OPTIONAL FORM 99 (7-90)

FAX TRANSMITTAL	# of pages ► 4
To Steve Zappe	From Kyle Rogers
Dept./Agency	Phone #
Fax # 505 827-1544	Fax # (202) 233-9626

NSN 7540-01-317-7368 5099-101 GENERAL SERVICES ADMINISTRATION

October 21, 1997

MEMORANDUM

TO: Kyle Rogers

FROM: Jonathan Cohen and Howard Finkel

SUBJECT: Review of Hypergeometric Sampling Procedure

SUMMARY

Real time radiography (RTR) of waste drums has a certain miscertification rate that occurs when a drum is determined to meet the WIPP-WAC and TRAMPAC criteria by the RTR operator but the drum does not meet those criteria based on a visual examination. The miscertification rate can be estimated by visually examining a random sample of drums and computing the fraction of miscertified drums in the sample. The DOE has proposed a simple random sampling method, whereby a random sample of drums is selected at random without replacement; the selection does not take into account any information about the waste streams for the drums.

In this memorandum, we review DOE's methodology and discuss the issue of whether or not a stratified sampling approach should be used to take into account the (type of) waste stream. Under such an alternative, separate random samples would be taken from each different waste stream. In this manner the overall sample would be guaranteed to have a certain number of drums from each waste stream; for example, a simple random sample might by chance have all the drums from the same waste stream.

Our overall conclusion is that DOE's procedure is valid, and properly reduces the chances of making an erroneous decision about the true, unknown, miscertification rate. However, a stratified scheme based on the waste stream would be more efficient, so that a smaller total sample size could be used to give the same control of the error rates. Equivalently, using a stratified approach and the same sample sizes as DOE's approach would lead to reduced probabilities of making an erroneous decision. Furthermore, a stratified sampling scheme has the advantage that the miscertification rates for different waste streams can be evaluated and could be separately controlled.

ANALYSIS

Simple Random Sampling

Consider this problem as determining whether or not the true miscertification rate is below 14 percent. There are two ways of making a wrong decision:

- Type I error: Deciding that the true rate is below 14 percent when it is above 14 percent, and

971032

- Type II error: Deciding that the true rate is above 14 percent when it is below 14 percent.

The sample size is selected in such a manner that the probabilities of making erroneous decisions are sufficiently small:

- The probability is at most 0.1 that the miscertification fraction is determined to be below 14 percent when the true miscertification fraction is 14 percent or higher (Type I error).

- The probability is at most 0.2 that the miscertification fraction is determined to be above 14 percent when the true miscertification fraction is no more than the percent miscertified in previous year(s) (Type II error).

For example, if the percentage miscertified in the previous year was 5 percent, and there were 400 drums, then DOE would visually examine 45 drums (their Table 5-1). Calculations by ICF using the hypergeometric distribution, show that if the 45 drums are selected at random without replacement, similar in principle to a lottery ticket drawing, then DOE would decide that the true miscertification fraction was below 14 percent if and only if 3 or fewer (less than 6.7 %) of the 45 sampled drums were shown by visual examination to be miscertified. The probability of wrongly determining that the miscertification rate is below 14 percent is at most 0.097. The probability of wrongly determining that the miscertification rate is above 14 percent when the true rate is 5 percent is 0.821. Thus the probabilities of Type I and II error are 9.7 % and 17.9 %, which meet the above criteria.

DOE's procedure uses the above method expressed in a different, but mathematically equivalent, formulation to determine sample sizes. If x sampled drums are miscertified out of a sample of n drums, and if the population of all drums has N drums and M miscertified drums, then a 90 percent upper confidence limit for M, M(x), is found by solving the following inequality for M:

$$\sum_{k=0}^{x} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \leq \alpha, \text{ where}$$

$$\alpha = 0.1.$$

Note that DOE (Appendix A) stated that M(x) is the largest M satisfying the inequality. In fact, the left hand side decreases as M increases, so M(x) should be defined as the smallest M satisfying the inequality.

DOE's method would deem acceptable any case where the upper confidence limit is below 14 percent of N. The sample size is determined by requiring that the left hand side is at least 0.80 when the true miscertification rate is the same as the previous year's rate and when x is any value such that M(x) <= 0.14N.

ICF calculations confirm the equivalence of the two formulations, and confirm the sample size determinations in Table 5-1. Note, however, that in Step 5 of DOE's algorithm, the possible case where $\alpha = a_i$ should be dealt with by stopping whenever: $a_{i-1} <= \alpha$ and $a_i > \alpha$, or; $a_{i-1} > \alpha$ and $a_i <= \alpha$.

Stratified Random Sampling

Alternative approaches might use knowledge of the waste stream type to select the sample of drums and appropriately determine if the overall miscertification rate was sufficiently low. A crucial point is that, although DOE's approach ignores this possibly useful information, the sample sizes will be large enough to control the Type I and II probabilities of making an incorrect decision. Stratified random sampling might be used to further reduce these error probabilities for the same sample sizes or to reduce the required sample sizes.

Consider the example case where waste stream A has 400 drums with 4 miscertified (1 %) and waste stream B has 100 drums with 21 miscertified (21 %). The overall miscertification rate is 25/500 = 5 %. Suppose that 5 percent were miscertified in the previous year. DOE's procedure would require 45 drums to be sampled at random without replacement. Because the waste stream B has fewer drums in the population, the simple random sample will most likely contain relatively few B drums. Thus the simple random sample will contain relatively few of the drums with the higher miscertification rate. This does not introduce a bias because the population also contains relatively few drums with the higher miscertification rate. Roughly speaking, the sample will contain about 80 % of waste stream A drums, i.e. about 36 drums from waste stream A and 9 from waste stream B. Of the 36 A drums, roughly 0 or 1 will be miscertified (1 % of 36 = 0.36). Of the 9 B drums, roughly 2 will be miscertified (21 % of 9 = 1.9). Thus, one would expect about 2 or 3 miscertified drums in the sample, which agrees with the population miscertification rate (5 percent). More exactly, the expected numbers of miscertified A and B drums in the sample are 0.36 and 1.89, which totals 2.25, (5 percent of 45).

Although DOE's approach does not take into account the waste stream type, the error probabilities are controlled. In this example case of 500 drums with 5 percent miscertified, the Type I and II error probabilities for a sample of 45 drums will be 9.7 and 17.9 percent, respectively, regardless of how many drums of each waste stream are in the population and their miscertification rates. This holds for the above case and for a more extreme case where waste stream A has 475 drums, with no miscertifications, and waste stream B has 25 drums, all miscertified. This extreme case shows that information about waste stream type can be used to reduce the required sample sizes, i.e. to sample more efficiently. If one knows that waste stream A is (almost) always correctly certified (as meeting the WIPP-WAC and TRAMPAC criteria), then one only needs to sample from the 25 waste stream B drums, a maximum sample size of 25.

In general, the optimum stratified sampling procedure (as shown in e.g., Cochran, W. G., "Sampling Techniques," Wiley, 1977) is to sample so that the fraction of sampled drums of a given waste stream type is approximately proportional to $\sqrt{(p(1-p)}$, where p is the miscertification rate for that waste stream. A separate random sample without replacement is taken from each waste stream's drums, and the estimated miscertification rate is a suitably weighted average of the waste stream miscertification rates. Obviously, p is not exactly known in practice, but this rule can be interpreted to mean that one should sample more from waste streams where the miscertification rate is closer to 0.5 (above or below). Using these techniques the variance of the estimated miscertification rate will be lower than for simple random samples of the same total size, provided that the miscertification rates vary sufficiently with the waste stream type. Thus one could either use DOE's total sample size estimates and obtain lower Type I and II error rates, or could obtain the same error rates with a smaller sample.

If necessary, the available data on waste stream miscertification rates could be examined to evaluate the potential increased efficiency of a stratified random sampling scheme. As pointed out in the last paragraph, if the waste stream miscertification rates are unknown and do not vary significantly, then a stratified sampling scheme will not offer an improved efficiency over DOE's simple random sampling scheme. Otherwise a more efficient sampling protocol can be designed. In many cases, even if the waste stream miscertification rates are completely unknown, a more accurate procedure would be to sample with equal sampling proportions from each set of waste stream drums. This proportional sampling method is often nearly as efficient as the optimum procedure and leads to much easier computations.