ESTIMATING MEASUREMENT PRECISION BY MEANS OF MEASUREMENT DIFFERENCES

V. Casella

DIET – Università degli Studi di Pavia, Pavia, Italy
email: casella@unipv.it

ISPRS Commission VI, Working Group 3

KEYWORDS: Precision estimation, variance estimation, estimators, statistics

ABSTRACT

When it is necessary to determine the precision of a certain instrument B, devoted to point position measurement for instance, but not necessarily, one way is to measure a certain number of points with B as well as with another instrument A, whose precision is known. Forming the differences between the B and A measurements, it is possible to calculate B precision by means of the dispersion of the differentiated quantities. The paper investigates how the precision of A influences the estimation of the precision of B. Even if the subject is basic, it is not easy to find in textbooks; for this reason the paper could have a certain interest for teaching purposes.

1. INTRODUCTION

Let's assume there are *n* points in the space and two instruments A and B able to measure their positions. The instrument called A has well known characteristics, while B has an unknown precision, which we want to determine by comparison between the measurements given by it and the ones given by A. This scheme is very common and applies, for instance, to the estimation of the precision of a DPW (Digital Photogrammetric Workstation) by means of the comparison with an analytical stereoplotter; the same scheme could also be applied to the study of the performances of a quick GPS mode, based on static measurements.

Our discussion will be limited to only one component, and will assume that the instruments have always the same precision, regardless of the position of the measured point. The measurement of the *i*-th point by the A instrument can be represented by a normal random variable (rv starting from now)

$$X_i = N\left[\bar{x}_i, s_A^2\right] \qquad (1)$$

and the measurement of the B instrument can be represented by the following normal rv

$$X_i = N\left[\bar{x}_i, s_B^2\right] \qquad (2)$$

where the *x* with the small line above it represents true values, while x_{Ai} and x_{Bi} represent the measured positions. The variance of A is supposed to be known, while the variance of B is unknown. The equation (1) and (2) also contains the hypothesis that both the instruments have no biases; this hypothesis should be checked in an actual case, but it doesn't damage the general value of our discussion: it only makes the job simpler.

It is possible to form the differences between the measurements given by the two instruments:

$$d_i = x_{Bi} - x_{Ai} \qquad (3)$$

which are formally extracted by *n* different but identical normal rvs (we will follow the convention of indicating one rv with an uppercase letter and extractions from it with the same letter, lowercased)

$$\Delta_i = N\left[0,\ s_A^2 + s_B^2\right]$$

The *n* differences can also be thought of as multiple extractions from a unique rv that will be called Δ

$$\Delta = N\left[0,\ s_A^2 + s_B^2\right] \qquad (4)$$

and this allows the estimation of the dispersion of Δ:

$$s_\Delta^2 = \frac{1}{n}\sum_{i=1}^{n} d_i^2$$

which is an extraction from the estimator rv

$$S_\Delta^2 = \frac{1}{n}\sum_{i=1}^{n} \Delta_i^2$$

The (4) means, among other things,

$$\begin{aligned} \sigma_\Delta^2 &= \sigma_A^2 + \sigma_B^2 \\ \sigma_B^2 &= \sigma_\Delta^2 - \sigma_A^2 \end{aligned} \tag{5}$$

which implies, passing from the true quantities to the estimated ones, that it is possible to give an estimation of the variance of the instrument B in the following way

$$S_B^2 = S_\Delta^2 - \sigma_A^2 \tag{6}$$

The equations (5) and (6) are always true, independently of the value of σ_A^2, and this seems to contradict our common-sense, which suggests that **the estimation of the precision of the instrument B can be effectively carried out by comparison with the measurements given by A only if the latter is better than the former** or, more formally, if the following inequality holds

$$\sigma_A < \sigma_B$$

The point is that (6) is an estimation; it is always true, but this only means that the mean values of the rvs involved coincide, that is

$$E\left[S_B^2\right] = E\left[S_\Delta^2 - \sigma_A^2\right] \tag{7}$$

Moreover the latter equality doesn't guarantee anything about the concentration of the rv S_B^2 around its mean value. In other words, it is necessary to investigate the goodness of the single estimation

$$s_B^2 = s_\Delta^2 - \sigma_A^2$$

and it will be demonstrated that it depends on the value of σ_A^2.

2. DISTRIBUTIONAL RESULTS

Let's consider again the estimator S_Δ^2. It is not too difficult to find how it is distributed. Indeed we have

$$\frac{S_\Delta^2\, n}{\sigma_A^2 + \sigma_B^2} = \sum_{i=1}^{n} \frac{\Delta_i^2}{\sigma_A^2 + \sigma_B^2} = \chi_n^2 \tag{8}$$

because the terms under the sum are in reality normal standardised rvs, as it can be easily demonstrated by the following equalities

$$\sum_{i=1}^{n} \frac{\Delta_i^2}{\sigma_A^2 + \sigma_B^2} = \sum_{i=1}^{n} \frac{\Delta_i^2}{\sigma_\Delta^2} = \sum_{i=1}^{n} \left(\frac{\Delta_i}{\sigma_\Delta}\right)^2 = \sum_{i=1}^{n} Z_i^2$$

Now, it is well known that the sum of n independent normal standardised rvs equals a chi square rv with n degrees of freedom, and this brings us to (8).

From relation (8) it is immediately deduced that

$$S_B^2 = \frac{\sigma_A^2 + \sigma_B^2}{n}\chi_n^2 - \sigma_A^2$$

which represents the first significant result of our discussion: the distribution of S_B^2 depends on the value of the variance of A. An immediate way to investigate this dependence formally and quantitatively is to consider the two most important descriptors of a probability distribution: mean value and variance. Remembering that for the chi square rv the following equalities hold

$$\begin{aligned} E\left[\chi_n^2\right] &= n \\ VAR\left[\chi_n^2\right] &= 2n \end{aligned}$$

whose demonstration can be found in every good statistics textbook, we have

$$E\left[S_B^2\right] = s_B^2$$

$$VAR\left[S_B^2\right] = \frac{2\left(s_A^2 + s_B^2\right)^2}{n} \qquad (9)$$

The first one simply confirms what has been observed about (6) and (7); it also shows that our formal development is correct; the second highlights the dependence of the variance of S_B^2 in respect of s_A^2. In other words: S_B^2 is a correct estimator of the variance of B because its mean value coincides with s_B^2, but the goodness of the estimation depends inversely on s_A^2.

Another well known way of understanding the properties of an estimator is to study its confidence interval at a certain confidence level a . From (8) it is easily obtainable

$$c^2_{n;\,a/2} \le \frac{S_\Delta^2\, n}{s_A^2 + s_B^2} \le c^2_{n;\,1-a/2} \qquad P = 1 - a$$

where $c^2_{n;\,a/2}$ and $c^2_{n;\,1-a/2}$ are two real numbers with the following properties

$$P\left([0, c^2_{n;\,a/2}] \,|\, c_n^2\right) = \frac{a}{2}$$

$$P\left([0, c^2_{n;\,1-a/2}] \,|\, c_n^2\right) = 1 - \frac{a}{2}$$

The confidence intervals for the estimated variance and standard deviation of B are easily calculated

$$c^2_{n;\,a/2}\frac{s_A^2 + s_B^2}{n} - s_A^2 \le S_B^2 \le c^2_{n;\,1-a/2}\frac{s_A^2 + s_B^2}{n} - s_A^2$$

$$\sqrt{c^2_{n;\,a/2}\frac{s_A^2 + s_B^2}{n} - s_A^2} \le S_B \le \sqrt{c^2_{n;\,1-a/2}\frac{s_A^2 + s_B^2}{n} - s_A^2}$$

The role played by the size of the variance of the instrument A is also shown in this case: the bigger it is, the larger the confidence interval becomes. So it confirms the rule that high values of s_A^2 mean low quality estimation of the variance of B. In the next section a practical example will clarify the orders of magnitude of the phenomenon.

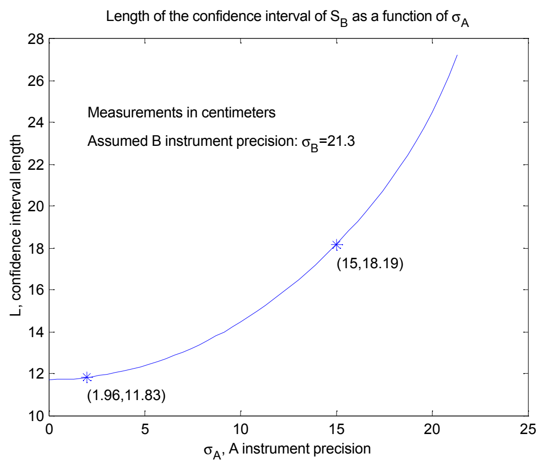
3. PRACTICAL EXAMPLE AND CONCLUSIONS

For a research job devoted to the determination of the precision of a DPW at different resolutions, the position of 25 control points has been measured with an analytical stereoplotter. The operator has completed not only one cycle of orientation and point measurements, but twelve cycles, so to determine the precision of the analytical measurements. For the X component the standard deviation was $s_X = 6.8$ cm; having twelve independent measurements, mean values have been calculated for each point

$$x_{Ai} = \frac{1}{12}\sum_{j=1}^{12} x_{Aij}$$

Their estimated standard deviation of the analytically measured coordinates is $s_A = s_X / \sqrt{12} = 1.92$ cm. The DPW measurements of the X component, at the resolution of 300 dpi, have shown an estimated standard deviation $s_B = 21.3$ cm; this will be assumed as the true value for the instrument B ($s_B = 21.3$).

The following picture shows the confidence interval width of the estimated standard deviation S_B as a function of s_A^2; the value of s_B and *n* are kept fixed; the confidence level is $a = 0.05$.



The interval width obviously increases with s_A; there is a lower limit, corresponding to the value $s_A = 0$, that can be improved only by increasing the number of the points, *n*; this limit is unfortunately high in our case, because the standard deviation estimation improves very slowly, in respect to the number of the measurements.

Therefore, when planning an experimental job, it is necessary to tune carefully, in a combined approach, the precision of the control instruments and the number of the measurements.

4. REFERENCES

Casella, V., 1998. Il sistema di fotogrammetria digitale Leica DVP per l'aggiornamento di cartografia a grande scala. Bollettino SIFET, n. 2/98, pagg. 19-36.

Mikhail, E.M., 1976. Observations and Least Squares. IEP, New York.

Mood A., Graybill F., Boes D., 1993. Introduzione alla statistica. McGraw-Hill, Milano, Italy.

Sansò, F., 1989. Il Trattamento Statistico dei Dati. Città Studi, Milano, Italy.

Stuart A., Ord K., 1994. Kendall's advanced theory of statistics. Distribution Theory. Edward Arnold, London, Great Britain.