USDA FOREST SERVICE RESEARCH NOTE NE-116

1970

N

FOREST SERVICE, U. S. DEPT. OF AGRICULTURE, 6816 MARKET STREET, UPPER DARBY, PA.

E

SEQUAN: A COMPUTER PROGRAM FOR SEQUENTIAL ANALYSIS

Abstract.—A description of a FORTRAN IV computer program for performing sequential analysis on four common distributions after the underlying probability distribution is known.

Decision-making and cost-reduction are key concepts in management research. Sequential hypothesis-testing procedures that were developed for quality control in industrial processes (*2, 3, 5*) have been applied in a limited way to biological problems. Forest entomologists have used the sequential probability ratio test of Sobel and Wald (2) to classify insect populations (*4*) and to evaluate the need for insect control (*1*). This note describes briefly a computer program that was developed to eliminate much of the computational work associated with sequential analysis.

Sequential analysis is popular with forest entomologists because, on the average, it requires fewer observations for fixed type-1 (α) and type-2 (β) errors than fixed-sample size plans require. The smaller the difference to be detected at a given type-1 (α) and type-2 (β) error, the larger the average sample size. In this method of analysis we assume that the observations are independent and the probability distribution is known beforehand.

To perform any sequential analysis, the following must be known:

1. A description of the underlying probability distribution.
2. A statement of the hypothesis that is to be tested.
3. The type-1 (α) and type-2 (β) errors.

The computer program was developed for the three-decision problem (light, medium, heavy) that is commonly encountered by forest entomologists. By repeating selected input data in certain columns of the second control card, the program can be used to obtain the required statistics for the two-decision problem (control *vs* no control).

The FORTRAN IV program was developed by the authors on a 7094 computer. The program will provide the following statistics for the binomial, negative binomial, poisson, and normal distributions:

1. Equations for the decision lines.
2. Table of decision boundaries.
3. Graph of decision lines.
4. List of selected points for the operating-characteristics (OC) curve.
5. List of selected points for the average-sample-number (ASN) curve.

For the normal and negative binomial distributions, it is assumed that the variance and k-value are known beforehand.

Description of Control Deck

The user must supply the following control cards when using the program. Two control cards are required for each analysis desired. The first control card contains the type-1 (α) and type-2 (β) errors, the distribution index, and the number of decision points to be printed.

Card 1	*Column*	*Format*	
	1 - 3	F3.2	Type-1 error, A
	4 - 6	F3.2	Type-2 error, B
	7 - 8	I2	Distribution index
		01 - Binomial	
		02 - Negative binomial	
		03 - Poisson	
		04 - Normal	
	9 - 11	I3	Number of decision points to be printed

The second control card contains the parameters needed for the analysis. The format of this control card depends upon the distribution. Binomial distribution:

H_0: $P<P_1$

H_1: $P_2 \leq P \leq P_3$

H_2: $P>P_3$

Column	*Format*	*Description*
1 - 8	F8.4	P_1
9 - 16	F8.4	P_2
17 - 24	F8.4	P_3
25 - 32	F8.4	P_4

Negative binomial:

H_0: $M < M_1$
H_1: $M_2 \leq M \leq M_3$
H_2: $M > M_4$

Column	*Format*	*Description*
1 - 8	F8.4	M_1
9 - 16	F8.4	M_2
17 - 24	F8.4	M_3
25 - 32	F8.4	M_4
33 - 40	F8.4	K - "Index of aggregation"

Poisson:

H_0: $\lambda < \lambda_1$
H_1: $\lambda_2 \leq \lambda \leq \lambda_3$
H_2: $\lambda > \lambda_4$

Column	*Format*	*Description*
1 - 8	F8.4	λ_1
9 - 16	F8.4	λ_2
17 - 24	F8.4	λ_3
25 - 32	F8.4	λ_4

Normal:

H_0: $\mu < \mu_1$
H_1: $\mu_2 \leq \mu \leq \mu_3$
H_2: $\mu > \mu_4$

Column	*Format*	*Description*
1 - 8	F8.4	μ_1
9 - 16	F8.4	μ_2
17 - 24	F8.4	μ_3
25 - 32	F8.4	μ_4
33 - 42	F10.4	σ - standard deviation

Example

Suppose we wish to calculate the sequential plan described by Waters (*4*) for the spruce budworm. The basic data for this plan were fitted to the negative binomial distribution. Waters listed the following parameters for his plan:

$\alpha = 0.10$ $\qquad$ $\beta = 0.10$
$M_1 = 3.00$ $\qquad$ $M_2 = 6.00$
$M_3 = 9.00$ $\qquad$ $M_4 = 12.00$

$k = 7.288$

Figure 1.—Program output of decision equations and sequential table.

```
SEQUAN        ALPHA = 0.100        BETA = 0.100

NEG BIN    P1 = 0.415    P2 = 0.830    P3 = 1.245    P4 = 1.660    K = 7.228

CASE 1        P0 VS P1        P0 .GT. P1        LOW VS MEDIUM

LOWER =  -5.040+     4.265 N

UPPER =   5.040+     4.265 N

CASE 2        P0 VS P2        P2 .GT. P0        MEDIUM VS  HIGH

LOWER = -18.612+    10.386 N

UPPER =  18.612+    10.386 N
```

N	LOWER: LOW VS MEDIUM — LOW	LOWER: LOW VS MEDIUM — HIGH	UPPER: MEDIUM VS HIGH — LOW	UPPER: MEDIUM VS HIGH — HIGH
1	-0	9	-7	29
2	3	14	2	39
3	8	18	13	50
4	12	22	23	60
5	16	26	33	71
6	21	31	44	81
7	25	35	54	91
8	29	39	64	102
9	33	43	75	112
10	38	48	85	122
11	42	52	96	133
12	46	56	106	143
13	50	60	116	154
14	55	65	127	164
15	59	69	137	174
16	63	73	148	185
17	67	78	158	195
18	72	82	168	206
19	76	86	179	216
20	80	90	189	226
21	85	95	199	237
22	89	99	210	247
23	93	103	220	257
24	97	107	231	268
25	102	112	241	278

Figure 2.—Graph of decision line equations.

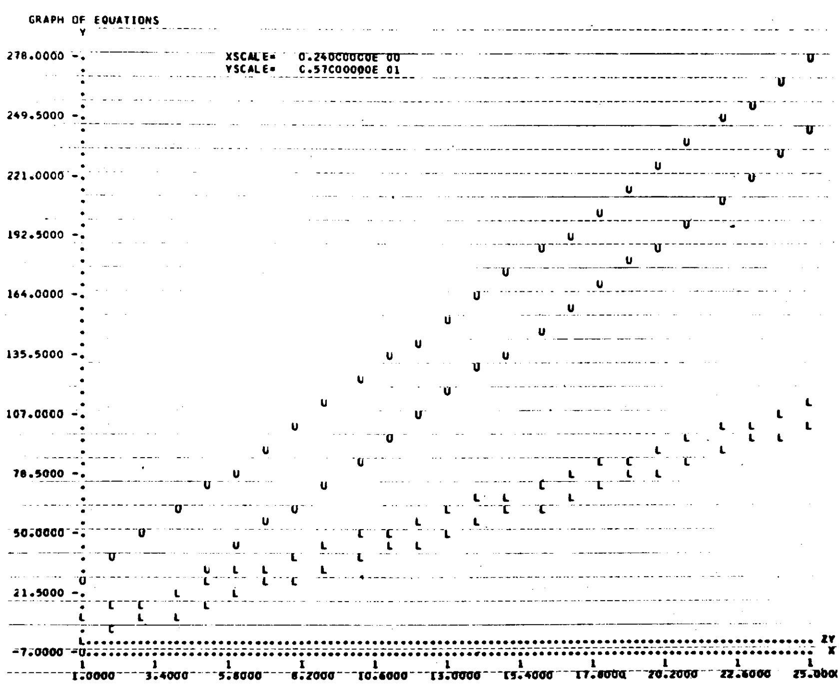

TABLE OF VALUES FOR THE OC AND ASN CURVES-- NEG BIN

CASE 1			CASE 2		
LOW VS MEDIUM			MEDIUM VS HIGH		
P	OC	ASN	P	OC	ASN
0.595	0.486	3.730	1.432	0.514	13.707
0.600	0.473	3.713	1.427	0.527	13.726
0.606	0.459	3.694	1.421	0.541	13.734
0.611	0.445	3.673	1.416	0.555	13.736
0.621	0.418	3.626	1.406	0.582	13.720
0.632	0.392	3.572	1.396	0.608	13.678
0.643	0.366	3.513	1.386	0.634	13.609
0.654	0.341	3.448	1.376	0.659	13.516
0.666	0.317	3.379	1.366	0.683	13.401
0.677	0.293	3.305	1.357	0.707	13.265
0.701	0.250	3.149	1.337	0.750	12.939
0.725	0.211	2.984	1.318	0.789	12.554
0.750	0.177	2.815	1.300	0.823	12.128
0.776	0.147	2.647	1.281	0.853	11.675
0.803	0.122	2.482	1.263	0.878	11.210
0.830	0.100	2.323	1.245	0.900	10.744
1.156	0.012	1.202	1.080	0.988	7.046
2.179	0.000	0.439	0.816	1.000	4.148

MEAN 1 = 3.0000 OC-VALUE = 0.900 MEAN 2 = 6.0000 OC-VALUE = 0.100
MEAN 3 = 9.0000 OC-VALUE = 0.900 MEAN 4 = 12.0000 OC-VALUE = 0.100

Figure 3.—Program output of OC and ASN values.

From this information the control deck would be prepared as follows:

COLUMN NO.

Card No. 1 2 3 4 5 6 7 8 9¹ 0 1 2 3 4 5 6 7 8 9² 0 1 2 3 4 5 6 7 8 9³ 0 1 2 3 4 5 6 7 8 9⁴ 0

1 0 1 0 0 1 0 0 2 0 2 5

2 0 0 0 3 0 0 0 0 0 0 0 6 0 0 0 0 0 0 0 9 0 0 0 0 0 0 1 2 0 0 0 0 0 0 0 7 2 2 8 0

Figures 1, 2, and 3 show the sequential plan produced by this program. The decision-line equations produced by this program differ slightly from those published by Waters, but these differences are probably due to rounding errors. The occasional discrepancies between values in the sequential tables are also the results of rounding errors.

A program deck with examples can be obtained from the authors.

Literature Cited

1. Ives, W. G. H., and G. L. Warren.
   1965. SEQUENTIAL SAMPLING FOR WHITE GRUBS. Can. Entomol. 97: 596-604.

2. Sobel, M., and A. Wald.
   1949. A SEQUENTIAL DECISION PROCEDURE FOR CHOOSING ONE OF THREE HYPOTHESES CONCERNING THE UNKNOWN MEAN OF A NORMAL DISTRIBUTION. Ann. Math. Statist. 20: 502-522.

3. Wald, A.
   1947. SEQUENTIAL ANALYSIS. 212 pp. J. Wiley & Sons, Inc., New York.

4. Waters, W. E.
   1955. SEQUENTIAL SAMPLING IN FOREST INSECT SURVEYS. Forest Sci. 1: 68-79.

5. Wetherill, G. B.
   1966. SEQUENTIAL METHODS IN STATISTICS. 218 pp. Methuen & Co., Ltd., London; J. Wiley & Sons, Inc., New York.

—ROBERT L. TALERICO and ROGER C. CHAPMAN[1]

[1] Mr. Talerico is a research entomologist at the Forest Insect and Disease Laboratory of the USDA Forest Service's Northeastern Forest Experiment Station, at Hamden, Conn. Mr. Chapman, formerly mathematical statistician on the staff of the Northeastern Station, Upper Darby, Pa., is now on the faculty of the School of Forestry at Duke University, Durham, N. C.