A Dynamic Programming Strategy for the Two Machine Problem

By Richard Steck

The two-machine problem may be stated as follows. Given two machines A and B with unknown, fixed probabilities of performing successfully, find a strategy of sequential sampling of the machines so as to converge on the machine with the higher success rate while minimizing the expected number of failures in the learning process. Based on information gained in earlier trials, the strategy supplied by dynamic programming gives the best choice of machines at each trial. The strategy developed is thus particularly valuable when the cost of failure is great. The two machine problem has gained attention in connection with the use of untried drugs in critical medical cases. In the case that two drugs labelled A and B are thought to be cures for a usually fatal disease, one might ask how should the drugs be sampled so as to converge on the better drug with a minimal sacrifice of patients in the meantime. In this paper the results of a particular application of a dynamic programming strategy to the two-machine problem are shown. A set of computed decision tables and a description of their use is presented.

Let r and r' represent the respective number of trials and m and m' the respective number of successes of machines A and B. With the a priori information that machine A has produced n successes out of s trials and machine B has produced n' successes out of s' trials we wish to maximize the expected number of future successes using an optimal policy. In the absence of information concerning the success rates of the machines we may take $n = n' = 1$, $s = s' = 2$. Defining $f_{r,r'}(m, m')$ as the expectation of the random variable

$$(1) \qquad w = \sum_{i=r+r'+1}^{\infty} a^{i-r-r'-1},$$

where a is a discount factor introduced to make w finite and to place emphasis on early trials, Gluss [1] arrives at the relation

$$(2) \quad f_{r,r'}(m, m') = \text{Max} \begin{bmatrix} \text{A:} \dfrac{m + n}{r + s} [1 + af_{r+1,r'}(m + 1, m')] \\[2mm] \qquad\qquad + \left(1 - \dfrac{m + n}{r + s}\right) af_{r+1,r'}(m, m') \\[3mm] \text{B:} \dfrac{m' + n'}{r' + s'} [1 + af_{r,r'+1}(m, m' + 1)] \\[2mm] \qquad\qquad + \left(1 - \dfrac{m' + n'}{r' + s'}\right) af_{r,r'+1}(m, m') \end{bmatrix}$$

Assuming that after a total of $r + r' = N$ trials we expect to make a decision as

to which is the better machine then we impose the constraints

$$(3) \qquad f_{r,r'}(m, m') = \text{Max} \begin{bmatrix} \text{A}: \left(\dfrac{m+n}{r+s}\right) \Big/ (1-a) \\ \text{B}: \left(\dfrac{m'+n'}{r'+s'}\right) \Big/ (1-a) \end{bmatrix}.$$

Using (3) to compute the values of $f_{r,r'}(m, m')$ for $r + r' = N$ the f's for $r + r' < N$ were computed sequentially using (2). In the example below N was chosen in accordance with the core capacity of the UNIVAC 1105, $N = 25$. The discount factor, a, was selected as $a = 0.95$.

A considerable reduction in computation and storage was affected by taking advantage of the symmetry between r and r'. When $r > r'$ we can consider the identities of the two machines to be interchanged, i.e., we consult the table entry for r', r, and reverse the indicated decision.

Presented below are computed strategy tables for N ranging from 1 to 25. In this example the initial choice of machines is arbitrary; after the first selection and trial we refer to the table for $N = 1$. Depending upon the success or failure of this trial the appropriate row and column are consulted. At the intersection a breakpoint M' is found. The selection rule for the next trial is as follows. If $m' \geq M'$ select B otherwise select machine A. Exceptions to this rule occur as footnotes to the individual tables. The appearance of the letters A or B imply that either A or B is to be selected respectively.

As an example let us assume that we have tried machine A seven times and machine B thirteen times. Let us further assume that machines A and B have been successful in four and nine trials respectively. We consult the table for $r + r' = N = 20$. In the column $r = 7$ and the row $m = 4$ the breakpoint $M' = 9$ is found. Since $m' = M'$ machine B is selected for the next trial. Had it been the case that $m = 3$, machine B would have been selected excepting the cases where $m' = 4, 5, 6$ as indicated in the footnote.

$N = 1$

$m \backslash r$	0	1
0	1	B
1	B	

$N = 2$

$m \backslash r$	0	1
0	1	B
1	1	

$N = 3$

$m \backslash r$	0	1	2
0	2	B	B
1		2	B
2			B

$N = 4$

$m \backslash r$	0	1	2
0	3	1	B
1		3	1
2			2

$N = 5$

$m \backslash r$	0	1	2	3
0	3	1	B	B
1		3	2	B
2			3	B
3				B

$N = 6$

$m \backslash r$	0	1	2	3
0	4	2	1	B
1		4	2	1
2			4	2
3				3

$N = 7$

m \ r	0	1	2	3	4
0	5	2	1	B	B
1		5	3	1	B
2			5	3	B
3				4	B
4					B

$N = 10$

m \ r	0	1	2	3	4	5
0	7	4	2	1	1	B
1		8	5	3	2	1
2			7	5	3	2
3				7	4	3
4					6	4
5						5

$N = 8$

m \ r	0	1	2	3	4
0	6	3	1	1	B
1		6	4	2	1
2			5	3	2
3				5	3
4					4

$N = 11$

m \ r	0	1	2	3	4	5	6
0	8	5	3	2	1	B	B
1		9	6	4	2	1	B
2			8	6	4	2	B
3				8	5	4	B
4					7	5	B
5						6	B
6							B

$N = 9$

m \ r	0	1	2	3	4	5
0	7	4	2	B	B	B
1		7	4	2	1	B
2			6	4	2	B
3				6	4	B
4					5	B
5						B

$N = 12$

m \ r	0	1	2	3	4	5	6
0	9	5	3	2	1	B	B
1		10	6	4	3	2	1
2			9	6	4	3	2
3				9	6	4	3
4					8	6	4
5						7	5
6							6

$N = 13$

m \ r	0	1	2	3	4	5	6	7
0	10	6	4	2	1	1	B	B
1		11	7	5	3	2	1	B
2			10	7	5	4	2	B
3				10	7	5	3	B
4					9	7	5	B
5						8	6	B
6							7	B
7								B

$N = 14$

m \ r	0	1	2	3	4	5	6	7
0	11	7	5	3	2	1	B	B
1		11	8	6	4	2	2	1
2			11	8	6	4	3	2
3				11	8	6	4	3
4					10	8	5	4
5						9	7	5
6							8	6
7								7

$N = 15$

m \ r	0	1	2	3	4	5	6	7	8
0	12	8	5	4	3	1	1	B	B
1		12	9	6	4	3	2	1	B
2			12	9	7	5	3	2	B
3				12	9	7	5	3	B
4					11	9	6	5	B
5						10	8	6	B
6							9	7	B
7								8	B
8									B

$N = 16$

m \ r	0	1	2	3	4	5	6	7	8
0	12	8	5	3	2	2	1	B	B
1		13	10	7	5	3	3	2	1
2			13	10	8	5	4	3	2
3				12	9	7	5	4	3
4					12	9	7	5	4
5						11	9	7	5
6							10	8	6
7								9	7
8									8

$N = 17$

m \ r	0	1	2	3	4	5	6	7	8	9
0	13	9	6	4	2	2	1	B	B	B
1		14	10	8	6	4	3	2	1	B
2			14	11	8	6	5	3	2	B
3				13	10	8	6	5	3	B
4					13	11	8	6	4	B
5						12	10	7	6	B
6							11	9	7	B
7								10	8	B
8									9	B
9										B

$N = 18$

m \ r	0	1	2	3	4	5	6	7	8	9
0	14	9	6	4	3	2	2	1	1	B
1		15	11	8	6	4	4	2	2	1
2			B[1]	11	9	7	5	4	2	2
3				14	11	9	7	5	4	3
4					14	11	9	7	5	4
5						13	11	8	6	5
6							12	10	8	6
7								11	9	7
8									10	8
9										9

[1] A if $6 \leqq m' \leqq 14$.

$N = 19$

m＼r	0	1	2	3	4	5	6	7	8	9	10
0	15	10	7	5	3	3	2	B	1	1	B
1		15	12	9	7	5	3	B[2]	2	1	B
2			B[1]	12	9	7	6	B[3]	3	3	B
3				15	12	10	8	B[4]	5	3	B
4					15	12	10	8	6	5	B
5						14	12	9	7	6	B
6							13	11	9	7	B
7								12	10	8	B
8									11	9	B
9										10	B
10											B

[1] A if $7 \leqq m' \leqq 15$.
[2] A if $m' = 1, 2$.
[3] A if $m' = 3, 4$.
[4] A if $m' = 3, 4, 5$.

$N = 20$

m＼r	0	1	2	3	4	5	6	7	8	9	10
0	16	11	8	5	4	3	2	B[2]	3	1	B
1		17	12	9	7	5	5	3[3]	3	2	1
2			B[1]	13	10	8	7	B[4]	4	3	2
3				16	13	10	9	B[5]	6	4	3
4					15	13	10	9	7	5	4
5						15	12	10	8	7	5
6							14	12	10	8	6
7								13	11	9	7
8									12	10	8
9										11	9
10											10

[1] A if $m' = 8 \leqq m' \leqq 14, 16$.
[2] A if $m' = 1$.
[3] A if $m' = 1, 2$.
[4] A if $m' = 4$.
[5] A if $m' = 4, 5, 6$.

$N = 21$

m＼r	0	1	2	3	4	5	6	7	8	9	10	11
0	17[1]	11	8[2]	6	4	3	2	2	1	6[6]	B	B
1		17	13	10	8	6	4	4	3	6[7]	1	B
2			B[3]	14	11	9	7	B[4]	5	6	2	B
3				17	14	11	9	B[5]	6	5	4	B
4					16	14	11	9	8	6	5	B
5						16	13	11	9	7	6	B
6							15	13	11	9	7	B
7								14	12	10	8	B
8									13	11	9	B
9										12	10	B
10											11	B
11												B

[1] B if $m' = 14$.
[2] B if $m' = 0, 1$.
[3] A if $10 \leqq m' \leqq 16$.
[4] A if $m' = 4$.
[5] A if $m' = 4, 5, 6$.
[6] B if $m' = 3$.
[7] B if $m' = 3, 4$.

$N = 22$

m\\r	0	1	2	3	4	5	6	7	8	9	10	11
0	17	12	8[1]	6	5	3[4]	2	2	1	1	B[7]	B
1		18	14	11	8	6[4]	6	4	3	3	B[8]	1
2			B[2]	14	12	9[4]	7	B[5]	5	4	3[9]	2
3				18	15	12[4]	10	B[6]	7	6	5	3
4					17[3]	16	12	10	8	7	6	4
5						16	14	12	10	8	7	5
6							16	13	11	10	8	6
7								15	13	11	9	7
8									14	12	10	8
9										13	11	9
10											12	10
11												11

[1] B if $m' = 0, 1, 2$.
[2] A if $12 \leqq m' \leqq 17$.
[3] B if $m' = 15$.
[4] A if $m' = 14, 15$.
[5] A if $m' = 5$.
[6] A if $m' = 5, 6, 7$.
[7] A if $m' = 1, 2, 4, 5$.
[8] A if $m' = 0, 1, 4, 5$.
[9] A if $m' = 4, 5$.

$N = 23$

m\\r	0	1	2	3	4	5	6	7	8	9	10	11	12
0	17	12	9	7	4	3	3[6]	2	2	1	B	B	B
1		18	14	11	9	7	8[7]	4	3	3	2	1	B
2			0[1]	15	12	10	8[3]	7	5	4	4	2	B
3				18	15	13	11[4]	9	8	6	5	4	B
4					18	15	12[5]	11	9	7	6	5	B
5						17[2]	16	12	10	9	7	6	B
6							16	14	12	10	9	7	B
7								16	13	12	10	8	B
8									15	13	11	9	B
9										14	12	10	B
10											13	11	B
11												12	B
12													B

[1] A if $m' = 15, 16, 17$.
[2] B if $m' = 15$.
[3] A if $m' = 14, 15$.
[4] A if $m' = 14, 15$.
[5] A if $m' = 14, 15$.
[6] A if $m' = 6, 7, 14, 15$.
[7] B if $m' = 5$,   A if $m' = 14, 15$.

$N = 24$

m \ r	0	1	2	3	4	5	6	7	8	9	10	11	12
0	16	11	8	7	5	4	3	3	2	2	2	B	B
1		18	14	11	9	7	6	5	4	3	3	1	2
2			19	15	12	10	8	7	6	5	4	4²	3
3				19	15	13	11	9	8	6	6¹	5	4
4					18	15	13	11	9	8	7	6	5
5						18	15	13	11	9	8	7	6
6							17	15	12	11	9	8	7
7								17	14	12	10	9	8
8									16	14	12	10	9
9										15	13	11	10
10											14	12	11
11												13	12
12													A

¹ B if $m' = 4$.
² B if $m' = 2$.

$N = 25$

m \ r	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	13	8	6	4	3	3	2	2	1	1	1	1	1	B
1		17	12	9	7	6	5	4	3	3	2	2	2	1
2			18	14	11	9	7	6	5	4	4	3	3	2
3				19	15	12	10	8	7	6	5	4	4	3
4					19	15	13	11	9	8	7	6	5	4
5						18	15	13	11	9	8	7	6	5
6							18	15	13	11	9	8	7	6
7								17	15	13	11	9	8	7
8									17	14	12	11	9	8
9										16	14	12	10	9
10											15	13	11	10
11												14	12	11
12													13	12
13														A

IIT Research Institute
Chicago, Illinois

1. BRIAN GLUSS, "A note on a computational approximation to the two-machine problem," *Information and Control*, v. 5, no. 3, September 1962.