The Philippine Statistician

Vol. xxvi, Nos. 1 & 2 1977

USE OF STATISTICS IN AGRICULTURAL RESEARCH AT THE ATOMIC RESEARCH CENTER, PAEC

IBARRA S. SANTOS*
Nuclear Research Supervisor
Agricultural Research Division

To play it safe, I am limiting this paper to a short narration on the use of statistics in *nuclear agricultural research* rather than to all of nuclear research. However, in doing so I might actually cover most of the statistics used in nuclear research. Actually, statistical methods needed in nuclear research are almost identical with those used in non-nuclear research. Also, agricultural research because of its very nature has to use statistical methods more than chemical research, for example.

One big difference between *conventional* agricultural research and *nuclear* agricultural research is the involvement of the latter with radioisotopes. In fertilizer experiments, this means taking data on the activity of the sample in *counts per minutes* which proportionately reflects the number of atomic disintegrations or the number of radioactive atoms that decay. There are several instruments used in counting; unfortunately, I am not familiar with the uses, characteristics, advantages and disadvantages of each. For our purpose, it is enough to realize that an atom either *disintegrates* or *does not disintegrate*. Hence, radioactivity should follow a binomial distribution. The probability (P) of any set of x atoms disintegrating is given in the following equation:

$$P(x) = \frac{N!}{x!\,(N-x)!}\, p^{x}\, q^{N-x}$$

where N = total number of atoms
q = probability of the remaining (N-x) atoms not disintegrating.

* Nuclear Research Supervisor, Agricultural Research Division, PAEC

The mean of the bionomial distribution is Np; the variance is Npq.

If p is small and N tends to infinity, the binomial distribution is approximated by the Poisson distribution given as follows:

$$P(x) = \frac{m^x}{X!} e^{-m}$$

where

m = the mean rate of disintegration of x atoms

In the Poisson distribution, m = Np = variance. The decay of radioactive atoms which is *completely random* and *independent* of any other event, temperature and pressure is very well described by the Poisson distribution. As a consequence, a single count may be taken as an estimate of the mean, provided counting time is sufficiently long.

The Poisson distribution is often cumbersome to use and is, therefore, literally poison to simple-minded people like us who are not statisticians. Thus, it is always tempting to assume that m is large enough to warrant the shift to our beloved normal distribution (mathematically described by the Gaussian distribution). Pearson's chi-square test may be used to determine the significance of the disagreement between the standard deviation computed assuming a Poisson distribution and that computed assuming a normal distribution. The ideal value of P from the chi-square test is 0.5 but a range of 0.1 to 0.9 is acceptable. If P is less than .02 or greater than 0.98, something must have gone wrong. It is therefore advisable to repeat the measurements or to examine the counting apparatus for possible malfunction.

Certain experimental values obtained may be rejected if they differ from the mean by three standard deviations. An alternative procedure in rejecting values is provided by *Chauvenet's* criterion which I will not elaborate on here.

Assuming a normal distribution, one can determine the number of counts that should be attained to obtain a certain level of accuracy by using the following formula:

$$\bar{x} = \left(\frac{1}{S_r}\right)^2$$

where, $\bar{x}$ = counts needed to obtain a relative deviation, S_r, which must be specified, for example, as 1% or 0.01.

Relative deviation, S_r, may be computed from experimental data as follows:

$$S_r = \frac{S}{\bar{x}}$$

where, S = standard deviation
x = mean of the experiment

Relative deviation, obviously, is the coefficient of variation expressed in decimals, rather than percentage.

In radioactivity counting, two variables are actually involved. These are (1) the *sample* count and (2) the *background* count. The background count originates from the existing natural radioactivity in the environment. This is a variable by itself rather than a source of error such as soil heterogeneity. In other words, each count is the sum of the background count (x_b) and the sample count (x_s), or $x_{s+b} = x_s + x_b$ and $\bar{x}_{s+b} = \bar{x}_s + \bar{x}_b$. The background count can be obtained directly with the use of the counting instrument while the sample count can obviously be obtained by difference.

Let us recall that these measurements follow a Poisson distribution; therefore, population mean M = population variance σ^2, or sample mean, $\bar{x}$ = sample variance, s^2 and $\bar{x} \pm \sqrt{\bar{x}}$ is the sample mean plus or minus the standard deviation. Each *count* or *mean count* should be corrected for background radioactivity.

The one-standard deviation counting range of sample and background count is given as:

$$R_{s+b} \pm \sqrt{s^2_{s+b}} = \frac{\bar{x}_{s+b}}{t_{s+b}} \pm \sqrt{\frac{\bar{x}_{s+b}}{t_{s+b}}} = \frac{\bar{x}_{s+b}}{t_{s+b}} \pm \sqrt{\frac{\bar{x}_{s+b}}{t^2_{s+b}}}$$

where,

R_{s+b} = uncorrected counting rate (count per unit time)
$\bar{x}_{s+b}$ = uncorrected count
t_{s+b} = time during which count (sample plus background) was taken

$$R_b \pm \sqrt{s_b^2} = \frac{\bar{x}_b}{t_b} \pm \sqrt{\frac{\bar{x}_b}{t_b}} = \frac{\bar{x}_b}{t_b} \pm \sqrt{\frac{\bar{x}_b}{t_b^2}}$$

where,

$\bar{x}_b$ = background

t_b = time during which background count was taken

and,

$$R_s \pm \sqrt{s_s^2} = \left[\frac{\bar{x}_{s+b}}{t_{s+b}} - \frac{\bar{x}_b}{t_b}\right] \pm \sqrt{\frac{\bar{x}_{s+b}}{t_{s+b}^2} + \frac{\bar{x}_b}{t_b^2}}$$

where,

R_s = sample counting rate

It is generally difficult to measure correctly a sample counting rate that is less than or equal to the background counting rate unless the latter is unduly large. Experiments should be planned to obtain a sample counting rate that is at least about ten times that of the background counting rate. The *times* spent in determining the background counting rate and the total (sample plus background) counting rate need not be the same. Ideally, the relative amount of time spent in counting is in inverse proportion to the square root of the counting rates, as indicated in the following equation:

$$\frac{t_{s+b}}{} = \frac{\sqrt{R_b}}{\sqrt{R_{s+b}}}$$

Generally, since background activitiy is much lower than sample activity, a longer time is needed in determining the background counting rate to obtain the same degree of precision.

At the risk of giving unnecessary details, I represented the preceding account of the *statistics of accounting* because in no other field of research or endeavor is *radioactivity* involved.

Let me now go into other aspects of nuclear agricultural research. As already mentioned earlier, there is hardly any difference between conventional agricultural research and nuclear agricultural research when it comes to the use of statistics. In greenhouse experiments, one still has the usual choice between the completely randomized design and randomized complete-block design, depending on the absence or presence of an environmental (light, wind, etc.) gradient. In most field experiments, the usual choice is the randomized

complete-block design. When there are only a few and the environmental gradient in the field goes into two directions perpendicular to each other, the Latin square design should be used although not many agricultural researchers take advantage of the usefulness of this design under those given conditions.

In nuclear entomology and food irradiation researches, either the completely randomized design or the randomized complete-block design is commonly used. The only notable thing in these two fields of research is the use of data transformation prior to the analysis of variance due to (a) the presence of several *O* data, (b) percentages ranging from 0 to 100 (or outside the 30 to 70% range) and (c) large differences in the "plot" values. The beloved Gaussian or normal distribution is also utilized in nuclear entomology instead of either the binomial or the Poisson distribution inspite of the fact that the variates are not continuous; that is, an egg either hatches or does not hatch or a larva, pupa or adult insect either dies or survives after having received a certain dose of ionizing radiation. Similarly, in food irradiation fruit either rots or does not rot wthin a given period of time.

Lattices and other incomplete designs are difficult to handle. Data analysis is difficult when there are several missing plots. Hence, the aversion of most agricultural researchers to incomplete block designs is very understandable.

However, there is one incomplete block design (Galiant and Everett, 1949) very similar to a 3 x 3 balanced lattice design which we have been using in our eating-quality tests at the Agricultural Research Division during the last eight years. In our mutation breeding, we are most particular about ending up with a food variety that is excellent in eating quality, even if the yield potential is not the best available but good enough. This is in contrast to the breeding philosophy in other research institutions concerned with plant breeding. This, I believe, is a healthy and desirable divergence for then the farmers will have a wider choice of varieties to plant.

One would not realize the utility of this design in eating quality tests unless one has actually tried comparing and rating more than three items at a time. To rank accurately five items at one time is an impossible task to me. I can say this because with only three test items to rank first, second and third I had to take at least ten minutes. I, therefore,

believe very strongly that not more than three items should be served to a taster at any given time.

The above-mentioned design has that feature. With only nine entries (one or more entries may be duplicated to make nine entries if less than nine entries are available), each entry is compared with two other entries in four incomplete blocks. There are twelve such incomplete blocks in each experiment. Each incomplete block is repeated four times. A minimum of four tasters is therefore necessary. The same four tasters or any number of tasters between four and forty-eight may be used to carry out the test on all twelve incomplete blocks. If 48 tasters are available, each of the twelve incomplete blocks is evaluated by a different set of four tasters.

Each entry is evaluated 16 times in the experiment since each of the four incomplete blocks to which it is included is given for evaluation to four different tasters. A ranking of first earns an entry a score of 0.85; a ranking of second, 0; and a ranking of third, a minus 0.85. No ties are allowed even if two entries happen to taste exactly the same. The highest possible score is 0.85 x 16 = 13.60; the highest possible negative score is -13.60. The minus entries may then be arranged in an array from the highest positive score down to the highest negative score. A just significant difference (j.s.d.) may be computed and applied to determine significance of differences. In practice though, hardly any difference even between the best and the worst-tasting entries turns out to be significant even though one can swear to the difference in eating quality between those two entries. Hence, for practical purposes the array itself and the total score of each entry are very useful in comparing relative eating of the 9 entries. The inclusion of a standard (or control) variety in the 9-entry test contributes tremendously to the usefulness of the test. Also, the size of the j.s.d. gives us an indication of whether or not something went wrong with the conduct of the test or of the statistical analysis.

I now come to an example to show how statistics is used in conventional breeding or genetics and in mutation breeding. A conventional breeder makes a cross between two parents with contrasting characters governed by a single pair of genes. He is interested in recovering the recessive trait (probably disease resistance) in the second generation. The probability of any one plant being of that recessive trait is 1/4. The probability of any one plant *not* being of that recessive trait

is 3/4. If he saves only one plant from the second generation population the probability of failure is 3/4 or 0.75. If he saves two plants, the probability of failure is $(3/4)^2$ or 0.5625; three plants, $(3/4)^3$ or 0.4219; four plants, $(3/4)^4$, or 0.3164; five plants, 0.2373; six plants, 0.1780; and ten plants, .0561. Probability of failure, of course, goes down as the number of plants saved is increased. One has to save about eleven plants to have a probability of failure of less than .05, our most worshipped probability level.

In the induction of a particular mutation desired, the estimate of success may be one in 1,000; one in 10,000; one in 100,000 or may be one in 1,000,000. Assuming it is one in 1,000 one has a probability of failure of (999/1,000) 10,000 = .000435 if he grows 10,000 plants. This probability of failure is very low and if one is willing to have a probability of failure of P = 0.10 then N or the number of plants to be grown can be computed as follows:

$$P = \left(\frac{999}{1000}\right)^N = 0.10$$

$$N \log 0.999 = \log 0.10$$

$$N = \frac{\log .10}{\log 0.999} = 229$$

However, if one is looking for a *particular mutation* the probability of success is closer to one in 100,000 than to one in 1,000 in most cases. Therefore, the number of second-generation plants needed to be grown to have a good chance of obtaining a desired mutation is usually very large. The mutation breeder just grows the number (about 10,000) he can afford to grow and then keeps his fingers crossed. However, the single-cell approach is now being put to use to increase the probability of recovering mutations and hence, the probability of recovering the mutation breeder will need the use of a field with uniform soil to enable him to identify certain types of mutations.

Statistics as applied in agricultural research including nuclear agricultural research, is a useful tool. For one thing it impresses on us the importance of *randomization* if conclusions are to be valid. For another thing, it keeps reminding us of the concept of *probability*. We should, however, not overlook the importance of the *control value* in evaluating re-

sults of experiments. Also, since many assumptions underlying most statistical procedures are not fulfilled, we should never surrender completely our common sense and intuition (most probably based on experience) in interpreting results of experiments.

BIBLIOGRAPHY

Kempthorne, Oscar. 1957. An Introduction to Genetic Statistics. John Wiley and Sons, Inc., New York and London. XVII + 545 pp.

Galiant, W.C. and H.L. Everett. 1949. A technique for testing flavor of sweet corn. Agron. Jour. 41: 443-445

Bantegui, Celia G. and L.R. de la Paz. 1975. Statistics of counting. In: Lectures on Nuclear Science for High School Teachers. First Revision. Published by the Philippine Atomic Energy Commission, Quezon City, Metro Manila, Philippines. 434 pp.

Gomez, Arturo A. 1961. Performance and eating-quality of sweet corn hybrids. Master's thesis, University of the Philippines College of Agriculture, College, Laguna, Philippines.

Santos, I.S., A.B. Asencion, L.M.V. Brewbaker and C.T. Villegas. Improvement of aromatic Milagrosa rice variety through mutation breeding. (manuscript ready for publication).