ELSEVIER

Journal of Microbiological Methods 35 (1999) 157–161

Journal of Microbiological Methods

Estimating microbial population counts by 'most probable number' using Microsoft Excel®

Aurelio M. Briones Jr.*, Wolfgang Reichardt

International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines

Received 8 June 1998; received in revised form 4 November 1998; accepted 11 November 1998

Abstract

A computer-assisted method for determining population counts using the 'most probable number' (MPN) was developed. The Microsoft Excel® spreadsheet and its Solver® tool were used to generate MPNs, error estimates and confidence limits. Our method was flexible, allowing the use of unbalanced replication schemes and varying replication numbers and inoculation volumes. Furthermore, it required no programming skills and generated fast results, which were comparable to those of standard MPN tables and MPN software.

Keywords: Computer software; Most probable number; Microbial populations; Counts; Microsoft Excel

1. Introduction

Many microbiologists rely on 'most probable number' (MPN) tables to derive population estimates in a dilution series. Such reliance imparts unnecessary rigidity to the experimental design, i.e., it involves a symmetrical dilution series with a constant, table-specified dilution factor. Moreover, these tables do not provide the associated probability values for a particular experimental outcome. An alternative involves the use of one or more customized computer programs (Clarke and Owens, 1983; Woomer et al., 1990) that may require some programming skills in order to tailor them to individual requirements.

Many of today's computers run versions of Windows® and Microsoft Office® (Microsoft). For researchers using these products, the convenience of retaining a familiar operating environment is usually welcome. In this report, we present a simple method for determining the most probable number using Microsoft Excel®. It requires no programming skills and only some familiarity with MS Excel and its associated tools.

2. Materials and methods

2.1. Software and hardware requirements

The formulas for deriving the MPN estimates are based on Halvorson and Ziegler (1933). Using Microsoft Excel (Microsoft) and the Solver (Frontline Systems) tool, MPN estimates can be generated efficiently using any IBM-compatible PC capable of running Windows and MS Office.

*Corresponding author. Tel.: +63-2-845-0563, ext. 310/448; fax: +63-2-761-2406; e-mail: abriones@irri.cgiar.org

0167-7012/99/$ – see front matter
PII: S0167-7012(98)00111-0

Table 1
Suggested spreadsheet format (the column headings are located in the first row, and therefore occupy A1 to K1)

	A	B	C	D	E	F	G	H	I	J	K
1	Sample	Dilution	No. positive	No. negative	No. of tubes	Vol. inoc.	x	L	R	P	Pop. est.
2	Example	0.00001	4	1	5	1	3.2597	5.55	5.55	0.25915	3.26E+06
3		1E-06	3	2	5	0.1					
4		1E-07	1	4	5	0.01					
5											

2.2. Configuring the spreadsheet

The following notations were used: a, volume of the dilution used for inoculation, assuming a single source of inoculum for all dilution levels; p, number of tubes showing positive growth; q, number of tubes showing negative growth; n, number of replicate tubes used in each dilution level; x, the most probable number.

Table 1 illustrates the suggested spreadsheet format. The sample label and dilutions scored were entered in the first and second columns, respectively. When comparing results with standard MPN table (Table 2) values, the formulas used three 'critical' dilutions. The range of relevant dilutions may be broadened, because the formulas are expandable. Comparisons with MPN estimates from other sources (Table 3) used six dilution levels. For the sake of brevity, the succeeding formulas take into account only three dilution levels; the expansion follows a logical extension of the first three elements in the series.

The actual volume inoculated (1 ml in the example) was entered in the cell corresponding to the first critical dilution (cell $F2$) and the succeeding volumes were derived from the formula

volume inoculated (succeeding)

$$= \left(\frac{\text{dilution inoculated}}{\text{dilution source}}\right) \times \text{volume inoculated (actual)}$$

In Excel notation this was entered as:

cell F3 = (B3/B2) · F2 and cell F4 = (B4/B2) · F2,

assuming a constant inoculation volume. If volumes differ between sets of tubes, these should be entered instead of $F2$.

The MPN (x) value was calculated from the equation

$$\frac{a_1 p_1}{1-e^{-a_1 x}} + \frac{a_2 p_2}{1-e^{-a_2 x}} + \frac{a_3 p_3}{1-e^{-a_3 x}} = a_1 n_1 + a_2 n_2 + a_3 n_3$$

The left side of the equation was entered in the L column of Table 1 as

cell H2 = ((F2*C2)/(1 − EXP(−F2*G2)))
+ ((F3*C3)/(1 − EXP(−F3*G2)))
+ ((F4*C4)/(1 − EXP(−F4*G2)))

while the right side was entered in the R column of Table 1 as

cell I2 = F2*E2 + F3*E3 + F4*E4

Solving for the MPN was achieved by the Excel Solver tool, located on the 'Tools' menu. Since the typical configuration of Excel may not include this tool, the user may need to reinstall the software to include Solver from among the various 'Add-Ins'. The solution was generated by specifying the target cell as H2 equal to the corresponding cell under the R column (5.55 in Table 1). The final step was to specify the cell to solve in order to attain the equality between H2 and I2, which is the corresponding cell under column x (the MPN estimate) of Table 1.

The probability value associated with combina-

Table 2
Comparison of Excel output with MPN values for use with decimal dilutions and five tubes per dilution (after Alexander (1982), from Cochran (1950))

Data for three 'critical' dilutions	Excel	Cochran (1950)
5-2-1	0.70	0.70
4-3-2	0.39	0.39
4-4-0	0.34	0.34
2-2-0	0.09	0.09
5-1-0	0.33	0.33

Table 3
Comparison of Excel output with MPN population estimates from other sources using six dilution levels.

Dilution ratio	n	Data	Excel	Woomer et al. (1990)	Other MPN (source)
4	2	2-1-1-0-0-0	42.6	42.6	36.0 (Vincent, 1970)
4	4	4-4-4-3-1-0	319.7	318.9	290 (Vincent, 1970)
5	4	4-4-0-2-0-0	70.7	70.7	71.0 (Brockwell, 1963)
10	5	5-3-1-0-0-0	106.9	106.9	110 (Alexander, 1982)
10	3	3-2-2-0-0-0	210.2	210.2	210 (de Man, 1975)
10	5	5-4-2-0-0-0	215.6	215.5	210 (de Man, 1975)
10	10	10-9-5-0-0-0	331.6	331.6	350 (de Man, 1975)

tions of p and q values given a certain value of x was entered under column P. This provides an estimate of frequency of a certain combination for a given probable number x, and was derived from the formula

$$P = \left(\frac{n_1!}{p_1!q_1!}\right)\left(\frac{n_2!}{p_2!q_2!}\right)\left(\frac{n_3!}{p_3!q_3!}\right)$$
$$(e^{-a_1x})^{q_1}(e^{-a_2x})^{q_2}(e^{-a_3x})^{q_3}(1 - e^{-a_1x})^{p_1}(1 - e^{-a_2x})^{p_2}(1 - e^{-a_3x})^{p_3}$$

which was entered in the spreadsheet as

cell J2 = (((FACT(E2))/(FACT(C2)*FACT(D2)))
*((FACT(E3))/(FACT(C3)*FACT(D3)))
*((FACT(E4))/(FACT(C4)*FACT(D4)))
*((EXP(−F2*G2))^{D2})
*((EXP(−F3*G2))^{D3})
*((EXP(−F4*G2))^{D4})
*((1 − EXP(−F2*G2))^{C2})
*((1 − EXP(−F3*G2))^{C3})
*((1 − EXP(−F4*G2))^{C4})) × 100

The population estimate (pop. est.) was solved by multiplying the calculated x value (the MPN) by the reciprocal of the single dilution source used in basing the values for the volumes inoculated. In the example given above, this is the first critical dilution (10^{-5}). In Excel notation, this was entered as

K2 = G2*(1/B2)

3. Results and discussion

MPN estimates generated by Excel were comparable to standard MPN tables (Brockwell, 1963; de Man, 1975; Alexander, 1982) and computer programs (Clarke and Owens, 1983; Woomer et al., 1990) (Tables 2–4). Excel estimates were the least similar to estimates that relied on Fisher and Yates

Table 4
Comparison of Excel output with MPN estimates from Clarke and Owens (1983)

Example no.	Dilution	Replicates	Score	Volume inoculated	MPN/unit volume	
					Excel	Clarke and Owens (1983)
1	1	5	5	1	78.1988	78.2001
	10	5	5	1		
	100	5	3	1		
	1000	5	0	1		
	10 000	5	0	1		
2	1	3	3	1.1	105.494	105.496
	10	3	3	1.05		
	100	10	8	0.95		
	200	10	3	0.98		
	400	5	0	1		

(1943) and Vincent (1970), which confirms the results of other computer software (Woomer et al., 1990). Most works, including the present article, relied on equations derived by Halvorson and Ziegler (1933). As seen from Table 4, the spreadsheet tool can also be used to calculate MPNs using an unbalanced replication scheme with varying replication numbers and inoculation volumes.

3.1. Limitations

The Solver program requires 'guess' estimates (starting values) when solving for x. Actually, because the spreadsheet calculates values of L and R from a, n, p and x, x can be solved iteratively, without using the Solver tool. We recommend that researchers solve for x initially by this trial-and-error method, which involves entering a starting value in the x column until L attains a value greater than or equal to R. The x estimate can then be resolved using the Solver tool. Entering an appropriate starting value is important because if the initial guess for x is too far off, Solver may encounter convergence problems; if so, it prompts the user. Although some nonlinear problems may have many locally optimum points, our experience has shown that a starting value of positive 1 is close enough to the global optimum for Solver to encounter minimal problems. The whole process takes only a few seconds, assuming the starting value is acceptable. Furthermore, the precision of the estimate can be adjusted up to nine decimal places.

The '#NUM' error message in Excel may sometimes arise when P is calculated. This invariably means that Excel has encountered a number beyond the range of $\pm 1 \times 10^{307}$. Because this happens only when the x values are large or when $p=n$ and $q=0$, this can be remedied by canceling out parts of the P formula using zero as the exponent, because

$$(e^{-a_z x})^0 = 1 \text{ and } (1 - e^{-a_z x})^0 = 1$$

3.2. Calculation of 95% confidence intervals

Confidence limits associated with a particular population estimate can be calculated through the confidence factor (CF), a number by which a population estimate is multiplied and divided to get the upper and lower limits, respectively (Cochran, 1950). The information required to calculate CF is the dilution ratio (dr) and the number of replications per dilution level (n). The dilution ratio refers to the ratio between the lower and higher dilutions in a particular series: in a 10-fold series, the ratio is 10; in a 5-fold series, it is 5, etc. The formula for obtaining CF at $P = 0.05$ is

$$CF = \text{antilog}_{10}(2 \cdot 0.55 \cdot \sqrt{[(\log_{10} \text{dr})/n]}) \text{ if dr} < 10 \text{ and}$$
$$= \text{antilog}_{10}(2 \cdot 0.58 \cdot \sqrt{[(\log_{10} \text{dr})/n]}) \text{ if dr} \geq 10$$

In Excel, this can be entered using the 'IF' function. Because n already occupies column E of Table 1, calculating for CF requires only an additional column for the dilution ratio. You can add these new parameters as additional columns in Table 1.

L	M
dr	CF
10	3.302

The equations for CF are entered as

cell N2 = IF(L2 < 10,10^(2*0.55
*(SQRT((LOG10(L2))/E2))),10^(2*0.58
*(SQRT((LOG10(L2))/E2))))

in Microsoft Excel. In the example, the value of 3.302 would be used to multiply and divide the population estimate to determine the upper and lower confidence limits. The value of CF decreases (confidence limits become narrower) when the number of replications increases or when the dilution ratio decreases.

4. Conclusions

Using an Excel tool to estimate MPN may allow greater freedom to design experiments with varying requirements. This flexibility allows us to specify dilution schemes (which need not be balanced), replicate unit numbers, and inoculation volumes while providing relevant error estimates, including 95% confidence intervals.

References

Alexander, M., 1982. Most probable number method for microbial populations. In: Page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis, Part 2, 2nd ed. American Society of Agronomy, Madison, WI, pp. 815–820.

Brockwell, J., 1963. Accuracy of a plant-infection technique for counting populations of *Rhizobium trifolii*. Appl. Microbiol. 11, 377–383.

Clarke, K.R., Owens, N.J.P., 1983. A simple and versatile microcomputer program for the determination of 'most probable number'. J. Microbiol. Methods 1, 133–137.

Cochran, W.G., 1950. Estimation of bacterial densities by means of the 'most probable number'. Biometrics 6, 105–116.

de Man, J.C., 1975. The probability of most probable numbers. Eur. J. Appl. Microbiol. 1, 67–78.

Fisher, R.A., Yates, F., 1943. In: Statistical Tables for Biological, Agricultural and Medical Research. Oliver and Boyd Ltd., London, pp. 6–8.

Halvorson, H.O., Ziegler, N.R., 1933. Application of statistics to problems in bacteriology. I. A means of determining bacterial population by the dilution method. J. Bacteriol. 25, 101–121.

Vincent, J.M., 1970. In: A Manual for the Practical Study of the Root Nodule Bacteria. Blackwell Scientific, Oxford and Edinburgh.

Woomer, P., Bennet, J., Yost, R., 1990. Overcoming the inflexibility of most-probable-number procedures. Agron. J. 82, 349–353.