NOTE ON "STATISTICAL PRINCIPLES OF ROUTINE WORK IN TESTING CLOVER SEED FOR DODDER"

BY M. SANKARAN

University of Madras

PRZYBOROWSKI AND WILENSKI (1935) considered the application of Neyman-Pearson theory of testing of hypotheses about the true mean content of dodder seeds in a sample, assuming that the occurrence of dodder seeds follows the Poisson law. For a justification of the law and other details the reader is referred to their paper. The testing of hypotheses arises as follows: A lot is accepted if the true mean content of the lot is less than or equal to m; otherwise it is rejected. A sample is taken and tested whether the number of seeds in the sample exceeds or falls below an acceptance number a. If it is less than or equal to a, the hypothesis that the true mean content is m, is accepted and hence the lot is accepted.

In any hypothesis testing we are likely to commit two kinds of error: we may reject a true hypothesis or we may accept a false hypothesis. Ordinarily we fix one type of error and minimise the other type of error. But this may unnecessarily penalise either the buyer or the seller in small samples. It is only this consideration that led Przyborowski and Wilenski to conclude that in large samples only it is possible to protect the interests of both the buyer and the seller; but this means an expensive analysis.

In this note we will be concerned only with the case of designing a proper test rule that may safeguard the interests of the buyer and seller in small samples. Throughout the paper, we shall assume that the variate under consideration follows the Poisson law given by $e^{-m}m^x/x!$.

2. In the case of small samples, the test which is employed as a basis for a best rule is that the sum of the probabilities of accepting seeds of true quality m_1 (which we may call acceptable quality) and rejecting seeds of true quality m_2 (which may be termed objectionable quality) be a maximum. Mathematically it amounts to finding an acceptance number a such that

$$P = Pr_{m_1}(A) + Pr_{m_2}(R) \tag{1}$$

is maximised. A similar approach for the binomial case was adopted by A. Golub (1953).

Now

$$Pr_{m_1}(A) = \sum_{x=0}^{a} \frac{e^{-m_1} m_1^x}{x!}$$

and

$$Pr_{m_2}(R) = 1 - \sum_{x=0}^{a} \frac{e^{-m_2} m_2^x}{x!}$$

For a to be the acceptance number such that (1) is maximum [provided (1) has only one maximum], the following relationships must be satisfied:

$$P(a) - P(a-1) > 0, \tag{2}$$

$$P(a+1) - P(a) < 0. \tag{3}$$

(2) and (3) reduce to

$$e^{-m_1} m_1^a - e^{-m_2} m_2^a > 0, \tag{4}$$

$$e^{-m_1} m_1^{a+1} - e^{-m_2} m_2^{a+1} < 0. \tag{5}$$

Taking logarithms and simplifying (4) and (5) it is easily seen that a is an integer satisfying (under the assumption $m_2 > m_1$)

$$\frac{(m_2 - m_1)}{\log \frac{m_2}{m_1}} - 1 < a < \frac{(m_2 - m_1)}{\log \frac{m_2}{m_1}}. \tag{6}$$

So a can be taken to be

$$a = -\tfrac{1}{2} + \frac{(m_2 - m_1)}{\log \frac{m_2}{m_1}}, \tag{7}$$

and rounded to the nearest integer. (7) can be also derived by the alternate approach considered by Golub (1953, pp. 281).

3. For the sake of illustration we have provided a table on next page giving the acceptance number a for a few values of m_1 and m_2.

Table giving the acceptance number a *for certain values of* m_1 *and* m_2 ($m_2 > m_1$). *For other values of* m_1 *and* m_2 a *can be obtained from formula* (7)

m_1 \ m_2	2	3	4	5	6	7	8	9	10	15	20
1	1	1									
2		2	2	3	3						
3			3	3	4	4	5				
4				4	4	5	5	6	6	8	
5					5	5	6	6	7	9	
6						6	6	7	7	9	
7							7	7	8	10	
8								8	8	11	13
9										11	13
10										12	14
11										12	15
12										13	15
13											16
14											16
15											17
16											17

In conclusion, my grateful thanks are due to the Government of India for the award of a scholarship.

REFERENCES

1. Golub, A. .. *Jour. Amer. Stat. Assn.*, 1953, **48**, 278.
2. Przyborowski, J. and Wilenski, H. *Biometrika*, 1935, **27**, 273.