SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853

TECHNICAL REPORT NO. 891

February 1990

OPTIMAL CHECKING PROCEDURES
FOR MONITORING LABORATORY ANALYSES

by

Lisa M. McShane and Bruce W. Turnbull

This research was supported in part by Grant GM 28364 for the U.S. National Institutes of Health, and in part by a fellowship from the U.S. Army Research Office through the Mathematical Sciences Institute at Cornell University.

OPTIMAL CHECKING PROCEDURES FOR MONITORING LABORATORY ANALYSES

Lisa M. McShane

Biometry and Field Studies Branch, National Institute for Neurological Disorders and Stroke

National Institutes of Health, Bethesda, Maryland 20892

and

Bruce W. Turnbull

School of Operations Research and Industrial Engineering

Cornell University, Ithaca, New York 14853

SUMMARY

Many clinical, environmental, and epidemiologic studies rely heavily upon biochemical data and the quality of this data is of paramount importance to the validity of study conclusions. Traditionally, far more attention has been given to the analysis of study data than has been given to monitoring the quality of the data. This paper draws an analogy between monitoring a laboratory system and an industrial production process and discusses limitations of industrial quality control plans, when applied in a laboratory setting. Methods are derived for computing optimal checking schedules for laboratory analyses. These schedules formalize traditional laboratory practices of periodic checking by providing guidelines for the frequency and placement of checks within a finite batch of analyses. Many of the schedules are easy to compute, thus further enhancing their practical applicability. An application is given concerning the measurement of selenium status in plasma samples using an electrothermal atomic absorption spectrophotometry procedure.

Key words: Laboratory quality control, Biochemical analyses, Optimal checking schedules

1. Introduction

The methods we describe in this paper were motivated by the monitoring of an electrothermal atomic absorption spectrophotometry (EAAS) laboratory procedure used to measure plasma selenium status in subjects participating in a double-blind clinical trial presently in progress. In this study, several chemical and biochemical analyses are being employed to investigate the effects of micronutrient status on the risk of new non-melanoma skin cancer. Quality control in scientific laboratories has recently become the subject of increased attention in the fields of biology, biochemistry, and medicine as laboratories in these fields are relied upon to perform analyses which provide data on environmental conditions, lifestyle factors, and medical conditions. One of the purposes of this paper is to stimulate more discussion about quality control in laboratories. The accuracy of laboratory analyses can directly influence the validity of conclusions drawn from clinical, environmental and epidemiologic studies or may result in changes in medical diagnoses and treatments.

With the recognition that scientific laboratory procedures have many similarities to industrial manufacturing processes, many standard industrial quality control procedures have been adapted for use in laboratories. Levey and Jennings (1950) described the use of control charts in laboratories, while Griffin (1968) adapted cumulative sum methods. Amenta (1968), Riddick, Flora, and Van Meter (1972) and others have applied analysis of variance techniques to laboratory settings. More complete reviews of historically and currently used clinical chemistry quality control procedures are given by Limonard (1980) and Nix, Rowlands, and Kemp (1987).

Many laboratories continue to employ informal methods of monitoring the quality of analyses by making periodic "checks". Here we will use the term "check" in a broad sense. It may mean inserting into a batch of samples a series of controls, standards, or blind replicate samples; it might mean reanalyzing certain samples, or it might mean a visual inspection of the laboratory equipment. Frequently inadequate consideration is given to questions of how many checks should be made and in what locations in the series of samples they should be placed. If the "checks" consist of analytic measurements made periodically on similar samples, and these measurements are compared to predetermined limits, then the checking procedure may be essentially a control chart procedure;

however, if the checks are analyses of different types of samples, or if the checks consist of visual inspections, say, then the checking procedure fits more into the general framework of problems of scheduling inspections (checks) for systems subject to failure. “Failure” of a laboratory system may have a very general meaning. It may refer to a mechanical failure of laboratory equipment, it may refer to a technician error resulting in inaccurate analyses, or it may refer to a sudden unexplained shift to a high degree of bias or variability in analytic measurements.

Here we present economic-based approaches for determining optimal checking schedules for laboratory systems. The schedules are optimal for a batch consisting of a finite number of samples in the sense of minimizing total expected cost when there are three types of costs encountered in the process — the cost of a check, the cost of allowing a “poor quality” analysis to escape detection, and the cost of repairing the laboratory system. It is only through repair that the process may leave the “poor quality” state and be returned to the “acceptable quality” state.

The quality control problem we discuss here arose in a biological application, but some related work has been published in the industrial quality control areas of economic control chart design and process control methods. Duncan (1956) is generally credited with first proposing economic design of control charts. Montgomery (1980) provides a review of much of the subsequent literature. More recent articles include those by Lorenzen and Vance (1976), Arnold (1987), Banerjee and Rahim (1988), and Pignatiello and Tsai (1988). Taguchi, Elsayed, and Hsiang (1989) present an economic model for the design of on-line process-control procedures. Their methods and others are discussed by Adams and Woodall (1989). The results in these papers and the references therein are only applicable if the laboratory is willing to follow the prescribed checking and control procedure, and the majority of control chart design procedures are based on minimizing long run cost per unit time; whereas, our interest is in minimizing cost for a finite batch of samples. There is also a large collection of literature dealing with optimal inspection schedules and an even larger body of literature dealing with the more general topic of maintenance models in which decisions are made regarding optimal inspection scheduling, optimal timing and degree of repairs and maintenance, and optimal timing of non-failure replacements. Surveys of this literature are given by McCall (1965), Pierskalla and Voelker (1976), and

Sherif and Smith (1981). We briefly mention some of the particular references which are most relevant to this laboratory system monitoring problem and what their shortcomings are when applied in a laboratory setting.

Barlow, Hunter, and Proschan (1963) were among the first to examine the problem of optimal inspection scheduling. They assume that the failure distribution of the system is known; system failure is known only through checking; checking does not degrade the system; the system cannot fail while being checked, and monitoring continues until system failure has been detected, at which time the problem ends. They employ a cost model in which there is a cost c_1 for each check (inspection), and there is a cost c_2 per unit time incurred during the time elapsed between system failure and its discovery. Based on an expression they derive for the expected total cost incurred up to the point when system failure is detected, they give conditions under which an optimal (minimum expected cost) inspection schedule exists and demonstrate an iterative method for obtaining that optimal schedule.

Derman (1961) finds the minimax optimal inspection schedule using a cost model similar to that of Barlow, *et al.* (1963). Klein (1962) uses Markov chain models and linear programming techniques to solve for the long run minimum cost inspection-maintenance-repair schedule. Luss and Kander (1974) consider the cases where inspection duration is non-negligible and inspections are subject to error. Kander (1978) uses numerical optimization techniques to solve for minimum cost schedules for deteriorating systems modelled by semi-Markov processes. Munford and Shahani (1972) propose an easily computable inspection policy which minimizes long run cost per unit time by spacing inspections so that the probability of failure between inspections is some optimal constant. Hayre (1982) departs from the cost-based optimality criterion and proposes a simple dynamic method for determining inspection schedules to control the probability that failure goes undetected for longer than some specified time.

All of the inspection scheduling methods just described have drawbacks for the laboratory check scheduling problem motivating this paper. One limitation is that they are applied to a single life cycle (time from system start-up to failure detection) or to an infinite time period composed of many life cycles. We are interested in monitoring the system for some finite period of time. During this time

period the system may be repaired or replaced several times and the monitoring may end in the middle of a life cycle. Laboratories typically perform many types of biochemical analyses and may analyze batches of samples for many different clients. In an effort to provide every client with high quality analyses at an affordable price, the laboratory may be more interested in optimal checking schedules for individual batches than in long run optimal checking schedules. A second disadvantage is that all of the methods, with the exception of the Munford and Shahani (1972) method and the Hayre (1982) method, are complex, thereby inhibiting their usefulness in the laboratory setting.

In this paper we propose methods for determining how many checks should be made during a finite series of analyses and how the checks should be scheduled to optimally balance the quality of batches of analyses with the cost of the checking schedule. The first schedule we derive assumes that system failure can be modelled in continuous time by an exponential distribution. Expressions for the optimal checking times are derived for the case of monitoring a system over a finite time period. The second checking schedule we propose is a randomized schedule in which analyses correspond to discrete time units and each analysis in a batch is checked with probability f where f is determined so as to minimize total cost of the monitoring process. The system is monitored over a finite time period corresponding to a finite batch of analyses. This checking schedule is derived assuming that the incoming quality follows a two state Markov process which is equivalent to modelling system failure by a geometric distribution. One state is an absorbing state representing “poor quality”, and the other state is an “acceptable quality” state. We show how both the exponential and geometric models give rise to optimal inspection schedules that can be computed easily. For situations in which neither the exponential model nor the Markov failure model is appropriate, we propose a dynamic programming formulation and solution algorithm for the problem of determining an optimal checking schedule for a system with a general failure distribution, monitored over a finite time period.

2. Example 1: Systems With Exponential Failure Distributions

Suppose we wish to monitor a laboratory system over a finite time period $[0, \tau]$ ($\tau > 0$), and it may be assumed that the failure distribution for the system can be approximated by an exponential

distribution with mean $1/\beta$ ($\beta > 0$). Laboratory analyses are performed at a rate of one per unit time. System failure can only be detected through checking. Following Barlow *et al.* (1963), we make the approximating assumption that the time required to perform a check is negligible, and checking does not degrade the system. Using a cost-based optimality criterion, we determine the optimal checking schedule.

We consider a cost model in which there is a cost per unit time for operation of a failed system, a cost for each check, and a cost to repair the system. We denote the costs as follows:

c_1 = cost per check;

c_2 = cost per unit time to operate failed system (i.e., cost per inaccurate analysis);

c_3 = cost to repair system.

The goal is to find the optimal number and placement of checks so as to minimize the total expected cost over the time period $[0, \tau]$.

If we never check the system, the expected cost is given by

$$C_0(c_2, \tau, \beta) = \int_0^\tau c_2(\tau - x)\beta e^{-\beta x} dx = c_2[\tau - \frac{1}{\beta} + \frac{e^{-\beta\tau}}{\beta}]. \quad (1)$$

If we make at least one check, the analysis proceeds as follows. Let $t_1 < t_2 < \dots < t_m$ denote a set of checking times. Define spacings

$$d_i = t_i - t_{i-1}; \quad i = 1, 2, \dots, m+1$$

where for notational convenience we define $t_0 = 0$, and $t_{m+1} = \tau$. (In some laboratories it is customary to perform checks at the beginning and end of a set of analyses in addition to interim checks. In that case $t_0 = 0$ and $t_{m+1} = \tau$ are checking times, and, as shown by McShane (1989), the optimal spacing of checks will be only slightly different than in the case we consider here.) Modelling the inter-failure times by independent exponential random variables with means $1/\beta$, by the memoryless property we can write the total expected cost associated with the checking schedule $\{t_1, t_2, \dots, t_m\}$ over the interval $[0, \tau]$ as

$$\begin{aligned} C(d_1, d_2, \dots, d_m) = & m c_1 + \sum_{k=1}^m \int_0^{d_k} [c_2(d_k - x) + c_3] \beta e^{-\beta x} dx \\ & + \int_0^{\tau - \sum_{k=1}^m d_k} c_2(\tau - \sum_{k=1}^m d_k - x) \beta e^{-\beta x} dx. \end{aligned} \quad (2)$$

Simplifying, we obtain

$$C(d_1, d_2, \dots, d_m) = m \left(c_1 + c_3 - \frac{c_2}{\beta} \right) + c_2 \tau - \frac{c_2}{\beta} + \left(\frac{c_2}{\beta} - c_3 \right) \sum_{k=1}^m \exp[-\beta d_k] + \frac{c_2}{\beta} \exp[-\beta (\tau - \sum_{k=1}^m d_k)]. \quad (3)$$

$C(d_1, d_2, \dots, d_m)$ depends upon τ , β , and the cost parameters, as well as the spacings. Note that if we define the empty sum to be zero, Equation (3) is equivalent to Equation (1) when $m = 0$.

We will define the optimal checking schedule as the schedule with spacings which minimize the total expected cost given by (3). Lemma 1 gives sufficient conditions under which an optimal checking schedule exists for each m , and Lemma 2 gives sufficient conditions for the existence of an optimal m and a method for computing that optimal m .

Lemma 1: For each $m \in \{m: m = 1, 2, 3, \dots\}$, $C(d_1, d_2, \dots, d_m)$ given by (3) achieves a unique minimum at the point

$$\vec{d}^{(m)} = \left(\frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta)}{m+1}, \frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta)}{m+1}, \dots, \frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta)}{m+1} \right)$$

when $c_2 > c_3 \beta$.

The proof is supplied in Appendix A.

Note that, for any $m \geq 1$, the optimal checking times are equally spaced except that the length of time between the final check and time τ is longer than the time between successive checks. This is evident from writing d_{m+1} (the time between the last check and time τ) as

$$d_{m+1} = \tau - \sum_{k=1}^m d_k = \frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta)}{m+1} - \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta).$$

Since $c_2 > c_3 \beta > 0$ we have $\frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta) < 0$ so that d_{m+1} is greater than the spacing between successive checks. It seems intuitively reasonable that the last spacing should be longer because while there is a greater risk of system failure, the cost consequences of failure are lessened by the fact that no further checks will be made and no additional repair expenses can be incurred.

Heuristically, the condition $c_2 > c_3\beta$ ($= c_3 \div \frac{1}{\beta}$) can be roughly interpreted as saying that the cost per unit time to operate a failed system is greater than the long run average cost per unit time incurred as a result of repair expenses. If the opposite were true, namely, $c_2 \leq c_3\beta$, then it would not make sense to ever repair the system. If the system is never repaired, then there is no point in ever checking it, and hence the optimal checking policy would be to never check.

From Lemma 1 we have the optimal spacings for fixed m , and now we want to find the optimal m . Substituting the point $\vec{d}^{(m)}$ back into Equation (3) we have a function of m given by

$$V(m) = c_2\tau - \frac{c_2}{\beta} + (c_1 + c_3 - \frac{c_2}{\beta})m + \frac{c_2(m+1)}{\beta} \left(1 - \frac{c_3}{c_2}\beta\right)^{\frac{m}{m+1}} \exp\left(-\frac{\beta\tau}{m+1}\right). \quad (4)$$

Now we differentiate Equation (4) with respect to m , treating m as a continuous variable, to obtain

$$\frac{dV(m)}{dm} = c_1 + c_3 - \frac{c_2}{\beta} + c_2 \left(1 - \frac{c_3}{c_2}\beta\right)^{\frac{m}{m+1}} \exp\left(-\frac{\beta\tau}{m+1}\right) \left(\frac{\tau}{m+1} + \frac{1}{\beta} + \frac{\frac{1}{\beta} \log(1 - \frac{c_3}{c_2}\beta)}{(m+1)}\right) \quad (5)$$

and

$$\frac{d^2 V(m)}{dm^2} = \frac{c_2\beta \left(1 - \frac{c_3}{c_2}\beta\right)^{\frac{m}{m+1}}}{(m+1)^3} \exp\left(-\frac{\beta\tau}{m+1}\right) \left(\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2}\beta)\right)^2.$$

Lemma 2 gives sufficient conditions for the existence of the optimal integer m^* which minimizes $V(m)$.

Lemma 2: If $c_2 > c_3\beta$, and there exists a number $m_0 \geq 0$ such that $\left.\frac{dV(m)}{dm}\right|_{m=m_0} = 0$, then there exists an integer $m^* \in \{m: m = 1, 2, 3, \dots\}$ such that m^* is the optimal number of checks when at least one check is made. If $0 \leq m_0 \leq 1$, then the optimal integer m^* that minimizes $V(m)$ is $m^* = 1$. If $m_0 > 1$, then m^* is the one of $[m_0]$ or $[m_0 + 1]$ which produces the smallest value when substituted into $V(m)$, where $[x]$ denotes the largest integer contained in x and $V(m)$ is given by (4). If $V([m_0]) = V([m_0 + 1])$, then $[m_0]$ and $[m_0 + 1]$ are equivalent optimal solutions.

The proof is supplied in Appendix B.

In summary, the procedure for finding the optimal checking schedule is as follows. First find the optimal m^* for the class of all schedules with at least one check. Lemma 2 provides sufficient conditions under which m^* exists, and if those sufficient conditions are satisfied, a method is given to

compute m^* . Then, according to Lemma 1, the optimal checks are scheduled at times

$$t_k = k \left(\frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2} \beta)}{m^* + 1} \right), \quad k = 1, 2, \dots, m^*. \quad (6)$$

The expected cost under that schedule is $V(m^*)$ where $V(m)$ is given by (4). Next compute $C_0(c_2, \tau, \beta)$. If $C_0(c_2, \tau, \beta) \leq V(m^*)$, the optimal plan is to perform no checks; otherwise, the optimal plan is to schedule checks at the times given by (6).

We now present an example of determining an optimal schedule for the EAAS selenium measurement procedure, mentioned in Section 1. In order to measure selenium in human plasma by EAAS, it is necessary to remove the protein to which Se is bound in plasma. The sample is heated in an electrically controlled graphite furnace. The analyte, in aqueous solution, is applied through a small hole in the top of the graphite tube to a graphite platform located inside the tube. A common problem in this procedure is the deterioration of the graphite tube which results in poor quality analyses. Approximately every 150-250 sample analyses, or about every day or two, the tube fails and must be replaced, at a cost of \$20. Checks must be made periodically to detect this problem. In this example, the laboratory system is checked by analyzing a control sample.

We would like to determine a schedule of quality control checks to detect system failure, with the major focus being on the failure of the graphite tube. Laboratory personnel estimated the important costs to be $c_1 = \$5$ for the cost of a check, $c_2 = \$15$ for the cost per inaccurate analysis, and $c_3 = \$20$ for the cost to replace the graphite tube. The cost c_1 was based on the expense required to prepare and analyze a control sample, and c_2 was based on costs study coordinators estimated for collecting plasma samples from subjects. The mean lifetime of a graphite tube is 150 - 250 firings. We will consider time units to correspond to sample analyses, and we will assume that the failure distribution can be approximated by an exponential distribution with mean $1/\beta = 200$, i.e., $\beta = 0.005$. Note that $c_2 > c_3\beta$. Each tray loaded onto the autosampler typically contains a batch of approximately 40 samples, so we will take $\tau = 40$.

First we determine the optimal number of checks assuming there will be at least one check.

Substituting the parameter values into (4) and (5), we obtain

$$V(m) = -2400 - 2975m + 3000(m+1)(0.99333)^{\frac{m}{m+1}} \exp\left(-\frac{0.2}{m+1}\right)$$

and

$$\frac{dV(m)}{dm} = -2975 + 15(0.99333)^{\frac{m}{m+1}} \left(200 + \frac{38.6622}{m+1}\right) \exp\left(-\frac{0.2}{m+1}\right).$$

By substituting several integer values of m into $\frac{dV(m)}{dm}$, we find that a sign change occurs in the interval $(2, 3)$. Therefore $2 < m_0 < 3$. Evaluating $V(m)$ at the neighboring integers, we find $V(1) = 35.8884$, $V(2) = 32.0822$, and $V(3) = 32.6031$, so we have $m^* = 2$. We now compare the cost of this plan to the cost of making no checks. The expected cost when no checks are made is given by Equation (1): $C_0(c_2, \tau, \beta) = C_0(15, 40, 0.005) = 56.1923$. Therefore, the optimal plan is to make 2 checks, and the optimal spacing between checks is 12.89. The checks are scheduled at times 12.89 and 25.77 (between the 12th and 13th samples and between the 25th and 26th samples).

It is interesting to note that if we perform similar calculations for $\tau = 30$, we find that the optimal checking schedule consists of just one check, scheduled at time 14.33. When $\tau = 80$, the optimal number of checks is 6, and the optimal spacing between checks is 11.24. This shows that the optimal spacings do depend on the length of the monitoring period (τ), and hence spacings derived assuming an infinite time horizon will not always be optimal for finite time horizon problems.

The solution we have presented in this section has the drawback that the spacings between checks might not necessarily be integers. If the inspections can only be made at integer timepoints, the proper method of solution would involve applying nonlinear integer programming techniques to the cost function given by (3), and these techniques are often very difficult to implement. However, the minimum expected cost computed without requiring integer spacings can be used as a “yardstick” against which plans with integer spacings could be measured. In Sections 3 and 4 we consider discrete time models for developing checking schedules which would be free of this drawback.

3. Example 2: Systems With Random Checking and Sensitivity/Specificity Errors

Suppose we wish to monitor a laboratory system over a finite time period $[0, \tau]$, and it may be assumed that the failure distribution for the system, in discrete time, has a geometric distribution with mean $1/\beta$. Here time will be measured in terms of number of analyses and we will take τ to be an integer representing the number of analyses in a batch. We will broaden our definition of “failure” from the previous section to mean that the system has failed when it switches from producing inaccurate analyses at a “normal” rate p_0 to a “failed” rate of p_1 where $p_0 < p_1$,

$$p_0 = \Pr[\text{inaccurate analysis in normal mode}], \text{ and}$$

$$p_1 = \Pr[\text{inaccurate analysis in failed mode}].$$

The quality of an analysis can only be determined through a check. To simplify terminology, we will henceforth use the term “defective” to mean an inaccurate analysis. By taking $p_0 = 0$ and $p_1 = 1$ we obtain the previous type of failure model in which a system either works perfectly or not at all. Once the system has switched into the failed mode, it remains in that mode until failure is discovered through checking and the system is repaired.

For convenience, we will speak of checking analyses instead of checking the laboratory system. Checking an analysis will mean reanalyzing that sample or analyzing a control or standard immediately following that analysis. In the former case, the analysis will be declared “defective” if the initial analysis disagrees with the reanalysis (assumed error free); in the latter case, an analysis will be declared “defective” if the control or standard analyzed immediately following it produces a reading outside of the known acceptable range of readings. When an analysis is found to be defective in the course of checking, the system is repaired and the sample is reanalyzed to obtain a correct reading. The “time clock” is assumed to stop during a check, resulting in the same approximating assumption made in Section 2, namely that the duration of a check is negligible and checking does not degrade the system.

The checking schedule will be randomly determined by a checking probability f , $0 \leq f \leq 1$. Each analysis is checked with probability f , and not checked with probability $1 - f$. On the average, for a sequence of τ analyses, this will result in approximately τf equally spaced checks. We note that

this randomized checking schedule could be determined in advance so that the checking times could be programmed into an automated system.

Randomized checking schedules have several potential advantages over fixed schedules. They are less likely to be affected by cyclic variations in a process, and if there is some part of the process controlled by a technician, the randomized schedule makes it easier to blind the technician with regard to the checking schedule. A major disadvantage is that there is always some possibility of large “gaps” occurring between checks during which time system failure could go undetected for a long period.

We will assume that there is no way to distinguish between a defective produced in normal mode and a defective produced in failed mode, so a repair will be made whenever a check indicates the presence of a defective. This means that it is possible that a system which has not failed will be “repaired”. A repair initiated by the detection of a defective at time m will guarantee that the system is in normal mode at time $m + 1$.

As in Section 2, we define three costs as c_1 = cost per check, c_2 = cost per undetected defective, and c_3 = cost to repair system. We describe the monitoring process by an eight state Markov chain where time m is said to occur when the m th analysis is either checked or allowed to pass by unchecked. Each state is described by a vector with three elements where the first element is 0 if the system is in normal mode and 1 if failed; the second element is 1 if the analysis is checked and 0 if not; the third element is 1 if the analysis is defective and 0 if the analysis is nondefective. For example, being in state $(0, 1, 0)$ at time m means that the system is in normal mode and the m th analysis is checked and found nondefective. We will denote the set of states by

$$\mathcal{S} = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}.$$

We define a set of transition probabilities over the set $\mathcal{S}$ as follows:

Let

$$P_{(i,j,k),(i',j',k')} = \Pr[\text{in state } (i',j',k') \text{ at time } m+1 / \text{in state } (i,j,k) \text{ at time } m].$$

Then

$$\begin{aligned}
P_{(0,0,0),(0,0,0)} &= (1 - \beta)(1 - f)(1 - p_0) & P_{(0,0,0),(0,0,1)} &= (1 - \beta)(1 - f)p_0 \\
P_{(0,0,0),(0,1,0)} &= (1 - \beta)f(1 - p_0) & P_{(0,0,0),(0,1,1)} &= (1 - \beta)fp_0 \\
P_{(0,0,0),(1,0,0)} &= \beta(1 - f)(1 - p_0) & P_{(0,0,0),(1,0,1)} &= \beta(1 - f)p_0 \\
P_{(0,0,0),(1,1,0)} &= \beta f(1 - p_0) & P_{(0,0,0),(1,1,1)} &= \beta fp_0 \\
P_{(0,0,1),(i,j,k)} &= P_{(0,0,0),(i,j,k)} & & \text{for } (i,j,k) \in \mathcal{Y} \\
P_{(0,1,0),(i,j,k)} &= P_{(0,0,0),(i,j,k)} & & \text{for } (i,j,k) \in \mathcal{Y} \\
P_{(0,1,1),(0,0,0)} &= (1 - f)(1 - p_0) & P_{(0,1,1),(0,0,1)} &= (1 - f)p_0 \\
P_{(0,1,1),(0,1,0)} &= f(1 - p_0) & P_{(0,1,1),(0,1,1)} &= fp_0 \\
P_{(0,1,1),(1,0,0)} &= P_{(0,1,1),(1,0,1)} = P_{(0,1,1),(1,1,0)} = P_{(0,1,1),(1,1,1)} = 0 \\
P_{(1,0,0),(0,0,0)} &= P_{(1,0,0),(0,0,1)} = P_{(1,0,0),(0,1,0)} = P_{(1,0,0),(0,1,1)} = 0 \\
P_{(1,0,0),(1,0,0)} &= (1 - f)(1 - p_1) & P_{(1,0,0),(1,0,1)} &= (1 - f)p_1 \\
P_{(1,0,0),(1,1,0)} &= f(1 - p_1) & P_{(1,0,0),(1,1,1)} &= fp_1 \\
P_{(1,0,1),(i,j,k)} &= P_{(1,0,0),(i,j,k)} & & \text{for } (i,j,k) \in \mathcal{Y} \\
P_{(1,1,0),(i,j,k)} &= P_{(1,0,0),(i,j,k)} & & \text{for } (i,j,k) \in \mathcal{Y} \\
P_{(1,1,1),(i,j,k)} &= P_{(0,1,1),(i,j,k)} & & \text{for } (i,j,k) \in \mathcal{Y}
\end{aligned}$$

We define the transition probability matrix P for this Markov chain to be the 8×8 matrix with these transition probabilities as elements, with the order of the rows and columns corresponding to the ordering of the states in the set $\mathcal{Y}$.

The expected cost associated with a checking plan that checks at rate f , assuming that the system in place at time 1 is new, is given by

$$\begin{aligned}
C(f) = & c_1 \sum_{n=1}^{\tau} [P_{(0,1,1),(0,1,1)}^n + P_{(0,1,1),(0,1,0)}^n + P_{(0,1,1),(1,1,1)}^n + P_{(0,1,1),(1,1,0)}^n] \\
& + c_2 \sum_{n=1}^{\tau} [P_{(0,1,1),(0,0,1)}^n + P_{(0,1,1),(1,0,1)}^n] \\
& + c_3 \sum_{n=1}^{\tau} [P_{(0,1,1),(0,1,1)}^n + P_{(0,1,1),(1,1,1)}^n]
\end{aligned} \tag{7}$$

where

$$P_{(i,j,k),(i',j',k')}^n = \text{Pr}[\text{in state } (i',j',k') \text{ at time } m+n / \text{in state } (i,j,k) \text{ at time } m].$$

By computing the spectral decomposition (Ayres 1962, p. 171) of the transition probability matrix, P , we can find expressions for the n -step transition probabilities. An expression for the n -step transition probability matrix, P^n , is given in Appendix C. Now we can compute the cost function $C(f)$ by substituting into (7) expressions for $P_{(0,1,1),(i,j,k)}^n$ calculated by the method in Appendix C. After much algebraic manipulation we obtain

$$\begin{aligned}
C(f) = & \frac{\tau}{\lambda - 1} \left\{ [(c_2 - c_3)p_0p_1(1 - \beta) + c_1(\beta p_0 - p_1)]f^2 \right. \\
& + [(c_2 - c_3)\beta p_1 - c_2p_0p_1(1 - \beta) - c_1\beta]f - c_2\beta p_1 \Big\} \\
& + \frac{(1 - \lambda^\tau)}{(1 - \lambda)^2} \beta (p_0 - p_1)(fp_0 - 1)[(c_2 - c_3)f - c_2]
\end{aligned} \tag{8}$$

where $\lambda = (\beta p_0 - p_1)f - \beta + 1$.

Numerical studies, performed for a variety of values for $c_1, c_2, c_3, p_0, p_1, \beta$, and τ , suggest that this function has no sharp peaks or valleys, and we should be able to find a close approximation to the $f \in [0, 1]$ which minimizes $C(f)$ by searching over a fine grid of values or by using some other more sophisticated numerical search procedure.

We have also found a simple method for finding the minimizing f which seems to provide very good approximations. The method only requires finding the roots of a cubic equation and has the advantage that it can easily be performed with just a handheld calculator as compared with the techniques mentioned above which would best be handled by a computer. Even if an exact optimum is desired, this approximation will provide a good starting point for a more extensive search.

The cubic equation used for the simple method was obtained by calculating the derivative of the cost function, $\frac{dC(f)}{df}$, setting it equal to zero, and dropping all terms which were multiplied by λ^τ , which tends to zero as τ tends to infinity because $|\lambda| < 1$. The result is a cubic equation given by

$$f^3 + a_1 f^2 + a_2 f + a_3 = 0$$

where

$$\begin{aligned} a_0 &= \tau(\beta p_0 - p_1)^2 \left\{ (c_3 - c_2) p_0 p_1 (1 - \beta) - c_1 (\beta p_0 - p_1) \right\} \\ a_1 &= 3 a_0^{-1} \tau(\beta p_0 - p_1) \beta \left\{ (c_2 - c_3) p_0 p_1 (1 - \beta) + c_1 (\beta p_0 - p_1) \right\} \\ a_2 &= a_0^{-1} \left\{ -c_2 p_0 p_1 \beta^2 \tau(p_0 - p_1 + 2 - 3\beta) - c_2 p_1^2 \beta \tau(p_1 - p_0 + \beta) \right. \\ &\quad \left. - c_2 p_0 (p_0 - p_1) \beta (\beta p_0 - p_1) + (c_2 - c_3) (p_0 - p_1) \beta (\beta p_0 + p_1) \right. \\ &\quad \left. - c_3 p_1 \beta^2 \tau(3 p_0 \beta - 2 p_0 - p_1) - 3 c_1 \beta^2 \tau(\beta p_0 - p_1) \right\} \\ a_3 &= a_0^{-1} \left\{ c_2 p_1 \beta^2 \tau(p_0 - \beta - p_1) + c_2 (p_0 - p_1) \beta (\beta p_0 - 2 p_1 - \beta) \right. \\ &\quad \left. + c_3 p_1 \beta^3 \tau + c_3 (p_0 - p_1) \beta^2 + c_1 \beta^3 \tau \right\}. \end{aligned}$$

The roots of this equation are given by (Spiegel 1968, p. 32)

$$\begin{aligned} f_1 &= S + T - \frac{a_1}{3} \\ f_2 &= -\frac{1}{2}(S + T) - \frac{a_1}{3} + \frac{i\sqrt{3}}{2}(S - T) \quad (i = \sqrt{-1}) \\ f_3 &= -\frac{1}{2}(S + T) - \frac{a_1}{3} - \frac{i\sqrt{3}}{2}(S - T) \end{aligned} \tag{9}$$

where

$$Q = \frac{1}{9}[3a_2 - a_1^2] \quad R = \frac{1}{54}[9a_1 a_2 - 27a_3 - 2a_1^3]$$

$$S = \left(R + \sqrt{Q^3 + R^2} \right)^{\frac{1}{3}} \quad T = \left(R - \sqrt{Q^3 + R^2} \right)^{\frac{1}{3}}.$$

Any of f_1, f_2, f_3 that are real numbers are approximate zeros of $\frac{dC(f)}{df}$ (for moderately large τ).

$C(f)$ is continuous everywhere on the interval $[0, 1]$. Its only discontinuity occurs at $f = \frac{\beta}{\beta p_0 - p_1}$ which is where $\lambda = 1$. Assuming that $\beta < 1$ (i.e., the system does not fail instantly), this value of f is negative since $p_0 < p_1$, so it is outside of the interval $[0, 1]$. Therefore $C(f)$ attains a global minimum on the interval $[0, 1]$, and this minimum must occur at one of the zeros of $\frac{dC(f)}{df}$ or at one of the endpoints, 0 or 1. Hence, to find the approximate value(s) of f which minimize $C(f)$, we need only check the points 0 and 1, and any of f_1, f_2, f_3 that are real numbers contained in the interval $[0, 1]$.

We now present an example of determining the optimal checking rate, f , for the EAAS selenium measurement procedure which was discussed in Section 2. We will consider the same costs and

parameters as in Section 2, namely $c_1 = \$5$, $c_2 = \$15$, $c_3 = \$20$, $\beta = 0.005$, $\tau = 40$. The probability of a defective was estimated to be $p_0 = 0.05$ when the system is in normal mode, and $p_1 = 0.95$ when the system is in failed mode.

We computed f in two ways. In the first way, we computed $C(f)$, given by (8), at all points on the grid $0.00, 0.01, 0.02, \dots, 0.99, 1.00$. The grid point producing the minimum was $f = 0.06$.

In the second way, we used our approximate method based on the roots of the cubic equation given by (8). The roots f_2 and f_3 were complex, and the real root was $f_1 = 0.063$. Evaluating $C(f)$ at the values $f = 0, 1$, and 0.063 , we found that $f = 0.063$ produced the minimum. In fact, we found that the approximate method gave a better value of f than the grid method we used with mesh size 0.01 .

$$C(0.063) = 70.39076 < 70.41258 = C(0.06)$$

Therefore, each analysis will be checked with probability 0.063 . On the average, we will expect to perform $40 \times 0.063 = 2.52$, or 2 - 3 checks. This is approximately the same answer obtained in Section 2 assuming a continuous time model. If p_0 were further from 0 and p_1 were further from 1, we might not expect the results from Sections 2 and 3 to be so similar.

4. A Dynamic Programming Method for Computing Optimal Checking

Schedules for Laboratory Systems

In this section we present a dynamic programming formulation for the problem of finding optimal checking schedules for monitoring systems, with some known failure distribution, over a finite time period. We discuss an algorithm for solving for the optimal schedule. General dynamic programming techniques are discussed, for example, in Bertsekas (1976).

The method of solution we present is computationally more difficult than the methods presented in Sections 2 and 3 and generally require the use of a computer. However, the dynamic programming method supplies an alternative when the exponential and geometric failure distributions and their implied "memoryless" properties are not appropriate.

We formulate the problem in discrete time with the same cost structure used in preceding sections, namely $c_1 = \text{cost per check}$, $c_2 = \text{cost per undetected defective}$, $c_3 = \text{cost to repair system}$. We assume that the failure time can be represented by a discrete random variable, T , defined over the positive integers. As in Section 3, it is assumed that the “time clock” stops during checks. When we speak of checking the system at a particular time, we will mean checking the analysis that was performed at that time, and checking will retain its general meaning given in Section 3. It is assumed that when a defective analysis is discovered through a check, the system is repaired and the sample is reanalyzed to obtain a correct reading.

It is assumed that the probability mass function for T is known, and we define the quantities

$$p_{j/k} = \Pr[T = j / T > k].$$

The system is monitored over the time period $[0, \tau]$. We define the quantities

$U_{k,n}(m) =$ expected cost of scheduling the next check in m time units when the age of the current system is k and n analyses remain to be monitored. ($m = 0$ corresponds to no further checks.)

Then the optimal value function is defined by the quantities

$$V_{k,n} = \min_{m \in Z_n} U_{k,n}(m)$$

where $Z_n = \{0, 1, \dots, n\}$. We also define quantities $m_{k,n}$ such that

$$V_{k,n} = U_{k,n}(m_{k,n}).$$

If the initial age of the system is k_0 , then k , m , and n have the ranges

$$k = 0, 1, \dots, k_0 + \tau; \quad m = 0, 1, \dots, \tau; \quad n = 0, 1, \dots, \tau.$$

The $V_{k,n}$ must satisfy

$$V_{k,0} = 0$$

and

$$V_{k,n} = \min \left\{ \begin{aligned} & c_2 \sum_{j=k+1}^{k+n} (k+n+1-j)p_{j/k} \\ & c_1 + \min_{m \in \mathbb{Z}_{n-\{0\}}} \left(c_2 \sum_{j=k+1}^{k+m} (k+m-j)p_{j/k} + (c_3 + V_{0,n-m}) \sum_{j=k+1}^{k+m} p_{j/k} \right. \\ & \quad \left. + V_{k+m,n-m} \left(1 - \sum_{j=k+1}^{k+m} p_{j/k} \right) \right) \end{aligned} \right. \quad (10)$$

This problem can be solved using the dynamic programming technique of backwards induction as follows:

1. Find $V_{k,1}$ and $m_{k,1}$, for $k = 0, 1, \dots, k_0 + \tau$.

If $c_2 p_{k+1/k} < c_1 + c_3 p_{k+1/k}$, then $V_{k,1} = c_2 p_{k+1/k}$ and $m_{k,1} = 0$.

Otherwise, $V_{k,1} = c_1 + c_3 p_{k+1/k}$ and $m_{k,1} = 1$.

2. The remaining $V_{k,n}$ are found using (10).

Thus, we have provided a general method for obtaining the optimal checking schedule, in discrete time, for a system with a general failure distribution. This method is computationally intensive, and for large values of $k_0 + \tau$, the time and storage requirements could become quite large, requiring a more sophisticated computer implementation than the algorithm presented above. Despite the increased computational difficulty, this method does provide an “all purpose” solution when more simple models and solutions are not appropriate.

5. Discussion

It is hoped that this paper will draw more attention to the issue of laboratory data quality control and encourage further research into the area of laboratory quality control monitoring procedures. Our experiences have shown that many laboratories employ only ad hoc quality control procedures or use quality control procedures without understanding them or assessing their effectiveness. In this paper we have proposed several methods for determining (minimum cost) optimal checking schedules which are well-suited to laboratory systems and are relatively simple to compute.

Poor quality laboratory analyses can have an impact on many different levels. On an individual level, inaccurate laboratory analyses may alter the medical diagnosis and treatment of a patient. On a larger scale, conclusions drawn from scientific studies which rely heavily upon laboratory data may be weakened or invalidated due to poor quality analyses. No amount of careful attention to study design or sophisticated statistical analyses can completely compensate for poor quality laboratory data.

ACKNOWLEDGMENTS

The authors would like to thank Professor John A. Muckstadt for his many helpful comments and Professors Larry C. Clark and Gerald F. Combs, Jr. for supplying us with an interesting laboratory application.

This research was supported in part by Grant GM 28364 for the U.S. National Institutes of Health, and in part by a fellowship from the U.S. Army Research Office through the Mathematical Sciences Institute at Cornell University.

REFERENCES

- Adams, B. M. and Woodall, W. H. (1989). An analysis of Taguchi's on-line process control procedure under a random walk model. *Technometrics* 31, 401-413.
- Amenta, J. S. (1968). Analysis of variance, control charts, and the clinical laboratory. *American Journal of Clinical Pathology* 49, 842-849.
- Arnold, B. F. (1987). Minimax principle and control chart design. *Frontiers in Statistical Quality Control* 3; H. J. Lenz, G. B. Wetherill, P. Th. Wilrich (editors), 229-241.
- Ayres, F., Jr. (1962). *Schaum's Outline Series: Theory and problems of matrices*. New York: Schaum Publishing Company.
- Banerjee, P. K. and Rahim, M. A. (1988). Economic design of $\bar{X}$ -control charts under Weibull shock models. *Technometrics* 30, 407-414.
- Barlow, R. E., Hunter, L. C. and Proschan, F. (1963). Optimum checking procedures. *Journal of the Society for Industrial and Applied Mathematics* 11, 1078-1095.
- Bertsekas, D. P. (1976). *Dynamic Programming and Stochastic Control*. New York: Academic Press.
- Derman, C. (1961). On minimax surveillance schedules. *Naval Research Logistics Quarterly* 8, 415-419.
- Duncan, A. J. (1956). The economic design of $\bar{X}$ charts when used to maintain current control of a process. *Journal of the American Statistical Association* 51, 228-242.

- Griffin, D. F. (1968). Systems control by cumulative sum method. *American Journal of Medical Technology* 11, 644-650.
- Hayre, L. S (1982). A simple dynamic method for obtaining inspection schedules. *Journal of Operational Research Society* 33, 1031-1034.
- Kander, Z. (1978). Inspection policies for deteriorating equipment characterized by N quality levels. *Naval Research Logistics Quarterly* 25, 243-255.
- Klein, M. (1962). Inspection-maintenance-replacement schedules under Markovian deterioration. *Management Science* 9, 25-32.
- Levey, S. and Jennings, E. R. (1950). The use of control charts in the clinical laboratory. *American Journal of Clinical Pathology* 20, 1059-1066.
- Limonard, C. B. G. (1980). *Quality control in clinical chemistry and dynamics of analytical processes*. Ph. D. Dissertation, De Katholieke Universiteit te Nijmegen.
- Lorenzen, T. J. and Vance, L. C. (1986). The economic design of control charts: A unified approach. *Technometrics* 28, 3-10.
- Luss, H. and Kander, Z. (1974). Inspection policies when duration of checkings is non-negligible. *Operational Research Quarterly* 25, 299-309.
- McCall, J. J. (1965). Maintenance policies for stochastically failing equipment: A survey. *Management Science* 11, 493-524.

McShane, L. M. (1989). *Statistical quality control procedures for monitoring laboratory analyses*. Ph.D. Dissertation, Cornell University, Ithaca, New York.

Montgomery, D. C. (1980). The economic design of control charts: A review and literature survey. *Journal of Quality Technology* 12, 75-87.

Munford, A. G. and Shahani, A. K. (1972). A nearly optimal inspection policy. *Operational Research Quarterly* 23, 373-379.

Nix, A. B. J., Rowlands, R. J., and Kemp, K. W. (1987). Internal quality control in clinical chemistry. *Statistics in Medicine* 6, 425-439.

Pierskalla, W. P. and Voelker, J. A. (1976). A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly* 23, 353-388.

Pignatiello, J. J., Jr. and Tsai, A. (1988). Optimal economic design of $\bar{X}$ -control charts when cost model parameters are not precisely known. *IIE Transactions* 20, 103-110.

Riddick, J. H., Flora, R. and Van Meter, Q. L. (1972). Computerized preparation of two-way analysis of variance control charts. *Clinical Chemistry* 18, 250-257.

Sherif, Y. S. and Smith, M. L. (1981). Optimal maintenance models for systems subject to failure - A review. *Naval Research Logistics Quarterly* 28, 47-74.

Spiegel, M. R. (1968). *Schaum's Outline Series: Mathematical handbook of formulas and tables*. New York: McGraw Hill.

Taguchi, G., Elsayed, E. A., and Hsiang, T. (1989). *Quality Engineering in Production Systems*.
New York: McGraw Hill.

APPENDIX A

Proof of Lemma 1

To find the spacings $\vec{d} = (d_1, d_2, \dots, d_m)$ which are optimal (minimum cost), we treat m as fixed and examine the gradient vector and Hessian matrix for $C(d_1, d_2, \dots, d_m)$. The elements of the gradient vector are

$$\frac{\partial C(\vec{d})}{\partial d_i} = (c_3\beta - c_2)\exp[-\beta d_i] + c_2\exp[-\beta(\tau - \sum_{k=1}^m d_k)]; \quad i = 1, 2, \dots, m,$$

and the Hessian matrix is given by

$$\begin{aligned} H(\vec{d}) = & \vec{1}_m \vec{1}_m' c_2 \beta \exp[-\beta(\tau - \sum_{k=1}^m d_k)] \\ & + (c_2 - c_3\beta) \beta \text{diag}(\exp[-\beta d_1], \exp[-\beta d_2], \dots, \exp[-\beta d_m]) \end{aligned}$$

where $\vec{1}_m$ is an $m \times 1$ vector with all elements equal to 1 and $H(\vec{d})$ is an $m \times m$ matrix defined at a point $\vec{d}_0$ by

$$\left((H_{ij}(\vec{d}_0)) \right) = \left(\left(\frac{\partial^2 C(\vec{d})}{\partial d_i \partial d_j} \right) \right) \Big|_{\vec{d} = \vec{d}_0}.$$

When $c_2 > c_3\beta$, $\vec{d}^{(m)} = \left(\frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2}\beta)}{m+1}, \frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2}\beta)}{m+1}, \dots, \frac{\tau + \frac{1}{\beta} \log(1 - \frac{c_3}{c_2}\beta)}{m+1} \right)$ is the point where the gradient vector is zero. To show that $C(d_1, d_2, \dots, d_m)$ achieves a unique minimum at the point $\vec{d}^{(m)}$ when $c_2 > c_3\beta$, it suffices to show that $C(d_1, d_2, \dots, d_m)$ is strictly convex. Strict convexity can be proven by showing that $H(\vec{d})$ is positive definite for all $\vec{d}$.

Let $\vec{x} = (x_1, x_2, \dots, x_m)$ be any $m \times 1$ nonzero vector. Using the expression given above for $H(\vec{d})$ we have

$$\begin{aligned} \vec{x}' H(\vec{d}) \vec{x} &= \vec{x}' \left\{ \vec{1}_m \vec{1}_m' c_2 \beta \exp[-\beta(\tau - \sum_{k=1}^m d_k)] \right. \\ &\quad \left. + (c_2 - c_3\beta) \beta \text{diag}(\exp[-\beta d_1], \exp[-\beta d_2], \dots, \exp[-\beta d_m]) \right\} \vec{x} \\ &= c_2 \beta \exp[-\beta(\tau - \sum_{k=1}^m d_k)] \left(\sum_{k=1}^m x_k \right)^2 + (c_2 - c_3\beta) \beta \sum_{k=1}^m x_k^2 \exp[-\beta d_k]. \end{aligned}$$

Observe that when $c_2 > c_3\beta$ the above quantity is strictly positive for all $\vec{x}$ and $\vec{d}$, so $\vec{x}' H(\vec{d}) \vec{x} > 0$.

This proves that H is positive definite for all $\vec{d}$ when $c_2 > c_3\beta$, and hence that $C(d_1, d_2, \dots, d_m)$ is strictly convex. Therefore $C(d_1, d_2, \dots, d_m)$ achieves a unique minimum at the point $\vec{d}^{(m)}$ when

$c_2 > c_3\beta$. $\square$

APPENDIX B

Proof of Lemma 2

When $c_2 > c_3\beta$, $\frac{d^2 V(m)}{dm^2} > 0$ in the interval $[0, \infty)$, so viewing m as a continuous variable, $V(m)$ is strictly convex in that interval. Let m_0 be a number such that $\left. \frac{dV(m)}{dm} \right|_{m=m_0} = 0$ and $m_0 \geq 0$. Then $V(m)$ assumes a minimum over $[0, \infty)$ at m_0 . If $0 \leq m_0 \leq 1$, then $V(m)$ is increasing on $[1, \infty)$, and the minimum over $\{m: m = 1, 2, 3, \dots\}$ is achieved at $m^* = 1$. If $m_0 \geq 1$, then $V(m)$ increases to the left and right of m_0 and the minimum over $\{m: m = 1, 2, 3, \dots\}$ is achieved at $[m_0]$ or $[m_0 + 1]$, where $[x]$ is defined as the largest integer less than or equal to x . If $V([m_0]) = V([m_0 + 1])$, then $m^* = [m_0]$ and $[m_0 + 1]$ are equivalent optimal solutions, otherwise m^* is the one of $[m_0]$ and $[m_0 + 1]$ which produces the smallest value of $V(m)$. $\square$

APPENDIX C

Expressions for n-Step Transition Probabilities

The n -step transition probability matrix can be decomposed as

$$\mathbf{P}^n = \mathbf{U}_1 + \lambda^n \mathbf{U}_2$$

where $\lambda = (\beta p_0 - p_1)f - \beta + 1$,

$$\mathbf{U}_1 = \frac{1 - \lambda - p_1 f}{\beta(f p_0 - 1)(1 - \lambda)} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -(p_0 - 1)p_1 f(f - 1) \\ p_0 p_1 f(f - 1) \\ (p_0 - 1)p_1 f^2 \\ -p_0 p_1 f^2 \\ \beta(p_1 - 1)(p_0 f - 1)(f - 1) \\ -\beta p_1(p_0 f - 1)(f - 1) \\ -\beta(p_1 - 1)(p_0 f - 1)f \\ \beta p_1(p_0 f - 1)f \end{bmatrix}^T$$

and

$$U_2 = \frac{1}{p_1 f \lambda (1 - \lambda)} \begin{bmatrix} f p_1 + (\lambda - 1)(1 - \beta) \\ f p_1 + (\lambda - 1)(1 - \beta) \\ f p_1 + (\lambda - 1)(1 - \beta) \\ \beta(f p_0 - 1) \\ f p_1 \\ f p_1 \\ f p_1 \\ \beta(f p_0 - 1) \end{bmatrix} \begin{bmatrix} -(p_0 - 1)p_1 f(f - 1) \\ p_0 p_1 f(f - 1) \\ (p_0 - 1)p_1 f^2 \\ -p_0 p_1 f^2 \\ p_1(p_1 - 1)f(f - 1) \\ -p_1^2 f(f - 1) \\ -p_1(p_1 - 1)f^2 \\ p_1^2 f^2 \end{bmatrix}^T,$$

where the symbol “ $\top$ ” indicates that the vector should be transposed.

$P^n_{(i,j,k),(i',j',k')}$ is given by the element of P^n located in the row corresponding to the state (i,j,k) and the column corresponding to the state (i',j',k') .