EUROPEAN COMMISSION
DIRECTORATE-GENERAL
REGIONAL POLICY

GUIDANCE NOTE ON SAMPLING METHODS FOR AUDIT AUTHORITIES (UNDER ARTICLE 62 OF REGULATION (EC) NO 1083/2006 AND ARTICLE 16 OF COMMISSION REGULATION (EC) N° 1028/2006)

DISCLAIMER:

"This is a Working Document prepared by the Commission services. On the basis of the applicable Community Law, it provides technical guidance to the attention of public authorities, practitioners, beneficiaries or potential beneficiaries, and other bodies involved in the monitoring, control or implementation of the Cohesion policy on how to interpret and apply the Community rules in this area. The aim of the working document is to provide Commission's services explanations and interpretations of the said rules in order to facilitate the implementation of operational programmes and to encourage good practice(s). However this guidance is without prejudice to the interpretation of the Court of Justice and the Court of First Instance or evolving Commission decision making practice."

TABLE OF CONTENTS

1. INTRODUCTION

This is a Working Document prepared by the Commission services. On the basis of the applicable Community law, it provides technical guidance to the attention of public authorities, practitioners, beneficiaries or potential beneficiaries, and other bodies involved in the monitoring, control or implementation of Cohesion Policy on how to interpret and apply the Community rules in this area. The aim of the working document is to provide Commission services' explanations and interpretations of the said rules in order to facilitate the implementation of operational programmes and to encourage good practices. However, this guidance is without prejudice to the interpretation of the Court of Justice and the Court of First Instance or evolving Commission decision making practice.

The present guide to statistical sampling for auditing purposes has been prepared with the objective of providing audit authorities in the Member States with an overview of the most commonly used sampling methods, thus providing concrete support in the implementation of the new regulatory framework for the programming period 2007-2013.

The selection of the most appropriate sampling method to meet the requirements of Article 62 of Council Regulation (EC) N° 1083/2006 and Article 16, including Annex IV, of Commission Regulation (EC) N° 1828/2006 is at the audit authority's own professional judgement. Accordingly, this guide is not an exhaustive catalogue nor are the sampling methods described therein prescribed by the European Commission. In annex VII, a list of reference material can be found which may be relevant when determining the sampling method to be used.

The selected method should be described in the audit strategy referred to in Article 62 (1) (c) of Regulation N° 1083/2006 which should be established in line with model of Annex V of the Commission Regulation (EC) N° 1828/2006 and any change in the method should be indicated in subsequent versions of the audit strategy

International auditing standards provide guidance on the use of audit sampling and other means of selecting items for testing when designing audit procedures to gather audit evidence.

The *Intosai* standards related to competence 2.2.37 state that "The SAI should equip itself with the full range of up-to-date audit methodologies, including systems-based techniques, analytical review methods, statistical sampling, and audit of automated information systems."

The Guideline number 23 of the European Implementing Guidelines for the Intosai auditing standards, issued by the European Court of Auditors, covers amongst others the factors affecting the decision to sample1, the stages of audit sampling and the evaluation of the overall results of substantive testing.

International Standard on Auditing 530 "Audit sampling and other means of testing" also provides indications about evaluating the sample results and examples of factors influencing sample size for tests of controls and for tests of details.

[1] Please see Annex VI – List of commonly used terminology

The Institute of Internal Auditors (IIA) refers to statistical sampling in the *International Standards for the Professional Practice of Internal Auditing (Standard 2100)* highlighting that the Practice advisory has been adopted from the Information Systems Audit and Control Association (ISACA) Guideline – Auditing Sampling, Document G10. This IS Auditing guideline was issued in March 2000 by ISACA.

2. REFERENCE TO THE LEGAL BASIS – REGULATORY FRAMEWORK

Article 62 of Council Regulation (EC) N° 1083/2006 of 11 July 2006 laying down the general provisions of the European Regional Development Fund, the European Social Fund and the Cohesion Fund refers to the responsibility of the audit authority to ensure the execution of audits of the management and control systems and of audits of operations on the basis of an appropriate sample.

Commission Regulation (EC) N° 1828/2006 of 8 December 2006 setting out rules for the implementation of Council Regulation (EC) N° 1083/2006 establishes detailed provisions in relation to sampling for audits of operations in Articles 16[2] and 17[3] and in Annex IV.

The two regulations define the requirements for the system audits[1] and audits of operations to be carried out in the framework of the Structural Funds, and the conditions for the sampling of operations to be audited which the audit authority has to observe in establishing or approving the sampling method. They include certain technical parameters to be used for a random statistical sample and factors to be taken into account for a complementary sample.

The principal objective of the systems audits and audits of operations is to verify the effective functioning of the management and control systems of the operational programme and to verify the expenditure declared[4]

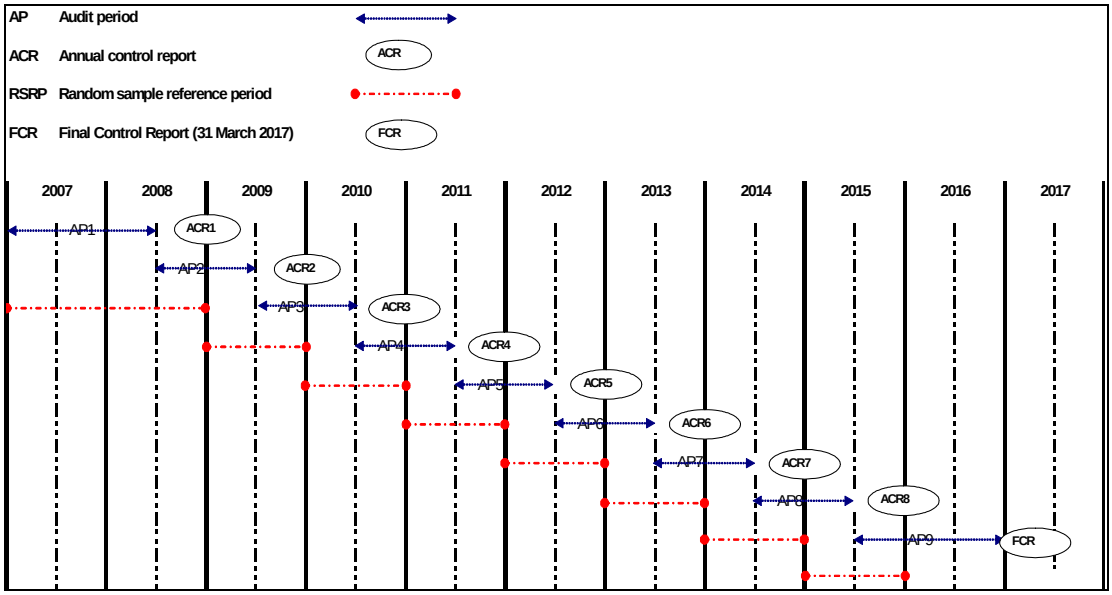
These Regulations also set out the timetable for the audit work and the reporting by the audit authority.

[2] Article 16.1 states " The audits referred to in point (b) of Article 62(1) of Regulation (EC) No 1083/2006 shall be carried out each twelve-month period from 1 July 2008 on a sample of operations selected by a method established or approved by the audit authority in accordance with Article 17 of this Regulation."

[3] Article 17.2 states " The method used to select the sample and to draw conclusions from the results shall take account of internationally accepted audit standards and be documented. Having regard to the amount of expenditure, the number and type of operations and other relevant factors, the audit authority shall determine the appropriate statistical sampling method to apply. The technical parameters of the sample shall be determined in accordance with Annex IV."

[4] Article 62 (1) (c) of Council Regulation (EC) No 1083/2006 (OJ L210/25)

Figure 1 Timeframe for Article 62 of Council Regulation (EC) N° 1083/2006



The audit authority has to report on the basis of the audit work carried out during the audit period 01/07/N to 30/06/N+1 as at 31/12/N+1[5]. The audits of operations are carried out on the expenditure declared to the Commission in year N (random sample reference period). In order to provide an annual opinion, the audit authority should plan the audit work, including system audits and audits of operations, properly. With respect to the audits of operations, the audit authority has different options in planning and performing the audits, as set out in section 6.8.

[5] The first annual control report and audit opinion (ACR1) must be provided by 31/12/2008 and will be based on audit work performed from 01/01/2007 to 30/06/2008. As expenditure is not expected to be incurred (or very little) in 2007, the first results of the sampling of operations are expected in the ACR2 to be reported by 31/12/2009, covering expenditure incurred from 01/01/2007 to 31/12/2008.

3. RELATIONSHIP BETWEEN AUDIT RISK AND SYSTEM AUDITS AND AUDITS OF OPERATIONS

Audit risk is the risk that the auditor issues (1) an unqualified opinion, when the declaration of expenditure contains material misstatements, or (2) a qualified or adverse opinion, when the declaration of expenditure is free from material misstatements.

Audit risk model and assurance model

The three components of audit risk are referred to respectively as inherent risk [IR], control risk [CR] and detection risk [DR]. This gives rise to the audit risk model of:

AR = IR x CR x DR, where

- IR, inherent risk, is the perceived level of risk that a material misstatement may occur in the client's financial statements (i.e. for the Structural Funds, certified statements of expenditure to the Commission), or underlying levels of aggregation, in the absence of internal control procedures. The inherent risk is linked to the kind of activities of the audited entity and will depend on external factors (cultural, political, economic, business activities, clients and suppliers, etc) and internal factors (type of organisation, procedures, competence of staff, recent changes to processes or management positions, etc). For the Structural Funds, the inherent risk is usually set at a high percentage.
- CR, control risk, is the perceived level of risk that a material misstatement in the client's financial statements, or underlying levels of aggregation, will not be prevented, detected and corrected by the management's internal control procedures. As such the control risks are related to how well inherent risks are managed (controlled) and will depend on the internal control system including application controls1, IT controls and organisational controls, to name a few.
- DR, detection risk, is the perceived level of risk that a material misstatement in the client's financial statements, or underlying levels of aggregation, will not be detected by the auditor. Detection risks are related to how adequately the audits are performed: competence of staff, audit techniques, audit tools, etc.

The assurance model is in fact the opposite of the risk model. If the audit risk is considered to be 5%, the audit assurance is considered to be 95%.

Audit planning

The use of the audit risk/audit assurance model relates to the planning and the underlying resource allocation for a particular operational programme or several operational programmes and has two purposes:

1. Providing a high level of assurance: assurance is provided at a certain level, e.g. for 95% assurance, audit risk is then 5%.
2. Performing efficient audits: with a given assurance level of for example 95%, the auditor should develop audit procedures taking into consideration the IR and CR. This allows the audit team to reduce audit effort in some areas and to focus on the more risky areas to be audited.

Illustration:

Low assurance: Given a desired, and accepted audit risk of 5%, and if inherent risk (=100%) and control risk (= 50%) are high, meaning it is a high risk entity where internal control procedures are not adequate to manage risks, the auditor should strive for a very low detection risk at 10%. In order to obtain a low detection risk the amount of substantive testing and therefore sample size need to be increased. In the formula= 1*0,5*0,1= 0,05 audit risk.

High assurance: In a different context, where inherent risk is high (100%) but where adequate controls are in place, one can assess the control risk as 12,5%. To achieve a 5% audit risk level, the detection risk level can be at 40%, the latter meaning that the auditor can take more risks by reducing the sample size. In the end this will mean a less detailed and a less costly audit. In the formula= 1*0,125*0,40=0,05 audit risk.

Note that both examples result in the same achieved audit risk of 5% within different environments.

To plan the audit work, a sequence should be applied in which the different risk levels are assessed. First the inherent risk needs to be assessed and, in relation to this, control risk needs to be reviewed. Based on these two factors the detection risk can be set by the audit team and will involve the choice of audit procedures to be used during the detailed tests.

Though the audit risk model provides a framework for reflection on how to construct an audit plan and allocate resources, in practice it may be difficult to quantify precisely inherent risk and control risk.

Assurance levels depend mainly on the quality of the system of internal controls. Auditors evaluate risk components based on knowledge and experience using terms such as LOW, MODERATE/AVERAGE or HIGH rather than using precise probabilities. If major weaknesses are identified during the systems audit, the control risk is high and the assurance level would be low. If no major weaknesses exist, the control risk is low and if the inherent risk is also low, the assurance level would be high.

In the context of the Structural Funds, Annex IV of Regulation (EC) No 1828/2006 states "In order to obtain a high level of assurance, that is, a reduced audit risk, the audit authority should combine the results of system audits (*which corresponds to the control assurance*) and audits of operations (*detection assurance*). The combined level of assurance obtained from the systems audits and the audits of operations should be high. The audit authority should describe in the annual control report the way assurance has been obtained". It is expected that the audit authority needs to obtain a 95% level of assurance in order to be able to state that it has "reasonable assurance" in its audit opinion. Accordingly the audit risk is 5%. The assumption contained in Regulation (EC) No 1828/2006 ("the Regulation") is that even a poorly functioning system will always give a minimum assurance (≥5%) and that the remaining assurance (90%) is obtained from the audit of operations.

In the exceptional case that the audit authority concludes that no assurance at all can be obtained from the system, the assurance level to be obtained from the audit of operations is 95%.

Relationship between audit risk, system audits and audits of operations

As indicated before, inherent risk is a factor that needs to be assessed first before starting detailed audit procedures. Typically this is performed by having interviews with management and key personnel, but also by reviewing contextual information (such as organisation charts, manuals and internal/external documents).

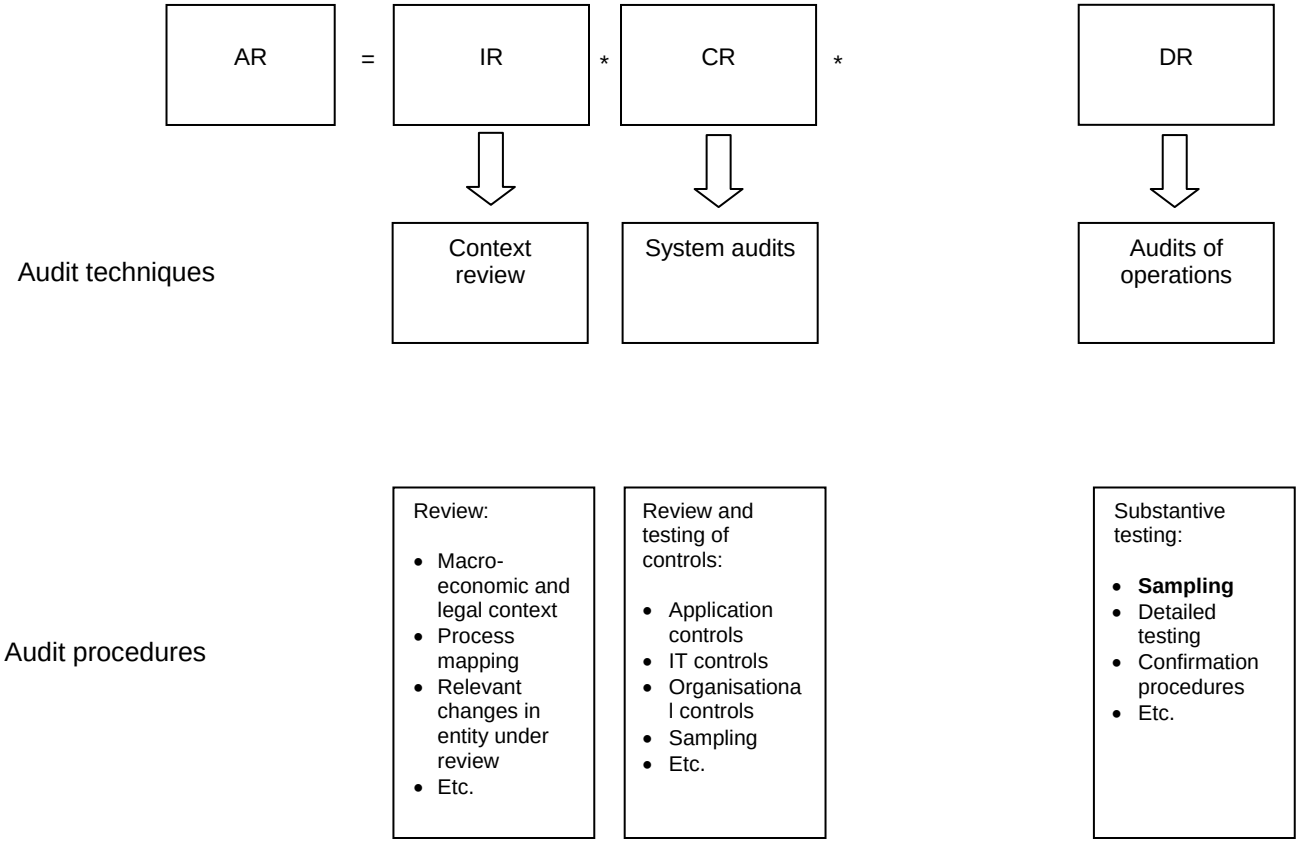
Control risks are evaluated by means of system audits[1], which consist of an internal controls review on processes and IT systems and include tests of controls. Effective control systems are based on control activities but also risk management procedures, the control environment, information and communication. For more details, reference can be made to Article 28a of the revised Financial Regulation[6] and to the COSO model[7].

Detection risks are related to performing audits of operations and underlying transactions. These include tests of details called substantive tests.

[6] Council Regulation (EC, Euratom) N° 1995/2006 of 13 December 2006 amending Regulation (EC, Euratom) No 1605/2002 on the Financial Regulation applicable to the general budget of the European Communities. OJ L390/1.

[7] COSO is one of the most important and well-known internal control frameworks. For further information please consult: www.coso.org.

Figure 2 Relationship between the different types of risks, audit techniques and audit procedures applied



The product of inherent and control risk (i.e. IR x CR) is referred to as the risk of material misstatement. The risk of material misstatement is related to the result of the system audits.

As previously indicated, if major weaknesses are identified during the systems audit, one can say that the risk of material misstatement is high (control risks in combination with inherent risks) and as such the assurance level would be low. Annex IV of Commission Regulation (EC) N° 1828/2006 indicates that if the assurance level is low the confidence level to be applied for sampling would be not less than 90%.

If no major weaknesses in the systems exist the risk of material misstatements is low, and the assurance level given by the system would be high meaning that the confidence level to be applied for sampling would be not less than 60%.

The implications of these strategic choices for the audit planning and sampling of operations are explained in the chapters that follow.

4. RELATIONSHIP BETWEEN THE RESULTS OF THE SYSTEM AUDITS AND THE SAMPLING OF OPERATIONS

Annex IV of Commission Regulation N° 1828/2006 states that substantive tests should be performed on samples, the size of which will depend on a confidence level determined according to the assurance level obtained from the system audit, i.e.

- not less than 60% if assurance is high;
- average assurance (no percentage corresponding to this assurance level is specified in the Commission Regulation);
- not less than 90% if assurance is low.

Annex IV also states that the audit authority shall establish criteria used for system audits in order to determine the reliability of the management and control systems. These criteria should include a quantified assessment of all key elements of the systems and encompass the main authorities and intermediate bodies participating in the management and control of the operational programme.

The Commission in collaboration with the European Court Auditors has developed a guidance note on the methodology for the evaluation of the management and control systems. It is applicable both to mainstream and ETC programmes. It is recommended that the audit authority takes account of this methodology.

In this methodology, four reliability levels[8] are foreseen:

- Works well, only minor improvements are needed
- Works, but some improvements are needed
- Works partially, substantial improvements are needed
- Essentially does not work.

In accordance with the Regulation, the confidence level for sampling is determined according to the reliability level obtained from the system audits.

As indicated above, the Regulation foresees only 3 levels of assurance on systems: high, average and low. The average level effectively corresponds to the second and third categories of the methodology, which provide a more refined differentiation between the two extremes of high/"works well" and low/"does not work".

8 Corresponding to the overall assessment of the internal control system.

The recommended relationship is shown in the table below[9]:

Assurance level from the system audits	Related reliability in the regulation/assurance from the system	Confidence level
Works well, only minor improvements are needed	High	Not less than 60%
Work, but some improvements are needed	Average	70%
Works partially, substantial improvements are needed	Average	80%
Essentially does not work	Low	Not below 90%

It is expected that at the beginning of the programming period, the assurance level is low as no or only a limited number of system audits will have taken place. The confidence level to be used would therefore be not less than 90%. However, if the systems remain unchanged from the previous programming period and there is reliable audit evidence on the assurance they provide, the Member State could use another confidence level (between 60 % and 90 %). The methodology applied for determining this confidence level will have to be explained in the audit strategy and the audit evidence used to determine the confidence level will have to be mentioned.

The confidence level is set by the Regulation for the purpose of defining the sample size for substantive tests. The sample size depends directly on three parameters:

1. The confidence level;
2. The variability of the population (i.e. a measure of how variable are the values of the population items, for instance a population with 100 operations of similar value is much less variable than a population of 100 operations made out of 50 very large value items and 50 very small value items);
3. The acceptable error set by the auditor (which is the maximum materiality level of 2%).

The sample size depends indirectly on the population size, through the variability of the population. A population of a larger size is likely to display more variability and therefore the corresponding sample size would be higher; the size of the corresponding sample continues to increase with larger populations, but at a decreasing rate. In other words, the sample required for a population of a certain size (say 5,000) would not be significantly larger than the one required for a population of half the size of the first (2,500).

As the sample size is directly affected by the confidence level, the objective of the Regulation is clearly to offer the possibility of reducing audit workload for systems with an established low error rate (and therefore high assurance), while maintaining the requirement to check a high number of items in the case a system has a potentially high error rate (and therefore low assurance).

---

9 In the sampling presentation to the MS, by way of illustration, 5 categories were shown. Following the preparation of the guidance for evaluation of the management and control systems, the Commission recommends MS to align their approach to the 4 categories.

4.1. Special considerations

Determination of the applicable assurance level when grouping programmes

The audit authority should apply **one** assurance level in the case of grouping of programmes.

In case the system audits reveal that within the group of programmes, there are differences in the conclusions on the functioning of the various programmes, the following options are available:

- to create two (or more) groups, for example the first for programmes with a low level of assurance (confidence level of 90%), the second group for programmes with a high level of assurance (a confidence level of 60%), etc. Consequently the number of controls to be performed will be higher, as a sample from each separate group will have to be taken;
- to apply the lowest assurance level obtained at the individual programme level for the whole group of programmes.

It is not acceptable within the group, to create a stratification between the programmes which present, for example, a level of assurance of 90% and the programmes which present a level of assurance of 60%, while maintaining a single sample, within which the layer at 90% will have a proportionally higher number of controls than the layer at 60%.

5. SAMPLING TECHNIQUES APPLICABLE TO SYSTEM AUDITS

Article 62 of Council Regulations (EC) N° 1083/2006 states: "The audit authority of an operational programme shall be responsible in particular for: (a) ensuring that audits are carried out to verify the effective functioning of the management and control system of an operational programme…". These audits are called system audits. System audits aim at testing the effectiveness of controls in the management and control system and concluding on the assurance level that can be obtained from the system. Whether or not to use a statistical sampling approach for the test of controls is a matter of professional judgement regarding the most efficient manner to obtain sufficient appropriate audit evidence in the particular circumstances.

Since for system audits the auditor's analysis of the nature and cause of errors is important, as well as, the mere absence or presence of errors, a non-statistical approach could be appropriate. The auditor can in this case choose a fixed sample size of the items to be tested for each key control. Nonetheless, professional judgement will have to be used in applying the relevant factors[10] to consider. If a non statistical approach is used then the results cannot be extrapolated.

Attribute sampling is a statistical approach which can help the auditor to determine the level of assurance of the system and to assess the rate at which errors appear in a sample. Its most common use in auditing is to test the rate of deviation from a prescribed control to support the auditor's assessed level of control risk. The results can then be projected to the population.

As a generic method encompassing several variants, attribute sampling is the basic statistical method to apply in the case of system audits; any other method that can be applied to system audits will be based on the concepts developed below.

Attribute sampling tackles binary problems such as yes or no, high or low, true or false answers. Through this method, the information relating to the sample is projected to the population in order to determine whether the population belongs to one category or the other.

The Regulation does not make it obligatory to apply a statistical approach to sampling for control tests in the scope of a systems audit. Therefore, this chapter and the related annexes are included for general information and will not be developed further.

For further information and examples related to the sampling techniques applicable to system audits, please refer to the specialized audit sampling literature included in Annex VIII of this guide.

10 For further explanation or examples see "Audit Guide on Sampling, American Institute of Certified Public Accountants, 01/04/2001".

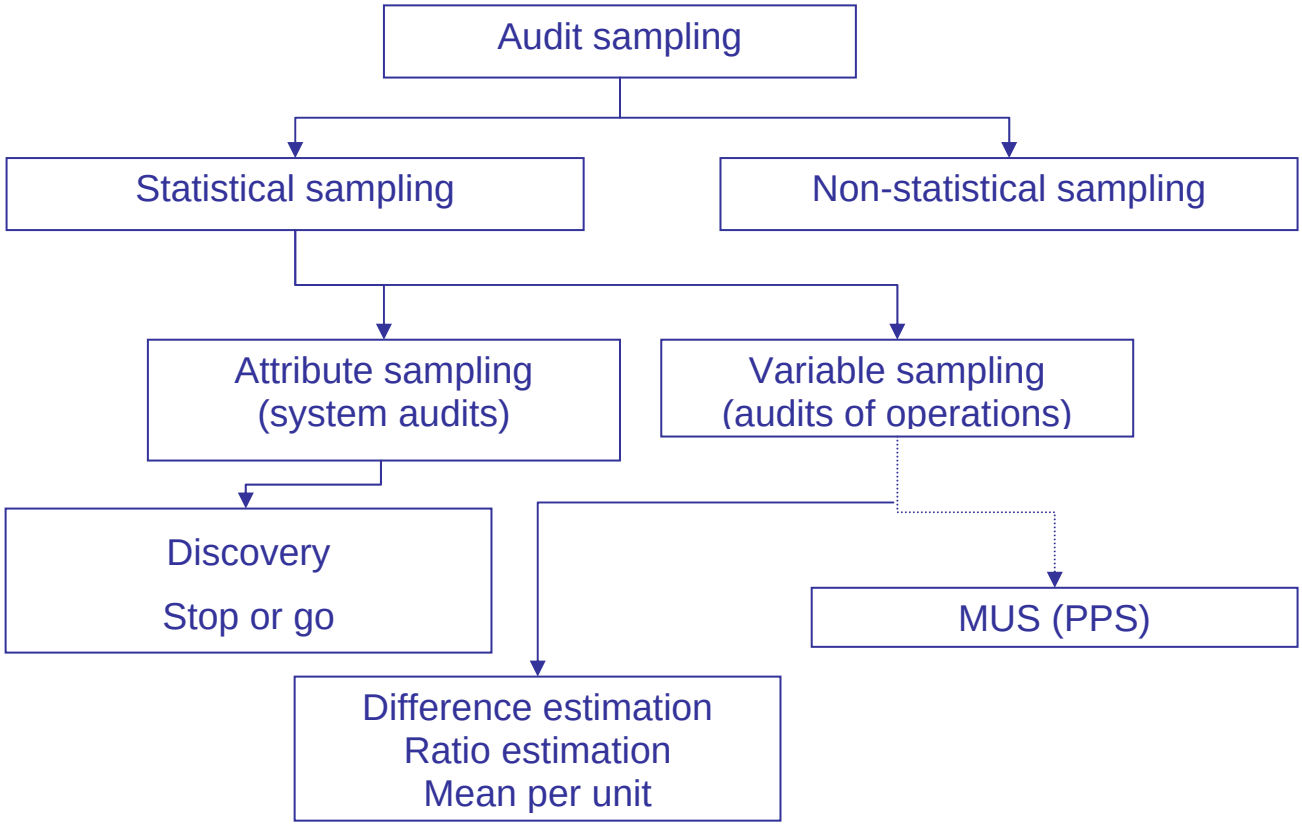
6. SAMPLING TECHNIQUES FOR THE SELECTION OF OPERATIONS TO BE AUDITED

Within the audit of operations, the purpose of sampling is to select the operations to be audited through substantive tests of details; the population comprises the expenditure certified to the Commission for operations within a programme/group of programmes in the year subject to sample ('random sample reference period' in Figure 1).

All operations for which declared expenditure has been included in certified statements of expenditure submitted to the Commission during the year subject to sample, should be comprised in the sampled population. All the expenditure declared to the Commission for all the selected operations in the sample must be subject to audit. The audit authority may decide to widen the audit to other related expenditure declared by the selected operations outside the reference period, in order to increase the efficiency of the audits. The results from checking additional expenditure should not be taken into account for determining the error rate from the sample.

Generally a distinction is made between statistical and non statistical sampling methods as shown in the overview below:

Figure 3 Audit sampling methods



Most statistical sampling methods covering the selection of operations belong to the category of variable sampling.

Variable sampling aims at projecting to the population the value of a parameter (the "variable") observed in a sample. The principal use of variable sampling in auditing is to determine the reasonableness of recorded amounts and to reach conclusions for the population in terms of whether or not it is materially misstated and, if so, by how much (an error amount). The "variable", in that sense, is the misstatement value of the sample item.

The only non-variable sampling method that can be applied to the selection of operations to be audited is monetary unit sampling (MUS), also labelled probability-proportional-to-size (PPS). It is also often classified as variable sampling because it serves the same objective of performing substantive tests.

As a preliminary remark on the choice of a method to select the operations to be audited, whilst the criteria that should lead to this decision are numerous, from a statistical point of view the variability of the population (large number of operations, operations with very different sizes…) and the expected error frequency (the expected number of misstatements, not their value) are the most relevant. The table below gives some indications on the most appropriate methods depending on the criteria.

Note that in the table below a low expected error frequency actually means an expected number of errors close to zero. Also, in the case of high variability and high error frequency (that is the most frequent case), the approach suggested is clustering or stratification of the population in the first instance. This means that clustering or stratification should be used to either minimise variability or isolate error-generating subsets of the population. The approach corresponding to the new situation (variable sampling or monetary unit sampling) should then be used. The rationale behind these approaches is detailed in the following sections of this guide.

Population variability	**Expected error frequency**	**Suggested approach**
Low	Low	Variable sampling – Monetary unit sampling
High	Low	Monetary unit sampling
Low	High	Variable sampling
High	High	Clustering or stratification (plus appropriate sampling method)

Note that "variable sampling" encompasses variable sampling as well as any variant methods, such as difference estimation.

It is also very important to stress once more the fact that in relation to all sampling methods, the application of the auditor's professional judgment is essential for choosing the most appropriate method and for evaluating correctly the results.

6.1. Selection methods

The concept of "sampling method" actually encompasses two elements: the selection method (statistical or non-statistical) and the actual sampling method, which provide the

framework for computing sample size and sampling risk and allowing for projection of the results.

A selection method can belong to one of two broad categories:

- Statistical (random) selection, or
- Non-statistical (non-random) selection.

This classification is mostly a naming convention, as some random methods do not rely on statistical concepts and some non-random methods provide some interesting statistical characteristics.

6.1.1. Statistical selection

Statistical selection covers two possible methods:

- Random sampling
- Systematic sampling

Random sampling is truly random, and randomness should be ensured by using proper random number generating software, specialised or not (e.g. MS Excel provides random numbers).

Systematic sampling picks a random starting point and then applies a systematic rule to select further items (e.g. each 20^{th} item after the first (random) starting item).

Random statistical sampling is required by Council Regulation (EC) N^{o} 1083/2006 and Commission Regulation (EC) N^{o} 1828/2006 for substantive tests (audit of operations). Both methods above fulfil the regulatory requirements if properly used.

6.1.2. Non-statistical selection

Non-statistical selection covers the following possibilities:

- Haphazard selection
- Block selection
- Judgement selection
- Risk based sampling combining elements of the three possibilities above

Haphazard selection is “false random” selection, in the sense of an individual “randomly” selecting the items, implying an unmeasured bias in the selection (e.g. items easier to analyse, items easily accessed, items picked from a list displayed particularly on the screen, etc…).

Block selection is similar to cluster sampling, where the cluster is picked non-randomly.
Judgement selection is purely based on the auditor’s discretion, whatever the rationale (e.g. items with similar names, or all operations related to a specific domain of research, etc…).

Risk-based sampling is a non-statistical selection of items based on various intentional elements, often taking from all three non-statistical selection methods.

Both statistical and non-statistical sampling is allowed by the Regulation for the complementary sample (see also section 6.8).

6.1.3. Cluster and stratified sampling

Cluster sampling, or clustering, is a random selection method of grouping items together in clusters. The whole population is divided into subsets, some subsets being sampled while others are not. Cluster sampling can be one-stage (randomly pick a cluster and analyse 100% of the items within), two-stage (randomly picking items in randomly picked clusters) or three-stage (randomly picking items in a randomly picked sub-group within a randomly picked cluster), depending on the size and complexity of the population. As a statistical sampling method must still be used, clustering may increase the sample size, and is therefore unlikely to be an efficient approach to follow.
Stratified sampling is a method which consists in sorting the population into several layers usually according to the value of the variable being audited (e.g. the value of expenditure per operation within the audited programme). Different methods can be used for each layer, for instance applying a 100% audit of the high-value items (i.e. no sampling), then applying a random statistical sampling method to audit a sample of the remaining lower-value items that constitute the second layer. This is useful in the event of a population with a few quite extraordinary items, as it lowers the variability in each layer and therefore the sample sizes for each layer. However, if by stratifying the variability does not decrease significantly, the sum of the sample sizes risks being above the sample size that would have been required for the population as a whole.

Stratification and clustering are methods to organise a population into smaller sub-sets. Randomness must be ensured: in clustering by randomly selecting clusters and/or items within clusters, and in the stratified approach by choosing 100% of a layer or a random sample in that layer.

Reaching conclusions for the whole population:

- for a stratified approach the resulting figures (expected misstatement and upper misstatement limit) from each layer are simply added together;

- for clustering, the resulting figures (expected misstatement and upper misstatement limit) from each cluster will be extrapolated to the level above it (the population, if one-stage clustering, or another cluster if several stages of clustering were used – in that case the figures are projected several times, with the risk of exaggerating the upper misstatement limit at the level of the population).

6.1.4. Special considerations

Materiality

The materiality level of 2% maximum is applicable to the expenditure declared to the Commission in the reference year. The audit authority can consider reducing the materiality for planning purposes.

Sampling unit

The population for sampling comprises the **expenditure certified** to the Commission for **operations** within a programme or group of programmes in the reference year subject to sample, and therefore not cumulative data.

The sampling unit is the Euro (or national currency) for Monetary Unit Sampling but the unit to be selected for audit is generally the operation/payment claim(s) submitted for the operation. Where an operation consists of a number of distinct projects, they may be identified separately for sampling purposes. In certain cases in order to counter the problem of a population being too small for statistical sampling, the unit to be selected for audit may be a payment claim by a beneficiary. In no case may the unit of audit be limited to an individual invoice.

For difference estimation, the sampling unit may be an operation or, in exceptional cases where the population is insufficiently large, a payment claim by a beneficiary.

It is expected that the sampling of operations will be carried out at programme level. However, it is not excluded, where the national system makes it more appropriate, that the population is established on the basis of intermediate bodies provided that the population is still sufficiently large to allow for statistical sampling and that the results can be used to support an opinion by the audit authority for each individual programme.

The terms "operation" and "beneficiary" are defined in Article 2 of Council Regulation (EC) No 1083/2006. For aid schemes, each individual project under the aid scheme is considered to be an operation.

Scope of testing of the selected operations

As already indicated above, all operations for which declared expenditure has been included in certified statements of expenditure submitted to the Commission in the reference year should be comprised in the population to be sampled.

Supporting documents should as a rule be checked at 100%. Where there is a large number of the same supporting documents such as invoices or proofs of payment, however, it is accepted audit practice to check a random sample of an adequate size rather than 100%. The sampling methodology should be recorded in the audit report or working papers in such cases. However, if the check reveals a significant level of errors by value or frequency, the sample should be widened to establish more accurately the extent of errors.

Small number of operations in a programme

According to Annex IV of the Regulation, a random statistical sampling method allows conclusions to be drawn from the results of audits of the sample on the overall expenditure from which the sample was taken, and hence provides evidence to obtain assurance on the functioning of the management and control systems. Therefore, it is considered important that the audit authority applies a random statistical sampling method in order to provide the most solid basis for the audit opinion.

However, where the number of operations in a programme is low (less than +/- 800), the use of a statistical sampling approach to determine the sample size may not always be appropriate. The Commission recommends in the first instance to use all possible means to

achieve a sufficiently large population by grouping programmes, when part of a common system, and/or by using as the unit the beneficiaries' periodic payment claims (e.g. quarterly claims will increase the number of items in the population). A statistical sampling method can then be used and the projection of the error rate should be carried out in line with the selected method.

Where it is concluded that the small size of the population makes use of a statistical sampling method not feasible, it is recommended to apply the procedures set out below.

In all cases the principle to be respected is that the size of the sample must be sufficient to enable the audit authority to draw valid conclusions (i.e. low sampling risk) on the effective functioning of the system.

OPTION 1

Examine whether a formal approach to non statistical sampling can be applied (see section 6.6). The advantage of this method is that it determines the size of the sample with reference to a precise confidence level and provides for evaluation of the sample results following a structured approach. The sampling risk is therefore lower than would be the case of informal non statistical methods. It is therefore recommended to apply this method where possible.

However, depending on the size and value of the population, and the number of individually significant amounts, the application of this method may produce a sample size which is disproportionate in the context of the multi-annual audit environment of structural actions programmes.

OPTION 2

Analyse the population and determine whether stratification is appropriate to take account of operations with high value.

Where stratification is applicable, a 100% audit of the high value items should be applied, although a strategy which ensures full coverage of these items over a number of years can be followed.

For the remaining population, determine the size of the sample necessary, taking account of the level of assurance provided by the system. This is a matter of professional judgment, having regard to the principle referred to above that the results must provide an adequate basis for the audit authority to draw conclusions. By way of guidance, it is considered that the number of operations selected would generally be not less than 10% of the remaining population of operations.

Where stratification is not applicable the procedure set out in the previous paragraph is applied to the whole population.

Once the sample size has been determined, the operations must be selected using a random method (for example by using spreadsheet random figures generator).

In practice, the number of operations in a programme may be lower than 800 during the initial stages of the implementation, but build up to a number higher than 800 later in the

programming period. Therefore, although the use of a statistical approach to determine the sample size might not be appropriate at the beginning of the programming period, it should be used as soon as it is feasible to do so.

European Territorial Cooperation (ETC) programmes

ETC programmes have a number of particularities: it will not normally be possible to group them because each programme system is different; the number of operations is frequently low; for each operation there is generally a lead partner and a number of other project partners.

The guidance set out above for the case of programmes with a small number of operations should be followed, taking into account the following additional procedures.

Firstly, in order to obtain a sufficiently large population for the use of a statistical sampling method, it may be possible to use as the sampling unit the underlying validated payment claims of each partner beneficiary in an operation . In this case the audit will be carried out at the level of each beneficiary selected, and not necessarily the lead partner of the operation.

In case a sufficiently large population cannot be obtained to carry out statistical sampling, option 1 or option 2 mentioned under the preceding section should be applied.

For the operations selected, the audit of the lead partners should always be carried out covering both its own expenditure and the process for aggregating the project partners' payment claims. Where the number of project partners is such that it is not possible to audit all of them, a random sample can be selected. The size of the combined sample of lead partner and project partners must be sufficient to enable the audit authority to draw valid conclusions.

Grouping of programmes

The regulation foresees the possibility to group programmes in the case of a common system[11]. This will reduce the number of operations selected per programme.

6.2. Audit planning for substantive tests

Auditing the operations through sampling should always follow the basic structure:

1. Define the objectives of the substantive tests, which corresponds to the determination of the level of error in the expenditure certified to the Commission for a given year for a programme based on projection from a sample.
2. Define the population, which corresponds to the expenditure certified to the Commission for a given year for a programme or for several programmes in

[11] A common system can be considered to exist where the same management and control system supports the activities of several operational programme. The presence of the same key control elements is the criteria to be considered for determining if the management and control systems are the same.

the case of common systems, and the sampling unit, which is the item to sample (e.g. the declared expenditure of the operations).

3. Define the tolerable error: the regulation defines a maximum 2% materiality; the maximum tolerable error and by definition the planning precision is therefore maximum 2% of the expenditure certified to the Commission for the reference year.
4. Determine the sample size, according to the sampling method used.
5. Select the sample and perform the audit.
6. Evaluate and document the results: this step covers the computation of the sampling error1, and the projection of the results to the population.

The choice of a particular sampling method refines this archetypal structure, by providing a formula to compute the sample size and a framework for evaluation of the results.

6.3. Variable sampling

Variable sampling is a generic method. It allows any selection method, and proposes simple projection of the results to the population. However, as it is not specific to the auditing of expenditure amounts and can be used for other purposes as well, it does not offer a specific framework for interpretation of the extrapolated results and the results may not give the appropriate conclusions. The method has been included in the guide for the sake of completeness.

Advantages	Disadvantages
Generic method Fits every type of population	No interpretation framework

6.3.1. Sample size

Computing the sample size **n** within the framework of (generic) variable sampling relies on the usual three values:

- Confidence level determined from system audits (and the related coefficient z from a normal distribution, e.g. 0.84 for 60%, 1.64 for 90% when referring to the parameters in the Commission Regulation (EC) N° 1828/2006)
- Tolerable error **TE** defined by the auditor (at the level of the operations)
- Standard deviation **σ** from the population (in this case the standard deviation of the operations value within a programme can be used).

The sample size is computed as follows:

$$n = \left(\frac{z \times \sigma}{TE} \right)^2$$

Note that the tolerable error (TE) is here defined at the level of the sampling unit (i.e. in most cases the operation). Assuming we name the tolerable error at the level of the population the tolerable misstatement (TM), we have TE = TM / N where N is the population size. Therefore the following formula is also a valid calculation, providing the exact same figure:

$$n = \left(\frac{N \times z \times \sigma}{TM} \right)^2$$

Note that the standard deviation for the total population is assumed to be known in the above calculations. In practice, this will almost never be the case and Member States will have to rely either on historical knowledge (standard deviation of the population in the past period) or on a preliminary sample (the standard deviation of which being the best estimate for the unknown value).

As with most statistical sampling methods, ways to reduce the required sample size include reducing the confidence level and raising the tolerable error.

6.3.2. Sampling error

Sampling implies an estimation error, as we rely on particular information to extrapolate to the whole population. This sampling error[1] (SE) is measured within the framework of variable sampling as follows, based on the sample size, population standard deviation and the coefficient corresponding to the desired confidence level:

$$SE = \frac{z \times \sigma}{\sqrt{n}}$$

Note that the sampling error is based on the actual sample size, which may not necessarily be the exact minimum sample size computed in the previous section. By taking a sample of the exact minimum size required, the sampling error will be equal to the tolerable error, which is a strong limitation because it means that any misstatement encountered in the sample will, through projection, breach the materiality threshold. In order to avoid this, it is wise to pick a sample of a larger size than the exact minimum computed.

6.3.3. Evaluation and projection

Variable sampling in the context of auditing operations of a programme uses the above concepts to estimate the misstatement in the total programme expenditure for the reference year. As observed misstatements are a by-product of auditing operations, the initial calculations (sample size, sampling error) are made based on the operations expenditures.

Based on a randomly selected sample of operations, the size of which has been computed according to the above formula, the average misstatement observed per operation in the sample can be projected to the whole population – i.e. the programme – by multiplying the figure by the number of operations in the programme, yielding the expected population misstatement.

The sampling error can then be added to the expected population misstatement to derive an upper limit to the population misstatement at the desired confidence level; this figure can then be compared to the tolerable misstatement at the level of the programme to draw audit conclusions.

6.3.4. Example of application

Let's assume a population of expenditure[12] certified to the Commission in a given year for operations in a programme or group of programmes. The system audits carried out by the audit authority have yielded a high assurance level. Therefore, sampling this programme can be done with a confidence level of 60%.

The characteristics of the population are summarised below:

Population size (number of operations)	10,291
Book value (sum of the expenditure in the reference year)	2,886,992,919
Mean1	280,536
Standard deviation	87,463

Size of the sample:

1. Applying variable sampling, the first step is to compute the required sample size, using the following formula:

$$n = \left(\frac{z \times \sigma}{TE} \right)^2$$

where z is 0.84 (coefficient corresponding to a 60%[13] confidence level), σ is 87,463 and TE, the tolerable error, is 2% (maximum materiality level set by the Regulation) of the book value divided by the population size, i.e. 2% x 2,886,992,919 / 10,291=5,611. The minimum sample size is therefore 172 operations. Let's assume we take a sample of size 200.

---

12 This data is based on programme data of the 2000-2006 period (cumulative information). The same population is used for the pilot sample in sections 6.4.1 and 6.4.4.

13 Note that with a 90% confidence level, the coefficient 1.64 would be used instead of 0.84, bringing the minimum sample size to 654.

2. The second step is to compute the sampling error associated to using variable sampling with the above parameters for assessing the population, using the following formula:

$$SE = \frac{z \times \sigma}{\sqrt{n}}$$

Where all the parameters are known and n is the size of the sample we have just computed. The sampling error is therefore 5,205.

Confidence level	60%
Tolerable error	5,611
Sample size	200
Sampling error	5,205

3. The third step is to select a random sample of 200 items (operations) out of the 10,291 that make up the population (expenditure declared).

Evaluation:

1. Auditing these 200 operations will provide the auditor with a total misstatement on the sampled items; this amount, divided by the sample size, is the average operation misstatement within the sample. Extrapolating this to the population is done by multiplying this average misstatement by the population size (10,291 in this example). This figure is the expected misstatement at the level of the programme.

Assume that the total misstatement on the sampled items amounts to 120,000€ and as a consequence the average misstatement per operation in the sample is 600€ (i.e. 120,000€ /200); the expected misstatement of the population would be 600 x 10,291 = 6,174,600€.

2. However, conclusions can only be drawn after taking into account the sampling error. The sampling error is defined at the level of the operation; therefore it has to be multiplied by the population size (i.e. 5,205x10,291=53,564,655). This amount is then added to the expected misstatement (see point 1) to find an upper limit to the misstatement within the programme.

3. The upper limit would therefore be the sum of both amounts, giving a total of 59,739,255€. This last amount is the maximum misstatement you can expect in the population based on the sample, at a 60% confidence level. This also means that you have an 80% chance of having a misstatement in the population below 59,739,255€, because a 60% confidence level leaves 40% uncertainty spread over the upper side and the lower side equally, therefore you have an 80% chance of being below that value of a normal probability distribution (see Annex I, I.4.).

5. Finally when compared to the materiality threshold of 2% of the total book value of the programme (2% x 2,886,992,919 = 57,739,858), the upper limit is higher, meaning that as an auditor you would conclude that there is enough evidence that significant (i.e. material) misstatements may exist in the programme, even though the expected misstatement (see point 1) is below the materiality threshold. The only conclusion you can draw is indeed that there is an 80% chance that the given misstatement is below the upper limit (a level that is above the materiality level).

Total misstatement in sample	120,000
Average misstatement in sample	600
Expected misstatement in population	6,174,600
Upper limit to the misstatement	59,739,255
Tolerable misstatement (materiality threshold)	57,739,858

6.4. Variable sampling - difference estimation

Difference estimation relies on the concepts of variable sampling, but provides an additional layer of analysis for projection of the results which makes it well-suited for auditing Structural Funds expenditure. This method, as its name implies, relies on computing the difference between two variables, e.g. in the case of Structural Funds the book value of the declared expenditure and the actual/audited value for all items in the sample. Based on the projection of these differences, an error rate can be determined. For the correct application of the method, it is necessary that sufficient differences are found in order to arrive at a realistic deviation. If there are no or insufficient differences, it is more efficient to use Monetary Unit Sampling (section 6.5).

Although the sample sizes determined under this method may be higher than those calculated using MUS, the projection of the errors is likely to be more accurate where many errors are found.

Advantages	Disadvantages
Interpretation framework Extrapolates book value	Sample size is higher

6.4.1. Sample size

The sample size **n** is computed according to the following formula:

$$n = \left(\frac{N \times Ur \times Sx}{A} \right)^2$$

Whereby:

n is the sample size,N is the population size in number of operations, A is the desired allowance for the sampling error and Sx the standard deviation of the individual differences between each audited value and the book value. The coefficient Ur is a value corresponding to the confidence level (1.64 for 90%, 0.84 for 60%).

Before this method can be applied, it is important to select a pilot sample and determine the standard deviation of the individual differences. This pilot sample can subsequently be used as a part of the sample chosen for audit. In general, a pilot sample of minimum 30 and maximum 50 operations should be drawn. Alternatively, historical data may be used to estimate the standard deviation in the population. This will generally provide more accurate data[14].

[14] The results of all the audits from the 2000-2006 period can be considered. However, the Commission expects that, in that case, the control system applied has not fundamentally changed and that all audit results are considered.

The standard deviation of the individual differences in the pilot sample can be calculated as follows:

SDd = SQRT (cumulative (individual difference – average difference) squared divided by sample size minus 1).

An example is provided below, the data of which is found in Annex II.

Step	Operation	Computation
1	Sample size (pilot or historical data)	30
2	Determine individual differences	See 4th column
3	Sum of Step 2	851,000
4	Step 2 ÷ Step 1	28,367
5	Sum of Square of (Step 2 differences – Step 4)	19,609,591,667
6	Step 5/(Step 1 – 1,0)	676,192,816
7	√(square root of) Step 6	26,004

6.4.2. Sampling error

The allowance for the sampling error (A) is first determined as a function of parameters decided by the auditor:

- the tolerable misstatement TM, defined at the level of the population (programme), which is maximum 2%
- a coefficient Z_α linked to the confidence level (1.64 for 90%, 0.84 for 60%), i.e. linked to type I risk[1] α (100% - confidence level, respectively 10% and 40%)
- a coefficient Z_β linked to the type II risk[1] β, usually set at 1.64 (β=10%)

$$A = \frac{TM}{1 + Z_\beta / Z_\alpha}$$

Note that for all practical aspects, A is actually equal to TM/2 at the level of 90% and close to TM/3 at the level of 60%, based on the parameters provided above. Some variants of the difference estimation method use directly A=TM. If the latter is used, the auditor must be aware that the achieved precision (see section 6.4.3.) may be higher than 2% (TM) and that additional work (i.e. extend sample) may be required in order to obtain an achieved precision equal to or below the allowance for sampling error (desired precision). It is recommended not to set A=TM in case the standard deviation is based on a pilot sample.

6.4.3. Evaluation and projection

Evaluation and projection using difference estimation requires the computation of two values.

First, the achieved sampling precision is defined as follows:

$$A' = \frac{N \times Ur \times Sx}{\sqrt{n}}$$

Sx = same calculation as that used to determine estimated standard deviation of individual differences (pilot sample in section 6.4.1) but applied to the results of the audit.

In principle, for the Structural Funds, the achieved precision (A') should be equal or lower than the tolerable misstatement (TM = 2% of declared expenditure).

Second, the extrapolated book value (EBV) is computed based on the actual book value (ABV):

$$EBV = ABV - N \times \frac{S}{n}$$

whereby S = the sum of the individual misstatements found.

Using the figures computed above, one can then evaluate the results of the sampling:

The first option compares an adjusted EBV to ABV, adjusting EBV with achieved sampling precision A'. If the ABV falls between EBV-A' and EBV+A' (called the precision interval), the population can safely be assumed to have a total misstatement below the materiality level. If that is not the case, it means a misstatement above the materiality level should be assumed.

EBV-A' EBV+A' ABV

EBV

OK if ABV within precision interval

The second option compares EBV to an adjusted ABV, adjusting ABV with tolerable misstatement (TM). If the EBV falls between ABV-TM and ABV+TM (called the decision interval), the population can safely be assumed to have a total misstatement below the materiality level. If that is not the case, it means a misstatement above the materiality level should be assumed.

EBV ABV-TM ABV+TM

ABV

OK if EBV within decision interval

Note that, in the special case of a variant method with S = TM, this decision interval is broader.

Both interval interpretations are valid and interchangeable; the results will always be in line and therefore conclusions can be drawn from both options.

6.4.4. Example of application

Let's assume the population[10] below is being analysed using difference estimation, at a level of confidence of 90%.

Population size (number of operations)	10,291
Actual Book Value (expenditure in a given year)	2,886,992,919

Size of the sample:

1. The first step is to select a pilot sample to determine the standard deviation. The pilot sample should cover between 30 and 50 files and must be randomly selected (see pilot sample calculation in 6.4.1).

2. The second step is to compute the tolerable misstatement, TM, which is 2% of the total book value (2%*2,886,992,919=57,739,858).

3. Then, the allowance for the sampling error (A) is computed: if the risk of incorrect acceptance (Z_β) is set at 10% and the risk of incorrect rejection (Z_α) is set at 20%, then, using the standard table[7] which gives a ratio of 0,50, A = (57,739,858x0,5) = 28,869,929.

4. From this information, a minimum sample size can be computed as $(10,291*1.64*26,004/28,869,929)^2$, which is rounded to 231 items.

Note that by lowering the type I and type II risks, the sample size decreases. Also, if we use a confidence level of 60% instead of 90% (Ur = 0.84) and if the sampling error A is 19,557,049 (about one third of the tolerable misstatement), the sample size required would be lower, or 132 items.

Let's assume that a sample of 231 items is randomly selected and audited, and that a total misstatement of 3,240,374 is found in that sample (i.e. an average misstatement per sampled operation of 14,028), with a standard deviation of the individual misstatements of 25,470.

Evaluation:

1. The first step after the actual audit is the determination of the achieved sampling precision, A', which in the present case amounts to 28,282,928 ($10.291x1.64x25,470 / \sqrt{231}$). As can be seen, the achieved precision is lower than the tolerable misstatement. Therefore, the audit objective has been reached and no additional audit work (i.e. extend the sample) is required.

2. For evaluating the results, the precision interval around the expected book value and the decision interval around the actual book value are described below.

The extrapolated book value is the difference between the declared expenditure (2.886.992.919) and the projected misstatement, i.e. in this case 144,362,148. The auditor's

best judgement is that the actual value is equal to 2,742,630,771 with a precision of an upper and lower bound of 28,282,928.

Precision interval		Actual book value
Lower bound	2,714,347,843	2,886,992,919
Upper bound	2,770,913,699	

Decision interval		Extrapolated book value
Lower bound	2,829,253,061	2,742,630,771
Upper bound	2,944,732,777	

The ABV does not fall within the precision interval and the EBV does not fall within the decision interval; therefore, based on the results of the sample, one can conclude, with a level of confidence of 90%, that there is a material misstatement within this population. In other words, the auditor can state that he is 90% certain that the maximum misstatement in this population is higher than the acceptable materiality level of 2%.

6.5. Monetary unit sampling

Monetary unit sampling (MUS) uses a monetary unit as the sampling unit, but the item containing the sampling unit is selected in the sample (i.e. the operation within the audited programme). This approach is based on systematic sampling (the item containing each n^{th} monetary unit is selected for examination).

MUS provides an implied stratification through systematic sampling, and usually provides a smaller sample size than other methods. Larger items have a much higher chance of being sampled, due to the systematic selection based on monetary interval. Therefore, MUS is also labelled "probability proportional to size" sampling, or PPS. This can be considered either a strength or a weakness, depending on the defined objective of the audit.

When misstatements are found, PPS evaluation may overstate the allowance of sampling risk at a given risk level. As a result, the auditor may be more likely to reject an acceptable recorded amount for the population.

Advantages	Disadvantages
Implied stratification	Assumes low error rate
Small sample size	Geared towards overstatements, not supporting the audit of understatement.
Focus on larger items	Neglects smaller items

6.5.1. Sample size

6.5.1.1. Anticipated misstatement is zero

When the anticipated misstatement is zero, the following simplified sample size formula is used:

$$n = \frac{BV \times RF}{TM}$$

The sample size **(n)** is based on the total amount (BV) of the book value of the expenditure declared for a selected year, the tolerable misstatement (TM) (at maximum acceptable error i.e. the materiality level) and a constant called the reliability factor (RF). The reliability factor is based on Poisson distribution for an expected zero misstatement, and represents at the same time the expected error rate and the desired confidence level:

3 at 95% confidence level

2.31 at 90% confidence level

0.92 at 60% confidence level.

These factors can be found from a Poisson table[13] or from software (e.g. MS Excel).

The sample size is not dependent on the number of items in the population.

The sample is then selected from a randomised list of all operations, selecting each item containing the x^{th} monetary unit, x being the step corresponding to the book value divided by the sample size. For instance, in a programme with Euro 10,000,000 book value, for which we take a sample of size 20, every operation containing the 500,000th Euro will be selected. This implies that in some cases an operation will be selected multiple times, if its value is above the size of the step.

6.5.1.2. Anticipated misstatement is not zero

When the anticipated misstatement is not zero, the following sample size formula is used:

$$n = \frac{BV \times RF}{TM - (AM \times EF)}$$

The anticipated misstatement (AM) or expected misstatement corresponds to an estimate of the Euro misstatement that exists in the population.

The expansion factor[15] (EF) is a factor used in the calculation of MUS sampling when misstatements are expected, which is based upon the risk of incorrect acceptance. It reduces the sampling error.

6.5.2. Evaluation and projection

When no misstatement is found in the sample, the auditor can conclude that the maximum misstatement in the population is the tolerable misstatement (TM). If compared with classical variable sampling and related methods such as difference estimation, this result just implies that our sampling error is equal to the tolerable error.

When misstatements are observed, the auditor must project the sample misstatements to the population. For each misstatement, a percentage of error is computed (e.g. 300€ overstatement on 1,200€ = 25%). This percentage is then applied to the MUS interval (e.g.

15 The Poisson table and values of the EF are extracted from standard tables. An example can be found in the Audit Guide on Audit Sampling, edition as of April 1, 2001 of the American Institute of Certified Public Accountants.

for steps of 4,000€x25%=1,000€). The projected misstatement is the sum of those intermediate results based on element of the lower stratum (value of each sample item is lower than the interval). In case the sample item is greater than the sampling interval (top stratum), the difference between the book value and the audited value is the projected misstatement for the interval (no percentage is calculated).

An upper misstatement limit should be calculated as the sum of the projected misstatements, the basic precision (=MUS step x reliability factor RF for zero or more errors as defined above) and an incremental allowance for widening the precision gap.

Calculation	
+ Basic precision + Most likely misstatement (projected errors from lower stratum plus known errors from top stratum) + Incremental allowance for the sampling error	
= Upper misstatement limit	

The auditor can also calculate an additional sample size needed by substituting the most likely misstatement from the sample evaluation for the original expected misstatement in the sample interval formula and determine the interval and total sample size based on the new expectations. The number of additional sample items can be determined by subtracting the original sample size from the new sample size. The new sampling interval can be used for the selection. Items should be selected that are not already included in the sample.

The incremental allowance is computed for each misstatement (in decreasing value order) as a function of reliability factors for increased number of overstatements at the same level of type I risk. More specifically, each allowance is calculated using the formula below, where RF(n) is the reliability factor for n misstatements at a given confidence level and RF(n+1) the reliability factor for n+1 misstatements at the same confidence level the projected misstatement is multiplied by the difference of reliability factors minus 1 (because already taken into account once).

(RF(n+1)-RF(n)-1)*projected misstatement

For instance, if we observe a single misstatement of 300€ (25%), i.e. a projected misstatement of 1,000€, with a TM of 5,000€ and a MUS step of 4,000€ at a 95% confidence level (confidence factor 3), we have a total of 13,750€ of upper misstatement limit. This figure is the sum of:

- the projected misstatement of 1,000€,
- the basic precision of 4,000€x3=12,000€ and
- the allowance of (4.75-3-1) x 1,000=750€ (4.75 is the RF for 1 misstatement at 95% confidence level, 3 is the RF for 0 misstatements at 95%).

This upper limit is greater than the tolerable misstatement; hence we conclude that the population misstatement is above the materiality threshold. We also conclude that we are 95% sure that the population misstatement is at most13,750€.

6.5.3. Example of application

Let's assume a population as expenditure certified to the Commission in a given year for operations in a programme or group of programmes. The system audits done by the audit authority have yielded a low assurance level. Therefore, sampling this programme can be done with a confidence level of 90%.

The population is summarised in the table below:

Number of operations	10,291
Book value (expenditure in a reference year) –Population size	2,886,992,919
Mean	280,536
Standard deviation	87,463

6.5.3.1. Anticipated misstatement is zero

Size of the sample:

1. Using monetary unit sampling, the first step would be to compute the sample size, using the following formula:

$$n = \frac{BV \times RF}{TM}$$

Where BV is the total amount (i.e. expenditure declared), TM the tolerable misstatement (i.e. 2% materiality level determined by the Regulation x the expenditure declared) and RF is the reliability factor corresponding to an expected 0 misstatement at the 90% confidence level (i.e. 2.31). Based on this information, we calculate the sample size at 115,5 or rounded to 116.

2. The MUS step is computed as the book value divided by the sample size, in that case 24,995,610.

Confidence level	90%
Reliability factor for 0 error	2.31
Sample size	116
MUS step	24,995,610
Tolerable misstatement	57,739,858

Note that with a confidence level of 60%, the reliability factor of 0.92 would be used instead of 2.31, yielding a sample size of 46.

3. The next phase of MUS is selecting the operations from the programme. The list of operations needs to be randomised (i.e. sorted in a random order), then every 24,995,610th Euro is looked up, and the operation containing this Euro is selected into the sample. When the sample is complete, audit procedures take place.

Evaluation:

1. First, if no misstatement was found in the sample, the auditor concludes that, at that level of confidence (90%), evidence shows that the maximum misstatement in the programme is below the materiality level.

However, if misstatements were found, the projection is more complex. Let's assume for the sake of this example that we found a single misstatement of 5,500€ in a 27,500€ item. It represents a 20% error, that has to be extrapolated to the MUS step to find the projected misstatement, here it is 20% x 24,995,610 = 4,999,122€. This is the expected misstatement at the level of the programme; however the auditor needs to compute the upper misstatement limit, which is the maximum misstatement he could find in the population at that level of confidence.

2. The basic precision is equal to the reliability factor used for zero error (2.31) times the MUS step, here 2.31 x 24,995,610 = 57,739,858.

3. The allowance is computed using the formula:

(RF(n+1)-RF(n)-1)*projected misstatement

Where RF(n) is the reliability factor for zero misstatement (2.31) and RF(n+1) is the reliability factor for one misstatement (3.89). The allowance is therefore 2,899,491.

4. The upper misstatement limit is the sum of the projected misstatement, the basic precision, and an allowance for widening the precision gap. The upper misstatement limit in this case is 65,638,471€. As this is above the tolerable misstatement (i.e. the materiality level), the auditor concludes in this example that there is enough evidence in the sample to indicate material misstatements at the level of the population. Additional conclusion is that the auditor is 90% sure the actual misstatement of the population is below 65,638,471€.

Number of misstatement in sample	1
Reliability factor for 1 error	3.89
Total misstatement in sample	5,500
Misstatement error	20%
Projected misstatement	4,999,122
Basic precision	57,739,858
Allowance for widening gap	2,899,491
Upper Misstatement Limit	65,638,471

6.5.3.2. Anticipated misstatement is not zero

Size of the sample:

1. Using monetary unit sampling, the first step would be to compute the sample size, using the following formula:

$$n = \frac{BV \times RF}{TM - (AM \times EF)}$$

Assuming the anticipated misstatement (AM) is calculated as 10% of the tolerable misstatement (TM), the sample size would be about 135[16] items.

Confidence level	90%
Reliability factor for 0 error	2.31
Tolerable misstatement	57,739,858
Anticipated (or expected) misstatement	5,773,986
Expansion factor	1,5
Sample size	135
MUS step	21,385,132

2. The MUS step is calculated as the population divided by the sample size = 21,385,132.

Evaluation:

1. Let's assume for the sake of this example that we found a single misstatement of 5,500€ in a 27,500€ item. It represents a 20% error that has to be extrapolated to the MUS step to find the projected misstatement. Here it is 20% x 21,385,132 = 4,277,026 €. This is the projected misstatement at the level of the programme; however, the auditor needs to compute the upper misstatement limit, which is the maximum misstatement he could find in the population at that level of confidence.

2. The basic precision is equal to the reliability factor used for zero error (2.31) times the MUS step, here 2.31 x 21,385,132 = 49,399,656.

3. The incremental allowance is computed using the formula below:

(RF(n+1)-RF(n)-1)*projected misstatement

Where RF(n) is the reliability factor for zero misstatement (2.31) and RF(n+1) is the reliability factor for one misstatement (3.89). The allowance is therefore 2,480,675.

4. The upper misstatement limit is the sum of the projected misstatement, the basic precision, and an allowance for widening the precision gap.

Calculation
+ Basic precision = 49,399,656 + Most likely misstatement (projected errors from lower stratum plus known errors from top stratum) = 4,277,026 + Incremental allowance for the sampling error = 2,480,675
= Upper misstatement limit = 56,157,357

The upper misstatement limit is 56,157,357 €; as this is below the tolerable misstatement (i.e. the materiality level), the auditor concludes in this example that there is enough evidence in the sample to indicate that there are no material misstatements at the level of the population. An additional conclusion is that the auditor is 90% sure the actual misstatement of the population is not higher than 56,157,357 €.

16 2,886,992,919x2,31/(57,739,858-(5,773,986x1,5))

6.6. Formal approach to non statistical sampling

A formal non statistical sampling plan uses a structured approach to calculate sample size and evaluate sample results. The methods of sample size calculation and sample results evaluation are based on the underlying mathematics of a statistical plan, but the selection of sample items and consideration of sampling risks are normally less rigorous than the statistical plan.

Before starting with the calculation of the sample size, the audit authority should first make a preliminary judgement about materiality. As already indicated before, the materiality level of maximum 2% is applicable to the expenditure declared to the Commission in the reference year. The audit authority can consider reducing the materiality level for planning purposes.

The materiality (or basic allowance) is used in essentially two ways in planning the extent of audit of operations:

1. to determine the cut off for items that are individually significant because of their nature or amount.

2. to calculate sample size for sampling applications.

In examining a specific population, the auditor will want to apply the planned audit procedure to all items that individually significant. The auditor is unwilling to accept any risk of failing to detect misstatements for these items. An item may be individually significant because of its nature or its amount.

To determine the cut off amount for individually significant items, a prudent approach is to divide the materiality (or basic allowance) by 3. The determination of the sample size for the remaining population is explained below.

6.6.1. Sample size

The formula to be applied is the following:

The sample size (monetary hits) is:

$$\frac{\text{Remaining population value} * \text{Confidence factor}}{\text{Planning materiality}}$$

Because the sample size determination is based on the MUS method, the auditor should use one of the two statistical selection methods (see section 6.1.1).

6.6.2. Evaluation and projection

The qualitative evaluation involves investigating the cause of misstatements. This can lead the auditor to apply additional audit procedures, to revise the judgement on the reliability of the management and control systems or to take actions as circumstances dictate.

The quantitative evaluation involves projection of the misstatements in order to determine how much misstatement the remaining population is likely to contain. The methodology is based on MUS and recognises that larger items are selected rather than smaller items. The formula to apply is the following:

$$\frac{\text{Sum of misstatement proportions} * \text{remaining population value}}{\text{Sample size}}$$

6.6.3. Example of application

Let's assume a population of 393 operations for which expenditure has been declared. There is no possibility of increasing this number by e.g. sampling on expenditure claims since beneficiaries send in one claim per reference year. All data used originates from real declarations of expenditure and actual audit results.

The auditors want to assess the validity of the expenditure declared. They consider that the systems work but that improvements are necessary. They wish to be 70% confident about their assessment of the legality and regularity of the expenditure declared. However, caution and due consideration need to be applied when evaluating this, given that their assessment will be based on a non-statistical approach.

The characteristics of the population are summarised below:

The total value of the population is €141.596.219

Materiality is set at 2% = 2.831.924

Determine the individually significant amounts

The first step the auditors will apply is to identify the operations which, individually, represent a significant amount or are significant because of their nature. For the benefit of this example, the individually significant amounts are determined as equal to materiality (2% of €141.596.219). The auditors can also choose to use a lower level of materiality as indicated above.

The selection gives the following results:

Project number	**Amount declared**
297	5.875.013
99	3.343.240
383	3.153.100
388	2.941.442

These projects will be excluded from sampling and will be treated separately. The total value of these projects is €15.312.795.

Sample size

From the remaining population (389 projects), a sample will have to be drawn with 70% confidence. The Confidence Factor to use is that of Monetary Unit Sampling which is, for the confidence level required: 1,21.

This results in a sample size of:

$$126.283.424 * 1,21 \quad = 60 \text{ hits}$$
$$2.525.668$$

The planning materiality used in this example is 1, 7% [17]

Select the sample

The sample should be selected in accordance with the principles systematic sampling (see point 6.1.1.). If other methods are used, it is generally considered appropriate that the sample size should be increased by, at least, 20%.

The sample selected for the operations to audit can be found in annex III.

Audit the sample

The results of the audit are shown in annex III.

The value of the sample is equal to €39.913.723. The total amount of errors in the sample is €1.063.137 (2,7%). The sum of the misstatement proportions amounts to 242,15%.

Evaluating sample results

When the auditor detects misstatements in selected items, two separate evaluations should be made: qualitative and quantitative as described above.

In the example given, the quantitative evaluation (the projection of the errors to the remaining population) leads to the following result:

$$\frac{2,42 * 126.283.424}{60}$$

= € 5.093.431 (4,03% of remaining population value).

The amount of projected errors must be added to the results of the audit of the 100% strata in order to determine the maximum amount of error in the population. In this example, no errors were found in the 100% strata.

The conclusion that can be derived from the exercise is that the auditor can reasonably conclude that the population contains a material error. The difficulty with the non-statistical approach is that the achieved precision cannot be determined. The auditor will therefore have to decide whether to apply additional audit procedures or alternative strategies to evaluate the declared expenditure.

For illustration purposes, the 100% audit of the 393 operations in the population showed an error amount of €5.529.496.

17 In this example, the planning materiality has been reduced.

6.7. Other sampling methods

6.7.1. Ratio estimation

Ratio estimation applies to estimating a ratio between two variables. It is similar to difference estimation, except for the fact that it is based on the ratio of two variables instead of the difference (for instance the ratio of observed value to book value instead of the misstatement which is the difference between observed value and book value).

Just as with difference estimation, ratio estimation provides a small sampling error due to the correlation between variables, and adjusts the sample results to known population data. Sharing the same logic, it has the same strengths and weaknesses as difference estimation, but difference estimation is actually closer to the needs of the Structural Funds from a logical point of view (computing a misstatement rather than a ratio) while being almost identical in all other aspects.

Examples can be found in the reference materials identified in annex VII.

6.7.2. Mean per unit

Mean per unit (MPU) applies to estimating unknown population values. It can therefore be used when the total book value or average misstatement per operation of a population is unknown, but it requires a low variability of the book value per operation because of usually large sample size requirements.

Theoretically, this method fits well the needs of the Structural Funds audit, but the reliance on low variability makes it a poor choice for most populations, while for those populations with low variability it is likely that MUS is the better choice because of the reduced sample size.

Examples can be found in the reference materials identified in annex VII.

6.8. Other considerations

<u>How to determine the expected error (expected misstatement/anticipated error).</u>

The expected error can be defined as the amount of error the auditor expects to find in the population. Factors relevant to the auditor's consideration of the expected error include the results of the test of controls, the results of audit procedures applied in the prior period and the results of other substantive procedures.

In MUS, one of the factors to be used is the expected error (also called anticipated error). In the examples included in the sampling guide, 10% of tolerable misstatement (materiality) has been used.

This is a typical approach generally used in those cases where the expected error is unknown and the use of 10% or 15% of materiality may be considered appropriate for planning purposes. If however the auditor has information on the error rates of previous years, it is recommended to use this figure as it may be more accurate and it will avoid carrying out additional work in case the most likely error from extrapolation is significantly different than the 10% (or 15%) expected error in the planning phase.

Evaluation of misstatements

When applying a statistical method, the audit authority will estimate most likely misstatement in the population and compare this to materiality in order to evaluate the results.

This evaluation of misstatement should be indicated to the Commission in the Annual Control Report[18].
It is expected that the actual known errors found will be corrected. The proof of these corrections should be available.

As indicated in the ISA 530[19] , the auditor should consider the sample results, the nature and cause of any errors identified, and their possible effect on the particular audit objective and on other areas of the audit. It is expected that the audit authority will perform a qualitative in-depth analysis of the misstatements.

In analyzing the misstatements discovered, the audit authority may observe that many errors have a common feature, for example type of transaction, location, responsible body, period of time, or may indicate possible fraud. In such circumstances, the auditor may decide to identify all items in the population that possess the common feature and extend the audit procedures in that stratum. A recommendation must be made for actions to correct all of the affected expenditure. Where there is evidence that earlier declared expenditure might also be affected by the same type of error, all affected expenditure must be identified and corrected.

Sometimes, the auditor may be able to establish that the error arises from an isolated event that has not occurred other than on specifically identifiable occasions and is therefore not representative of errors in the population (an anomalous error). To be considered an anomalous error, the auditor has to have a high degree of certainty that such error is not representative of the population. When an anomalous error has been established, it may be excluded when projecting sample errors to the population. The effect of such an error, if uncorrected, still needs to be considered, as a known error, in addition to the projection of the non anomalous errors.

The audit authority has to report on the actions which have been carried out by the responsible authorities to address the risk of error, which the Commission will then assess.

These actions could for example include:

- Additional testing of operations, leading to the correction of all affected expenditure. This additional testing can be performed by the managing authority under the supervision of the audit authority. These actions aim at error detection.
- Strengthening of controls, providing evidence of effective implementation by way of reduction of errors in subsequent years. These actions aim at error prevention.

---

[18] See Article 62 (1) (d) (i) of Council Regulation No 1083/2006 (OJ L210/25) and Article 18 (2) of the Commission Regulation No 1828/2006 (OJ L45/3)

[19] International Standards Auditing 530 (IFAC)

A high level of errors might also be an indication that the assumptions used when planning the sampling were not correct, e.g. the expected error rate assumption is too low or the confidence level is too high. The audit sample may need to be extended using more appropriate parameters and appropriate action taken in light of the results. Future sampling should take account of the more appropriate parameters from experience gained.

Assessment of results of sampling covering several programmes

The application of the results of an audit from a sample covering several programmes, in the case of grouping of the programmes, will require some special attention. Where the error rate is low, the audit authority should be able to apply the results to all the programmes concerned. However, there may be cases where a concentration of errors is detected in only one part of the system or in only one programme which would require further analysis. Where the error rate exceeds 2%, the audit authority has to analyse the results to establish in which programmes or parts of programmes the irregularities were detected and draw appropriate conclusions. It should however be noted that the results of the sample are valid for the whole population and therefore no separate error rates can be drawn for the individual programmes included.

Complementary sampling

In Article 17 § 5 of the Council Regulation (EC) No 1828/2006, reference is made to complementary sampling.

The results of the random statistical sampling have to be assessed in relation to the results of the risk analysis of each programme and to the coverage of priorities, type of operations, beneficiaries etc in the programme. Where it is concluded from this comparison that the random statistical sample does not address the high risk areas and/or coverage, it should be completed by a further selection of operations, ie a complementary sample.

The audit authority should make this assessment on a regular basis during the implementation period.

The results of the audits covering the complementary sample are analysed separately from the results of the audits covering the random statistical sample. In particular, the errors detected in the complementary sample are not taken into account for the calculation of the error rate resulting from the audit of the random statistical sample. However, a detailed analysis must also be done of the errors identified in the complementary sample, in order to identify the nature of the errors and to provide recommendations to correct them.

The results of the complementary sample should be reported to the Commission in the Annual Control report immediately following the audit of a complementary sample.

Sampling carried out during the year

Based on the timeframe fixed by the Regulations as described in chapter 2, the audit authority has the following options on how to plan the audits of operations:

a) to wait until 01/01/N+1 to start the audit of operations covering the expenditure declared to the Commission in year N;

b) to start the audit as at 01/07/N and take all the expenditure incurred for the period from 01/01/N to 30/06/N as one population. This means that there will be a need to cover the expenditure as from 01/07/N in a second phase through a second population taken at 01/01/N+1. As a result, this option will increase the overall workload;

c) to start the audit as at 01/07/N on the basis of a first sample by determining the total population for the whole of year N by adding the expenditure already declared to the Commission (01/01/N) – 30/06/N) and an estimation of the expenditure to be declared for the second semester of year N. This method has a number of possible risks that should be considered, arising from the possible inaccuracies in the estimation so that the actual final population differs substantially from that estimated.

Therefore, one of the preconditions to applying this approach is that the estimation of the expenditure declared for the second semester of year N can be made accurately. The difference between the estimation and the actual final population should be minor.

When using MUS, it would be necessary to establish the population on the basis of the payment claims submitted by beneficiaries, to determine the total expenditure and to apply the interval to the randomly sorted population of the first half of the year and in a second stage to the actual population of the second half of the year. This could lead to the selection of the same operations.

Using difference estimation, it may be easier to apply sampling during the year. In difference estimation, the sample size will be based on items, i.e. operations/expenditure declarations. Considering that the number of approved projects for which expenditure will be incurred during the reference year may be more stable, a sample may be randomly selected, based on expenditure declarations (in case more than one declaration of expenditure is required). This sample will be a good estimator to determine the expected standard deviation of the population and will serve as a basis for calculating the sample size of operations to audit once the population is known.

The Regulation foresees that the Annual Control Report presented on 31 December of year N+1 relates to the audit work done during the period 01/07/N to 30/06/N+1. This means that the field work should be finalized by 30/06/N+1 and validation of findings may be completed during the period 30/06/N+1 and 31/12/N+1 (date of reporting to the Commission).

Change of sampling method during the programming period

If the audit authority is of the opinion that the sampling method initially selected is not the most appropriate one, it could decide to change the method. However, this should be notified to the Commission in the framework of the Annual Control Report or in a revised audit strategy.

Sampling of operations in consecutive years

In practice, it could happen that the same operations are selected for sampling in consecutive years. There can be no derogation from maintaining the operation in the sample since otherwise the results drawn from the statistical sample will be prejudiced. Therefore the operation should be audited again. Every operation is potentially auditable every year as regards expenditure relating to the particular year, and beneficiaries should be aware of this.

However, the scope of the audit will be different from one year to another. During a second audit the horizontal aspects, such as public procurement would not need to be covered again, and it would therefore be a lighter process.

Where it is expected that some operations are likely to be selected every year due to their high value, the audit authority should consider the use of stratification.

7. TOOLS FOR SAMPLING

The complexity of audit sampling methods underlines the need to rely on appropriate software. A broad range of software can help the auditor apply sampling methods, from standard office software, such as MS Excel, to specific data management/data mining software, like SPSS and SAS, and the obvious audit-dedicated software, such as ACL (Audit Command Language) or IDEA (Interactive Data Extraction and Analysis).

Regarding sampling methods, audit-specialised software can perform data stratification, sample extraction and statistical analysis. Non-dedicated software can provide the same features, though the most basic tools such as Excel or Access only provide a basic structure through formulas.

The most useful formulas to be used for sampling and included in Excel are mentioned in Annex IV.

The advantages offered by these tools are many. First of all, auditors do not need to remember many complicated formulas. These statistical formulas are already embedded in the software and by inputting the necessary parameters, the system provides reliable calculations. Secondly, these tools are fast allowing auditors to save time. Thirdly, selections operated by the software are not influenced by subjective factors that could, on the other hand influence the auditor in a manual selection. Furthermore, audit-dedicated software available in the market offers many audit-specific features, and provides documentation of each test performed that can be used as documentation in the audit working papers.

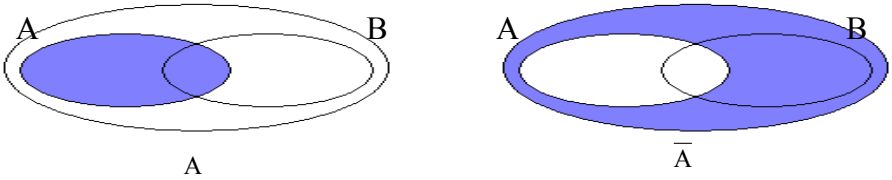
A disadvantage is that auditors may tend to use mechanically the software. The purpose of presenting audit sampling methods with a high level of detail, knowing that most technicalities can be handled by a computer, lies in demonstrating that understanding what the tool performs is key to using it correctly. For instance, monetary unit sampling is a powerful method when expecting few misstatements, but extrapolating the observed misstatements to the population may produce unreliable results when the number of errors rises; the calculations are performed by the computer, but the computer does not indicate how the successive intermediate projections impact on the reliability of the final result (see the computation of the upper misstatement limit in the MUS section 6.5).

Annexes

Annex I Theory of statistical and non statistical sampling methods

I.1. Probability theory

Probabilities are associated to events; events can easily be compared using set notation.



Event A and its complement (non-A)

Example:
A: projects with value more than € 1.000.000
Non-A: projects with value less than or equal to € 1.000.000

Assuming two events, A and B, we can further define intersection and union of events.

Example:
A: projects with value more than €1.000.000
B: projects in the field of wind power
A∩B: wind power projects valued at more than €1.000.000
AUB: projects with high value and/or in wind power field

The complement of "A or B" is "Non-A and Non-B", while the complement of "A and B" is "Non-A or Non-B".

Probabilities are formally defined by the three following rules.

Rule 1: a probability is a number between 0 and 1 (i.e. 0% and 100%)
e.g.: P(even roll on a die toss)=50%; P(odd roll on a die toss)=50%
Rule 2: the probabilities of all possible events add up to 100%
e.g.: P(even)+P(odd)=50%+50%=100%; P(1)+P(2)+P(3)+P(4)+P(5)+P(6)=100%
Rule 3: the probability of one of several mutually exclusive events happening is the sum of the probabilities of those events

e.g.: P(roll is less than 4)=P(1)+P(2)+P(3)=50%

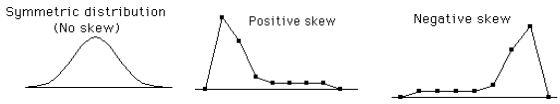
I.2. Descriptive statistics

Descriptive statistics are figures summarising a data set in a few facts, e.g. the average value of a set is a descriptive statistic.

Measures of central tendency are figures describing the expected values in a random environment. This includes the mean (i.e. the average value), the median (the middle value) and the mode (the most frequent value).

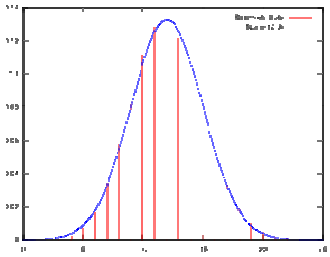
Measures of dispersion try to summarise the variability of data in a set. The usual measure of volatility is the variance, which measures the "distance" between data and their mean. The standard deviation is the square root of the variance; it is useful because it is on the same scale as the data; it can therefore be used to express volatility as a ratio (e.g. standard deviation divided by the mean will provide a percentage of variability).

The skewness is a measure of the symmetry of data around the mean; in other words, it measures whether data are equally spread on both sides of the average value (zero skewness) or if extreme values appear on the upper side (positive skewness) or lower side (negative skewness).



I.3. Normal probability distribution

The most widely used distribution is the normal distribution, due to its theoretical value and the many shortcuts it offers in many areas of statistics – including sampling. The normal distribution is a bell-shaped curve (vertical axis is frequency, horizontal axis is the value observed) defined by its mean and its variance.

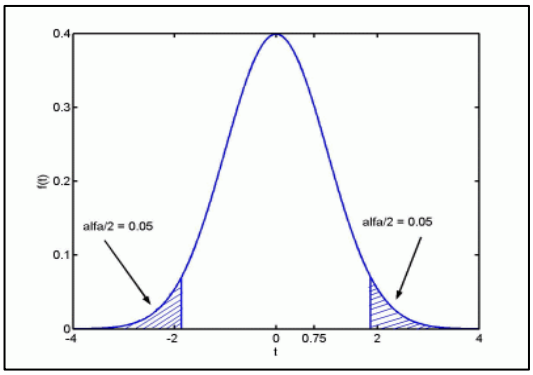


The standardised normal distribution is a special case with mean 0 and variance 1; it is used in many areas as a reference curve, since every existing normal distribution can be compared to the standardised curve quite easily through what is called "standardisation" (a transformation of any normal distribution into a standardised normal distribution with mean 0 and variance 1).

I.4. Confidence level

Assumptions and sample results lead to uncertainty, because of the particular analysis that is made of the actual population. This is reflected in a confidence level: the probability of having the looked-up value falling within a proper interval. The confidence level is often denoted “1-alpha”, where alpha is often called the significance level.

The most frequent use of a confidence level is in a confidence interval, i.e. an interval for which the probability of having the looked-up value falling within its boundaries equals the confidence level. The picture below displays such an interval on a normal curve (the white area) at a level of confidence of 90%, leaving 5% above and below the boundaries.



As a rule of thumb, assuming a bell-shaped distribution, an empirical rule says that, based on the mean μ and the standard deviation σ,

μ±σ contains approximately 68% of the data
μ±2xσ contains approximately 95% of the data
μ±3xσ contains almost all data.

Annex II Example for the calculation of a standard deviation

(for the pilot sample in section 6.4.1.)

Item	Book value	audit value	difference	average sample difference	(Di - Dn)2
1	150.000	116.000	34.000	28.367	31.734.444
2	342.000	308.250	33.750	28.367	28.980.278
3	456.000	401.000	55.000	28.367	709.334.444
4	768.000	768.000	0	28.367	804.667.778
5	23.000	0	23.000	28.367	28.801.111
6	1.090.500	1.075.500	15.000	28.367	178.667.778
7	387.500	350.000	37.500	28.367	83.417.778
8	2.500.000	2.500.000	0	28.367	804.667.778
9	5.000.000	4.935.000	65.000	28.367	1.342.001.111
10	450.000	450.000	0	28.367	804.667.778
11	2.350.000	2.290.000	60.000	28.367	1.000.667.778
12	459.500	436.500	23.000	28.367	28.801.111
13	980.000	980.000	0	28.367	804.667.778
14	7.000.500	6.950.500	50.000	28.367	468.001.111
15	105.000	70.000	35.000	28.367	44.001.111
16	56.000	43.500	12.500	28.367	251.751.111
17	12.000.000	11.920.000	80.000	28.367	2.666.001.111
18	750.000	705.000	45.000	28.367	276.667.778
19	250.000	217.500	32.500	28.367	17.084.444
20	500.000	480.000	20.000	28.367	70.001.111
21	100.000	100.000	0	28.367	804.667.778
22	230.000	230.000	0	28.367	804.667.778
23	650.000	620.500	29.500	28.367	1.284.444
24	450.000	450.000	0	28.367	804.667.778
25	850.000	828.750	21.250	28.367	50.646.944
26	10.000.000	9.910.000	90.000	28.367	3.798.667.778
27	980.000	980.000	0	28.367	804.667.778
28	440.000	429.000	11.000	28.367	301.601.111
29	650.000	583.000	67.000	28.367	1.492.534.444
30	275.000	264.000	11.000	28.367	301.601.111
∑	**50243000**	**49.392.000**	**851.000**	**∑(Di - Dn)2**	**19.609.591.667**
				Standard deviation	**26.004**
				Sample size	**231**

Annex III Example of non-statistical sampling

Sample selected

Project number	Amount declared
11	166.902
16	545.000
31	237.916
40	821.048
44	643.775
53	310.868
61	195.049
71	386.178
78	455.591
84	542.841
88	770.601
90	396.414
93	1.066.056
95	1.384.586
97	652.783
100	1.427.310
105	308.857
109	484.994
118	75.728
129	375.476
136	98.040
146	1.217.661
149	1.012.018
152	682.857
154	133.073
163	374.368
167	1.140.535
174	870.830
176	348.378
192	917.940
197	189.185
205	377.540
215	987.676
218	1.954.073
219	1.596.306
221	880.190
226	387.768
233	319.035
245	464.511
253	470.482
259	99.140
273	535.165
280	501.653

286	231.974
295	121.929
306	234.388
316	518.008
324	587.694
330	527.649
338	680.629
343	459.577
350	419.180
359	289.467
368	573.957
370	1.425.054
375	830.078
382	2.323.053
385	282.995
386	2.660.525
393	941.169

Results of the audit:

Project number	**Error rate**	**Error**	**Amount declared**
11	0,00%	0	166.902
16	0,00%	0	545.000
31	0,00%	0	237.916
40	0,14%	1.109	821.048
44	4,35%	28.000	643.775
53	0,00%	0	310.868
61	0,00%	0	195.049
71	1,42%	5.492	386.178
78	0,00%	0	455.591
84	0,00%	0	542.841
88	0,00%	0	770.601
90	1,38%	5.462	396.414
93	6,33%	67.452	1.066.056
95	0,00%	0	1.384.586
97	0,00%	0	652.783
100	0,00%	0	1.427.310
105	0,87%	2.679	308.857
109	0,59%	2.840	484.994
118	0,00%	0	75.728
129	0,00%	0	375.476
136	29,60%	29.020	98.040
146	9,48%	115.439	1.217.661
149	0,00%	0	1.012.018
152	0,00%	0	682.857
154	0,00%	0	133.073
163	0,00%	0	374.368
167	1,84%	20.969	1.140.535

174	0,00%	0	870.830
176	20,54%	71.571	348.378
192	1,76%	16.126	917.940
197	1,86%	3.513	189.185
205	0,26%	975	377.540
215	0,00%	0	987.676
218	1,90%	37.115	1.954.073
219	0,00%	0	1.596.306
221	0,00%	0	880.190
226	0,00%	0	387.768
233	0,00%	0	319.035
245	1,61%	7.497	464.511
253	0,87%	4.093	470.482
259	0,00%	0	99.140
273	1,46%	7.840	535.165
280	0,00%	0	501.653
286	1,98%	4.596	231.974
295	58,56%	71.396	121.929
306	0,00%	0	234.388
316	23,14%	119.868	518.008
324	0,00%	0	587.694
330	0,00%	0	527.649
338	0,03%	210	680.629
343	53,87%	247.594	459.577
350	0,02%	84	419.180
359	0,00%	0	289.467
368	0,56%	3.206	573.957
370	0,00%	0	1.425.054
375	15,47%	128.391	830.078
382	0,00%	0	2.323.053
385	0,00%	0	282.995
386	2,28%	60.600	2.660.525
393	0,00%	0	941.169
	242,15%	1.063.137	39.913.723

Annex IV MS Excel formulas to assist in sampling methods

The formulas listed below can be used in MS Excel to assist in computing the various parameters required by the methods and concepts detailed in this guidance note. For further information on the way these formulas work, you can refer to the Excel "help" file that provides the details of the underlying mathematical formulas.

=AVERAGE(.) : average of a data set

=VAR(.) : variance of a sample

=VARP(.) : variance of a population

=STDEV(.) : standard deviation of a sample

=STDEVP(.) : standard deviation of a population

=NORMSDIST(.) : probability from standardised normal distribution

=NORMSINV(.) : value from standardised normal distribution

=NORMDIST(.;.;.;.) : probability from normal distribution, for which you specify a mean and variance

=NORMINV(.;.;.) : value from normal distribution, for which you specify a mean and variance

=CHIDIST(.;.) : probability from chi squared distribution

=CHIINV(.;.) : value from chi squared distribution

=RAND() : random number between 0 and 1, taken from a uniform distribution

Annex V Symbols reference guide

The symbols listed below are standard notations used in sampling and statistics.

N : population size

n : sample size

x_i : values from a population or sample

$\bar{x}$: average from a sample

μ : average from the population

σ : standard deviation of the population

s : standard deviation from a sample

z : standardised normal distribution

k : (= Ur) coefficient for confidence intervals

Annex VI List of commonly used terminology

Application controls: Refer to the transactions and data relating to each computer-based application system and are therefore specific to each such application. The objectives of application controls, which may be manual, or programmed, are to ensure the completeness and accuracy of the records and the validity of the entries made therein resulting from both manual and programmed processing. Examples of application controls include data input validation, agreement of batch totals and encryption of data transmitted.

Audits of operations: Synonymous with substantive tests of details or transactions relating to operations in a programme.

Central limit theorem: A foundation rule of estimation, stating the sampling distribution converges to a standardised normal distribution when the sample size is large enough.

Confidence level: A measure of the likelihood of a test or result.

Descriptive statistics: Collective name associated to figures summarising a set of data, e.g. the mean, variance, etc…

Estimator: A value taken from a sample used to estimate a population's parameter.

Inferential statistics: Area of statistics devoted to estimating unknown population parameters.

Mean: The average value of a set of data.

Median: The middle value of a set of data (in the sense of being e.g. the 4th value in an ordered list of 7 items).

Mode: The most frequently observed data in a set.

Population: The population is the entire set of data (all the items constituting a class of transactions or account balance) from which the auditor wishes to sample in order to reach a conclusion on the population. Therefore the population from which the sample is drawn has to be appropriate and verified as complete for the specific audit objective.

Sample: Any subset of the population. The items selected for examination.

Sampling distribution: Theoretical distribution of sample parameters used as estimators of population values e.g. when trying to estimate the mean of a population, you take a sample – the sample's mean is your best estimator. If you take a second sample, you will very likely get a different figure, which is as good an estimator as the first. Assuming you can take many samples, all the sample averages you can get will be distributed along what is called the sampling distribution.

Sampling error: Sampling implies an estimation error, as we rely on particular information to extrapolate to the whole population. This sampling error can be measured and represents the inaccuracy related to selecting a sample of a certain size to represent the population.

Synonymous with allowance for sampling risk and desired precision.

Significance level: The type I risk incurred (among others) when sampling, always equals to 100% minus the confidence level.

Standard deviation: A measure of dispersion of the data, which is calculated as the square root of the variance; useful because it has the same scale as the data and can be compared directly.

Standardised normal distribution: A normal distribution with mean 0 and variance 1; very useful for defining confidence intervals, sampling error etc.

System audits: Detailed tests of controls and reporting, which are intended to provide evidence about the effectiveness of the design and operation of a control system in preventing or detecting material misstatements and about the organisation's ability to record, process, summarize and report data.

Tolerable deviation rate: The maximum population rate of deviations from a prescribed control that the auditor will tolerate without modifying the planned assessment of control risk.

Type I risk(α): The risk of incorrect rejection; in auditing, it is the risk of incorrectly stating that the population does not conform to expectations, based on a sample.

Type II risk (β): The risk of incorrect acceptance; in auditing, it is the risk of incorrectly stating that the population does conform to expectations based on a sample.

Variance: A measure of dispersion of the data, function of the "distance" between data and their mean.

Misstatement: The overstatement of an amount of expenditure. Used synonymously with error and irregularity.

Irregularity: Any infringement of a provision of Community law resulting from an act or omission by an economic operator which has, or would have, the effect of prejudicing the general budget of the European Union by charging an unjustified item of expenditure to the general budget. [14]

Operation: A project or group of projects selected by the managing authority of the operational programme concerned or under its responsibility according to criteria laid down by the monitoring committee and implemented by one or more beneficiaries allowing achievement of the goals of the priority axis to which it relates. [20]

20 Article 2 (Definitions) of Council Regulation (EC) N° 1083/2006

Annex VII List of reference material

"Mathematical Statistics with Applications", Wackerly, Mendenhall and Sheaffer

"AICPA Audit Guide on audit Sampling" -American Institute of certified public accountant

"IS Auditing Guideline – Audit sampling" - Information Systems Audit and control Association (ISACA)

"Audit sampling – an introduction", Dan M. Guy, Douglas R. Carmichael, Ray Whittington

"Sampling – a guide for internal auditors", the iia research foundation handbook series

Intosai standard on Competence 2.2.37 -Approved by the XVIIth Congress of INTOSAI in Seoul 2001

International standard on auditing 530 "AUDIT SAMPLING AND OTHER MEANS OF TESTING" – IFAC

International Standards for the Professional Practice of Internal Auditing (Standard 2100) - The Institute of Internal Auditors (IAA)

"Risk assessment standards" - AICPA's Auditing Standards Board (ASB), March 2006.

Annex VIII Sampling techniques applicable to system audits

VIII.1. Audit planning for system audits

When applying attribute sampling in a system audit, the following generic six-step plan is applied.

1. Define the test objectives: for instance, determine whether the error frequency in a population meets the criteria for a high assurance level;

2. Define the population and sampling unit: for instance the invoices allocated to a programme;

3. Define the deviation condition: this is the attribute being assessed, e.g. the presence of a signature on the invoices allocated to a operation within a programme;

4. Determine the sample size, according to the formula below;

5. Select the sample and carry out the audit (the sample should be selected randomly);

6. Evaluate and document the results.

VIII.2. Determining the sample size in attribute sampling

The required sample size depends on three parameters:

1. The risk of assessing control risk too low, set usually at 5% or 10% (implied confidence level of respectively 100%-5% = 95% or 100%-10% = 90%); note that this parameter is not determined by the Regulation;

2. The tolerable deviation rate, **T**, determined by the auditor; the tolerable levels are set by the Member State audit authority (e.g. the number of missing signatures on invoices under which the auditor considers there is no issue);

3. The expected population deviation rate, **e**, estimated or observed from a preliminary sample. Note that the tolerable deviation rate should be higher than the expected population deviation rate, as, if that is not the case, the test has no purpose (i.e. if you expect an error rate of 10%, setting a tolerable error rate of 5% is pointless because you expect to find more errors in the population than you are willing to tolerate).

The sample size **n** is computed as follows, with "**k**" being a coefficient corresponding to the confidence level based on a normal distribution (e.g. 1.96 at 95%, 1.64 at 90%).

$$n = \frac{k^2 \times e \times (1-e)}{T^2}$$

Example: assuming a confidence level of 95% (k=1.96), a tolerable deviation rate (T) of 12% and an expected population deviation rate (e) of 6%, the minimum sample size would be $(1.96)^2 x(0.06)x(1-0.06)/(0.12)^2 = 15.05$, rounded up to 16 items.

Note that the population size has no impact on the sample size; the calculation above slightly overstates the required sample size for small populations, which is accepted. Ways to reduce the required sample size include reducing the confidence level (i.e. raising the risk of assessing the control risk too low) and raising the tolerable deviation rate.

For large populations (above 5,000 items), statistical tables can be used instead of the above standard formula; For smaller populations, these tables however overestimate the required size.

VIII.3. Evaluating the results

The number of deviations observed in the sample divided by the number of items in the sample (i.e. the sample size) is the sample deviation rate. This is also the best estimation of the population deviation rate one can obtain from the sample.

The achieved upper deviation limit is a theoretical figure based on the sample size and the number of errors encountered. It represents the maximum error rate of the population at the defined confidence level and results from binomial tables (for instance, for sample size 150 and an observed amount of deviations of 3, the maximum error rate (or achieved upper deviation limit) at a 95% confidence level is 5.1%).

If this percentage is higher than the tolerable deviation rate, the sample does not support the assumed expected error rate of the population at that confidence level. The logical conclusion is therefore that the population does not meet the criterion set of high assurance level and must be classified as having an average or low assurance level. Note that the threshold at which low, average or high assurance is reached is defined by the Member State audit authority.

The achieved precision is defined in attribute sampling as the difference between the achieved upper deviation limit and the sample deviation rate; it is in other words the "distance" between the observed deviation rate and the maximum rate that could have been observed. Assuming we observed a sample deviation rate of 2% (3 deviations observed from 150 items), the achieved upper deviation limit is, as stated above, 5.1%, and therefore our achieved precision is 3.1%.

VIII.4. Specialised methods of attribute sampling

Attribute sampling is a generic method, and therefore some variants have been designed for specific purposes. Among those, discovery sampling and stop-or-go sampling serve specialised needs.

Discovery sampling aims at auditing cases where a single error would be critical; it is therefore particularly geared towards the detection of cases of fraud or avoidance of controls. Based on attribute sampling, this method assumes a zero (or at least very low) rate of error and is not well suited for projecting the results to the population, should errors be found in the sample. Discovery sampling allows the auditor to conclude, based on a sample, whether the assumed very low or zero error rate in the population is a valid assumption. It is not a valid method for assessing the level of assurance of internal controls, and therefore is not applicable to system audits.

Stop-or-go sampling comes out of the frequent need to reduce the sample size as much as possible. This method aims at concluding that the error rate of the population is below a predefined level at a given confidence level by examining as few sample items as possible – the sampling stops as soon as the expected result is reached. This method is also not well-suited for projecting the results to the population, though it can be useful for assessing system audit conclusions. It can be used when the outcome of system audits is questioned, to check whether the criterion is indeed reached for the assurance level provided.