TECHNICAL NOTE

Alan B. Clark,[1] M.S., MBA and Charles C. Clark,[1] B.S.

Sampling of Multi-Unit Drug Exhibits

REFERENCE: Clark, A. B. and Clark, C. C., **"Sampling of Multi-Unit Drug Exhibits,"** *Journal of Forensic Sciences*, JFSCA, Vol. 35, No. 3, May 1990, pp. 713–719.

ABSTRACT: The total number of drug exhibits submitted to forensic science laboratories continues to grow markedly, and increasingly, these exhibits are composed of ever larger numbers of units. These facts make the use of a sampling plan, which applies some limitations to the number of units sampled, highly desirable. A sampling plan assumes that the characteristics of the nonexamined units are the same as those in all of the examined units. This paper examines the validity of this assumption by the use of mathematical concepts. Various sampling plans, exhibit sizes, and sample sizes are used as examples. The effect of sample size on quantitative accuracy is also discussed.

KEYWORDS: toxicology, sampling, drug identification, drug exhibits

Drug exhibits submitted for analysis often consist of a single unit in some sort of container, for example, a foil package, a vial, or perhaps a kilo package. In these instances, an analysis is performed on a small amount of material removed from the unit and the analytical result is considered representative of the entire contents of the unit. Sometimes an exhibit consists of several units of material. Each of the units can be presumptively tested, and a portion from each combined into a composite for the actual analysis. Again, the analytical result is considered representative of the entire contents of all units in the exhibit. In either of the above two cases all units have been examined; thus, the probability that all units contain the ingredient identified is 100%.[2]

When the situation becomes more complicated, as in much larger seizures of 100, 1000, 5000 units or more, what testing procedure is to be followed? Resource constraints on the analytical laboratory and sentencing guidelines based on possession limits necessitate some sort of alternative testing procedure wherein not all units in a given submission are sampled. Obviously, if all units are not sampled, the mathematical probability that all units contain the identical ingredient is less than 100%. This leads to questions regarding the confidence level attained using limited sampling procedures and the effect this use might have on the criminal justice system which relies on our analyses.

To turn limited sampling scenarios into questions which can be dealt with mathematically, we ask

1. What is the probability of finding the only unit that is different in a population of N units when n units are sampled, where $n < N$?

Received for publication 1 April 1989; revised manuscript received 3 July 1989; accepted for publication 6 July 1989.

[1]Supervisory chemist and forensic chemist, respectively, Drug Enforcement Administration, Miami, FL.

[2]For purposes of this paper, no differentiation is made between conclusive analytical techniques and presumptive testing. Thus, a given unit is considered as being positive for the controlled substance in question if it reacts positively to any sort of testing technique, whether presumptive (for example, with color tests) or conclusive (for example, with infrared).

2. What is the probability of finding one or more of the different units in a population of N units when only n units are sampled and there is more than one different unit present?

Permutations and Combinations

Before discussing examples of the treatment of these two questions, brief consideration should be given to the underlying mathematical concepts of permutations and combinations.

Consider, for example, an exhibit consisting of two black units and two white units. Figure 1 shows the 24 orders in which these four units could be examined. These orders are called permutations (P) or variations (V). Generalizing, for a population of N units, the number of possible permutations is equal to $N!(N{\cdot}(N - 1){\cdot}(N - 2)\cdots 1)$. Suppose, however, that one wishes to examine only a portion of the population for example, three of the four units in our exhibit. Figure 2 shows the twenty-four orders in which these three units could be selected. Again generalizing, the number of different orders in which r distinct objects can be selected from N distinct objects is $P(N,r) = N!/(N - r)!$. If the order in which the three units are selected is not important, then, as shown in Fig. 3, only four combinations are possible. Each of three units which can be selected from the exhibit in six different orders. Again generalizing, the number of combinations of r units taken from a total of N units without regard to order, can be written as:

$$C(N,r) = \frac{N!}{(N-r)!r!} \tag{1}$$

Using our example of two black units and two white units, this combination rule could be used to determine the probability of selecting exactly two white units and one black unit. This probability = (the number of possible combinations containing exactly two white units and one black unit)/the total number of possible combinations of three units. Since the example contains only two white units, they can be selected in $C(2,2)$ ways. The black unit can be either of the two and can be selected in $C(2,1)$ ways. Remembering

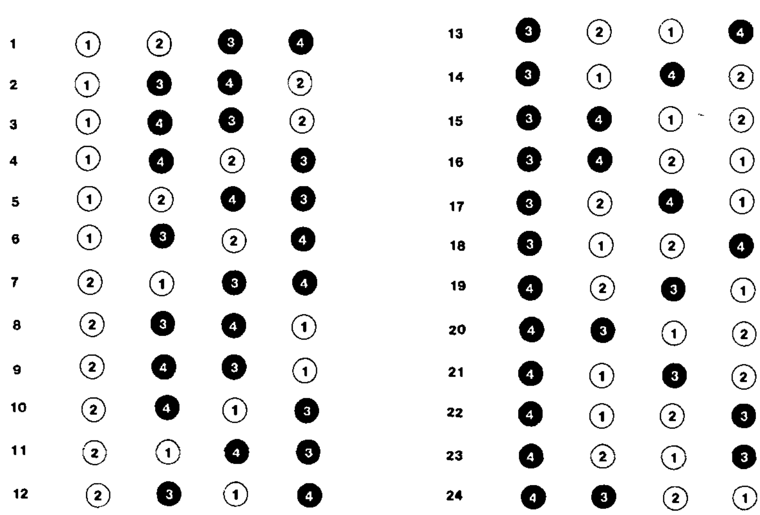

FIG. 1—*Number of ways (permutations) four distinct objects can be chosen for examination.*

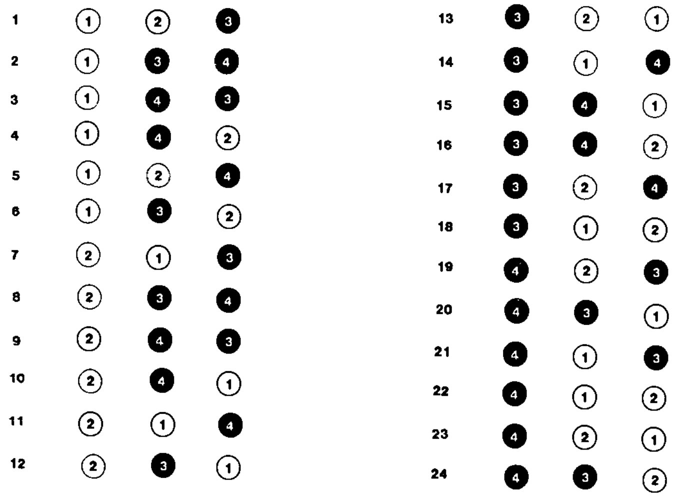

FIG. 2—*Number of ways (permutations) three out of four distinct objects can be chosen for examination.*

that the number of ways in which two independent events can occur together is the product of the number of ways each can occur separately, the probability (P) of selecting exactly two white units and one black unit from a group of two each =

$$\frac{C(2,2)C(2,1)}{C(4,3)} = \frac{2!}{2!(2-2)!} \cdot \frac{2!}{1!(2-1)!} \frac{4!}{3!(4-3)!} = \frac{2}{4} = 0.5$$

(Note ! by definition $0! = 1$). Examination of Fig. 3 shows that, indeed, two of four combinations in the example contain exactly two white and one black unit.

In general terms, the probability (P) of selecting exactly A majority units and B minority

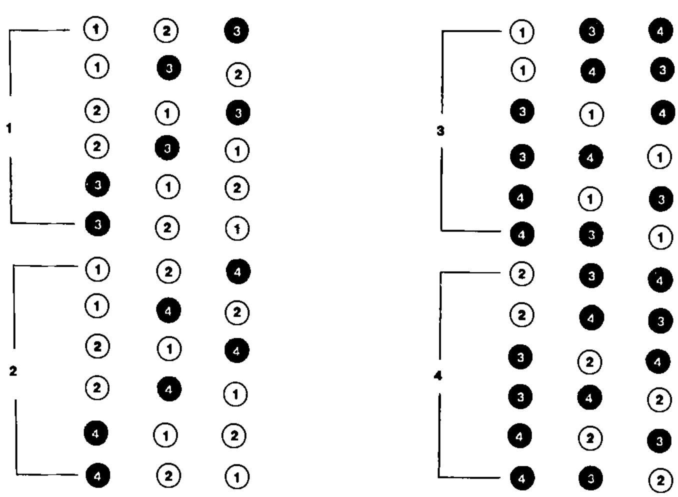

FIG. 3—*Number of ways (combinations) three out of four distinct objects can be chosen for examination (order not important).*

units from a population N consisting of X majority units and Y minority units can be written as:

$$\frac{C(X,A)C(Y,B)}{C(N,A+B)} \tag{2}$$

Returning to Question 1, this equation can be used to calculate the probability of finding the 1 different unit when 10 of 100 units are examined:

$$P = \frac{C(99,9)C(1,1)}{C(100,10)} = \frac{99!}{9!(99-9)!} \cdot \frac{1!}{1!(1-1)!} \bigg/ \frac{100!}{10!(100-10)!}$$

$$= \frac{99!}{9!90!} \bigg/ \frac{100!}{10!(90!)}$$

$$= \frac{10}{100}$$

$$= 0.1$$

This means that only 10% of all possible combinations of 10 units each will contain the odd unit. One can also use Eq 2 to calculate the probability of finding at least 1 odd unit when more than 1 odd unit is present, for example, the probability of detecting at least 1 "cut" unit when 5 units are examined from a population consisting of 190 "uncut" units and 10 "cut" units. The sum of all probabilities must = 1. Therefore, the summed probabilities of detecting 5, 4, 3, 2, or 1 "cut" units in the screened units is 1 minus the probability of finding 0 "cut" units or

$$P = 1 - \frac{C(190,5)C(10,0)}{C(200,5)} = 0.229$$

Thus, even when only 5 units are examined from this population, almost 1 in 4 of all possible combinations of 5 units will contain at least 1 of the odd units [*1*,*2*].

Four Sampling Approaches

To deal with the first question posed earlier concerning the chances of finding the one different unit in a population when all units are not sampled, we will arbitrarily consider four different approaches to sampling populations of various sizes: (1) sampling five units regardless of the population's size, (2) sampling 10% of the population, (3) sampling a number of units equal to the square root of the population, and (4) sampling 50% of the population. Table 1 shows the result of these four sampling approaches for populations of 10, 50, 100, 500, 1000, 1500, and 2000 units.

For Approach 1, when five units are sampled, the probability of finding the only different unit present is one in two for a population of ten but, as the population increases in size, the probability of finding the one different unit decreases rapidly and quickly approaches zero. This is shown graphically in Fig. 4. For Approach 2, where 10% of the units are sampled regardless of population size, the probability of finding the only different unit is constant at one in ten. This is also shown in Fig. 4. For Approach 3 in which the square root of the population is sampled, the probability of finding the one odd unit present decreases rapidly and approaches a zero likelihood, as a practical matter, very

TABLE 1—*Probability of finding the single odd unit in a population of N units.*

	Probability of Detection, %			
Population, N	$n = 5$	$n = 10\%N$	$n = \sqrt{N}$	$n = 50\%N$
10	50	10	40	50
50	10	10	16	50
100	5	10	10	50
500	1	10	4.6	50
1000	0.5	10	3.2	50
1500	0.3	10	2.6	50
2000	0.25	10	2.2	50

quickly. Even for a relatively small population, for example, ten units, the probability is less than one in two. For Approach 4, where 50% of the population is sampled, the probability of finding the odd unit is constant at one in two even though the number of sampled units is relatively large.

In each of the above cases, the probability of finding the single odd unit in a population never exceeds 0.5 for a population >10. More importantly, it can be seen that increasing the number of units sampled does not greatly increase the assurance of finding the odd unit. As shown in Fig. 4, for Approach 2 when a much greater number of units is sampled at larger population sizes, the chance of finding the odd unit still is equal to 1 in 10. Thus, for a population of 1000 units where 100 are sampled, the chance of not finding the odd unit is 90%. Clearly, dramatically increasing the number of units sampled does not increase the likelihood of finding the odd unit to an extremely high probability. In other words, the dividend in terms of probability of finding the odd unit is not commensurate with the extra time and effort involved in sampling more units. Furthermore, of what significance is it that a single odd unit in a population may not be discovered? The overwhelming majority of the units in these cases would still consist of the units of interest and the single odd unit would be insignificant in practical terms, because there is a limit to which criminal sentences are linked to the amount possessed. For amounts beyond certain weights, the sentence is not increased.

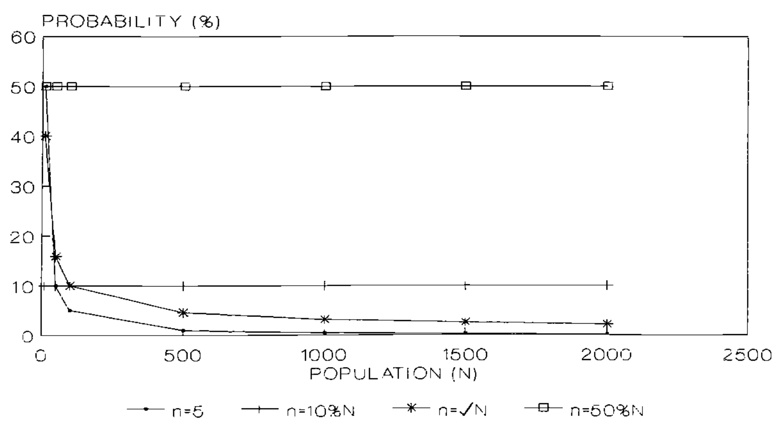

FIG. 4—*Probability of detection (various populations).*

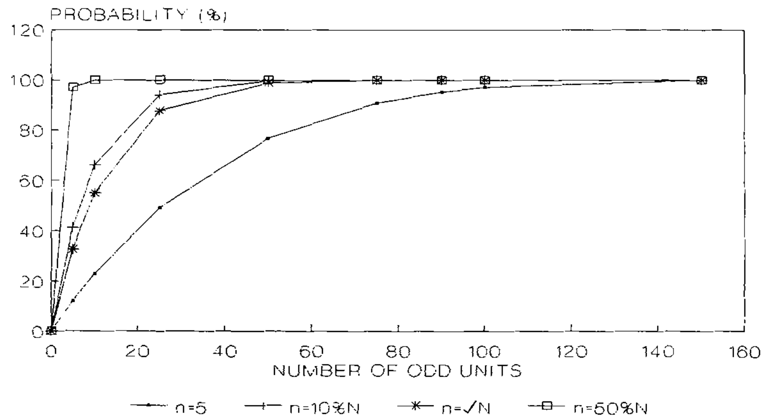

FIG. 5—*Probability of detection (population of 200).*

More Than One Unit in a Population

Perhaps a more useful question to ask is "what is the chance of finding one odd unit when a more significant number of units are odd?" This situation can be addressed mathematically as posed earlier in Question 2: What are the chances of finding one or more of the units different from those of interest in a population of N units when n are sampled and there is more than one unit that is different?

As an example, consider the situation in which the population size is 200 units, the number to be screened is $n = 5$, $n = 10\%N$, $n = \sqrt{N}$, $n = 50\%N$, and the number of different units varies from 0 to 150. What, then, are the chances of finding one or more of the odd units in each? This would at least alert the analyst to nonuniformity in the population. As shown in Table 2, as the number of odd units in the population increases, the chance of finding at least one of those units increases dramatically to a virtual certainty even when the odd unit population is less than half of the entire population. This is further demonstrated in Fig. 5, where the probability of finding at least one of the odd units is shown to increase rapidly with increasing numbers of odd units present. Even with limited sampling the graph shows that there is little room left for dramatic improvement in the probability of detection by sampling a greater number of units.

TABLE 2—*Probability of finding at least one odd unit in a population of 200 units.*

	Probability of Detection, %			
No. of Odd Units	$n = 5$	$n = 10\%N$	$n = \sqrt{N}$	$n = 50\%N$
0	0	0	0	0
5	12.1	41.3	32.6	97.0
10	22.9	66.1	55.0	99.9
25	49.1	94.0	87.6	99.9
50	76.7	99.8	98.9	99.9
75	90.8	99.9	99.9	99.9
90	95.2	99.9	99.9	99.9
100	97.1	99.9	99.9	99.9
150	99.9	99.9	99.9	99.9

Conclusion

It is easy to recognize that, in a given population, if all units are sampled and are positive the probability is 100% that all units contain the identified ingredient.[3] Similarly, if all units in a population are not sampled, the probability that all units contain the identified ingredient is less than 100%. Given the practical resource limitations present in most analytical laboratories, it is neither feasible nor is it necessary to have a 100% probability that all of the units in the submission are the same.

Furthermore, the guilt or innocence of a defendant is not affected by a limited sampling procedure as long as a positive identification has resulted from analysis of those units which were sampled. If the identity of each and every unit in a multi-unit submission is proven to a probability of less than 100%, the only area in which this might be a factor is in sentencing, assuming there is a statutory relationship between the sentence and the amount of contraband possessed.[4] In this situation, the sentencing authority may have to determine if the amount of material in possession of the defendant was satisfactorily proven.

Acknowledgments

We would like to express our appreciation to the Forensic Sciences Section of the Drug Enforcement Administration as well as the staff of the Southeast Laboratory for their review of this article. We would also like to thank Dr. Sidney W. Hinkley for his extremely helpful suggestions and Mrs. Yvonne Harder for typing the manuscript.

References

[*1*] Daniel, W. W., *Biostatistics: A Foundation For Analysis In The Health Services*, 2nd ed., John Wiley and Sons, New York, 1978, pp. 42–50.

[*2*] Bauer, E. L., *A Statistical Manual For Chemists*, 2nd ed., Academic Press, New York, 1971, pp. 131–136.

[*3*] Natrella, M. G., "Experimental Statistics," in *NBS Handbook 91*, U.S. Department of Commerce, National Bureau of Standards, U.S. Government Printing Office, 1963.

[*4*] Taylor, J. K., *Quality Assurance of Chemical Measurements*, Center for Analytical Chemistry, National Bureau of Standards, Gaithersburg, MD, 1984.

Address requests for reprints or additional information to
Alan B. Clark
Drug Enforcement Administration
5205 NW 84th Ave.
Miami, FL 33166

[3]Here we assume that the criteria (or screening procedure) for determining the identity of each is adequate.

[4]However, at least in the Federal Statutes, there is only a finite relationship between the amount possessed and sentence imposed; beyond certain possession limits, the sentence does not increase.