DRUGS WORKING GROUP

Guidelines on Sampling of Illicit Drugs for Qualitative Analysis

Second Edition

(2016)

INDEX

Foreword to the Second Edition

In 2012, revised versions of the Excel sampling calculator and the guidance document 'Hypergeometric Sampling Tool (version) 2012): Background of Calculation and Validation' were published for the forensic community on the ENFSI website.
As the next step it was necessary to revise the corresponding guidelines on representative drug sampling as well. The changes which were made can be summarised as follows:

Statistical sampling by Hypergeometric distribution:

- **Rounding down versus rounding up**
- **Handling very large populations with very high 'k' and 'CL' values.**
- **Splitting HPD into separate 'proportion' & 'number' worksheets**
-

Statistical sampling by Bayesian approach:

- **Splitting Bayesian into separate 'N<50' & 'N>=50' worksheet**

The ENFSI Drugs Working Group is proud to present hereby the second edition of the guidelines.

Dr Udo Zerell
ENFSI DWG Chairman 2013-2016

We would like to express our special thanks to the members of the Hypergeometric Sampling Subcommittee for taking over the scientific part of the revision:

Dr Sonja KLEMENC
Ministry of the Interior RS
National Forensic Laboratory
Vodovodna 95
1000 Ljubljana
e-mail: sonja.klemenc@policija.si

Hugh COYLE
Forensic Science Ireland
Garda Headquarters,
Phoenix Park, Dublin 8, Ireland
e-mail: hjcoyle@fsl.gov.ie

List of further contributors:

Dr Michael Bovens, Forensic Science Institute, Zurich, Switzerland.

Tomislav Houra, Dr Maja Jelena Cop and Dr Ines Gmajnick, Forensic Science Centre, Zagreb, Croatia.

Dr Laurence Dujourdy, National Forensic Science Institute, LPS Lyon, France.

Second Edition

DWG Contact
John Power, Forensic Science Ireland, Garda Headquarters,
Phoenix Park, Dublin 8, Ireland.

Reviewer

Prof. C G G Aitken, University of Edinburgh, Scotland (UK).

Foreword to the First Edition

In the ENFSI Working Group (WG) Drugs, quality assurance (QA) and best practice are important topics. QA is an extensive field where a lot of documentation and various guidelines are already available. For this reason, the WG Drugs meeting in Krakow in 1999 decided to focus on a number of other topics that were more or less specific for the WG Drugs. So far, two targets were chosen, the first one being sampling (and sampling strategy), the second one being drugs reference compounds.

In the Madrid meeting of 2001, many members of the WG showed their interest and offered help with the realisation of a manual or a discussion document on sampling. Out of these members the steering committee composed a subcommittee with - for practical reasons - a limited number of colleagues. The composition was such that members had own practical experience with at least one of the sampling methods. A first draft was presented in March 2002 by email and later in May 2002 at the WG meeting in Oslo. The comments, discussions and responses resulted in the document that is presented here and was adopted at the WG meeting in Istanbul 2003.

The primary task of the subcommittee was to identify and describe common sampling procedures. From this, a discussion could be initiated whether there was a sampling method that was superior to the other ones.

In the meantime it became clear that a mere collection of possible sampling strategies was not sufficient. So, on the European level a sampling proposal was submitted by the Spanish presidency (2002) to the Drugs Trafficking Working Party; it comprised the use of an arbitrary sampling method as the European standard; later the proposal was extended with another (different) method. Since ENFSI was working on sampling, the discussion of the proposal was postponed. During the Greek presidency, an ENFSI advice on this matter was

urgently requested, as decisions had to be made by June 2003. The WG steering committee was pleased to have, as requested, the sampling document ready in spring 2003. This document formed the basis of the advice, that was formulated and accepted by both the sampling subcommittee and the WG steering committee; next the (draft) document and the advice were presented by our chairman to the Police Cooperation Working Group on June 17th 2003 in Brussels.

The steering committee is proud of the realisation of this sampling document.

We wish to express our thanks to the chairperson and members of the subcommittee on sampling for their excellent work.

The ENFSI WG Drugs,

Dr Henk Huizer (Chairperson 2000-2002)

Dr Erkki Sippola (Chairperson 2002-2006)

LIST OF CONTRIBUTORS to First Edition

Sergio SCHIAVONE (Chairman of the Sampling Subgroup)
Raggruppamento Carabinieri Investigazioni Scientifiche
Reparto di Roma, Sezione di Chimica
Via Aurelia 511, 00165 Roma, Italia

Martine PERRIN
Institut de Recherche Criminelle de la Gendarmerie Nationale
Department Toxicologie
1, Boulevard Theophile Sueur, F-93111, Rosny Sous Bois Cedex, France
Phone 0033-1-49355079, Fax 0033-1-49355027, e-mail

Hugh COYLE
Forensic Science Ireland
Department of Justice, Garda Headquarters
Phoenix Park, Dublin 8, Ireland

Henk HUIZER
Netherlands Forensic Institute
Volmerlaan 17, 2288 GD Rijswijk, Netherlands

Annabel BOLCK
Netherlands Forensic Institute
Volmerlaan 17, 2288 GD Rijswijk, Netherlands

Bruno CARDINETTI
Raggruppamento Carabinieri Investigazioni Scientifiche
Reparto di Roma, Sezione di Balistica
Via Aurelia 511, 00165 Roma, Italia

Chapter: INTRODUCTION

The sampling document itself describes a number of sampling methods, from arbitrary methods to methods with a statistical background. The document focuses on sampling in cases where large numbers of relatively homogeneous material are available. It does not deal with so-called tactical sampling, which may be applied for house-searches or in clandestine laboratory investigations. These cases are characterised by different materials, sometimes in different amounts, different packages and/or sometimes with different suspects; these cases are considered as so specific and so dependent on the situation (also in legal aspects) that a guideline would be inadequate in many cases.

Thus, the document contains a number of sampling strategies for cases with large numbers of items of relatively homogeneous material. However, from the descriptions of the sampling methods, it is not automatically clear which strategy should be preferred (or would be optimum). This is mainly due to the fact that it is not possible to define a sampling strategy, if the requirements have not been defined. This is the main reason why it was decided to refrain from giving advice at local, regional or national level. ENFSI cannot give such a fine-tuned advice as is possible in a specific agreement between prosecutor, police, chemist and laboratory management.

However, at the European level, advice was requested and the steering committee felt that if one group were competent to produce one, it would be the ENFSI WG Drugs. The resulting advice for international cases is mainly based on a number of aspects, which are discussed in the chapter on 'Considerations'. Here, the advantages and disadvantages of the various methods, also in relation with sampling practice, are brought up. It seems that a Bayesian approach is a reasonable one in many cases, but its complexity might be a major drawback, especially for court. Luckily enough, the hypergeometric and Bayesian

approaches appear to show more or less the same results in cases where no prior probability is used.

Since sampling is often carried out by police and customs, we did not want to introduce an advice with the number of samples to be calculated for each separate case; this would bother them with computers or lists with Bayesian and hypergeometric tables. Therefore the final sampling advice just mentions the number of samples to be taken (5, 8 or 11, the number of samples being dependent on the circumstances). The final evaluation and probability calculations can, if necessary, be performed by the forensic laboratory.

Many WG members have contributed to this document. In the first place the subcommittee on sampling, who have studied, considered, drafted, discussed and finalised it. In addition, many members have contributed either by their written response or comments in any other form; their comments, support and enthusiasm was a fruitful input for the subcommittee. We wish to express our thanks to the contributors and their management. We are convinced that the members of the ENFSI WG Drugs and other colleagues working in the field of drug analysis will benefit from this work; it provides them with a reference on which they can develop an appropriate good working practice.

Dr. Sergio Schiavone

Chairperson of the Sampling Subcommittee

November 2003

Chapter: DEFINITIONS

1. Seizure

The entire quantity of items seized. This may consist of a single population or a number of populations.

2. Population

The collection of items under discussion. A population may be real or hypothetical; finite or infinite; homogeneous or heterogeneous. For the purposes of this booklet, the term population will refer to a real, finite homogeneous population unless otherwise specified.

3. **Package**

A container for a single unit, a number of units or a number of other sub-packages.

4. Unit

A single individual element of a population (e.g. a single tablet or a single package containing powder etc.).

5. Sample

A unit or a number of units selected from a population.

6. Mean

This is the average value of a set of measurements. The mean can refer to either:

i) The arithmetic mean of a *population*. This is the true mean calculated from the entire population. It is denoted by μ. Or

ii) The arithmetic mean of a *sample*. This is an estimate of μ calculated from a sample of the population. It is denoted by $\bar{X}$.

Unless otherwise stated, the term 'mean' will refer to the arithmetic mean of a sample as described in 6 (ii).

7. Standard Deviation

This is a measure of the variation in the values of a set of measurements. The standard deviation can refer to either:

i) The standard deviation of a *population*. This is the true standard deviation calculated from the entire population. It is denoted by σ. Or

ii) The standard deviation of a *sample*. This is an estimate of σ calculated from a sample of the population. It is denoted by s.

Unless otherwise stated, the term 'standard deviation' will refer to the standard deviation of a sample as described in 7 (ii).

Symbols

$P =$	probability
$N =$	population size
$N_1 =$	number of positives in the population
$n =$	sample size
$X =$	number of positives in the sample
$x =$	the value of number of positives in the sample
$r = n - x =$	the value of the number of negatives in the sample
$\theta = \frac{N_1}{N} =$	proportion of positives in the population
$K =$	threshold number of positives guaranteed in the population
$k = K/N =$	proportion of positives at least guaranteed in the population (hypergeometric sampling)
$\alpha =$	probability of rejecting H_0 when H_0 is true
1- α =	probability of not rejecting H_0 when H_0 is true
$(1-\alpha)100\% =$	confidence level

H_0 =	null hypothesis
H_1 =	hypothesis alternative (opposite) to H_0
M_0 =	the highest integer lower than K at which H_0 is tested (hypergeometric distribution)
a =	first parameter of beta function
b =	second parameter of beta function
Y =	number of positives in the unexamined units
μ =	the arithmetic mean in the population
$\bar{X}$ =	the arithmetic mean in the sample
σ =	the standard deviation in the population
s =	the standard deviation in the sample
w =	the total weight in the sample
W =	total estimated weight in the population
P_{corr} =	correction factor in weight estimation
Q_{corr} =	correction factor in weight estimation
RoundUp	rounding number up (for example: RoundUp (to zero decimals) 10.1 = 11)
Trunc	Truncates a decimal fraction to its integer for example: Trunc (18.9) = 18

Chapter: REPRESENTATIVE SAMPLING TECHNIQUES

A representative sampling procedure can be performed on a population of units with sufficient similar external characteristics (e.g., size, colour). The decision on how to perform it is left to the discretion of the examiner. An example about what is meant by similar external characteristics is very important. Considering a group of heroin street doses, which are packed in similar packaging, we can apply a sampling rule to this population. So, if you have 100 street doses with different groups of external characteristics, you have to separate your 100 street doses in as many groups as dissimilarities. Each group will be considered as a whole population and will be sampled alone. In some rare cases, although the external characteristics look the same, when we open the units (sampling), we may notice huge differences in the powder appearance among the units. In this case, you have to stop the sampling procedure according to the above-mentioned criteria. In general it happens when you don't look thoroughly at the external characteristics of the packages.

The theoretical way to select a truly random, unbiased representative sample from a population is to individually number each item in the population and then use a random number generator to choose which item to select. This is not possible in practice, especially for large populations containing many thousands of units.

When sampling, we must ensure that two principles are maintained:

The properties of the sample are a true reflection of the properties of the population from which the samples were taken.

Each unit in the population has an equal chance of being selected.

In reality, it is more difficult to adhere to these principles than it first seems. As was mentioned before, the decision in selecting the samples is left to the discretion of the examiner because, when the population is high, it is impossible to number all the units and use a protocol based on a random selection of numbers. So, considering a subjective choice, it happens that sometimes the expert tends to choose similar sized units, instead of running a real random sampling.

The practical solution is quite easy: after having observed that the external characteristics are the same, you can put all the units in a "black box" (plastic bag or any other idea) and take out your sample without looking. This kind of solution can be applied to practical cases such as seizures of a thousand heroin street doses in similar external packages or a thousand tablets. In this case you can apply this "black box" sampling method to eliminate (or at least reduce to a minimum) any bias that may be introduced by the person selecting the samples. When we refer to a "black box" method we mean any method that will prevent the sampler from consciously selecting a specific item from the population. These methods are not standardized yet and we can refer to the example given above.

Chapter: ARBITRARY SAMPLING

The following are various arbitrary sampling methods. They are often used in practice and work well in many situations. However they have no statistical foundation and may lead to very large samples in case of large seizures. Not all existing sampling procedures are given. Some laboratories use variations of these.

1. All **(** $n = N$ **)**

Advantage(s): 100% certainty about the composition of the population.

Disadvantage(s): Excessive sample sizes for larger populations.

2. $n = 0.05N$ **,** $n = 0.1N$, etc.

Advantage(s): Simple approach.

Disadvantage(s): Excessive sample sizes for larger populations.

3. $n = \sqrt{N}$ **,** $n = 0.5\sqrt{N}$ **,** $n = \sqrt{\frac{N}{2}}$, etc.

Advantage(s): Widely accepted approach.

Disadvantage(s): The number of samples may be too small when the population is small.

4. $n = 20 + 10\%(N - 20)$ *(where* $N > 20$*)*

Advantage(s): Heterogeneous populations likely to be discovered before analysis is complete.

Disadvantage(s): Excessive sample sizes for larger populations.

5. $n = 1$

Advantage(s):	Minimum amount of work.
Disadvantage(s):	Least amount of information on the characteristics of the seizure.

Chapter: STATISTICAL SAMPLING METHODS

INTRODUCTION

The methods discussed in this chapter provide statistically founded ways to determine the sample size. The first two methods concern a frequentist approach, while the third method describes a Bayesian approach.

The assumption behind a frequentist approach is that a fixed but unknown proportion of the seizure contains drugs. The proportion of drugs in a sample can estimate this seizure proportion. The sample proportion will, however, vary over different samples. One sample will give a higher proportion than another sample. Therefore, the frequentist methods provide a confidence, $(1-\alpha)100\%$ (for instance 95% if α is selected to be 0.05), that with a given sample proportion the seizure proportion is at least $k100\%$ (for instance 90% if k is selected to be 0.9). This means that if the sampled proportion is indeed found to be as assumed, one would be correct about a seizure containing at least 90% drugs in 95 out of 100 cases.

The assumption behind a Bayesian approach is that the sample proportion is known and fixed. This proportion is used to calculate probabilities on certain values of the unknown seizure proportion, that at that point is still assumed variable. With this approach it is possible to incorporate some knowledge about the seizure that you may possibly have. The seizure proportion is not known but often some ideas about this proportion exist. For instance, if all plants in a hemp nursery appear similar they probably are all hemp plants. It is also possible that there is no clue about the amount and type of drugs in a seizure. These various forms of prior information will result in different mathematical models to estimate a desired sample size in the Bayesian approach.

THE HYPERGEOMETRIC DISTRIBUTION

Application

The probability that a sample of size n contains exactly X positives (units containing illegal drugs), given that the population of size N contains N_1 positives, can be calculated by

$$P(X = x \mid N_1, N, n) = \frac{\binom{N_1}{x}\binom{N - N_1}{n - x}}{\binom{N}{n}}.$$

This is the hypergeometric distribution. On this distribution the first (and most used) frequentist method is based.

In sampling drug units, the actual number of positives, N_1, and negatives, $N - N_1$ are unknown. To determine these numbers exactly, the whole seizure has to be analyzed. If some uncertainty is allowed, the hypergeometric distribution, combined with hypothesis testing, can be used to calculate the sample size of n units that must be analyzed such that **at least** K ($K = kN$) units, or at least the proportion k from the population of N samples are positive with (1-α)100% confidence. For instance, calculate n such that, with 95% confidence, at least 90% of the units contain illegal drugs. The choice of values for α and k (or K) depend on laboratory guidelines, costs, legal requirements and so on.

If the choices about α and k (or K) are made and if an assumption is made about the number of positives to be expected in the sample (usually n), the sample size n can be solved from the formulas shown in the next section. When sampling strategy (i.e. sample size calculation) is based on a threshold proportion (k) of drug positive items in the population, it is not necessary to do the calculations over and over again as sample sizes for given criteria can be calculated and tabulated. However, when the calculation is based directly on a threshold *number* of positives K, then the calculation has to be performed for each case individually.

Table 1 provides the required sample sizes for some standard choices of α and k at different population sizes, if all sampled units are believed to be positive. Table 2 provides the same information if 1 or 2 of the sampled units are expected to be negative (contain no drugs).

Table 1: Hypergeometric distribution.

Required sample size to guarantee with 95% or 99% confidence that the seizure contains at least a proportion of k drugs, if it is expected that all sampled units contain drugs.

Population size N	95% confidence			99% confidence		
	k=0.5	k=0.75	k=0.9	k=0.5	k=0.75	k=0.9
10	**3**	**6**	**8**	**4**	**7**	**9**
20	**4**	**7**	**12**	**5**	**10**	**15**
30	**4**	**9**	**15**	**6**	**12**	**20**
40	**4**	**9**	**18**	**6**	**12**	**23**
50	**4**	**9**	**19**	**6**	**14**	**26**
60	**4**	**9**	**20**	**6**	**14**	**28**
70	**5**	**10**	**21**	**7**	**14**	**30**
80	**5**	**10**	**22**	**7**	**14**	**31**
90	**5**	**10**	**23**	**7**	**15**	**32**
100	**5**	**10**	**23**	**7**	**15**	**33**
200	**5**	**10**	**26**	**7**	**15**	**38**
300	**5**	**11**	**27**	**7**	**16**	**40**
400	**5**	**11**	**27**	**7**	**16**	**41**
500	**5**	**11**	**28**	**7**	**16**	**41**
600	**5**	**11**	**28**	**7**	**16**	**42**
700	**5**	**11**	**28**	**7**	**16**	**42**
800	**5**	**11**	**28**	**7**	**16**	**42**
900	**5**	**11**	**28**	**7**	**16**	**43**
1,000	**5**	**11**	**28**	**7**	**16**	**43**
5,000	**5**	**11**	**29**	**7**	**16**	**44**
10,000	**5**	**11**	**29**	**7**	**16**	**44**

Example 1

Suppose that a population contains 100 packages. To guarantee with 95% confidence that at least 90% of the packages contains illegal drugs, a sample of 23 packages has to be drawn and all of these packages have to contain illegal drugs (see Table 1).

The assumption that all sampled units contain drugs is often made. This assumption can be made because this is learned from many years of experience in the field, or simply by reasoning that it makes no sense to mix the drugs with no-drugs, apart from maybe a layer of distraction material on top to avoid detection. However, occasionally one or more units in the sample may not contain drugs. In that case, the guaranteed confidence or the minimum proportion drugs in the population drops. Figure 1 shows that the confidence to guarantee a proportion of drugs of at least 90% drops from 95% to 77% if 1 sampled unit did not contain drugs instead of 0 (N=100). Figure 2 shows that the guaranteed proportion at a confidence of 95% drops from 90% to 84 %. (1 negative instead of 0, N=100). Table 2 shows that a sample of 36 was needed to guarantee with 95% confidence that at least 90% of the population contained drugs if one negative in the sample was assumed beforehand.

It is statistically not correct to sample another 13 units on top of 23 if one of these 23 does not contain drugs. Before sampling a decision should be made how many negatives in the sample are expected. Afterwards, when one or more sampled units are found to be negative this has consequences for the confidence and/or the proportion guaranteed. This property makes the sampling with the hypergeometric distribution (and other frequentists methods) hard to understand intuitively.

Table 2: Hypergeometric distribution

Required sample size to guarantee with 95% or 99% confidence that the seizure contains at least a proportion of k drugs, if is expected that either 1 or 2 sampled units do not contain drugs (1 or 2 negatives).

Population size *N*	95% confidence						99% confidence					
	k=0.5		*k*=0.75		*k*=0.9		K=0.5		k=0.75		*k*=0.9	
	1 neg	2 neg	1 neg	2 neg	1 neg	2 neg	1 neg	2 neg	1 neg	2 neg	1 neg	2 neg
10	**5**	**7**	**9**	**10**	**10**	**--**	**6**	**7**	**9**	**10**	**10**	**--**
20	**6**	**8**	**11**	**14**	**17**	**20**	**8**	**10**	**13**	**16**	**19**	**20**
30	**7**	**9**	**13**	**17**	**22**	**27**	**8**	**11**	**16**	**20**	**25**	**29**
40	**7**	**9**	**14**	**18**	**26**	**32**	**9**	**11**	**17**	**21**	**30**	**35**
50	**7**	**10**	**15**	**19**	**29**	**36**	**9**	**12**	**19**	**24**	**34**	**41**
60	**7**	**10**	**15**	**19**	**31**	**39**	**9**	**12**	**19**	**24**	**38**	**45**
70	**7**	**10**	**16**	**20**	**32**	**41**	**10**	**12**	**20**	**25**	**40**	**48**
80	**7**	**10**	**15**	**20**	**34**	**43**	**10**	**12**	**20**	**25**	**42**	**51**
90	**7**	**10**	**16**	**21**	**35**	**45**	**10**	**13**	**21**	**26**	**44**	**54**
100	**7**	**10**	**16**	**21**	**36**	**46**	**10**	**13**	**21**	**26**	**46**	**56**
200	**8**	**10**	**17**	**22**	**40**	**53**	**10**	**13**	**22**	**28**	**54**	**67**
300	**8**	**10**	**17**	**23**	**42**	**55**	**10**	**13**	**23**	**29**	**57**	**71**
400	**8**	**11**	**17**	**23**	**43**	**57**	**10**	**13**	**23**	**30**	**58**	**74**
500	**8**	**11**	**17**	**23**	**44**	**58**	**10**	**14**	**23**	**30**	**59**	**75**
600	**8**	**11**	**17**	**23**	**44**	**58**	**10**	**14**	**24**	**30**	**60**	**76**
700	**8**	**11**	**17**	**23**	**44**	**59**	**11**	**14**	**24**	**30**	**61**	**77**
800	**8**	**11**	**17**	**23**	**44**	**59**	**11**	**14**	**24**	**30**	**61**	**77**
900	**8**	**11**	**17**	**23**	**45**	**59**	**11**	**14**	**24**	**30**	**61**	**78**
1,000	**8**	**11**	**17**	**23**	**45**	**59**	**11**	**14**	**24**	**30**	**62**	**78**
5,000	**8**	**11**	**17**	**23**	**46**	**61**	**11**	**14**	**24**	**31**	**64**	**81**
10,000	**8**	**11**	**17**	**23**	**46**	**61**	**11**	**14**	**24**	**31**	**64**	**81**

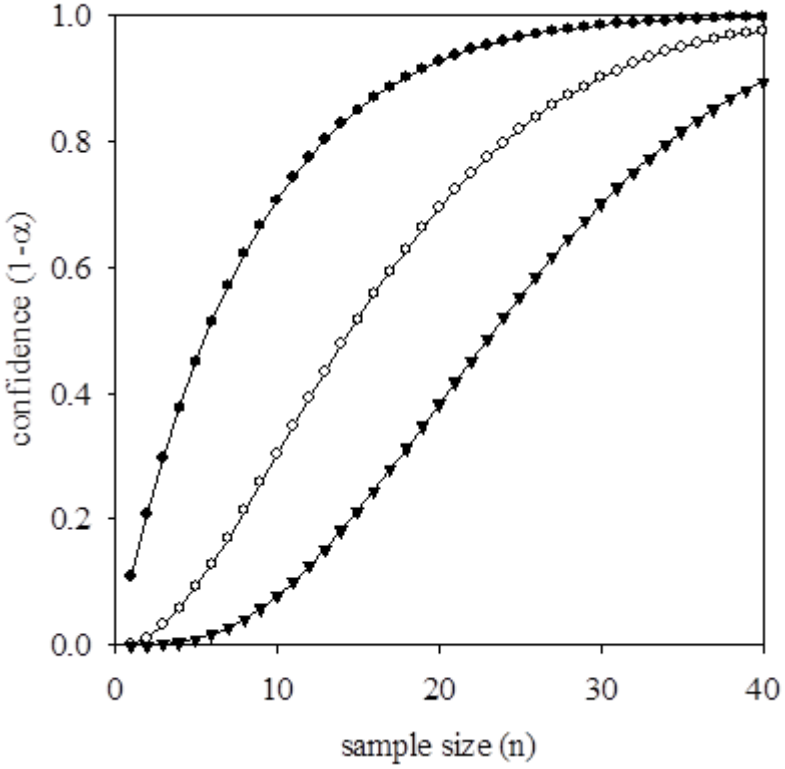

Figure 1: Confidence against sample size (N = 100; k = 0.9) for 0, 1, and 2 negatives expected. Lines -●- for 0 negatives; - □ - for 1 negative; -▼- for 2 negatives

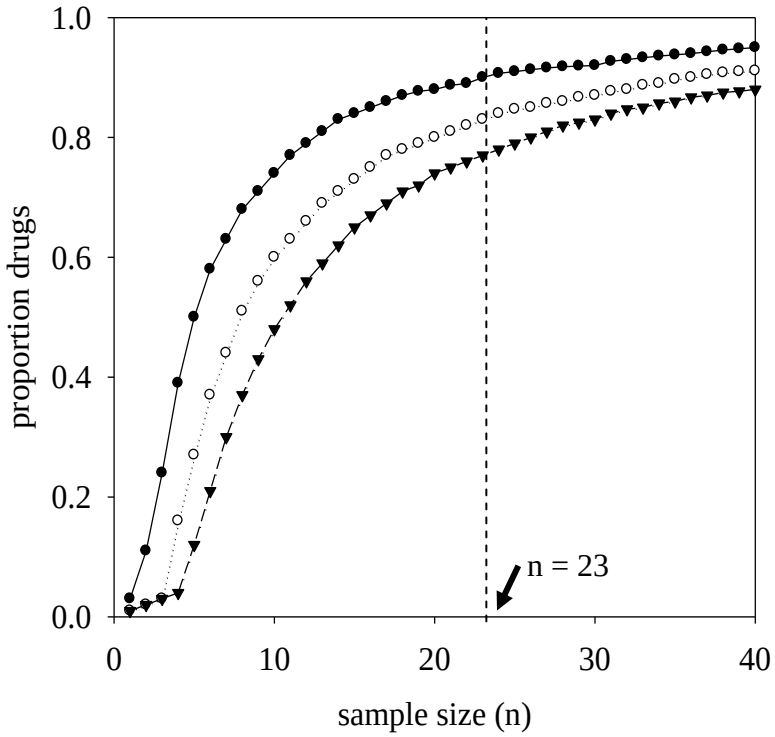

Figure 2: *Proportion against sample size (N = 100; k = 0.95) for 0, 1, and 2 negatives expected. Lines -●- for 0 negatives; - o - for 1 negative; -▼- for 2 negatives*

Example 2

If it is sufficient to guarantee with a high probability (say 95%) that drugs are present in the majority (> 50%) of the exhibit (of 100), then only a sample of 5 is necessary provided that no negative is found (see Table 1).

Example 3

In Table 3 all steps of hypergeometric sample size n calculation for the sampling strategy based on the threshold proportion of positives k are shown for the following threshold parameters chosen by the laboratory: $k \geq 0.90$, confidence level at least 0.95 and number of expected negatives $r = 0$, for the population sizes from N=10 to N=50. For theory see the next section and for in-depth explanation see also ENFSI DWG document
"Validation of the Guidelines on Representative Sampling" DWG-SGL-001 version 002"

Table 3: Hypergeometric distribution

Population size N is shown in the first column. Number of positives corresponding to the proportion k is calculated first (see column 2), afterwards M_0 is calculated (see column 3). In the column 4 calculated sample n is shown, while calculated actual confidence level CL and actual proportion k` guaranteed by calculated sample size n are shown in columns 5 and 6, respectively.

population size N	number of positives (calculated) *K=k*N*	number positives for H_0 test M_0= *RoundUp(k*N)-1*	calculated sample size n	Actual CL (1-α)	actual proportion k` = RoundUp(k*N)/N
10	9	8.00	8	0.9778	0.9000
11	9.90	9.00	9	0.9818	0.9091
12	10.80	10.00	9	0.9545	0.9167
13	11.70	11.00	10	0.9615	0.9231
14	12.60	12.00	11	0.9670	0.9286
15	13.50	13.00	12	0.9714	0.9333
16	14.40	14.00	12	0.9500	0.9375
17	15.30	15.00	13	0.9559	0.9412
18	16.20	16.00	14	0.9608	0.9444
19	17.10	17.00	15	0.9649	0.9474
20	18	17.00	12	0.9509	0.9000
21	18.90	18.00	13	0.9579	0.9048
22	19.80	19.00	14	0.9636	0.9091
23	20.70	20.00	14	0.9526	0.9130
24	21.60	21.00	15	0.9585	0.9167
25	22.50	22.00	16	0.9635	0.9200
26	23.40	23.00	16	0.9538	0.9231
27	24.30	24.00	17	0.9590	0.9259
28	25.20	25.00	18	0.9634	0.9286
29	26.10	26.00	18	0.9548	0.9310
30	27	26.00	15	0.9502	0.9000
31	27.90	27.00	16	0.9566	0.9032
32	28.80	28.00	17	0.9620	0.9062
33	29.70	29.00	17	0.9555	0.9091
34	30.60	30.00	18	0.9608	0.9118
35	31.50	31.00	18	0.9545	0.9143
36	32.40	32.00	19	0.9596	0.9167
37	33.30	33.00	19	0.9537	0.9189
38	34.20	34.00	20	0.9585	0.9211
39	35.10	35.00	20	0.9529	0.9231
40	36	35.00	18	0.9600	0.9000
41	36.90	36.00	18	0.9551	0.9024
42	37.80	37.00	18	0.9500	0.9048
43	38.70	38.00	19	0.9558	0.9070
44	39.60	39.00	19	0.9511	0.9091
45	40.50	40.00	20	0.9565	0.9111
46	41.40	41.00	20	0.9520	0.9130
47	42.30	42.00	21	0.9571	0.9149
48	43.20	43.00	21	0.9528	0.9167
49	44.10	44.00	22	0.9577	0.9184
50	45	44.00	19	0.9537	0.9000

Theory

This section is for those who want more background information on the hypergeometric distribution and the calculation of the table values.

The hypergeometric distribution, and thus the theory below, assumes that samples are taken without replacement. The hypergeometric distribution is discrete and all numbers applied in calculations must be non-negative integers. The theory presented here is fit for situations where sampling size calculation is based on either threshold number (*K*) or threshold proportion (*k*) of expected drugs positive items in the population as specified by the laboratory.

If the choices about α and *K (or k*) are made and if an assumption is made about the number of positives to be expected in the sample, the minimum a sample size (*n*) is calculated by testing of the null hypothesis[1] that **number of positives** in the population is less than *K* against the alternative hypothesis that the number of positives is at least *K:*

H_0: $N_1 < K$ against $H_1 : N_1 \geq K$

To prosecute people for all the seized units it is desired that $N_i \geq K$. Evidence has to be found to reject the null-hypothesis. However, no big mistakes are allowed. This means that the probability that the null-hypothesis is rejected, while it is true, should be small, say α. This provides a confidence level of $(1-\alpha)100\%$. The hypotheses are tested with the **number of positives** in the sample, *X*, as the test statistic. The null-hypothesis is rejected when *X* is larger than a certain number. If this number is taken as the number of positives expected in the sample, *x*, then, *n* should be selected such that:

$$P(X \geq x|N, N_1 < K) \leq \alpha.$$

[1] To understand the background of hypothesis testing one should note that: Guaranteeing with (1-α)100% confidence that at least proportion *k* ×100% (or particular number *K*) of population are drugs is the same as guaranteeing that the probability on finding only (or mostly) drugs in the sample will be less than α when the proportion of drugs in the population is less than *k* (or number of drug items less than K).

Equation 1

Intuitively, P(Reject H_0 | N_1) increases as the number of positives N_1 in the population increases. (H_0 is N_1 < K.) Therefore, to find the smallest sample size (n) which guarantees at least proportion of positives *k* (or number of positives *K*) the null hypothesis (H_0) is tested at the highest possible $\boldsymbol{M_0}$ (integer) that is lower than *K*. If H_0 is rejected (H_1 is accepted), then the calculated sample size n will give the smallest number of samples for analyses, which guarantees at least ***k*** proportion (or corresponding K) of positives in the population, at a confidence level at least equal or greater than (1- α). Equation 1 may be rewritten as:

$$P(X \geq x|\mathrm{N}, M_0 < K) \leq \alpha.$$

So given that $M_0 < K$ the minimal sample size (*n*) is the smallest value of *x* for which $P(X \geq x|M_0 < K) \leq \alpha$. The smallest sample size *n* is actually calculated with the consecutive use of Equation 2 or Equation 3 or Equation 4 below (dependent on the number of negatives at most allowed by increasing n, until the cumulative hypergeometric probability ≤ α or at maximum to n = N.

When all sampled drug units are expected to contain drugs (i.e. $x = n$ which is equivalent to zero negatives expected ($r = 0$)), X follows a hypergeometric distribution:

$X \sim \mathrm{HYP}(n, M_0, N)$.

Formally, interest lies in the tail area probability $P(X \geq x \mid M_0 < K) \leq \alpha$. However, it is not possible for *X* to be greater than *n* so $P(X \geq x \mid M_0 < K, x = n)$ can be written as $P(X = n \mid M_0 < K)$ resulting in:

$$P(X \geq x / M_0 < K, x = n) = P_0 = P_{r=0} = \frac{\binom{M_0}{x}\binom{N - M_0}{n - x}}{\binom{N}{n}} = \frac{\binom{M_0}{n}\binom{N - M_0}{0}}{\binom{N}{n}} = \frac{\binom{M_0}{n}}{\binom{N}{n}}.$$

Equation 2

When at most one sampled drug unit is expected not to contain drugs (i.e. $x \geq n$-1 which means that x = n-1 or x = n are possible; hence, the number of negatives *r* may be at most 1, i.e.: r = 0 or r = 1 are possible), X is distributed as a mixture of two hypergeometric random variables:

$$P(X \geq x / M_0 < K, x \geq n-1) = P_1 = P_{r=0} + P_{r=1} = P_0 + \frac{\binom{M_0}{n-1}\binom{N-M_0}{1}}{\binom{N}{n}} \quad .$$

Equation 3

When at most two sampled drug units are expected not to contain drugs (i.e. $x \geq n$-2 which means the number of negatives can be at most two, i.e.: r = 0 or r = 1 or r = 2 are possible), X is distributed as a mixture of three hypergeometric random variables:

$$P(X \geq x / M_0 < K, x \geq n-2) = P_2 = P_{r=0} + P_{r=1} + P_{r=2} = P_1 + \frac{\binom{M_0}{n-2}\binom{N-M_0}{2}}{\binom{N}{n}} \quad ,$$

Equation 4

and so on for higher number of negatives at most allowed.

How to find M_0 (the highest integer lower than K for H_0 test) in practice? Two different situations can appear:

a) When the threshold **number** of expected positives K (always an integer), confidence level CL and number of negatives *r* at most allowed are specified by the laboratory, the M_0 is simply:

$$M_0 = K\text{-}1.$$

b) When the threshold proportion *k* of expected positives, confidence level CL and number of negatives at most allowed are specified by the laboratory, the threshold proportion *k* shall first be converted to number of expected positives $K = k \times N$ (note in the table 3, that

the calculated *K*s can be integers or non-integers). Afterwards, M_0 can be calculated along general formula:

$$M_0 = RoundUp(K)-1,$$

which is actually valid for integer and non-integer *K and is especially useful when calculations performed by the computer software. For the calculations by hand one can use*

$$M_0 = K-1, \text{ for integer } Ks$$

and

$$M0 = Trunc(K) = RoundUp(K)-1, \text{ for non integer } Ks.$$

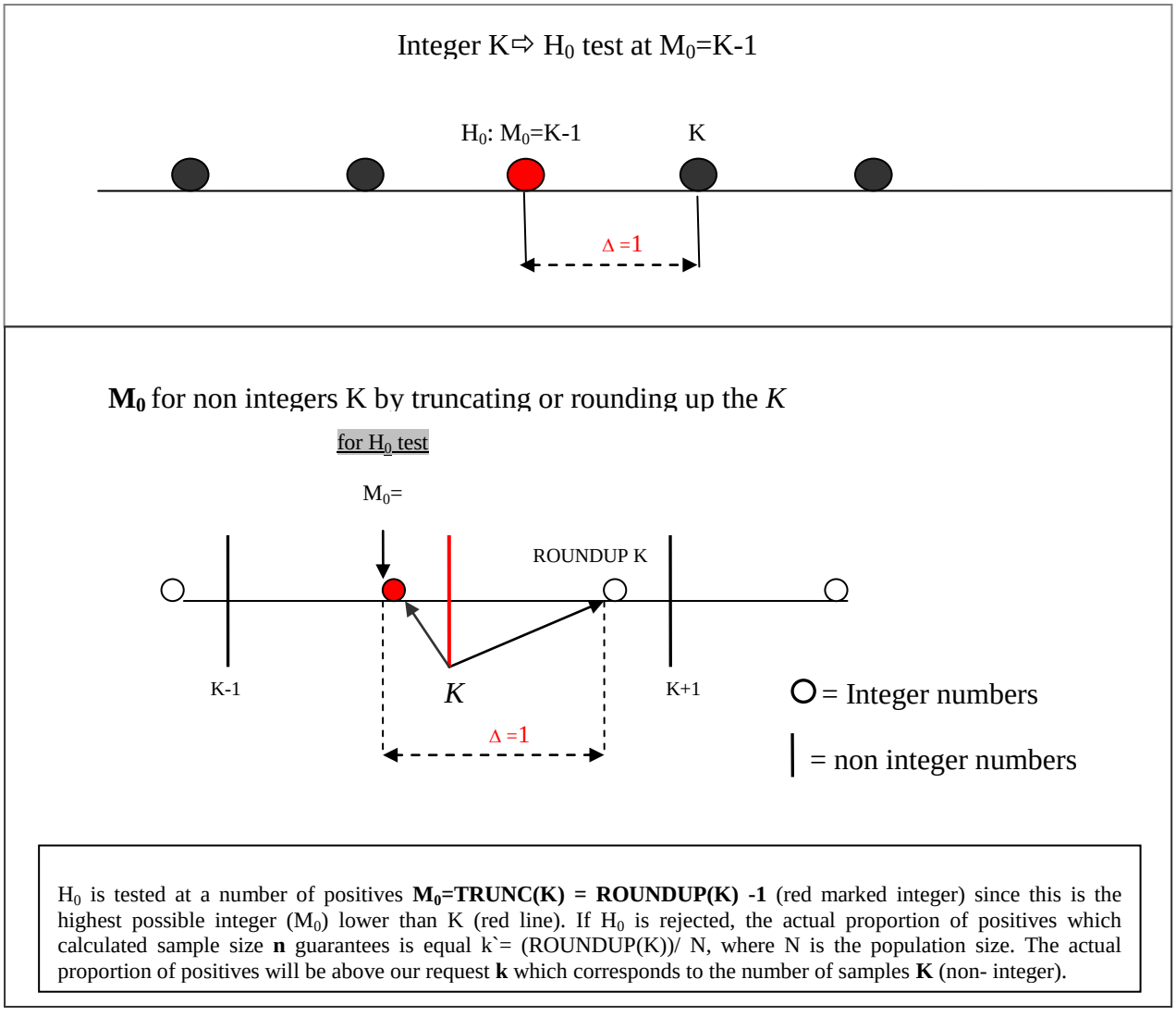

Figure 3: Calculation of M_0 from integer and non-integer *K*`s.

How to calculate the actual proportion of positives guaranteed by calculated sample size n?

a) For integers Ks actual proportion k` match laboratory request (k) exactly (see figure below) and is expressed as:

$$k` = k = K/N = \text{RoundUp}(K)/N.$$

b) For non-integer Ks actual proportion k` is always slightly above the laboratory request (see figure below) and is expressed as:

$$k` = \text{RoundUp}(K)/N.$$

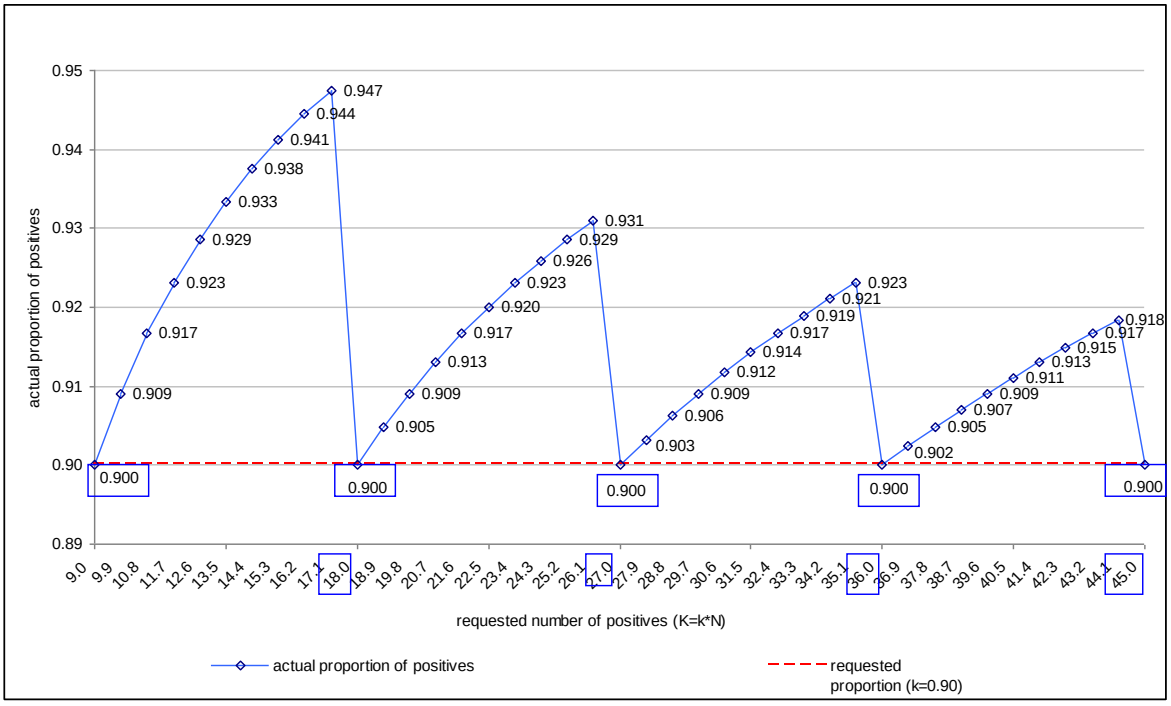

Figure 4: Actual proportion of positives k` for integers and non-integers K (data: k=0.90, CL=0.95, population sizes from N=10 to N=50, see data in Table 3). When *K* is an integer (note the numbers in blue rectangles) the actual and the threshold proportion *k* defined by the laboratory match exactly. For non-integers *Ks* the actual proportion *k`* is higher than the proportion *k* requested by the laboratory. Note that for the given example the *k* = 0.90 (red line) was specified by the laboratory.

THE BINOMIAL DISTRIBUTION

Application

This is the second method using a frequentist approach. It is an easier method, but can only be used in special cases. The binomial distribution assumes sampling with replacement. This means that a unit is placed back after it is sampled and analyzed before the next unit is sampled. Of course this is not practiced in drugs sampling. However, in situations where the seizure is very large (at least 50, preferable larger) and the sample is relatively small the hypergeometric distribution can be approximated by the simpler binomial distribution. In that case, the probability that a sample of size n contains X positives (units containing illegal drugs), given that the population of size N contains a proportion of $\theta = \frac{N_1}{N}$ positives, is

$$P(X = x \mid \theta, n) = \binom{n}{x} \theta^x (1-\theta)^{n-x}.$$

Similarly, as with the hypergeometric distribution, the binomial distribution can be used to calculate a sample size n such that with $(1-\alpha)100\%$ confidence can be stated that at least a proportion of $k100\%$ is positive. The calculations with the binomial distribution are easier than the ones with the hypergeometric distribution. However, it should be kept in mind that the binomial distribution is an approximation. The sample size estimated with it will be slightly overestimated. Only in very large seizures (sometimes of several thousands) the sample sizes calculated from both distributions will be exactly equal.

If no negatives are expected the sample size n, that with $(1-\alpha)100\%$ confidence can be stated that at least a proportion of $k100\%$ is positive, can be calculated by the minimum value for which

$$n \leq \frac{\log \alpha}{\log \theta},$$

regardless of the population size. If negatives are found in the sample conclusions have to be adapted in a similar way as with the hypergeometric distribution. Again tables (see **Table 4**) or software can be used.

Table 4: Binomial distribution.

Required sample size to guarantee with 95% or 99% confidence that the seizure contains at least a proportion of k drugs if is expected that 0, 1 or 2 sampled units do not contain drugs (0, 1 or 2 negatives). Use this only for large seizures.

Population size N	95% confidence			99% confidence		
	k=0.5	*k*=0.75	*k*=0.9	*k*=0.5	*k*=0.75	*k*=0.9
0 negatives	5	11	29	7	17	44
1 negative	8	18	46	11	24	64
2 negatives	11	23	61	14	31	81

Example 1

To guarantee with 95 % confidence that at least 90% of the pills contain drugs a sample of 29 should be drawn (if no negatives in the sample are assumed). Compare this with the hypergeometric distribution when a sample has to be drawn from a population of 100. Then the sample size is only 23. Only when the population is as large as 1600 the results from the binomial distribution coincides with that of the hypergeometric distribution for this $(1-\alpha)100\%$ and k.

Example 2

A large seizure is found. Experienced police people can see that this is most probably all heroin. Even if only half of it is heroin this will be a large seizure. Therefore a sample that guarantees with 95% confidence that at least 50% of the seizure is drugs is sufficient. **Table 4** shows that in that case the sample size will be 5, if no negatives are assumed.

Theory

The theory behind the binomial distribution is similar to that of the hypergeometric distribution. The hypotheses are

$$H_0 : \theta < k$$
$$H_1 : \theta \geq k$$.

To select n, the equation to be solved is

$$P(X \geq x \mid \theta = k) = \sum_{i=x}^{n} \binom{n}{x} \theta^x (1-\theta)^{n-x} \leq \alpha .$$

Thus in case that $x = n$, the equation to be solved is

$$\theta^n \leq \alpha .$$

That is, find the minimum value for which

$$n \leq \frac{\log \alpha}{\log \theta}.$$

The binomial distribution is an approximation of the hypergeometric distribution. The value for n found with the binomial distribution will always be equal to or greater than the value found with the hypergeometric distribution.

BAYESIAN APPROACH

Application

Within the Bayesian approach (like the frequentist approach) a distinction can be made between sampling with replacement and sampling without replacement. Again sampling with replacement is simpler and can be used as an approximation for situations where the population size is at least 50 and the sample relatively small. Here overestimation is not such a problem as with the binomial distribution. That is why the sampling with replacement approximation is much more used in the Bayesian approach.

Bayesians assume that, although the population proportion is not known, there may be some ideas about the size of this proportion. These ideas are represented by a probability distribution $p(\theta)$, the so-called prior distribution of the proportion. This uncertain knowledge is combined with the information provided by the sample to a so-called

posterior distribution of the proportions, given the sample results. With this posterior distribution it is possible to calculate directly the probability that the proportion of drugs is at least *k* (given the sample results) without using tests or confidence intervals. This is because Bayesians calculate $P(\theta > k \mid x, n)$ directly instead of $P(X > x \mid \theta > k, n)$ as the frequentists do.

Seizure containing more than 50 units

If a population is large ($N > 50$) and the sample is relatively small compared to the population, the probability density function for the proportion θ of positives, given that a sample of size *n* contains *x* positives is

$$f(\theta \mid x, n, a, b) = Be(x + a, n - x + b) = \frac{\theta^{x+a-1}(1-\theta)^{n-x+b-1}}{B(x + a, n - x + b)}.$$

This is the beta distribution with parameters $x + a$ and $n - x + b$. The parameters *a* and *b* have to be selected beforehand based on prior knowledge or assumptions about θ. The prior knowledge together with the information about the data (the sample size *n* and number of positives in the sample *x*) form the above presented posterior distribution. *Be* stands for the beta distribution and *B* stands for the beta function. For more details see the theory section.

The probability that the population proportion is larger than *k* can be calculated with $P(\theta > k \mid x, n)$. This can be used to select a sample size *n* such that the probability is $(1-\alpha)$ that $\theta > k$. For instance, select *n* such that the probability is 95% that at least 90% of the pills contain illegal drugs. The calculations are independent of the population size.

Calculations on the beta distribution to find such an *n* can best be carried out with the aid of a computer. **Table 5** is based on computer calculations. Like in the frequentist methods you have to assume beforehand what the number of positives in your sample will be, and adapt your conclusions if afterwards this number is not correct. Again in most cases no negatives will be expected.

Besides the expected number of positives in the sample, a prior distribution has to be selected. In general this is a beta distribution One suggestion is to take both parameters *a* and *b* equal to 1, if there is no prior idea about the contents of the pills. The prior distribution then equals the uniform distribution. Another suggestion is to take them both equal to ½ if there is a prior that either all pills contain drugs or no pills at all contain drugs. Take $b = 1$, and $a = 3$ (or even higher) if there is a prior belief, based on visual inspection and experience or so, that probably all is drugs. For instance, 100 packages of white powder are found, all similarly packed, all having the same weight, and all smelling of illicit cocaine. Sampling a hemp nursery may be an even more extreme case.

Table 5: Beta distribution (with parameters $x + a$ and $n - x + b$).

Required sample size to guarantee with a probability of 95% or 99% that the seizure contains at least a proportion of k drugs if expected that 0, 1, or 2 sampled units do not contain drugs (0, 1 or 2 negatives). A large seizure is assumed (N ≥ 50). Use (a=1, b=1) if no prior information is known, (a=0.5, b=0.5) if it is reasonable to assume that either everything is drugs or nothing is drugs, (a=3, b=1, or more extreme values) is there are reasons to believe that all or most of the seizure contains drugs.

a = 1 **b = 1**	95% confidence			99% confidence		
	k=0.5	*K*=0.75	*K*=0.9	*k*=0.5	*k*=0.75	*k*=0.9
0 negatives	4	10	28	6	16	43
1 negative	7	17	47	10	24	64
2 negatives	10	23	63	13	31	83

a = 3 **b = 1**	95% confidence			99% confidence		
	k=0.5	*K*=0.75	*K*=0.9	*k*=0.5	*k*=0.75	*k*=0.9
0 negatives	2	8	26	4	14	41
1 negative	5	15	45	8	22	62
2 negatives	8	21	61	11	29	81

a = 0.5 **b = 0.5**	95% confidence			99% confidence		
	k=0.5	*K*=0.75	*K*=0.9	*k*=0.5	*k*=0.75	*k*=0.9
0 negatives	3	7	18	5	12	32
1 negative	7	15	39	9	21	55
2 negatives	10	21	56	12	28	75

Example 1

To be sure, without any prior knowledge (see **Table 5** with a=1, b=1, 0 negatives), with 95% probability that at least 90% of all pills contain illegal drugs, a sample of size 28 is needed within the Bayesian approach. This is higher than when the hypergeometric distribution is used, because then only 23 (see **Table 1**) samples are needed. However, if it is very clear that we are dealing with drugs, and we combine this with the practical knowledge that then probably all are drugs the sample size drops to 26 (a = 3, b =1) or even 19 (a = 10, b = 1).

Example 2

To guarantee with a probability of 95% that at least half of the seizure contains drugs, only a sample size of 4 is needed (when no negatives are expected in the sample). In very extreme cases this number can be reduced or increased by one or two. In general, to guarantee at least 50% of drugs (with a probability of 95%) a sample size of 4 is an easy guideline.

Seizure containing less than 50 units

If the consignment is small (N <50), it is better to consider the number of positives in the unexamined units instead of the proportion of positives. The probability density function for the number of positive in the unexamined units Y, given that a sample of size n contains x positives is

$$f(Y \mid x, n, (N-n), a, b) = \frac{\Gamma(n+a+b)\binom{N-n}{y}\Gamma(y+x+a)\Gamma(N-x-y+b)}{\Gamma(x+a)\Gamma(n-x+b)\Gamma(N+a+b)}.$$

This is the beta-binomial distribution.

The probability that the number of positives in the unexamined pills is larger than y can be calculated with $P(Y \geq y \mid x, n, N)$. This can be used to select a sample size n such that the probability is $(1-\alpha)$ that $Y > y$. Calculations on the beta-binomial distribution to find such

an *n* have to be done with the computer (statistical software, or Excel for example), or at least a scientific calculator. Like in the frequentist methods you have to assume beforehand what the number of positives in your sample will be, and adapt your conclusions if afterwards this number is not correct. Again in most cases no negatives will be expected.
In contrast to the Bayesian method for large consignments the calculated sample size depends on the consignment size. Furthermore, calculations on the proportion cannot be very precise, because of the small numbers. Therefore it is probably best to use the hypergeometric distribution for small consignments or use the calculated sample sizes calculated with the Bayesian method for large consignments as approximation for small consignments.

Theory

This section is for those who want to know where the numbers in the tables come from.

The Bayes approach allows the use of prior information about a parameter (such as the proportion drugs in a seizure); by combining this prior information with the results from the sampling, it leads to a posterior information about that parameter. Let θ be the parameter of interest and *x* the data from the sample; the Bayes theorem is then:

$$P(\theta \mid x) = \frac{P(x \mid \theta)\, p(\theta)}{P(x)}.$$

This is often rewritten as Bayes formula

$$P(\theta \mid x) \propto L(\theta \mid x)\, p(\theta),$$

Where $L(\theta \mid x)$ is the likelihood function? This function contains information about the data. Formally it is of the same form as a corresponding probability mass function for discrete data or probability density function for continuous data. However, the likelihood function is a function of θ rather than *x*. For example, for a binomial distribution,

$$\Pr(X = x \mid n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}; x = 0,1,...n$$

whereas

$$L(\theta \mid, n, x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}; 0 < \theta < 1$$

Here, $p(\theta)$ is the prior distribution, representing the uncertainty about the knowledge of θ. If no knowledge or ideas exist about θ, any value (between 0 and 1, if θ is a proportion) is as likely as any other. Then $p(\theta)$ is a uniform distribution. This is a special case of the beta distribution. In general, a beta distribution with parameters *a* and *b* is assumed.

The beta distribution *Be*(*a*,*b*) is given by

$$f(\theta \mid a,b) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)},$$

with the beta function $B(a,b) = \int_0^1 y^{a-1}(1-y)^{b-1}dy$. This can also be written as $B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$, where we have used the gamma function Γ given by

$$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx,$$

and for integer $n > 0$, $\Gamma(n) = (n-1)!$ And $\Gamma(1/2) = \sqrt{(\pi)}$.

In case of no prior belief about the seizure *a* and *b* both equal 1 (the uniform distribution). In case more information is available, for instance, all units of the seizure show the same (visual) characteristics, other values of *a* and *b* have to be used. If all pills look similar it is most likely that all pills contain drugs or no pills at all contain drugs, then $a = 1/2$ and $b = 1/2$. If there is a founded suspicion that drugs is involved, so that it is very likely that θ is high, *a* could be 3 and $b = 1$, or even stronger: $a = 10$, and $b = 1$.

The likelihood function combines with the prior information to the posterior distribution of the proportion θ given the data

$$f(\theta \mid x,n,a,b) = Be(x+a, n-x+b) = \frac{\theta^{x+a-1}(1-\theta)^{n-x+b-1}}{B(x+a,n-x+b)}.$$

If all sampled pills contain drugs ($x = n$) this is

$$f(\theta \mid n,n,a,b) = Be(n+a,b) = \frac{\theta^{n+a-1}(1-\theta)^{b-1}}{B(n+a,b)}.$$

To calculate the sample size *n* such that with a probability of (1-α)% at least k% of all pills contains drugs, the equation

$$P(\theta > k \mid n, n, a, b) = \int_k^1 \theta^{n+a-1}(1-\theta)^{b-1} d\theta / B(n+a, b) = (1-\alpha)\%,$$

has to be solved.

The same Bayesian theory concerning Bayes theorem is true for the case of small consignments. Then the distribution of $P(Y \mid N-n, \theta)$ is binomial. When this is combined with the prior beta distribution for θ the resulting posterior distribution of $P(Y \mid n, N-n, \theta, a, b)$ is beta-binomial

Chapter: CONSIDERATIONS

In the previous chapters a number of sampling strategies were (briefly) described. Although advantages and disadvantages of certain methods were given, no real preference was mentioned. This chapter attempts to bring up a number of considerations about the use of (one of) the methods, and to mention and discuss a number of related aspects, with the aim to support laboratories in the selection of their recommended method.

THE BASIS OF SAMPLING

The basis of sampling is that the composition found in the samples taken reflects, in principle, the composition of the whole lot.
As a consequence, only a fraction of the total packages in a seizure can be investigated. Sampling is an intentional choice to refrain from doing things to (unnecessary or impossible) perfection, for reasons of efficiency and cost effectiveness.

As an example, if one item is taken at random out of a population of 10 items in which 10% (one and only one) contains cocaine, then the probability the item taken contains cocaine is 0.1 (10%). In contrast, if one item is taken at random out of a population of 10 items in which 50% (five and only five) contain cocaine, and then the probability the item taken contains cocaine is 0.5 (50%).

THE AIM OF SAMPLING

Actually, a sampling strategy is fully dependent on the question and thus the problem that has to be solved. There may be different needs for possession, production, or trafficking. The question usually arises from the national law, or from a national policy (habit) or sometimes directly from the prosecutor's opinion or from the police staff. Simplified, in a sequence of increasing workload:

(i) Is a drug present? Minimal sampling (this may require 1 positive result).
(ii) Is a drug present in (more than) a specified proportion of the items? Increased sampling.
(iii) Is a drug present in *all* the items? Maximum sampling (this may require full analysis of all items, which will lead to unrealistic costs, especially for large numbers of units).

It is clear that, for large seizures, situation (ii) is widely considered as a reasonable approach, often allowing a scientist to include a statistical approach. In this case, we can choose the desired confidence level. An increase in confidence from 95% to 99% will result in an increase of the number of samples to be taken; depending on the conditions, it could mean more than a doubling. In statistics 95% is very common and widely accepted; for this reason we advise to adopt this 95% confidence as the standard.

Table 6: Hypergeometric distribution

Number of samples to be taken for describing (with 95% confidence) a certain proportion of drugs in a seizure, assuming 0 negatives in the sample.

Proportion of seizures at least positive on drugs	For a seizure consisting of 100 units	For a seizure consisting of 1000 units
50%	5	5
60%	6	6
70%	8	9
80%	12	14
90%	23	28
95%	39	56

The higher the proportion, the larger the increase in sample size needs to be to increase the minimum proportion with a fixed percentage. Equilibrium has to be found between the costs of exponential increasing sample sizes and the increase in the guaranteed drugs proportion gained from this.

Although many different methods are in use, the hypergeometric approach seems to be the most widely accepted one; it has been well described and is recommended by the UNDCP and SWGDRUG. This does not mean that this approach should automatically be adopted by ENFSI. In the first place, it is heavily influenced by practices in the USA, whereas the aim of the ENFSI manual is to look for methods that are 'fit for use' in Europe. Secondly, a number of laboratories choose the Bayesian approach because this method allows the use of other relevant, so-called prior information (e.g. external characteristics).

The main problem with the hypergeometric is that it is blind. It does not take into account additional aspects. Visual inspection, smelling etc. can contribute to the investigation of the seizure, but there is no way to incorporate this in the hypergeometric approach. This problem can be best demonstrated with an example. When investigating a hemp field of 1000 plants, hypergeometric tables show a number of 28 samples to be taken. That seems a bit much, especially for an expert who has been working with hemp for years, he smells it, notices the lamps, the nutrition, the books about hemp nursery and so on. And the suspect admits that he is breeding hemp. And if another room is found with again 1000 the same and identical plants? And with 5 of these rooms? The sampling will have to be done, considering a population, respectively, of 2,000 and 5,000 plants.

Actually, the hypergeometric model gives the absolute minimum proportion of drugs to be present in cases where no other information is available at all. More abstractly formulated: in cases where more information is contributing, the strict use of the hypergeometric approach leads to a too an unrealistic high number of samples.

The Bayesian approach can incorporate above-mentioned additional information in its model, by the use of a prior distribution. In general the prior distribution is a beta distribution with parameters ‘a’ and ‘b’. The more additional information, in the sense that is clear that we are dealing with drugs and that all units contain drugs, the higher the parameter ‘a’ should be chosen. When the plants can be visually identified as hemp, it can be seen that the plants are all the same, and the suggestion that all the plants are of something other than hemp is really unrealistic, a very high value for ‘a’ may be selected (e.g. 40). Then the number of samples to be taken will be 1 indeed. The choice of the exact value of ‘a’, however, may be an argument for discussion since there is no standard rule available.

A similar but less evident situation is in the case of a body packer seized at the airport, coming from a South-American destination, with 80 plastic and rubber wrapped packages. Upon collection they all seem to be similar. Opening of two of them shows a white powder. Both are sent for laboratory investigation. The difference with the hemp field is a lower information value of the powder, the similarity lies in the conditions and situations. Within the framework of the Bayesian approach, a prior distribution with a high value for ‘a’, but much lower than in the previous case, can be chosen. The hypergeometric

distribution can be used in court in a case like that of the body packer. The defense may argue that maybe the 78 other packages that were not measured do not contain drugs. However, the probability that only the two measured packages contain drugs is

$$\frac{\binom{2}{2}\binom{78}{0}}{\binom{80}{2}} = 0.000316$$

,

about 3 in 10000. This is a very small probability. If the fact that all packages of all body packers measured always contained drugs is incorporated and the Bayesian approach is used, this probability will be even much smaller.

In general, it can be stated that Bayesian methods should be preferred when much prior information is available, even though one can argue that they imply subjective prior beliefs. In situations where one wants to be completely free of subjective hypotheses or where there is hardly any prior information available, frequentist methods (hypergeometric and binomial) seem attractive because they are easier to understand and to explain. However, they always provide sample sizes on the safe side. This has the advantage that the defence team in court can hardly object against it, but the disadvantage of often too many samples analysed, as shown the above two examples (hemp and body packer).

The hypergeometric distribution is especially valuable for small seizures (N<50) because then other methods easily overestimate sample sizes. This includes the binomial distribution, which is not commonly used.

When the majority (at least 50%) of all units should be guaranteed to contain drugs the results of the hypergeometric distribution and the Bayesian method do not differ that much. Only in very extreme cases (like with the hemp plants) the Bayesian method provides lower sample sizes. In most other cases the sample size will be around 5.

The importance of experience in a profession is generally recognized; this expertise cannot be linked to the hypergeometric distribution. So, already Sutherland in 1990 mentioned that in cases with large numbers of packages, containing similar material upon visual inspection, they always all appeared to contain the same drug. (Note: This consideration is in qualitative analysis only!). In import/export cases, by its nature the seizure is logically

composed of drugs; experience in The Netherlands shows that mixtures with non-drugs were extremely rare; as an indication, in many thousands of cases only one case was found where some negative samples were present. This experience can be linked to the Bayesian approach; however, there are no standard rules for it (yet).

The sampling of tablets may give some specific complications. What is a realistic sampling of 2000 tablets, all in one bag, all with the same external characteristics including all the same logo? Again the hypergeometric approach would lead to 29 samples (for 90% proportion and 95% probability). Intuitively, this is a large number, and intuitively it is very unlikely that negative samples will be present in the whole lot.

A question to be considered is the previous situation, but now the 2000 similar tablets not in one bag, but in 4 bags with each 500 tablets. Does this mean 4 times 29 analytical samples need to be selected, giving a total of 116 separate analyses? From a purely statistical standpoint, probably yes. From a practical standpoint, probably not. From the standpoint of cost effectiveness, also, probably not. The statistically correct approach would be to combine the 4 packages (only allowed with similar material) and then sample accordingly; this approach has also disadvantages.

In addition to the collection of (numerous) samples it has been discussed how to treat these many samples in the laboratory. In some laboratories is common practice to do a spot test on all, maybe then TLC on all or on a large selection, and then –when no differences have been found - end with a very selective analytical technique on only a small number of samples. SWGDRUG however, recommends the full analysis of all analytical samples selected for analysis, if statistical conclusions must be drawn (rather than, say, combining them together for a single analysis). In addition, another strategy is mentioned where selected analytical samples are all tested with a screening technique; followed by a full analysis of one individual and a mixture of all others. The very first strategy maybe preferred, but so far, a solid statistical basis has not yet been presented. However, it can be expected that the approach fits with the Bayesian approach. If so, much laboratory work can be avoided.

‘Bulking‘ of samples may be described as the preparation of one mixture, composed of a number of samples. If bulking can be arranged in such a way that the composition of the

mixture reflects the total composition, then it seems to be a very effective strategy to reduce workloads. Such a mixture may be easy to prepare. A disadvantage will appear in relatively inhomogeneous lots; by definition bulking shows the average and no information about the specific item (although some improvement in this aspect could be obtained by a prior investigation with spot tests).

Sampling strategies must be relatively easy in order to be practical. From a table, a number of samples must be read, and more samples must be taken when it is expected that one or more will not contain drugs. The basis of such an expectation is unclear. So, it would probably mean that a first sample set is collected, analysed, and that, if negative samples have been found, a resampling will be done. That seems rather complicated and even impossible if the seizure is destroyed immediately after the sampling. And always using a standard sample strategy as if 2 negatives are expected leads to an increase in the number of samples; this may look a bit exaggerated when in almost all cases no negatives are found.
Especially when police or customs are doing the sampling, they should be guided by easy-to-understand instructions. In that context, tables or computer programs are less attractive. Some colleagues have solved the problem by the instruction to always take a fixed number of samples, (e.g. 25).

ENFSI ADVICE ON SAMPLING

On the national / regional / laboratory level

Sampling is a strategy; the intensity of sampling is highly dependent on the purpose of the sampling, the question, and the aims. National laws and legal practices will dictate most of them. In practice, there is often some freedom, which means that regional police forces, courts and the laboratories may have the possibility to develop their own strategy. This should be appropriate for their needs, satisfactory for the customer, easy to understand, sensitive for the laboratories' workload, and be cost effective. Further, experience with the local drugs market can be included. For this 'regional' or national level, a general rule will seldom yield the perfect solution, in other words, it results by definition in too few or too many samples; too few samples being insufficient, and too many samples have the disadvantage of waste of time and money. A general ENFSI advice cannot compete with the fine-tuning that can be obtained on the national or regional level.

Regarding this, the ENFSI sub-committee on sampling and the ENFSI steering committee has decided that no specific sampling procedure will be recommended. It is left to the decision of the specific chemists, who will, together with the laboratory management, choose and develop an appropriate strategy that will be satisfactory for and agreed with their customers (police, courts). It is, however, strongly recommended to document the strategy and, if appropriate, to give written instructions for use by the police and/or customs.

On the international level

ENFSI has also to consider sampling of large seizures with clearly international aspects, which means that suspects will be found in one or more countries. It was felt as necessary to have a reasonable strategy that will meet broad support by (most) forensic chemists in EU countries and that can be used as a guideline for police and custom officers.

Also here the starting point is that sampling is a strategy; the intensity of sampling is highly dependent on the purpose of the sampling, the question, and the aims.
Since this is unknown and may even vary from case to case, only a general strategy can be recommended.

The previous chapters have shown that there is no single perfect solution; a sampling strategy it is by definition a compromise between level of perfection and workload, and strongly driven by the needs of the customer. As a consequence, there is no single strategy getting full support from all chemists. Nevertheless, ENFSI seeks a proposal with a broad support, giving individual laboratories the possibility to do more in cases where they consider it appropriate. In specific cases it is the specific chemist who will have to explain to the court that s(he) did (or instructed to perform) a realistic sampling. This factor is important since especially the explanation of a Bayesian approach may be difficult to courts.

The ENFSI subcommittee on sampling and the ENFSI steering committee advise that a sampling strategy for ‘international’ cases:

a) Must have a basis easily to explain in terms of statistics;
b) Must be easily understood and be practical, also for use by police and custom officers;
c) Must be realistic, and not result in an increase in workload for the laboratories (resulting in unacceptable long turn-over times). This means a minimum approach and no maximum approach;
d) Must still be reasonably defendable in court.

As a result of these requirements, it is proposed to advice as the minimum standard for large international cases:

1. A detailed report on the seizure by the police / customs officers, or forensic experts, for use by the experts and the court. To include description, numbers, weighing, packaging, origin, external characteristics, appearance, pictures, etc.).

2. A sampling technique with a hypergeometric or Bayesian basis.

To choose a 95% confidence level.
To choose a 50% proportion level (at least half of the items).

This means that 5 samples must be taken for chemical investigation (if it is expected that all sampled units contain drugs).

If a re-sampling is not possible, 8 samples are recommended.
Note: These 8 samples are based on the possible (but unlikely) finding that 1 of these samples appears to be negative. In that case still 50% of the packages can be guaranteed to be positive for drugs.
If the material gives rise to some doubt, at least 11 samples are recommended. Note: these 11 samples are based on the possible (but unlikely) finding that 2 of these samples appear to be negative. In that case still 50% of the packages can be guaranteed to be positive for drugs.

Note: If a forensic laboratory is doing the sampling or the sub-sampling, the number of samples can be influenced by the actual findings of the chemical analysis. Hypergeometric or Bayesian tables can be used to calculate the sample size.

Chapter : ESTIMATION OF WEIGHT AND TABLET NUMBERS

The Student *t*-distribution, relative to df degrees of freedom (see **Table 7**), can be used to calculate an interval that contains with $(1-\alpha)100\%$ probability the weight of a drug unit in a population.

Application

Using the Student *t*-distribution theory, we can estimate the average weight of a drug unit in a population, along with its associated uncertainty, within a given confidence level $(1-\alpha)100\%$.

Table 7: Student t-distribution.

Critical values for some degrees of freedom df and a confidence coefficient α equalling either 0.05 or 0.01.

df	A	
	0.05	0.01
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898

df	A	
	0.05	0.01
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845
21	2.080	2.831
22	2.074	2.819
23	2.069	2.807
24	2.064	2.797
25	2.060	2.787
26	2.056	2.779
27	2.052	2.771
28	2.048	2.763
29	2.045	2.756
30	2.042	2.750
40	2.021	2.704
60	2.000	2.660
120	1.980	2.617
∞	1.960	2.576

This can be expressed by the following relation:

$$\overline{X} - \frac{s}{\sqrt{n}} t_\alpha \leq \mu \leq \overline{X} + \frac{s}{\sqrt{n}} t_\alpha ,$$

where:

$\mu =$ the average weight of the drug unit in the population;

$\overline{X} =$ the average weight of the drug unit in the sample;

$s =$ the standard deviation of the measurements;

$n =$ the sample size;

and t_α is the critical value of the Student *t*-distribution with $df = n - 1$ degrees of freedom within the confidence coefficient α (**Table 7**).

The uncertainty of the calculated average weight, $u_{\bar{X}}$, is represented by the term $\frac{s}{\sqrt{n}}$.

Therefore the above expression becomes:

$$\overline{X} - u_{\bar{X}}\, t_\alpha \leq \mu \leq \overline{X} + u_{\bar{X}} t_\alpha$$

In addition to the uncertainty associated with the calculated average weight ($u_{\bar{X}}$), there is also an uncertainty associated with the weighing balance used (u_w).

The combined uncertainty (u_c) is calculated as follows:

$$u_c = \sqrt{u_{\bar{X}}^{\,2} + u_w^{\,2}}$$

giving the following confidence interval:

$$\overline{X} - u_c\, t_\alpha \leq \mu \leq \overline{X} + u_c t_\alpha$$

In practice, an appropriate software application can be used to assist with the determination of the confidence interval applied to the estimated weight of the drug unit.

In common practice, an acceptance criterion is that the sampling results are taken into consideration if the ratio between the standard deviation *s* and the average weight $\overline{X}$ of a drug unit in the sample is less than 0.1 (RSD<10%). Otherwise, an increase of the sample size is required in order to reach the target percentage. (If this cannot be reached because the sample weight is not a normally distributed random variable, we could be forced to weigh the entire exhibit, not using statistical inference any more).

The estimation of the total weight of the exhibit (W) can be obtained by multiplication by N of the average value and the combined uncertainty as follows.

If $w = N\overline{X}$ and $U_T = Nu_c$ then the estimation of the total weight W is:

$$w - U_T t_\alpha \leq W \leq w + U_T t_\alpha .$$

The same approach can be used for the estimation of the total weight of illicit drug in an exhibit, after quantification of the drug present in each sample unit.

If r negative results are obtained after the analysis of the drug units, for the estimation of the weight of the total (positive) drug exhibit, a corrector factor $P_{corr} = \frac{n-r}{n}$, should be used.

Also, if r negative results are obtained, this will change the uncertainty associated with the calculated average ($u_{\bar{x}}$), thus requiring a corrected uncertainty $u_{\bar{x}(corr)}$

$$u_{\bar{X}(corr)} = \frac{s}{\sqrt{n-r}}$$

which gives a corrected combined uncertainty

$$u_{c(corr)} = \sqrt{u_{\bar{X}(corr)}{}^2 + u_w{}^2}$$

So the new estimation for W becomes:

$$P_{corr} w - P_{corr} U_{T(corr)} t_\alpha^* \leq W \leq P_{corr} w + P_{corr} U_{T(corr)} t_\alpha^* .$$

where $U_{T(corr)} = Nu_{c(corr)}$

Moreover, for a population where $\frac{n}{N} > 0.1$, a further correction factor $Q_{corr} = \sqrt{\frac{N-n}{N}}$ should be applied, giving:

$$P_{corr} w - Q_{corr} P_{corr} U_{T(corr)} t_\alpha^* \leq W \leq P_{corr} w + Q_{corr} P_{corr} U_{T(corr)} t_\alpha^* ,$$

where t_{α}^{*} is the critical value of the Student *t*-distribution with *(n-r-1)* degrees of freedom (see Stoel & Bolck, 2010). Please note, that if the uncertainty in P_{corr} is not taken into account, then more optimal confidence intervals may exist (Alberink, Bolck and Stoel, 2010) and that weight estimation could also be approached from a Bayesian perspective (see Aitken & Lucy, 2002).

Example 1

Let's suppose that an exhibit of suspected heroin is contained in 100 packages. We want to estimate the average weight of a drug unit in the population with a probability of 95%.

According to the applied representative sampling theory, following the example indicated in the chapter about the hypergeometric distribution, a sample of 23 units is taken and each of them weighed and analysed.

The average net weight of the powder in the 23 units is $\overline{X} = 0.265g$ with the standard deviation *s* of 0.023 g . Since the error is 8.7%, the acceptance criterion is satisfied.

The value of t_α taken from the **Table 7** is 2.074, the corrector factor Q_{corr} is 0.877 and the estimated weight for the total exhibit W is:

$$(26.500 - 0.873)\,g \leq W \leq (26.500 + 0.873)\,g\,.$$

If one negative result is obtained after the analysis of the drug units, and a reduction in confidence and/or guaranteed percentage of positives because of this is accepted, then with the same values of the mean and standard deviation the corrector factor is $P_{corr}=22/23$, t_α^* equals 2.08, and Q_{corr} remains 0.877. Assuming, for the sake of this example, that the values of $\overline{X}$ and *s* remain unchanged, then the estimated weight for the total positive drug exhibit is W_1 is:

$$(25.348 - 0.856)\,g \leq W_1 \leq (25.348 + 0.856)\,g\,.$$

In the same way, if two negatives results are obtained, the corrector factor is $P_{corr}=21/23$, t_α^* equals 2.0860, and again Q_{corr} remains 0.877. So we have

$$(24.196 - 0.839)\,g \leq W_2 \leq (24.196 + 0.839)\,g\,.$$

Theory

The Student *t*-distribution theory may solve problems of estimation of the average of a number of measurements n. The definition of the Student *t*-distribution, relative to df degrees of freedom, is:

$$f(t)=\frac{\Gamma\left[\frac{1}{2}(df+1)\right]}{\Gamma\left(\frac{df}{2}\right)\sqrt{\pi df}}\left(1+\frac{t^2}{df}\right)^{-\frac{1}{2}(df+1)}.$$

If α is a threshold index, the value t_α according to which the probability calculated between $-t_\alpha$ and t_α is equal to $1-\alpha$, can be calculated from the following equation:

$$P_\alpha=1-\alpha=\frac{\Gamma\left[\frac{1}{2}(df+1)\right]}{\Gamma\left(\frac{df}{2}\right)\sqrt{df\pi}}\int_{-t_\alpha}^{t_\alpha}\left(1+\frac{t^2}{df}\right)^{-\frac{1}{2}(df+1)}dt.$$

The critical values of the equation for some values of df and α are listed in **Table 7**.

Estimation of Tablet Numbers

The same process can be applied to the estimation of tablet numbers (in a single population) by obtaining the average weight and combined uncertainty of a number of individual tablets. The estimated total number of tablets can be obtained by dividing the total weight of tablets in the population by the average tablet weight. The extrapolated uncertainty is obtained by multiplying the combined uncertainty by the estimated total number of tablets *(note: the extrapolated uncertainty must also incorporate the uncertainty associated with all weighing balances when more than one balance is used).*

REFERENCES

Aitken C.G.G., Sampling - How big a sample?, *Journal of Forensic Sciences,* JFSCA, 1999, 44(4), 750-760.

Aitken C., Bring J., Leonard T., Papasouliotis O., Estimation of quantities of drugs handled and the burden of proof, *Statist. Soc.,* 1997, 160(2), 333-350.

Aitken, C.G.G & Lucy, D. Estimation of the quantity of a drug in a consignment from measurements on a sample, *Journal of Forensic Sciences,* JFSCA, 2002, 47, 968-975.

AlberInk I, Bolck A, Stoel RD, Comparison of frequentist methods for estimating the total weight of consignments of drugs, *J Forensic Science 2010 Mar 1;55(2):508-12.*

Amraoui Y., Allio I., Garcia C., Perrin M., Echantillonnage et interprétation: application aux produits de saisie analysés par un laboratoire de toxicologie, ATA, 2001, vol XIII, n°4, 265-274.

Azoury M., Grader-Sageev D., Avraham S., Evaluation of Sampling Procedure for Heroin Street Doses, *Journal of Forensic Sciences,* JFSCA, 1998, 43 (6), 1203-1207.

Clark A.B, Clark C.B., Sampling of Multi-unit Drug Exhibits, *Journal of Forensic Sciences,* JFSCA, 1990, 35 (3), 713-719.

Colon M., Rodriguez G., Diaz R.O., Representative Sampling of "Street" Drug Exhibits, *Journal of Forensic Sciences,* JFSCA, 1993, 38(3), 641-648.

Coulson S.A., Coxon A., Buckleton J.S., How many Samples from a Drug Seizure Need to be analysed, *Journal of Forensic Sciences,* JFSCA, 2001, 46(6), 1456-1461.

ENFSI DWG document, DWG-SGL-002 version 001, Hypergeometric Sampling Tool (version 2012) Background of Calculation and Validation, 2012, www.enfsi.eu.

ENFSI DWG document, DWG-SGL-001 version 002, Validation of the 'Guidelines On Representative Drug Sampling', www.enfsi.eu.

Frank, R.S., Hinkley, S.W. and Hoffman, C.G., Representative Sampling of Drug Seizures in Multiple Containers, *Journal of Forensic Sciences,* JFSCA, 1991, 36 (2), 350-357.

Masson (Ed.), Initiation aux méthodes de la statistique et du calcul des probabilités, Paris, 1996, 179-180.

Miller J., Statistics and chemometrics for analytical chemistry, 4th Edition, Pearson education limited, Harlow, 2000.

Stoel, R.D. and Bolck, A., A correction to Tzidony and Ravreby (1992): 'A Statistical Approach to Drug Sampling: A Case Study', *Journal of Forensic Sciences* JFSCA, 2010, 55(1):213-4.

Sutherland G.,J., Sampling and Identifying Multiple Discrete Objects containing Drugs, Analog, Australasian Forensic Drug Analysis Bulletin, Vol. 14, No. 1, January 1992, 9-12.

SWGDRUG, Sampling Seized Drugs for Qualitative Analysis, www.swgdrug.org

SWGDRUG, Supplemental document SD-3 for Part IVC – Quality Assurance/Uncertainty: "Measurement Uncertainty for Weight Determinations in Seized Drug Analysis". www.swgdrug.org

SWGDRUG, Supplemental document SD-6 for Part IVC – Quality Assurance/Uncertainty: "Measurement Uncertainty for Extrapolations of Net Weight and Unit Count". www.swgdrug.org

Tzidony D., Ravreby M., Statistical Approach to Drug Sampling: A Case Study, *Journal of Forensic Sciences,* JFSCA, 1992, 37(6), 1541-1549.