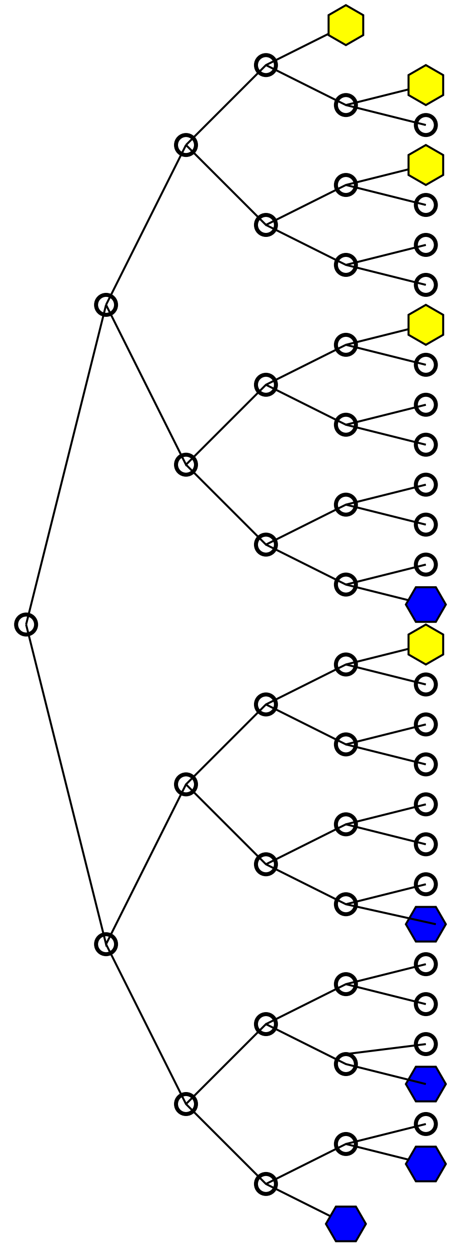
We counted elements in a list, looked at the probability of outcomes when tossing coins, introduced the concept of a ***possibility tree*** (which shows all possible combinations of sequential events), and introduced the ***multiplication rule*** for independent events.

Recap: Key Concepts — Possibility Tree

The previous figure shows all the possible ways the world series can play out, ***but*** there are multiple ways to reach some (most) states; *e.g.* the scenario "A wins, B wins" and "B wins, A wins" end up in the same **state** (one win for each team).

In a possibility tree, these two paths are differentiated; the possibility tree for the first 5 games looks like this:

Figure (to the right:) The possibility tree for the first 5 games of the world series. Note that 2 (out of 16) paths terminate after 4 games. An additional 8 paths terminate after 5 games...



Recap: Key Concepts — Independence / Multiplication Rule

If we have a sequence of events which are ***independent*** (note that this does not apply to the world series, since depending on the outcome of previous games, games #5, #6, and #7 may not be played) the multiplication rule applies:

Theorem: ***Multiplication Rule*** **—**
If an operation consists of k steps and step $\#i$ can be performed in n_i ways $i = 1, 2, \ldots, k$, then the entire operation can be performed in $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ ways.

If all 7 games of the world series were played no matter what the outcome of the previously played games:

$$k = 7, \quad n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = n_7 = 2$$

$\mathbf{2^7 = 128}$ **possibilities.**

A **permutation** of a set of objects is an **ordering** of the objects. For example, the set {**a,b,c**} has six permutations:

abc, **acb**, **bac**, **bca**, **cab**, **cba**

Question: How many permutations does a set with n elements have???

The first element can be selected in n ways, the second in $(n-1)$ ways, the third in $(n-2)$ ways, ...

$$\#\textbf{Permutations}(n) \quad = \quad n \cdot (n-1) \cdot \ldots \cdot 1 \quad = \quad \mathbf{n}!$$

Theorem: For any integer $n \geq 1$, the number of permutations of a set with n elements is $n!$ (n-factorial).

Example: Permuting at the Dinner Table

We are to seat six dinner guests around a table:

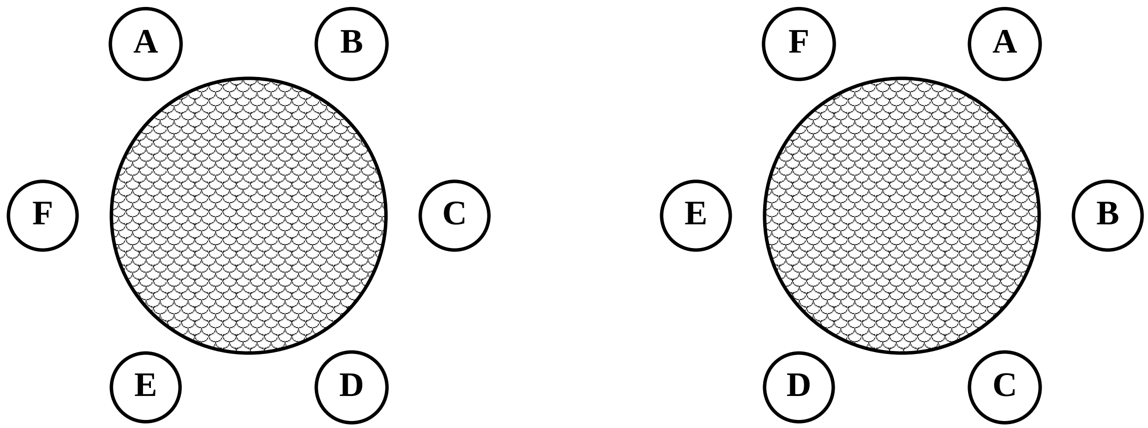


Figure: Two seating arrangements are considered the same if they are just a rotation of each other.

Question: How many seating arrangements are there, taking ***rotational symmetry*** into consideration?

Solution: We can take one guest and put him/her in a fixed position; — then the other five can be seated in $\mathbf{5! = 120}$ different ways relative to the first guest.

Example: Permuting at the Dinner Table, Take #2

We are to seat six dinner guests around a table:

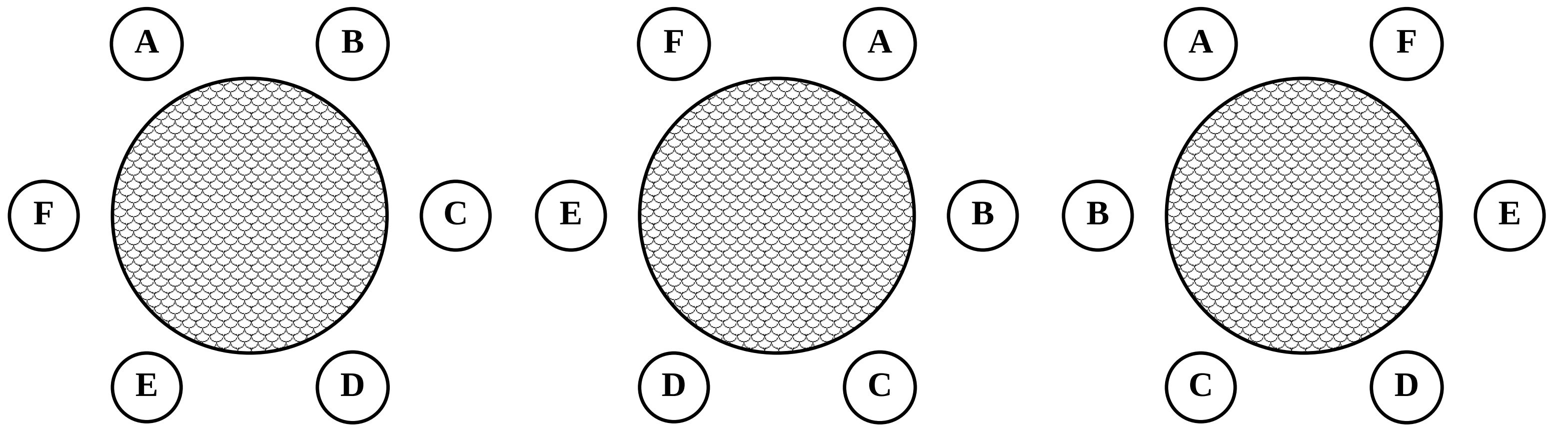


Figure: Two seating arrangements are considered the same if they are a rotation of each other and/or a reflection of each other.

Question: How many seating arrangements are there, taking ***rotational and reflective symmetry*** into consideration?

Solution: Since each seating arrangement has a mirror image, we now effectively have $\frac{1}{2} \cdot 5! = 60$ different seating arrangements.

Permutations of Selected Elements

Given the set $\mathbf{S} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ there are six ways to select two letters from S and write them in order:

$$ab \quad ac \quad ba \quad bc \quad ca \quad cb$$

Each such ordering of 2 elements of S is called a ***2-permutation*** of S.

> ***Definition: r-permutation —***
> An ***r-permutation*** of a set of n elements is an ordered selection of r elements taken from the set. The number of r-permutations of a set of n elements is denoted $\mathbf{P(n, r)}$.

I wonder if we could create a game using a set with 52 elements, and consider the 5-permutations...

$\mathbf{P(n,r)}$ — Counting the r-permutations

Theorem: If n and r are integers and $1 \leq r \leq n$, then the number of r-permutations of a set of n elements is given by the formula

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

Proof: There are $(n-0)$ ways to make the first choice, $(n-1)$ ways to make the second choice, ..., $(n-(r-1))$ ways to make the r$^{\text{th}}$ choice, therefore the number of combinations are

$$n\cdot(n-1)\cdot\ldots\cdot(n-(r-1)) = n\cdot(n-1)\cdot\ldots\cdot(n-r+1)$$

Now we notice

$$\frac{n!}{(n-r)!} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-r+1)\cdot(n-r)!}{(n-r)!} = n\cdot(n-1)\cdot\ldots\cdot(n-r+1). \square$$

Evaluating $\mathbf{P(n, r)}$

Problem: How many 4-permutations are there of a set of 15 elements?

Solution #1: We can just plug in and evaluate

$$P(15, 4) = \frac{15!}{(15-4)!} = \frac{15!}{11!} = \frac{1,307,674,368,000}{39,916,800} = 32,760$$

However, this can become problematic if n is large (my, quite ancient, calculator can only compute up to 69! $(1.711\ldots \times 10^{98})$...

Solution #2: Think about what the denominator does, *i.e.* canceling the “tail” of the factorial in the numerator:

$$P(15, 4) = \frac{15!}{(15-4)!} = \frac{15!}{11!} = 15 \cdot 14 \cdot 13 \cdot 12 = 32,760$$

Example: Proving $P(n,2) + P(n,1) = n^2$

> ***Proposition:*** For all integers $n \geq 2$,
>
> $$P(n,2) + P(n,1) = n^2$$

Proof: Let n be an integer ≥ 2, and use the theorem on slide 9, *i.e.*

$$\begin{aligned} P(n,2) &= \frac{n!}{(n-2)!} = n(n-1) \\ P(n,1) &= \frac{n!}{(n-1)!} = n \end{aligned}$$

and therefore

$$P(n,2) + P(n,1) = n(n-1) + n = n^2 - n + n = n^2$$

$\forall n \geq 2. \quad \square$

The Addition Rule — Counting Elements of Disjoint Sets

Theorem: If $\{\mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_n}\}$ is a partition of $\mathbf{A}$, then

$$n(A) = n(A_1) + n(A_2) + \ldots + n(A_n)$$

The formal proof (exercise ***Epp-6.3.33***) uses mathematical induction. Intuitively it is clear: each element in A is a member of exactly one of the sets A_i, so the element count on both the left- and right-hand-side must be the same.

Example: Counting 4-6 Digit PINs

Problem: In order to use the Big Bank's ATMs the user must have a 4–6 digit PIN (each digit is an integer 0–9). How many such PINs are there?

Solution:

4-digit PINs	$10^4 = 10,000$
5-digit PINs	$10^5 = 100,000$
6-digit PINs	$10^6 = 1,000,000$
4–6-digit PINs	$1,110,000$

The Difference Rule

Theorem: If A is a set with finitely many elements, and B a subset of A, $B \subset A$, then

$$n(A - B) = n(A) - n(B)$$

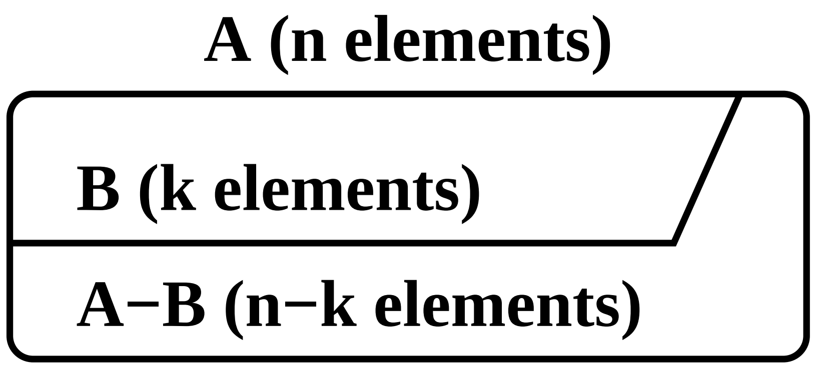


Figure: The Difference Rule — Illustration.

Example: Counting 4-6 Digit PINs (with/without repetition)

Problem: In order to use the Big Bank's ATMs the user must have a 4–6 digit PIN (each digit is an integer 0–9). How many such PINs are there? — How many PINs have no repeating digits; How many PINs have repeating digits?)

Solution:

PINs	Any Digits	No Repetition	With Repetition
4-digit	$10^4 = 10,000$	$P(10,4) = 5,040$	$4,960$
5-digit	$10^5 = 100,000$	$P(10,5) = 30,240$	$69,760$
6-digit	$10^6 = 1,000,000$	$P(10,6) = 151,200$	$848,800$
4–6-digit	$1,110,000$	$186,480$	$923,520$

Thus requiring non-repeating passwords limits the password space quite severely.

Example: Random PINs

Problem: When you get your ATM card from the Big Bank, you are assigned a random 4–6-digit PIN. What is the probability that the PIN will have repeated digits? What is the probability that it will not?

Solution:

$$P(\textbf{repeated}) = \frac{n(\textbf{repeated})}{n(\textbf{any})} = \frac{923,520}{1,110,000} = 0.832 = 83.2\%$$

$$P(\textbf{no repeated}) = \frac{n(\textbf{no repeated})}{n(\textbf{any})} = \frac{186,480}{1,110,000} = 0.168 = 16.8\%$$

We notice

$$P(\textbf{repeated}) + P(\textbf{no repeated}) = 1$$

This true in general for complementary events...

Probability of the Complement of an Event

Formula for the Probability of the Complement of an Event

If S is a finite sample space and E is an event in S, then

$$\mathbf{P(E^c) = 1 - P(E)}$$

Since $S = E \cup E^c$, and $E \cap E^c = \emptyset$ we have

$$n(S) = n(E) + n(E^c) \quad \Leftrightarrow \quad n(E^c) = n(S) - n(E)$$

and

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

$$P(E^c) = \frac{n(E^c)}{n(S)} = \frac{n(S) - n(E)}{n(S)} = \frac{n(S)}{n(S)} - \frac{n(E)}{n(S)} = 1 - P(E)$$

The Inclusion/Exclusion Rule

Theorem: If A and B are any finite sets, then

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

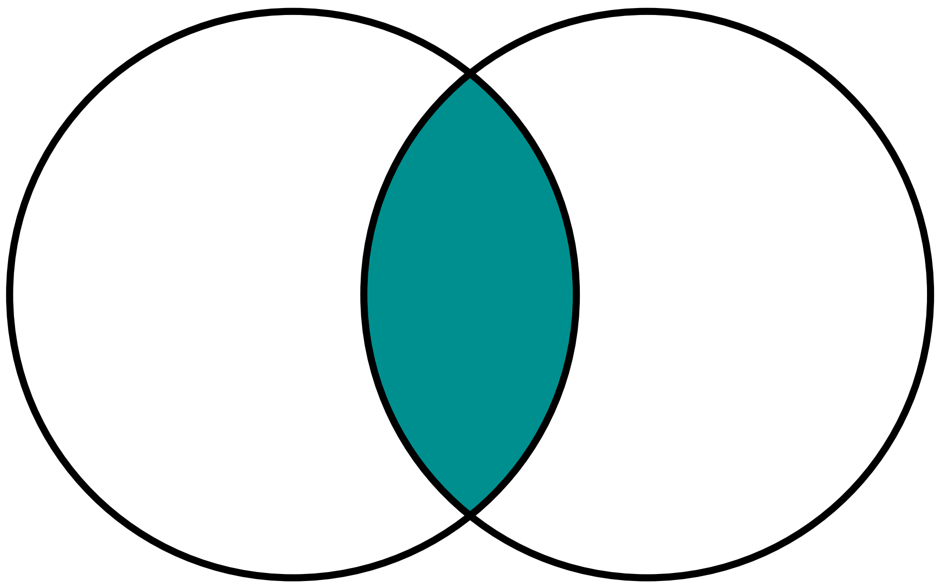


Figure: If we count the elements in A and add the elements in B, then the elements in the intersection $(A \cap B)$ are counted twice. The statement in the theorem subtracts one instance of the elements in the intersection, making the count correct.

Inclusion/Exclusion for Three Sets

Given the Inclusion/Exclusion rule for two sets, we can find rules for more sets: Let A, B, and C be any sets

$$\begin{aligned}
\mathbf{n}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{C}) &= n(A) + n(B \cup C) - n(A \cap (B \cup C)) \\
&= n(A) + n(B) + n(C) - n(B \cap C) - n(A \cap (B \cup C)) \\
&= n(A) + n(B) + n(C) - n(B \cap C) - n((A \cap B) \cup (A \cap C)) \\
&= n(A) + n(B) + n(C) - n(B \cap C) \\
&\quad - (n(A \cap B) + n(A \cap C) - n((A \cap B) \cap (A \cap C))) \\
&= \mathbf{n}(\mathbf{A}) + \mathbf{n}(\mathbf{B}) + \mathbf{n}(\mathbf{C}) - \mathbf{n}(\mathbf{B} \cap \mathbf{C}) \\
&\quad - \mathbf{n}(\mathbf{A} \cap \mathbf{B}) - \mathbf{n}(\mathbf{A} \cap \mathbf{C}) + \mathbf{n}(\mathbf{A} \cap \mathbf{B} \cap \mathbf{C})
\end{aligned}$$

See exercise ***Epp-6.3.36*** for the general inclusion/exclusion rule for n sets.

Example: Knowledge of Computer Languages

Problem: 50 students replied to a survey of what computer programming languages they knew:

$$
\begin{aligned}
A &= \{\textbf{Students that know Java}\}, & n(A) = 30 \\
B &= \{\textbf{Students that know Fortran}\}, & n(B) = 18 \\
C &= \{\textbf{Students that know C}\}, & n(C) = 26
\end{aligned}
$$

Further the survey reveals

$n(A \cap B) = 9, \quad n(A \cap C) = 16, \quad n(B \cap C) = 8, \quad n(A \cup B \cup C) = 47$

Using the difference rule we find that the number of students that do not know any of the 3 languages:

$$n(U) - n(A \cup B \cup C) = 50 - 47 = 3$$

Here U, our "universe," is the set of all students who replied to the study.

Example: Knowledge of Computer Languages

$$n(A) = 30,\ n(B) = 18,\ n(C) = 26,\ n(A \cup B \cup C) = 47,$$

$$n(A \cap B) = 9,\ n(A \cap C) = 16,\ n(B \cap C) = 8$$

Using our derived inclusion/exclusion formula for three set we find that the number of student that know all three languages are

$$\begin{aligned} n(A \cup B \cup C) &= n(A) + n(B) + n(C) - n(B \cap C) \\ &\quad - n(A \cap B) - n(A \cap C) + n(A \cap B \cap C) \\ 47 &= 30 + 18 + 26 - 9 - 16 - 8 + n(A \cap B \cap C) \end{aligned}$$

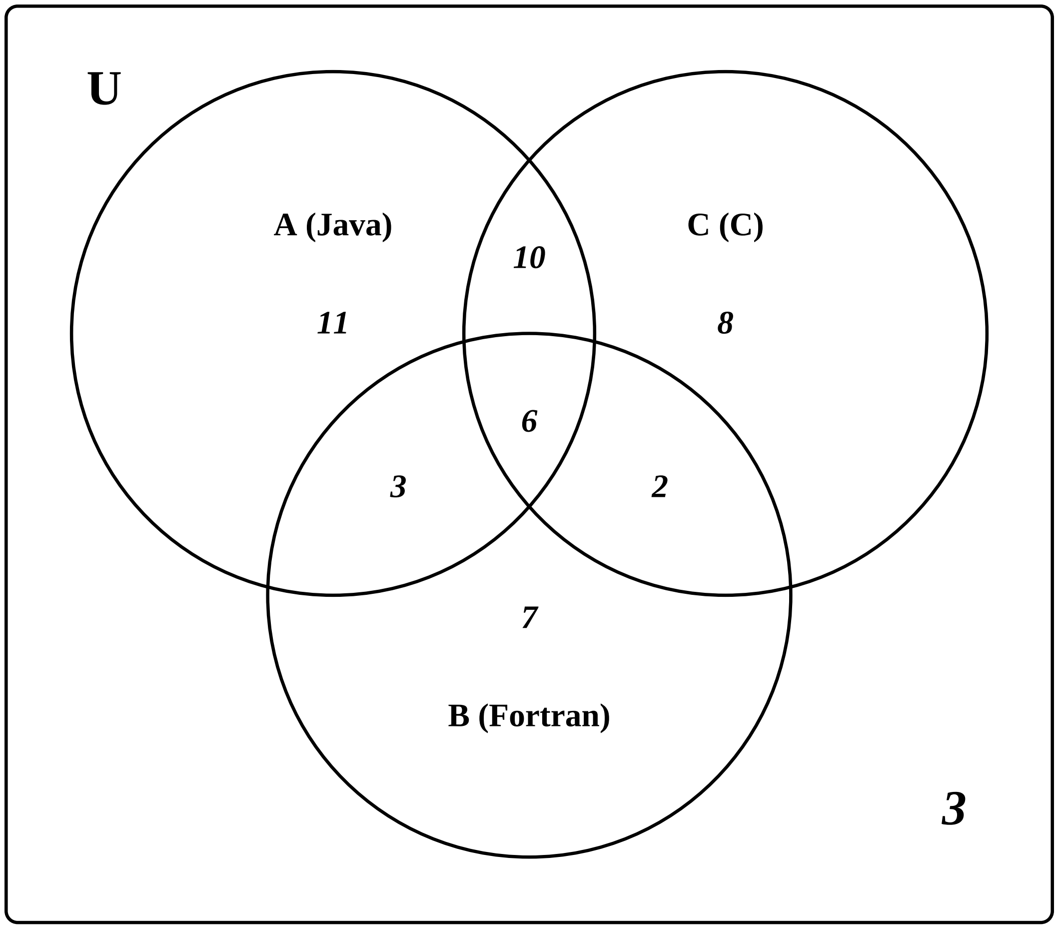
This gives us $n(A \cap B \cap C) = 6$.

We now have a complete picture...

Example: Knowledge of Computer Languages

$$n(A) = 30,\ n(B) = 18,\ n(C) = 26,\ n(A \cup B \cup C) = 47,$$

$$n(A \cap B) = 9,\ n(A \cap C) = 16,\ n(B \cap C) = 8,\ n(A \cap B \cap C) = 6$$



(Epp-v3.0)

Epp-6.3.4, Epp-6.3.6, Epp-6.3.11, Epp-6.3.18, Epp-6.3.25, Epp-6.3.26, Epp-6.3.28

Write down the inclusion/exclusion principle for 4 sets (hint: ***Epp-6.3.36***)

(Epp-v2.0)

Epp-6.3.4, Epp-6.3.6, Epp-6.3.11, Epp-6.3.18, —, Epp-6.3.23, Epp-6.3.25

Write down the inclusion/exclusion principle for 4 sets (hint: ***Epp-6.3.33***)