2.2 Combinatorial analysis

Department of Electrical and Computer Engineering
McGill University, Montréal, CANADA

September 2006

What is combinatorial analysis?

- Part of mathematics dealing with the study and development of systematic methods for counting.
- Find applications in many areas of sciences and engineering: probability and statistics, information theory, data compression, genetics, etc.
- The calculation of probabilities often amounts to counting the number of elements in various sets.
- Combinatorial techniques will be of great help in the solution of these problems.

2.2.1 Basic counting techniques

Generalized counting principle

Definition:

- A r-tuple is an ordered list (or vector) of elements, of the form $(x_1, x_2, ..., x_r)$, or simply $x_1 x_2 ... x_r$
- Two r-tuples are equal ($=$) if and only if each of the corresponding elements are identical.

Theorem 2.4: Let A be a set of r-tuples, $\{x_1\, x_2 \dots x_r\}$, such that there are:

- Firstly: n_1 different ways in which to chose x_1,
- Secondly: n_2 different ways in which to chose x_2,
- ...
- Finally: n_r different ways in which to chose x_r.

Then A contains $N(A) = n_1\, n_2 \dots n_r$ different r-tuples.

Note:

- The theorem specifies only the number of possible choices available at each step
- the specific choices in the rth step may depend on previous choices, but not their number n_r.

Corollary: Suppose the sets $A_1, A_2, ..., A_r$ contain $n_1, n_2, ..., n_r$ elements, respectively. Then the product set

$$A_1 \times A_2 \times ... \times A_r = \{(a_1, a_2, ..., a_r) : a_i \in A_i\} \tag{1}$$

contains $n_1 n_2 ... n_r$ elements.

Example

- In Quebec, license plate numbers are made up of 3 letters followed by 3 digits, that is $l_1 l_2 l_3 d_1 d_2 d_3$ where l_i is any one of 26 possible letters from a to z, and d_i is any one of the possible digits from 0 to 9.
- Thus there are, in principle,

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 26^3 \times 10^3 = 17,576,000 \quad (2)$$

different license plate numbers.

Number of subsets

Theorem 2.5: A set S containing n elements has 2^n different subsets, or equivalently, its power set $\mathcal{P}_S$ contains 2^n elements.

Proof: ...

Tree diagrams

- Useful when counting principle does not apply directly.
- For example, when the number of ways of selecting a second element depends on the choice made for the first element, and so on.
- Tree diagram provides systematic identification of all possibilities.

Example

- In a certain binary coding scheme, individual pieces of information are represented by specific sequences of 0s and 1s, called codewords.
- List all possible codewords that terminate upon the occurrence of symbol 0 or a maximum of 3 bits, whatever comes first?

2.2.2 Permutations

Permutations

Definition: An ordered arrangement of r elements taken without replacement from a set A containing n elements ($0 < r \leq n$) is called an r-element permutation of A.

The number of such permutations is denoted $P(n, r)$.

- For example, the possible 2-element permutations from $A = \{a, b, c\}$ are: ab, ac, ba, bc, ca, cb The number of these permutations is $P(3, 2) = 6$.
- Repetitions are not allowed in a permutation (once a has been selected, the remaining choices are b or c).
- A permutation is an ordered arrangement of r elements, i.e. an r-tuple. Thus the order does matter: $ab \neq ba$

Factorial notation

- For any positive integer n, we define

$$n! = n(n-1)(n-2)...1 \tag{3}$$

  It is also convenient to define $0! = 1$.

- Alternatively, factorials may be defined (and computed) recursively as $n! = n\,(n-1)!$, with initial condition $0! = 1$.
- Factorials grow surprisingly fast: $10! = 3628800$, $20! \approx 2.4329 \times 10^{18}$, etc.
- For large values of n, may use Stirling's approximation:

$$n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n} \tag{4}$$

Number of permutations

Theorem 2.6: The number of r-element permutations from a set A containing n elements is given by the product

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

Theorem 2.7: The number of distinguishable permutations of n objects of k different types, where n_1 are alike, n_2 are alike, ..., n_k are alike ($n_1 + n_2 + ... + n_k = n$) is given by

$$\frac{n!}{n_1!n_2!...n_k!} \tag{5}$$

Example

How many different "words" can we form:

- (a) with the 4 letters P H I L?
- (b) with the 6 letters P H I L I P?

2.2.3 Combinations

Combinations

Definition: An *unordered* arrangement of r objects taken without replacement from a set A containing n elements ($0 < r \leq n$) is called an r-element combination of A.

The number of such combinations is denoted $C(n, r)$.

- For example, the possible 2-element combinations from $A = \{a, b, c\}$ are: ab, ac, bc The number of such combinations is $C(3, 2) = 3$.
- As for permutations, repetitions are not allowed.
- However, order does not matter: $ab = ba$
- Conceptually, an r-element combination of A is the same as an r-element subset of A.

Number of combinations

Theorem 2.8: The number of r-element combinations of a set A containing n elements, is given by

$$C(n,r) = \frac{n!}{(n-r)!\,r!} \tag{6}$$

Proof: All the r-element permutations of A can be obtained by first selecting an r-element combination and then permuting the r selected elements. Therefore:

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

Corollary: A set S containing n elements has $C(n,r)$ different subsets of size r.

Binomial coefficients

Definition: For any integers r and n, with $0 \le r \le n$,

$$\binom{n}{r} \triangleq \frac{n!}{r!(n-r)!} = C(n,r) \tag{7}$$

The expression $\binom{n}{r}$ (read "n choose r") is also called binomial coefficient.

Theorem 2.9:

$$\binom{n}{0} = \binom{n}{n} = 1 \tag{8}$$

$$\binom{n}{r} = \binom{n}{n-r} \tag{9}$$

$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1} \tag{10}$$

Example: 6/49 lottery

In a 6/49 lottery, players pick 6 different integers between 1 and 49, without repetition, the order of the selection being irrelevant. The lottery commission then selects 6 winning numbers in the same manner.

- A player wins the first prize if his/her selection matches the 6 winning numbers.
- The player wins the second prize if exactly 5 of his/her chosen numbers match the winning selection.

How many different winning combinations are there?

2.2.4 Sampling problems

Introduction

- Many counting problems can be viewed as sampling problems, in which objects are selected from a finite set.
- We define four types of sampling problems:

Type	Replacement	Ordering
1	with replacement	with ordering
2	without replacement	with ordering
3	without replacement	without ordering
4	with replacement	without ordering

- In each case, a selection of r objects is made from a set A initially containing n distinct objects.
- We provide a general counting formula for each case.

1. *Sampling with replacement and with ordering:*
   - After selecting an object from A and noting its identity in an ordered list, the object is put back into A.
   - The number of distinct ordered lists is $N_1(n,r) = n^r$
2. *Sampling without replacement, with ordering:*
   - After selecting an object from A and noting its identity in an ordered list, the object is discarded A.
   - The number of distinct lists is equal to the number of r-element permutations from the set A:

$$N_2(n,r) = P(n,r) = \frac{n!}{(n-r)!} \tag{11}$$

3. Sampling without replacement, without ordering

- After selecting an object from A and noting its identity in a non-ordered list, the object is discarded.
- The number of distinct lists is equal to the number of r-element combinations from set A:

$$N_3(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{12}$$

4. With replacement, without ordering

- After selecting an object from A and noting its identity in a non-ordered list, the object is put back in A.
- In order to count the number of possibilities, we need to specify the way in which the observations are recorded.

- The standard approach consists in listing for each object how many times it is selected.
- For example, suppose $n = 6$ and $r = 5$. A possible observation is then $(3, 0, 0, 1, 0, 1)$, which can also be represented as

$$xxx|\,|\,|x|\,|x$$

- The number of distinct possible observations (or lists) is equal to the number of distinguishable permutations of $n + r - 1$ objects of two different types, of which r are alike (the x's) and $n - 1$ are alike (the $|$'s).
- Therefore (Theorem 2.7), we have

$$N_4(n, r) = \frac{(n + r - 1)!}{r!(n - 1)!} \tag{13}$$

Useful expansions

Theorem 2.10 (binomial): For any integer $n \geq 0$,

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$$

Theorem 2.11 (multinomial): For any integer $n \geq 0$,

$$(x_1 + x_2 + \ldots + x_k)^n = \sum_{n_1+n_2+\ldots+n_k=n} \frac{n!}{n_1!\, n_2! \ldots n_k!} x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}$$