# Packing House Sampling high degree of accuracy possible by use of special machine and calculation 

Roy J. Smith

The use of sampling methods in fruit and vegetable marketing and processing operations to determine the type of product which is received from each grower is becoming a common practice.

There is a type of sampling that may be considered as ideal even though it does contain a degree of error, but the error is the least possible in sampling and is calculable-which is not possible with some sampling. practices.
The error which may be anticipated in this representative sampling is indicated in the accompanying table. The error shown is the per cent of difference to be found one time out of three between the value indicated by the sample and the true value for the whole lot. This is: The sample usually varies a little in its proportion of fruit classes from the proportions to be found in the whole lot. Consequently when each class is given a price and the value per box calculated, the value based on the sample differs a little from the value based on the whole lot.

The figures in the table show how great these differences may be. The error is shown in terms of per cent of value of the main lot because value constitutes the grower's real interest in the matter. Also, the use of a percentage figure may be applied to the value of either an average box of fruit or of the whole lot of fruit. The error is that which is due to sampling, not that which may be due to grading or to other practices. It cannot be eliminated or reduced below these per cents.

The meaning of the table may be illustrated in more detail as follows: If a $1 \%$ sample is taken of 25,000 boxes of fruit, an error equal to or greater than 0.13 of $1 \%$ of the value of the fruit may be expected one time out of three. Two times out of three the error will be less. Again: If a $4 \%$ sample is taken of 100 boxes, an error equal to or greater than $1.08 \%$ of the value of the fruit will occur one time out of three. Two times out of three the error will be less.
An error twice that shown will occur one time out of about 20 . Nineteen times out of 20 it will be less. Also: An error three times that shown will occur one time out of about 360 .

Detailed tests made on a lemon sampling machine designed by the Division of Agricultural Engineering, University of California, and installed at a citrus fruit association in Santa Paula, have proved
that the above figures are accurate. For instance, 25 boxes of fruit were run over the sampler 83 times. Each size and grade of fruit was painted with a specific color so that a fruit would always be put in the same size and grade. Thus, no grading or sizing error was allowed.

## Formula Machine Proved

The average sample was $4.7 \%$ of the lot. According to formula the error to be expected or exceeded one time out of three was $2 \%$ of the true value of the fruit. The test showed an error one time out of three of $2.05 \%$, virtually identical with that calculated by formula.

In another test 125 boxes were run 55 times. Again each lemon was painted so that identification by size and grade was always exact. The average sample was $1.28 \%$ of the lot. According to formula the per cent of error to be expected or exceeded one time out of three was 1.71. The test showed an error of 1.92. Considering the volume of fruit to be handled and the need for a greater number of runs in order to attain average results, these results can also be regarded as evidence of reliability.

Any error shown in representative samples is far less important than the im-


Pifof model sampling machine in operation on test runs with citrus fruit of various colors.
provement which will result in having grower credits based on the judgment of one or two highly skilled operators in place of the large crews now relied upon where traditional methods are used.

## Representative Sampling

The sample must be representative. It must be made up of the very smallest seg. ments possible. Each segment, for instance, must not be a box of fruit but one or a very few fruits. The segments must be distributed equally throughout the lot. The taking of the sample must not be selective as to the characteristics of the fruit.

The spread of values among the various classes of fruit must be close to that usedthe proportions and credits of which are shown in the second accompanying table.

The many possible combinations of proportions and prices make possible widely different values for the various classes of fruit. If the spread of values

Continued on page 14

Per cent of error in value of fruit which will be exceeded one time out of three for various percentages of lot taken as sample for various sizes of lots.

| Per cent of lot in sample |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| No. Fd. oxs. in lot | 0.1 | 0.5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 |
|  | \% | \% | \% | \% | \% | \% | \% | \% | \% | \% | \% |
| 100,000 | 0.22 | 0.10 | 0.07 | 0.05 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | . 02 |
| 50,000 | 0.31 | 0.14 | 0.10 | 0.07 | 0.06 | 0.05 | 0.04 | 0.04 | 0.04 | 0.03 | 0.03 |
| 25,000 | 0.44 | 0.19 | 0.13 | 0.10 | 0.08 | 0.07 | 0.06 | 0.05 | 0.05 | 0.05 | 0.04 |
| 10,000 | 0.70 | 0.32 | 0.21 | 0.15 | 0.12 | 0.10 | 0.10 | 0.09 | 0.08 | 0.07 | 0.07 |
| 5,00 | 0.99 | 0.44 | 0.31 | 0.22 | 0.18 | 0.15 | 0.14 | 0.12 | 0.11 | 0.11 | 0.09 |
| 2,500 | 1.39 | 0.62 | 0.44 | 0.32 | 0.25 | 0.22 | 0.19 | 0.18 | 0.16 | 0.15 | 0.13 |
| 1,000 | 2.22 | -0.99 | 0.70 | 0.49 | 0.40 | 0.34 | 0.30 | 0.28 | 0.25 | 0.24 | 0.21 |
| 750 | 2.56 | 1.14 | 0.81 | 0.56 | 0.47 | 0.39 | 0.35 | 0.32 | 0.29 | 0.27 | 0.24 |
| 500 | 3.14 | 1. | -0.98 | 0.69 | 0.56 | 0.48 | 0.43 | 0.39 | 0.36 | 0.34 | 0.29 |
| 250 | 4.44 | 1.97 | 1.38 | -0.98 | 0.80 | 0.68 | 0.62 | 0.55 | 0.51 | 0.48 | 0.42 |
| 100 | 7.02 | 3.20 | 2.21 | 1.55 | 1.26 | 1.08 | -0.96 | _0.87 | 0.81 | 0.72 | 0.66 |
| 50 | 9.92 | 4.27 | 3.12 | 2.19 | 1.78 | 1.52 | 1.37 | 1.24 | 1.14 | 1.06 | 0.94 |
| 25 | 14.05 | 6.27 | 4.41 | 3.18 | 2.50 | 2.17 | 1.93 | 1.75 | 1.61 | 1.49 | 1.32 |
| 10 | 22.21 | 9.91 | 6.99 | 4.91 | 4.00 | 3.42 | 3.05 | 2.77 | 2.55 | 2.37 | 2.10 |

MOSAIC

## Continued from page 6

In the tests at the Experiment Station the original sugar-beet-mosaic virus was obtained from a field of naturally infected sugar beets near San Pablo. Mechanical inoculation of healthy sugar beets grown under cover in the greenhouse was carried out to obtain a virus supply and this was maintained by continuous inoculations during the experiments.

The green peach aphid was used in most tests. Noninfective aphid were obtained by transferring mature, wingless aphids from populations collected in the field to favorable healthy host plants.

On the following day the offspring from the mature aphids were transferred to a second healthy plant and allowed to multiply. No symptoms appeared and the disease was not produced in any case.

In one instance aphids recovered the virus from a sugar beet infected with
the virus one day before symptoms of the disease developed; in another instance, on the same day after the first symptom appeared; while in still others, one to two days after the earliest symptom developed.

Virus transmission by lots of 20 erigeron root aphid, pea aphid and the green peach aphid reared on mosaic beets were compared with that by mechanical inoculation.

Infections obtained with these three aphid species were $20 \%, 60 \%$ and $56 \%$ respectively, as compared with $96 \%$ by mechanical inoculation of the virus extract from the plants on which they were reared.

The transmission of the virus by 10 aphid species reared on other host plants varied from $8 \%$ to $76 \%$, as compared with $88 \%$ to $100 \%$ by mechanical inoculation of juice expressed from the same mosaic beets on which the aphids were forced to feed.

With the green peach aphid, the percentages of infections produced were observed to increase with the number of aphids per plant.

Short feeding time of winged aphids on mosaic and healthy beets may be of significance in the natural spread of the disease, since lots of one, two, three, four and five green peach aphids gave infections averaging $0 \%, 25 \%, 25 \%, 40 \%$ and $45 \%$, respectively, after having fed five minutes on mosaic and five minutes on healthy beets.

The retention of the virus by lots of 20 infective aphids varied from one to three hours under greenhouse conditions.

[^0]The above progress report is based upon Research Project No. 657.

## SAMPLING

## Continued from page 7

with which a house is concerned is greater than used here, the per cent of error to be expected will be greater than that shown in the table. If the spread of values is less, the error will be less. To illustrate: If the prices used were lower as, say, $\$ 1.40, \$ 1.26, \$ 0.70, \$ 1.00, \$ 0.90, \$ 0.50$ and $\$ 0.40$, for the successive classes respectively, the per cent of error for 1,000 boxes with a $1 \%$ sample would be $0.62 \%$ as compared with $0.70 \%$ in the table. Since all prices are lower, the proportionate spread of values is less and the per cent of error to be expected is less.

If the prices of only the last two classes of fruit were lower, say to $\$ 0.30$ and $\$ 0.20$ respectively, the per cent of error in our example would be $0.81 \%$. It is greater than the $0.70 \%$ error shown in the table below because the lower value in the two classes increased the proportionate
spread of value among the various classes.
These two illustrations represent about the usual variations to be expected by lemon houses because of variations in proportions of fruit in various classes and because of variations in prices. In practical application the difference may be ignored.

## Application For Accuracy

A given degree of error in the table slopes downward from the left to the right. This characteristic is shown by the sloping line of dashes in the table, which represents the approximate position of an error of $1 \%$. It is clear that a very low proportion of fruit may be taken as the sample for large lots, but that a much greater proportion is required in the sample for smaller lots. Roughly, a $0.1 \%$ sample is adequate for 5,000 or more boxes if an accuracy of $1 \%$ two times out of three is desired.
$\left.\begin{array}{|llllll|}\hline & & \text { Description of fruit by classes, proportions and credits } \\ \text { used in caculating data. }\end{array}\right]$

A similar accuracy is obtained for a lot of 500 boxes if a $1 \%$ sample is taken, or for a lot of 100 boxes if a $4 \%$ sample is taken. If comparable accuracy is to be obtained for all growers, regardless of size of lot, it would appear necessary for a packing house to take varying percentages from different sized lots. The smaller the lot the larger the percentage needed for a sample.

There seems to be little reason for taking more than about $4 \%$ of a lot for a sample. Beyond that percentage the increase in accuracy is very slight. To go beyond $4 \%$ would add greatly to the machine and labor requirements. That a material decrease in accuracy would occur for very small lots, as say 25 boxes, can be disregarded because under present practices in marketing and processing the error that exists in the credits given for such small lots is far greater than that shown here:

While more information is required than is now available to make a final decision, it seems likely that there will be little reason to take less than a $0.5 \%$ sample. Moreover, it would seem that labor which is employed in handling the sample should be fully employed at the task which is a rigorous, exact one and should have the operator's undeviating attention to assure accuracy.

[^1]
[^0]:    Henry H. P. Severin is Entomologist in the Experiment Station, Berkeley.

    Roger M. Drake, who was a graduate student in Entomology and Parasitology at the time of these studies, assisted with the research reported above.

[^1]:    Roy J. Smith is Associate Professor of Agricultural Economics, Associate Agricultural Economist in the Experiment Station and on the Giannini Foundation, Los Angeles.
    Mathematical formulae for determining the per cent of error were developed by Prof. G. M. Kuznets, Division of Agricultural Economics, College of Agriculture, and Prof. P. G. Hoel, Department of Mathematics, Los Angeles.

    The above progress report is based upon Research Project No. 1331.

