Data Analysis
数据分析

张凡

新东方国际教育

zhangfan@xdf.cn

GRE 冲分班数学

2022 年 6 月 9 日

Counting Methods

Presentation Overview for Counting Methods

1. Counting Methods
   - Lists
   - Sets
   - Addition Principle
   - Multiplication Principle
   - Permutations
   - K-Permutation
   - Combinations
2. Probability
3. Descriptive Data Analysis (Exploratory Data Analysis)
4. Random Variables and Probability Distribution

Lists

Lists

列表：有顺序可重复

定义

- In a list, the members are ordered—that is, rearranging the members of a list makes it a different list.
- Elements can be repeated in a list and the repetitions matter.

例

1, 2, 3 and 2, 3, 1 and 1, 1, 2, 3 are all different list.

Sets

Sets

集合：无顺序不可重复

定义

- In a set, repetitions are not counted as additional elements.
- the order of the elements does not matter.

例

1, 2, 3 and 2, 3, 1 and 1, 1, 2, 3 are the same set.

Subsets

空集是任意集合的子集

定义

If A and B are sets and all of the members of A are also members of B, then A is a subset of B. By convention, $\emptyset$ is a subset of every set.

例

2, 8 is a subset of 0, 2, 4, 6, 8.

定理

If a set has n elements, then the number of subset of the given set is 2^n

We will prove the above theorem by induction when we talk about Combinations.

The Operations of Sets

交集并集

定义

If S and T are sets, then the intersection of S and T is the set of all elements that are in both S and T and is denoted by $S \cap T$

例

2, 8 is a subset of 0, 2, 4, 6, 8.

定理

If a set has n elements, then the number of subset of the given set is 2^n

We will prove the above theorem by induction when we talk about Combinations.

Addition Principle

Addition Principle
加法原则

定义

- Suppose there are two choices to be made sequentially.
- Suppose also that there are k different possibilities for the first choice and m different possibilities for the second choice for each possibilities of the first choice.

Then, under those conditions, there are $k \cdot m$ different possibilities for the pair of choices.

例

For a 6 digit Bank PIN, there could be 10^6 different passwords.

Multiplication Principle

Multiplication Principle
乘法原则

定义

- Suppose there are two choices to be made sequentially.
- Suppose also that there are k different possibilities for the first choice and m different possibilities for the second choice for each possibilities of the first choice.

Then, under those conditions, there are km different possibilities for the pair of choices.

例

For a 6 digit Bank PIN, there could be 10^6 different passwords.

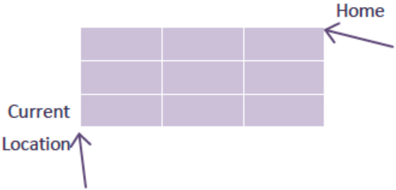
The Decision Tree
决策树

What if a 6 digit Bank PIN can not start or end with 0?

Can you draw the decision tree?

Have a try!

A man walks to his home from his current location on the rectangular grid shown. If he may choose to walk north or east at any corner, but may never moves south or west. How many different paths can the man take to get home?



Answer 2^3

GMAT-OG In a meeting of 3 representatives from each of 6 different companies, each person shook hands with every person not from his or her own company. If the representatives did not shake hands with people from their own company, how many handshakes took place?

- (A) 45
- (B) 135
- (C) 144
- (D) 270
- (E) 288

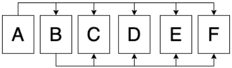


图: Handshakes for Company A should be $3 \times 15 = 45$;Handshakes for Company B should be $3 \times 12 = 36$.

$3 \times 15 + 3 \times 12 + 3 \times 9 + 3 \times 6 + 3 \times 3 = 135$

Answer **B** 135

Permutations

A Vanilla Question
排排坐的排法

How many kinds of different lists can be constructed 3 Students — student A, student B and student C.

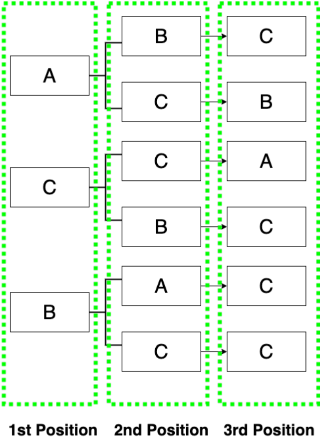


图: 6 Different Lists can be formed from 3 different students

A Vanilla Question
排排坐的排法

What about the 4 different students?

5 different students

6 different students

$\vdots$

$$4 \times 3 \times 2 \times 1 = 24$$

Permutation
全排

定义

Suppose n objects are to be ordered from 1st to nth, then the number of ways the objects can be ordered is $n(n-1)(n-2)(n-3)\ldots 1 = n!$

og-p463-1.7.16 Suppose that 10 students are going on a bus trip, and each of the students will be assigned to one of the 10 available seats. What is the number of possible different seating arrangements of the students on the bus?

$10! = 10 \times 9 \times 8 \times 7 \ldots 1 = 3,628,800$

Answer **3,628,800**

Have a try!

A gardener wishes to plant 5 bushes in straight row. Each bush has flowers of a different solid color (white, yellow, pink, red, and purple). How many ways can the bushes be arranged so that middle bush is the one with red flowers?

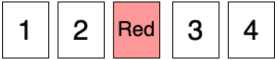


图: There are 4 vacancy.

$$^{5}P_4 = \frac{5!}{(5-4)!} = 5! = 120$$

Answer 120

A Real QR Problem!

A father purchased theater tickets for 6 adjacent seats in the same row of seats for himself, his wife, and their 4 children. How many seating arrangements are possible if the father and mother sit in the 2 middle seats?

- ◯ 24
- ◯ 36
- ◯ 48
- ◯ 120
- ◯ 240

图: 8-Sec3-8

$$Children's\ \ Permutation = 4! = 24$$
$$Parents'\ \ Permutation = 2! = 2$$
$$PossibleArrangements =$$
$$(Children's\ \ Permutation) \cdot (Parents'\ \ Permutation) = 4! \cdot 2! = 48$$

Answer **C** 48

K-Permutation

Suppose that there are 8 students and 5 students are selected going on a bus trip, and each of the students will be assigned to one of the 5 available seats. What is the number of possible different seating arrangements of the students on the bus?

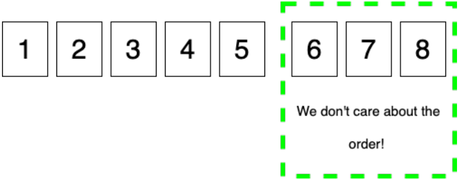


图: The order about the final 3 students is irrelevant.

$$\frac{8!}{3!} = \frac{8 \times 7 \times 6 \cdots \times 1}{3 \times 2 \times 1} = 6,720 \tag{1}$$

K-Permutation

K-全排

定义

Suppose that k objects will be selected from a set of n objects, where , and the k objects will be placed in order from 1st to kth, the number of ways to select and order k objects from a set of n objects is

$$^{n}P_k = \frac{n!}{(n-k)!}$$

Demo: K-Permutation
难题演示

How many distinguished way to arrange the word “ORDER”if there are at least two letters between two “R”s?

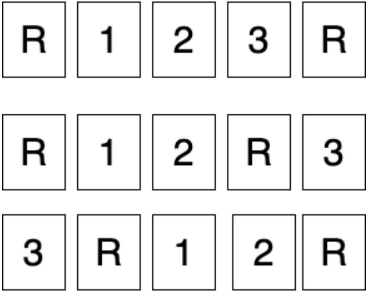


图: Three possible arrangements with blanks to fill

$$
\begin{aligned}
&{}^3P_3 + {}^3P_2 \cdot 2 \\
&= 3! + \frac{3!}{2!} \cdot 2 \\
&= 18
\end{aligned}
$$

Combinations

A Vanilla Question
选部分人不讲顺序的排法

Suppose that there are 8 students and 5 students are selected going on a bus trip. What is the number of possible combinations of the students on the for the bus trip?

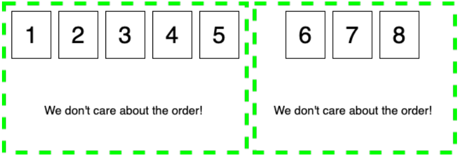


图: The order about the 3 students abandoned and the 5 students chosen is irrelevant.

$$\frac{8!}{5!3!} = \frac{8 \times 7 \times 6 \cdots \times 1}{(5 \times 4 \times 3 \times 2 \times 1)(3 \times 2 \times 1)} = 56 \tag{2}$$

Combinations

n 选 k，不讲顺序

定义

Suppose that k objects will be chosen from a set of n objects, where , but that the k objects will not be put in order. The the number of combinations of n objects taken k at a time and is

$$^{n}C_k = \frac{n!}{k!(n-k)!}$$

证明.

$$(number\ \ of\ \ ways\ \ to\ \ select\ \ without\ \ order) \times (number\ \ of\ \ ways\ \ to\ \ order)$$
$$=(number\ \ of\ \ ways\ \ to\ \ select\ \ with\ \ order)$$
$$(number\ \ of\ \ ways\ \ to\ \ select\ \ without\ \ order) = \frac{(number\ \ of\ \ ways\ \ to\ \ select\ \ with\ \ order)}{(number\ \ of\ \ ways\ \ to\ \ order)}$$

Have a try!

Suppose you want to select a 3-person committee from a group of 9 students. How many ways are there to do this?

$$^{9}C_3 = \frac{9!}{3!6!} = 84$$

Answer 84

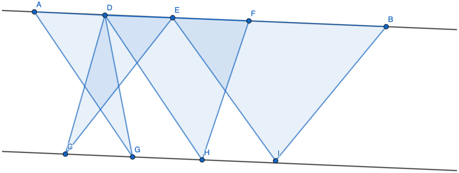
Rules of Combination

快速计算规则

- ${}^{n}C_0 = 1$
- ${}^{n}C_1 = n$
- ${}^{n}C_r = {}^{n}C_{n-r}$

Have a try!

There are five points on line l_1 and four points on l_2. If l_1 and l_2 are parallel, how many different triangles can be constructed based on these nine points?



$$^2C_5 \cdot ^1C_4 + ^1C_5 \cdot ^2C_4 = 70$$

Answer 70

Demo

反其道而行之：用总数去减

GMAT-鸡精 In a certain class of 20 students with different initials of surname, the roster is arranged by the order of initial of surnames. If three students are selected to join a seminar, how many different ways to select them whose initial of surnames are not near to each other?

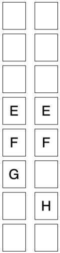


图: Two combinations that are not allowed.

Case 1: Three consecutive initials

$Combination_1 = 18$

Case 2: Only two consecutive initials

Case 2-1: the two consecutive initials occupies the start or the end of the roster

$Combination_{2-1} = 2 \times 17 = 34$

Case 2-2: the two consecutive initials occupies neither the start nor the end of the roster

$Combination_{2-2} = 17 \times 16 = 186$

$\therefore {}^3C_{20} - Combination_1 - Combination_{2-1} - Combination_{2-2} = 1140 - 18 - 34 - 186 = \mathbf{902}$

A Real QR Problem!

A store is shipping 7 items to a single address. The items can be packed into 1, 2, or 3 different containers for shipping. If a container is used for shipping, it must be packed with at least 2 items. Which of the following statements about the different combinations of packing must be true?

Indicate all such statements.

- ☐ There will be exactly 3 items in one container.
- ☐ There will be at most 3 items in each container.
- ☐ There will be an odd number of items in one container.

图: 9-Sec1-12

Answer **B**

Probability

Presentation Overview for Probability

1. Counting Methods

2. Probability
   - Sets
   - Events
   - Mutually Inclusive Events
   - Independent Events

3. Descriptive Data Analysis (Exploratory Data Analysis)

4. Random Variables and Probability Distribution

Sets

Events

Mutually Inclusive Events

Independent Events

Descriptive Data Analysis (Exploratory Data Analysis)

Presentation Overview for Descriptive Data Analysis (Exploratory Data Analysis)

1. Counting Methods
2. Probability
3. Descriptive Data Analysis (Exploratory Data Analysis)
    - Data Type
    - Methods for Presenting Data
    - Measures of Central Tendency
    - Measures of Position
    - Measures of Dispersion
4. Random Variables and Probability Distribution

To Begin With

QR Mathematical Convention 5

When graphical data presentations, such as bar graphs and line graphs, are shown with scales, you should read, estimate, or compare quantities by sight or by measurement, according to the corresponding scales.

QR Mathematical Convention 4

Scales, grid lines, dots, bars, shadings, solid and dashed lines, legends, etc., are used on graphs to indicate the data. Sometimes scales that do not begin at 0 are used, and sometimes broken scales are used..

看图要点

- Carefully understand title, the labels for x-axis and y-axis, and legends. 看标题，XY 轴，图例找到数据的含义
- Carefully examine the scales and unit. 看单位
- Carefully examine the the starting point. 看起点

How to read a graph
读图要点

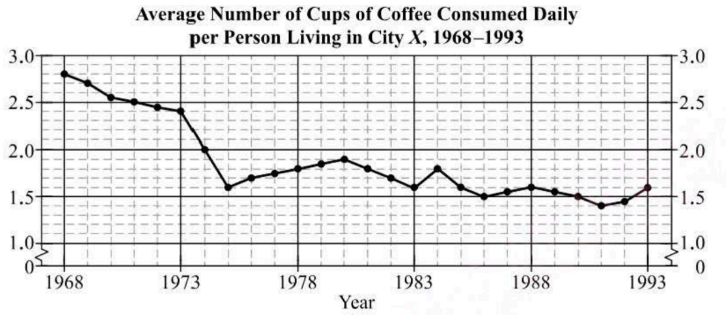


图: The graph shows The trends of The average number of cups of coffee consumed daily per person living in city X by year. Y suggest number of cups. X suggest the year. Y use a broken scales.

Data Type

Methods for Presenting Data

Measures of Central Tendency

Measures of Position

Measures of Dispersion

Random Variables and Probability Distribution

Presentation Overview for Random Variables and Probability Distribution

Random Variables

Probability Distribution

Normal Distribution

1 Min Break

Questions? Comments?