



RESEARCH ARTICLE

A COST EFFECTIVE METHOD FOR ESTIMATING A RARE DISEASE INFECTION RATES IN PLANTS

*Nyongesa, L. K.

Department of Mathematics, Masinde Muliro University of Science and Technology, Kenya

ARTICLE INFO

Article History:

Received 12th October, 2017
Received in revised form
17th November, 2017
Accepted 03rd December, 2017
Published online 19th January, 2018

ABSTRACT

Group testing involves more than one unit with simple-test for cost effectiveness. The study constructs a maximum likelihood estimator of a rare disease in plants based on group testing framework. The properties of the estimator, bias, MSE, and asymptotic variance are also discussed. The procedure is viable if the group size is relatively small but as the group size increases the estimator is relatively poor.

Key words:

Asymptotic variance, Bias, Group size and MSE.

Copyright © 2017, Nyongesa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Nyongesa, L. K. 2017. "A cost effective method for estimating a rare disease infection rates in plants", *International Journal of Current Research*, 10, (01), 63770-63771.

INTRODUCTION

The idea of sampling plant tissues together herein referred to as a group was proposed by Dorfman (1943) during world war II as an economical method of testing blood specimen of army inductees in order to detect the presence of infection. A population is divided into groups of equal sizes each of size k and a test is performed on each group rather than testing each individual constituents of the group. The main benefit of group testing algorithm is that it reduces the number of tests if the infection rate is low. Dorfman (1943) showed that if the prevalence rate is low then the procedure leads to worthwhile savings of upto 80%. More effort have been put in determining group size, Chiang and Reeves (1962) recommended a group size which gives optimal savings. Recently, Juan and Wenjun (2015) have developed algorithm of determining group size. Tebbs *et al.* (2003) proposed a group testing procedure that is done in an increasing order of probabilities in order to reduce the bias. Hepworth and Watson (2008) investigated the bias of MLE when testing group sizes using fixed and sequential procedures. They were able to correct the bias for fixed procedures. Nyongesa (2012) proposed hierarchical estimation procedures thereby improving the efficiency of the estimators. Plants are planted in rows or columns in farms which may act as groups in group testing algorithms. To this end group testing procedure can be used to estimate the infection rate of disease if it occurs. Therefore the purpose of this study is to propose a method of estimating disease transmission rates in plants with the use of Dorfman (1943) group testing procedure.

The Model

Suppose we have a large number of plant tissues say N where $N \rightarrow \infty$ we are interested in estimating the infection rate of a rare disease in these N plants. For simplicity, we denote the infection rate by H . The obvious method is to test each of the N plant tissues. Since $N \rightarrow \infty$, one at a time testing is time consuming, costly, tedious, and may result into poor estimator of H . Therefore we apply group testing algorithm. Split N into n homogeneous groups each of size k for simplicity the model of interest is

$$f(X; H) = \begin{cases} \binom{n}{x} (1 - (1 - H)^k)^x ((1 - H)^k)^{n-x}, & x = 0, 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where x is the number of groups that are positive. This model is of interest.

Estimation of infection rate

The binomial model (1) will provide the distribution of the groups that test positive on test. The likelihood function is

$$L(H | X) = \binom{n}{x} (1 - (1 - H)^k)^x ((1 - H)^k)^{n-x} \quad (2)$$

from which the infection rate is obtained as

$$\hat{H} = 1 - \left(1 - \frac{x}{n}\right)^{\frac{1}{k}} \quad (3)$$

*Corresponding author: Nyongesa, L. K.

Department of Mathematics, Masinde Muliro University of Science and Technology, Kenya.

Table 1. Asymptotic variance for various group sizes k and infection rate H

H	K				
	1	5	10	15	20
0.01	0.000495	0.000101	0.000052	0.000035	0.000027
0.02	0.000980	0.000204	0.000108	0.000076	0.000060
0.03	0.001455	0.000310	0.000168	0.000121	0.000099
0.04	0.00192	0.000417	0.000232	0.000173	0.000145

Table 2. Bias and MSE for various group sizes

H	Bias					MSE				
	1	5	10	15	20	1	5	10	15	20
0.01	0.000	-0.0080	-0.0090	-0.009	-0.0095	0.000	0.000	0.0001	0.0001	0.0001
0.02	0.000	-0.0161	-0.0181	-0.018	-0.0190	0.001	0.000	0.0004	0.0004	0.0004
0.03	0.000	-0.0242	-0.0271	-0.028	-0.0286	0.001	0.000	0.0009	0.0009	0.0009
0.04	0.000	-0.0323	-0.0362	-0.037	-0.0381	0.001	0.001	0.0015	0.0015	0.0016

The asymptotic variance of the estimator (3) is

$$Var(\hat{H}) = \frac{1-(1-H)^k}{(1-H)^{k-2}} \tag{4}$$

We compute the estimator of H when k is fixed but this might not be the case in practice. To demonstrate our computation, we simulate the asymptotic variance of \hat{H} for various values of k .

From Table 1, it is clear that as the prevalence rate H increases the variance also increases. This is true in practice because the model is only viable when the prevalence rate is small, Dorfman (1943). The variance decreases with increase in k when H is held constant. A worthwhile saving is achieved when relatively small groups are used.

Bias and Mean Square Error

The bias of the estimator \hat{H} measures the accuracy of the estimator and is determined by

$$Bias(\hat{H}) = E(\hat{H}) - H \tag{5}$$

To compute (5), we require the expected value of \hat{H} , hence

$$E(\hat{H}) = 1 - E\left(1 - \frac{X}{H}\right)^{\frac{1}{k}} \tag{6}$$

Implying that for $k = 1$, the estimator is unbiased Nyongesa (2012) and for $k > 1$, the estimator is biased. The bias can be shown to be

$$Bias(\hat{H}) = o(k^{-1}) + o(n^{-2}) \tag{7}$$

Therefore we are in a position to provide the MSE of the estimator \hat{H} is given by

$$MSE(\hat{H}) = \frac{1-(1-H)^k}{nk(1-H)^{k-2}} + o(k^{-1}) + o(n^{-2}) \tag{8}$$

Next, we provide some simulations of the MSE and Bias. The results are provided in Table 2 above;

From the Table (2), it is clear that for $k = 1$, the estimator \hat{H} is unbiased estimator for H . The bias increases with increase in k and vice versa. For the MSE, it reduces with increase in group size. We note that the group size 5 provides optimal results although optimal group sizes can be obtained by minimizing the asymptotic variance for instance see Juan and Wenjun (2015).

Conclusion

In this study we have constructed a maximum likelihood estimator of a rare disease infection rate in plants. The properties of the estimator such as bias, MSE, and asymptotic variance obtained. It was observed that a relative group size of S would provide optimal results. This concurs with Shallow (1985) who recommended that relatively small group sizes should be used to obtain optimal results. Also, there tends to be a loss of sensitivity in group testing procedures if relatively high groups are employed cf. Nyongesa (2012).

REFERENCES

Dorfman, R. 1943. The detection of defective members of large population. *Annals of Mathematical Statistics*, 14, 436-440.

Hepworth, G. and Watson, R. 2009. Debaised Estimation of Proportions in Group Testing. *Journal of the Royal Statistical Society, Applied Statistics*, 58, 105-121.

Juan, D. and Wenjun, X. 2015. Robust group testing for multiple traits with misclassifications. *Journal of Applied Statistics*, 42(10), 2115-2015.

Nyongesa L. K. 2012. Estimating the rate of defectiveness by Group Testing in the presence of test errors. *Journal of Statistical Science*, 12(1), 1-6.

Swallow W.H. 1985. Group testing for estimating infection rates and probability of disease transmission. *Phytopathology*, 75, 882-889. 18.

Tebbs M.J. and Swallow, H.W. 2003. Estimating ordered binomial proportions with the use of group testing. *Biometrika*, 90, 471-477.
