Yonathan Bard

Estimation of State Probabilities Using the Maximum Entropy Principle

A simple method is derived for computing state probabilities of a system when the probabilities of certain aggregate states are known. The method is based on maximizing the system entropy. It is shown that the results obtained by the method satisfy certain assumptions on statistical independence between events. The method is applied to a problem arising in computer performance analysis.

Introduction

The method of maximum entropy has long been used to determine state probabilities in fields such as statistical mechanics [1], queuing theory [2], and computer performance analysis [3]. The method and many of its applications are discussed in [4]. Use of the method is usually justified in information-theoretic terms, since the maximum entropy solution introduces the minimum possible extraneous information beyond what is implied in the problem formulation. In large population combinatorial problems, the maximum entropy solution can also be justified as being the distribution which can be realized in overwhelmingly more ways than others. It is interesting, however, to see what the maximum entropy solution means in purely probabilistic terms, without reference to information or sampling theories.

We shall attempt to apply the method to the estimation of discrete state probabilities when the probabilities of certain events (*i.e.*, aggregate states) are known. The problem arises because the number of unknown state probabilities often exceeds the number of given event probabilities, thus leaving the former indeterminate. We shall justify the maximum entropy procedure by showing that, in many important cases, the estimated probabilities satisfy certain assumptions of statistical independence. We shall then use the method to solve a problem relating to the performance of peripheral storage devices for computers. In that problem, one knows the probabilities that given access paths are free, and one seeks the probability that a given path is free when a device is ready to receive data.

Definitions

Consider a stochastic system Q which may assume any one of a set of mutually exclusive and exhaustive states $S_1, S_2, \cdots, S_n$. Let p_i be the probability that the system is in state S_i.

An *event* E is defined as a specified aggregate of states. The probability of event E is

$$q = \sum_{i|S_i \in E} p_i .$$

Event E is said to *cover* state S_i if $S_i \in E$. A *partition* of the system is a set of mutually exclusive and exhaustive events. Two partitions are *orthogonal* if every event in one partition has a nonempty intersection with every event in the other partition.

Let $E_1, E_2, \cdots, E_m$ be a set of not necessarily exclusive or exhaustive events. Let $S_1, S_2, \cdots, S_n$ be the minimal set of states required to represent these events; *i.e.*, any two states covered by identical subsets of events have been amalgamated into a single state. We shall assume henceforth that we are always dealing with a minimal set of states relative to the defined events. We say in this case that the S_i comprise the partition of Q induced by the E_j.

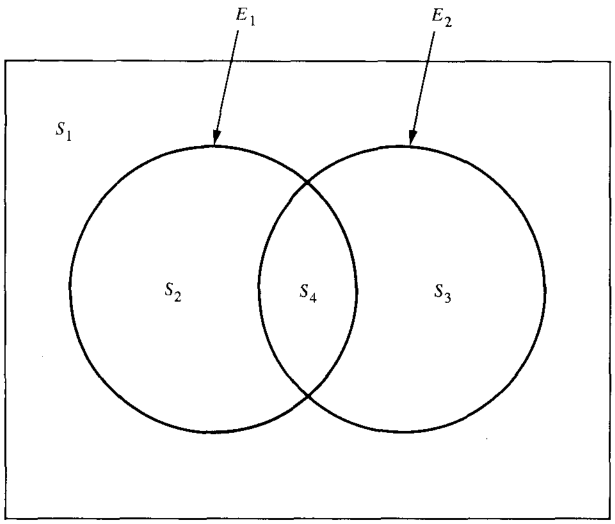

Figure 1 Two-event example.

Problem statement

Given a set of *defined* events $E_1, E_2, \cdots, E_m$ with known probabilities $q_1, q_2, \cdots, q_m$, let $S_1, S_2, \cdots, S_n$ be the partition of Q induced by the E_j. Assume that events are statistically independent as far as possible. Compute the vector of state probabilities $\mathbf{p} = [p_1, p_2, \cdots, p_n]$.

Notes:

1. The independence assumption has been left deliberately vague at this point. We shall proceed to formulate a solution, then discuss what independence assumptions it actually satisfies.
2. Among the defined events one must include the universal event which covers all states and has probability 1.

Solution

Let $F(\mathbf{p})$ be the negative entropy function, *i.e.*,

$$F(\mathbf{p}) = \sum_{i=1}^{n} p_i \log p_i.$$

Find the value of $\mathbf{p}$ which minimizes $F(\mathbf{p})$ subject to the constraints

$$p_i \geq 0 \qquad i = 1, 2, \cdots, n,$$

$$\sum_{i|S_i \in E_j} p_i = q_j \qquad j = 1, 2, \cdots, m.$$

Minimizing $F(\mathbf{p})$ is equivalent to minimizing $G(\mathbf{p}) = F(\mathbf{p}) - 1 = \sum(p_i \log p_i - p_i)$. Following standard procedures, we form the Lagrangian

$$L(\mathbf{p}) = G(\mathbf{p}) - \sum_j v_j \left(\sum_{i|S_i \in E_j} p_i - q_j \right),$$

where the v_j are unknown Lagrange multipliers. We equate to zero the derivatives $\partial L / \partial p_i$:

$$\frac{\partial L}{\partial p_i} = \log p_i - \sum_{j|S_i \in E_j} v_j = 0 \qquad i = 1, 2, \cdots, n,$$

from which we get

$$p_i = \prod_{j|S_i \in E_j} u_j \qquad i = 1, 2, \cdots, n, \tag{1}$$

where $u_j = \exp(v_j)$. Now the constraints

$$\sum_{i|S_i \in E_j} \left(\prod_{k|S_i \in E_k} u_k \right) = q_j \qquad j = 1, 2, \cdots, m \tag{2}$$

yield equations to be solved for the m unknown u_j. The latter can then be substituted in (1) to evaluate the required p_i. In words, the procedure can be described as follows:

1. Associate an unknown u_j with each defined event E_j.
2. The probability p_i of state S_i is the product of all the u_j for events E_j that cover S_i.
3. Substitute these expressions for the p_i in the equations defining the probabilities q_j.
4. Solve for the u_j and substitute in the expressions for the p_i to evaluate the latter. It should be noted that the equations are usually nonlinear, and finding their solutions may be a nontrivial task.

Before justifying this procedure, we illustrate with the simplest possible nontrivial case (Fig. 1): There are only two events E_1 and E_2 for which the states $S_1 = \sim E_1 \cap \sim E_2$, $S_2 = E_1 \cap \sim E_2$, $S_3 = \sim E_1 \cap E_2$, and $S_4 = E_1 \cap E_2$ all have finite probabilities. In addition, there is the universal event $E_3 = S_1 \cup S_2 \cup S_3 \cup S_4$, with $q_3 = 1$. There are three constraints:

E_1: $p_2 + p_4 = q_1$,

E_2: $p_3 + p_4 = q_2$,

E_3: $p_1 + p_2 + p_3 + p_4 = 1$.

Then

$p_1 = u_3$ (because p_1 appears only in E_3),

$p_2 = u_1 u_3$ (because p_2 appears only in E_1 and E_3),

$p_3 = u_2 u_3$,

$p_4 = u_1 u_2 u_3$,

so that

$$u_1(1 + u_2)u_3 = q_1,$$

$$u_2(1 + u_1)u_3 = q_2,$$

$$(1 + u_1 + u_2 + u_1 u_2)u_3 = 1.$$

The third equation may be rewritten as $(1 + u_1)(1 + u_2)u_3 = 1$. Multiplying the first equation by $1 + u_1$ and dividing by the third, one finds $u_1 = q_1(1 + u_1)$, or

$$u_1 = \frac{q_1}{1 - q_1}.$$

Analogously,

$$u_2 = \frac{q_2}{1 - q_2}$$

and

$$u_3 = \frac{1}{(1 + u_1)(1 + u_2)} = (1 - q_1)(1 - q_2),$$

so that $p_4 = u_1u_2u_3 = q_1q_2$. This solution, then, renders E_1 and E_2 statistically independent.

The existence and uniqueness of the maximum entropy solution are well known (see for example [5]). All that remains to be done is to show what statistical independence assumptions are satisfied by the maximum-entropy solutions. It should be made clear that the choice of independence assumptions which can be made in the formulation of problems of this nature is not at all unique. For example, in the case shown in Fig. 2 we can demand that E_a and E_b be independent, or we can demand that they be independent conditionally on the occurrence of E_c. In the first case, $\Pr(E_a \cap E_b) = q_aq_b$. In the second case, the requirement is that $\Pr(E_a \cap E_b|E_c) = \Pr(E_a|E_c)\Pr(E_b|E_c)$, which reduces to $\Pr(E_a \cap E_b) = q_aq_b/q_c$. We feel that the latter condition is the more natural one, and are not surprised to discover that this is the one satisfied by the maximum-entropy solution, as demonstrated in the following.

Theorem 1

Let E_a, E_b, and E_c be events satisfying the following conditions:

1. The sets $U_1 = (\sim E_a \cap \sim E_b) \cap E_c$, $U_2 = (E_a \cap \sim E_b) \cap E_c$, $U_3 = (\sim E_a \cap E_b) \cap E_c$, and $U_4 = (E_a \cap E_b) \cap E_c$ are all nonempty.
2. The four sets U_1, U_2, U_3, U_4 are partitioned by all other defined events in identical manners. That is, for every state $S_{ij} \in U_j$, $j = 1, 2, 3, 4$, there exist unique states $S_{ik} \in U_k$, $k \neq j$, such that the sets of defined events other than E_a, E_b, and E_c covering each of the four S_{ik} are identical.

Then, in the maximum-entropy solution, E_a and E_b are statistically independent conditionally on E_c. In other words, $\Pr(E_a \cap E_b|E_c) = \Pr(E_a|E_c)\Pr(E_b|E_c)$, which in this case is equivalent to

$$\Pr(U_4) = \Pr(U_2 \cup U_4)\Pr(U_3 \cup U_4)/q_c.$$

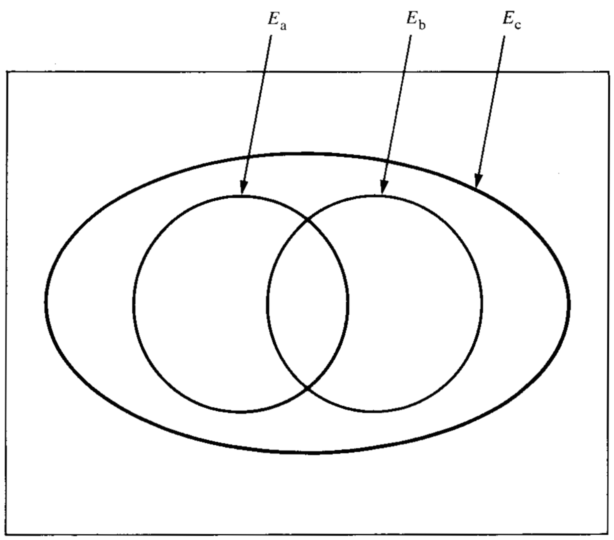

Figure 2 Three-event example.

Notes:
1. Condition 2 essentially states that knowledge of whether any other event has or has not occurred conveys no information on whether E_a, E_b, and/or E_c have occurred.
2. The events E_a, E_b, and/or E_c need not be among the E_j whose probabilities are given *a priori*. For suppose q_a is not given. We can solve the problem using all given events, then evaluate q_a from the computed state probabilities, then append E_a to the list of E_j using the computed q_a. The solution of the new problem is obviously identical to that of the old one. We may, then, assume without loss of generality that E_a, E_b, and E_c are among the events with specified probabilities.

Proof

Let S_{ik}, $i = 1, 2, \cdots, r$, be the set of states included in U_k, $k = 1, 2, 3, 4$. Let $Q_i = \Pr(U_i)$. Then, for instance,

$$Q_1 = p_{11} + p_{21} + \cdots + p_{r1},$$

where $p_{ik} = \Pr(S_{ik})$. Since $U_1 \cap E_a = \varnothing$ and $U_1 \cap E_b = \varnothing$, the expressions for the p_{i1} in terms of the u_j do not contain either u_a or u_b. Now, by assumption, the coverage of S_{i2} differs from the coverage of S_{i1} only by also including E_a. Hence $p_{i2} = u_a p_{i1}$, $i = 1, 2, \cdots, r$. We conclude, then, that

$$Q_2 = u_aQ_1. \tag{3}$$

Similar arguments lead to

$$Q_3 = u_bQ_1, \tag{4}$$

$$Q_4 = u_au_bQ_1. \tag{5}$$

The constraint imposed by q_a may be written as

$$Q_2 + Q_4 + \Pr(E_a \cap \sim E_c) = q_a$$

or

$$Q_2 + Q_4 = Q_a, \tag{6}$$

where $Q_a = \Pr(E_a \cap E_c) = q_a - \Pr(E_a \cap \sim E_c)$. Similarly,

$$Q_3 + Q_4 = Q_b, \tag{7}$$

where $Q_b = \Pr(E_b \cap E_c) = q_b - \Pr(E_b \cap \sim E_c)$; and

$$Q_1 + Q_2 + Q_3 + Q_4 = q_c. \tag{8}$$

Substituting (3), (4), and (5) into (6), (7), and (8) we find

$$Q_1u_a + Q_1u_au_b = Q_a,$$

$$Q_1u_b + Q_1u_au_b = Q_b,$$

$$Q_1 + Q_1u_a + Q_1u_b + Q_1u_au_b = q_c.$$

These equations may be factored into

$$Q_1u_a(1 + u_b) = Q_a,$$

$$Q_1u_b(1 + u_a) = Q_b,$$

$$Q_1(1 + u_a)(1 + u_b) = q_c.$$

Multiplying the first equation by $1 + u_a$ and substituting from the third equation yields

$$u_aq_c = (1 + u_a)Q_a;$$

i.e.,

$$u_a = \frac{Q_a}{q_c - Q_a}.$$

Analogously,

$$u_b = \frac{Q_b}{q_c - Q_b}$$

and

$$Q_1 = \frac{q_c}{(1 + u_a)(1 + u_b)} = \frac{(q_c - Q_a)(q_c - Q_b)}{q_c},$$

so that

$$Q_4 = u_au_bQ_1 = \frac{Q_aQ_b}{q_c} = \frac{(Q_2 + Q_4)(Q_3 + Q_4)}{q_c}, \tag{9}$$

as was to be shown.

Corollary

If E_a and E_b are events such that $V_1 = \sim E_a \cap \sim E_b$, $V_2 = E_a \cap \sim E_b$, $V_3 = \sim E_a \cap E_b$, and $V_4 = E_a \cap E_b$ are all nonempty and are all identically partitioned by all other defined events, then the maximum-entropy solution renders E_a and E_b statistically independent.

Proof

Take E_c to be the entire system, so that $q_c = 1$, $Q_a = q_a$, $Q_b = q_b$, $Q_4 = \Pr(V_4)$. Then Eq. (9) becomes $\Pr(E_a \cap E_b) = q_aq_b$. *Q.E.D.*

• *Theorem 2*

Let $(E_{1\cdot}, E_{2\cdot}, \cdots, E_{r\cdot})$ and $(E_{\cdot 1}, E_{\cdot 2}, \cdots, E_{\cdot s})$ be two sets of events which partition Q orthogonally, *i.e.*,

$$E_{i\cdot} \cap E_{j\cdot} = \varnothing \qquad (i \neq j),$$

$$E_{\cdot i} \cap E_{\cdot j} = \varnothing \qquad (i \neq j),$$

$$\bigcup_i E_{i\cdot} = Q,$$

$$\bigcup_j E_{\cdot j} = Q,$$

$$E_{i\cdot} \cap E_{\cdot j} \neq \varnothing \qquad (\text{all } i, j).$$

Let the event probabilities $p_{i\cdot} = \Pr(E_{i\cdot})$ and $p_{\cdot j} = \Pr(E_{\cdot j})$ be given, and let the probabilities p_{ij} of states $S_{ij} = E_{i\cdot} \cap E_{\cdot j}$ be unknown. Then, in the maximum entropy solution for the p_{ij}, the events $E_{i\cdot}$ and $E_{\cdot j}$ are independent for all i, j. In other words,

$$p_{ij} = p_{i\cdot}p_{\cdot j}.$$

Note: In this theorem, the $p_{i\cdot}$ and $p_{\cdot j}$ play the role of the q_j in the general case, and the p_{ij} play the role of the p_i. The constraints that the p_{ij} must satisfy are

$$p_{i\cdot} = \sum_j p_{ij} \text{ and } p_{\cdot j} = \sum_i p_{ij}.$$

Proof

We associate the variables $u_{i\cdot}$ and $u_{\cdot j}$ with the events $E_{i\cdot}$ and $E_{\cdot j}$, respectively. Equation (2) for this case then reduces to

$$p_{ij} = u_{i\cdot}u_{\cdot j}.$$

Hence,

$$p_{i\cdot} = \sum_j p_{ij} = u_{i\cdot}\sum_j u_{\cdot j}; \qquad i.e., \; u_{i\cdot} = p_{i\cdot}\Big/\sum_j u_{\cdot j},$$

$$p_{\cdot j} = \sum_i p_{ij} = u_{\cdot j}\sum_i u_{i\cdot}; \qquad i.e., \; u_{\cdot j} = p_{\cdot j}\Big/\sum_i u_{i\cdot}.$$

Now,

$$1 = \sum_i p_{i\cdot} = \sum_j p_{\cdot j} = \sum_i u_{i\cdot}\sum_j u_{\cdot j}.$$

Therefore,

$$p_{ij} = u_{i\cdot}u_{\cdot j} = \frac{p_{i\cdot}\,p_{\cdot j}}{\sum_i u_{i\cdot}\sum_j u_{\cdot j}} = p_{i\cdot}p_{\cdot j}.$$

Q.E.D.

Remarks:

1. The proof is easily extended to the case where the partition is not of the entire system Q, but only of some event E. In this case, independence is conditional on E.
2. The proof is easily extended to any number of orthogonal partitions.

3. An immediate application is the estimation of contingency table entries when only the margins are known [6, 7]. All cases discussed in [6] are covered. Theorem 2 states that if the maximum entropy principle is used to fill in the empty slots in a contingency table, the result is as though the marginal distributions were assumed to be independent of each other.

Discussion

The maximum entropy principle is usually applied in the following form: maximize the entropy, given the expected values of certain observable quantities. Such conditions translate into linear constraints on the state probabilities. Conversely, any linear function of the state probabilities can be viewed as an expectation of a certain random variable. For example, in our case, the probability of an event E is the expected value of a random variable which takes the value one in states covered by E, and zero elsewhere. Thus, our problem is a special case of the maximum entropy formalism.

In solving maximum entropy problems, it is common practice to omit the universal event from the set of explicit constraints. Instead, one normalizes by changing (1) into

$$p_i = \prod_{j|S_i \in E_j} u_j \Big/ \sum_k \prod_{j|S_k \in E_j} u_k .$$

While this reduces the number of unknowns by one, the form of the equations becomes more complicated, and explicit solutions are more difficult to derive.

Application

We now provide an example which constitutes a nontrivial application of the method. The example is sufficiently intractable to make calculation of the required probability by direct methods beyond the author's capabilities. For more details, see [8]. The following analysis is presented in a somewhat simplified form.

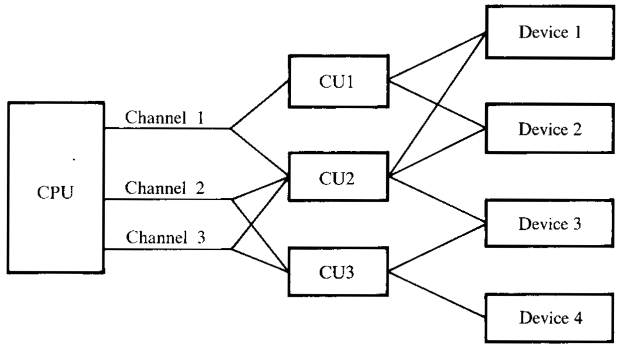

Figure 3 Sample I/O configuration.

Table 1 System states.

System state no.	*Channel state*	*Control unit state*	*Device x state*
1	0	y	0
2	y	0	0
3	y	y	0
4	y	z	0
5	x	0	x
6	y	0	x
7	x	x	x
8	x	y	x
9	y	x	x
10	y	y	x
11	y	z	x
12	0	0	x
13	0	x	x
14	0	y	x
15	0	0	0

Key:
0: Component free.
x: Component serving device x.
y: Component serving device y.
z: Component serving device z.
(*Note:* y and z represent any distinct devices other than x.)

A computer system wishes to transfer data to and from several peripheral storage *devices*. A data transmission *path* from the system to a device (actually, to a string of devices) consists of a *channel* and a *control unit*. There may be several control units connected to each channel, and several devices to each control unit. Conversely, a device may be attached to several control units, and a control unit to several channels. A possible configuration is depicted in Fig. 3.

Assume that the data flow rate over each path is given. That is, we are told what fraction of the time each channel/control unit combination is busy transmitting data to each device. Any component (channel, control unit, or device) can be transmitting for only one path at a time. In order to evaluate the performance of this system, one must be able to answer the question: given that a specified device is free (not busy), what is the probability that a given path to it is free?

For our purposes, we concentrate on a specific path, consisting of a specific channel, control unit, and device. We neglect the fine structure of events not relating to the selected path. Let x designate the specific device, and let y and z designate any other distinct devices. Then we can distinguish fifteen different states, which are listed in Table 1. Referring to Fig. 3, suppose channel 2, control unit 2, and device 2 are the specific components. Then, for example, in state S_1 the control unit (CU2) is transmitting for another channel (channel 1 or 3) and another device (device 1 or 3). In state S_{11} the channel and control unit are each serving a separate device other than x (*i.e.*, channel 2 is serving device 4 via CU3, and CU2 is serving device 3 via channel 3, while device 2 is being served by CU1 via channel 1).

The path utilization data are expressible as probabilities for the following events:

E_C: channel is busy.
E_U: control unit is busy.
E_D: device is busy.
E_{CU}: channel and control unit are busy serving same device.
E_{CD}: channel is busy serving specified device.
E_{UD}: control unit is busy serving specified device.
E_{CUD}: channel and control unit are serving specified device.

In terms of the states defined in Table 1, the event probabilities are

$$q_C = p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 + p_9 + p_{10} + p_{11},$$

$$q_U = p_1 + p_3 + p_4 + p_7 + p_8 + p_9 + p_{10} + p_{11} + p_{13} + p_{14},$$

$$q_D = p_5 + p_6 + p_7 + p_8 + p_9 + p_{10} + p_{11} + p_{12} + p_{13} + p_{14},$$

$$q_{CU} = p_3 + p_7 + p_{10},$$

$$q_{CD} = p_5 + p_7 + p_8,$$

$$q_{UD} = p_7 + p_9 + p_{13},$$

$$q_{CUD} = p_7,$$

and, of course,

$$1 = \sum_{i=1}^{15} p_i.$$

Through a set of simple linear transformations one can reach a somewhat simpler but equivalent set of events whose probabilities can be computed directly from the given probabilities:

$$q_1 = q_C - q_{CU} - q_{CD} + q_{CUD} = p_2 + p_4 + p_6 + p_9 + p_{11},$$

$$q_2 = q_U - q_{CU} - q_{UD} + q_{CUD} = p_1 + p_4 + p_8 + p_{11} + p_{14},$$

$$q_3 = q_D - q_{CD} - q_{UD} + q_{CUD} = p_6 + p_{10} + p_{11} + p_{12} + p_{14},$$

$$q_4 = q_{UD} - q_{CUD} = p_9 + p_{13},$$

$$q_5 = q_{CD} - q_{CUD} = p_5 + p_8,$$

$$q_6 = q_{CU} - q_{CUD} = p_3 + p_{10},$$

$$q_7 = 1 - q_{CU} - q_{CD} - q_{UD} + 2q_{CUD} = p_1 + p_2 + p_4 + p_6 + p_{11} + p_{12} + p_{14} + p_{15}.$$

In addition, the value of p_7 is known directly ($= q_{CUD}$). To apply our algorithm, we associate the variables $u_1, u_2, \cdots, u_7$ with the above constraints and obtain the following expressions for the state probabilities:

$p_1 = u_2u_7,$	$p_9 = u_1u_4,$
$p_2 = u_1u_7,$	$p_{10} = u_3u_6,$
$p_3 = u_6,$	$p_{11} = u_1u_2u_3u_7,$
$p_4 = u_1u_2u_7,$	$p_{12} = u_3u_7,$
$p_5 = u_5,$	$p_{13} = u_4,$
$p_6 = u_1u_3u_7,$	$p_{14} = u_2u_3u_7,$
$p_8 = u_2u_5,$	$p_{15} = u_7.$

For instance, the expression for p_1 can be derived from Eq. (1) by noting that p_1 appears only in the equations for q_2 and q_7. The constraints are now transformed into

$$u_1(u_7 + u_2u_7 + u_3u_7 + u_2u_3u_7 + u_4) = q_1,$$

$$u_2(u_7 + u_1u_7 + u_5 + u_1u_3u_7 + u_3u_7) = q_2,$$

$$u_3(u_1u_7 + u_6 + u_1u_2u_7 + u_7 + u_2u_7) = q_3,$$

$$u_4(u_1 + 1) = q_4,$$

$$u_5(u_2 + 1) = q_5,$$

$$u_6(u_3 + 1) = q_6,$$

$$u_7(1 + u_2 + u_1 + u_1u_2 + u_1u_3 + u_1u_2u_3 + u_3 + u_2u_3) = q_7.$$

The first equation can be factored into

$$u_1u_7(1 + u_2)(1 + u_3) + u_1u_4 = q_1.$$

Multiply by $1 + u_1$ and substitute from the fourth equation

$$u_1u_7(1 + u_1)(1 + u_2)(1 + u_3) + u_1q_4 = (1 + u_1)q_1.$$

The seventh equation can also be factored into

$$u_7(1 + u_1)(1 + u_2)(1 + u_3) = q_7.$$

Substituting this into the previous expression yields

$$u_1q_7 + u_1q_4 = (1 + u_1)q_1,$$

which has the solution

$$u_1 = \frac{q_1}{q_7 + q_4 - q_1} = \frac{q_C - q_{CU} - q_{CD} + q_{CUD}}{1 - q_C}.$$

From the fourth equation, then,

$$u_4 = \frac{q_4}{1 + u_1} = \frac{(q_{\mathrm{UD}} - q_{\mathrm{CUD}})(1 - q_{\mathrm{C}})}{1 - q_{\mathrm{CU}} - q_{\mathrm{CD}} + q_{\mathrm{CUD}}}.$$

Analogous procedures yield

$$u_2 = \frac{q_2}{q_7 + q_5 - q_2} = \frac{q_{\mathrm{U}} - q_{\mathrm{CU}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}}}{1 - q_{\mathrm{U}}},$$

$$u_3 = \frac{q_3}{q_7 + q_6 - q_3} = \frac{q_{\mathrm{D}} - q_{\mathrm{CD}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}}}{1 - q_{\mathrm{D}}},$$

$$u_5 = \frac{q_5}{1 + u_2} = \frac{(q_{\mathrm{CD}} - q_{\mathrm{CUD}})(1 - q_{\mathrm{U}})}{1 - q_{\mathrm{CU}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}}},$$

$$u_6 = \frac{q_6}{1 + u_3} = \frac{(q_{\mathrm{CU}} - q_{\mathrm{CUD}})(1 - q_{\mathrm{D}})}{1 - q_{\mathrm{CD}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}}},$$

and

$$u_7 = \frac{q_7}{(1 + u_1)(1 + u_2)(1 + u_3)}.$$

From these expressions, all the state probabilities may be computed. We are specifically interested in the probability of the path being free, given that the device is free. This probability is

$$q = \frac{p_{15}}{1 - q_{\mathrm{D}}} = \frac{u_7}{1 - q_{\mathrm{D}}} = \frac{q_7}{(1 - q_{\mathrm{D}})(1 + u_1)(1 + u_2)(1 + u_3)}$$

$$= \frac{(1 - q_{\mathrm{CU}} - q_{\mathrm{CD}} - q_{\mathrm{UD}} + 2q_{\mathrm{CUD}})(1 - q_{\mathrm{C}})(1 - q_{\mathrm{U}})}{(1 - q_{\mathrm{CU}} - q_{\mathrm{CD}} + q_{\mathrm{CUD}})(1 - q_{\mathrm{CU}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}})(1 - q_{\mathrm{CD}} - q_{\mathrm{UD}} + q_{\mathrm{CUD}})}.$$

Validation of the results of this analysis against observed data is presented in [8].

Conclusion

The foregoing discussion is incomplete in several respects. Principally, we are lacking the following:

1. A precisely formulated complete set of independence conditions for which the maximum-entropy method provides the unique solution.
2. A set of conditions under which the constraint equations have a rational solution, as in all the cases examined in this paper.
3. A systematic algorithm for solving the constraint equations, other than general nonlinear-equation solution algorithms such as the Newton-Raphson method. For a discussion of that method in the maximum-entropy context, with a proof of convergence, see [5].

Even without specialized algorithms, the maximum-entropy method provides an easily implemented systematic method for solving problems which often baffle an analyst who attempts to use direct methods.

References

1. A. Katz, *Principles of Statistical Mechanics—The Information Theory Approach*, W. H. Freeman Co., New York, 1967.
2. J. E. Shore, "Derivation of Equilibrium and Time-Dependent Solutions to M/M/∞//N and M/M/∞ Queueing Systems Using Entropy Maximization," *Proceedings*, National Computer Conference, Anaheim, CA, 1978, pp. 483-487.
3. A. E. Ferdinand, "A Statistical Mechanical Approach to Systems Analysis," *IBM J. Res. Develop.* **14**, 539-547 (1970).
4. R. D. Levin and M. Tribus, Eds., *The Maximum Entropy Formalism*, MIT Press, Cambridge, MA, 1979. In particular, see E. T. Jaynes, "Where do we stand on maximum entropy?," pp. 15-118.
5. N. Agmon, Y. Alhassid, and R. D. Levine, "An Algorithm for Determining the Lagrange Parameters in the Maximum Entropy Formalism," Ref. 4, pp. 207-210.
6. R. W. Chapman, "Contingency Tables with Given Marginals: A Unification," *Research Report RJ1930*, IBM Research Laboratory, San Jose, CA, 1977.
7. D. V. Gokhale and S. Kullback, *The Information in Contingency Tables*, M. Dekker, New York, 1978.
8. Y. Bard, "A Model of Shared DASD and Multipathing," *Technical Report G320-2130*, IBM Scientific Center, Cambridge, MA; also presented at Performance 80, Toronto, Ont., May 1980.

Received January 10, 1980

The author is located at the IBM Cambridge Scientific Center, 545 Technology Square, Cambridge, Massachusetts 02139.