66

1949 PREPRINT.—This paper will be published by the **American Society for Testing Materials,** 1916 Race St., Philadelphia 3, Pa. It is issued in preprint form primarily to stimulate discussion. Discussion, either for publication, or for the attention of the author, is invited and may be transmitted to the Executive Secretary. The paper is subject to modification and is not to be republished as a whole or in part pending its release by the Society through the Executive Secretary.

ASTM 80

SOME PROBLEMS IN THE SAMPLING OF BULK MATERIALS

By Louis Tanner[1] and W. Edwards Deming[2]

Synopsis

The sampling of lots of bulk materials for the purpose of estimating the total weight and some quality characteristic presents various kinds of difficulties, both theoretical and practical. The fundamental character of a statistical plan is demonstrable and controllable precision. In order to use the statistical theory for the sampling of finite populations, it is necessary for the shipment to be divisible into sampling units in such a manner that every particle of the shipment lies in one and only one sampling unit, and that all of any unit and no material from any other may be drawn in a routine manner from the lot. This condition is met by packaged materials, an example being wool in bales, from which cores may be drilled in specified positions. A sample of bales may be drawn from the shipment, and a sample of cores may be taken from the sample bales, these cores then to be tested to furnish an estimate of the average quality of the shipment. The total weight of the shipment may be estimated from the weights of the sample bales. Statistical theories have been developed and are here applied in the development of procedures for sampling sugar, tobacco, and wool, and for regulating the precisions of the estimates. Such methods have wide applicability. However, other types of statistical procedures, possibly involving theories for the analysis of an infinite supply, must be developed for materials such as a pile of coal which, for purposes of routine sampling, cannot readily be divided into definite small units. Indications point to distinct advantages of sampling such materials at the source of supply. Hazardous biases and uncontrolled errors of sampling may be encountered through failure to regard statistical principles.

Two Types of Problems

In facing the problems of the sampling of bulk materials one encounters two broad types of physical problems:

1. Materials like piles of coal or ore in which no unique subdivisions may be specified as sampling units that will be practicable for a routine sampling operation.

2. Materials in packages, bales, or bags, or any material that may be subdivided into unique sampling units that are practicable for a routine sampling operation.

A material may change from type A to type B. For example, a type A material, if arrangements are made to draw samples periodically from a conveyor belt during a loading or unloading operation, is during this operation no longer type A but type B.

Statistical theories and a growing

[1] Chief Chemist, U. S. Customs Laboratory, 408 Atlantic Ave., Boston, Mass.

[2] Adviser in Sampling, Executive Office of the President, Bureau of the Budget, Washington 25, D. C.

background of experience are being developed by which many materials, particularly those of type B, may be sampled for estimating the weight and quality of a lot. Much new research is being carried on, with the result that costs of sampling and testing are constantly being lowered and some progress is being made in devising new methods for converting certain type A materials to type B.

The particular theory to be applied to the materials of type B is known as the theory for sampling finite populations. This theory has heretofore been developed and applied mostly in the sampling of human populations and farms, and for wholesale and retail business establishments. The miniature monthly census of the population in the United States and a similar quarterly census in Canada are examples: out of these studies come very important figures for business and government, such as the monthly or quarterly report on employment and unemployment, indexes of the prices of food and clothing, and indexes of business activity.

Extension and modification of these procedures to the samplng of physical materials have commenced only recently. The next part of this paper describes an application of statistical theory to some of the problems of sampling shipments of sugar, wool, and tobacco.

The outstanding characteristic of a statistical method of sampling is the controllable and demonstrable precision of its results, which are attained at minimum cost. But before a statistical theory can be applied, certain conditions of sampling must be met. Briefly, sampling depends on the theory of probability. The theory for sampling finite populations demands that every particle of a lot have a definite and known probability of being drawn into the sample: otherwise the calculation of the precision of the results cannot be carried out. This condition is met by materials satisfying the definition given for type B. For such materials it is possible to perform random operations of sampling and to specify by number or code that certain particular sampling units are to be included in the sample, with reasonable assurance that in a routine procedure of sampling, all of the material contained in any of the units so designated for the sample will be included, and all else will be excluded.

Type A materials are those for which no practical and unequivocal sampling units can be specified in routine operations. One might specify, for example, that a pile or car of coal is to be sampled by extracting certain cubes or parallelepipeds, defined as located so many feet from the top and having certain other spacial coordinates. It would be impossible, though, as a regular routine operation, for a worker to draw into the sample the contents of any such sampling unit and to exclude all material not intended for this unit. Under such circumstances a material fails to meet the fundamental criterion of a unique probability for every particle, and it is therefore type A.

The sampling of materials of type A is thus by definition not open to the methods that are used for materials of type B. When samples are drawn from a material of type A, the drawing must too often depend on human judgment. Such dependence, experience demonstrates, is hazardous, as severe biases often develop. As a matter of fact, biases are the rule, not the exception. No statistical measure of precision can be calculated for a judgment selection.

The sampling of materials of type A is not hopeless, however, but calls for ingenuity and a new attack. If a material can be sampled off a conveyor belt at specified intervals, it is, in effect, a material of type B, no matter how it is piled up afterward.

In this paper we intend merely to state some of the problems of bulk sampling and to describe certain particular solutions to some of the problems that are encountered with materials of type B in the Bureau of Customs.

SOME APPLICATIONS OF STATISTICAL THEORY TO BULK MATERIALS OF TYPE B

Theory.—As stated earlier, a statistical plan for sampling a lot of product is expected to provide estimates of quality and weight with known and controllable errors of sampling, and to provide this control at minimum cost. The errors of sampling are controlled by the size of the sample and the procedure of drawing it. One sample design is better than another if it yields the desired precision at a lower cost than the other plan. Mathematical formulas are available which give the sampling errors corresponding to different designs, and for the determination of the most economical design.

In the theory to be applied here it is assumed that the material to be sampled has been divided into distinct units in some suitable manner—bales or packages, hourly production, etc.: in other words, the material is of type B. The problem is to sample a lot composed of a number of such units. In the general case, each of these primary units is composed of a set of secondary units—shovelfuls, cores, sub-packages, cuts, or the smallest portion that will be drawn out for testing. Tertiary or higher order units may also occur. In the special case, the primary unit itself is the test portion, in which case the general sampling theory reduces to the well-known simple theory of sampling.

Statistical theory tells us that the size of sample and the most economical design may be computed from the following equations:

Most economical number of secondary units per primary unit:

$$k = \frac{\sigma_w}{\sigma_b}\sqrt{\frac{c_1}{c_2}} \qquad (1)$$

Number of primary units to be drawn into the sample:

$$n = \frac{N(\sigma_w^2 + k\sigma_b^2)}{Nk\left(\frac{E}{t}\right)^2 + k\sigma_b^2} \qquad (2)$$

For the special case, where there are only primary units:

Where previous measurements have not been made:

$$n = \left(\frac{tV_x}{E}\right)^2 \qquad (3)$$

Where previous measurements have been made:

$$n = \left(\frac{tV_c}{E}\right)^2 \qquad (4)$$

where:

c_1 = cost of preparing a primary unit for sampling,

c_2 = cost of taking a secondary unit from a primary unit,

E = allowable uncertainty of the sample mean,

k = number of secondary units to be drawn from each sampled primary unit,

N = number of primary units in the lot,

n = number of primary units to be sampled,

t = probability factor,

V_x = coefficient of variation of the quality between primary units,

V_c = coefficient of variation of the ratio of the second quality measurement to the first,

σ_b^2 = variance of the quality between primary units, and

σ_w^2 = variance of the quality between secondary units, within a primary unit, averaged over all primary units.

Often the required information may be reliably estimated from existing data. If not, it may be developed by suitable investigation. Continuous evaluation of the pertinent statistics is possible by the adoption of suitable ways of keeping records.

Illustrations of Application.—The sampling theory given above has been successfully applied on a large scale to a variety of bulk commodities. Several cases are here described briefly as illustrations.

The simplest application of sampling theory is the case where the primary unit is the test portion. An example of this occurs in the weighing of imported refined sugar. This commodity arrives in large lots, about 43,000 standard size bags containing approximately 100 lb. each. The weighing is usually done in drafts of 8 bags at the warehouse after damaged bags, about 1.4 per cent of the total, have been segregated for separate handling. A study showed that the coefficient of variation, V_x, of the draft weights of sound bags of refined sugar was stable and averaged about 0.1 per cent. From Eq. 3 it is found that the number of random drafts that must be weighed in order that errors due to chance may be no greater than ± 0.1 per cent when $t = 3$ (a risk considered acceptable for this commodity) is 9. On the basis of this information a convenient sample size of 25 drafts was specified to insure obtaining an average bag weight for the entire cargo with a precision of better than ± 0.1 per cent.

The potential time and labor savings with a sample weighing method such as this is obviously considerable.

Had segregation of the damaged, partially empty bags not been possible, as is the case with commodities weighed at the time of discharge from the vessel, the coefficient of variation of the draft weights would have been higher, perhaps so high as to require an uneconomically large sample size to achieve the desired precision. In such cases, and in general whenever individual package weights are not standard or differ considerably, a "calibration" technique may be employed. For most commodities there is available the producer's individual package weights on the basis of which commercial transactions have been conducted. Due to changes that usually occur in transit (moisture absorption, damage, etc.), the received or settlement weights may differ considerably from the original weights, necessitating reweighing. In such cases the average ratio of the second weight of a sample to the original weight of the same packages may be used as a conversion factor to be applied to the total original weight of the lot. The coefficient of variation of the conversion factor, V_c, must be known in order to calculate the size of sample required to achieve a prescribed precision.

An example of the use of this "calibration" technique is the weighing of imported cigarette leaf tobacco, of which some 100,000,000 lb. are imported annually. This commodity arrives in lots of several thousand bales of various sizes ranging in weight from 10 to 300 lb., according to source. Previous Customs practice was to weigh each package individually—a very accurate but slow and expensive procedure. Individual bale weights at the time of shipping are available, but they are not considered sufficiently close to the landed weight to serve as the basis for the assessment of duty, nor for some commercial purposes.

It was found that the coefficient of variation of the bale weights averaged about 3.5 per cent and ranged from 1.5

per cent to 9.4 per cent. This statistic is therefore not satisfactory because so large a sample would be required for any reasonable degree of accuracy that material savings in weighing would be impossible. The coefficient of variation of the conversion factor, however, was found to be stable, averaging about 0.7 per cent except for bales of Greek tobacco for which the average was about 2.5 per cent. On the basis of these values, Eq. 4 shows that a sample size of 100 from shipments of most cigarette tobacco (900 for lots of Greek tobacco) is adequate to achieve a precision of ± 0.25 per cent, a level considered satisfactory in protecting both the government and the importer against a material loss of revenue or overcharge in duty on any one importation. On the average, these sample sizes represent a saving of from 75 to 95 per cent in the number of weighings.

These two illustrations are obviously simple applications of familiar sampling theory involving only primary units. An instructive case involving both primary and secondary units, and therefore a more general illustration of the application of statistical theory to the sampling of bulk materials, is the sampling of raw wool for the determination of its clean wool content. This case is but one of many similar instances covering a wide variety of products.

Raw wool appears in commerce in lots of various sizes, sometimes as high as 3000 bales or bags, averaging about 100 units. The packages range in weight from 200 lb. to 1200 lb. but are fairly uniform in this respect for any one source. The wool contains a number of impurities, the percentage of which may vary widely, as much as 30 per cent, both between bales in a lot and between portions within a bale. The evaluation of a lot of this commodity for any given grade and quality depends on the quantity of wool contained therein. Customarily this factor is estimated visually by qualified appraisers, but the need for a reliable objective method for determining the clean wool content without processing all or a large fraction of the lot has long been felt. The major obstacle to the realization of such a method has until recently been the unreliability of the only type of small sample that could be taken, usually a few handfuls of fleeces selected as representative on the basis of the sampler's judgment. This particular difficulty has been eliminated by the development of a practicable sampling tool that can penetrate into a bale of wool and remove therefrom a relatively small core. The question arises as to how many such cores should be taken, and from how many bales.

In the sampling of this commodity the primary unit is the package, the secondary unit the core. When reliable estimates of the within-bale and between-bale variances are available, different sampling plans may be followed, each of which yields a sample of the same precision. Thus, k_1 cores may be taken from each of n_1 randomly selected bales in the lot, or k_2 cores may be taken from each of n_2 bales. For any chosen value of k, the corresponding value of n is given by Eq. 2.

Systematic studies were made of the magnitudes and stability of σ_w and σ_b for many commercial lots of most types and grades of wool from different sources throughout the world. As was to be expected, no one single pair of values covered all the wools investigated. However, it was found that, when grouped in a relatively small number of broad classifications according to origin, end use, and type of package, these statistics were fairly stable and could serve as sufficiently reliable estimates for sampling purposes. Some typical vaues are shown in Table I.

By using these estimates of σ_w and σ_b in Eq. 2, schedules may be prepared giving, for each class of wool, the number of packages to be sampled for various lot sizes and for various numbers of cores per package, at any required level of precision. Table II illustrates such a schedule.

Each of the four plans illustrated in Table II will result in a sample of the same precision, but not at the same cost. There are two operations in obtaining a core of wool: first, positioning the bale, and second, boring it. Each operation

TABLE I.—CLASSIFICATION OF PACKAGED GREASY APPAREL WOOLS FOR SAMPLING PURPOSES.

	σ_w	σ_b
Argentina	2.5	2.5
Australia	1.5	4.0
Chile	2.0	5.0
Uruguay	2.0	1.5
Domestic	4.5	2.0

σ_w = the estimated standard deviation of the per cent clean wool content of individual cores within packages in a lot.

σ_b = the estimated standard deviation of the per cent clean wool content between packages in a usual commercial lot.

has its own cost,[3] c_1 and c_2 respectively, and their ratio may vary widely depending on the existing conditions. For example, if sampling takes place at the time the individual bales are being weighed or moved to storage, there is virtually no additional cost for positioning the bale, and $\frac{c_1}{c_2}$ might be, for example, 1.0. On the other hand, if the bales are already stored in piles in a warehouse and the piles have to be broken down to permit sampling, the cost ratio might easily exceed 25. The most economical plan of samplng is the one which uses that number of cores per bale given by Eq. 1.

The usefulness of this method has been proved in actual practice. One testing organization is regularly called upon to sample lots of domestic wool stored in warehouses. For this wool $\frac{\sigma_w}{\sigma_b}$ is about 2.25, and under the conditions of sampling, $\frac{c_1}{c_2}$ is approximately 20. Accord-

TABLE II.—SAMPLING SCHEDULE SHOWING NUMBER OF PACKAGES TO BE SAMPLED FROM A LOT FOR WHICH $\sigma_w = \sigma_b = 2.5$, FOR A PRECISION OF 1.0 AT A PROBABILITY OF 0.95.

Number of packages in the lot	Number of cores per sampled package			
	1	2	4	6
25	25	19	16	15
50	34	25	21	20
75	38	29	24	22
100	40	30	25	24
150	43	33	27	25
200	45	34	28	26
300	47	35	29	27
500	48	36	30	28
1000	49	37	31	29

ingly, the sampling plan followed is to take 10 cores from each package, the number of packages being determined from Eq. 2. A second organization is able to sample its imported carpet wool as it is being trucked into the mill. In this case $\frac{c_1}{c_2}$ is about 1 and $\frac{\sigma_w}{\sigma_b}$ also averages about 1, so that the sampling plan used is one core per sampled bale. The two plans differ considerably for the same precision, yet each is the most economical plan possible under the different conditions faced by the two organizations.

[3] The cost of testing should also be included in deciding on the most economical plan. If each secondary unit is tested individually, c_2 should include the unit test cost. In the case of wool, the cores are composited and a single test is performed on the composite sample.