STUDY ON SAMPLING INSPECTION SCHEME TO DIGITAL PRODUCTS IN GIS

LIU Chun
LIU Dajie

KEY WORDS sampling inspection; digital products; GIS

ABSTRACT Adopting a principle of "check-accept for the first rank, inspection for the second rank", this paper briefly discusses the rationale of the sampling inspection and the sampling inspection schemes to digital products in GIS. The OC curve is drawn to explain the deficiency of the percent sampling inspection. Meanwhile, the method of One Time Limiting Quality of count selection is presented as the inspection scheme for production departments while the method of One Time After-inspection Mean Percent Defective Upper Limit of count selection is for acceptance departments.

1 Introduction

Spatial data is the basis of the Geographic Information System (GIS) and non-spatial data is the intension and the description of geographic cell. The quality of data affects greatly on the result of application analysis and the realization of the practical target in GIS. The quality of the GIS data must be considered in the procedures of data collection, process and operation.

Today, many research achievements have been made on the data quality in research of GIS. Yan (1998) presented the elements of the data quality from the point of information standardization. Chen and Gong analyzed the sources of data errors concretely in GIS[4]. Shi analyzed the error propagation during the data process[9]. In respect to inspection and testing to geometric error, the positional error of point, line and area features are also discussed respectively[8,9], and a fundamental research on the thematic maps and the sources of attribute errors was given as well[4].

With the development of digital products and industrialization of GIS, more and more digital products appear on the market of surveying and become the basis of applications. In order to ensure the quality of digital products in line, it is necessary to conduct scientific inspection and evaluation to the quality of digital products. Meanwhile, with the development of the standard of quality management and quality guarantee, the requirement of scientific inspection to digital products should be updated. At present, there are about 20 published national standards on sampling inspection which play a fine directorial role in practice. The theory of quality control and inspection can not be used for GIS production perfectly unless the concept, range of the application, method and procedure of these standard have been further studied, which make a good basis for the development of the industrialization in GIS.

The sampling inspection can provide reliable information for the quality management of products. As a basic means of quality control, it is also very important to study the sampling inspection of quality management of GIS products. Because the digital surveying product is a key component in GIS products, this paper will discuss the sampling inspection
methods to GIS products. Taking the digital surveying products as example, first, this paper briefly introduces some key concepts on quality standard for sampling inspection on digital surveying products; then, analyzes the Probability of Acceptance of some sampling inspection schemes; finally, gives some modified suggestions to the sampling inspection schemes.

2 Key points of quality standards for sampling inspection on digital surveying products

The frame of quality standards and the rule of digital surveying products are the basis for investigating and testing analysis in some relevant departments. The framing procedure plays an important role in the development of GIS product, the general requirement of quality control and the development of national GIS industrialization.

The quality of digital surveying products must be inspected with sampling inspection theory based on these quality standards and prescriptions. The key points are:

1) Using the method of “defectiveness” to quantify errors among the inspection of digital surveying products. As the basis for analyzing the quality characteristics of digital surveying products, the degree of influence on the quality should be classified, and the degree of defectiveness be ascertained.

2) Re-drawn map must be inspected by comparison with the real or digital topographic maps. Recording the errors and leaks on the inspection note, the number of defectiveness, evaluated the degree of defectiveness and giving the score of decreasing of the defectiveness.

3) Using the principle of “check-accept for the first acceptance inspection and supervision inspection for the second rank”. The first acceptance inspection is the sampling inspection for users, the inspection for a production group needs 100% total inspection, the figure of procedure quality control is drawn to find out and solve issues timely. The quantity of sampling inspection for production departments should be no less than 30% while the quantity of acceptance inspection for users should be no less than 10%.

4) After given the assessment of the quality to every products unit, remark as “after inspection, the batch is judged as eligibility or ineligibility” should be given according to the sample quality.

From above, we can see both the sampling inspection for production departments and users need sampling inspection. The method used is a simple inspection, and named as percent inspection.

The framing of quality standards can promote the development of digital surveying products, and the simple sampling inspection has virtues of facility and simplicity.

3 The probability of acceptance of sampling inspection methods to digital surveying products

While the sampling inspection is performed to a batch of products, the result of acceptance or rejection should be given. The probability that the batch of products may be accepted is called Probability of Acceptance. The Probability of Acceptance has an affinity with percent defective $p$ to a batch of products, so that the Probability of Acceptance is expressed as $L(p)$. To a batch of products which has the lot size of $N$, the One Time Sampling Inspection scheme is expressed as $(n, c)$, where $n$ is the size, and $c$ is the number used to judge eligibility.

After the ineligibility number $d$ is gained with the inspection, we can make a decision according to the rules: if $d \leq c$ then the batch is accepted, if $d > c$ then the batch is rejected.

The number of unaccepted product is $D = N \cdot p$ to a batch of $N$, and the number of unaccepted product sampled from sample size of $n$ obeys the Hypergeometric Distribution, the probability is:

$$P(d) = \frac{C_d C_{N-d} C_N}{C_n}$$

the Probability of Acceptance which the batch of products are judged as eligibility according to the scheme of One Time Sampling Inspection $(n, c)$ is:

$$L(p) = \sum_{d=0}^{c} P(d) = \sum_{d=0}^{c} \frac{C_d C_{N-d} C_N}{C_N}$$
When $N$ is larger, and $N/N \leq 0.10$, $L(p)$ can be calculated according to the Binomial Distribution:

$$L(p) = p(d \leq c) = \sum_{d=0}^{c} \binom{n}{d} p^d (1 - p)^{n-d}$$

(3)

The Percent Defective $p$ is a decreasing function of the Probability of Acceptance $L(p)$. The functional relation between $L(p)$ and $p$ is usually in rectangular coordinates system, and expressed as sampling inspection characteristic curve (named as OC curve). OC curve has the virtues of sectionalizing the ability of acceptance or rejection to different sampling schemes because different sampling schemes have different OC curves, and it is one of the references to establish and select a good sampling inspection scheme.

According to the above discussion, for digital surveying products, it is stipulated that the lot size of sampling inspection for production departments is no less than 30% while the lot size of sampling inspection for users is no less than 10%. However, the decision number $c$ is still not regulated. In reality, it is often used with the methods of percent sampling that fetch $n = N \cdot 10\%$ from sample number after giving the decision number $c$ and double percent sampling that fetch $c = n \cdot k\%$ (e.g. $k = 5$) from decision number.

From Fig. 1, we can see both the methods of percent sampling inspection and double percent sampling inspection have the defect of "strictness for large lot size, toleration for small lot size". References have more discussions about this defect. However, for digital surveying products, it is regular to take "a map" as a product unit in Reference [6]. In practice, the lot size will not be very large because a map contains quite a lot of contents. So the issue that strictness for large lot size will not be very serious when using the method of percent sampling inspection. When the lot size is a small one, the issue of toleration for small lot size will be improved when selecting decision number $c$ properly. However, from the point of modern quality management of digital products in GIS, it is necessary to select more scientific sampling inspection schemes.

4 Selection of sampling inspection schemes

The sampling inspection schemes to products have three categories: standard form, adjustable form and selection form. The adjustable form is not suitable for today's GIS products because it only suits the continuation of the production and acceptance, and represents its superiority only used in long term and sequence. The scheme of standard form needs the quality level of products $p_0$ and $p_1$, which must be given by production department and users. But the production department and users are lack of experience and reference in giving the quality level of $p_0$ and $p_1$.

The scheme of selection form has the feature of accepting the batch if its sample is judged as eligi-
Study on Sampling Inspection Scheme to Digital Products in GIS

65

ability, and will conduct full inspection to the batch if the sample is judged as ineligibility. The national standard GB/T15346 has given the sampling schemes and procedure about the One Time Sampling Inspection of count selection.

Two kinds of schemes of One Time Sampling Inspection of count selection are as follows.

4.1 One Time Limiting Quality Sampling Inspection of count selection

To a batch of products whose lot size is \( N \), assuming that the mean percent defective of these products is \( \bar{p} \), if the One Time Sampling scheme is \(( n, c)\), the Probability of Acceptance can be calculated by Eq. (3) when \( N \) is larger, and \( \frac{n}{N} \leq 0.10 \), \( \bar{p} \leq 0.10 \). According to the concept of mean, the mean inspection number of One Time Sampling Inspection scheme of count selection is:

\[
I = n + (N - n)[1 - \sum_{d=0}^{c} b(d, n, \bar{p})]
\]

if a limiting percent defective \( p_t \) is given, the Probability of Acceptance of the limiting percent defective is:

\[
L(p_t) = \sum_{d=0}^{c} b(d, n, \bar{p})
\]

Where \( L(p_t) \) is regarded as the risk of user, so the criteria of One Time Limiting Quality Sampling Inspection of count selection is given as the risk of user \( L(p_t) \). And limiting percent defective \( p_t \), the \( c \) and \( n \) will be computed by given arbitrary \( N \) and \( \bar{p} \), meantime, making the given \( I \) minimum. Table 1 is the sampling schemes for different limiting quality, different lot size using the method of One Time Limiting Quality Sampling Inspection of count selection.

Form Table 1, we can see:

1) When the lot size \( N \) is very small \((N < 25 \sim 50)\), full-inspection is required.

2) When the limiting quality \( p_t = 5\% \), \( N = 91 \sim 150, (n, c) = (37, 0) \), or \( p_t = 3.15\% \), \( N = 151 \sim 280, (n, c) = (63, 0) \). Lot size \( n/N \) is approximately 30%.

3) When \( N \) is more large, \( n/N < 30\% \).

4.2 One time after-inspection mean percent defective upper limit of count selection

Assuming the percent defective for the totality of product is \( p \), using the scheme \(( n, c)\) to inspect the products whose lot size is \( N \), \( AOQ \) is used to express the After-inspection Mean Percent Defective, then we have:

\[
AOQ = \frac{N - n}{N} \cdot L(p) \cdot 0 \cdot [1 - L(p)]
\]

When \( L(p) \) is the sample Probability of Acceptance of percent defective \( p \), when \( N \) is more large, and \( \frac{n}{N} \leq 0.10, p \leq 0.10 \), \( AOQ \) can be expressed by using the form of Poisson Distribution:

\[
AOQ = \frac{N - n}{N} \cdot L(p)
\]

To the scheme \(( n, c)\), learned from Eq. (6), \( AOQ \) is the function of \( p \), and defining \((AOQ)_p = 0\), we have:

\[
\sum_{d=0}^{c} \frac{(np)^d}{d!} e^{-np} = \frac{1}{c!} e^{-np}
\]

If define \( x = np \), then have:

Table 1 Sampling schemes for One Time Limiting Quality Sampling Inspection of count selection

| Limiting quality \( p_t/\% \) | Lot size \( N \) | Sampling schemes \((n, c)\) |
|-------------------------------|----------------|-----------------------------|
| 1.25                         | 150–280        | Full-inspection (122,0)     |
| 2.00                         | 150–280        | Full-inspection (89,0)      |
| 3.15                         | 51–90          | Full-inspection (48,0)      |
| 5.00                         | 51–90          | Full-inspection (33,0)      |
| 8.00                         | 51–90          | Full-inspection (22,0)      |
Define \( y = x \cdot \sum_{d=0}^{\infty} \frac{x^d}{c!} e^{-x} \),
then the upper limit \( p_L \) of AOQ is:
\[
p_L = \left( \frac{1}{n} - \frac{1}{N} \right) y
\]
We can obtain from above:
\[
y = \frac{x^{e+2}}{c!} e^x
\]
\[
n = \frac{N \cdot y}{y + N \cdot p_L}
\]
so the criteria of One Time After-inspection Mean Percent Defective Upper Limit of count selection is given for after-inspection mean percent defective upper Limit \( p_L \) and \( N \), fixing the \( c \) arbitrarily, \( x \) is computed from Eq. (9) and \( y \) is computed from Eq. (11), then determine \( n \) from Eq. (12), from the group of schemes \((n, c)\), look for the scheme which can make the given \( I \) minimum, where the \( I \) is determined by the given mean percent defective \( \bar{p} \) and from Eq. (4). Table 2 is the sampling schemes for different after-inspection mean percent defective upper limit and different lot size when using the method of One Time After-inspection Mean Percent Defective Upper Limit of count selection.

Form Table 2, we can see:
1) When the lot size \( N \) is very small \((N < 16 \sim 25)\), full-inspection is required.
2) When the After-inspection Mean Percent Defective Upper Limit is \( p_L = 2\% \), \( N = 91 \sim 150 \), \((n, c) = (15, 0)\), and when \( p_L = 5\% \), \( N = 51 \sim 90 \), \((n, c) = (7, 0)\). The lot size \( n/N \) is approximately 10%.
3) When \( N \) is more large, \( n/N < 10\% \).

### 4.3 Selection of sampling inspection schemes to digital data

From above results, we can see when the lot size \( N \) is not very small or very large, the quantity using the Limiting Quality Sampling Inspection Scheme approaches to 30%, and the quantity using After-inspection Mean Percent Defective Upper Limit approaches to 10%. They all can overcome the defect of "strictness for large lot size, toleration for small lot size". Therefore it is suggested that the sampling scheme of Limiting Quality of count selection is used to inspect sample for production department, and the Limiting quality can be selected as 3% \sim 5%. The sampling scheme of After-inspection Mean Percent Defective Upper Limit is used for acceptance inspection by user, and the after-inspection percent defective upper limit can be selected as 2% \sim 5%.

| After-inspection mean percent defective upper limit \( p_L /\% \) | Lot size \( N \) | Sampling schemes \((n, c)\) |
|-----------------------------|-----------------|-----------------|
| 1 - 90                      | Full-inspection |                |
| 0.2                         | 91 - 150        | (83, 0)         |
| 151 - 280                   | (111, 0)        |                |
| 1 - 50                      | Full-inspection |                |
| 51 - 90                     | (41, 0)         |                |
| 0.5                         | 91 - 150        | (50, 0)         |
| 151 - 280                   | (58, 0)         |                |
| 1 - 25                      | Full-inspection |                |
| 26 - 50                     | (24, 0)         |                |
| 0.8                         | 51 - 90         | (31, 0)         |
| 91 - 150                    | (35, 0)         |                |
| 151 - 280                   | (40, 0)         |                |
| 1 - 16                      | Full-inspection |                |
| 17 - 25                     | (11, 0)         |                |
| 26 - 50                     | (14, 0)         |                |
| 2.0                         | 51 - 90         | (15, 0)         |
| 91 - 150                    | (16, 0)         |                |
| 151 - 280                   | (17, 0)         |                |
| 1 - 16                      | Full-inspection |                |
| 17 - 25                     | (6, 0)          |                |
| 26 - 50                     | (6, 0)          |                |
| 5.0                         | 51 - 90         | (7, 0)          |
| 91 - 150                    | (7, 0)          |                |
| 151 - 280                   | (7, 0)          |                |

During the procedures of sampling inspection, two kinds of miscarriage of justice may occur when the lot judged as qualification or disqualification according to the \( c \): 1) Taking the qualification lot as disqualification lot, the producer will expense for this kind of miscarriage. 2) Taking the disqualification lot as qualification lot, then the user will expense for this kind of miscarriage. Usually, the probability of the first kind of miscarriage is marked as \( a \), and named as producer’s risk; the probability of the second kind of miscarriage is marked as \( \beta \), and named as user’s risk.

Two rat of reject \( (p_0, p_1, p_0 < p_1) \) are often provided to control the two risks. The \( p_1 \) is called limit
quality. A lot of digital data is regarded as qualification when its rate of reject $p_0 < p_1$, and is accepted with high probability or rejected with low probability, the probability of reject is usually controlled within $\alpha$; A lot of digital data is regarded as disqualification when its rate of reject $p_1 < p_0$, and is rejected with high probability or accepted with low probability. A certain relation between $\alpha$, $\beta$ and $p_0$, $p_1$ is in existence, and it connects with the sampling size $n$ and the judge eligibility number $c$. The two kinds of miscarriage can not be avoided during the procedure of sampling. So the schemes of sampling, $\alpha$, $\beta$ and $p_0$, $p_1$ should be given in reason in order to restrict the two risks to a certain scope.

The sampling scheme of count selection should be such for controlling the user’s risk and making the total inspection number minimal, or the total charge minimal. So the count selection is adapted to the check-up inspection and check-accept inspection, and both the user’s risk and controlling the sampling error under the minimal charge, will be considered.

5 Summary and conclusion

The collection of spatial data is often in order and in batch in GIS. Some sampling methods corresponding to the national standards can be used to inspect quality of digital products because the procedure of the collection of spatial data can be regarded as a working process of digital product. This paper discussed the rationale and method for the sampling inspection to digital product in GIS, and explain the deficiency by using the traditional percent sampling to inspect digital products. Then, Limiting Quality Sampling scheme is presented for production department as the inspection at second rank, and the sampling scheme of After-inspection Mean Percent Defective Upper Limit is used for users as acceptance inspection. These two schemes are in agreement with exiting percent sampling inspection scheme in certain circumstances, and have the advantage that they can be easily conducted. More contents and quality characteristics have been involved in the digital product in GIS, and the model of sampling level and suited for variant complexity of digital products, and ensure the reliability of sampling while make the fee of sampling minimum.

References

1 Yu S Q. Sampling inspection and quality control. Beijing: Beijing University Press, 1991 (in Chinese)
2 Ma R, Wang Z C. The count of sampling inspection. Harbin: The Press of Harbin University of Industry, 1993 (in Chinese)
3 Ma Y L, Yu Z F, Yu S Q. Sampling inspection for product quality. Beijing: China Press of Standard, 1997 (in Chinese)
4 Chen J, Gong P. The practice geographic information system. Beijing: The Scientific Press, 1998 (in Chinese)
5 Yan Z. The tutor of standardization in GIS for city. Beijing: The Scientific Press, 1998 (in Chinese)
6 The State Bureau of Surveying and Mapping. Quality Standard for Digital Topographic Map Product. (in Chinese)
7 Li Z Q. The explanation for quality standard for Digital Topographic Map Product. The State Bureau of Surveying and Mapping (in Chinese)
8 Liu D J. Accuracy analysis and quality control of spatial data in geographic information systems. Shanghai: Shanghai Press of Scientific Literature, 1999 (in Chinese)
9 Shi W Z. Modeling positional and thematic uncertainties in integration of remote sensing and geographic information systems. ITC Publication, Enschede, 1994