On Statistical Comparison of Efficiencies of Fishing Gear

A. K. KESAVAN NAIR

Central Institute of Fisheries Technology, Cochin--682 029

The distribution of fish caught by experimental gill nets has been found to be in the Poisson or Negative binomial form. Using this information, application of Chi-square test as suggested by Mood *et al.* (1974) has been illustrated, for comparing the efficiencies of gill nets. This test provides an alternative to Anova F-test especially in the context of significance of nonadditivity for the two-way model. Based on the present work and the findings by Nair (1982) and Nair & Alagaraja (1982, 1984) an outline approach for statistical comparison of the efficiencies of fishing gear is presented.

A study of the distribution of the data is important in view of developing test procedures. If the form of the distribution is known, that information could be used to construct a test to compare the location. With this in view the gill–net catch data were examined. This assumes importance because of the fact that non-additivity in the two-way model was found to be present when the experimental data for comparing the efficiencies of gill-nets were examined. Further, Nair (1982) and Nair & Alagaraja (1982, 1984) have investigated the applicability of some tests to compare the efficiencies of trawl nets. The difficulties caused by lack of satisfaction of relevant assumptions for applying parametric tests and some approaches to obviate some of these difficulties are discussed by them. It is also the purpose of this communication to use these findings along with the present work to suggest an outline for a practical approach for the statistical comparison of the efficiencies within trawl nets and within gill-nets.

Materials and Methods

Data on catches of different types of gill-nets, for instance (Kunjipalu *et al.,* 1984) obtained under comparable conditions for different days were used to compare the efficiencies. Frequency distribution of the numbers of fish caught according to the frequency (in terms of number of hauls) of occurrence of 0, 1, 2 etc fish in the catch was made for different types of gill–nets. As the largest frequencies corresponded to occurrence of 0 or 1 fish and the frequencies decrease for increasing numbers of fish, the Poisson, Negative binomial and Geometric distributions were considered for the data. The theoretical frequencies were calculated using the densities,

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} I_{(0,1,\ldots)}(x) \quad \text{(Poisson)} \qquad (1)$$

$$f(x) = \binom{r+x-1}{x} p^{r} q^{x} I_{(0,1,\ldots)}(x); \quad 0 < p \leq 1,\ r > 0 \qquad (2)$$

$(q = 1 - p)$ (Negative binomial)

and $f(x) = pq^{x} I_{(0,1,\ldots)}(x); 0<p\leq 1;$ $(q=1-p)$ (Geometric) (3)

as given by (Mood et al., 1974). λ, p, q ($=1-p$) and r are parameters of the distributions. The goodness of fit was tested by chi-square. Further, the chi-square test (Mood *et al.,* 1974),

$$Q_{2k} = \sum_{i=1}^{2} \sum_{j=1}^{k+1} \frac{(N_{ij} - n_i p_j)^2}{n_i p_j} \qquad (4)$$

with degrees of freedom equal to (2k-the number of parameters estimated) was used to test whether two given samples are drawn from the same population such as the Poisson, the Negative binomial or the Gamma. Here k + 1 refers to the number of classes and i=1 and 2 for two samples.

On the basis of this and the results obtained in the studies mentioned above, an approach for the statistical comparison of the efficiencies within trawl nets and within gill–nets is listed. Information obtained by applying Quade (1979) test and rank transform test (Lemmer & Stocker, 1967; Conover & Iman, 1976; Hora & Conover, 1984) as presented in Iman *et al.* 1984, to data on trawl and gill–net catches has also been utilized to indicate this approach.

Table 1. *Distribution of fish caught by 4 gill–nets A, B, C and D*

A		B	
No. of fish caught	Frequency	No. of fish caught	Frequency
0	27	0	13
1	26	1	18
2	14	2	24
3	4	3	9
4	3	4	5
5	–	5	2
6	1	6	3
7	–	7	1
8	–	8	–
9	1	9	–
10	–	10	1
11	–	11	1
12	1		
Total	77	Total	77
C		D	
0	9	0	6
1	9	1	12
2	6	2	11
3	3	3	4
4	5	4	3
5	2	5	1
6	–	6	–
7	–	7	1
8	3	8	2
9	1	9	1
10	–	10	–
11	1	11	–
12	1	12	–
13	1	13	–
Total	41	Total	41

Results and Discussion

The frequency distribution of the numbers of fish caught by gill–nets A, B, C and D (per equal area) are presented in Table 1. (The frequencies are the number of operations of equal duration). The maximum frequencies correspond to the occurrences of 0 or 1 fish in the catch and the frequencies decrease with occurrences of increasing numbers of fish, as already mentioned. The comparisons made were between A and B and between C and D. The Poisson, Negative binomial and Geometric distributions fitted to these data alongwith the observed frequencies are presented in Table 2. The chi-square values with the respective degrees of freedom for the goodness of tests are also presented in Table 2. It can be seen from Table 2 (A and B) without any test itself that the geometric distribution does not fit any set of the data. Therefore, this distribution was not fitted for sets C and D. The chi-square goodness of fit for Poisson and Negative binomial distributions as presented by Mood *et al.* (1974) showed Poisson and negative binomial to be a good fit for the sets A and B and Negative binomial for sets C and D (Table 2). Poisson distribution was however found to be satisfactory for set D, though not for set C. From the mean and variance presented in Table 2, it can be seen that they are not widely different for sets A, B and D, so that Poisson distribution too fitted these data. But for set C, variance is very much larger than the mean, which made the Poisson distribution, a poor fit. Negative binomial distribution fitted all the four sets of data. However, for any of these distributions, the chi-square test as given by (4) can be used to test whether the samples came from the same Poisson or Negative binomial populations (Mood *et al.,* 1974).

The application of this test for the two distributions, that is, to test whether sets A and B came from the same Poisson distribution and sets C and D came from the same Negative binomial distribution is illustrated below.

(1) Comparison of gear A and B:

Frequency distribution of the number of fish caught by the two nets is

No. of fish	0	1	2	3	4 or more	Total
Net A	27	26	14	4	5	76
Net B	13	18	24	9	12	76
Total	40	44	38	13	17	152

Table 2. *Fit of Poisson, negative binomial and geometric distributions to the data given in Table 1*

A

No. of fish caught	Observed frequency	Poisson	Negative binomial	Geometric
0	27	22	31	41
1	26	27	21	19
2	14	17	12	9
3	4	7	6	4
4	3	2	3	2
5	–			
6	1			
7	–	1	3	1
8	–			
9	1			
10	–			
Total	76	76	76	76

Mean = 1.22, Variance = 2.25
Test for goodness of fit (chi-square)

		1.80 N.S	2.87 N.S.
d.f.		2	2

(Frequencies in classes 3 and above were pooled for Poisson and 4 and above for Negative binomial, to computer chi-square)

B

No. of fish caught	Observed frequency	Poisson	Negative binomial	Geometric
0	13	10	15	46
1	18	20	19	18
2	24	20	16	7
3	9	14	11	3
4	5	7	7	1
5	2	3	4	
6	3		2	
7	1			
8				
9		2	2	1
10	1			
Total	76	76	76	76

Mean = 2,08., Variance = 3.38
Test for goodness of fit chi-square

		5.06 N.S	5.38 N.S
d.f		4	3

(Frequencies in classes 5 and above were pooled to compute chi-square)

C

No. of fish caught	Observed frequency	Poisson	Negative binomial
0	9	1.76	9.37
1	9	5.55	7.55
2	6	8.73	5.85
3	3	9.15	4.47
4	5	7.20	3.40
5	2		
6	–		
7	–		
8	3		
9	1	8.61	10.36
10	–		
11	1		
12	1		
13	1		
Total	41	41	41

Mean = 3.15, Variance = 12.28
Goodness of fit (chi-square)

		37.92** ($p < 0.005$)	1.71 N.S
d.f		4	3

D

No. of fish caught	Observed frequency	Poisson	Negative binomial
0	6	8.91	4.04
1	12	9.43	9.36
2	11	7.55	10.85
3	4	5.40	8.38
4	3	3.63	4.85
5	1		
6	–		
7	1		
8	2	6.08	3.52
9	1		
10	–		
11	–		
12	–		
13	–		
Total	41	41	41

Mean = 2.32, Variance = 5.07
Goodness of fit (chi-square) 5.31N.S 3.89N.S

d.f		4	3

N.S – Not significant; ** –Highly significant; d.f – degrees of freedom

Now, the parameter, namely, the mean of the Poisson population is to be estimated. The maximum likelihood estimate of the sample mean is

$$\frac{0(40) + 1(44) + 2(38) + 3(13) + 4(8) + 5(2) + 6(4) + 7(1) + 9(1) + 10(1)}{152} = 1.6513$$

From (1), the expected number in each group of the population is given by

No. of fish	0	1	2	3	4 or more
Expected number	14.58	24.07	19.95	10.97	6.43

Thus the chi-square given by Mood, Graybill and Boes (1974) is

$$Q_{2k} = \sum_{i=1}^{2} \sum_{j=1}^{k+1} \frac{(Nij - nipj)^2}{nipj} = \frac{(27-14.58)^2}{14.58} + .. + \frac{(12-6.43)^2}{6.43}$$

$= 24.96^{**}$ with $2k-1 = 8-1 = 7$ degrees of freedom (as 1 parameter is estimated). The significance of the chi-square shows that the two samples are not from the same population which means that the catches by the two gear are not equal. This can be generalized to several samples, that is, to catches by more than two gear also.

(2) Comparison of gear C and D:

Assuming that the catches by nets C and D are distributed in the Negative binomial form (as found already) whether they came from the same Negative binomial distribution is tested by the chi-square test discussed and applied above. Frequency distributions of the number of fish caught by the two nets are as under.

No. of fish	0	1	2	3	4	5 and above	Total
Net C	9	9	6	3	5	9	41
Net D	6	12	11	4	3	5	41
Total	15	21	17	7	8	14	82

Here two parameters r and p are to be estimated from the combined data. Estimation of these parameters by the method of moments ($\hat{p} = \frac{\text{mean}}{\text{variance}}$ $\hat{r} = \text{mean} \times \frac{\hat{p}}{\hat{q}}$), gave $\hat{p} = 0.3125$, $\hat{q} = 0.6875$ and $\hat{r} = 1.2427$. Thus from (2), the expected number in each group of the population is given by

No. of fish	0	1	2	3 and above	Total
Expected number	9.67	8.26	6.36	16.71	41

(Frequencies in the last three classes were pooled to form a single class '3 and above' to make all the expected values greater than 5, for computing chi-square)

$$Q_{2k} = \frac{(9-9.67)^2}{9.67} + .. + \frac{(17-16.71)^2}{16.71} +$$

$$\frac{(6-9.67)^2}{9.67} + .. + \frac{(12-16.71)^2}{16.71}$$

$= 7.94$ (N.S.) with $2k-2$, that is, 4 d.f., as two parameters are estimated.

Thus the hypothesis that the two catches come from the same population is not rejected.

The above illustrations show that a test based on the distribution of fish catch data (for gill-nets) can be constructed. The distribution has been found to be either Poisson or negative binomial. Negative binomial fitted three sets out of the four when all the observations were considered and the same distribution fitted all the four sets when one observation in the extreme class after some discontinuity was omitted. Poisson distribution fitted 3 sets with and without the omission of the observation in the extreme class. Geographical and species difference may attribute to the difference in the distribution. However fitting Poisson or negative binomial is easy and can be tried for any set of gill-net data. Depending on the adequacy of the fit either of these distributions may be assumed and the difference between the samples tested by employing chi-square test. But it is important to test the goodness of fit because, when the fit is not adequate, that itself will contribute to the significance of chisquare, vitiating the results of the test for difference of the two samples.

An outline of approach for statistical comparison:

(i) When the efficiency of two trawl nets or two gill-nets are to be compared, Wilcoxon matched-pairs signed-rank test (WSR test) may be used, as this has been found to be more efficient for the data. Also, its application is simple. For normal distributions this test is 95.5 percent as efficient as the parameteric F or t-test (Siegel, 1956) but, for other types of distribution (for instance, some long tailed ones) this test may be more than 100% efficient compared to the F or t-test (Snedecor & Cochran, 1968). The superiority of WSR test over F-test for trawl catches has been demonstrated by Nair & Alagaraja (1982) and the test has been applied in Narayanappa *et al.* (1982). Moreover, the nonnormality of the data has been indicated by Nair (1982) as revealed by the dependence of the mean on the variance. Lack of satisfaction of other assumptions like nonadditivity for ANOVA, has also been established by applying Tukey's test and the presence of outliers have been observed by Nair & Alagaraja (1984). Finally among the nonparamteric methods for paired comparisons, except for randomization test, only Wilcoxon test seems to use interblock information. But randomization test is unwieldy for even moderately large samples (say, when the number of pairs exceeds 12) and as Siegel (1956) has observed, Wilcoxon test (WSR test) is a very efficient alternative to the randomization test because it is a randomization test on ranks.

(ii) When the efficiency of more than two trawl nets are to be compared, Friedman test and ordinary ANOVA F-test may be tried first. Applications showed Friedman test to be as sensitive as F-test, though no higher sensitivity was observed in any case. As Friedman test depends on fewer assumptions than does F-test, as a practical procedure, if either of these tests brings out the difference in the efficiency, there is no need to test further. If both the tests are not found to be sensitive and if the probability for an observed difference is close to the significance level, the Quade (1979) test and if still inconclusive the combination procedure as demonstrated in Nair & Alagaraja (1984) may be applied. The latter, though not simple, may bring out the real difference, if any, in this case. Recently, Iman *et al.* (1984), while making a comparison of Friedman test, Quade (1979) test and rank transform test (Lemmer & Stoker, 1967; Conover & Iman, 1976; Hora & Conover, 1984) found Quade test to be a better choice than Friedman test for normal data for the number of treatments, $k \leq 6$ and vice versa for $k > 6$. For the nonnormal settings the result favoured the Quade test for uniform case and lognormal case (when $k = 3$), while Friedman test showed more power than the Quade test in the remaining 11 of the 16 nonnormal cases, they examined. They found Quade test to be favourable for light tailed uniform distributions while the Friedman test and the rank transform test for heavy tailed distributions. Application of Quade test and rank transform test to trawl catches showed the same result as when Friedman test was applied. However, Friedman and Quade tests showed more or less the same sensitivity but rank transform test showed a little less sensitivity.

(iii) When the efficiency of more than two gill-nets are to be compared, Friedman test and ordinary ANOVA F-test may be used. Friedman test helps to confirm the result as its applicability for the data is more valid and as applications (Kunjipalu *et al.*, 1984) have shown Friedman test to be as sensitive as F-test. The performance of Quade test, Friedman test and rank transform test were compared for gill-net catches too. All the tests showed the same result. However, Quade test and rank transform test showed a little more sensitivtiy than Friedman test. Therefore it is advisable to apply Quade test and rank transform test when the probability for an observed difference is close to the significance level. Another alternative to confirm the results would be the test illustrated in this paper. Fitting of the Poisson or negative binomial for this purpose is simple. So also the application of chi-square test for goodness of fit and then for testing equality of samples from the same Poisson or same negative binomial populations. In fact this test can be applied to compare the efficiencies of two or more gill–nets.

The author is grateful to Shri M. Rajendranathan Nair, Director, Central Institute of Fisheries Technology, Cochin for permission to publish this paper, to Shri M.K. Kandoran, Scientist-in-Charge of Extension, Information and Statistics Division for encouragement and Dr. K. Alagaraja of Central Marine Fisheries Research Institute and S/Shri H. Krishna Iyer and K. Krishna Rao of Central Institute of Fisheries Technology for critical comments on the manuscript.

References

Conover, W.J. & Iman, R.L. (1976) *Commun. Statist. Theor. Meth.*, A5 (14), 1349

Hora, S.C. & Conover, W.J. (1984) *J. Amer. Statist. Assoc.* **79**, 387

Iman, R.L., Hora, S.C. & Conover, W.J. (1984) *J. Amer. Statist. Assoc.* **79**, 387

Kunjipalu, K.K., Boopendranath, M.R., Kuttappan, A.C., Subramonya Pillai, N. Gopalakrishnan, K. & Nair, A.K.K. (1984) *Fish. Technol.* **21**, 51

Lemmer, H.H. & Stoker, D.J. (1967) *South African Statistical J.* **1**, 67

Mood, A.M., Graybill, F.A. & Boes, D.C. (1974) *Introduction to the Theory of Statistics*. III Edn, McGraw Hill International Book Co., Tokyo

Nair, A.K.Kesavan (1982) *Fish. Technol.* **19**, 49

Nair, A.K.Kesavan & Alagaraja, K. (1982) *Proc. Symp. Harvest and Post-harvest Technology of Fish Organised by Society of Fisheries Technologists (India)*, Nov. 24–27, 1982

Nair, A.K.Kesavan & Alagaraja, K. (1984) *Fish. Technol.* **21**

Narayanappa, G., Ramarao, S.V.S., Sitaraman Rao, J., Naidu, R.M. & Satyanarayana, A.V.V. (1982) *Proc. Symp. Harvest and Post-harvest Technology of Fish organised by Society of Fisheries Technologists (India)*, Nov. 24–27, 1982

Quade, D. (1979) *J. Amer. Statist. Assoc.* **74**, 680

Siegal, S. (1956) *Nonparametric Statistics for the Behavioural Sciences*, McGraw Hill Book Co. Inc. Japan

Snedecor, G.W. & Cochran, W.G. (1968) *Statistical Methods*. 6th edn. Oxford and IBH Publishing Company, New Delhi