Journal of Research of the National Bureau of Standards Vol. 46, No. 2, February 1951 Research Paper 2180

A Statistical Solution of a Problem Arising in the Sampling of Leather

John Mandel and Charles W. Mann

The evaluation of the relative merits of different tanning procedures requires that a number of hides or sides [1] of each tannage be subjected to a variety of physical and chemical tests. The destructive nature of these tests makes it desirable to keep to a minimum the number of hides or sides required for test purposes. Moreover, the total number of test specimens for each property should be kept as small as practicable for reasons of economy in leather and labor. Because of the nature of a leather hide, a considerable saving can be realized through a judicious choice of the location on the side from which test specimens for any given property are cut. In this paper a criterion is developed for the evaluation of the suitability of any given side location as a source of test specimens for a given property. It is shown that the coefficient of correlation between the test result on a given location and the average of the test results, on the same property, over the entire side, determines the suitability of this location as sampling location, both from the viewpoint of economy in the number of hides or sides and in the number of specimens required. It is further shown that for any particular property, the number of sides required to detect a given difference between two tanning procedures is inversely proportional to the square of the coefficient of correlation corresponding to the block chosen as sampling location and directly proportional to the square of the coefficient of variation of side averages for the property considered.

I. The Problem

The evaluation of the relative merits of different tanning procedures requires that a number of sides of each tannage be subjected to a variety of physical and chemical tests. It is known, however, that the results of such tests vary appreciably from location to location on the same side, and also, of course, from side to side for the same relative location.

Studies of this variability have been made by Beek [1] [2] and by Beek and Hobbs [2]. Their studies, however, did not include an attempt to define, on the basis of the observed variability, an optimum sampling location on the side. Moreover, their work was limited to a single physical property, namely tensile strength.

For any particular test, the problem naturally arises of finding the "best" location on the side for the selection of a test specimen, or, more generally, of evaluating every location on the side with regard to its suitability for that purpose. As the various physical and chemical tests reflect different properties of the leather, it is reasonable to treat this problem individually for each test.

Information of this type is needed as a basis for sampling leather for research, for specification purposes, and for general testing.

This paper is concerned with the theoretical aspects of the problem of sampling leather. More specifically, the following three parts will be discussed:

1. The planning of an experiment designed to furnish the data, for each of the tests considered, from which the desired information can be extracted. Such an experiment was performed at the National Bureau of Standards.

2. The development of a statistical criterion for the characterization of different positions on the side from the viewpoint of their suitability as test specimen locations.

3. The determination of the number of sides required for the detection, with a prescribed degree of confidence, of variations in the hide properties arising from variations in the tanning procedure.

While the discussion is mainly concerned with the leather problem described above, the statistical development may well apply to a variety of sampling problems in other fields. As this paper deals with principles only, no data will be presented. These will be found, together with the conclusions drawn from them, in a publication that has been submitted to the Journal of the American Leather Chemists Association.

II. The Experimental Plan

Thirty sides were tanned by the same chrome tanning procedure. Each side was then cut into 21 blocks in a rectangular pattern of three rows, parallel to the backbone, and seven columns, perpendicular to the backbone (see fig. 1). Each of the 21 blocks was cut into a number of specimens, such that one or two test specimens of appropriate size

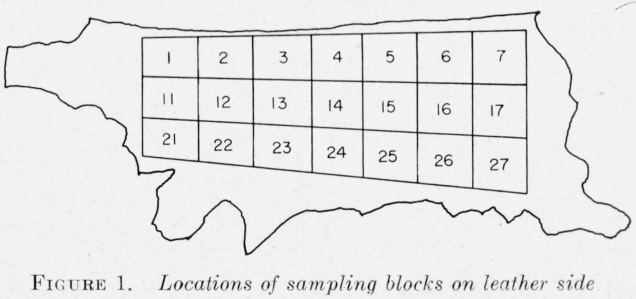

FIGURE 1. *Locations of sampling blocks on leather side*

[1] A side is either one of the halves obtained by cutting a hide along the backbone line.

[2] Figures in brackets indicate the literature references at the end of this paper.

were available for each property. A total of 10 tests was included in the project. Each specimen was subjected to the test for which it was intended. Thus the data represented the results of 10 tests, on each of 21 blocks of 30 sides.

III. Criteria for the Ranking of Locations

The data from each of the 10 tests were treated separately, inasmuch as a location may be "good" for one test and "poor" for another. Thus the following discussion applies to each test individually. The first object consists in defining a good location. One might think that the best location is the one for which the result varies least from side to side. However, a criterion based on such a consideration disregards an essential desideratum, namely that the characteristic, when measured on the selected location, bear some relationship to the relevant value of this characteristic for the entire side. If, in practice, all 21 locations, that is, the entire side, could be tested, it would be reasonable to characterize the side by the average of the 21 results thus obtained.[3] Such an average will be denoted in this paper by h and referred to as a "side average". Then, the average of a set of h-values, corresponding to a random sample of sides from the lot, is a reasonable measure for the value, characterizing the lot, for the test under consideration.

On the other hand, if it is impractical to carry out the test on all 21 positions, so that a single position must be selected, it seems logical to choose the particular position that best represents the average of the side from which it is taken. Let us denote the result of the test on a particular position P as p. Then, according to this principle, for a good position P, p must represent h in the best possible way.

Consequently the first task is to determine, on the basis of the experimental results, what type of relationship exists between p and h, for each of the 21 positions. Such a study was made, using the data obtained in the experiment described in the preceding section, and it revealed the existence of two types of cases. For some positions, a plot of p versus h, for the 30 sides, shows a mere scattering of points without any discernible pattern or trend. For other positions, however, a definite linear trend is apparent.

Figure 2 shows a typical example taken from the stitch-tear data for each of these two cases. The lower plot represents location 22 and shows a mere random scattering of points, while the upper plot, representing location 14, clearly suggests the existence of an approximate straight-line relationship.

It is evident that locations belonging to the first category are undesirable for sampling purposes, as the position value does not represent in any way the quality of the side for the test under consideration. Attention will therefore be focused on the locations belonging to the second category.

[3] Which value is relevant (e. g. average, largest, smallest) depends primarily on the use of the side in which one is interested. In view of the general nature of our problem the relevant value was taken to be the average of the test results of all 21 locations.

It appears intuitively plausible to select for sampling purposes the position of this category for which the scatter of the experimental points on the p versus h graph is smallest. This, however, is not a sufficient criterion for an adequate sampling position; indeed, in order to detect relatively small differences between side values, the corresponding differences between the values for the selected sampling location should be relatively large; that is, the p versus h straight line should be steep.

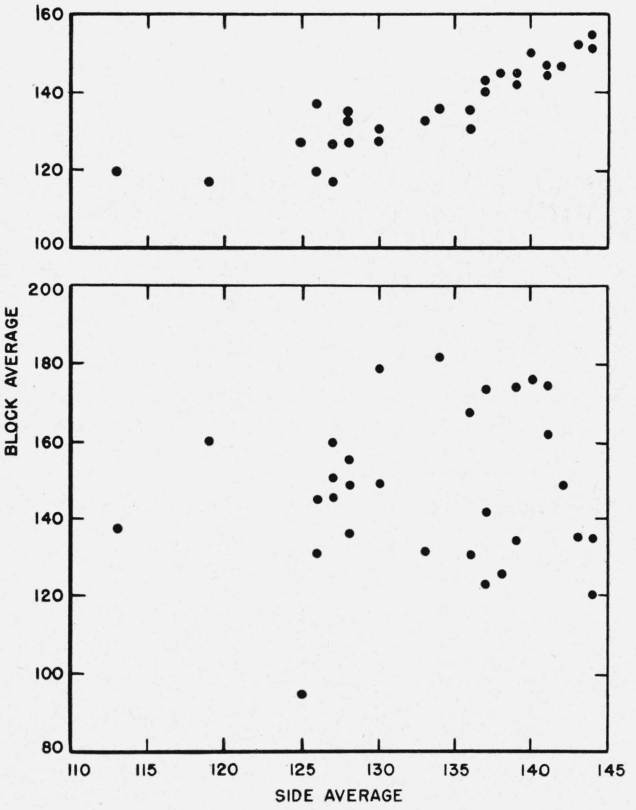

FIGURE 2. *Observed relationship between block average and side average.*

Upper graph, good sampling position: relationship is distinctly linear; lower graph, poor sampling position: no definite relationship exists.

Summarizing, it can be stated that a desirable sampling location P should fulfill the conditions (1) the slope of the p versus h line should be relatively large, and (2) the scatter of the experimental points (h, p) about this line should be small.

IV. Statistical Analysis

In comparing the merits of the various possible sampling locations, the question arises as to how the two criteria formulated in the preceding section should be weighted. For example, the choice between two locations P and P' will appear difficult if P has the higher slope, while P' has the smaller scattering around the p versus h line. It will now be shown that for each physical or chemical test, a single index can be defined for each location that

entirely determines its suitability for sampling purposes. To this end the following definition will be adopted.

A sampling location is said to be better than another if it requires a smaller number of sample sides, for an equal certainty in the final results.

Besides this definition the derivation requires the assumption that the relationship between p and h is linear.

In figure 3, the (p, h) plot is shown for a particular position P. Let p be the value observed, in position P, on a side for which the average value of the characteristic under consideration is h_0. Among all the sides for which the value h happens to be the same (h_0), the values p, in position P, will nevertheless vary, because of biological differences, tanning effects, and test errors. Geometrically, this will result in the fact that p will, in general, not lie exactly on the line, but rather at a variable distance e from it. Likewise, the average h_0 of any particular side will, in general, not be identical with the lot average M. Let us denote by d, the difference between h_0 and M; that is, $d=h_0-M$. The relation between the observed position value p, and the lot average M for which it is the experimental estimate, can be found by noting that the straight line has the equation

$$p_0=b+mh_0$$

where b ls the intercept, m the slope, and p_0 the "theoretical" value of p corresponding to the value h_0, that is the value corresponding to the point on the line.

Now, since $p=p_0+e$, we have

$$p=b+mh_0+e$$

and since $h_0=M+d$, we obtain the desired relation

$$p=b+m\ (M+d)+e$$
$$p=b+mM+(md+e) \qquad (1)$$

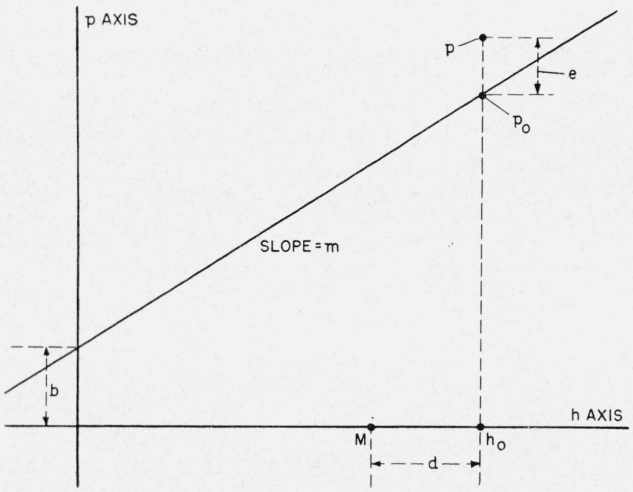

FIGURE 3. *Theoretical relationship between block average, side average, and lot average.*

In this expression, the quantity $(md+e)$ represents the total random fluctuation, analagous to the "error" in the theory of errors. Indeed, if this quantity were known, the value M could be derived from p on the basis of the two parameters b and m. The order of magnitude of this random part is characterized by its standard deviation, which, in accordance with the laws of propagation of errors [3] is given by the relation

$$\sigma=\sqrt{m^2(\sigma_d)^2+(\sigma_e)^2},$$

where σ_d and σ_e are the standard deviations associated with the random fluctuations d and e, respectively.

Suppose now that N sides are taken from a lot, and that the test is carried out in position P of each of these N sides. Then the average of these N test results, which we will denote by $\overline{p}_1$, has a standard deviation equal to $\sigma/\sqrt{N}$, that is

$$\sigma_1=\sqrt{\frac{m^2(\sigma_d)^2+(\sigma_e)^2}{N}}.$$

If, for the purpose of comparing tanning procedures, a sample of N sides is also taken from a second lot, of different tannage, and the test carried out in the same position P on each of these sides, an average $\overline{p}_2$ will be obtained, with a standard deviation:

$$\sigma_2=\sqrt{\frac{m^2(\sigma_d)^2+(\sigma_e)^2}{N}}.$$

It is seen that $\sigma_1=\sigma_2$, provided that the slope m, corresponding to position P, is the same for the two lots and that the fluctuations d and e in the second lot have the same statistical distributions as the corresponding fluctuations in the first lot. Although there is no conclusive evidence for the validity of these assumptions, it seems reasonable to accept them on a tentative basis. It is readily seen that the general principle of the proposed statistical procedure is not dependent on these assumptions, and that situations in which these simplifications do not apply will require only slight modifications in the formulae.

The difference $\overline{p}_1-\overline{p}_2$ constitutes the experimental evidence for any effect of tannage on the test considered. According to the laws of propagation of errors [3], the standard deviation for this difference is given by

$$\sigma_{\overline{p}_1-\overline{p}_2}=\sqrt{\sigma_1^2+\sigma_2^2}=\sqrt{\frac{2}{N}\left[m^2(\sigma_d)^2+(\sigma_e)^2\right]}. \qquad (2)$$

This relation shows that for any observed difference between two lots, the precision of this difference increases (the standard deviation decreases) as the number of sides, N, taken from each lot, increases. From our definition of a "better" sampling location it then follows that the most suitable position is that one for which the quantity $\overline{p}_1-\overline{p}_2$ is known with the greatest relative precision, that is the position for which the ratio of the standard deviation of $\overline{p}_1-\overline{p}_2$

to its expected value, for any given N, is smallest. The problem thus becomes that of finding the position P for which the ratio

$$\frac{\sigma_{\bar{p}_1-\bar{p}_2}}{\text{Expected value of } (\bar{p}_1-\bar{p}_2)}$$

is a minimum.

The numerator of this expression is given by eq 2. In order to calculate the denominator, let us denote the lot averages of the two lots by M_1 and M_2, respectively. Then, in accordance with eq 1, the expected value of $\bar{p}_1-\bar{p}_2$ is equal to $(b+mM_1)-(b+mM_2)=m(M_1-M_2)$.[4] Accordingly, the ratio that must be made a minimum is equal to

$$\frac{\sigma_{\bar{p}_1-\bar{p}_2}}{\text{Expected value of } (\bar{p}_1-\bar{p}_2)}=\frac{\sqrt{\frac{2}{N}\left[m^2(\sigma_d)^2+(\sigma_e)^2\right]}}{m(M_1-M_2)}$$

$$=\frac{\sqrt{\frac{2}{N}}}{M_1-M_2}\sqrt{(\sigma_d)^2+\left(\frac{\sigma_e}{m}\right)^2}. \tag{3}$$

Now, for a given number of sides N, the only quantity of this expression that depends on the position P is $(\sigma_e/m)^2$, as M_1 and M_2 are lot values and, σ_d represents the variability of side averages. The minimum of the expression corresponds to a minimum of the quantity σ_e/m.

It is clear now that the single index σ_e/m is the criterion according to which positions must be classified in terms of their suitability as sampling locations. Remembering that σ_e is a measure of the scatter of the (p, h) points about the straight line, and that m represents the slope of that line, it is seen that the statistical analysis confirms the adequacy of the two criteria formulated in the preceding section, and at the same time provides the manner in which they are to be combined for the selection of the best sampling location.

It can also be observed in the expression just given that the larger the index σ_e/m, the larger will have to be the number of sides N to obtain the same relative precision. Thus, the index σ_e/m actually determines the amount of testing required in order to detect tanning effects with a given degree of confidence.

V. The Coefficient of Correlation as a Criterion

It is interesting to note that the index, σ_e/m, which must be chosen as small as possible, is closely related to the coefficient of correlation ρ between p and h.

For a chosen sampling location P, every pair of p and h values can be considered as a random selection from a population of such pairs, since it corresponds to a side selected at random from a population of sides.[5] If it is assumed that this two-variable population is of the Gaussian type, then the coefficient of correlation ρ between p and h can be expressed by the relation ([4])

$$\rho=m\,\frac{\sigma_h}{\sigma_p},$$

where m has the same meaning as above. The standard deviation σ_h refers to the variability of side averages in the lot, that is, $\sigma_h=\sigma_d$, as defined above. The standard deviation σ_p, on the other hand, refers to the total variability of the p values, that is, $\sigma_p=\sqrt{m^2(\sigma_d)^2+(\sigma_e)^2}$ as shown earlier. Thus,

$$\rho=m\,\frac{\sigma_d}{\sqrt{m^2(\sigma_d)^2+(\sigma_e)^2}}=\frac{\sigma_d}{\sqrt{(\sigma_d)^2+\left(\frac{\sigma_e}{m}\right)^2}}. \tag{4}$$

This relation shows that to a small value of σ_e/m corresponds a large value of ρ. The best sampling location, according to our criterion, is therefore that for which ρ is largest.

VI. Ranking of Sampling Locations

If, for every physical and chemical test, the coefficient ρ were known for every location on the side, there would be a simple solution to the problem of the selection of the most suitable sampling location for each such test. An experiment such as the one described in this paper, based on the testing of 30 sides, permits the calculation of a correlation coefficient for each test and each position. However, the coefficient thus computed, which we will denote by r, is only an estimate for the "true" coefficient ρ, and this estimate will vary from experiment to experiment, because of the chance fluctuations of the p,h points and therefore of the exact location of the p,h line and of the average scatter about this line. Consequently, in comparing positions on the basis of these observed coefficients of correlation, account must be taken of the sampling fluctuations of these estimates, when based on such a limited number of points as 30. This can be done most effectively by grouping the correlation coefficients, for any given property, into a number of groups, such that within each group the differences would not be considered statistically significant, while the differences between successive group averages are significant. One of the difficulties arising in this problem lies in the fact that the precision of an estimated coefficient of correlation depends not only on the sample size (number of points) but also on the value of the corresponding true coefficient. For large values of ρ, the

[4] Indeed, the term $(md+e)$ occurring in eq 1, being a random fluctuation, has an expected value equal to zero.

[5] The randomness in the selection of sides from the lot is essential for the applicability of the method presented in this paper. As has been pointed out by Berkson [5], a coefficient of correlation calculated on a systematically selected sample will depend on the range of values encompassed by the sample, and increase with this range. A coefficient of correlation thus obtained is devoid of any sensible interpretation as a measure of the real degree of interdependence.

precision of r is considerably larger than for small values of ρ. This difficulty can be overcome to a large extent by transforming each r-value into a z-value by means of the transformation ([4])

$$z=\frac{1}{2}\log_e\frac{1+r}{1-r}.$$

It is known that for any value ρ the sample estimates r, when transformed into z-values, have an approximately Gaussian frequency distribution, the mean of which corresponds to the transformed ρ-value and the standard deviation of which is equal to $1/\sqrt{N-3}$, where N is the size of the sample from which r was computed. Since a positive r-value [6] can lie anywhere between 0 and 1, the corresponding z-value can lie anywhere between 0 and ∞. Thus, for any value of ρ, it is theoretically possible to obtain a sample estimate r such that the corresponding z-value will be exceedingly large. However, even for relatively large values of ρ, the sample estimate r will rarely exceed 0.99, and therefore the corresponding z-value will rarely exceed $1/2\log_e(1+0.99)/(1-0.99)=2.65$. Consequently, from a practical viewpoint, this value can be considered as an upper limit for z.

On the other hand, the standard deviation of z, regardless of the value of ρ, is always equal to $1/\sqrt{N-3}$, which for the case of 30 sides, becomes $\sigma=1/\sqrt{30-3}=0.19$.

Consider now, in the range extending from 0 to 2.65, a sequence of z-intervals, S_0, S_1, S_2, . . . the midpoints of which are spaced 2×0.675 standard deviations, that is, $2\times0.675\times0.19$ units apart. Then, in accordance with the properties of the Gaussian distribution assumed to hold for the z's, the probability is 50 percent that a z-value corresponding to a population centered in the midpoint of S_i, will actually fall in the interval S_i. The probability that such a value will fall in S_{i-1} or in S_{i+1}, is, in each case, approximately 23 percent. Thus the probability for such a value to fall in its correct group, or, at worst, in one of the two adjacent groups, is 96 percent. As these probabilities also apply to the corresponding r-values, the z-intervals thus defined can be made the basis of a grouping procedure of correlation coefficients computed on 30 points. To this end it is merely necessary to apply the inverse transformation $r=(e^{2z}-1)/(e^{2z}+1)$ to the end-points of the S-intervals. The values thus obtained will be the endpoints of the groups for the r-values. For the unique determination of the S_1-intervals, the upper limit of S_1 was chosen to be $z=2.65-0.605\sigma$ or $2.65-0.13=2.52$. In this way, an extra interval S_0 is formed extending from $z=2.52$ to 2.65. The probability for a point belonging to the population S_1, to fall in S_0 is 23 percent, and this includes values of the correlation coefficient up to 0.99.

[6] It is unlikely that for any position on the side, the correlation ρ between position value and side average, for any given property, be negative. However, for small positive values of ρ, the sample estimate r will occasionally be negative, due to chance fluctuation. When this occurs, the value of ρ is best taken equal to zero.

The grouping procedure just described is not an ideal solution for the problem outlined at the beginning of this section. In fact, the statistical literature does not seem to contain a satisfactory practical method for a grouping problem of this nature, and it should be noted that any grouping procedure will result in a certain amount of misclassification. The proposed method must be considered as a practical working rule that has some theoretical plausibility.

Table 1 shows the groups into which the r-values were classified, on the basis of the corresponding S-intervals for z. It is noted that the intervals S_0, S_1, S_2, and S_3 were combined to give a single r-group, denoted group zero. This was done because no experimental r-value exceeding 0.94 was found. Furthermore, the lower limit in group 6 has been raised from 0.42 to 0.46, the latter being the 1-percent significance level. It is considered that a value of r that fails to be significantly different from $\rho=0$ at the 1-percent level corresponds to a very poor sampling location.

TABLE 1. *Ranking of sampling location according to r*

z-interval	r-group	Range of r	
		Lower limit	Upper limit
S_0	0	0.988	0.990
S_1	0	.979	.987
S_2	0	.965	.978
S_3	0	.941	.964
S_4	1	.91	.940
S_5	2	.85	.90
S_6	3	.75	.84
S_7	4	.61	.74
S_8	5	.46	.60
S_9	6	Negative [a]	.45

[a] See footnote 6.

VII. Determination of Sample Size

When two lots are compared, with respect to a physical or chemical characteristic, on the basis of a limited number of measurements on each, the conclusions can be affected by two types of error. The effect of chance fluctuations could produce an apparently large difference between the two observed averages, while the real difference between the lots is actually inconsequential. On the other hand, a real difference of practical significance could escape notice if, by the interplay of chance effects, the observed values were sufficiently alike so that one lot would not be considered different from the other. The statistical method of selecting a sample size—in this case, the number of sides N to be tested from each lot—consists in taking predetermined risks with respect to both types of error. (It is obviously impossible to eliminate either type entirely.)[7]

It was shown (eq 3) that the ratio of the standard deviation of $\overline{p}_1-\overline{p}_2$ to its expected value is equal to

$$\frac{\sqrt{\frac{2}{N}}}{M_1-M_2}\sqrt{(\sigma_d)^2+\left(\frac{\sigma_e}{m}\right)^2}.$$

[7] The two types of error in statistical inference are discussed in (6). An example of their use as a basis for the determination of sample sizes is given by Curtiss [7].

In view of the formula derived for the coefficient of correlation ρ (eq 4), this expression can be written $[(\sqrt{2/N})/(M_1-M_2)]\cdot\sigma_d/\rho$
Thus

$$\frac{\sigma_{\overline{p}_1-\overline{p}_2}}{\text{Expected value of } (\overline{p}_1-\overline{p}_2)}=\frac{\sqrt{\frac{2}{N}}}{M_1-M_2}\cdot\frac{\sigma_d}{\rho}.$$

In order to simplify the notation, let us write $\overline{p}_1-\overline{p}_2=\delta$ and denote the expected value of any quantity Y by the symbol $E(Y)$.
Then

$$\frac{\sigma_\delta}{E(\delta)}=\frac{\sqrt{\frac{2}{N}}\,\sigma_d}{\rho(M_1-M_2)}.$$

Taking reciprocals, we have

$$\frac{E(\delta)}{\sigma_\delta}=\frac{\rho(M_1-M_2)\sqrt{\frac{N}{2}}}{\sigma_d}. \tag{5}$$

Now, the observed difference δ consists of two portions: a fixed portion, equal to $E(\delta)$; and a variable part, due to chance fluctuations, which we will, for brevity, call "error of δ", and denote by the symbol $\epsilon(\delta)$. Thus

$$\delta=E(\delta)+\epsilon(\delta).$$

Expression 5 thus becomes

$$\frac{\delta-\epsilon(\delta)}{\sigma_\delta}=\frac{\rho(M_1-M_2)\sqrt{\frac{N}{2}}}{\sigma_d}. \tag{6}$$

The values for the risks corresponding to the two types of error mentioned earlier can now be assigned in the following manner:

1. The risk of inferring the existence of a difference between the two lots where there is actually none.

Make $M_1=M_2$ in eq 6 in order to express the fact that no actual difference exists. Then the equation becomes

$$\frac{\delta}{\sigma_\delta}=\frac{\epsilon(\delta)}{\sigma_\delta},$$

that is, the ratio of a chance fluctuation to its standard deviation. It is known that for Gaussian variables the absolute value of this ratio will exceed the value 1.96 only five times in a hundred. (Cf. any table giving the "areas", i. e. the cumulative frequencies of the "normal" curve, e. g. [4]). Thus if the risk in question is to be kept at the 5-percent level, the existence of a difference should only be inferred whenever

$$\frac{|\delta|}{\sigma_\delta}>1.96.$$

2. The risk of not detecting a real difference between the two lots.

Suppose that this real difference equals D. Then $M_1-M_2=D$, and eq 6 becomes

$$\frac{\delta}{\sigma_\delta}=\frac{\rho D\sqrt{\frac{N}{2}}}{\sigma_d}+\frac{\epsilon(\delta)}{\sigma_\delta}.$$

Now, in order to keep the risk of the first type at 5 percent, a difference between the two lots was considered to exist only when

$$\frac{|\delta|}{\sigma_\delta}>1.96,$$

that is, whenever

$$\left|\frac{\rho D}{\sigma_d}\sqrt{\frac{N}{2}}+\frac{\epsilon(\delta)}{\sigma_\delta}\right|>1.96,$$

that is, whenever the first member is either smaller than -1.96 or greater than $+1.96$. The first possibility is remote, as it would require that the standardized error term be considerably smaller than -1.96.

Therefore, in practice, the rule will result in detecting the existence of D whenever

$$\frac{\epsilon(\delta)}{\sigma_\delta}>1.96-\frac{\rho D}{\sigma_d}\sqrt{\frac{N}{2}},$$

and consequently in committing the error of failing to recognize its existence whenever

$$\frac{\epsilon(\delta)}{\sigma_\delta}<1.96-\frac{\rho D}{\sigma_d}\sqrt{\frac{N}{2}}.$$

As in the case of the error of the first type, the probability that this inequality will hold can be given any preassigned value, by using the table of areas of the normal curve. For example, if the second member is made equal to -1.64, (the value that has 5 percent of the area under the curve to its left), this inequality will hold five times in a hundred, so that the risk for the error of considering D as negligible is limited to 5 percent. Then

$$1.96-\frac{\rho D}{\sigma_d}\sqrt{\frac{N}{2}}=-1.64,$$

hence

$$N=2\,\frac{(1.64+1.96)^2}{\rho^2D^2}(\sigma_d)^2. \tag{7}$$

This procedure illustrates the general method of determining sample size on the basis of predetermined risks. By changing the numerical values inside the parentheses, the risks can be changed to any desired values.

It should be noted that N is inversely proportional to the square of the coefficient of correlation, a

relation that permits the comparison of different sampling locations from the viewpoint of economy in sampling. Furthermore, the formula shows the rapid increase in the required number of sides as the difference D to be detected becomes smaller.

By dividing in eq 7 both $(\sigma_d)^2$ and D^2 by the square of the average of the lot means M_1 and M_2, it is seen that the number of sides N is proportional, for a given *percentage* difference D to be detected, to the square of the coefficient of variation of side averages, $\sigma_d/\frac{1}{2}(M_1+M_2)$.

VIII. Conclusion

A rational solution to the problem of selecting the best sampling location, for any particular property, on a leather side has been obtained by theoretical considerations combined with facts derived from data obtained in a statistically designed experiment. This statistical method, which may result in considerable savings both in labor and in material, can readily be applied to other situations in which the properties to be measured vary with location on the sampling unit in a systematic way.

IX. References

[1] J. Beek, Jr., The probable error in the measurement of tensile strength of heavy leather, J. Am. Leather Chem. Assoc. **32**, 4 (1937).

[2] J. Beek, Jr. and R. B. Hobbs, Some applications of statistical methods to sampling of leather, J. Am. Leather Chem. Assoc. **36**, 190 (1941).

[3] A. G. Worthing and J. Geffner, Treatment of experimental data (John Wiley and Sons, Inc., New York, N. Y., 1943).

[4] G. W. Snedecor, Statistical methods (The Iowa State College Press, Ames, Iowa, 1946).

[5] Joseph Berkson, Are there two regressions?, J. Am. Statistical Assoc. **45,** 164 (1950).

[6] A. M. Mood, Introduction to the theory of statistics (McGraw-Hill Book Co., Inc., New York, N. Y., 1950).

[7] J. H. Curtiss, Acceptance sampling by variables, J. Research, NBS **39,** 271 (1947) RP 1827.

Washington, July 27, 1950.