A CONDITIONAL TEST CRITERION FOR JUDGING INTERNAL SPREAD OF DISEASE AMONG PLANTS

By

N. Sundararaj And H.S. Rajeshwari

University of Agricultural Sciences, Bangalore.

(Received : May, 1979)

Summary

A statistical criterion for judging internal spread of disease among plants was proposed earlier by working out the expected number of doublets based on the assumption of randomness in the incidense of disease and constancy of the probability that a plant exposed to the risk would be attacked. In the present paper this result has been extended under a few additional features. The conditional test criterion based on the number of doublets from all the *M* plots has been proposed.

Introduction

Van der Plank (1960) investigated a problem of great importance to ascertain how infection would spread from plant to plant ; and more specifically if such a spread was from 'within' the crop or from atmosphere. In the case of spread from 'within', this tendency is revealed by a nest of diseased plants around an infected plant. Van der Plank proposed a measure of this tendency by counting the number of 'doublets'—a doublet consisting of two adjacent diseased plants ; a run of three diseased plants counted as two doublets and so on. In an earlier paper (Sundararaj 1976) a statistical criterion for judging internal spread of disease among plants was proposed by working out the expected number of doublets based on the assumption of randomness in the incidence of the disease and constancy of the probability that a plant, exposed to the risk, would be attacked.

The present paper extends this result under a few additional features, *viz.*,

(*a*) Every plant has the same probability of being attacked by the disease, the incidence being random in nature (*b*) In a large field, a random sample of M plots is selected, each plot containing exactly n plants @

(*c*) In a randomly selected plot, the observed number d of diseased plants is a random variable having a Poisson distribution with parameter λ.

In this paper we propose a conditional test criterion based on the number of doublets from all the M plots.

Some Preliminary Results

In this section we shall derive a few prelimiminary results for later use.

(*i*) *Conditional Joint Distribution of Number of Diseased Plants.* Under the assumption (c) above, the observed number of diseased plants is a Poisson variate with parameter λ.

When d_i, the number of diseased plants in the i^{th} plot ($i=1, 2 \ldots M$) are independently and identically distributed Poisson variates with common parameter λ, it may be shown that the joint distribution of d_i, given $\sum_1^M d_i = t$ will be a multinomial distribution with parameters $\left(t, \frac{1}{M}\right)$. That is

$$P\left[d_1, d_2, \ldots d_M \;\Big|\; \sum_1^M d_i = t\right] = \frac{\underline{|t}\; M^{-t}}{\underline{|d_1} \cdot \underline{|d_2} \ldots \underline{|d_M}} \qquad \ldots(2.1)$$

Writing π for $\frac{1}{M}$, the moments of d_i ($i=1, 2 \ldots M$) about the origin are (Johnson & Kotz 1969)

$$\left.\begin{aligned}
E(d_i) &= t\pi \\
E(d_i^2) &= t\pi + t(t-1)\,\pi^2 \\
E(d_i^3) &= t\pi + 3t(t-1)\,\pi^2 + t(t-1)(t-2)\,\pi^3 \\
E(d_i^4) &= t\pi + 7t(t-1)\,\pi^2 + 6t(t-1)(t-2)\pi^3 \\
&\quad + t(t-1)(t-2)(t-3)\pi^4 \\
E(d_i^{(2)} \cdot d_j^{(2)}) &= t(t-1)(t-2)(t-3)\pi^4 \\
\text{with } d_i^{(2)} &= d_i(d_i-1),\ d_j^{(2)} = d_j(d_j-1)\ i \neq j
\end{aligned}\right\} \qquad \ldots(2.2)$$

@The assumption of n being the same from plot is made for to plot mathematical simplicity.

(ii) The Moments for the Number of Doublets

In an earlier paper (Sundararaj 1976) the conditional mean and the conditional variance for the number of doublets (X) in a random arrangement of d diseased plants and $(n-d)$ healthy plants were shown to be :

$$\text{Mean :} \quad E(X/d)=\frac{d(d-1)}{n}$$

$$\text{Variance :} \quad V(X/d)=\frac{d(d-1)\,(n-d)\,(n-d+1)}{n^2(n-1)}$$

Since the number of diseased plants could very from plot to plot it is useful to obtain the unconditional moments for the number of doublets X_i, independent of individual d_iS' $(i=1, 2 \ldots M)$. Denoting the three moments—mean, variance and covariance—by $E(X_i)$, $V(X_i)$, Cov (X_i, X_j) $(i \neq j)$, they are easily obtained by making use of the moments given in (2.2) and the following relationships $(i, j=1, 2 \ldots M\,;\, i \neq j)$:

$$E(X_i)=\underset{d_i}{E}\,[E(X_i \mid d_i)]$$

$$V(X_i)=\underset{d_i}{V}\,[E(X_i \mid d_i)]+\underset{d_i}{E}\,[V(X_i \mid d_i)]$$

$$\text{Cov}\,(X_i, X_j)=\text{Cov}\,[E(X_i \mid d_i), E(X_j \mid d_j)]$$

The covariance component between X_i & X_j mainly comes from the covariance component between d_i & d_j in the multinomial distribution given in (2.1). Wherefrom, for all $i, j=1, 2 \ldots M$, $i \neq j$ we obtain

$$E(X_i)=\frac{t(t-1)\,\pi^2}{n} \qquad \ldots(2.3)$$

$$V(X_i)=\frac{t(t-1)}{n^2(n-1)}\Big\{[t^2-(4n+1)t+6n]\pi^4+2n(t-2)\pi^3+n(n-1)\pi^2\Big\} \ldots(2.4)$$

$$\text{Cov}\,(X_i, X_j)=-\frac{2t(t-1)\,(2t-3)\pi^4}{n^2} \qquad \ldots(2.5)$$

which are functions of known quantities t, n & M.

(iii) The Distribution of the Average Number of Doublets

Since the number of doublets (X_i) $(i=1, 2 \ldots M)$ are distributed with common mean, variance and covariance given by (2. 3), (2. 4) & (2.5) respectively, their average defined by

$$\bar{X}=\frac{1}{M}\sum_{1}^{M} X_i \qquad \ldots(2.6)$$

will then have the mean and variance given by

$$E(\bar{X})=\frac{t(t-1)\pi^2}{n} \qquad ...(2.7)$$

$$V(\bar{X})=\frac{t(t-1)\pi}{n^4(n-1)}\left\{(t-2)(t-3)\pi^2-2\,[(n-2)(t-1)+1]\pi+n(n-1)\right\} \qquad ...(2.8)$$

And if n & t are large enough to ignore $1/t$, $2/t$, $1/n$ etc., $V(\bar{X})$ may be approximated by the formula

$$V(\bar{X})\simeq\frac{[n-t\pi]^2t^2\pi}{n^5} \qquad ...(2.9)$$

For large n & d_i values (*i.e.*, as $n\to\infty$ & $d_i\to\infty$ such that $d_i/n\to$ a constant for each $i=1, 2, \ldots M$), each X_i tends to be normally distributed so that their mean $\bar{X}$ will be asymptotically normal with mean and variance given by (2.7) & (2.8), respectively.

A Conditional Test

With these preliminary results a conditional test (conditional on t value being fixed) may be proposed to test the hypothesis whether the spread of the disease from plant to plant is random in nature against the alternative that it is from 'within'. When the latter hypothesis is true, there will be a tendency for a nest of diseased plants appearing together so that the number of doublets will be much larger than may be expected under chance factors. The case of 'too few' number of doublets than may be expected on chance factors or of no doublets at all, though could be an evidence against the Null Hypothesis of random incidence, does not lend any support in favour of the internal spread of the disease ; in fact, it could be due to some other systematic cause (s). Hence the test to be proposed should be obviously one-sided.

In section 2, we have seen that the distribution of the average number of doublets, when the Null Hypothesis is true, could be reasonably assumed to be asymptotically normal with mean and variance given by (2.7) and (2.8) respectively. In otherwords, the conditional distribution of

$$T=\frac{\bar{X}-E(\bar{X})}{\sqrt{V(\bar{X})}} \qquad ...(2.10)$$

for $\sum_{1}^{M} d_i=t$ fixed, is $N(0,1)$, and may be used as a test statistic for testing the Null Hypothesis. That is, the Null Hypothesis may be rejected at 100α percent level whenever the observed value of T exceeds $Z_{(1-\alpha)}$ where $Z_{(1-\alpha)}$ is the $100\,(1-\alpha)$th percentile point in the normal table.

The procedure for testing the hypothesis of the internal spread of the disease is :

(*i*) Choose randomly M plots in the fixed, each plot having exactly n plants ; (*ii*) Count the number of diseased plants d_i and the number of doublets X_i ($i=1, 2 \dots M$) for each selected plot as described by Van der Plank (1960) ; (*ii*) Compute the average number of doublets per plot using (2.6), and also the mean and the variance from (2.7) & (2.8), respectively ; and (*iv*) Finally compute the value for the test statistsic T from (2.10) and reject the Null Hypothesis of random mode of incidence in favour of the alternative hypothesis of the internal spread of the disease if the computed value of T exceeds the table value $Z_{(1-\alpha)}$, where $Z_{(1-\alpha)}$ is the $100(1-\alpha)$th percentile point in the standard normal table corresponding to the 1000α percent level of significance : otherwise, retain the Null Hypothesis of random incidence of the disease.

REFERENCES

The authors are greateful to Dr. R. Narayana, the Director of Instruction (BSH), College of Basic Sciences and Humanities, University of Agricultural Sciences, Bangalore, for his constant encouragement and keen interest in this work. The authors are also grateful to the referees for their valuable comments which have improved the present version of the paper.

REFERENCES

Johnson, N.L. (1969) : *Discrete Distributions*, Houghton Mifflin Co., Boston, U.S.A.

Sundararaj, N. 1976) : "A Statistical Criterion for Judging Internal Spread of Diseases Among Plants". *Current Research* **5**, : pp 13-15. Published by the University of Agricultural Sciences, Bangalore, India.

Van der Plank (1960) : *Analysis of Epidemics* : Chapter VII in *Plants Pathology*, Vol III, The Disease Population : Epidemics and control. Edited by Horsfall, J.G., and Dimond. A.E. Published by the Connecticut Experimental Station, New Haven, Conn. pp 242-243.