Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Letter to the Editor—Parameter Estimation in a Markov Dependent Firing Distribution

J. S. Rustagi, R. C. Srivastava,

PARAMETER ESTIMATION IN A MARKOV DEPENDENT FIRING DISTRIBUTION

J. S. Rustagi and R. C. Srivastava

The Ohio State University, Columbus, Ohio

(Received February 8, 1968)

In the study of tank weapon systems, the probability distribution of the number of trials required to obtain a preassigned number of successes in a Markov chain with two states arises naturally. This paper studies this distribution, tabulates it for a few values of the parameters, and then discusses the problems of estimating the parameters and of applications to tank-weapon systems.

IN DESIGNING a new weapon system, its lethality characteristic is of great importance. In a tank-weapon system, for example, the number of rounds required to destroy a target may be taken to describe its lethality characteristics The number of rounds to obtain a preassigned number of hits so as to destroy a target is important in the study of weapons of this type. Since in a tank gun different adjustments are needed depending on the results of previous rounds, the sequence of hits and misses does not form a sequence of independent Bernoulli trials. Assuming that the probability of a hit given that there was a miss on the previous round and also the probability of a hit given a hit on the previous round remain constant, we can describe this firing process by a two-by-two Markov chain. The probability distribution of the number of rounds required to get a specified number of hits so as to destroy a target therefore becomes an important quantity.

This note studies the statistical problems associated with this probability distribution. It discusses estimating the parameters involved, considers an application of this model, and tabulates some values of the distribution. Although the distribution arises naturally in applications of the above type, the results of the paper have much wider applications, for example, in transportation engineering, computer failures, and hospital admissions.

THE PROBABILITY DISTRIBUTION

CONSIDER A Markov Chain with two states, success (S) and failure (F) whose transition probability matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} S & F \end{array} \\ \begin{array}{c} S \\ F \end{array} & \begin{pmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{pmatrix}. \end{array}$$

Let the initial probabilities that the chain is in state S or F be given by p and $1-p$ respectively. Let Z_m be the number of trials required to obtain exactly m successes and $f_m(n) = P(Z_m = n)$ be the probability distribution of Z_m. Let $f_m(n|S)$ denote the conditional probability that $Z_m = n$ given that the result of first trial is a success and similarly $f_m(n|F)$. Expressions for $f_m(n|S)$ and $f_m(n|F)$ are given by GNEDENKO,[4] page 139. BONDER[2] has also given these expressions in his paper. We have then

$$f_m(n) = \begin{cases} p\alpha^{m-1}, & (n=m) \\ pf_m(m+1|S)+(1-p)f_m(m+1|F), & (n=m+1) \\ pf_m(n|S)+(1-p)f_m(n|F). & (n \geqq m+2) \end{cases} \quad (1)$$

The probability generating function of Z_m can be derived in a simple manner by noting that Z_m is the sum of m independent random variables.

Let X denote the number of trials required up to and including the first success given that the initial trial, which is not counted in X, resulted in a success. The probability distribution of X is given by

$$\begin{aligned} &P(X=1)=\alpha, \\ &P(X=n)=(1-\alpha)(1-\beta)^{n-2}\beta. \qquad (n \geqq 2) \end{aligned} \quad (2)$$

Then we have

$$Z_m = Z_1 + X_1 + \cdots + X_{m-1}, \quad (3)$$

where the X_i's are independent and identically distributed as X and the probability distribution of Z_1 is given by

$$\begin{aligned} &P(Z_1=1)=p, \\ &P(Z_1=n)=(1-p)(1-\beta)^{n-2}\beta. \qquad (n \geqq 2) \end{aligned} \quad (4)$$

The probability generating function of Z_m is then easily obtained as follows:

$$\begin{aligned} Q(t) &= E(t^{Z_m}) = E(t^{Z_1})[E(t^X)]^{m-1} \\ &= \{tp+(1-p)\beta t^2[1-(1-\beta)t]^{-1}\}\cdot\{t\alpha+(1-\alpha)\beta t^2[1-(1-\beta)t]^{-1}\}^{m-1} \quad (5) \\ &= t^m\{p+(1-p)\beta t(1-t+\beta t)^{-1}\}\{\alpha+(1-\alpha)\beta t(1-t+\beta t)^{-1}\}^{m-1}. \end{aligned}$$

For tabulating the distribution, we see that there are four parameters α, β, p, and m involved in it. To give the listing for even ten values of each will require 10,000 entries, but a few typical values are given in Table I. The FORTRAN program for an IBM 7094 utilized in computing the table is available for extensive tabulation.

The table gives the three decimal figures for the probability that $Z_m = a + y$ where a and y are listed for each combination of α, β, p, and m reported. The values are correct to three place decimal and probabilities less than 0.001 are not listed.

ESTIMATING THE PARAMETERS

WE CONSIDER here the problem of estimating the parameters p, α, and β that occur in the probability distribution. In view of the importance of studying any physical, biological, or social phenomenon with the aid of stochastic models of this type, the problem of estimation becomes important.

Let $S_1, S_2, \cdots, S_r$ be r independent sequences of successes and failures having total number of $n_1, n_2, \cdots, n_r$ trials respectively. These sequences are observed so as to obtain the preassigned number m of successes in each sequence. Let n^i_{SS} denote the number of transitions $S \rightarrow S$ in the sequence S_i. Similarly we define n^i_{SF}, n^i_{FF}, and n^i_{FS}. Let $Y_i = 1$ if the first trial of the sequence S_i is a success and zero otherwise. Then $t = \sum_{i=1}^{i=r} Y_i$, is the number of sequences S_i that have success in their first trial. Then using the arguments of BILLINGSLEY,[1] we have the following result:

TABLE I

THE MARKOV DEPENDENT FIRING DISTRIBUTION: $P\{Z_m = a + y\}$

α	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
β	0.1	0.1	0.5	0.5	0.5	0.5	0.5	0.5	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
p	0.9	0.9	0.1	0.1	0.5	0.5	0.9	0.9	0.1	0.1	0.1	0.5	0.5	0.9	0.9	0.9
m	5	5	5	10	5	10	5	10	5	10	15	5	10	5	10	15
y/a	4	34	4	9	4	9	4	9	4	9	14	4	9	4	9	14
1	590	003	066	039	328	194	590	349	066	039	023	328	194	590	349	206
2	033	003	310	194	237	194	164	194	558	349	217	426	349	295	349	309
3	030	002	222	188	161	167	100	145	272	331	302	180	262	089	192	247
4	028	002	150	160	105	132	060	104	081	179	235	051	128	021	077	140
5	026	002	097	125	067	099	036	072	019	071	131	011	048		025	063
6	024	002	061	093	041	071	021	049		023	059		015			024
7	022	002	038	066	025	049	012	032			022					008
8	020	002	023	046	015	033	007	021			007					
9	019	001	014	031	009	022		013								
10	017	001	008	021	005	015		008								
11	016	001	005	013		009		005								
12	014	001		009		006										
13	013	001		006												
14	012	001														
15	011															
16	010															
17	009															
18	009															
19	008															
20	007															
21	007															
22	006															
23	006															
24	005															
25	005															
26	004															
27	004															
28	004															
29	003															
30	003															

TABLE I—*Continued*

THE MARKOV DEPENDENT FIRING DISTRIBUTION: $P\{Z_m=a+y\}$

α	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.9	0.9
β	0.5	0.5	0.5	0.5	0.9	0.9	0.9	0.9	0.9	0.5	0.5	0.5	0.5	0.5	0.9	0.9	0.9	0.9	0.1	0.1
p	0.1	0.1	0.9	0.9	0.1	0.1	0.1	0.5	0.5	0.1	0.1	0.5	0.5	0.9	0.1	0.1	0.9	0.9	0.1	0.1
m	5	10	5	10	5	10	15	5	15	5	10	5	10	10	5	10	5	10	5	5
y/a	7	18	6	17	7	16	25	6	24	5	12	4	11	11	5	11	4	10	4	34
1	011	008	012	010	054	031	017	021	009	041	016	031	013	020	073	019	056	014	066	008
2	036	017	045	021	231	104	053	128	032	086	030	078	027	038	220	062	208	054	062	007
3	074	029	097	035	438	225	121	349	083	121	047	117	044	057	308	131	314	123	058	007
4	109	044	135	052	196	289	199	326	158	137	064	137	061	075	237	194	252	192	055	006
5	129	059	147	069	057	200	229	129	216	134	079	137	076	089	111	211	120	215	052	006
6	132	073	138	082	013	095	182	035	210	119	088	123	087	096	035	174	038	181	049	005
7	121	083	117	091		035	109		149	098	093	103	093	097	008	112	009	118	046	005
8	102	089	092	094		011	052		082	077	092	081	093	093		057		061	043	005
9	081	089	069	091			021		037	057	086	060	088	084		024		026	040	004
10	062	086	049	085			007		015	041	078	044	080	074		009		009	038	004
11	045	079	034	075						029	068	031	070	062					035	004
12	032	070	023	064						019	057	021	059	051					033	003
13	022	060	015	054						013	047	014	049	040					031	003
14	015	049	010	043						008	037	009	039	031					029	003
15	010	040	006	034						005	029	006	030	024					027	003
16	006	032		026							022		023	018					025	002
17	004	024		020							017		017	013					023	002
18		018		014							012		013	009					022	002
19		014		010							009		009	006					020	002
20		010		007							006		007	005					019	002
21		007		005							004		005						017	001
22		005		004									003						016	001
23		004																	015	001
24																			014	001
25																			013	001
26																			012	001
27																			011	
28																			010	
29																			009	
30																			009	

THEOREM. *The set* $(t, \sum_{i=1}^{i=r} n_{SS}^i, \sum_{i=1}^{i=r} n_{SF}^i, \sum_{i=1}^{i=r} n_{FF}^i, \sum_{i=1}^{i=r} n_{FS}^i)$ *is a set of sufficient statistics for the family of probability distributions given in (1). The maximum likelihood estimates are then obtained as follows:*

$$\hat{p}=t/r,$$
$$\hat{\alpha}=\sum n_{SS}^i/\sum (n_{SS}^i+n_{SF}^i), \qquad (6)$$
$$\hat{\beta}=\sum n_{FS}^i/\sum (n_{FS}^i+n_{FF}^i).$$

Since the maximum likelihood estimates $\hat{p}$, $\hat{\alpha}$, and $\hat{\beta}$ are based on the sums of independent and identically distributed random variables Y_i, n_{SS}^i, n_{SF}^i, n_{FS}^i, n_{FF}^i and the expression satisfies the standard regularity conditions, it follows that these estimates are consistent and their limiting joint distribution is multivariate normal.

APPLICATIONS

IN EVALUATING a tank-weapon system, one is interested in estimating the average number of rounds needed to kill a target (Bonder[2]). Assuming that the model discussed here holds for the system, we are interested in estimating the expected number of rounds needed to get a preassigned number, m, of hits. The maximum likelihood estimate of the expected value, μ, is given by $\hat{\mu}=[m(1-\hat{\alpha}+\hat{\beta})+\hat{\alpha}-\hat{p}]/\hat{\beta}$, where $\hat{p}$, $\hat{\alpha}$, $\hat{\beta}$ are given by (6).

Situations are not infrequent in practice where only the total number of rounds in each sequence is available. In other words, we have the set of $(n_1, n_2, \cdots, n_r)$ given by the sequences $S_1, S_2, \cdots, S_r$. In such circumstances, p, α, and β can be estimated by the method of moments. The estimate of the expected value is given by the average, $\bar{n}=(1/r)\sum_{i=1}^{i=r} n_i$. It is well known that such estimates are inferior to the estimates that are functions of sufficient statistics.

The estimate of the mean of the distribution may also be utilized in estimating the total time to kill a target. In military operations, the time element is a very important quantity. Suppose the time to adjust the gun is t_1 if there is a hit and is t_2 if there is a miss. Then the expected time to kill a target is given by $t_0+t_1(m-1)+t_2(\mu-m)$, where μ is the expected number of rounds to get m hits and t_0 is the time to fire the first round. An unbiased estimate of this quantity is given by $t_0+t_1(m-1)+t_2(\hat{\mu}-m)$. This estimate is obtained in terms of $\hat{p}$, $\hat{\alpha}$, and $\hat{\beta}$.

If the times to adjust the gun are random variables, the problem may be considered as the study of series of events occurring in time. An excellent exposition of such situations has been given by COX AND LEWIS.[3]

Statistical tests of hypotheses about the mean of the number of rounds may provide ways of deciding which of the two given weapon systems is more effective. One obvious way would be to use the statistic $\bar{n}$. Distribution theory of $\bar{n}$ and other associated statistics is needed to develop such tests.

ACKNOWLEDGMENTS

THE PROBLEM arose in discussion with SETH BONDER, to whom the authors are greatly indebted. DAVID MAYER wrote the IBM 7094 program for calculating the tables.

This research was partially sponsored by the Air Force Office of the Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant No. 1305-67.

REFERENCES

1. PATRICK BILLINGSLEY. "Statistical Methods in Markov Chains," *Ann. Math. Stat.* **32,** 12–40 (1961).

2. SETH BONDER, "The Lanchester Attrition-rate Coefficient," *Opns. Res.* **15**, 221–232 (1967).
3. D. R. COX AND P. A. W. LEWIS, *The Statistical Analysis of Series of Events*, Methuen, London, 1966.
4. B. V. GNEDENKO, *The Theory of Probability*, Chelsea Publishing Co., New York, 1962.

MODIFICATION OF THE PRIMAL-DUAL ALGORITHM FOR DEGENERATE PROBLEMS

Harold Greenberg

Naval Postgraduate School and Stanford Research Institute

(Received August 10, 1967)

This note presents a modification of the primal-dual algorithm for solving linear programs that speeds convergence in degenerate problems; it also gives some computational results.

DEGENERACY in a linear programming problem may cause very slow convergence to the solution when one is using a simplex method. Furthermore, care must be taken to avoid cycling by the iterative procedure. For a full description of degeneracy and for methods to prevent cycling, *see* DANTZIG.[1]

In this note, we present a modified primal-dual algorithm that speeds convergence in degenerate problems and usually requires no special means to prevent cycling. In the absence of degeneracy, the method reverts to the usual primal-dual algorithm.

THE STANDARD PRIMAL-DUAL ALGORITHM

WE CONSIDER first the primal-dual algorithm as it appears in Dantzig.[1] Using matrix notation, the linear programming problem is: minimize $z=z_0+cx$, given $Ax=b$, $x_j \geqq 0$, $j=1, \cdots, n$, where

z_0 is a constant;
c is a $1\times n$ vector with components c_j, $j=1, \cdots, n$;
b is an $m\times 1$ vector with components $b_i \geqq 0$, $i=1, \cdots, m$;
A is an $m\times n$ matrix with elements a_{ij}, $i=1, \cdots, m$; $j=1, \cdots, n$;
x is an $n\times 1$ vector with components x_j, $j=1, \cdots, n$.

The standard primal-dual algorithm consists in augmenting the constraint equations with artificial variables as,

$$Ax+y=b, \tag{1}$$

where y is an $m\times 1$ vector with components consisting of the artificial variables y_i, $i=1, \cdots, m$, and forming $w=w_0+dx$, where