Lubomír Kubáček; Ludmila Kubáčková

Ferrero formula and triangle misclosures revisited

Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica, Vol. 36 (1997), No. 1, 99--116

Persistent URL: `http://dml.cz/dmlcz/120357`

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **36** (1997) 99–116

Ferrero Formula and Triangle Misclosures Revisited *

LUBOMÍR KUBÁČEK, LUDMILA KUBÁČKOVÁ

*Department of Mathematical Analysis, Faculty of Sciences,
Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kubacekl@risc.upol.cz*

(Received January 14, 1997)

Abstract

The accuracy of geodetic triangulation network is characterized by the Ferrero estimator of the standard deviation in measurements of angles or directions and by misclosure histograms. Statistical properties both of these characteristics are investigated.

Key words: Quadratic estimator, histogram, Ferrero formula, triangulation network, triangle misclosure, Pearson statistic.

1991 Mathematics Subject Classification: 62J05, 86A30

Introduction

The time of large triangulation networks in geodesy is over. However the coordinates of network points have been utilized. The characteristics of accuracy of triangulation measurements are topical.

Statistical methods developed and applied in geodesy in the time after Ferrero formula enable us to make a look back on some statistical properties of this formula and on a statistical behaviour of the triangle misclosure histogram, which together with the Ferrero formula had been used for an analysis of the accuracy of triangulation networks.

*Supported by grant No 201/96/0436 of The Grant Agency of Czech Republic

1 Formulation of problems and auxiliary lemmas

Let in R^2 (two-dimesional Euclidean space) an $n-$tuple of triangles be considered. These triangles must not be crossed by any side of another triangle; only common sides are admissible. In each triangle all three angles are measured with an accuracy characterized by a standard deviation σ in a single direction. Thus $3n$ angles are measured. Let $\eta_{i,1}, \eta_{i,2}, \eta_{i,3}$ be random variables the realizations of which give the result of measurement of angles in the ith triangle. Then $U_i = \eta_{i,1} + \eta_{i,2} + \eta_{i,3} - \pi$ is the misclosure of the ith triangle.

A preliminary analysis (i.e. an analysis performed before the optimum processing the measured data within a geodetical network) of a set of triangle misclosures consists usually of

a) calculating the Ferrero formula and

b) constructing a "histogram" of misclosures of triangles included into the trigonometric network.

Definition 1.1 The Ferrero estimator (cf. also [2], p. 43 and [9], p. 163) of a standard deviation σ in a measurement of a direction, characterizing the whole described triangulation network, is

$$\tilde{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} U_i^2}{6n}}.$$

This formula has been used also in cases when the accuracy of measurements of directions is not the same in different triangles.

The statistical behaviour of the vector $U = (U_1, \dots, U_n)'$ ($'$ denotes the transposition) depends on the method applied for the measurement of angles. Therefore the statistical properties of $\tilde{\sigma}$ from Definition 1.1 depends on it as well.

In the following two methods are considered, i.e., the method of measurement in sets and the Schreiber method (measurement of angles in all combinations, cf. also [4]) and also the vertex method is mentioned. The aim, in the case of normally distributed errors, is

1. to recognize, whether different methods of measurement make the statistical properties of the Ferrero estimator significantly different and

2. to investigate the behaviour of the Pearson statistic used for goodness-of-fit test in the case of the misclosures histogram.

For solving the first problem the following lemmas are useful.

Lemma 1.2 *Let $Y \sim N_n(\mu, \Sigma)$ and A be a symmetric $n \times n$ matrix. Then $Y'AY + 2b'Y + c$ has a $\chi_r^2(\delta)-$distribution (r are degrees of freedom and δ is the parameter of non-centrality) if and only if*

(i) $\Sigma A \Sigma A \Sigma = \Sigma A \Sigma$,

(ii) $\mathcal{M}[\Sigma(A\mu + b)] \in \mathcal{M}(\Sigma A \Sigma)$, *and*
(iii) $(A\mu + b)'\Sigma(A\mu + b) = \mu' A\mu + 2b'\mu + c$.
In this case $r = Tr(A\Sigma)$ *and* $\delta = (b + A\mu)'\Sigma A \Sigma (b + A\mu)$.

Proof Cf. [8] p. 171.

Remark 1.3 If the random variable $Y'AY + 2b'Y + c$ satisfies the conditions from Lemma 1.2, then the parameter of non-centrality δ can be written as $\delta = [E(Y)]'AE(Y) + 2b'E(Y) + c$.

Lemma 1.4 *Let* $Y \sim N_n(0, \Sigma)$ *and* B *be a symmetric* $n \times n$ *matrix. Then*

$$Var(Y'BY|\Sigma) = 2Tr(B\Sigma B\Sigma).$$

Proof Cf. [3], p. 327 (the notation $|\Sigma$ means that the variation is determined for the actual value $Var(Y) = \Sigma$ of the observation vector).

Lemma 1.5 *Let* A *be an* $n \times n$ *symmetric matrix and let* Y *be an* $n-$*dimensional random vector such that* $E(Y) = \mu$ *and* $Var(Y) = \Sigma$ *(the normal distribution of it is not assumed). Then*

$$E(Y'AY|\mu, \Sigma) = \mu' A\mu + Tr(A\Sigma).$$

Proof Proof is obvious (the notation $|\mu, \Sigma$ means that the mean value is determined at the point μ, Σ, where μ and Σ are the actual values of the mean value and the variance matrix of the observation vector, respectively). □

Lemma 1.6 *Let* $Y \sim N_n(0, \sigma^2 V)$, *where* V *is positively definite. The uniformly best unbiased estimator of* σ^2 *is*

$$\hat{\sigma}^2(Y) = Y'V^{-1}Y/n.$$

Proof It is based on Theorem 3.1. and Example 3.4 in [5].

Lemma 1.7 *Let* $\eta_i \sim N_1(\mu, \sigma^2)$, $i = 1, \ldots, n$, *be identically and independently distributed random variables. The points* $t_1, \ldots, t_{k-1}$ *such that* $-\infty < t_1 < t_2 < \ldots < t_{k-1} < \infty$, *divide the interval* $(-\infty, \infty)$ *into* k *disjunctive intervals* $T_1 = (-\infty, t_1), T_i = [t_{i-1}, t_i)$, $i = 2, \ldots, k-1$, *and* $T_k = [t_{k-1}, \infty)$. *If* n_i *is the number of realizations of the* n*-tuple* $\eta_1, \ldots, \eta_n$ *into the interval* T_i $(n_1 + \ldots + n_k = n)$ *and*

$$p_i = \int_{T_i} n(x, \mu, \sigma^2) dx, \quad i = 1, \ldots, k,$$

then the random variable

$$\sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \tag{1}$$

possesses assymptotically a central chi-square probability distribution with $k - 1$ *degrees of freedom.*

Proof It is a special case of the statement given, e.g., in [7], p. 330.

2 Basic statistical properties of the Ferrero estimator

In the first step the estimator

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n} U_i^2}{6n} \tag{2}$$

is considered. The statistical properties of (2) are studied within a regular hexagonal network, starting with its simplest form and ending with its infinite form, i.e., in a sequence of growing hexagonals, the first, central ($p = 1$), of them being composed of 6 triangles rounding the central point P_0, the second ($p = 2$) being formed by the further 18 triangles rounding the central ($p = 1$) hexagon, etc. (Fig. 1)

Definition 2.1 The measurement of angles in sets in the framework of hexagonals described is called Scheme 1.

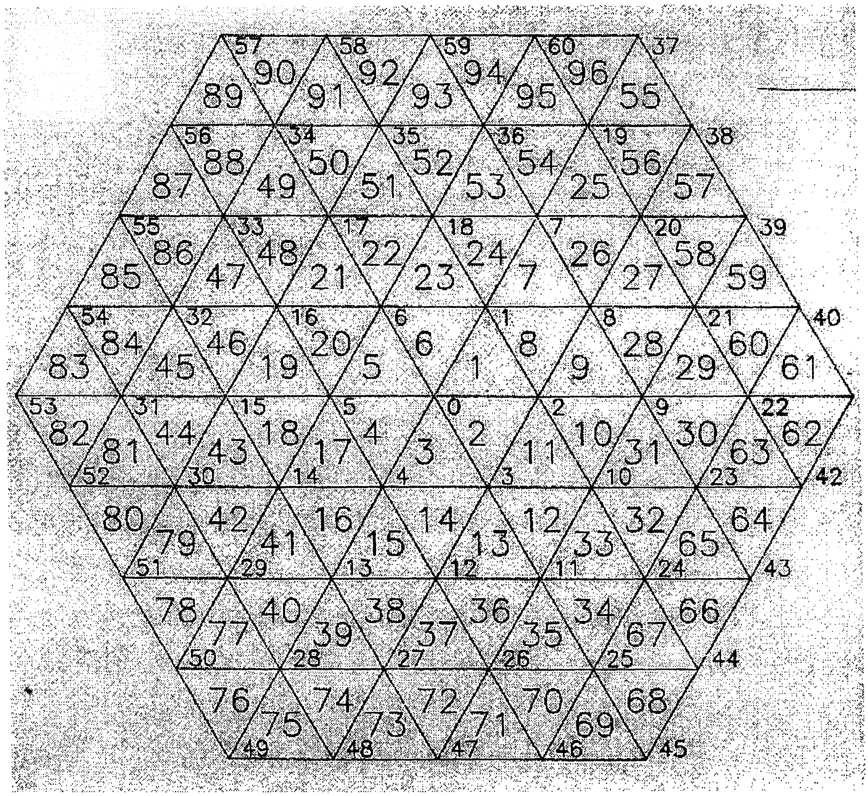

Fig. 1

Lemma 2.2 *In Scheme 1*

(i) the number of vertices is $1 + 3p(p+1), p = 1, 2, \ldots$*;*

(ii) the number of triangles is $6p^2, p = 1, 2, \ldots$*;*

(iii) the number of directions is $6[4 + (p-1)(3p+4)], p = 1, 2, \ldots$.

Proof is obvious.

Example 2.3 Let us consider the case $p = 1$. In this case 24 directions at 7 vertices are to be measured. For the sake of simplicity 1 set in both positions of the telescope is considered (it is the basic unit of replications in several sets). Thus the observation vector ${}_iY$ modelling the measurements in the ith, $i = 1, 2$, position of the telescope is of the form

$$ {}_iY_{(24)} = ({}_iY'_{0(6)},\ {}_iY'_{1(3)},\ {}_iY'_{2(3)},\ {}_iY'_{3(3)},\ {}_iY'_{4(3)},\ {}_iY'_{5(3)},\ {}_iY'_{6(3)})', $$

where

$$ {}_iY'_{0(6)} = ({}_iY_{0,1},\ {}_iY_{0,2}, \ldots,\ {}_iY_{0,6})',\ {}_iY'_{j(3)} = ({}_iY_{j,j+1},\ {}_iY_{j,0},\ {}_iY_{j,j-1})', $$

$$ i = 1, 2, j = 2, \ldots, 5, $$

$$ {}_iY_{1(3)} = ({}_iY_{1,2}, {}_iY_{1,0}, {}_iY_{1,6})' $$

and

$$ {}_iY_{6(3)} = ({}_iY_{6,1}, {}_iY_{6,0}, {}_iY_{6,5})' $$

(the first index shows the number of the station, the second the number of the target, the numbers in parenthesis denote the dimension of the vector, if necessary). The result of the described direction measurements is a 48-dimensional vector, a realization of a random vector $({}_1Y', {}_2Y')'$ characterized by the covariance matrix

$$ Var[({}_1Y', {}_2Y')'|\sigma^2] = 2\sigma^2 I_{(48,48)}. $$

The input for data processing is a 24-dimensional random vector

$$ Y = ({}_1Y + {}_2Y)/2 $$

characterized by a covariance matrix

$$ Var(Y|\sigma^2) = \sigma^2 I_{(24,24)}. $$

6-dimensional vector of triangle misclosures within the central hexagon and its covariance matrix are

$$ U = TY - \pi\mathbf{1},\ Var(U) = \sigma^2 V, $$

where

$$ T_{(6,24)} = \left(\begin{array}{cccccccccccc} -1 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 1 & 0 & 0 & 0 \end{array} \right. $$

$$\left.\begin{array}{rrrrrrrrrrrr}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0
\end{array}\right),$$

$\mathbf{1}_{(24)} = (1,1,1,1,\dots,1)'$, $V = TT'$ and σ^2 is the actual value of the unit dispersion (in direction).

The estimator of the unit dispersion according to the Ferrero formula is $\tilde{\sigma}^2(Y_{(24)}) = U'U/(6 \times 6)$. This estimator is unbiased because of

$$E[\tilde{\sigma}^2(Y_{(24)}|\sigma^2)] = (1/36)E[U'U|\sigma^2] = (1/36)Tr[Var(U|\sigma^2)] = \sigma^2.$$

Its dispersion under the condition $U \sim N_6(0, \sigma^2 V)$ is $Var[\tilde{\sigma}^2(Y_{(24)}|\sigma] = 2\sigma^4 \times 36^{-2}Tr(V^2) = (11/27)\sigma^4$.

The unbiased invariant and uniformly best estimator of the unit dispersion is $\hat{\sigma}^2(Y_{(24)}) = U'V^{-1}U/6$ and its dispersion under the condition $U \sim N_6(0, \sigma^2 V)$ reads $Var[\hat{\sigma}^2(Y)|\sigma^2] = (2\sigma^4/6^2)Tr(I_{(6,6)}) = \sigma^4/3$.

The efficiency of the Ferrero formula is

$$\frac{Var[\hat{\sigma}^2(Y)|\sigma^2]}{Var[\tilde{\sigma}^2(Y)|\sigma^2]} = 0.818.$$

The Ferrero estimator $\tilde{\sigma}\sqrt{2} = \sqrt{2\tilde{\sigma}^2}$ of the standard deviation $\sqrt{2}\sigma$ in the measurement of an angle is obviously biased; with respect to the Jensen inequality, cf. [7], p. 46,

$$\sqrt{E(\xi^2)} > E(|\xi|) > 0$$

for any regular random variable ξ.

If $\xi^2 = U'U/(6 \times 3)$, then obviously $E(\sqrt{2}\tilde{\sigma}|\sigma) < \sqrt{2}\sigma$.

In the case of normality of the random vector U the efficient estimator $\widehat{2\sigma^2} = U'V^{-1}U/3$ is distributed as follows $\widehat{2\sigma^2} \sim 2\sigma^2\chi_6^2/6$.

If $Y = \widehat{2\sigma^2}$, then the probability density of Y is

$$f(y|\sigma^2) = \frac{27}{16\sigma^6}y^2 \exp\left(-\frac{3y}{2\sigma^2}\right), \quad y > 0,$$

and the probability density of $Z = \sqrt{Y} = \hat{\sigma}\sqrt{2}$ is

$$g(z|\sigma) = \frac{27}{8\sigma^6}z^5 \exp\left(-\frac{3z^2}{2\sigma^2}\right), \quad z > 0.$$

Thus

$$E(\hat{\sigma}\sqrt{2}|\sigma) = \int_0^\infty zg(z|\sigma)dz = \sqrt{2}\sigma\frac{15}{16}\sqrt{\frac{\pi}{3}} = 0.959\sqrt{2}\sigma.$$

For the estimator of σ it is obviously $E(\hat{\sigma}|\sigma) = 0.959\sigma$. For greater degrees of freedom the bias will be even smaller.

Theorem 2.4 *The efficiency of the Ferrero estimator of σ^2 for a given p in the Scheme 1 is*

$$\frac{Var(\hat{\sigma}^2|\sigma^2)}{Var(\tilde{\sigma}^2|\sigma^2)} =$$
$$= \frac{Var\left\{[U'V^{-1}U/(6p^2)]|\sigma^2\right\}}{Var\left\{[U'U/(36p^2)]|\sigma^2\right\}} = \frac{9p}{12p-1}, \quad p = 1, 2, \ldots.$$

Proof Let $p = 2$. The input observation vector is of the form

$$Y_{(84)} = (Y'_{0(6)}, \ldots, Y'_{6(6)}, Y'_{7(3)}, Y'_{8(4)}, \ldots, Y'_{17(3)}, Y'_{18(4)})', \; Var(Y) = \sigma^2 I_{(84,84)}$$

(the observation vectors at stations with serial numbers 0-6 are 6-dimensional, at stations with serial numbers $2k-1$ are 3-dimensional and at stations with serial numbers $2k$ are 4-dimensional, $k = 4, 5, \ldots, 9$. The 24-dimensional vector U of triangle misclosures is $U_{(24)} = T_{(24,84)}Y_{(84)} - \pi\mathbf{1}_{(84)}$ and $TT' = V$ is of the form

$$V = \left(\begin{array}{cccccccccccc}
6 & -2 & 0 & 0 & 0 & -2 & 0 & -2 & 0 & 0 & 0 & 0 \\
-2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\
0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2 & 0 & 0 & 0 & -2 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 6 & -2 & 0 & 0 & 0 & 0 \\
-2 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 \\
0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\
0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0
\end{array}\right.$$

$$
\left.\begin{array}{rrrrrrrrrrrr}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
-2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
-2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6 & -2\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 6
\end{array}\right)
$$

This matrix can be writen directly using Scheme 1 by the help of the rule: each diagonal element equals 6, elements outside the diagonal are zeros with exception of 60 elements being -2; they are dislocated symmetrically with respect to the diagonal in those rows of the ith column ($i = 1, 2, \dots, 24$), whose serial number is equal to the serial number of triangles in the scheme which has a common side with the ith triangle. Thus $Tr(V^2) = 24 \times 36 + 60 \times 4 = 1\ 104$.

This rule may be generalized for an arbitrary number p of layers of triangles. The reader can easily confirm himself that for an arbitrary positive integer p a network consisting of $1 + 3p(p+1)$ points contains $6p^2$ triangles (cf. also Lemma 2.2); the observation vector Y in this case is $6[4 + (p-1)(3p+4)]$-dimensional. The matrix TT' being of the type $6p^2 \times 6p^2$ in addition to $6p^2$ diagonal elements equalling 6 contains further $6p(3p-1)$ elements equalling -2; $Tr(V^2) = 24p(12p-1)$. □

Remark 2.5 With growing p the value of the ratio

$$Var[\hat{\sigma}^2(Y)|\sigma^2]/Var[\tilde{\sigma}^2(Y)|\sigma^2]$$

tends to 0.750.

Thus the efficiency of the Ferrero formula lies in the interval (0.750, 0.818]. For different p see the efficiency in the following table.

p	1	2	...	10	...	100	...	∞
efficiency	0.818	0.783	...	0.756	...	0.751	...	0.750

Definition 2.6 The Schreiber scheme is called a procedure of measurements of angles in the hexagonal network which requires the measurement of angles in all combinations (cf. [4], p. 299 and 333).

Theorem 2.7 *In the Schreiber scheme the Ferrero formula for a dispersion σ^2 is not unbiased.*

Proof It is sufficient to find such p that $E(\tilde{\sigma}^2|\sigma^2) - \sigma^2 \neq 0$.

Let $p = 1$. The observation vector Y for determining the set of 6 misclosures within the central hexagon when the Schreiber method of measurement is used is of the form

$$Y_{(66)} = (Y'_{0(30)}, Y'_{1(6)}, Y'_{2(6)}, Y'_{3(6)}, Y'_{4(6)}, Y'_{5(6)}, Y'_{6(6)})',\ Var(Y_{(66)}) = \sigma^2 I_{(66,66)}.$$

The input for processing is formed by a 33-dimensional random vector

$$\Delta Y_{(33)} = T_{(33,66)} Y_{(66)} = \begin{pmatrix} T_{0(15,30)} Y_{0(30)} \\ T_{1(3,6)} Y_{1(6)} \\ T_{2(3,6)} Y_{2(6)} \\ T_{3(3,6)} Y_{3(6)} \\ T_{4(3,6)} Y_{4(6)} \\ T_{5(3,6)} Y_{5(6)} \\ T_{6(3,6)} Y_{6(6)} \end{pmatrix},$$

$$var(\Delta Y) = T var(Y) T' = 2\sigma^2 I_{(33,33)},$$

where

$$T_{i(3,6)} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix},\ i = 1, 2, \ldots, 6,$$

the form of the matrix $T_{0(15,30)}$ is analogous.

The unbiased and efficient estimator of the 18-dimensional vector of angles forming 6 misclosures within the central hexagon is

$$\begin{pmatrix} \hat{\Theta}_1(\Delta Y) \\ \hat{\Theta}_2(\Delta Y) \end{pmatrix} = \begin{pmatrix} (A'A)^{-1}A' \\ 0 \end{pmatrix} \Delta Y - \begin{pmatrix} (A'A)^{-1}B'Q_{11} \\ Q_{21} \end{pmatrix} [b + B(A'A)^{-1}\Delta Y]$$

and its variance reads

$$Var \begin{pmatrix} \hat{\Theta}_1(\Delta Y) \\ \hat{\Theta}_2(\Delta Y) \end{pmatrix} =$$
$$= 2\sigma^2 \begin{pmatrix} (A'A)^{-1} - (A'A)^{-1}B'Q_{11}B(A'A)^{-1}, & -(A'A)^{-1}B'Q_{12} \\ -Q_{21}B(A'A)^{-1}, & -Q_{22} \end{pmatrix}$$

(cf. [6], p. 144). Here

$$\Theta_{1(17)} = (\omega'_{0(5)}, \omega'_{1(2)}, \omega'_{2(2)}, \ldots, \omega'_{6(2)})',\ \omega_{0(5)} = (\omega_{201}, \omega_{302}, \omega_{403}, \omega_{504}, \omega_{605})',$$

$\omega_{i(2)}$, $i = 1, 2, \dots, 6$, are angles whose vertex lies at stations with the serial number i, $\Theta_{2(1)} = \omega_{106}$, the matrix $A_{(33,17)}$ consists of diagonal block matrices

$$A_{(33,17)} = \begin{pmatrix} A_{0(15,5)} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{1(3,2)} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{1(3,2)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{1(3,2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{1(3,2)}0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & A_{1(3,2)} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{1(3,2)} \end{pmatrix},$$

$$A_0' = \begin{pmatrix} 1&1&1&1&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&1&1&1&1&1&1&1&0&0&0&0&0&0 \\ 0&0&1&1&1&0&1&1&1&1&1&1&0&0&0 \\ 0&0&0&1&1&0&0&1&1&0&1&1&1&1&0 \\ 0&0&0&0&1&0&0&0&1&0&0&1&0&1&1 \end{pmatrix},$$

$$A_1' = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

$B = (\mathbf{1}'_{(5)}, \mathbf{O}'_{(12)})$, $C = 1$, $b = -2\pi$ (the condition for determining the angle ω_{106} has the form $-2\pi + \mathbf{1}'\omega_0 + \omega_{106} = 0$) and

$$\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{pmatrix} B(A'A)^{-1}B' & C \\ C' & 0 \end{pmatrix}^{-1}.$$

After realizing that $Q_{11} = 0$, $Q_{12} = Q_{21} = 1$, $Q_{22} = -0.3333$,

$$(A_0'A_0)^{-1} = \begin{pmatrix} 0.3333 & -0.1667 & 0 & 0 & 0 \\ -0.1667 & 0.3333 & -0.1667 & 0 & 0 \\ 0 & -0.1667 & 0.3333 & -0.1667 & 0 \\ 0 & 0 & -0.1667 & 0.3333 & -0.1667 \\ 0 & 0 & 0 & -0.1667 & 0.3333 \end{pmatrix},$$

$$(A_1'A_1)^{-1} = \begin{pmatrix} 0.6667 & -0.3333 \\ -0.3333 & 0.6667 \end{pmatrix},$$

and

$$(A'A)^{-1}B'Q_{12} = (0.1667, 0, 0, 0, 0.1667, \mathbf{0}'_{(12)})',$$

we can immediately write the matrix $Var[(\hat{\Theta}_1'(\Delta Y), \hat{\Theta}_2(\Delta Y))']$ in its numerical form. Further

$$(U_1, U_2, U_3, U_4, U_5, U_6)' = K \begin{pmatrix} \hat{\Theta}_1(\Delta Y) \\ \hat{\Theta}_2(\Delta Y) \end{pmatrix} - \mathbf{1}\pi,$$

where

$$K = \begin{pmatrix} 1&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&1&0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&1&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&1 \end{pmatrix}$$

and

$$Var(U_{(6)}) =$$

$$2\sigma^2 \begin{pmatrix} 1.6667 & -0.5000 & 0 & 0 & 0 & -0.5000 \\ -0.5000 & 1.6667 & -0.5000 & 0 & 0 & 0 \\ 0 & -0.5000 & 1.6667 & -0.5000 & 0 & 0 \\ 0 & 0 & -0.5000 & 1.6667 & -0.5000 & 0 \\ 0 & 0 & 0 & -0.5000 & 1.6667 & -0.5000 \\ -0.5000 & 0 & 0 & 0 & -0.5000 & 1.6667 \end{pmatrix} =$$

$$= 2\sigma^2 V.$$

In accordance with the Ferrero formula the estimator of the unit dispersion in a direction is $\tilde{\sigma}^2(Y_{(66)}) = U'U/(6 \times 6)$ and its dispersion is

$$Var(\tilde{\sigma}^2(Y_{(66)})|\sigma^2) = \frac{2}{(6 \times 6)^2} Tr[Var^2(U)|\sigma^2] = 0.1214\sigma^4.$$

This estimator is biased, because of

$$E[\tilde{\sigma}^2(Y_{(66)})|\sigma^2] = 0.556\sigma^2.$$

□

Remark 2.8 The mean square error of the biased estimator $\tilde{\sigma}^2(Y_{(66)})$ from the last relationship is

$$E[(\tilde{\sigma}^2(Y_{(66)}) - \sigma^2)^2|\sigma^2] = 0.431\sigma^4.$$

The unbiased and uniformly best estimator of σ^2 in the same case is

$$\hat{\sigma}^2(Y_{(66)}) = \frac{U'V^{-1}U}{2 \times 6}, \quad Var[\hat{\sigma}^2(Y_{(66)})|\sigma^2] = \frac{1}{3}\sigma^4.$$

Thus

$$0.431\sigma^4 = E[(\tilde{\sigma}^2(Y_{(66)}) - \sigma^2)^2|\sigma^2] \approx Var[\hat{\sigma}^2(Y_{(66)})|\sigma^2] = \frac{1}{3}\sigma^4.$$

This seems to be not a very bad property of the Ferrero estimator in the case of the Schreiber scheme, if only accuracy of one network is characterized.

If more networks are compared with respect to their accuracy by the Ferrero formula, then its bias in the case of the Schreiber scheme can cause problems in an interpretation of the comparison.

Remark 2.9 The vertex method means the measuring each single angle separately. It means that the covariance matrix of the vector U is of the form $6\sigma^2 I$. Thus the Ferrero formula for σ^2 is efficient; when σ is estimated, then a bias, as shown in Example 2.3, must be expected.

3 Some statistical properties of the misclosures histogram

The other method of analysing a network accuracy is based on the misclosure histogram. This is approximated by the Gaussian curve $n(.,\mu,\sigma^2)$, $x \in R^1$, with properly chosen μ and σ^2; the value μ is considered as a measure of some systematic influences (e.g., the horizontal refraction) and σ^2 characterizes the accuracy of the network similarly as (2).

The coincidence of the Gaussian curve and the histogram is checked either by the values of random variables Γ_1 and Γ_2, defined in the following, or by the Pearson statistic (Lemma 1.7).

In the previous section it was shown that in the case of the $n-$tuple of misclosures $U_1,\dots,U_n$ the covariance matrix is not in each case equal to $6\sigma^2 I$ (i.e., the misclosures are not stochastically independent) and thus the assumptions of Lemma 1.7 cannot be satisfied.

What can be said on a statistical behaviour of Γ_1 and Γ_2 and of the Pearson statistic (1) in the case of a set of the triangle misclosures ?

At first the random variables Γ_1 and Γ_2 will be investigated.

Let Y be an n-dimensional random vector with the mean value equal to zero vector. Then

$$\Gamma_1 = \frac{\frac{1}{n}\sum_{i=1}^n Y_i^3}{\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)^{3/2}} \quad \text{and} \quad \Gamma_2 = \frac{\frac{1}{n}\sum_{i=1}^n Y_i^4}{\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)^2} - 3.$$

Let ξ be an n-dimensional random vector, i.e. $\xi = (\xi_1,\dots,\xi_n)'$ and let $\xi_1,\dots,\xi_n$ be stochastically independent and let they have the same distribution function, i.e. ξ is a random sample of the size n. If the mean value of the random vector ξ is equal to zero vector, then random variables Γ_1 and Γ_2, belonging to ξ, are estimators of skewness and kurtosis of the random variable ξ_i, $i = 1,\dots,n$.

In practice the n-dimesional vector of misclosures was considered to be a random sample of some random variable with the mean value equal to zero and with the normal distribution. If the number n is sufficiently large then realizations of Γ_1 and Γ_2 should vary around zero. In fact the random vector of misclosures does not satisfy our assumption, i.e. it is not a vector of random sample. Thus the statistical properties of the random variable Γ_1 and Γ_2 in the case of misclosures must be investigated.

Lemma 3.1 *Let $Y \sim N_n(0,\Sigma)$. Then*

(i)

$$E\left(\frac{1}{n}\sum_{i=1}^n Y_i|0,\Sigma\right) = 0, Var\left(\frac{1}{n}\sum_{i=1}^n Y_i|0,\Sigma\right) = \frac{1}{n^2}\mathbf{1}'\Sigma\mathbf{1},$$

(ii)

$$E\left(\frac{1}{n}\sum_{i=1}^n Y_i^2|0,\Sigma\right) = \frac{1}{n}\sum_{i=1}^n \sigma_{ii},$$

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^2|0,\Sigma\right) = \frac{1}{n^2}\left(2\sum_{i=1}^{n}\sigma_{ii}^2 + 4\sum_{i<j}\sigma_{ij}^2\right),$$

(iii)

$$E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3|0,\Sigma\right) = 0,$$

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3|0,\Sigma\right) = \frac{1}{n^2}\left[15\sum_{i=1}^{n}\sigma_{ii}^3 + 6\sum_{i<j}(2\sigma_{ij}^3 + 3\sigma_{ii}\sigma_{ij}\sigma_{jj})\right],$$

(iv)

$$E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^4|0,\Sigma\right) = \frac{3}{n}\sum_{i=1}^{n}\sigma_{ii}^2,$$

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^4|0,\Sigma\right) = \frac{1}{n^2}\left[96\sum_{i=1}^{n}\sigma_{ii}^4 + 48\sum_{i<j}(\sigma_{ij}^4 + 3\sigma_{ii}\sigma_{ij}^2\sigma_{jj})\right],$$

(v)

$$cov(Y_i^3, Y_j^2|0,\Sigma) = 0, \quad i \neq j, \quad i,j = 1,\ldots,n,$$

(vi)

$$E(Y_i^4 Y_j^2|0,\Sigma) = 3\sigma_{ii}^2\sigma_{jj} + 12\sigma_{ii}\sigma_{ij}^2.$$

Proof The characteristic function of the vector $Y \sim N_n(0,\Sigma)$ is

$$\phi_Y(t) = \exp\left(-\frac{1}{2}t'\Sigma t\right), \quad t \in R^n.$$

With respect to [1], p. 110

$$E(Y_i^k Y_j^l) = \frac{1}{(\sqrt{-1})^{k+l}}\frac{\partial^{k+l}\phi_Y(t)}{\partial t_i^k \partial t_j^l}\Big|_{t=0};$$

further it is valid

$$cov(Y_i^k, Y_j^l|0,\Sigma) = E(Y_i^k, Y_j^l|0,\Sigma) - E(Y_i^k|0,\Sigma)E(Y_j^l|0,\Sigma).$$

The last two relationships are used for $k,l = 1,2,3,4$. □

Corollary 3.2 *Let Y be the vector from Lemma 3.1. Then*

$$cov\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3, \frac{1}{n}\sum_{i=1}^{n} Y_i^2\middle|0,\Sigma\right) = 0,$$

$$cov\left(\frac{1}{n}\sum_{i=1}^{n} Y_i^4, \frac{1}{n}\sum_{i=1}^{n} Y_i^2\middle|0,\Sigma\right) = \frac{1}{n^2}\left(\sum_{i=1}^{n} 12\sigma_{ii}^3 + 12\sum_{i\neq j}\sigma_{ii}\sigma_{ij}^2\right).$$

Proof It is a direct consequence of Lemma 3.1.

Theorem 3.3 *Let Y be the vector from Lemma 3.1 and let*

$$\Gamma_1 = \frac{\frac{1}{n}\sum_{i=1}^n Y_i^3}{\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)^{3/2}} \quad and \quad \Gamma_2 = \frac{\frac{1}{n}\sum_{i=1}^n Y_i^4}{\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)^2} - 3.$$

Then

(i)

$$E(\Gamma_1|0,\Sigma) \approx \frac{E\left(\frac{1}{n}\sum_{i=1}^n Y_i^3\right)}{\left[E\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)\right]^{3/2}} = 0,$$

$$Var(\Gamma_1|0,\Sigma) \approx$$

$$\approx \frac{1}{n}\frac{1}{\left(\frac{\sum_{i=1}^n \sigma_{ii}}{n}\right)^3}\left\{15\left(\frac{\sum_{i=1}^n \sigma_{ii}^3}{n}\right) + 6\sum_{i<j}\left[2\left(\frac{\sigma_{ij}^3}{n}\right) + 3\frac{\sigma_{ii}\sigma_{ij}\sigma_{jj}}{n}\right]\right\}$$

and

(ii)

$$E(\Gamma_2|0,\Sigma) \approx \frac{E\left(\frac{1}{n}\sum_{i=1}^n Y_i^4\right)}{\left[E\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right)\right]^2} - 3 = 3\frac{\sum_{i=1}^n \sigma_{ii}^2/n}{(\sum_{i=1}^n \sigma_{ii}/n)^2} - 3,$$

$$Var(\Gamma_2|0,\Sigma) \approx$$

$$\approx \frac{1}{n}\frac{1}{\left(\frac{\sum_{i=1}^n \sigma_{ii}}{n}\right)^6}\left\{\left(\frac{1}{n}\sum_{i=1}^n \sigma_{ii}\right)^2\left[96\left(\frac{\sum_{i=1}^n \sigma_{ii}^4}{n}\right) + \frac{48}{n}\sum_{i<j}(\sigma_{ij}^4 + 3\sigma_{ii}\sigma_{ij}^2\sigma_{jj})\right]\right.$$

$$-12\left(\frac{\sum_{i=1}^n \sigma_{ii}}{n}\right)\left(\frac{\sum_{i=1}^n \sigma_{ii}^2}{n}\right)\left(\frac{12}{n}\sum_{i=1}^n \sigma_{ii}^3 + \frac{12}{n}\sum_{i\neq j}\sigma_{ii}\sigma_{ij}^2\right) +$$

$$\left. + 36\left(\frac{\sum_{i=1}^n \sigma_{ii}^2}{n}\right)^2\left(\frac{2}{n}\sum_{i=1}^n \sigma_{ii}^2 + \frac{4}{n}\sum_{i<j}\sigma_{ij}^2\right)\right\}.$$

Proof The statements on $E(\Gamma_1|0,\Sigma)$ and $E(\Gamma_2|0,\Sigma)$ are direct consequences of Lemma 3.1.

Let U, V be random variables. If the relationship

$$Var[f(U,V)] \approx A^2Var(U) + 2ABcov(U,V) + B^2Var(V),$$

where

$$A = \frac{\partial f(u,v)}{\partial u}\Big|_{(u,v)=(E(U),E(V))}, B = \frac{\partial f(u,v)}{\partial v}\Big|_{(u,v)=(E(U),E(V))},$$

$f(u,v) = v/u^{3/2}$, and the substitutions $U = \frac{1}{n}\sum_{i=1}^{n} Y_i^2, V = \frac{1}{n}\sum_{i=1}^{n} Y_i^3$, are taken into account, then with respect to Lemma 3.1 and Corollary 3.2 it is easy to obtain the statement on $Var(\Gamma_1|0,\Sigma)$ in (i). The relationship in (ii) can be obtained analogously. □

Corollary 3.4 *Let in Theorem 3.3 $\sigma_{ij} = 0, i,j = 1,\dots,n, i \neq j$. Then*
(i)

$$E(\Gamma_1|0,\Sigma) \approx 0, \quad Var(\Gamma_1|0,\Sigma) \approx \frac{15}{n}\frac{1}{\left(\frac{\sum_{i=1}^{n}\sigma_{ii}}{n}\right)^3}\frac{\sum_{i=1}^{n}\sigma_{ii}^3}{n},$$

(ii)

$$E(\Gamma_2|0,\Sigma) \approx 3\frac{(\sum_{i=1}^{n}\sigma_{ii}^2/n)}{(\sum_{i=1}^{n}\sigma_{ii}/n)^2} - 3,$$

$$Var(\Gamma_2|0,\Sigma) \approx$$

$$\approx \frac{1}{n}\frac{1}{\left(\frac{\sum_{i=1}^{n}\sigma_{ii}}{n}\right)^6}\left[96\left(\frac{\sum_{i=1}^{n}\sigma_{ii}}{n}\right)^2\frac{\sum_{i=1}^{n}\sigma_{ii}^4}{n} - \right.$$

$$\left. -144\frac{\sum_{i=1}^{n}\sigma_{ii}}{n}\frac{\sum_{i=1}^{n}\sigma_{ii}^2}{n}\frac{\sum_{i=1}^{n}\sigma_{ii}^3}{n} + 72\left(\frac{\sum_{i=1}^{n}\sigma_{ii}^2}{n}\right)^2\frac{\sum_{i=1}^{n}\sigma_{ii}^2}{n}\right].$$

Corollary 3.5 *Let in Theorem 3.3 $\sigma_{11} = \ldots = \sigma_{nn}$ and let $\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}$, $i,j = 1,\dots n, i \neq j$. Then*
(i)

$$E(\Gamma_1|0,\Sigma) \approx 0, \quad Var(\Gamma_1|0,\Sigma) \approx \frac{1}{n}\left[15 + \frac{6}{n}\sum_{i<j}(2\rho_{ij}^3 + 3\rho_{ij})\right],$$

(ii)

$$E(\Gamma_2|0,\Sigma) \approx 0, \quad Var(\Gamma_2|0,\Sigma) \approx \frac{1}{n}\left(24 + \frac{48}{n}\sum_{i<j}\rho_{ij}^4\right).$$

Remark 3.6 From Corollary 3.4 and Corollary 3.5 it can be seen that the stochastical dependence characterized by the correlation coefficients $\rho_{i,j}$ does not influence in practice the "skewness" (Γ_1) and the "kurtosis" (Γ_2) significantly when the number n is sufficiently large (i.e. in the case $\sigma_{11} = \ldots = \sigma_{nn}$ the random variables Γ_1 and Γ_2 behave approximately as estimators of skewness and kurtosis). However the heteroscedasticity of the triangle misclosure vector (i.e. $\sigma_{ii} \neq \sigma_{jj}$ for $i \neq j$) can influence both of them in a non-negligible way.

Theorem 3.7 *Let $U \sim N_n(0, \Sigma)$ be a vector of triangle misclosures. If U is realized $r-$times, $r \to \infty$, then the histogram of misclosures can be approximated by the density*

$$h(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{1}{2\sigma_i^2} x^2), \quad x \in (-\infty, \infty),$$

where $\sigma_i^2 = \{\Sigma\}_{ii}$.

Proof The analogy of the empirical distribution function generated by r-times repeated realization $(u_{i1}, \ldots, u_{in})'$ of the vector U is

$$F_{n,r}(a) = \frac{1}{nr} \sum_{i=1}^{n} \sum_{j=1}^{r} I_{ij}(-\infty, a),$$

where

$$I_{ij}(-\infty, a) = \begin{cases} 1 \;\; if \;\; u_{ij} \;<\; a, \\ 0 \;\; if \;\; u_{ij} \;\geq\; a. \end{cases} .$$

With respect to Gliwenko theorem [1], p. 327

$$\lim_{r \to \infty} \frac{1}{r} \sum_{j=1}^{r} I_{ij}(-\infty, a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{1}{2\sigma_i^2} x^2) dx;$$

thus

$$F_{r,n}(a) \to \frac{1}{n} \sum_{i=1}^{n} \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{1}{2\sigma_i^2} x^2) dx, \quad a \in (-\infty, \infty).$$

□

Remark 3.8 If (cf. [7], p.86) the series

$$\sum_{j=1}^{\infty} \frac{\alpha_j}{j!} t^j \tag{3}$$

is absolutely convergent for some $t > 0$, then the sequence of moments $\alpha_1, \alpha_2, \ldots$ defines a unique distribution function.

Let

$$\alpha_j = E\left(\frac{1}{n} \sum_{k=1}^{n} Y_k^j\right), \quad j = 1, 2, \ldots.$$

If there exists such real numbers $a, b, a < b$ that $\sigma_{ii} \in [a, b], i = 1, \ldots, n$, then obviously (3) is absolutely convergent. As α_j is the $j-$th moment of the function $h(.)$ from Theorem 3.7, this can be considered as another reason why the misclosure histogram can be approximed by the function $h(.)$.

In general, the difference between the function $h(.)$ and the histogram constructed from the vector $U \sim N_n(0, \Sigma)$ can be formidable; e.g., if $\Sigma = \frac{1}{n}\mathbf{1}\mathbf{1}'$, then the histogram consists of a single point and $Var(\Gamma_1|0, \Sigma) = 15, Var(\Gamma_2|0, \Sigma) = 24$. However in practice such a great influence of the non-diagonal elements of the matrix Σ cannot be expected.

Remark 3.9 In order to show the influence of the discrepancy between the vectors $p = (p_1, \dots, p_k)'$ and $p^* = (p_1^*, \dots, p_k^*)'$ in the Pearson statistic (1) let us consider the following.

The mean value of the statistic $\sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}$ is

$$E\left[\sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}\Big|p^*\right] = \sum_{i=1}^{k} \frac{p_i^*(1-p_i^*)}{p_i} + n\sum_{i=1}^{k} \frac{(p_i^* - p_i)^2}{p_i}.$$

If $p^* \neq p$, then for sufficiently large n obviously

$$\sum_{i=1}^{k} \frac{p_i^*(1-p_i^*)}{p_i} + n\sum_{i=1}^{k} \frac{(p_i^* - p_i)^2}{p_i} >> k-1$$

and thus the standardly used statistic varies not arround the number $k-1$ (the mean value of the random variable χ^2_{k-1}); in such a case the test based on the misclosure histogram has to refuse the hypothesis on the normality of the set of triangle misclosures.

4 Conclusion

On the basis of the statements and the simple examples given in Section 2 it can be concluded that even the Ferrero estimator is not, in general, the efficient estimator of σ^2 in a homogeneous triangulation network and sometimes (the Schreiber scheme) gives the biased estimates, its efficiency is not so bad. It gives a realistic information on σ^2 in the form of a preliminary estimator. If σ is to be estimated, then some caution is useful because of the bias.

As far as the misclosure histogram is concerned it can be stated, on the basis of Section 3, that it could be a good tool for investigation of the accuracy of the network in the case $Var(U) = \sigma^2 W, W_{11} = \dots = W_{nn}$ only. If the last relationship is not fulfilled, then the use of the skewness Γ_1 and the kurtosis Γ_2 or the Pearson statistic leads to the refusing of the hypothesis on the normality of the vector U for a sufficiently large n. This can explain the fact that sometimes, in the past, the triangulation network was considered as affected by systematic effects, however a mixture of several normally distributed sets of misclosures could be actual reason of the refusing of the hypothesis on the normality.

References

[1] Fisz, M.: *Wahrscheinlichkeitsrechnung und mathematische Statistik. VEB Deutscher Verlag der Wissenschaften.* Berlin, 1966.

[2] Hirvonen, R. A.: *Adjustment by Least Squares in Geodesy and Photogrammetry.* Frederic Ungar Publishing Co., New York, 1971.

[3] Humak, K. M. S.: *Statistische Methoden der Modellbildungen III.* Akademie–Verlag, Berlin, 1984.

[4] Jordan/Eggert/Kneissl: *Handbuch der Vermessungskunde, Band IV, Mathematische Geodäsie, Erste Hälfte.* Metzlersche Verlagsbuchhandlung, Stuttgart, 1958.

[5] Kubáček, L.: *Foundations of Estimation Theory.* Amsterdam–Oxford–New York–Tokyo, Elsevier, 1988.

[6] Kubáčková, L.: *Foundations of Experimental Data Analysis.* CRC Press, Boca Raton–Ann Arbor–London–Tokyo, 1992.

[7] Rao, C. R.: *Linear Statistical Inference and Its Applications.* John Wiley, New York, 1965.

[8] Rao, C. R., Mitra, K. S.: *Generalized Inverse of Matrices and Its Application.* J. Wiley, New York 1971.

[9] Wolf, H.: *Ausgleichungsrechnung Formeln zur praktischen Anwendungen.* Ferd. Dümmlers Verlag, Bonn, 1975, p. 163.