A Method for Testing Results of Infectivity Tests with Plant Viruses for Compatibility with Hypotheses

BY A. KLECZKOWSKI

Rothamsted Experimental Station, Harpenden, Hertfordshire

SUMMARY: A statistical method is described to test the compatibility between results of local-lesion counts and hypotheses relating changes in infectivity of plant-virus preparations to treatments applied to them. The method allows the variance-ratio test to be applied: it involves establishing a relation between numbers of lesions and virus concentration for each experiment and a logarithmic transformation of lesion counts to make variance independent of the magnitude of the variate. Its use is illustrated with results obtained to see whether inactivation of viruses by ultraviolet radiation is a first order reaction.

When the infectivity of plant-virus preparations is assayed by the local-lesion method, exact conclusions are difficult to draw because there is no simple relationship between numbers of lesions and virus concentration; also, standard errors of individual numbers are high and depend on the magnitude of the number (Kleczkowski, 1949). With no method to overcome these difficulties, compatibility between experimental results and a hypothesis predicting quantitative changes in infectivity could be judged only by intuition and not tested statistically. The purpose of this paper is to describe a method whereby objective statistical tests can replace subjective judgement.

THE METHOD

Experimental design

When infectivity of several virus preparations is to be assayed, each of them can be inoculated at one convenient dilution for comparison with a standard preparation inoculated at several dilutions. Each inoculum is rubbed on a number of half-leaves (or whole leaves) distributed in such a way as to eliminate as many causes of variation as possible. Thus each inoculum must occur an equal number of times on each of a number of plants, or of blocks of plants, and, if the leaves occupy different positions on the stem, each inoculum must occur an equal number of times on each leaf position or on each leaf level comprising not more than two neighbouring positions. Some convenient distributions are given by Kleczkowski (1950).

The statistical treatment to be described can be applied only if *all* the inocula produce on the average at least about five lesions per half-leaf (or per leaf) and not more than can be conveniently counted.

Statistical test

Let us assume that the hypothesis under test predicts that a given treatment alters the magnitude of virus activity by a factor p which is a function f of a variable v, so that

$$p = f(v, k), \tag{1}$$

where k is an adjustable constant. The variable v can be any quantitative aspect of a treatment, such as time of its duration, concentration of a reagent, etc. There can be more than one function and/or more than one variable or one constant involved, but this does not alter the principle of the procedure.

The procedure is based on two facts. First, it is usually possible, by suitably adjusting the values of three constants, N, ξ and λ, to express the relationship between the numbers of lesions (Y) and the concentration of a virus preparation in the inoculum (x) by the equation

$$Y = \frac{N}{\lambda \sqrt{(2\pi)}} \int_{-\infty}^{u} \exp\left\{-\frac{1}{2}\left(\frac{r-\xi}{\lambda}\right)^2\right\} dr, \tag{2}$$

where $u = \log_{10} x$ if the virus preparation has not been subjected to the treatment under test or $u = \log_{10} px$ if the preparation has been subjected to the treatment which has altered its activity by the factor p of equation (1). (The value of Y/N can be obtained from a table of 'probits'.)

Secondly, if individual numbers of lesions (y) on each half-leaf (or whole leaf) are transformed by the formula

$$z = \log_{10}(y + c), \tag{3}$$

(where c is a constant whose value can be determined from experimental data or taken as anything between 5 and 15), the values of z are approximately normally distributed with a standard error independent of their magnitude (Kleczkowski, 1949). It is essential that the mean value of y, obtained with any inoculum, should not be smaller than about 5, otherwise the transformation (3) may not make the standard error of z independent of its magnitude.

Thus all the numbers of lesions per half-leaf (or whole leaf) are transformed according to the formula (3) and the values of z thus obtained are used throughout.

An estimate of the variance (s^2) of single values of z is then obtained by means of the customary analysis of variance, when use is made of the experimental design eliminating a number of known sources of variation in mean values of z for each inoculum.

Further procedure resembles that used to test different equations for expressing the relationship between numbers of lesions and virus concentrations (Kleczkowski, 1950).

Mean values of z are decimal logarithms of geometric means of experimental values $(y + c)$, whereas the values of Y (given by equation (2)) are expected arithmetic means of the values of y. Thus the values of Y must be transformed into decimal logarithms of expected geometric means of values $(y + c)$, and the logarithms can then be directly compared with mean values of z. The transformation is

$$Z = \log_{10}(Y + c)_1 - 1 \cdot 149 s^2, \tag{4}$$

where Z is an estimate of the logarithm of the geometric mean and s^2 the estimate of the variance of single values of z obtained from the analysis of variance (Kleczkowski, 1950).

The values of the constants k, N, ξ and λ of equations (1) and (2) are adjusted by the method of least squares, i.e. by minimizing the value

$\Sigma((\bar{z}_i - Z_i)^2$, where $\bar{z}_i$'s are the mean values of z obtained with different inocula and Z_i's are the corresponding values of Z of equation (4). As the values of z can be assumed to be normally distributed around the corresponding values of Z with the same variance, the method of the least squares of adjusting the values of the constants is here equivalent with the method of maximum likelihood.

If each $\bar{z}$ is a mean of n values of z, $n\Sigma(\bar{z}_i - Z_i)^2/W$ (where W is the number of degrees of freedom equal to the number of inocula minus the number of adjustable constants), should give an estimate of the variance of single values of z independently of that given by s^2, obtained from the customary procedure of the analysis of variance. If the hypothesis on which equation (1) is based is true, the two estimates should be compatible, and this can be tested by the variance-ratio test.

Adjusting the values of the constants is usually the most laborious part of the procedure. There may be several ways of doing this but all start by guessing likely values of the constants, substituting them and computing the results. Successive improvements are then made. Sometimes the two estimates of the variance can be made compatible by a few successive improvements of the values of the constants simply by guesses helped by inspecting the last results. Various graphical methods of adjusting the constants can also be used. If the two estimates of the variance still remain incompatible, several successive improvements of the constants can finally be made by computing increments to be added to the constants by solving the 'normal' equations:

$$\Delta N \Sigma \left(\frac{\partial Z_i}{\partial N}\right)^2 + \Delta \xi \Sigma \left(\frac{\partial Z_i}{\partial N}\frac{\partial Z_i}{\partial \xi}\right) + \Delta \lambda \Sigma \left(\frac{\partial Z_i}{\partial N}\frac{\partial Z_i}{\partial \lambda}\right) + \Delta k \Sigma \left(\frac{\partial Z_i}{\partial N}\frac{\partial Z_i}{\partial k}\right) = \Sigma \left\{\frac{\partial Z_i}{\partial N}(\bar{z}_i - Z_i)\right\},$$

$$\Delta N \Sigma \left(\frac{\partial Z_i}{\partial \xi}\frac{\partial Z_i}{\partial N}\right) + \Delta \xi \Sigma \left(\frac{\partial Z_i}{\partial \xi}\right)^2 + \Delta \lambda \Sigma \left(\frac{\partial Z_i}{\partial \xi}\frac{\partial Z_i}{\partial \lambda}\right) + \Delta k \Sigma \left(\frac{\partial Z_i}{\partial \xi}\frac{\partial Z_i}{\partial k}\right) = \Sigma \left\{\frac{\partial Z_i}{\partial \xi}(\bar{z}_i - Z_i)\right\},$$

$$\Delta N \Sigma \left(\frac{\partial Z_i}{\partial \lambda}\frac{\partial Z_i}{\partial N}\right) + \Delta \xi \Sigma \left(\frac{\partial Z_i}{\partial \lambda}\frac{\partial Z_i}{\partial \xi}\right) + \Delta \lambda \Sigma \left(\frac{\partial Z_i}{\partial \lambda}\right)^2 + \Delta k \Sigma \left(\frac{\partial Z_i}{\partial \lambda}\frac{\partial Z_i}{\partial k}\right) = \Sigma \left\{\frac{\partial Z_i}{\partial \lambda}(\bar{z}_i - Z_i)\right\},$$

$$\Delta N \Sigma \left(\frac{\partial Z_i}{\partial k}\frac{\partial Z_i}{\partial N}\right) + \Delta \xi \Sigma \left(\frac{\partial Z_i}{\partial k}\frac{\partial Z_i}{\partial \xi}\right) + \Delta \lambda \Sigma \left(\frac{\partial Z_i}{\partial k}\frac{\partial Z_i}{\partial \lambda}\right) + \Delta k \Sigma \left(\frac{\partial Z_i}{\partial k}\right)^2 = \Sigma \left\{\frac{\partial Z_i}{\partial k}(\bar{z}_i - Z_i)\right\}.$$

The partial derivatives are

$$\frac{\partial Z}{\partial N} = \frac{0\cdot4343}{(Y+c)\lambda\sqrt{(2\pi)}} \int_{-\infty}^{u} \exp\left\{-\frac{1}{2}\left(\frac{r-\xi}{\lambda}\right)^2\right\} dr,$$

$$\frac{\partial Z}{\partial \xi} = -\frac{0\cdot4343 N}{(Y+c)\lambda\sqrt{(2\pi)}} \exp\left\{-\frac{1}{2}\left(\frac{u-\xi}{\lambda}\right)^2\right\},$$

$$\frac{\partial Z}{\partial \lambda} = -\frac{0\cdot4343 N(u-\xi)}{(Y+c)\lambda^2\sqrt{(2\pi)}} \exp\left\{-\frac{1}{2}\left(\frac{u-\xi}{\lambda}\right)^2\right\},$$

$$\frac{\partial Z}{\partial k} = \frac{0\cdot1886 N}{(Y+c)p\lambda\sqrt{(2\pi)}} \exp\left\{-\frac{1}{2}\left(\frac{u-\xi}{\lambda}\right)^2\right\}\frac{\partial p}{\partial k}.$$

As soon as the two estimates of the variance of z do not differ significantly, it can be concluded that the hypothesis under test is compatible with the

experimental data. If the two estimates of the variance remain incompatible, the adjustment of the constants can be discontinued when it becomes obvious that the value $\Sigma(\bar{z}_i - Z_i)^2$ cannot be further reduced to any appreciable extent. It can then be concluded that the hypothesis is incompatible with the data.

The course of inactivation of two viruses by ultraviolet radiation

As an example of application of the method results of three experiments made by Bawden & Kleczkowski (1953) with ultraviolet-irradiated viruses (Exps. 1, 2 and 4 in Table 3 of their paper) will be tested for compatibility with the hypothesis that inactivation by ultraviolet radiation is a first-order reaction. If it is, the fraction p of the original activity left after the time t of irradiation will be

$$p = e^{-kt}, \tag{5}$$

where k is an adjustable constant whose value depends on experimental conditions. Equation (1) takes the form of (5), the variable v becoming t.

Table 1 shows the results of analyses of variances carried out on the values of z obtained by (3) using 5 as the constant c, and Tables 2–4 show the results of testing the experimental data for compatibility with the hypothesis.

In Exp. 4 (Table 4) tobacco mosaic virus was irradiated at two different concentrations, so that two different values of k (k_1 and k_2) had to be used. Had it been necessary to use the 'normal' equations for adjusting the values of the five constants involved, there would have been five such equations.

The results of Exps. 1 and 4 (Tables 2 and 4) are compatible with the hypothesis, whereas those of Exp. 2 (Table 3) are not. However, although the deviations of the values of $\bar{z}$ from the corresponding values of Z are statistically significant, they are small, so that it can be concluded that if the inactivation is not exactly a first-order reaction, its course approaches that of a first-order reaction quite closely. It is, of course, also possible that the inactivation is exactly a first-order reaction and that Exp. 4 contained some experimental inaccuracies that were not allowed for in the treatment of the data.

REFERENCES

BAWDEN, F. C. & KLECZKOWSKI, A. (1953). The behaviour of some plant viruses after exposure to ultraviolet radiation. *J. gen. Microbiol.* 8, 145.

KLECZKOWSKI, A. (1949). The transformation of local lesion counts for statistical analysis. *Ann. appl. Biol.* 36, 139.

KLECZKOWSKI, A. (1950). Interpreting relationships between the concentrations of plant viruses and numbers of local lesions. *J. gen. Microbiol.* 4, 53.

Table 1. *Analyses of variances*

	Exp. 1 Rothamsted tobacco necrosis virus French bean			Exp. 2 Rothamsted tobacco necrosis virus French bean			Exp. 4 Tobacco mosaic virus *Nicotiana glutinosa*		
Virus ... Host plant ...	Degrees of freedom	Sums of squares	Mean square (s^2)	Degrees of freedom	Sums of squares	Mean square (s^2)	Degrees of freedom	Sums of squares	Mean square (s^2)
Between inocula	7	12·1369	—	7	17·6574	—	11	12·9707	—
Between leaf levels*	—	—	—	—	—	—	2	1·6568	—
Between plants (or blocks of plants)	13	0·4990	—	13	0·7670	—	8	5·1102	—
Left–right half-leaves	1	0·0163	—	1	0·0181	—	—	—	—
Residual	90	1·4004	0·0156	90	0·9797	0·0109	86	4·4219	0·0514
Total	111	14·0526		111	19·4222		107	24·1596	

Nicotiana glutinosa plants used in Exp. 4 had six leaves each. The six leaf positions were divided into three levels and each inoculum occurred three times in each level.

Table 2. *Comparison of experimental with computed values for Exp.* 1

Rothamsted tobacco necrosis virus irradiated as a 1 % solution

Time of irradiation t (min.)	$\log_{10}$ of concentration (in %) at which the virus was inoculated	u	Z	$\bar{z}$	Expected fraction of the original activity
2·480		−3·495	2·1524	2·1314	0·32
4·243		−3·848	2·0017	1·9957	0·14
6·006	−3·0	−4·200	1·8023	1·8286	0·063
7·769		−4·552	1·5591	1·5071	0·029
0	−3·7	−3·7	2·0713	2·0657	
0	−4·2	−4·2	1·8057	1·8014	
0	−4·7	−4·7	1·4472	1·4729	1·00
0	−5·2	−5·2	1·0620	1·1036	

Constants: $N=260$; $\xi=-3·6$; $\lambda=0·83$; $k=0·46$.

$\Sigma(\bar{z}_i - Z_i)^2 = 0·006314$; $n=14$; $W=4$.

Variance ratio: $R=1·42$ $\begin{cases} n_1=4. \\ n_2=90. \end{cases}$

$P>0·2$.

Table 3. *Comparison of experimental with computed values for Exp.* 2

Rothamsted tobacco necrosis virus irradiated as a 0·1 % solution.

Time of irradiation t (min.)	$\log_{10}$ of concentration (in %) at which the virus was inoculated	u	Z	$\bar{z}$	Expected fraction of the original activity
0·58		−3·736	2·0313	1·9807	0·18
1·00		−4·268	1·6835	1·6221	0·054
1·42	−3·0	−4·801	1·2903	1·3443	0·016
1·84		−5·333	0·9334	0·9600	0·0047
0	−3·7	−3·7	2·0512	2·0943	
0	−4·2	−4·2	1·7315	1·7714	
0	−4·7	−4·7	1·3650	1·3700	1·000
0	−5·2	−5·2	1·0185	0·9921	

Constants: $N=580$; $\xi=-2·8$; $\lambda=1·03$; $k=2·92$.

$\Sigma(\bar{z}_i - Z_i)^2 = 0·0141$; $n=14$; $W=4$.

Variance ratio: $R=4·5275$ $\begin{cases} n_1=4. \\ n_2=90. \end{cases}$

$P=0·001$ (approximately).

Table 4. *Comparison of experimental with computed values for Exp. 4.*

	Tobacco mosaic virus					Expected fraction of the original activity
	Time of irradiation t (min.)	$\log_{10}$ of concentration (in %) at which the virus was inoculated	u	Z	z	
Virus irradiated as a 0·5 % solution	3·85		−3·002	2·0876	1·9633	0·32
	7·70	−2·5	−3·503	1·8383	1·8933	0·099
	11·55		−4·005	1·5595	1·5178	0·031
	15·40		−4·506	1·2644	1·3111	0·0098
Virus irradiated as a 0·01 % solution	0·575		−3·163	2·0038	1·9333	0·22
	1·150	−2·5	−3·822	1·6619	1·7800	0·047
	1·725		−4·485	1·2784	1·1989	0·0104
	2·300		−5·147	0·9304	0·9711	0·0023
Unirradiated virus	0	−3·0	−3·0	2·0876	2·1456	1·000
	0	−3·5	−3·5	1·8383	1·7689	
	0	−4·0	−4·0	1·5595	1·5833	
	0	−4·5	−4·5	1·2697	1·2711	

Constants: $N = 850$; $\xi = -1·6$; $\lambda = 1·4$;

$k_1 = 0·3$ (for 0·5 % virus solution);

$k_2 = 2·65$ (for 0·01 % virus solution).

$\Sigma(\bar{z}_i - Z_i)^2 = 0·0589$; $n = 9$; $W = 7$;

Variance ratio: $R = 1·473$ $\begin{cases} n_1 = 7. \\ n_2 = 86. \end{cases}$

$P = 0·2$ (approximately).