

## チェック・ディジットによるエラー検出率の導出\*

吉田 敬一\*\*

### Abstract

It is generally known that the check-digit method is effective for the detection of key-punch errors of code numbers.

This paper denotes that the derivation of the detection rate by the check-digit can be found by the mathematical method without applying simulation or statistical method.

### 1. ま え が き

情報システムを設計する際、処理すべき対象物をコード化することは、よく用いられる手段である。社員コード、品名コードなどはその一例である。

こうしたコードを紙テープやカードに穿孔するとき、誤ってコードの一部ないしすべてを他の数字や文字で穿孔してしまうことがある。そのようなとき、こうしたエラーを発見する方法のひとつとして、check-digit を用いるやり方がある。この方法によるエラーの発見率は、「エラーの種類と、ウエイトに用いられた数字」によって Table 3~Table 6 のように差異が生じることはよく知られているところである。

ところで、この表の数値の導出については文献 1)、2) の論文があるが、これらの方法とは全くべつの新しい方法で数値の導出ができることを示すのが、本論文の目的である。

本論文でとりあげるエラーの種類は

- ① Transcription Error
- ② Transposition Error
- ③ Double Transposition Error

の3つである。また、check-digit を求める2つの方法

- ① コードとウエイトの各桁の積和による方法
- ② コードとウエイトの各桁の積を、一位の数と十

位の数の和に分解する方法 (付録参照)のうち、本論文では①の方法にしたがった。また、Modulus は 10 と 11 の両方の場合について扱った。

### 2. Modulus 10 の場合

#### 2.1 Transcription Error

このエラーは、仮りに5桁の正しいコードを  $c_1c_2c_3c_4c_5$  としたとき、このうちの1桁を誤って (たとえば、 $c_1c_2c_3c_4c_5$  というように) 穿孔してしまう誤りである。

いま、桁数  $n$  のコード  $c_1c_2\cdots c_n$  の  $i$  桁目の数字  $c_i$  を、誤って  $c_i'$  と穿孔したとする。また、このときの  $c_i$  に対するウエイトを  $w_i$  とする。すなわち、

- 正しいコード  $\rightarrow c_1c_2\cdots c_i\cdots c_n$
- 誤って穿孔したコード  $\rightarrow c_1c_2\cdots c_i'\cdots c_n$
- ウエイト  $\rightarrow w_1w_2\cdots w_i\cdots w_n$

である。

このとき、コードとウエイトの積はそれぞれ

$$c_i w_i = 10d_1 + d_2 \quad (1)$$

$$c_i' w_i = 10d_1' + d_2' \quad (2)$$

とあらわすことができる。但し、 $0 \leq c_i \leq 9$ ,  $0 \leq c_i' \leq 9$ ,  $1 \leq w_i \leq 9$ ,  $0 \leq d_1 \leq 8$ ,  $0 \leq d_1' \leq 8$ ,  $0 \leq d_2 \leq 9$ ,  $0 \leq d_2' \leq 9$  で、すべて整数である。

いま、 $c_i$  を誤って  $c_i'$  と穿孔したときこの誤りの発見が不可能であるということは (1), (2) において

$$c_i \neq c_i' \text{ で、かつ } d_2 = d_2' \text{ となるような } c_i, c_i' \text{ が存在すること}$$

である。

そこでいま、 $d_2 = d_2'$  を仮定しそのような  $c_i, c_i'$  が

\* The Derivation of the Detection Rate of Errors by the Check-digit by Keiichi YOSHIDA (College of Engineering, Shizuoka University Dept. of Computer and Information Science).

\*\* 静岡大学工業短期大学部情報工学科

存在するかどうかを考察してみる。  $d_2=d_2'$  とすると (1), (2) より

$$w_i(c_i - c_i') = 10(d_1 - d_1') \quad (3)$$

がえられる。そこで、(3)において  $w_i$  をいろいろとかえてみて、そのときの  $c_i, c_i'$  の存在の有無を調べることとする。

①  $w_i=1$  のとき

$$c_i - c_i' = 10(d_1 - d_1')$$

ところが、 $1 \leq |c_i - c_i'| \leq 9, 0 \leq |d_1 - d_1'| \leq 8$  であるから、上式の左辺および右辺がとりうる値は Table 1 のようになる。Table 1 の左辺と右辺に等しいものがないので、上式を満たすような  $c_i, c_i'$  は存在しない。ゆえに、発見率は 100% である。

②  $w_i=2$  のとき

$$2(c_i - c_i') = 10(d_1 - d_1')$$

$$\therefore c_i - c_i' = 5(d_1 - d_1')$$

ところが、 $1 \leq |c_i - c_i'| \leq 9, 0 \leq |d_1 - d_1'| \leq 8$  であるから、上式の左辺および右辺がとりうる値は Table 2 のようになる。Table 2 から明らかのように、 $|c_i - c_i'| = 5$  のときに限り、上式を満たす  $c_i, c_i'$  が存在する。たとえば、 $c_i=6, c_i'=1$  のように 6 と穿孔すべきところを誤って 1 を穿孔してしまったときがその一例である。 $|c_i - c_i'|$  は 1 から 9 までの 9 通りの整数値をとることができ、 $|c_i - c_i'| = 5$  のときに限ってエ

Table 1 The possible values on the left and the right side when  $W_i=1$  in Eq. (3).

左 辺	右 辺
1	0
2	10
3	20
4	30
5	40
6	50
7	60
8	70
9	80

Table 2 The possible values on the left and the right side when  $W_i=2$  in Eq. (3).

左 辺	右 辺
1	0
2	⑤
3	10
4	15
⑤	20
6	25
7	30
8	35
9	40

Table 3 The detection rate of the transcription errors (Modulus 10).

ウ エ イ ト	発 見 率 (%)
1	100
2	88.9
3	100
4	88.9
5	55.6
6	88.9
7	100
8	88.9
9	100

ラーの発見が不可能となるのであるから求める発見率は

$$\frac{8}{9} \times 100\% = 88.9\%$$

となる。 $|c_i - c_i'| = 5$  となるような  $c_i$  と  $c_i'$  の組合せ ( $c_i, c_i'$ ) は、前記の (6, 1) のほかに (1, 6), (0, 5), (5, 0), (2, 7), (7, 2), (3, 8), (8, 3), (4, 9), (9, 4) がある。

以上のような考察を、 $w_i=3$  以降 9 までにつづけることにより、Table 3 の数値をうることができる。

### 2.2 Transposition Error

このエラーは、仮りに 5桁の正しいコードを  $c_1c_2c_3c_4c_5$  としたとき、このうち互いに隣り合う数字同志を誤って入れかえて (たとえば、 $c_1c_3c_2c_4c_5$  というように) 穿孔してしまう誤りである。

いま、桁数  $n$  のコード  $c_1c_2 \dots c_i c_{i+1} \dots c_n$  の  $i$  桁目と  $i+1$  桁目を見まちがえて、 $c_1c_2 \dots c_{i+1}c_i \dots c_n$  と穿孔してしまったとする。このとき、左から  $i$  番目、 $i+1$  番目のウェイトをそれぞれ、 $w_i, w_{i+1}$  とすると  $c_i w_i + c_{i+1} w_{i+1}$ 、 $c_{i+1} w_i + c_i w_{i+1}$  はそれぞれ

$$c_i w_i + c_{i+1} w_{i+1} = 10d_1 + d_2 \quad (4)$$

$$c_{i+1} w_i + c_i w_{i+1} = 10d_1' + d_2' \quad (5)$$

とあらわすことができる。但し、 $0 \leq d_1 \leq 16, 0 \leq d_1' \leq 16, 0 \leq d_2 \leq 9, 0 \leq d_2' \leq 9$  の整数であり、 $c_i, c_{i+1}, w_i$  については (1), (2) の場合と同様である。

さて、この場合も Transcription Error のときと同様の考察により

$$(w_i - w_{i+1})(c_i - c_{i+1}) = 10(d_1 - d_1')$$

をうる。 $w_i - w_{i+1}$  をあらためて  $w'$  とおくと

$$w'(c_i - c_{i+1}) = 10(d_1 - d_1') \quad (6)$$

をうる。この式の型は (3) 式と同じものである。したがって、これから先の考察は Transcription Error の場合に帰着させることができる。但し、 $w'$  は  $0 \leq w' \leq 9$  である。Table 4 (次頁参照) は、かくしてうることができる。

**Table 4** The detection rate of the transposition errors and double transposition errors (Modulus 10).

ウェイトの差	発見率 (%)
0	0
1	100
2	88.9
3	100
4	88.9
5	55.6
6	88.9
7	100
8	88.9
9	100

### 2.3 Double Transposition Error

このエラーは、仮りに5桁の正しいコードを  $c_1c_2c_3c_4c_5$  とするとき、このうち1桁とんで左右の数字を入れ違えて(たとえば、 $c_1c_4c_3c_2c_5$  というように)穿孔してしまう誤りである。

このことを長さ  $n$  桁のコードで考えれば、 $i$  桁目と  $i+2$  桁目(但し、 $1 \leq i \leq n-2$ )を入れ違えて穿孔してしまうことに相当する。したがって、エラーの発見率を求めるため考察に必要な数式は、(6)において  $c_{i+1}$  の代りに  $c_{i+2}$  としたものになる。すなわち

$$w'(c_i - c_{i+2}) = 10(d_1 - d_1'), \text{ 但し } w' = w_i - w_{i+2}$$

前と同様の理由で、これから先の考察は Transcription Error の場合に帰着させることができる。したがって、エラーの発見率は Table 4 のようになる。

## 3. Modulus 11 の場合

### 3.1 Transcription Error

2.1 の場合と同様の考察により

$$w_i(c_i - c_i') = 11(d_1 - d_1')$$

をうる。ところが、 $1 \leq w_i \leq 9$ ,  $1 \leq |c_i - c_i'| \leq 9$ ,  $0 \leq |d_1 - d_1'| \leq 8$  であるので、 $w_i$  を上記の範囲内でどんなにかえても上式を満たす  $c_i, c_i'$  は存在しない。ゆえに、発見率は100%である (Table 5 参照)。

**Table 5** The detection rate of the transcription errors (Modulus 11).

ウェイト	発見率 (%)
1	100
2	100
3	100
4	100
5	100
6	100
7	100
8	100
9	100

**Table 6** The detection rate of the transposition errors and double transposition errors (Modulus 11).

ウェイトの差	発見率 (%)
0	0
1	100
2	100
3	100
4	100
5	100
6	100
7	100
8	100
9	100

### 3.2 Transposition Error, Double Transposition Error

2.2 および 2.3 と同様の考察により

$$w'(c_i - c_i') = 11(d_1 - d_1')$$

をうる。但し、 $w' = w_i - w_{i+1}$  または  $w_i - w_{i+2}$ ,  $c_i' = c_{i+1}$  または  $c_{i+2}$  である。これから先は 2.2 と同様の考察をつづけることにより Table 6 をうる。

## 4. おわりに

ここで示された「エラーの発見率」の導出方法の特徴は

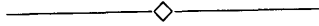
- ① Transcription, Transposition, Double Transposition のどの場合のエラーにおいても、統一された推論の方法が用いられていること
- ② その推論の方法が、データの集計結果や経験、シミュレーションによらず、数式だけにたよっていること
- ③ Transposition, Double Transposition の各場合を Transcription の場合に帰着させていることである。

最後に、貴重な文献やアドバイスを下さった MSK システムズの小沢暢夫氏、松下通信工業の植山 喬氏、NCR の Mr. K. L. Thomas に心から感謝する次第である。

## 参考文献

- 1) 小沢暢夫: 電子計算機インプット研究, p. 509, (株)日本経営出版会, 東京 (1966).
- 2) 植山 喬: チェック・ディジットについて(第I報) 基本的事項, National Technical Report, Vol. 15, No. 3, pp. 309~313 (1969).
- 3) NCR 299: Check Digit Verification Appendix.
- 4) NCR 399: Check Digit Calculation and Verification.

- 5) NCR: Chart of the Most Commonly Used  
Check Digit Systems.  
6) IBM: Self-Checking Number Feature



### 付 録

- ① コードとウェイトの各桁の和積による方法

例で示す:   コード→6382

              ウェイト→1357

$$\underline{6+9+40+14=69}$$

この 69 を Modulus  $M$  でわって, その剰余を

check-digit とする.  $M=10$  のとき, このコードの check-digit は 9 になる.

- ② コードとウェイトの各桁の積を, 一位の数と十位の数に分解する方法  
例で示す;

上記の  $6+9+40+14$  を

$$6+9+4+0+1+4=24$$

として, この 24 を Modulus  $M$  でわって, その剰余を check-digit とする.  $M=10$  のとき, 上記のコードの check-digit は 4 である.

(昭和 52 年 4 月 4 日受付)