International Journal of Pure and Applied Mathematics
Volume 78 No. 8 2012, 1099-1109
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu

ijpam.eu

WHAT'S IN A BARCODES?: DUPLICATED COMBINATIONS

Yutaka Nishiyama

Department of Business Information
Faculty of Information Management
Osaka University of Economics
2, Osumi Higashiyodogawa Osaka, 533-8533, JAPAN

Abstract: This article explains how barcodes are read mathematically. One numeral is constructed from 7 modules in a barcode. Firstly, a counting method is presented through which the number of different characters which can be expressed is obtained. After touching upon formulae for permutations and combinations, duplicated combinations are expressed explicitly. The use of a modulus of 10 is mentioned with regard to the calculation of check digits.

AMS Subject Classification: 00A09, 90C08, 97A20
Key Words: barcode, tree diagram, permutation, combination, duplicated permutation, duplicated combination, check digits, modulo 10

1. How are Barcodes Read?

I'm sure you know that commercial goods usually have a barcode marked on them somewhere. These barcodes include information related to maker and product codes, and they can in an instant be read-in and so reduce the burden on cashiers. Barcodes are also essential when operating the Point-of-Sale (POS) system (see Nishiyama, 1991) [1].

We all know that barcodes are composed of a pattern of black and white stripes, but do you know what they mean? The truth is they have a close

Received: March 5, 2012

relationship with the combinations and permutations studied in counting patterns. Mathematics therefore has application to the manufacture of goods and its related technologies. Standard barcodes are composed of a total of 13 digits. From the leftmost digit, the first 2 represent the country code (Japan is 49), the next 5 are the maker code, the following 5 are the product code, and the last 1 is a check digit. In the example in Figure 1, the country code is 49, the maker code is 01306, the product code is 04282, and the check digit is 3 (see Japanese Standards Association, 1985) [2].

Figure 1: An example barcode

Let's try the following investigation to discover the relationship between a 13 digit barcode and its black and white striped pattern. Firstly, the width of the lines varies, but how many lines are there in total? Careful counting reveals that there are 30. Assuming that these 30 lines represent 13 digits, there are 2.3 lines for each digit. However, doesn't such an inexact number raise a few questions? Examining the lines carefully, it can be seen that there is a sequence of long thin lines, with 2 at the left edge, 2 in the center, and 2 at the right hand edge. These show where the barcode begins, its midpoint, and where it ends, and do not represent actual numbers. Excluding these long thin lines leaves $30 - 2 \times 3 = 24$ lines.

The first digit, 4, of the country code 49 is not expressed within the barcode. Instead this 4 is printed just to the side. Compiling these facts, 24 lines represent 12 numerical digits, and since $24 \div 12 = 2$, so two lines represent each digit.

Next, let's look at the thickness of the lines. How many times wider is the thickest line compared to the thinnest line? Careful observation reveals that the widths differ by a factor of 4. It is reasonable to suppose that a pair of two such thick or thin lines represents a single digit. The number underneath the barcode is provided so that people can read it and confirm the barcode. Machines themselves do not read this number.

By reading-in the maker code and product code, computers can look up and display a corresponding price from a database. There are also some barcodes which include the price directly.

When a barcode is enlarged with a magnifying glass, the result resembles the image shown in Figure 2. For example, the numeral 3 has a striped pattern with 1 white, 4 black, 1 white and 1 black lines. The numeral 4 has a striped pattern with 2 white, 3 black, 1 white, and 1 black lines. One numeral occupies 7 units of width (known as modules), which are arranged in stripes in a white, black, white, black order. Denoting their widths by a, b, c and d, and relating these mathematically, the problem is to find integer solutions satisfying $a+b+c+d=7$ under the condition that $1 \leq a, b, c, d \leq 4$.

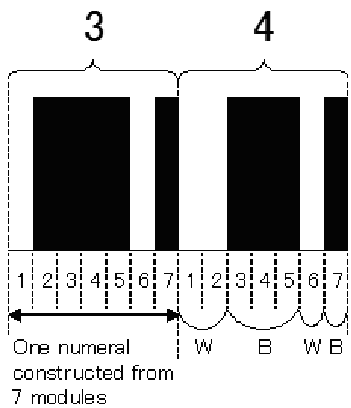

Figure 2: The structure of the symbols

2. Enumeration Method

I mentioned that 7 modules are used to represent one numeral, so let's try and think about how many different characters can actually be expressed with 7 modules.

This answer can be obtained effortlessly by using the formula for 'duplicated combinations' which is presented as a 'compiled application' and covered under the topic of combinations and permutations in the chapter on probability and case counting in high-school mathematics A. One cannot, however, guarantee to remember this formula, so after revising it, let's try thinking about the problem from scratch as if we had absolutely no knowledge of it.

What should you do when you have completely forgotten a formula? In such a situation there's nothing for it but to write out all the possible cases that can occur. Figure 3 shows a tree diagram produced by counting them up. This appears simple, but unless the system is written out in an orderly way, oversights will occur. The number of modules, 7, acts as a bound when writing out the cases. One can pretty much give up on trying to write out how many patterns could be produced with 8 or 9 modules of white, black, white, black stripes.

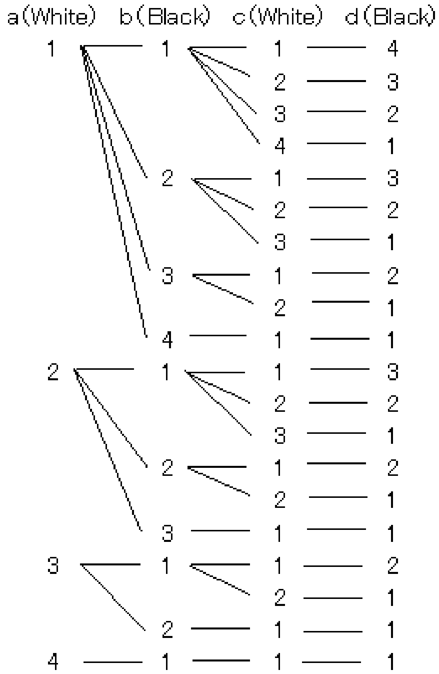

Figure 3: Tree diagram

Before we move on and discuss duplicated combinations, let's revise combinations and permutations.

[Question 1] How many ways can 3 people be selected from a group of 10 people and lined up?

There are 10 ways of choosing the first person, then the next is chosen from a group reduced by 1 person, thus containing 9 people, and the next from a group of 8, so there are $10 \times 9 \times 8 = 720$ different ways. If this is expressed using factorials, it may be then be rewritten as a permutation equation.

$$10 \times 9 \times 8 = \frac{10 \times 9 \times \cdots \times 1}{7 \times 6 \times \cdots \times 1} = \frac{10!}{7!} = \frac{10!}{(10-3)!} = {}_{10}P_3$$

In general, the total number of ways of selecting and lining up r things from a collection of n is,

$${}_nP_r = \frac{n!}{(n-r)!} \quad .$$

[Question 2] How many ways can a combination of 3 people be selected from a group of 10 people?

The result of Question 1 can be used to answer this second question since it reveals the number of permutations by which 3 people can be selected *and lined up* from among 10 people. When the 3 people selected are placed in order, the result is a permutation, but when they are merely selected, the order is not important. The number of ways of ordering 3 people, 3!, is the number of duplicated combinations in each case, so the number of permutations ${}_{10}P_3$ divided by 3! is equal to the number of combinations, *i.e.*, there are $\frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$ different combinations. If this is expressed using factorials, it may then be written as a combinatorial equation, $\frac{{}_{10}P_3}{3!} = \frac{10!}{(10-3)!3!} = {}_{10}C_3$. In general, the total number of ways of extracting a combination of r things from among n things is

$${}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{(n-r)!r!}.$$

3. Duplicated Permutations and Duplicated Combinations

The discussion above covers some basic points learned in high-school mathematics, but combinations and permutations each have extensions known as duplicated combinations and duplicated permutations respectively, i.e., one may seek the number of combinations or permutations when duplicates are permitted.

While some topics are included in the syllabus as fundamentals and should be learned, others are summarized as extension problems. Since the barcodes we are discussing this time constitute a combinatorial problem with duplicates, allow me to state this issue in detail.

First let's think about duplicated permutations.

[Question 3] How many ways are there to select and line up 3 people from a group of 10 when duplicates are permitted?

If it's difficult to imagine duplicating people, it should be enough to suppose that there are 3 different draws and all the people may participate each time. The number of possible selections each time includes all 10 people, so there are $10 \times 10 \times 10 = 10^3 = 1000$ different ways. In general, there are a total of n^r ways of selecting and lining up r things from among n things.

[Question 4] How many ways can a combination of 3 people be selected from a group of 10 when duplicates are permitted?

As an example, suppose that the 3 people are chosen so that the 3rd person is selected twice, and the 8th person is selected once, then a route like that shown in Figure 4 is conceivable. Looking at it like this unifies the number of ways of selecting 3 people from among the 10 while permitting duplicates, with the route between S and G. There are $(10-1)+3 = 12$ routes between S and G, and the problem becomes one of deciding how to make the 3 steps according to the upwards pointing arrows, or 9 steps according to the right-facing arrow. There are ${}_{12}C_3 = {}_{12}C_9 = \dfrac{12 \times 11 \times 10}{3 \times 2 \times 1} = 220$ different ways.

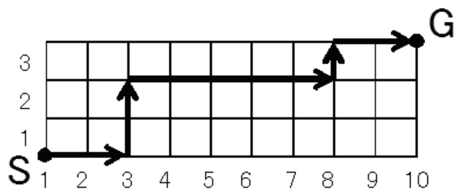

Figure 4: How many ways are there to make the route between S and G?

Pay attention to the positions of the numerals written in the bottom line in Figure 4. Since the numerals are on lines extended from a grid, the breadth between 1 and 10 is 9. Understanding the expression $(10-1)+3 = 12$ is particularly important, so let's consider an explanation from another perspective. Question 4 is equivalent to the problem of placing 3 indistinguishable balls into a set of boxes numbered from 1 to 10. Suppose that for the example in Figure 5, there are 2 balls in the '3' box, and 1 ball in the '8' box. Removing the box's outer frame reveals that there are 9 empty partitions. With 9 empty partitions and 3 balls, there are a total of 12 locations among which the locations of the 3 balls, or alternatively, the 9 empty partitions, must be chosen. This is a combinatorial problem solved by, ${}_{12}C_3 = {}_{12}C_9 = 220$. .

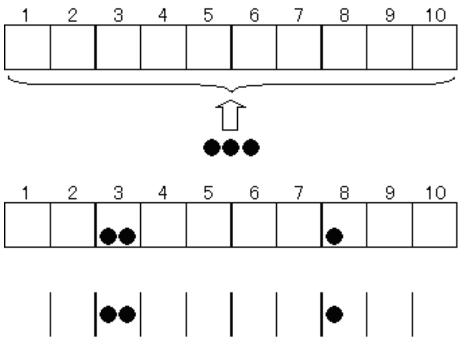

Figure 5: Duplicated combinations (alternative solution)

In general, when performing r selections from a group of n people, when duplications are permitted, the following number of ways are possible.

$$ {}_nH_r = {}_{n+r-1}C_r = \frac{(n+r-1)!}{(n-1)!r!} $$

The number of combinations including duplicates is sometimes written ${}_nH_r$.

All the necessary preparation is now complete. Let's now think about the combinations involved in barcodes again. Each numeral in a barcode is formed from 7 modules. One numeral is represented by a white + black + white + black pattern of lines. Denoting the widths of these four lines as a, b, c and d modules, the problem is to find the integer solutions of

$$ a + b + c + d = 7. $$

However, the width of each black and white line must be at least 1 module, so the maximum module width is constrained. The maximum module width is 4, which may expressed mathematically by seeking solutions constrained by

$$ 1 \leq a, b, c, d \leq 4. $$

Subtracting the minimum module width from each of a, b, c and d reveals that in the end the problem is to allocate the remaining $7 - (1 + 1 + 1 + 1) = 7 - 4 = 3$ modules. This is equivalent to the problem of finding the total number of ways of selecting a combination of 3 elements from the 4 alternatives a, b, c and d when duplications are permitted. From the formula for combinations with duplicates, the answer is

$$ {}_4H_3 = {}_{4+3-1}C_3 = {}_6C_3 = \frac{6!}{(6-3)!3!} = 20. $$

There is nothing wrong with looking at the construction of the symbol in Figure 2, and immediately applying the formula for duplicated combinations as shown above. However, many people might not remember the equation without making an error. Therefore, in order to find the solution without making a mistake, wouldn't it be good to count up the different possibilities according to a tree diagram, such as that shown in Figure 3, so as to obtain the solution reliably, without using the formula for duplicated combinations?

4. 40 Different Patterns are Possible

I explained that using 7 modules, characters can be expressed in 20 different ways, but the inversion of the white, black, white, and black pattern (which is a black, white, black, and white pattern), can also be used to represent characters in 20 different ways. In total, there are 40 different possible ways of representing characters. Since there are only 10 numerals between 0 and 9, this 4-fold excess may be thought unnecessary, but the leeway is used to prevent mistakes when barcodes are read.

The case when the total number of black modules is an odd number is known as 'odd parity', and the even case is known as 'even parity'. There are two types of pattern for the maker code (on the left side), *i.e.*, even or odd parity. Only even parity is used for the product code (on the right side). For example, the same number, 3, may be represented with white (1), black (4), white (1), and black (1), in which case there are a total of 5 black modules and the parity is odd, or it may be represented as white (1), black (1), white (4), and black lines (1) in which case there are a total of 2 black modules and the parity is even. 30 different patterns from JAN: Japanese Article Number code, are shown in Figure 6.

The maker code (left side) has a white, black, white, black pattern, and the product code has a black, white, black, white pattern. This is related to the

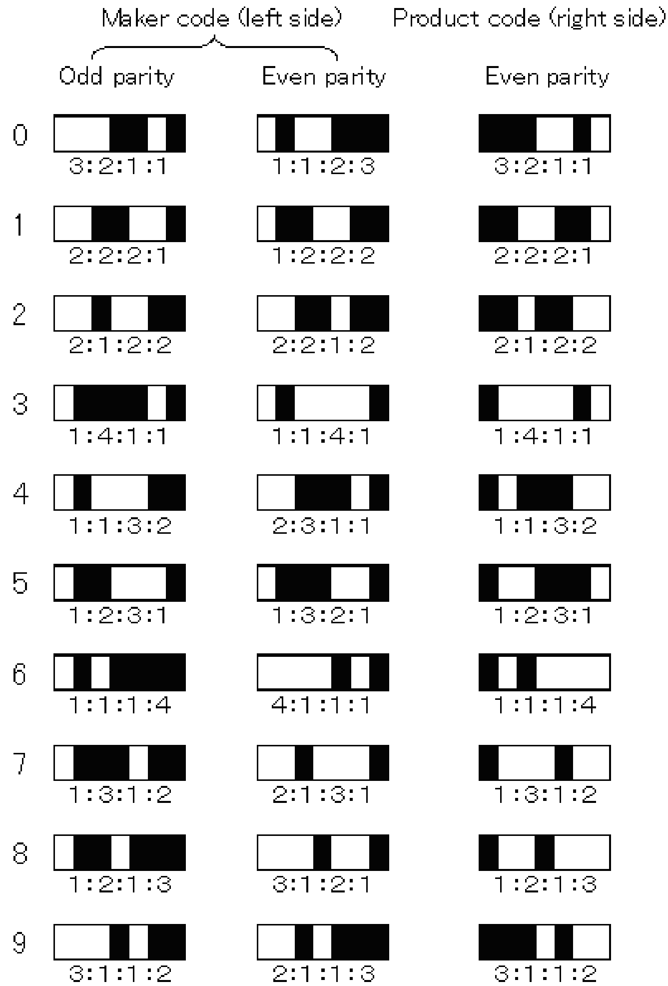

Figure 6: Table showing the correspondence between numerals and patterns

reading of the barcode by a scanner, and allows the barcode to be read from either side.

When the number of modules used to represent 1 character is 7, there are 20 possible combinations. If there are 6 modules, then there are 10 combinations. This can be written as a tree diagram, and it is not particularly time consuming to do so. However, when solving the problem of the number of ways characters can be represented when there are 8 modules, it is necessary to rely on the formula after all. The formula is not intended to make examination candidates suffer; it is intended to save the considerable amount of time that would otherwise be spent writing out the cases.

If you are liable to forget the formula, or can only remember it vaguely, it is enough to remember the process behind the introduction of the formulae for permutations, combinations, duplicated permutations and duplicated combinations, as demonstrated through Questions 1 to 4.

In the case of 8 modules the integer solutions of

$$a+b+c+d=8$$

should be found according to the constraints

$$1 \leq a, b, c, d \leq 5.$$

Since each of a, b, c and d must have at least 1 module each, the result is the number of duplicate combinations of the $8-1\times 4=4$ modules among a, b, c and d, *i.e.*, the number of ways is

$${}_4H_4 = {}_{4+4-1}C_4 = {}_7C_4 = \frac{7!}{(7-4)!4!} = 35.$$

Considering the black, white, black, white pattern in addition to the white, black, white, black pattern yields 70 alternatives.

4 901306 042823

The 13th digit of a barcode is a check digit. This is used to confirm that the 12 digits composed of the country code (2 digits), the maker code (5 digits), and the product code (5 digits) are read-in correctly, so it does not represent a numerical value.

The check utilizes a modulus of 10, and I would like to explain how it is calculated. From the 12 digits, the sum of all those in an even position is obtained and then multiplied by 3. The sum of all the digits in odd numbered positions is then added. The result is then subtracted from the smallest multiple of 10 which exceeds its value. The result is the check digit. Let's look at a concrete example, and calculate the specific value corresponding to Figure 1. The 13 digits of this number are 4901306042823, and the meaningful digits are 490130604282. The sum of the digits in even positions is $9+1+0+0+2+2=14$. Multiplying this number by 3 yields $14\times 3=42$. The sum of the digits in odd positions is $4+0+3+6+4+8=25$, and adding these together yields $42+25=67$. The smallest multiple of 10 larger than 67 is 70, and subtracting 67 from 70 yields $70 - 67 = 3$. The check digit is therefore 3. Owing to this check, the ratio of correctly read-in barcodes is kept high.

References

[1] Y. Nishiyama, Bakodo Shinboru [Barcode symbols], In *Saiensu no Kaori [The Scent of Science]*, Tokyo, Nihon Hyoronsha (1991), 1-9.

[2] Japanese Standards Association, *Barcode Symbols for Common Use as Product Codes in Japan* (1985).