

## **WHY THE SQUARE SUM MUST BE DIVIDED BY $N - 1$ EVEN FOR A POPULATION VARIANCE**

**MARTIN BACHMAIER**

Agricultural Systems Engineering  
Technische Universität München  
Am Staudengarten 2  
D-85354 Freising  
Germany  
e-mail: bachmai@wzw.tum.de

### **Abstract**

The objective of this article is to address the frequently found academic opinion that in order to calculate the population variance, the sum of the squared deviations from the mean should be divided by the non-reduced number of summands, and not, as is the case for the sample variance, by one less. In reality, the question regarding the subtraction of one has nothing to do with whether it concerns a population or a sample; the decisive factor is whether the sum of squares refers to all measurement values or to characteristic values that are equally probable (or frequent). In the latter case, the full number is the divisor. However, if the sum of squares contains all measurement values of a sample or a population, one must be deducted from the overall number before it can be considered suitable as a divisor. Nevertheless, and precisely because of this subtraction, the variance becomes an average spread in the sense that it is not subject to a trend with regard to the number of data points, so that variances of populations of different sizes can be compared to each other. To better understand the latter part, an alternative variance definition is used, which instead of deviations from the mean refers to pairwise differences and in this respect characterizes the nature of variance more accurately. Therefore, it should be taught as the actual variance definition.

---

2010 Mathematics Subject Classification: 60, 62.

Keywords and phrases: variance definition, population variance, sample variance, unbiased estimate, sampling techniques, semivariance.

Received September 19, 2010

### 1. Motivation

A thesis looked at the player ages prevalent in competitive and team sports. No one would be surprised to hear people would still play snooker at a much older age than European football, for example. However, the result that significantly more ice hockey teams featured forwards, whose age was higher and more widely dispersed than that of defensive players did give the advisor some cause for thought. Do forward players require a higher level of mental maturity for creativity, while the defense attracted powerful and impulsive young players? While he was thinking about it and turning the pages of the thesis, he could not help, but notice that the age data seemed quite high. He asked the student, which statistical measures he had used. The student said that he selected the maximum as the measure of location; it varied the most in view of all sports types and was in a way characteristic for the respective sports type. Accordingly, he used the range as the measure of statistical dispersion. The advisor became very angry and said that these measures were totally unsuitable for comparing populations of different sizes. It is easier to obtain a high maximum or high range for three forwards than two defensive players. The results were misleading. One could use average values to fairly compare populations of different sizes. The advisor instructed the student to calculate the mean and variance instead of the maximum and the range.

As the student started to calculate the variances, he noticed that there were different formulae: One divided the square sum by the number of addends, while the other divided by that number less one. Which one should he choose? He was not keen on getting berated again for unfairness with regard to the different volume of data. Let us first look at the literature.

### 2. Population and Sample Variance in the Literature

With regard to the population variance, many authors avoid the question of the divisor of the deviation square sum from the mean by indicating only the variance of a random variable (or its distribution) as well as a sample variance. However, these authors already cannot agree

with regard to the sample variance. In this vein, Bosch [4], Cox and Cohen [8], Kreyszig [16], Snedecor and Cochran [23], as well as Warren and Grant [26] use divisor  $n - 1$ , while Bortz [3], Burkschat et al. [5], Fahrmeir et al. [11], and Fowler [12] favour divisor  $n$ , the number of sample values. The latter apparently interpret the variance as an average square deviation from the mean, while the former invoke arguments such as unbiasedness and the associated degrees of freedom, whereby their unbiasedness refers to the variance  $\sigma^2$  of a random variable or its distribution, but not of a finite population.

There is also disagreement among authors, who do not differentiate between population and sample. For example, Crawshaw and Chambers [9] calculate the variance of a “set of numbers” with its number as the divisor, while in Moore et al. [19], the divisor of the variance of a “set of observations” is smaller by 1. Wonnacott and Wonnacott [27] allow for both, but speak about the former as “mean squared deviation” and only call the latter “variance”.

The two types of definitions for variance are more frequently found with finite populations, which were not explicitly considered by the authors that have been mentioned so far. Cochran [7], Hansen et al. [13], and Kish [15] define its variance both via the divisor  $N$ , the number of elements of the population, as well as via the divisor  $N - 1$ . All of these authors describe it as  $S^2$  in the first case, and  $\sigma^2$  in the latter. A preference for one of the two definitions is not provided. Divisor  $N - 1$  seems to derive its justification from consistency with the sample variance, where these authors favoured by using divisor  $n - 1$ , or, as expressed by Cochran [7], to “approach sampling theory by means of the analysis of variance”. On the other hand, however, the argument of an unbiased estimate, on the basis of which the divisor of the sample variance has been reduced by 1, has fallen by the wayside, since no estimate is required in the case of a population, so why reduce the divisor  $N$ ?

In contrast, Devore and Peck [10], Lohr [18], and Thompson [25] follow a consistent line. Similar to Moore et al. [19], the divisor of the square sum is 1 less than the data number, regardless of whether that number originates from a sample or a population. The latter criterion only decides whether the divisor is described as  $N - 1$  or  $n - 1$ . No reason is indicated as to why 1 is subtracted for the population variance. Divisor  $N - 1$  is also used by Sampath [21], who however does not mention the sample variance.

Those authors, who use different definitions for population and sample variance without compromise garner the most attention. These include Anderson et al. [1], Clarke et al. [6], Levy and Lemeshow [17], Sachs [20], Scheaffer et al. [22], as well as Som [24]. They use  $N$  as the divisor as soon as it concerns a population, and  $n - 1$  if it refers to a sample. For Scheaffer et al. [22], it is only possible to detect the use of  $N$  on the basis of an example on page 53. Sachs [20] effectively abandons the subtraction of 1 in the divisor, if the average value is known. However, in practice, this only applies to populations. The use of different divisors is surprising with regard to Levy and Lemeshow [17] and Som [24] to the extent that, they mention that  $[(N - 1)/N]s^2$  provides an unbiased estimate of their  $\sigma^2$  defined with the divisor  $N$ , or  $s^2$  estimates  $[N/(N - 1)]$ -times of their  $\sigma^2$  without bias. At the same time, Som [24] also mentions that some textbooks also use the divisor  $N - 1$  for population variances. Clarke et al. [6] on the other hand, justify divisor  $n - 1$  for a sample variance with unbiasedness for the variance of a random variable, but in doing so they do not provide a reason for why they use divisor  $N$  to calculate the variance of a finite population. On the other hand, Anderson et al. [1] contend that divisor  $n - 1$  of a sample variance leads to an unbiased estimate for the population variance as defined with divisor  $N$ . They do not mention that this only applies to the rather unusual case of sampling with replacement. Hedayat and Sinha [14] also use a divisor  $N$  for the population variance, without providing any reasons; the estimation of this variance using a sample is not part of the topic of their book.

The less than unanimous variance definitions found in this overview of literature are unlikely to assist our student, especially since his desire for a fair variance comparison with different data numbers has not even been touched.

Moreover, this overview of literature leaves one with the impression as if there were no definition per se; each definition seems to have advantages and disadvantages. And yet it exists. All the discrepancies mentioned above are easily solved. There is always a way of appropriately defining variance. To demonstrate this, we begin with an alternative definition of variance that is based on pairwise differences and just on this basis already offers better insights into the nature of the variance.

### 3. Variance Definitions Based on Pairwise Differences

The idea of a measure of statistical dispersion is better defined by way of pairwise differences, since this does not require an estimate of a nuisance parameter, such as the mean. However, this clarifying definition seems to be little known; with regard to all the literature referenced in the present article, it is only mentioned by Hedayat and Sinha [14].

The following applies to the usual sample variance  $s^2$  and the maximum likelihood estimate  $s_*^2$ :

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1)$$

$$= \frac{1}{2} \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)^2 = \frac{1}{2} \frac{1}{\binom{n}{2}} \sum_{i=1}^n \sum_{j=i+1}^n (x_i - x_j)^2, \quad (2)$$

$$s_*^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3)$$

$$= \frac{1}{2} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2. \quad (4)$$

If one refers to a population  $\{x_1, x_2, \dots, x_N\}$  with  $N$  individual items, usually the definition with a divisor  $N$  is considered as the actual one; however, for the sake of consistency with the above definitions, we will refer to it as  $\sigma_*^2$ , while the actual variance name  $\sigma^2$  receives the definition with the divisor  $N - 1$ :

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (5)$$

$$= \frac{1}{2} \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N (x_i - x_j)^2 = \frac{1}{2} \frac{1}{\binom{N}{2}} \sum_{i=1}^N \sum_{j=i+1}^N (x_i - x_j)^2, \quad (6)$$

$$\sigma_*^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (7)$$

$$= \frac{1}{2} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (x_i - x_j)^2. \quad (8)$$

A proof for the identity between (7) and (8), from which the identities between (5) and (6), (1) and (2), and (3) and (4) immediately follow, can be found in Hedayat and Sinha [14].

**Example 1.** The difference between the divisors  $N$  and  $N - 1$  will now be demonstrated by means of the average age of six ice hockey players at Landshut Cannibals. For the sake of simplicity, we will use the birth date of the players as a basis, so that their ages will only be accurate to the year.

**Table 1.** Squared pairwise age differences of six Landshut Cannibals players

	Goalie	Defencemen		Forwards		
Squared age differences	Hähnel *1992	Moburg *1977	Pufal *1981	Michl *1989	Welz *1976	Möhle *1989
Hähnel *1992	–	225	121	9	256	9
Moburg *1977	225	–	16	144	1	144
Pufal *1981	121	16	–	64	25	64
Michl *1989	9	144	64	–	169	0
Welz *1976	256	1	25	169	–	169
Möhle *1989	9	144	64	0	169	–

As the team currently playing on the ice, the selection of  $N = 6$  players forms a population. On the other hand, it constitutes a sample of size  $n = 6$ , for instance, to make inferences about the variance in ages of the entire team. These two perspectives suffice to make it inappropriate to calculate the variance for the population and the sample in different ways.

The table illustrates the difference between the two variance calculations. While for  $s_*^2$  and  $\sigma_*^2$  in (4) and (8), the age differences between self-identical players, which are zero by default, are considered just as all other differences — whereby the dash “—” in the table is to be interpreted as zero for this purpose — these zero differences are not considered for  $s^2$  and  $\sigma^2$  in (2) and (6). The dash “—” in this case means “does not apply”. At the same time, zero differences that result from players of the same age flow into the variance calculation. While the fact that Möhle is the same age as himself does not offer any information with regard to the spread of the age, the same age of forwards Möhle and Michl points to the age homogeneity of the team, which should bring the variance closer to zero.

Table 1 shows that the share of self-identical pairwise comparisons on the main diagonal increases as fewer players are compared. The larger the share of these zero differences that are devoid of information, the

more downward pressure it exerts on the variance  $\sigma_*^2$ . If only one player is examined, e.g., the goal keeper, this share even increases to 100% and the variance  $\sigma_*^2$  of the age is 0, even though no information regarding a dispersion is available with one player. As a result, variance  $\sigma^2$  for  $N = 1$  remains undefined. If we now examine the two defenders playing on the ice as a population, the mentioned share still makes up half, while for the three forwards only a third of comparisons result in self-identical differences. If we expand our view to all of the world's ice hockey teams, we would conclude that with regard to age, a forward line varies more than a defensive line, as long as the "varying" is measured with the  $\sigma_*^2$  variance. Debates could start as to why trainers use more homogeneous players as regards age in defensive rather than offensive lines. However, if one measures the heterogeneity of age with  $\sigma^2$ , these discrepancies do not occur.

We can see that the question regarding a fair variance comparison for different data numbers clearly speaks in favour of a divisor  $N - 1$ . The advisor of a thesis, which offers a variance calculation using divisor  $N$  instead of  $N - 1$  should not be any less angry than he was about the range for the measurement of the dispersion. Qualitatively, the resulting misrepresentation would be the same in both cases.

#### **4. Which Population Variance is Estimated without Bias by the Sample Variance?**

Is the population variance  $\sigma_*^2$  in (7) at least justified by the fact that it is estimated without bias by the usual sample variance  $s^2$  in (1)? Can a sample of  $n = 5$  result in an unbiased estimate for the variance of a population with  $N = 6$  by dividing the sum of squares of the sample variance by  $n - 1 = 4$ , and that of the population variance by  $N = 6$ ? Of course not, unless the sample is taken with replacement. As has been shown by many authors (e.g., Cochran [7], Kish [15], Levy and Lemeshow [17], and Som [24]), the variance  $s^2$  of a sample without replacement

does not estimate  $\sigma_*^2$ , with the divisor  $N$ , in an unbiased manner, but  $\sigma^2$ , where the divisor is also 1 less than the data quantity. Hence, for a normal case of sampling without replacement, the argument of unbiasedness also clearly speaks for divisor  $N - 1$ .

### **5. The Variance of a Discrete Distribution with Equal Probability Mass on each Support Point**

Are we not creating a conflict with the probability theory when using the divisor  $N - 1$  even for the population variance? The variance of a random variable  $X$  or its distribution is defined by  $\text{Var}(X) = E([X - E(X)]^2)$ , which in a discrete distribution with equal probability mass on each support point<sup>1</sup> leads to a division of the sum of squares by the number of support points, and there is no hint of a subtraction of 1!

However, a probability distribution does not relate to frequencies of a characteristic of a finite population, but to a set of characteristic values, whose occurrences are described by a probability measure, and to approach a probability by a frequency, the population size must tend to infinity. In particular, a discrete distribution with equal probability mass on each support point  $x_i, i = 1, \dots, M$ , is not comparable to the characteristic values of a finite population,  $x_1, \dots, x_M$ , but to those of an infinitely large imaginary population,  $x_1, \dots, x_M, x_1, \dots, x_M, x_1, \dots, x_M, \dots$ , where each  $x_i$  occurs an infinite number of times and with the same frequency. The variance of this imaginary population of size  $N = \infty$  can be computed as the following limit of the population variance formula in (5), which is based on divisor  $N - 1$ :

---

<sup>1</sup>To avoid misunderstandings, such a distribution should not be called a discrete uniform distribution, unless the support points  $x_i$  are equidistant.

$$\text{Var}(\text{imag. pop.}) = \lim_{k \rightarrow \infty} \frac{1}{\underbrace{kM}_N - 1} \sum_{i=1}^M k(x_i - \bar{x})^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \bar{x})^2 = \text{Var}(X). \quad (9)$$

This shows that the population variance formula based on divisor  $N - 1$  is compatible with the variance of a random variable. As far as the theory of probability is concerned, it does not play a role whether divisor  $N$  or  $N - 1$  is used.

### 6. The Significance of $x_i$ — One-Time Measurements or Recurring Occurrences of Characteristics

In an introductory statistics course, I taught the conventional mean calculation  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , but also how to calculate the empirical mean on the basis of a frequency table,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n f_i x_i$ , whereby  $f_i$  describes absolute frequencies. Some exasperated students approached me to ask when this formula should be used. One time the factor  $f_i$  is included, and another time it is not. At first, I was a little surprised at the low level of understanding of these students, as they were already finding it difficult to calculate the mean, but I then realized that I did not sufficiently emphasize the significance of the values  $x_i$  and  $n$ . They refer to different things in the two formulae. In the usual formula,  $x_i$  in sequence indicate all measurement values, and  $n$  for their number, while for a mean calculation using a frequency table,  $x_i$  only indicates the different occurrences of the characteristic, and  $n$ , which in subsequent lectures I designated as  $m$  for better differentiation, indicates the number of different occurrences.

Why am I telling you all this? Because the discrepancies with regard to different variance definitions are also based on the disregard of the exact significance of  $x_i$ , and yet no one seems to be aware of this. It seems to me as if many authors imagined a finite population as a discrete distribution with equal probability mass on each support point, and

hence, they adopt the variance formula of this distribution when calculating the variance of a finite population by using divisor  $N$  instead of  $N - 1$ . The reason for this error is nothing else than that they do not differ if  $x_i$  denote one-time measurements, which can equal each other, or different occurrences, where each of them occurs an infinite number of times.

The next example also concerns the same difference on the basis of a sample instead of a population. In this case, nobody would make this mistake of disregarding the number of occurrences.

**Example 2.** We are examining the coin toss. Heads win one Euro, while tails lose one Euro. Let us assume someone tossed a coin 2000 times and has obtained approximately 1000 tails and 1000 heads. The sample only recognizes  $m = 2$  different wins,  $x_1 = +1$  and  $x_2 = -1$ . Both appear frequently and at approximately the same frequency, so that the sample variance of the win can be approximated with  $s^2 \approx \frac{1}{m}$

$\sum_{i=1}^m (x_i - \bar{x})^2 = \frac{1}{2} [(+1 - 0)^2 + (-1 - 0)^2] = 1$ . Calculating with  $1/(m - 1)$  instead of  $1/m$  would have yielded 2 as a result, which is completely wrong. To accurately obtain the sample variance, one cannot bypass the formula in (1), which includes all  $n = 2000$  measurement values  $x_i$  in the sum and hence has the divisor  $n - 1 = 1999$ .

## 7. Sampling with Replacement

When sampling with replacement, the ordinary sample variance,  $s^2$  in (1), unbiasedly estimates  $\sigma_*^2$  in (7) with divisor  $N$ , and at a first glance, this seems to favour defining the population variance with divisor  $N$ . However, drawing and replacing involves pretending that each of the  $N$  individuals in the population is present an infinite number of times, and hence, we have arrived at the imaginary population or discrete distribution in Section 5 with the only exception that now equal  $x_i$  are also possible. Thus, what one unbiasedly estimates when applying the ordinary sample variance to a sample based on drawing and replacing is

not the population variance, but the variance of a discrete distribution, whose probabilities,  $P(X = x_i)$ , equal the frequencies of the corresponding  $x_i$  in the population. The latter variance can be computed according to the last equation in (9), where the divisor  $M$  equals the size  $N$  of the finite population, whose variance is to be estimated.

The population variance continues to be represented by  $\sigma^2$  in (5). Its divisor is  $N - 1$ , and hence, the factor by which it is greater than the variance of the aforementioned discrete distribution is  $N / (N - 1)$ . Therefore, to obtain an unbiased estimate of the population variance with data from sampling with replacement, the sample variance must be corrected by this factor:

$$s_{\text{with repl.}}^2 = \frac{N}{N-1} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (10)$$

## 8. Summary of Results and Conclusions

In the case of a finite number of data, only  $s^2$  and  $\sigma^2$  in (1) and (5), where the sum of deviances from the mean is divided by  $n - 1$  or  $N - 1$ , are appropriate for a definition of a variance. There are three reasons for this:

- They are the true average squared dispersion measures, as their averaging of pairwise squared differences in the alternative formulae in (2) and (6) is not distorted by a nuisance parameter.
- They enable a fair comparison of samples or populations of different size.
- They match with respect to unbiasedness in the usual case of sampling without replacement: Neither is  $E(s_*^2) = \sigma_*^2$  nor  $E(s^2) = \sigma_*^2$ , but  $E(s^2) = \sigma^2$ .

Moreover, it is a lot easier to teach students to use the same variance formula for sample and population than to use different formulae, especially, as it is not always easy to differ between sample and population. The variance estimate in (10) is an exception, since drawing

and replacing involves different conditions for sampling and finite population. The question of whether the sum of squares is to be divided by the number of summands or by one less,

- does not depend on the existence of either a population or sample,
- but rather depends on whether the sum of squares refers only to the different occurrences of a characteristic or to all measurement values.

If it refers to the latter, the divisor is one less than the number of summands. However, where a characteristic is discrete and its  $M$  different values occur at the same frequency and in great number, and more accurately even in infinite number, which can be described by a discrete distribution with equal probability for each occurrence, the variance calculation can be shortened by exclusive reference to the different occurrences pursuant to the limit result in (9). In that case, however, the divisor is not  $M - 1$  but rather  $M$ .

An introductory statistics course should not begin with the formulae in (1) and (5), which refer to the mean, but rather with their alternatives in (2) and (6), whereby the differences are calculated on a pairwise basis. The pairwise differences remind one of the distance tables for cities, something students are familiar with, which reduces the fear of variance calculations. Furthermore, only the latter definitions are recognizable as average values, because only here is the sum divided by the number of summands and not by one less. The factor  $\frac{1}{2}$  contained in these definitions is justified in accordance with the standard deviation of the normal distribution density; its inflection points should be at  $\mu \pm \sigma$  and not at  $\mu \pm \sigma / \sqrt{2}$ .

The definition in (2) also forms the foundation for the calculation of empirical variogram values. To this end, it only needs to be restricted to those value pairs, which were measured at a spatial separation  $h$ . Students who encounter geo-statistics later in their studies, should already be taught this connection during the definition of the variance. Only then can they understand why variogram values should not be referred to as semi-variance, which some might be tempted to do owing to the factor  $\frac{1}{2}$  (Bachmaier and Backes [2]).

**References**

- [1] D. R. Anderson, D. J. Sweeney and T. A. Williams, *Statistics for Business and Economics*, 8th ed., South Western, a division of Thomson Learning, Cincinnati, 2002.
- [2] M. Bachmaier and M. Backes, Variogram or semivariogram? — Understanding the variances in a variogram, *Precision Agriculture* 9 (2008), 173-175.
- [3] J. Bortz, *Statistik für Human- und Sozialwissenschaftler*, 6th ed., Springer Medizin Verlag, Heidelberg, 2005.
- [4] K. Bosch, *Statistik-Taschenbuch*, R. Oldenbourg Verlag, München, Wien, 1993.
- [5] M. Burkschat, E. Cramper and U. Kamps, *Beschreibende Statistik, Grundlegende Methoden*, Springer-Verlag, Berlin, Heidelberg, 2004.
- [6] G. M. Clarke and D. Cooke, *A Basic Course in Statistics*, 5th ed., Hodder Arnold, Oxford University Press, London, New York, 2005.
- [7] W. G. Cochran, *Sampling Techniques*, 3rd ed., 14th print, John Wiley & Sons, Inc., New York, 1977.
- [8] B. G. Cox and S. B. Cohen, *Methodological Issues for Health Care Surveys*, Marcel Dekker, Inc., New York, Basel, 1985.
- [9] J. Crawshaw and J. Chambers, *A Concise Course in Advanced Level Statistics*, 4th ed., Nelson Thornes Limited, Cheltenham, 2001.
- [10] J. Devore and R. Peck, *Introductory Statistics*, 2nd ed., West Publishing Company, Minneapolis, New York, 1994.
- [11] Fahrmeir, Künstler, Pigeot and Tutz, *Statistik, Der Weg zur Datenanalyse*, Springer-Verlag, Berlin, Heidelberg, 2007.
- [12] F. J. Fowler, *Survey Research Methods*, 4th ed., Sage, Los Angeles, 2009.
- [13] M. H. Hansen, W. N. Hurwitz and W. G. Madow, *Sample Survey Methods and Theory, Volume I, Methods and Applications*, John Wiley & Sons, Inc., New York, 1993.
- [14] A. S. Hedayat and B. K. Sinha, *Design and Inference in Finite Population Sampling*, John Wiley & Sons, Inc., New York, 1991.
- [15] L. Kish, *Survey Sampling*, John Wiley & Sons, Inc., New York, 1995.
- [16] E. Kreyszig, *Statistische Methoden und ihre Anwendungen*, Vandenhoeck & Ruprecht in Göttingen, 1988.
- [17] P. S. Levy and S. Lemeshow, *Sampling of Populations, Methods and Applications*, 3rd ed., John Wiley & Sons, Inc., New York, 1999.
- [18] S. L. Lohr, *Sampling: Design and Analysis*, Duxbury Press, Pacific Grove, California, 1999.
- [19] D. S. Moore, G. P. McCabe and B. A. Craig, *Introduction to the Practice of Statistics*, 6th ed., W. H. Freeman and Company, New York, 2009.

- [20] L. Sachs, *Angewandte Statistik*, 1992 7th ed., Springer-Verlag, Berlin, Heidelberg, New York, 1989.
- [21] S. Sampath, *Sampling Theory and Methods*, 2nd ed., Alpha Science International Ltd., Harrow, UK, 2005.
- [22] R. L. Scheaffer, W. Mendenhall and R. L. Ott, *Elementary Survey Sampling*, 6th ed., Duxbury, Thomson Books/Cole, Belmont, USA, 2006.
- [23] G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed., Iowa State University Press / Ames, 1989.
- [24] R. K. Som, *Practical Sampling Techniques*, 2nd ed., Marcel Dekker, Inc., New York, 1996.
- [25] S. K. Thompson, *Sampling*, John Wiley & Sons, Inc., New York, 2002.
- [26] J. E. Warren and G. R. Grant, *Statistical Methods in Bioinformatics*, Springer-Verlag, Berlin, Heidelberg, New York, 2001.
- [27] R. J. Wonnacott and T. H. Wonnacott, *Introductory Statistics*, 4th ed., John Wiley & Sons, Inc., New York, 1985.

