United States
Environmental Protection
Agency

Environmental Monitoring
Systems Laboratory
Cincinnati OH 45268

Research and Development

EPA/600/S4-89/002 Sept. 1989

EPA

Project Summary

Experimental Design and Data Analysis Applicable to Assays for Monitoring Waterborne Viruses

Larry J. Wymer

Suitable statistical methods applicable to assays for monitoring waterborne viruses as described in the "USEPA Manual of Methods for Virology" (EPA-600/4-84-013) are presented. These methods have been selected to show the non-statistician how measurement evaluations should be made to analyze collected waterborne virus data. The specific experimental situations that have been included pertain to relative frequencies of virus types, estimates of viral titer, assessing the precision of these estimates, and comparing results among subsamples. Also included are numerous references for additional information regarding statistical theory and appropriate statistical tables, many of which are not commonly available from other sources but are essential to performing the analyses suggested in this report.

This Project Summary was developed by EPA's Environmental Monitoring Systems Laboratory, Cincinnati, OH, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

Currently, the "USEPA Manual of Methods for Virology" makes no mention of the statistical treatment of accumulated data. Given the effort and expense of performing viral assays, it seems appropriate that guidelines for the evaluation of data obtained from these assays be available.

Single titrations are commonly performed to monitor viruses in environmental samples. The basic unit of infection, observable by its cytopathic effects within the cell sheet, resulting in the appearance of a plaque, is the plaque forming unit (pfu). A pfu may consist of a single virus, or it may be an aggregation of two or more viruses. A pfu, regardless of its exact composition, is taken as the minimum level of exposure of an organism to the virus.

Two critical assumptions are made with respect to the formation and observation of plaques obtained from these assays. One assumption is that the plaque count itself is not subject to error. Potential sources of error in plaque counting arise from the existence of "false positives" among plaques counted in the assay and overlapping of plaques on the cell sheet. False positives may be eliminated through the confirmation of each plaque as being caused by a virus, while the problem of overcrowding may be minimized by using only those results obtained at suitable dilutions of the test material.

The second assumption made is that a single pfu is sufficient to infect a cell. This assumption of single hit kinetics implies that the mean number of plaques at any level of concentration is directly proportional to the amount of test material used in the inoculum. Cases have been reported for which, apparently, one or two virus particles may be required for infection of the cell to occur.

The single hit model may still be a reasonable approximation to such data, however, if the amount of test material used for each inoculation is small enough that the response is still approximately proportional to the dosage.

Results

The Poisson distribution has widespread application in the modeling of experimental data consisting of the number of times some event occurs over a fixed interval of time, length, area, or volume. When such counts follow a Poisson distribution, the probability of x occurrences of the event in a fixed interval is given by the probability distribution function (pdf):

$$f(x) = e^{-\theta}\theta^x/x!$$

The pdf for a Poisson process is completely specified by its mean, θ. The true value of θ will depend on the amount of material used, which for viral assays corresponds to the volume, v, used in the inoculum before dilution. Thus, $\theta = \tau v$, where τ is the true mean density of pfu's per unit volume of eluate ("pfu titer").

The assumption that plaque counts follow a Poisson law is often made in practice. However, little has been done to validate this assumption. To this end, plaque count data from 65 raw and treated sewage sample titrations (Table 1) and from a round-robin soil study involving 29 sandy loam and 29 sand samples (Table 2) were utilized to determine whether the Poisson adequately characterizes viral assay data. Fisher's index of dispersion (D) was used as the test statistic. Under the null hypothesis, that the plaque counts within each trial follow a Poisson process, D is approximately distributed as a χ^2 variable with k-1 degrees of freedom, where k is the number of independent counts obtained from that trial. Generally, the number of independent counts is equal to the number of cell culture bottles used; however, in some cases the results from two or more bottles are combined in order to ensure that the expected count for each grouping is at least five — a requirement in order for the χ^2 approximation to apply. Of the 29 sewage samples titrations inoculated in BGM cell cultures, only 2 failed the test for a Poisson distribution (San Lorenza influent under the virus adsorption elution cartridge filter method and Guaynabo influent under the beef extract-celite precipitation method). Even if all trials were truly Poisson distributed, one would expect one or two rejections of the null hypothesis (H_0) among 29 such tests, simply due to the level of type I error (0.05) used. Thus, these results indicate excellent agreement with a Poisson assumption.

Of the 36 sewage sample titrations inoculated in bovine kidney (MDBK) cell cultures, 9 led to rejection of a random dispersion of plaques throughout the medium. This is much higher than the rejection rate that would be expected if all trials were Poisson processes and is, therefore, reliable evidence that at least some of the data are non-Poisson. The MDBK cell line used in this assay, however, was later found to be contaminated; although MDBK cells are not susceptible to coxsackie virus infection, plaques of this type were identified among those found on the cell sheet. Thus, the non-random dispersion of plaques was affected by the distribution of cell types within the culture. These results illustrate the use of a test for randomness in identifying results that may be suspect.

Among the 29 sand sample trials performed in the soil round-robin study, none failed the test for a Poisson distribution. However, 8 of the 29 sandy loam trials resulted in rejection of the null hypothesis that the distribution of the plaques follows a Poisson law.

These results lead to the conclusion that the assumption of a Poisson distribution is not unreasonable for plaque count data, although testing of the assumption whenever possible is warranted. Failure to obey a Poisson distribution may be due to the method used in processing the sample, which differs among water, sludge, and soil samples, or may result from distribution of pfu's in the sample itself.

Conclusions

Standard statistical reporting that should be incorporated as part of every viral monitoring assay includes:

1. Test for Poisson distribution of plaques at a 0.05 critical level.
2. Confirmed virus pfu titer and associated 95% confidence interval.
3. Titers by virus type and associated 95% joint confidence intervals.
4. Relative proportions of virus types and associated 95% joint confidence intervals.

When plaques are shown to follow a Poisson distribution, 95% confidence intervals for the Poisson parameter, θ, may be calculated from the cumulative Poisson distribution function. Tables of lower and upper 95% confidence lim for Poisson counts and for lower a upper simultaneous 95% confiden limits for up to five virus types a included in the full report. Because t Poisson distribution has only one para eter, confidence limits for a Poiss mean are completely determined by t total count.

Tables of 95% confidence limits f binomial proportions and for 95% sim taneous confidence limits for multinom proportions for up to five virus types a also included in the full report. Thes may be used in constructing confiden intervals for proportions of virus type found in the sample.

A significant departure from customa practice is recommended in the way which plaques are selected for confi mation and identification. It is recon mended that all plaques be confirme and identified to the extent possible; if is not practical to perform this assay o every plaque found in the experimen then a sufficient number of cell cultu bottles should be randomly selected, an all plaques in these bottles so assayed. this recommendation is followed, a dire estimation of titer by virus type can b made. In addition, this procedure elim inates bias that may result from allowin the operator to select the specifi plaques to be confirmed and typed.

Because all plaques in a bottle are t be picked, confirmation and identificatio may be performed on each day, a plaques are counted on the cell sheets.

Given the widespread availability o desktop computers to lab personnel, th development of software to perform th analysis of viral assay data is recom mended. The existence of such softwar would eliminate the reliance on statistica tables, diminish the possibility of huma error, enable standardization of th analysis and reporting of the results, an reduce the time required to perform th analyses. A full range application i recommended; this would incorporate no only data analysis but also data entry data management, file handling, an reporting features.

Table 1. *Texts for Poisson Distribution of Plaque Counts Based on Fisher's Index of Dispersion: Puerto Rico Sewage Treatment Plant Study*

	CF[a] Method				PPT[b] Method			
	BGM[c] Cells		MDBK[d] Cells		BGM Cells		MDBK Cells	
	DF[e]	D[f]	DF	D	DF	D	DF	D
Aibonito Effluent			9	8.702				
Aibonito Influent	9	2.392	9	14.147			5	1.857
Barranquitas Effluent			9	4.882				
Barranquitas Effluent (2)[g]	9	4.667	9	11.479				
Caguas Effluent	9	11.871	9	7.364				
Caguas Influent	9	13.101	9	44.735*	9	10.842	5	1.432
Cazey Effluent	9	5.213	9	8.890	9	8.769	2	2.632
Cazey Influent	9	5.529	4	2.122	9	11.750		
Cidra Effluent			9	21.006*				
Cidra Effluent(2)	9	4.446	9	20.889*			2	2.804
Cidra Influent	9	4.357	9	15.769	4	3.667	2	0.080
Comerio Effluent	9	10.492	4	1.510	4	3.348		
Comerio Influent	9	6.366	9	18.000*	4	5.750		
Guaynabo STP Effluent	4	0.500	9	7.346	2	5.804	2	1.625
Guaynabo STP Influent	9	7.352	9	14.404	1	7.364*	1	0.818
Gurabo Influent	4	3.238	4	3.024				
Gurabo Influent(2)	19	20.760						
Juncos Effluent	9	8.385	9	14.045			2	1.333
Juncos Influent	9	4.083	9	22.170*			1	0.600
Narangito Effluent			2	0.028				
Pueblito El Rio Influent							4	5.545
San Lorenzo Effluent	1	0.818	2	13.500*				
San Lorenzo Effluent(2)	5	4.267						
San Lorenzo Influent			9	23.034*				
San Lorenzo Influent(2)	3	13.0*						
Villalba Effluent			8	37.462*	9	12.455		
Villalba Effluent(2)	9	11.289	9	8.836				
Villalba Influent			9	31.291*				
Vista Monte Influent			2	2.848			2	3.000

*Indicates significant departure from Poisson distribution at 0.05 critical level
[a]Virus adsorption elution (VIRADEL) cartridge filter
[b]Beef extract - celite precipitation method
[c]Buffalo green monkey kidney cells
[d]Madin and Dorby bovine kidney
[e]Degrees of freedom for chi-square test
[f]Index of dispersion
[g](2) = Retrial

Larry J. Wymer is with Computer Sciences Corporation, Fall Church, VA 22042.
Robert S. Safferman is the EPA Project Officer (see below).
The complete report, entitled "Experimental Design and Data Analysis Applicable to Assays for Monitoring Waterborne Viruses," (Order No. PB 89-148 571/AS; Cost: $15.95, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Environmental Monitoring Systems Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

BULK RATE
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Official Business
Penalty for Private Use $300

EPA/600/S4-89/002