CHAPTER 16 – IDENTIFICATION NUMBERS

Consider the UPC code on a can of RO★TEL tomatoes

The scanner is not working so the clerk enters the numbers by hand as

0 64144 28263 2

and this is invalid even though the product code for the mild version of this is 28263. What happened?

The UPC codes use a ***check digit*** to minimize scanning errors. A check digit is a digit included in a code to help detect errors.

For the UPC code $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}$, the check digit, a_{12}, is chosen so that S a multiple of 10 where

$$S = 3a_1 + a_2 + 3a_3 + a_4 + 3a_5 + a_6 + 3a_7 + a_8 + 3a_9 + a_{10} + 3a_{11} + a_{12}.$$

What is the check digit for the mild RO★TEL if the first eleven digits are 0 6 4 1 4 4 2 8 2 6 3?

When talking about check digits, modular arithmetic will be helpful.

Definition: Congruence Modulo m

Let a, b, and m be integers with $m \geq 2$. Then a is congruent to b modulo m, written

$$a \equiv b \bmod m$$

if $(a - b) \div m$ has a remainder of 0. This means $a - b$ is a multiple of m. One way to find a value for b is to find the remainder when a is divided by m.

Determine if the congruences below are true or false:

$25 \equiv 1 \bmod 6$

$100 \equiv 20 \bmod 10$

$52 \equiv 0 \bmod 13$

$75 \equiv 7 \bmod 5$

Find the following values:

(a) 34 mod 5 = ___________

(b) 78 mod 11 = ___________

(c) 13 mod 15 = ___________

(d) 12 mod 2 = ___________

Some types of errors when dealing with identification numbers are

- Replacing one digit with a different digit (single digit error)
- Transposing two adjacent digits (adjacent transposition error)
- Transposing two digits that are separated by another digit (jump transposition error)

Assume that the correct code was 5678 and provide an example of these errors:

Single digit error:

Adjacent Transposition Error:

Jump Transposition Error:

Note that some of the digits in the UPC code are multiplied by 3. Those digits had a ***weight*** of 3. Other codes use different weights.

A code $a_1a_2a_3a_4a_5$ uses the last digit as a check digit. The check digit is found using the formula

$$a_5 = (a_1 + 7a_2 + a_3 + 7a_4) \text{ mod } 10$$

(a) What is the check digit for the code 2374?

(b) Find the value of the missing digit x in the code $468x3$

(c) Will this code find an error if a single digit is entered incorrectly?

Let's look at an error in the first digit, a_1.

Correct Code: $a_1a_2a_3a_4$

Incorrect Code: $e_1a_2a_3a_4$

So the correct check digit is

$(a_1 + 7a_2 + a_3 + 7a_4)$ mod 10

and the incorrect check digit is

$(e_1 + 7a_2 + a_3 + 7a_4)$ mod 10

The error will NOT be caught if

$(a_1 + 7a_2 + a_3 + 7a_4) - (e_1 + 7a_2 + a_3 + 7a_4)$ is a multiple of 10.

This simplifies to

$a_1 - e_1$ is a multiple of 10,

Multiples of 10 are $\pm\{0,\ 10, 20, 30,\ \ldots\}$.

If $a_1 - e_1 = 0$, then $a_1 = e_1$ so we did not really make an error.

If $a_1 - e_1 = \pm 10$ (which can also be written as $|a_1 - e_1| = 10$), then a_1 and e_1 are digits that are separated by 10 units. The digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are never separated by more than 9 units, so this and all higher multiples of 10 are impossible.

Therefore, all single-digit errors in the first digit would be caught.

The same logic also applies to the third digit.

Therefore, we need to check the even-numbered positions.

Let's look at an error in the second digit, a_2.

Correct Code: $a_1a_2a_3a_4$

Incorrect Code: $a_1e_2a_3a_4$

So the correct check digit is

$(a_1 + 7a_2 + a_3 + 7a_4)$ mod 10

and the incorrect check digit is

$(a_1 + 7e_2 + a_3 + 7a_4)$ mod 10

The error will NOT be caught if

$(a_1 + 7a_2 + a_3 + 7a_4) - (a_1 + 7e_2 + a_3 + 7a_4)$ is a multiple of 10.

This simplifies to

$7a_2 - 7e_2$ is a multiple of 10,

which means $7(a_2 - e_2)$ is a multiple of 10,

which means, $a_2 - e_2$ is a multiple of $\frac{10}{7}$.

Multiples of $\frac{10}{7}$ are $\pm\left\{0, \frac{10}{7}, \frac{20}{7}, \frac{30}{7}, \frac{40}{7}, \frac{50}{7}, \frac{60}{7}, 10, \frac{80}{7}, \ldots\right\}$.

If $a_2 - e_2 = 0$, then $a_2 = e_2$ so we did not really make an error.

If $a_2 - e_2 = \pm\frac{10}{7}$ (which can also be written as $|a_1 - e_1| = \frac{10}{7}$), then a_1 and e_1 are digits that are separated by $\frac{10}{7}$. The difference between digits is always an integer, so this and all other non-integer multiples of $\frac{10}{7}$ are impossible.

If $|a_2 - e_2| = 10$, then then a_1 and e_1 are digits that are separated by 10 units. The digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are never separated by more than 9 units, so this and all higher multiples of $\frac{10}{7}$ are impossible.

Therefore, all single-digit errors in the second digit would be caught.

The same logic also applies to the fourth digit.

We have now checked all four digits for single-digit errors and have not found any that would not be detected. Therefore, this scheme detects all single-digits errors.

(d) Will this code find all adjacent transposition errors?

Let's look at an adjacent transposition of the first two digits, a_1 and a_2.

Correct Code: $a_1a_2a_3a_4$

Incorrect Code: $a_2a_1a_3a_4$

So the correct check digit is

$(a_1 + 7a_2 + a_3 + 7a_4)$ mod 10

and the incorrect check digit is

$(a_2 + 7a_1 + a_3 + 7a_4)$ mod 10

The error will NOT be caught if

$(a_1 + 7a_2 + a_3 + 7a_4) - (a_2 + 7a_1 + a_3 + 7a_4)$ is a multiple of 10.

This simplifies to

$a_1 + 7a_2 - a_2 - 7a_1$ is a multiple of 10,

which means $6a_2 - 6a_1 = 6(a_2 - a_1)$ is a multiple of 10,

which means, $a_2 - a_1$ is a multiple of $\frac{10}{6} = \frac{5}{3}$.

Multiples of $\frac{5}{3}$ are $\pm\left\{0, \frac{5}{3}, \frac{10}{3}, 5, \frac{20}{3}, \frac{25}{3}, 10, \frac{35}{3}, \ldots\right\}$.

If $a_2 - a_1 = 0$, then $a_2 = a_1$ so we did not really make an error.

If $a_2 - a_1 = \pm\frac{5}{3}$ (which can also be written as $|\,a_2 - a_1| = \frac{5}{3}$), then a_2 and a_1 are digits that are separated by $\frac{5}{3}$. The difference between digits is always an integer, so this and all other non-integer multiples of $\frac{5}{3}$ are impossible.

If $|\,a_2 - a_1| = 5$, then then a_2 and a_1 are digits that are separated by 5 units. The digits that are separated by 5 units are 0 and 5, 1 and 6, 2 and 7, 3 and 8, and 4 and 9. These errors will NOT be caught.

If $|\, a_2 - a_1| = 10$, then then a_2 and a_1 are digits that are separated by 10 units. The digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are never separated by more than 9 units, so this and all higher multiples of $\frac{5}{3}$ are impossible.

Therefore, all adjacent transposition errors of the first two digits will be caught unless the digits are separated by 5 units. (This is the same as saying $|\, a_2 - a_1| = 5$).

The same logic also applies to the remaining adjacent digits.

Therefore, this scheme detects all adjacent transposition errors other than the interchange of 0 and 5, 1 and 6, 2 and 7, 3 and 8, or 4 and 9.

(e) Will this code find all jump transposition errors?

Data can be encoded in identification numbers.

An Illinois driver's license contains a portion (Y-YDDD) that represents the last two digits of the year of birth of the driver and codes the birthday according to the following formulas where m represents birth month and d represents the birth date.

$$\text{Male} = 31(m-1) + d$$
$$\text{Female} = 31(m-1) + d + 600$$

(a) What would the Y-YDDD digits of an of Illinois driver's license number look like for a man born on February 12, 1967?

(b) What do you know about a person whose Y-YDDD digits are 1-0642?

(c) What do you know about a person whose Y-YDDD digits are 9-0373?

SAMPLE EXAM QUESTIONS FROM CHAPTER 16

1. Determine the check digit that should be appended to the identification number 634498, if the check digit is the number needed to bring the total of all the digits to a multiple of 10.

(A) The code is invalid (B) 6 (C) 8 (D) 4

(E) None of these

2. Which, if any, of the statements below are true? Mark all correct answers.

(A) $101 \equiv 1 \bmod 2$

(B) $77 \equiv 0 \bmod 11$

(C) $49 \equiv 1 \bmod 12$

(D) $39 \equiv 5 \bmod 5$

(E) None of these are true.

3. The number 4320 is accidentally entered as 4321.
What type of error is this?

(A) A transposition error

(B) A jump transposition error

(C) A single digit error

(D) A baseball error

(E) None of these

4. The last three digits of a person's ID are calculated based on their birthday where m represents birth month and d represents the birth date.

$$\text{Male} = 35(m-1)+d$$

$$\text{Female} = 35(m-1)+d+500$$

(a) What are the last three digits of a man's ID number if he was born on October 8$^{\text{th}}$?

(b) What do you know about a person if the last three digits of the person's ID number are 603?

(c) What do you know about a person if the last three digits of the person's ID number is 320?

5. A code is given by $a_1a_2a_3a_4$ where a_4 is the check digit. The check digit is $a_4 = 7a_1 + 2a_2 + 5a_3 \bmod 9$.

(a) Determine the value of x in the code $2x45$, given that the check digit is valid.

(b) Determine if the check digit will find all single digit errors in the second position.

(c) Determine if the check digit will find all transposition errors in the second and third positions.