PROBABILITY OF DETECTING NEMATODE INFESTATIONS IN QUARANTINE SAMPLES†

R. McSorley[1] and R. C. Littell[2]

Entomology and Nematology Department[1] and Statistics Department,[2] University of Florida, Gainesville, FL 32611, U.S.A.

ABSTRACT

McSorley, R., and R. C. Littell. 1993. Probability of detecting nematode infestations in quarantine samples. Nematropica 23:177–181.

In quarantine and certification programs, intensive sampling may be needed to determine if lots or shipments of plants, pots, cuttings, or other units are free of plant-parasitic nematodes and other plant pests. It is essential for each quarantine program to define an acceptable tolerance limit for each pest sampled, recognizing that zero tolerance requires sampling everything in the shipment, which is usually impractical. The binomial and hypergeometric probability distributions were used as bases for determining probabilities of detecting various infestation levels as increasing numbers of samples are collected. Probabilities of detecting infestations of 50%, 10%, 5%, 1%, or 1 unit in lot sizes of 100, 1 000, 10 000, and 100 000 units are provided for selected numbers of samples. According to the binomial probability distribution, if 5% of the units in a lot are infested, 59 and 90 samples, respectively, would be required to detect the infestation 95% and 99% of the time. When only 1% of the units are infested, sample sizes must be increased to 300 and 500 for 95% and 99% detection, respectively. Subject to certain assumptions, methodology is provided to estimate the probability of detection for various lot and sample sizes.

Key words: binomial distribution, certification, hypergeometric distribution, nematodes, quarantine, sampling.

RESUMEN

McSorley, R. y R. C. Littell. 1993. Probabilidad de detectar infestaciones de nematados en muestreos de cuarentena. Nematrópica 23:177–181.

En programas cuarentenarios y de certificación los muestreos intensivos podrían ser requeridos para determinar si los lotes o envíos de plantas, maceteros, virutas o cualquier otra parte vegetal está libre de nematodos u otras plagas. Es esencial para cada programa cuarentenario definir un límite de tolerancia aceptable para cada plaga muestreada, en el entendido que cero tolerancia implica muestrear todo en el envío lo cual es impráctico. Las probabilidades de las distribuciones binomial e hipergeométrica fueron utilizadas como bases para determinar las probabilidades de detectar varios niveles de infestación en la medida que se incrementa el número de muestras recolectadas. Las probabilidades para detectar infestaciones de 50%, 10%, 5% y 1% o una unidad en lotes de 100, 1 000, 10 000 y 100 000 unidades se suministran para números seleccionados de muestras. De acuerdo a la probabilidad de la distribución binomial, si 5% de las unidades de un lote están infestadas, se requerirían 59 y 90 muestras para detectar infestaciones de 95% y 99% respectivamente, para un tiempo dado. Cuando sólo el 1% de las unidades están infestadas, el tamaño de muestra debe incrementarse a 300 y 500 para detectar el 95% y 99%, respectivamente. Bajo ciertos supuestos se da una metodología para estimar la probabilidad de detección para varios lotes y tamaños de muestras.

Palabras clave: certificación, cuarentenas, distribución binomial, distribución hipergeométrica, muestreo, nematodos.

INTRODUCTION

Quarantine and certification programs are important in limiting the spread of plant-parasitic nematodes and other plant pests (3). Fundamental to such programs is the inspection of lots or shipments of plants, pots, soil, cuttings, or other units for detection of pests, which may occur

†Florida Agricultural Experiment Station Journal Series No. R-03200.

at low incidences. Unfortunately, complete certainty of pest absence cannot be assured unless every unit in the lot is inspected. Since examination of every unit is often economically unfeasible, it is usually necessary to sample a portion of the units in the lot, and to accept or reject the entire lot based on the results. In determining the efficacy of such sampling protocols, the fundamental question is "What is the probability that the sampling procedure will detect an infestation?" The related question "How many samples are needed (to detect an infestation)?" cannot be answered unless some tolerances are specified. If a sampling program is to be used, the inspector must assume some risk and set limits defining "freedom from infestation," *e.g.*, less than 1% of units infested, less than 10 units infested, zero units infested, *etc.* To guarantee no risk and zero units infested, all units must be examined. The objective of the following analysis was to determine the probability of detecting various infestation levels as increasing numbers of samples are collected.

MATHEMATICAL ASSUMPTIONS AND CALCULATIONS

If a proportion of a group or lot of pots, plants, or other units is infested with a pest, and if reliable detection methodology exists, then it is possible to determine the probability of finding an infestation when a specified number of units are inspected. To develop a mathematical approach for determining the probability level, some further assumptions are necessary:

1. The distribution of the pest among units is random, and (or) the sampling of units is random.
2. There is 100% probability of detecting a pest within a unit that is infested. In nematology, this assumption is more easily met for foliar nematode symptoms than for soil sampling, since the latter often does not approach 100% efficiency (4).

Given these assumptions, and since it is possible to characterize units as infested or not, the probability of any combination of infested and non-infested units in a sample can be determined from the binomial or from the hypergeometric distribution (1,6,7). If a proportion (p) of units are infested, then $q = 1 - p$ represents the proportion of pots that are free of the pest. If a sample size of n pots is inspected from a lot size (L) which is large relative to the sample size (n), then the probability that all n units will be free of the pest is given by the term of the binomial expansion, q^n (6). Since q^n represents the probability that none of the n units contains the pest, $1 - q^n$ gives the probability that one or more of the n units will be infested. For example, if 5% of the units in a shipment of $L = 2\ 000$ units are infested with a pest, then the probability of detecting the pest in a sample of $n = 50$ units is given by $1 - (0.95)^{50} = 0.9230$. The calculation is essentially independent of the lot size, as long as L is large relative to n.

If the lot size (L) is not large relative to the sample size (n), then the hypergeometric distribution (1) must be used to calculate the detection probability, and the computation becomes more difficult. The hypergeometric probability of finding zero infested units in a sample of n units from a lot of L units, in which K are infested, is

$$\frac{(L-K)!\,(L-n)!}{L!\,(L-K-n)!}$$

where ! denotes the factorial of the number preceding it (*e.g.*, 5! = 5 × 4 × 3 × 2 × 1 = 120). In the above example, with L = 2 000, K = (0.05) × (2 000) = 100, and n = 50, the exact probability of finding no infested units is 0.0745, which is very close to the approximation $(0.95)^{50}$ = 0.0770 using the binomial distribution. Probabilities of detecting an infested unit for a range of lot sizes and sample sizes were computed for a range of infestation levels, using both the binomial distribution and the hypergeometric distribution.

Probabilities of detecting an infested unit are summarized for five levels of pest infestation (Table 1). Note that the probabilities calculated using the binomial distribution are always less than the exact probabilities calculated using the hypergeometric distribution. Thus, the binomial calculations provide a lower bound to the exact probabilities. Also note that the two probabilities differ at a practical level only for $n/L \geq 0.2$, and that this difference becomes greatest at the lowest infestation levels. If 50% of the pots were infested, it would not be difficult to detect the pest with any of the sampling schemes shown (Table 1). More typical is the need to sample to detect a low infestation level, such as 5%. Within a lot size, the probability of detection increases as the number of units sampled increases. For example, with an infestation level of 5% and a lot size of 1 000, there is a rapid increase in detection as numbers of samples increase up to 100, but there is little gain in detection in sampling more than 100 units (Table 1), even though much additional labor may be required. From a practical standpoint, the probability of detecting a 5% infestation remains the same when 100 units are sampled from lots of 1 000, 10 000, and 100 000 pots, because the proportion of infested units is the same. (The number of infested targets increases from 50 to 500 to 5 000.) According to the binomial probability distribution, when 5% of the units in a lot are infested, a sample size of n = 90 would be needed to provide a 99% chance of detection. When 1% of the units are infested, considerably more samples are needed, approximately 300 samples for a 95% chance of detection, and 500 samples for a 99% chance (Table 1).

The most difficult situation is when only one unit in the entire lot is infested. In this instance, p varies with lot size and is given by $1/L$. When only one unit is infested, detection is extremely unlikely, even with very high numbers of samples (Table 1).

DISCUSSION

For quarantine sampling, the definition of an acceptable level of infestation is fundamental to determining the numbers of samples required. The definition of an acceptable level will vary with inspector and program, since it is a subjective matter and not mathematically determined. Zero will not be an acceptable level, unless resources are available to sample every pot or unit. A level of one unit in a lot is also very laborious to detect. Infestation levels of 5% could be detected 95% of the time with a sample of 59 units or 99% of the time with 90 units. Detection of infestation levels of 1% with similar frequency would require sample sizes of 300–500. Clearly, the probability of detection must be weighed against the cost of the additional labor required.

In addition to increasing the number of samples, detection may be improved by changing the sampling pattern. Maas (3) suggests non-random sampling directed at plants showing symptoms, which may be practical if obvious symptoms occur.

The probability table presented (Table 1) represents a situation in which the

Table 1. Probability of detecting at least one infested unit at each of five levels of infestation, using selected numbers of units sampled per lot. Probabilities were calculated using the binomial (B) and hypergeometric (H) distributions.

Lot size[z] (*L*)	Number of samples (*n*)	Probability of detecting infested unit if the lot sampled is: 50% infested B	50% infested H	10% infested B	10% infested H	5% infested B	5% infested H	1% infested B	1% infested H	1 infested unit B	1 infested unit H
100	5	0.969	0.972	0.410	0.416	0.226	0.230	0.049	0.050	0.049	0.050
100	10	0.999	>0.999	0.651	0.670	0.401	0.416	0.096	0.100	0.096	0.100
100	20	>0.999	>0.999	0.878	0.905	0.642	0.681	0.182	0.200	0.182	0.200
100	30	>0.999	>0.999	0.958	0.977	0.785	0.839	0.260	0.300	0.260	0.300
100	40	>0.999	>0.999	0.985	0.996	0.872	0.928	0.331	0.400	0.331	0.400
100	50	>0.999	>0.999	0.995	>0.999	0.923	0.972	0.395	0.500	0.395	0.500
1 000	10	0.999	>0.999	0.651	0.653	0.401	0.403	0.096	0.096	0.010	0.010
1 000	20	>0.999	>0.999	0.878	0.881	0.642	0.645	0.182	0.184	0.020	0.020
1 000	50	>0.999	>0.999	0.995	0.996	0.923	0.928	0.395	0.403	0.049	0.050
1 000	100	>0.999	>0.999	>0.999	>0.999	0.994	0.996	0.634	0.653	0.095	0.100
1 000	200	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.866	0.894	0.181	0.200
1 000	300	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.951	0.972	0.259	0.300
1 000	400	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.982	0.994	0.330	0.400
1 000	500	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.993	>0.999	0.394	0.500
10 000	10	0.999	0.999	0.651	0.652	0.401	0.401	0.096	0.096	0.001	0.001
10 000	20	>0.999	>0.999	0.878	0.879	0.642	0.642	0.182	0.182	0.002	0.002
10 000	50	>0.999	>0.999	0.995	0.995	0.923	0.923	0.395	0.396	0.005	0.005
10 000	100	>0.999	>0.999	>0.999	>0.999	0.994	0.994	0.634	0.636	0.010	0.010
10 000	200	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.866	0.869	0.020	0.020
10 000	500	>0.999	>0.999	>0.999	> 0.999	>0.999	>0.999	0.993	0.994	0.049	0.050
10 000	1 000	>0.999	>0.999	>0.999	> 0.999	>0.999	>0.999	>0.999	>0.999	0.095	0.100
10 000	2 000	>0.999	>0.999	>0.999	> 0.999	>0.999	>0.999	>0.999	>0.999	0.181	0.200
10 000	5 000	>0.999	>0.999	>0.999	> 0.999	>0.999	>0.999	>0.999	>0.999	0.394	0.500
100 000	100	>0.999	>0.999	>0.999	> 0.999	0.994	0.994	0.634	0.634	0.001	0.001
100 000	1 000	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	0.010	0.010
100 000	10 000	>0.999	>0.999	>0.999	> 0.999	>0.999	>0.999	>0.999	>0.999	0.095	0.100

[z]Total number of units in the lot.

presence of a pest within a pot or unit is readily detected. In nematology, one of the most applicable examples would be the presence of *Aphelenchoides fragariae* (Ritzema-Bos) Christie on susceptible hosts such as hibiscus or chrysanthemum on which foliar symptoms (2,8,9) could be readily observed by non-destructive inspection of each unit. The actual probability of detection could be lower than the tabled values (Table 1) for soilborne nematodes, since the assumption of 100% detection of an infection within a unit may be violated. If only a core or portion of the soil from a pot is removed for analysis, then there is not a 100% certainty of obtaining a nematode from an infested pot if the pot contains only one nematode. This error increases as small cores are used or as pot size increases. For example, a core 15 cm long and 2.5 cm in diameter (73.6 cm^3) represents only 2.78% of a volume 15 cm high × 15 cm in diameter (2 651 cm^3). This source of error is alleviated to some extent by the typical occurrence of large numbers of nematodes, rather than a few individuals, in infested pots. In this above example, the core would pick up one nematode per 73.6 cm^3 of soil if 36 nematodes were present in the pot and evenly distributed. An additional source of error is the fact that, once the soil core is obtained, the methodology for extracting the nematodes often may be only 30–80% efficient (4,5).

Both of these sources of error contribute to the difficulty in recognizing an infested unit. Therefore, Table 1 represents a best-case scenario, which must be adjusted by increasing numbers of samples when these sources of error are considered to be important. If the nematode population density within a unit is relatively high, these errors become less important and the scheme in Table 1 should be adequate. Further research is needed to determine the complicated effects of these sources of error on the probability of detection. Nevertheless, the binomial and hypergeometric probabilities can provide useful guidelines for determining sample numbers, especially when the entire sampled unit can be readily observed, as with foliar diseases or foliar nematodes.

LITERATURE CITED

1. COCHRAN, W. G. 1977. Sampling Techniques. John Wiley and Sons: New York.
2. HAGUE, N. G. M. 1972. Nematode diseases of flower bulbs, glasshouse crops and ornamentals. Pp. 409–434 *in* J. M. Webster, ed. Economic Nematology. Academic Press: New York.
3. MAAS, P. W. T. 1987. Physical methods and quarantine. Pp. 265–291 *in* R. H. Brown and B. R. Kerry, eds. Principles and Practice of Nematode Control in Crops. Academic Press: Sydney.
4. McSORLEY, R. 1987. Extraction of nematodes and sampling methods. Pp. 13–47 *in* R. H. Brown and B. R. Kerry, eds. Principles and Practice of Nematode Control in Crops. Academic Press: Sydney.
5. McSORLEY, R., and J. J. FREDERICK. 1991. Extraction efficiency of *Belonolaimus longicaudatus* from sandy soil. Journal of Nematology 23: 511–518.
6. SNEDECOR, G. W., and W. G. COCHRAN. 1967. Statistical Methods. The Iowa State University Press: Ames, Iowa, U.S.A.
7. SOKAL, R. R., and F. J. ROHLF. 1969. Biometry. W. H. Freeman and Company: San Francisco.
8. STOKES, D. E. 1967. *Aphelenchoides fragariae*, a foliar nematode pathogenic to *Hibiscus rosa-sinensis*. Plant Disease Reporter 51:988–990.
9. STOKES, D. E. 1968. *Aphelenchoides fragariae* infected *Hibiscus rosa-sinensis* in Florida. Plant Disease Reporter 52:118.

Received: *9.VI.1993* *Recibido:*

Accepted for publication: *23.VII.1993* *Aceptado para publicación:*