www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529–542.

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

M&MS

STANDARD DEVIATION OF THE MEAN OF AUTOCORRELATED OBSERVATIONS ESTIMATED WITH THE USE OF THE AUTOCORRELATION FUNCTION ESTIMATED FROM THE DATA

Andrzej Zięba, Piotr Ramza

AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Mickiewicza 30, 30-059 Krakow, Poland
(✉ *Andrzej.Zieba@fis.agh.edu.pl*, +48 12 617 3551)

Abstract

Prior knowledge of the autocorrelation function (ACF) enables an application of analytical formalism for the unbiased estimators of variance s_a^2 and variance of the mean $s_a^2(\bar{x})$. Both can be expressed with the use of so-called effective number of observations n_{eff}. We show how to adopt this formalism if only an estimate $\{r_k\}$ of the ACF derived from a sample is available. A novel method is introduced based on truncation of the $\{r_k\}$ function at the point of its *first transit* through *zero* (FTZ). It can be applied to non-negative ACFs with a correlation range smaller than the sample size. Contrary to the other methods described in literature, the FTZ method assures the finite range $1 < \hat{n}_{eff} \leq n$ for any data. The effect of replacement of the standard estimator of the ACF by three alternative estimators is also investigated. Monte Carlo simulations, concerning the bias and dispersion of resulting estimators s_a and $s_a(\bar{x})$, suggest that the presented formalism can be effectively used to determine a measurement uncertainty. The described method is illustrated with the exemplary analysis of autocorrelated variations of the intensity of an X-ray beam diffracted from a powder sample, known as the particle statistics effect.

Keywords: autocorrelated data, time series, effective number of observations, estimators of variance, measurement uncertainty.

1. Introduction

An analysis of series of observations represents perhaps the most common procedure in applied statistics. The well-known formulae for the mean $\bar{x}$ and the estimators of variance s^2 and variance of the mean $s^2(\bar{x})$ are unbiased and with maximum efficiency when the observations $\{x_i\}$ are independent (uncorrelated), equivalent, and normally distributed.

Let us now assume that the observations are *autocorrelated* while the two remaining assumptions remain unchanged. The equivalence of the observations x_i implies that the expected value μ and variance σ^2 are the same for all x_i, whereas the correlation coefficient relating x_i and x_j depends only on the difference $|i - j|$ [1]. Consequently, the correlations are fully specified by the discrete one-dimensional autocorrelation function (ACF) denoted as $\{\rho_k\}$, $k = 0, 1, \ldots, n - 1$. Alternatively, such data can be considered as a finite sample obtained from a *stationary* time series or as an effect of sampling of a *stationary* stochastic process.

When the autocorrelation function $\{\rho_k\}$ is known, it is possible to develop an analytical formalism in order to derive the mean and the estimators of variance. The arithmetic mean $\bar{x}$ remains to be the unbiased estimator of the expected value, but it is no more the *best linear*

Article history: received on Sept. 11, 2011; accepted on Nov. 23, 2011; available online on Dec. 12, 2011;
DOI: 10.2478/v10178-011-0052-x.

unbiased estimator (BLUE). Its common use is justified because it remains asymptotically BLUE (in the limit $n \to \infty$) and the loss of efficiency for a finite sample size is small [2, 3]. The resulting unbiased estimators of variance and variance of the mean are, respectively, given by

$$s_a^2 = \frac{n_{eff}}{n(n_{eff}-1)} \sum_{i=1}^{n} (x_i - x)^2 \tag{1}$$

and

$$s_a^2(\bar{x}) = \frac{s_a^2}{n_{eff}}. \tag{2}$$

Both estimators are expressed as functions of the *effective number of observations*,

$$n_{eff} = \frac{n}{1 + 2\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k}, \tag{3}$$

which depends on the number of observations n and elements ρ_k of the autocorrelation function. The dispersion of variance and standard deviation estimators depends on a quantity

$$\nu_{eff} \cong \frac{n}{1 + 2\sum_{k=1}^{n-1} \rho_k^2} - 1, \tag{4}$$

termed the *effective degree of freedom*. Equations (1)–(4) and the above concept of 'effective' numbers were first introduced by [4] and presented independently in several other works (as reviewed in [1]).

This formalism can be directly used when the autocorrelation function is known from other sources of information. This assumption occurs quite frequently in practice. Nevertheless, the implementation of this method should also involve the estimation of quantities (1)–(4) when only an estimate $\{r_k\}$ of the ACF, calculated from the analyzed data $\{x_i\}$ is available. This could not be done by merely replacing $\{\rho_k\}$ by $\{r_k\}$ in Eqs. (3)–(4). The main objective of this work is to propose suitable algorithms and to investigate their properties by using both analytical methods and *Monte Carlo* (MC) simulations.

Of course, the use of n_{eff} and ν_{eff} is not a necessity and, in fact, is not widely used. It allows to express the formulae for both variances in a compact manner which reveals the similarities and differences with respect to the case of independent observations. Regardless of the notation, the stochastic properties of estimators of variance (1) and (2) can be analyzed bearing in mind the fact that they are the products of the sum of squares (with the stochastic properties defined by ν_{eff}) and a multiplicative factor depending on the autocorrelation function and the sample size. When the autocorrelation function is known this factor is fixed, whereas it becomes a random number when the estimator of the ACF is used.

2. Estimating the effective number of observations: truncating the tail of the autocorrelation function estimate

The estimators $\hat{n}_{eff}$ investigated in this paper are derived from (3) defining n_{eff}. This formula has been modified by truncating the summation in (4) at the *limiting lag* n_c, which should be smaller than $n-1$. The resulting formula reads

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

$$\hat{n}_{eff} = \frac{n}{1+2\sum_{k=1}^{n_c}\left(1-\frac{k}{n}\right)r_k}. \tag{5}$$

The effect that the choice of n_c may have, is investigated in this section assuming the use of just one standard estimator $\{r_k\}$ of the ACF. The effect of the replacement of $\{r_k\}$ with other estimators of the ACF will be discussed in Sec. 3. Quantitative data for addressing both these issues is obtained by means of Monte Carlo simulation. The results of simulations presented in Figs. 2 and 3 concern in fact the variable $1/\hat{n}_{eff}$ instead of $\hat{n}_{eff}$ (i) to cover conveniently the case when $\hat{n}_{eff}$ goes to infinity or becomes negative, and (ii) because $1/\hat{n}_{eff}$ is a proportionality coefficient in relation (2) between $s_a^2(\bar{x})$ and s_a^2.

2.1. The specifics of Monte Carlo simulation

Two different models of the stationary time series are used to generate autocorrelated numbers. The *simple moving average* (SMA) is defined as the arithmetic mean

$$x_i = \frac{u_i + u_{i-1} + \ldots + u_{i-m}}{m} \tag{6}$$

of m successive uncorrelated Gaussian numbers $\{u_i\}$. The *first-order autoregressive model* (AR(1)) is defined by

$$x_i = a\,x_{i-1} + u_i \tag{7}$$

with the parameter $0 < a < 1$.

Both models belong to two general categories: *moving average* (MA) and *autoregressive* (AR) time series. In the present study, we set $m = 5$ for the simple moving average model, whereas parameter a of the autoregressive AR(1) model was adjusted to obtain the same theoretical n_{eff} that was calculated using the SMA for the given sample size. The autocorrelation function for both models is nonnegative ($\rho_k \geq 0$) and of finite range. The difference is that, in the case of the SMA, the function $\{\rho_k\}$ equals zero above a well-defined k, whereas for the AR(1) model, it approaches zero asymptotically. In both cases, though, the autocorrelation function becomes practically zero for k significantly smaller than the sample size n.

The sample size chosen for the modelling was $n = 15$, 60, and 240 for the following reasons. In the textbook [5] (p. 32) it is proposed that 'in practice, to obtain a useful estimate of the ACF, we would need at least 50 observations'. A sample of $n = 60$ elements, which is slightly larger than that limit, was chosen. The simulations for the four times larger value of $n = 240$ provide us with information concerning the convergence of estimators, whereas the results for $n = 15$ allow us to check the case of a small sample. The total number of MC duplicates used was $N_{MC} = 250{,}000$ for each case. This assures two significant decimal digits for estimators' values derived through the simulations.

2.2. Standard estimator of the autocorrelation function

The estimator

$$r_k = \frac{\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{n}(x_i - x)^2} \tag{8}$$

is named *standard* because it is the most frequently discussed in literature one and widely implemented in computer programs. An important reason of its widespread use is that the correlation matrix composed of the elements r_k remains positive definite.

Example estimates $\{r_k\}$ are shown in Fig. 1. These discrete functions seem to be rich in details but only a limited number of initial points corresponding to non-zero values of the autocorrelation function $\{\rho_k\}$ are relevant. The remaining points, the so-called *tail*, represents merely autocorrelated noise, which is further characterized by dispersion $s(r_k)$. The formula for the variance in the tail was given by [5] (p. 33). Utilising the concept of effective degrees of freedom (4) , it can also be expressed as $s(r_k) \cong (\nu_{eff}+1)^{-1/2}$. This explains the rather large amplitude fluctuations in the tail of the $\{r_k\}$ function for a sample of finite size.

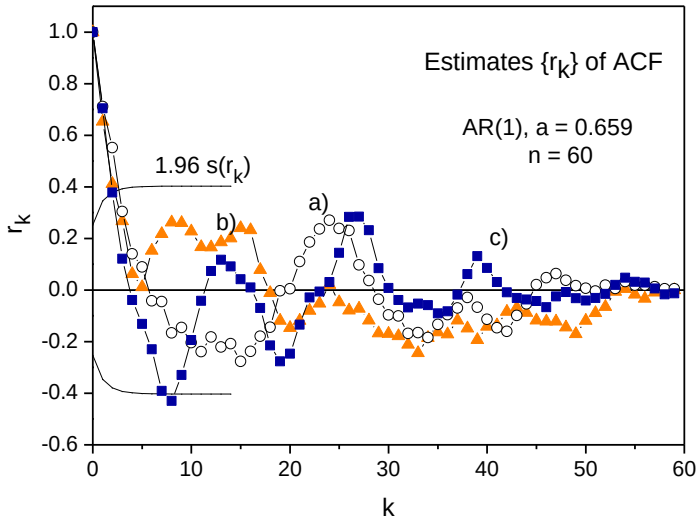

Fig. 1. Exemplary estimates a), b) and c) of the autocorrelation function calculated from three 60-element samples generated with the use of the AR(1) model. Solid lines denote the function $\pm 1.96 s(r_k)$ used in the LSN method (pt. 2.4).

2.3. The failure of first attempts to estimate n_{eff} from an autocorrelation function estimate

The investigation into the effects that the choice of n_c may have, can start by taking into account all elements of $\{r_k\}$, *i. e.*, $n_c = n - 1$ in (5). This option assures that $\hat{n}_{eff} > 0$ because the denominator of (5) represents the sum of all elements of the correlation matrix, which is positive definite for this specific estimator of the ACF. The corresponding MC simulated probability density function $g(1/\hat{n}_{eff})$ is represented by curve a) of Fig. 2. One can see the large negative bias: almost the whole area of function $g(1/\hat{n}_{eff})$ is located leftwards with respect to the theoretical value of $1/n_{eff}$.

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

The rule of thumb has been expressed in the literature [5] that reasonably good estimates of the ACF can be obtained only for $k < n/4$. One may try to estimate the effective number of observations using a limiting lag of $n_c = n/4$ (curve b) in Fig. 2). However, this estimator is characterised by the largest dispersion. Even more troublesome is the quite large probability of $1/\hat{n}_{eff} < 0$ leading to $s_a^2(\bar{x}) < 0$. The negative $\hat{n}_{eff}$ can occur because the correlation matrix composed of the truncated $\{r_k\}$ estimate ($r_k = 0$ for $k > n_c$) has not to be positive-definite. Hence, the option of n_c being equal to $n/4$ cannot be accepted nor another fixed fraction of the sample size.

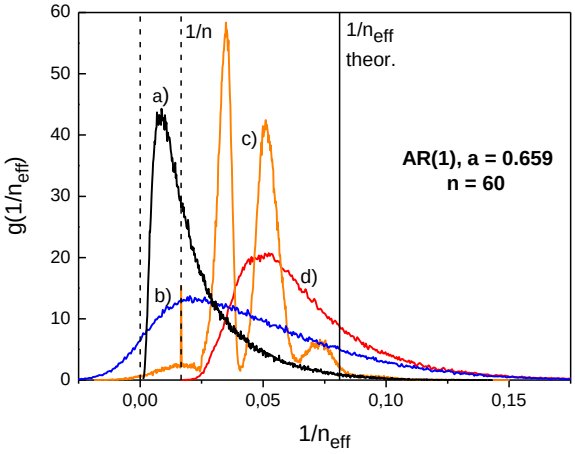

Fig. 2. MC probability density functions for different approaches to estimate the effective number of observations shown for $1/\hat{n}_{eff}$ variable.The presented curves correspond to: a) $n_c = n - 1$, b) $n_c = ¼\ n$, c) the LSN method, and d) the FTZ method (as described in Sec. 2.3 – 2.5). Vertical lines indicate: $1/\hat{n}_{eff} = 0$, $1/\hat{n}_{eff} = 1/n$, and the theoretical value $1/n_{eff} = 1/12.33$.

2.4. Limiting lag determined by last significant nonzero element of the autocorrelation function estimate (LSN method)

Statistical fluctuations in the tail seem to be the main source of the undesirable properties of the two estimators outlined above. Zhang [6] was the first to introduce a well-defined statistical procedure to address this issue by introducing the limiting lag n_c, corresponding to the *last significant nonzero* value of $\{r_k\}$ (LSN method). The algorithm defining n_c is given by a set of formulae

$$n_c = \max\left\{k \,\middle|\, |r_k| > 1.96\, s(r_k)\right\}, \tag{9a}$$

$$s(r_1) = \frac{1}{\sqrt{n}} \text{ and } s(r_k) = \sqrt{\frac{1 + 2\sum_{j=1}^{k-1} r_j^2}{n}} \text{ for } k \geq 2, \tag{9b}$$

$$n_c = \min\{n_c, n/4\}. \tag{9c}$$

It can be described as follows:

(a) n_c is first defined as the maximum value of k for which $|r_k| > 1.96\ s(r_k)$ (a confidence level of 0.95 is assumed).
(b) Eq. (9b) defines $s(r_k)$ for $k = 1$ and for $k \geq 2$. For further details see [6].
(c) if the value of n_c thus obtained is larger than $n/4$, it is fixed at $n_c = n/4$.

Two features of the method can be qualitatively understood by visual inspection of Fig. 1 and curve c) in Fig. 2. First, the limit of $1.96\, s(r_k)$ intersects with the estimate $\{r_k\}$ at a relatively high level, leading in most cases to excessively low values of n_c. Underestimation of n_c, in turn, leads to a lower value of $1/\hat{n}_{eff}$ and to the peculiar shape of $g(1/\hat{n}_{eff})$ with its successive maxima corresponding, from left to right, to n_c = 1, 2, 3, and 4. The finite probability of obtaining $n_c = 0$ leads to a small sharp peak for $1/\hat{n}_{eff} = 1/n$. Furthermore, there remains a finite probability that fluctuations in the tail may exceed the limit of $\pm 1.96 s(r_k)$, as illustrated by curve c) in Fig. 1. As a result, a small fraction of $\hat{n}_{eff}$ can take values larger than the number of observations or even become negative.

Selected statistical parameters of the quantity $1/\hat{n}_{eff}$ for two time series models and samples of three different sizes are summarized in Table 1.

Table 1. The statistical parameters of the quantity $1/\hat{n}_{eff}$, which have been obtained using two different methods of estimating the effective number of observations based on MC simulations. These correspond to two time series models and samples of three different sizes n.

		SMA model			**AR(1) model**		
Model parameter		m = 5			a = 0.634	a = 0.659	a = 0.665
n		15	60	240	15	60	240
n_{eff} theoretical		3.36	12.33	48.32	3.36	12.33	48.32
$1/\hat{n}_{eff}$ LSN method	$bias_r$	–0.58	–0.28	–0.16	–0.69	–0.44	–0.26
	s_r	0.14	0.21	0.32	0.13	0.19	0.31
	$P(1/\hat{n}_{eff} < 1/n_{eff})$	1.00	0.92	0.74	0.99	0.98	0.84
$1/\hat{n}_{eff}$ FTZ method	$bias_r$	–0.38	–0.04	0.08	–0.53	–0.19	0.01
	s_r	0.19	0.31	0.32	0.17	0.33	0.36
	$P(1/\hat{n}_{eff} < 1/n_{eff})$	0.98	0.68	0.55	0.99	0.78	0.61

The relative bias and relative dispersion are defined as

$$bias_r(\hat{\theta}) = \frac{\mathrm{E}(\hat{\theta}) - \theta}{\theta} \quad \text{and} \quad s_r(\hat{\theta}) = \frac{\sqrt{\mathrm{Var}(\hat{\theta})}}{\theta}. \tag{10}$$

Symbols θ and $\hat{\theta}$ denote the statistical parameter and its estimator. In addition, the table provides the quantity $P(1/\hat{n}_{eff} < 1/n_{eff})$ indicating the probability that the estimate is lower than the true value of $1/n_{eff}$.

2.5. Limiting lag determined based on the first transit of the autocorrelation function estimate through zero (FTZ method)

In order to truncate the tail of the ACF estimate, one can subjectively decide where the borderline is between the significant nonzero values of the $\{r_k\}$ function and the tail. To make this process more objective we propose to define the cut-off lag n_c as corresponding to the last positive value of the $\{r_k\}$ function before its first transit through zero,

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

$$n_c = \min\left\{k \,\middle|\, (r_k > 0 \ \wedge \ r_{k+1} < 0)\right\}. \tag{11}$$

The value $n_c < n - 1$ is found in every data set, even when all ρ_k are nonnegative, because of two reasons. First, negative r_k values will appear due to statistical fluctuations. More precise is another argument. Percival [7] has proved that for another estimator of ACF,

$$r_k^* = \frac{\frac{1}{n-k}\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2}, \tag{12}$$

the sum of all its elements equals zero,

$$n + 2\sum_{k=1}^{n-1}(n-k)\, r_k^* = 0\,. \tag{13}$$

To fulfill this condition at least one element r_k^* has to be negative. It follows from (8) and (12) that $r_k^* = [n/(n-k)]r_k$. Hence at least one r_k is also negative and (11) defines $n_c < n - 1$ for any data.

FTZ method assures $1 < \hat{n}_{eff} \le n$ since the effective number of observations is calculated exclusively based on the positive values of r_k. The case $\hat{n}_{eff} = n$ is possible because of a finite probability that $r_1 < 0$ for nonnegative ACF. The inspection of Fig. 2 suggests that this probability is usually insignificantly small.

The range of allowed values of $\hat{n}_{eff}$ for the FTZ method, $1 < \hat{n}_{eff} \le n$, is finite, contrary to three other options discussed in pt. 2.3 and 2.4. When all elements of $\{r_k\}$ are used in (5) $\hat{n}_{eff}$ can take values from the range $1 < \hat{n}_{eff} < \infty$. For the LSN method, or when a fixed $n_c < n$ is used in (5), this range is even wider, $-\infty < \hat{n}_{eff} < \infty$.

The results of MC modelling based on this FTZ method are presented in Table 1 and Fig. 2. When compared to the results obtained with the LSN method, one can observe that anomalies of $g(1/\hat{n}_{eff})$ are eliminated and the bias is reduced. The entire procedure is simpler than the LSN method and independent of the choice of a tuning parameter (the factor 1.96 in (9a)).

However, the FTZ method may only be applied when all the ρ_k elements of the ACF are nonnegative. This is the case in the majority of experimental situations. On the contrary, the LSN method remains more general because it can be applied regardless of the correlation coefficients' sign. Both methods could be interpreted as using an estimator of the ACF composed of two parts, of the nonzero r_k for $k \le n_c$ and of $r_k = 0$ for $k > n_c$.

The inspection of Table 1 also reveals that, assuming the same value of n and n_{eff}, both bias and dispersion are larger in the case of the AR(1) model when compared to the SMA. The function $\{\rho_k\}$ gradually approaches zero, meaning the uncertainty of the determination of limiting lag n_c is higher than with the SMA method. In the following section, our investigations are narrowed down to the more difficult AR(1) case, which is also more frequently encountered in experiments.

3. Estimating the effective number of observations: alternative estimators of the autocorrelation function

In the search for the potentially most accurate estimator of the effective number of observations, we first concentrated on methods of eliminating the fluctuating tail (Section 2).

In this section, we will check the possibility of using alternative estimators of the ACF that are less biased.

There are two sources of bias for the elements of $\{r_k\}$ function. The first is the unequal numbers of terms in sums defining the numerator and denominator of Eq. (8). (These are equal to $n - k$, and n, respectively). To correct this bias, one may use an alternative estimator of the ACF, defined by (12). The replacement of r_k in (5) *with* $r_k^* = [n/(n-k)]r_k$ leads to a simple formula

$$\hat{n}_{eff}^* = \frac{n}{1 + 2\sum_{k=1}^{n_c} r_k}. \tag{14}$$

As stressed by Percival [7], using $\{r_k^*\}$ does not assure a substantial reduction in the bias because the more important source of the one is application the sample mean $\bar{x}$ instead of the expected value μ in both (8) and (12). An estimator of the ACF that is unbiased up to $O(n^{-2})$ with respect to both sources of bias was introduced by Quenouille ([8], see also [9]). To obtain this estimator, one must calculate $\{r_k\}$ for the whole sample (assumed its size to be even) and separately for its two halves, leading to two estimates $r_k^{(1)}$ and $r_k^{(2)}$. The combination $r_k^{(Q)} = 2r_k - (r_k^{(1)} + r_k^{(2)})/2$ defines a new estimator of the ACF. The resulting estimator of the effective number of observations reads

$$\hat{n}_{eff}^{(Q)} = \frac{n}{1 + 2\sum_{k=1}^{n_c}\left(1 - \frac{k}{n}\right)\left(2r_k - \frac{r_k^{(1)} + r_k^{(2)}}{2}\right)}. \tag{15}$$

Another timator $\{r_k^+\}$ aiming to reduce the bias of the standard estimator of $\{r_k\}$ is proposed by the authors [10]. The resulting *bias-reduced* estimator of the effective number of observations is given by

$$\hat{n}_{eff}^+ = \frac{n - 2n_c - 1 + n_c(n_c+1)/n}{1 + 2\sum_{k=1}^{n_c} r_k} + 1. \tag{16}$$

Contrary to (15) it is defined solely by standard $\{r_k\}$ function and the numbers n and n_c.

It appears that the use of different estimators of the ACF has a moderate influence on its shape, which remains qualitatively similar and markedly non-symmetric (Fig. 3). This is in contrast to the dramatic effect that the choice of n_c has (Fig. 2).

A rather small difference between the estimators $\{r_k\}$ and $\{r_k^*\}$ result in the fact that the functions $g(1/\hat{n}_{eff})$ and $g(1/\hat{n}_{eff}^*)$ are quite close to each other. A more substantial transfer of probability density from the region below the theoretical value $1/n_{eff}$ to the region above this limit is evident in the case of estimators $1/\hat{n}_{eff}^+$ and $1/\hat{n}_{eff}^{(Q)}$.

The value of selected statistical parameters for all four estimators are summarized in Table 2, where one observes a reduction in bias and improvement in $P(1/\hat{n}_{eff} < 1/n_{eff})$. Inevitably, this occurs at the cost of increased dispersion.

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

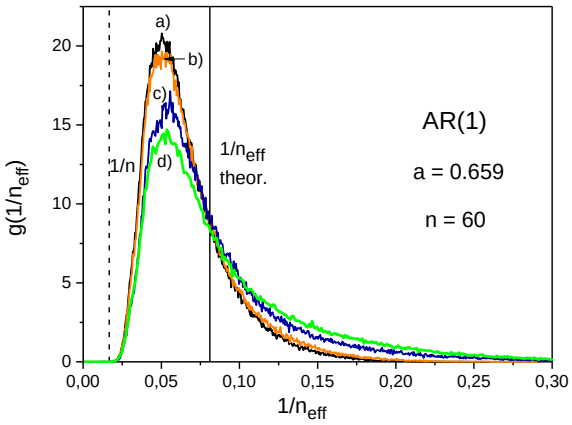

Fig. 3. The probability density function for four estimators of the inverse effective numbers of observations, defined using the same FTZ method and different estimators of the ACF: a) $\{r_k\}$, b) $\{r_k^*\}$, c) $\{r_k^+\}$ and d) $\{r_k^{(Q)}\}$.

4. Estimators of the standard deviation and the standard deviation of the mean

Estimators of standard deviation s_a and standard deviation of the mean $s_a(\bar{x})$ are defined by Eqs. (1) and (2) when parameter n_{eff} is replaced by one of its estimators. Our investigations are focused on only two of these estimators, namely, $\hat{n}_{eff}$ calculated using the standard estimator of the ACF and $\hat{n}_{eff}^+$, which is defined by using the same quantities $\{r_k\}$ and n_c but ensures an appreciable reduction in the bias. In both cases, the limiting lag n_c defined by means of the FTZ method is used.

The statistical properties of the estimators of standard deviation and standard deviation of the mean can be conveniently characterised using the dimensionless ratios s_a/σ and $s_a(\bar{x})/\sigma(\bar{x})$ (or s_a^+/σ and $s_a^+(\bar{x})/\sigma(\bar{x})$). Selected probability density functions and statistical parameters obtained by means of the MC simulation method are shown in Fig. 4 and summarized in Table 3.

4.1. Estimators of standard deviation

In the case of uncorrelated observations, the statistical properties of the relative estimators s/σ and $s(\bar{x})/\sigma(\bar{x})$ are described by the same statistical variable z_ν which is related through equation $z_\nu = \sqrt{\chi_\upsilon^2/\nu}$ to the chi-square variable χ_ν^2 where the degrees of freedom are equal to $\nu = n - 1$. Its probability density function is

$$g(z_\nu) = \frac{2^{1-\nu/2}\nu^{\nu/2}}{\Gamma(\nu/2)} z^{\nu-1} \exp\left(-\nu z^2/2\right), \tag{17}$$

where symbol Γ denotes the Euler gamma function.

Table 2. The properties of four estimators of the effective number of observations, estimated based on by (5), (14), (15), and (16) and the same FTZ method of determining n_c. The parameters of the AR(1) model and the sample size are the same as in Table 1.

		AR(1) model		
Model parameter		$a = 0.634$	$a = 0.659$	$a = 0.665$
n		15	60	240
$1/\hat{n}_{eff}$	$bias_r$	−0.53	−0.19	0.01
	s_r	0.17	0.33	0.36
	$P(1/\hat{n}_{eff} < 1/n_{eff})$	0.99	0.78	0.61
$1/\hat{n}^{*}_{eff}$	$bias_r$	−0.50	−0.16	0.02
	s_r	0.19	0.36	0.38
	$P(1/\hat{n}^{*}_{eff} < 1/n_{eff})$	0.99	0.75	0.60
$1/\hat{n}^{+}_{eff}$	$bias_r$	−0.38	0.00	0.11
	s_r	0.29	0.55	0.51
	$P(1/\hat{n}^{+}_{eff} < 1/n_{eff})$	0.88	0.64	0.53
$1/\hat{n}^{(Q)}_{eff}$	$bias_r$	−0.31	0.11	0.20
	s_r	0.37	0.67	0.69
	$P(1/\hat{n}^{(Q)}_{eff} < 1/n_{eff})$	0.81	0.58	0.50

Table 3. Selected parameters of the estimators of relative standard deviations for the same models of autocorrelated data as in Table 2.

MODEL	AR(1) $A = 0.634$	AR(1) $A = 0.659$	AR(1) $A = 0.665$
n	15	60	240
theoretical ν_{eff}	5.4	22.7	91.8
$(2\nu_{eff})^{-1/2}$	0.304	0.149	0.073
$bias_r(s_a)$	−0.11	−0.02	0.00
$s_r(s_a)$	0.25	0.14	0.07
$bias_r[s_a(\bar{x})]$	−0.38	−0.12	0.00
$s_r[s_a(\bar{x})]$	0.26	0.28	0.21
$bias_r[s_a^{+}x)]$	−0.08	−0.01	0.00
$s_r[s_a^{+}(x)]$	0.27	0.15	0.07
$bias_r[s_a^{+}(\bar{x})]$	−0.27	−0.02	0.04
$s_r[s_a^{+}(\bar{x})]$	0.36	0.37	0.25

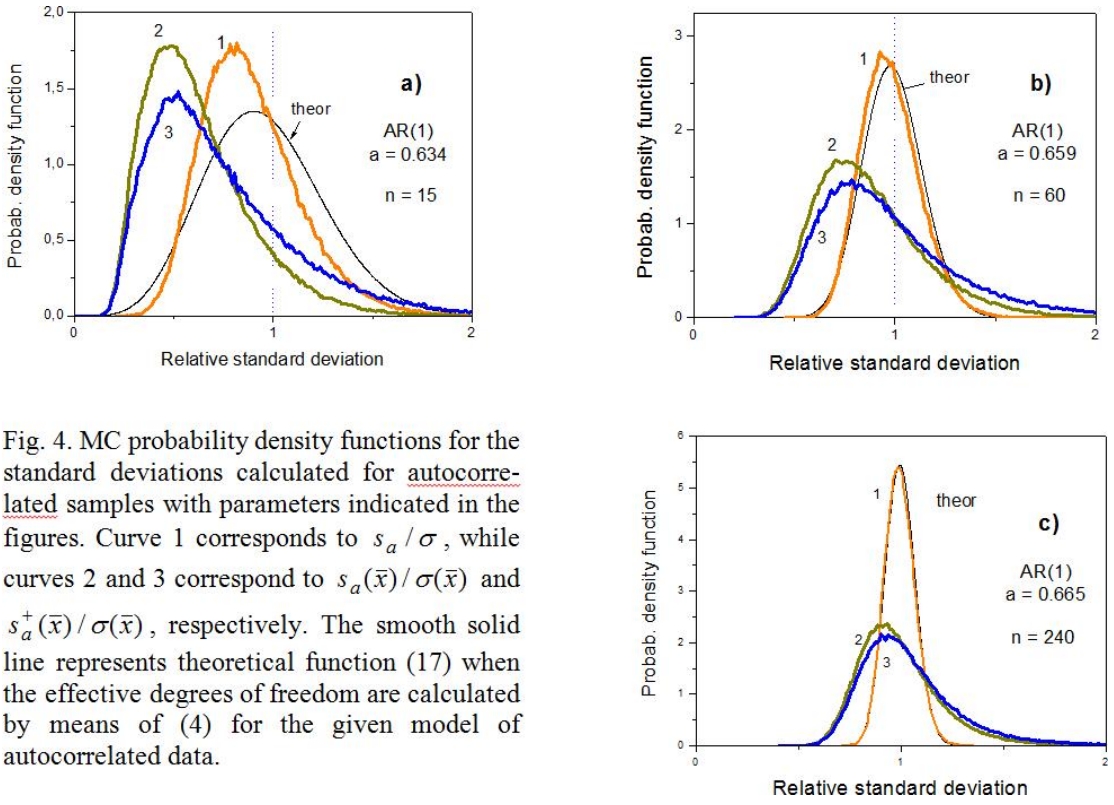

Fig. 4. MC probability density functions for the standard deviations calculated for autocorrelated samples with parameters indicated in the figures. Curve 1 corresponds to s_a/σ, while curves 2 and 3 correspond to $s_a(\bar{x})/\sigma(\bar{x})$ and $s_a^{+}(\bar{x})/\sigma(\bar{x})$, respectively. The smooth solid line represents theoretical function (17) when the effective degrees of freedom are calculated by means of (4) for the given model of autocorrelated data.

Visual inspection of Fig. 4 shows that $g(s_a/\sigma)$ approximates the theoretical function $g(z_\nu)$ with ν replaced by the suitable effective degrees of freedom ν_{eff}. This is because $g(s_a/\sigma)$ is mainly determined by probabilistic properties of the sum $(1/n)\sum(x_i-\bar{x})^2$ in (1), whereas the ratio $n_{eff}/(n_{eff}-1)$ tends to unity with increasing sample size n (hence the discrepancy is most visible for n = 15). This conclusion is reflected in Table 3 by the fact that the relative dispersion $s_r(s_a)$ and $s_r(s_a^{+})$ remain in agreement with the theoretical value

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

$$s_r(s_a) \cong (2\nu_{eff})^{-1/2}. \tag{18}$$

4.2. Standard deviation of the mean

Contrary to the case of uncorrelated observations, the relative dispersions of estimators of the standard deviation and the standard deviation of the mean are different (Fig. 4 and Table 3). The value of $s_a^2(\bar{x})$ is now estimated as the product of two random variables, $1/\hat{n}_{eff}$ and $(1/n)\sum(x_i - \bar{x})^2$ of comparable relative dispersion. Under the simplifying assumption that the correlation between the two aforementioned random factors is neglected, the approximate formulae

$$bias_r[s_a(\bar{x})] \approx bias_r(s_a) + (1/2)\, bias_r(1/\hat{n}_{eff}) \tag{19}$$

and

$$s_r[s_a(\bar{x})] \approx \left\{[s_r(s_a)]^2 + [(1/2)\, s_r(1/\hat{n}_{eff})]^2\right\}^{1/2}, \tag{20}$$

can be obtained from the law of propagation of uncertainty ((12) in [11]). Relations (19) and (20) are approximately satisfied by the numerical data provided in Table 3 and 2.

The results obtained for n = 15 demonstrate that the crude estimation of standard deviation of the mean $s_a(\bar{x})$ is possible even for this rather small sample. In general, the formalism presented here ((4) and (18)) makes it possible to determine objectively the sample size n required to achieve a given accuracy for the estimates of standard deviation.

5. An experimental example of autocorrelated data

5.1. Particle statistics effect in X-ray diffraction

X-ray diffraction is an important analytical technique used in various areas of science and technology [12]. Among others, it allows the phase content of powder samples to be determined. In this method, the reflection of X-rays occurs only for a small fraction of crystallites of the powder sample for which the so-called Bragg condition is met. When the sample is gradually inclined at a small angle ω (while the positions of the X-ray source and detector remain fixed) certain crystallites cease to reflect the X-rays, while others start to diffract the radiation. This leads to variation in the detected signal N_i (number of registered photons per single observation) as a function of the angle ω (Fig. 5).

The detector signal is a weighted sum of the contributions of crystallites that are oriented in such a way that the reflected radiation reaches the detector. The number of 'diffracting' crystallites is finite, hence the detector signal is a random variable. On the other hand, successive measurements are autocorrelated because the same crystallite contributes to the signal for a few successive values of the angle ω.

The variation observed in the intensity of the diffracted beam, which is the result of a finite number of diffracting crystallites, is called 'particle statistics'. Its magnitude increases with increasing particle size, because a corresponding decrease in the number of diffracting grains leads to stronger fluctuations in the signal detected. On the contrary, the term 'counting statistics' refers to uncorrelated Poisson fluctuations resulting from a finite number N of registered photons which are characterized by a relative standard deviation equal to $zs(N)/N = N^{-1/2}$.

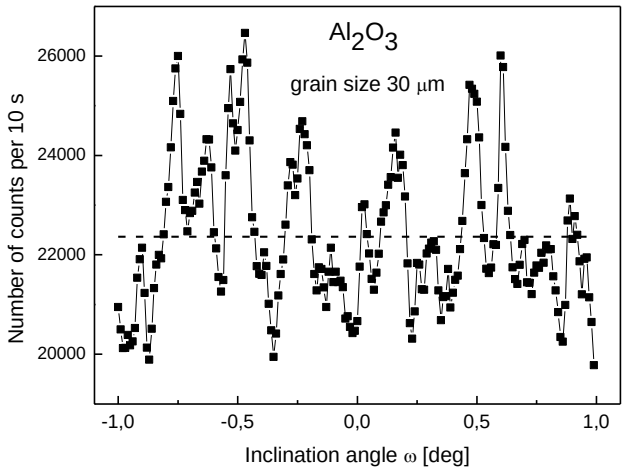

Fig. 5. Intensity of selected Bragg reflection of corundum (Al_2O_3) as a function of the inclination angle ω. The conditions under which the experimental data were collected were as follows: X-ray diffractometer X'Pert, CuKα radiation, powder sample with mean grain size of 30 μm. The experimental data were supplied by Dul (2006) in private communication.

5.2. Processing the data

The mean value of the signal detected, $\overline{N} = 22.36 \cdot 10^3$ is indicated in Fig. 5 by the dashed line. The estimate $\{r_k\}$ calculated based on these data is shown in Fig. 6.

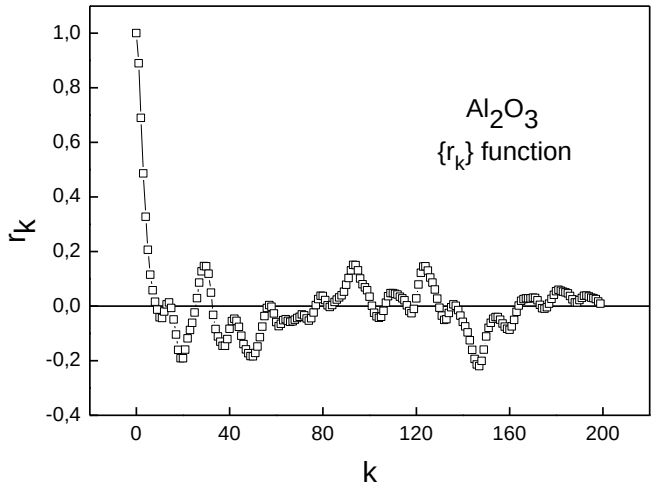

Fig. 6. Sample ACF function for the data shown in Fig. 5.

The cut-off lag of $n_c = 8$ is obtained by counting the number of values of $\{r_k\}$ preceding the first transit through zero. For $n = 200$ the effective number of observations is estimated using (15) as:

$$\hat{n}_{eff}^{+} = \frac{200 - 2 \cdot 8 - 1 + 8 \cdot 9/200}{1 + 2\,(0.889 + 0.690 + 0.486 + 0.327 + 0.206 + 0.114 + 0.057 + 0.016)} + 1 = 28.8\,.$$

The absolute and relative values of the standard deviation and standard deviation of the mean are calculated using (1) and (2) as follows:

Metrol. Meas. Syst., Vol. XVIII (2011), No. 4, pp. 529-542.

$$s_a = \sqrt{\frac{28.8\,(200-1)}{200\,(28.8-1)}} \times 1.47 \cdot 10^3 = 1.49 \cdot 10^3, \quad \frac{s_a}{\overline{N}} = 6.6\%.$$

$$s_a(\overline{N}) = \frac{1.49 \cdot 10^3}{\sqrt{28.8}} = 0.28 \cdot 10^3, \quad \frac{s_a(\overline{N})}{\overline{N}} = 1.2\%.$$

For the data analyzed, the relative standard deviation resulting from Poisson fluctuations is equal to $s_a(N)/N = 1/\sqrt{22.36 \cdot 10^3} = 0.7\%$. Hence, we verified that the effect of particle statistics dominates over the effect of counting statistics for this specific size of crystallites.

The uncertainty in the peak intensity propagates directly into the uncertainty of determining the phase content in a multiphase sample. A sample standard deviation of $s_a(I)/I = 6.6\%$ allows the uncertainty resulting from the particle statistics effect to be known *a posteriori* for a routine measurement with the sample position fixed. Averaging the signal detected over the ±1° range of angle ω lowers this uncertainty by a factor equal to $\sqrt{28.8} = 5.5$. The fact that $\hat{n}^{+}_{eff} = 28.8$ is small when compared to $n = 200$ suggests that a smaller number of data points (say, 100 or 50) in the same range ±1° of angle ω is sufficient to obtain a comparable reduction in the measurement uncertainty.

6. Conclusions

The way to estimate the standard deviations for a series of autocorrelated observations depends on the available knowledge of the autocorrelation function. Prior knowledge of the ACF allows the application of an analytical formalism. The present work focused on the case where only an estimate of the ACF can be obtained from the available data. MC simulations were carried out to investigate different estimators of the effective number of observations and the resulting estimators of the standard deviation of the data and the standard deviations of their mean.

The results obtained through the Monte Carlo simulations are not general and, for the problem examined here, depend on (a) the chosen model of autocorrelated data, (b) the sample size, (c) the estimator of the ACF, and (d) the necessary truncation of sums defining the estimators of n_{eff} at the limiting lag n_c. To obtain meaningful results within reasonable limits for the applied work, our strategy was based on consecutive checking of a few variants in succession for each of the four mentioned factors.

The two models used in this work (the simple moving average SMA and the first-order autoregressive model AR(1)) are the simplest and most representative of two basic categories of stationary time series: the autoregressive (AR) and the moving average (MA). Starting with Section 3, quantitative results are obtained only with the AR(1) model, for which different estimators tend to exhibit larger bias and dispersion when compared to the SMA model with the same effective number of observations. Hence,, limits of its applicability should also apply for MA models. Moreover, the AR(1) model is most often used to interpret the autocorrelated data collected during real-world experiments.

Numerical investigations were performed for three sample sizes: 15, 60 and 240, with corresponding effective numbers of observation equal to 3.36, 12.3 and 48.3. The simulations show that estimates of different statistical parameters can be obtained even for surprisingly small samples.

Truncation of the sums defining the effective number of observations n_{eff} and effective degree of freedom ν_{eff} can be alternatively considered as using estimates of the ACF, which are effectively zero above the limiting lag n_c. Different options for the choice of n_c were investigated, and the FTZ method introduced in this work was found to have the most

favorable properties. This method defines n_c as corresponding to the last positive element of the ACF estimates before the *first transit* through *zero*. It can be used when we know that the true correlation coefficients are nonnegative.

The use of the standard estimator $\{r_k\}$ of the ACF and three alternative estimators was examined. However, the choice of the ACF's estimator was found to be less critical than the choice of the method of defining n_c. Nonetheless, a reduction of bias can be obtained by using a proposed estimator of the effective number of observations calculated using a closed formula from the standard estimator $\{r_k\}$, the sample size n and the limiting lag n_c.

The final MC simulations concerned standard deviation of the data and standard deviation of their mean. They confirmed that their statistical properties depend on a suitably defined effective degree of freedom.

As codified by [11], the type-A uncertainty is defined as the square root of the unbiased estimator of the variance of the mean. This statement emphasises the need for estimators of standard deviation, which are possibly unbiased. The formalism presented in this paper represents a generalization of the common formulae used for independent observations. For practical purposes, exemplary calculations for real-world data are presented. An immediate extension of this work will be to apply the estimators of the standard deviation of the mean and the effective degrees of freedom to calculate the confidence interval.

Acknowledgement

P. Ramza has been partially supported by the EU Human Capital Operation Program, Polish Project No. POKL.04.0101-00-434/08-00. This work was also supported by the Polish Ministry of Science and Higher Education (MNiSW).

References

[1] Zięba, A. (2010). Effective number of observations and unbiased estimators of variance for autocorrelated data – an overview. *Metrol. Meas. Syst.*, 17, 3-16.

[2] Chipman, J.S., Kadiyala, K.R., Madansky, A., Pratt, J.W. (1968). Efficiency of the sample mean when residuals follow a first-order stationary Markoff process. *J. Amer. Statist. Assoc.*, 63, 1237-1246.

[3] Pham, T.D., Tran, L.T. (1992). On the best unbiased estimate for the mean of a short autoregressive time series. *Econometric Theory*, 8, 120-126.

[4] Bayley, G.V., Hammersley, J.M. (1946). The "effective" number of independent observations in an autocorrelated time series. *J. R. Stat. Soc. Suppl.*, 8, 184-197.

[5] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control 3rd ed*. New Jersey: Prentice Hall, Englewood Cliffs.

[6] Zhang, N.F. (2006). Calculation of the uncertainty of the mean of autocorrelated measurements. *Metrology*, 43, 276-281.

[7] Percival, D.B. (1993). Three curious properties of the sample variance and autocovariance for stationary processes with unknown mean. *The American Statistician*, 47, 274-276.

[8] Quenouille, M.H. (1949). Approximate tests of correlation in time-series. *J. R. Statist. Soc. B*, 11, 68-84.

[9] Marriott, F.H.C. Pope, J.A. (1954). Bias in the estimation of autocorrelations. *Biometrika*, 41, 390-402.

[10] Zieba, A., Ramza, P., to be published.

[11] ISO/IEC. (1995). *Guide to the Expression of Uncertainty in Measurement*. Geneva: ISO.

[12] Dinnebier, R.E., Billinge, S.J.L. Eds. (2008). *Powder Diffraction: Theory and Practice*. Cambridge: RSC Publishing.