Journal of Applied Sciences

ISSN 1812-5654

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 7 (23): 3618-3627, 2007
ISSN 1812-5654

A Decision Making Framework in Production Processes Using Bayesian Inference and Stochastic Dynamic Programming

Seyed Taghi Akhavan Niaki and Mohammad Saber Fallah Nezhad
Department of Industrial Engineering, Sharif University of Technology,
P.O. Box 11155-9414, Azadi Ave., Tehran, Iran

Abstract: In order to design a decision-making framework in production environments, in this study, we use both the stochastic dynamic programming and Bayesian inference concepts. Using the posterior probability of the production process to be in state λ (the hazard rate of defective products), first we formulate the problem into a stochastic dynamic programming model. Next, we derive some properties for the optimal value of the objective function. Then, we propose a solution algorithm. At the end, the applications and the performances of the proposed methodology are demonstrated by two numerical examples.

Key words: Production processes, Bayesian inference, stochastic dynamic programming, gamma distribution

INTRODUCTION

While a production process produces items, an operator or controller observes it over time and by the quality of the output classifies the process to be either in a good or in a bad state. At the beginning of each period the operator must make one of two decisions; do nothing (continue) and accept having defective products or renew (replace) the system and pay a fixed cost (halt). The process is stochastically deteriorating over time, i.e., during one period if the process is in the good state, there is a constant probability that it will be in the bad state during the next period. Decisions and state transitions are considered instantaneous. The objective is to maximize the expected discounted value of the total future profits. This model, which represents a partially observable Markov decision problem, has been discussed by many researchers and its applications in many areas can be found by Monahan (1982), Ross (1983), White (1988), Valdez-Flores and Feldman (1989), Scarf (1997) and Wang (2002).

We intuitively expect that the quality of the output in the good state to be higher than in the bad state. There are two popular ways of modeling this notion: (a) stochastic dominance, i.e., the quality of the output is stochastically higher in good state than in the bad state and (b) dominance in expectation, i.e., the expected quality of the output in the good state is higher than in the bad state. In the application of these models, it can be shown that the optimal policy initiates a maintenance (or a replacement) of the operating device if the degree of its deterioration is greater than or equal to a critical level. Such a policy is usually called control-limit policy (Kyriakidis and Dimitrakos, 2006). In other words, the optimal policy has a control limit and the optimal decision is continue if and only if the probability that the process is in the good state exceeds the control limit.

While Albright (1979), Bertsekas (1976), Lovejoy (1987) and White (1979) used the stochastic dominance condition in their modeling, Grosfeld-Nir (2007) showed that the dominance in expectation suffices for the optimality of a control limit policy; making the partially observable Markov decision problem more applicable.

Tagaras (1988) studied the joint process control and machine maintenance problem of a Markovian deteriorating machine. Assuming that sampling and preventive maintenance were performed at fixed intervals, he searched the best $\bar{X}$ control chart limits, preventive maintenance interval and sampling interval to minimize the time average maintenance and quality control related cost numerically.

Kuo (2006) studied the joint machine maintenance and product quality control problem in which both the timing of the sampling action and the sample size were directly included in the action space of the dynamic programming model of the system. Unlike previous studies in this area, he did not impose a mandatory fixed sample size and fixed sampling intervals on the system. Instead, he let the dynamic programming mechanism dictated the best sample size and sampling epoch based on the current state of the system. He derived some properties of the objective function that minimized the

Corresponding Author: Seyed Taghi Akhavan Niaki, Department of Industrial Engineering, Sharif University of Technology, P.O. Box 11155-9414, Azadi Ave., Tehran, Iran Tel: (+9821) 66165740 Fax: (+9821) 66022702

expected total discounted system cost in the value iteration algorithm of the dynamic programming model.

In many realworld decision-making problems (like the ones in which we either continue or halt a production process, replace or repair a specific machine, whether or not the available data comes from a certain probability distribution and so forth), first we divide all of the probable solution space into smaller subspaces (the solution is one of the subspaces), then we assign a probability measure to every subspace considering our experiences and finally based on the current situation we update the probabilities and make the decision. In production environments, one of the probability measures in a decision-making problem is the time between producing defective products. Assuming a certain probability distribution, based upon the information from a sample taken from the process at any time, we may estimate the parameters of this distribution. If the value of this parameter is less than a given threshold, we will halt the process.

PROBLEM STATEMENT

Assuming that the true state of a production process at any stage can be indirectly measured in term of number of defective items produced by the process (Sinuany-Stern *et al*., 1997) and that the time between defective products follows an exponential distribution, in this study, first we estimate the parameter of this distribution (λ as hazard rate) by a sequential decision-making framework. Then, if λ is less than a threshold, the production process continues and we accept to have a cost associated with producing defective products. Otherwise, we halt the process; accepting to pay the unsatisfied customers' cost together with corrective maintenance costs.

In order to estimate the hazard rate (λ) at any stage of the sampling process, in this research, we propose a sequential decision-making framework for high-yield production processes such that not only the total costs of unsatisfied customers, defective products and the corrective maintenance will be minimized, but also the probability of making correct decision is maximized. In other words, at the beginning of each period we want to either continue the production or supply the customers' demands based on the current process condition or to halt the process and do maintenance action while not being able to satisfy the demands.

In order to make the proposed method more realistic, we assume that λ is a random variable that follows a gamma distribution. Then based upon an objective function definition based on cost and risk we derive several properties of the optimal value function, which help us to find the optimal policy for a vendor to either satisfy the customer order or to halt the production process and consider corrective maintenance action to be taken in any period. This policy is derived based on a stochastic dynamic programming and Bayesian estimation approach that develops an optimal framework for the decision-making process at hand.

THE MODEL

In production processes, in cases where we are to decide between producing and not producing a batch, we are in stochastic state and we never can surely say that a batch should be produced or should not be produced. Since the stochastic state of the process may be dynamic, we may use the concept of the stochastic dynamic programming to model such problems.

Some researchers have developed sequential analysis inference in combination with optimal stopping problem to determine the probability of making correct decision. One of these researches is a new approach in probability distribution fitting of a given statistical data that Eshragh and Modarres (2001) named it Decision On Belief (DOB). In this decision-making method, a sequential analysis approach is employed to find the best underlying probability distribution of the observed data. Moreover, Eshragh and Niaki (2006) applied the DOB concept as a decision-making tool in response surface methodology. In this study, we use the concept of DOB to model the problem. However, before doing so, first we need to have some notations and definitions.

Notations and definitions: We will use the following notations and definitions in the rest of the study:

We illustrate an application of the proposed approach by specifying the distribution of the time to produce defective products as an exponential distribution with hazard rate λ.

Let t_i denote the time between productions of (i-1)st and (i)th defective products in a production cycle. During these failures if m defective products are produced, to use a non-informative prior by assuming that parameters of gamma converge to zero, i.e., the prior distribution of λ is gamma (0,0). Then, using Bayesian inference, the posterior distribution of λ is also gamma with parameters of m and

$$\sum_{i=1}^{m} t_i$$

(Nair *et al*., 2001). In other words:

$$f(\lambda) \sim \Gamma(\alpha = m, \beta = \sum_{i=1}^{m} t_i) \quad (1)$$

where:

f	:	The probability density function of λ.
R	:	Defined as the cost of halting production process (it includes cost of not satisfying customer order and cost of maintenance actions).
C	:	The cost of having one defective product in an order.
$V_n(\lambda)$	:	The cost associated with λ when there are n remaining stages to make the decision.
$W_n(\lambda)$	:	Defined as the probability of correct choice associated with λ when there are n remaining stages to make the decision.
d_n	:	The upper threshold for λ. If the hazard rate is more than d_n, then we halt the production process.
d'_n	:	Defined as the lower threshold for λ. If the hazard rate is less than d'_n, we continue the production process.
δ_1	:	The maximum acceptable level of the batch quality (Accepted Quality Level (AQL)).
δ_2	:	Defined as the minimum rejectable level of the batch quality (Lot Tolerance Proportion Defective (LTPD)).
λ_1	:	The maximum acceptable level of the hazard rate.
λ_2	:	Defined as the minimum rejectable level of the hazard rate.
CS	:	Is the event of making the correct decision.
ε_1	:	The size of type-one error in making a decision.
ε_2	:	Defined as the size of type-two error in making a decision.
H	:	The default time to produce the product.
D	:	The total number of products in an order.

Derivations: We may model described the decision-making problem as an optimal stopping problem in which in each stage of the decision-making process we take a sample from a batch and based on the information obtained from the sample we want to decide whether to halt or to continue the production or take more samples.

We mentioned that the hazard rate (λ) could be modeled as

$$G(\alpha = m, \beta = \sum_{i=1}^{m} t_i) \ .$$

Hence, $P(\lambda \geq d_n)$ shows the probability of halting a production process and $P(\lambda \leq d'_n)$ shows the probability of continuing a production process. Then, by use of the total probability theorem $[1-P(\lambda \geq d_n)-P(\lambda \leq d'_n)]$ shows the probability of neither halting nor continuing and hence taking more samples. We note that for the third probability not to be negative we need to have $d_n \geq d'_n$.

If we define n to be the index of the decision-making stage and λ to be the state variable, then $RP(\lambda \geq d_n)$ shows the cost when we halt the production process, $CH\lambda P(\lambda \leq d'_n)$ represents the cost when we continue the production process and $\alpha' V_{n-1}(\lambda)$ shows the cost when we continue to the next stage. It is obvious that we need the discount factor a′ to evaluate the cost of the next stage in the current stage (according to the approach of stochastic dynamic programming). Hence, we can define the stochastic dynamic equation of the cost as:

$$\begin{aligned} &E(\text{cost})=E(\text{cost}|\text{Halt})P(\text{Halt})+E(\text{cost}|\text{Continue}) \\ &P(\text{Continue})+E\begin{pmatrix}\text{cost}|\text{Going to the next} \\ \text{decision making stage}\end{pmatrix} \\ &P(\text{Going to the next decision making stage}) \end{aligned} \quad (2)$$

Then the cost associated with λ when there are n remaining stages to make the decision is:

$$V_n(\lambda)= \underset{d_n, d'_n}{\text{Min}} \left\{ \begin{aligned} &R\, P(\lambda \geq d_n)+CH\lambda\, P(\lambda \leq d'_n)+ \\ &(1-P(\lambda \geq d_n)-P(\lambda \leq d'_n))\alpha'\, V_n(\lambda) \end{aligned} \right\} \quad (3)$$

However, we defined CS as the event of correct decision, so we will have:

$$\begin{aligned} &P(CS)=P(CS|\text{Halt})P(\text{Halt})+P(CS|\text{Continue}) \\ &P(\text{Continue})+P\begin{pmatrix}CS|\text{Going to the next} \\ \text{decision making stage}\end{pmatrix} \\ &P(\text{Going to the next decision making stage}) \end{aligned} \quad (4)$$

It is obvious that $\frac{H\lambda_1}{D} = \delta_1$ and $\frac{H\lambda_2}{D} = \delta_2$ Hence, we have

$$P(CS|\text{Halt})=\int_{\lambda_2}^{\infty} f(\lambda)d\lambda$$

$$P(CS|\text{Continue})=\int_{0}^{\lambda_1} f(\lambda)d\lambda$$

Now we can define the stochastic dynamic equation of making the correct decision as:

$$W_n(\lambda)= \underset{d_n, d'_n}{\text{Max}} \left\{ \begin{aligned} &P(\lambda \geq d_n)\int_{\lambda_2}^{\infty} f(\lambda)d\lambda+P(\lambda \leq d'_n)\int_{0}^{\lambda_1} f(\lambda) \\ &d\lambda+(1-P(\lambda \geq d_n)-P(\lambda \leq d'_n))\alpha'\, W_{n-1}(\lambda) \end{aligned} \right\} \quad (5)$$

Since we are to minimize the objective function given in Eq. 3 and maximize the objective function of Eq. 5 simultaneously, based on the ratio of the cost to (1-risk) criterion we combine these two equations in one as:

$$H_n(\lambda) = \underset{d_n, d'_n}{\text{Min}} \left\{ \frac{V_n(\lambda)}{W_n(\lambda)} \right\} \quad (6)$$

It is obvious that this function should be minimized. In theorem 1, we will show that the minimum value of $H_n(\lambda)$ occurs at the boundary limits of d_n and d'_n.

Theorem 1: The optimal value of $H_n(\lambda)$ in Eq. 6 occur at the boundary limits of d_n and d'_n.

Proof: We take the first derivatives of $H_n(\lambda)$ in Eq. 6 with respect to d_n and d'_n and set them both equal to zeros. That is,

$$\text{(a)} \quad \frac{\partial H_n(\lambda)}{\partial d_n} = 0 \Rightarrow$$

$$\frac{W_n(\lambda)(-f(d_n))(R - \alpha' V_{n-1}(p)) - V_n(\lambda)(-f(d_n))\left(\int_{\lambda_2}^{\infty} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right)}{(W_n(\lambda))^2} = 0$$

$$\text{(b)} \quad \frac{\partial H_n(\lambda)}{\partial d'_n} = 0 \Rightarrow$$

$$\frac{W_n(\lambda)(f(d'_n))(H\lambda C - \alpha' V_{n-1}(\lambda)) - V_n(\lambda)(f(d'_n))\left(\int_0^{\lambda_1} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right)}{(W_n(\lambda))^2} = 0$$

In other words:

$$\text{(a)} \quad \frac{W_n(p)}{V_n(p)} = \frac{\left(\int_{\lambda_2}^{\infty} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right)}{(R - \alpha' V_{n-1}(p))} \quad (7)$$

$$\text{(b)} \quad \frac{W_n(p)}{V_n(p)} = \frac{\left(\int_0^{\lambda_1} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right)}{(H\lambda C - \alpha' V_{n-1}(\lambda))} \quad (8)$$

As Eq. 7 and 8 share a unique left hand side, their right hand sides must be equal. However, we notice that in general the right hand sides cannot be equal. Hence, we conclude that at most one of the derivatives can be equal to zero.

Assume the derivative in (a) is equal to zero, hence the equation in (b) is not equal to zero and we conclude that the optimal values of d'_n is in its boundary limits. However, if we expand equation (7), we will have:

$$P(\lambda \le d'_n) = \frac{(R - \alpha' V_{n-1}(p))\alpha' W_{n-1}(p) - \left(\int_{\lambda_2}^{\infty} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right)\alpha' V_{n-1}(p)}{(CH\lambda - \alpha' V_{n-1}(p))\left(\int_{\lambda_2}^{\infty} f(\lambda)d\lambda - \alpha' W_{n-1}(\lambda)\right) - (R - \alpha' V_{n-1}(p))\left(\int_0^{\lambda_1} f(\lambda)d\lambda - \alpha' W_{n-1}(p)\right)}$$

This is contradiction, because, we showed that the optimal values of d'_n is in its boundary limits. Hence we can conclude that none of the derivatives (7) and (8) is equal to zero and hence the optimal values of d_n and d'_n are in their boundary limits. For the condition when the derivative in (b) is equal to zero, the reasoning is similar.

In order to determine the boundary limits of d_n and d'_n, we use the concept of the first and the second type errors. First type error shows the probability of halting the production process when the hazard rate of production process is acceptable and second type error is the probability of continuing the production process when the hazard rate of production process is not acceptable. Then on one hand if $\lambda \le \lambda_1$, the probability of halting the production process will be smaller than ε_1 and on the other hand, in cases where, $\lambda \ge \lambda_2$, the probability of continuing the production process will be smaller than ε_2. Hence, as the mean of the gamma distribution is $\frac{\alpha}{\beta}$, for a process being in a good sate we have:

$$f(\lambda) \sim \Gamma(\alpha = m, \beta) \quad \Rightarrow \quad \frac{\alpha}{\beta} = \lambda_1 \quad \Rightarrow \quad \beta = \frac{\alpha}{\lambda_1}$$

In this case, the probability of halting the production process (type-one error) is

$$P(\lambda \ge d_n) = \int_{d_n}^{\infty} f(\lambda)d\lambda \le \varepsilon_1 \quad \Rightarrow$$
$$1 - F(d_n) \le \varepsilon_1 \quad \Rightarrow \quad d_n \ge F^{-1}(1 - \varepsilon_1)$$

where, $f(\lambda)$ is the probability density function of a Gamma distribution with parameters of α and $\beta = \frac{\alpha}{\lambda_2}$ and $F(d_n)$ is the cumulative probability distribution function of λ evaluated at d_n.

However, if we define th_1 to be the boundary limit of d_n, $F(d_n)$ is an increasing function and we have:

$$d_n \ge F^{-1}(1-\varepsilon_1) = th_1 \quad (9)$$

Similarly, defining th_2 to be the boundary limit of d'_n, for a production process being in a bad state we have:

$$f(\lambda) \sim \Gamma(\alpha = m, \beta) \quad \Rightarrow \quad \frac{\alpha}{\beta} = \lambda_2 \quad \Rightarrow \quad \beta = \frac{\alpha}{\lambda_2}$$

In this case, the probability of continuing the production process (type-two error) is

$$P(\lambda \le d'_n) = \int_0^{d'_n} f(\lambda)d\lambda \le \varepsilon_2 \quad \Rightarrow$$
$$F(d'_n) \le \varepsilon_2 \quad \Rightarrow \quad d'_n \le F^{-1}(\varepsilon_2)$$

where, $f(\lambda)$ is the probability density function of a gamma distribution with parameters of α and $\beta = \frac{\alpha}{\lambda_2}$ and $F(d_n)$ is the cumulative probability distribution function of λ evaluated at d'_n.
Hence

$$d'_n \le F^{-1}(\varepsilon_2) = th_2 \tag{10}$$

Now, since the optimal values of d_n and d'_n are in the boundary limits, in order to make the optimum decision, we can consider the framework given in Eq. 11 to make a decision. In this Eq.

$$\frac{m}{\sum_{i=1}^{m} t_i}$$

is the mean of the gamma distribution for λ given in Eq. 1.

In order to identify the signs of the derivatives in Eq. 11, we note that $W_n(\lambda)$ and $V_n(\lambda)$ are required to be evaluated. Besides, to obtain these functions the values of d_n and d'_n are needed. We showed that these values are at their boundaries, resulting in four cases. These cases are different combinations of the values for d_n and d'_n as $d_n = \infty$, $d_n = th_1$, $d'_n = 0$ and $d'_n = th_2$. We evaluate the objective function given in Eq. 6 by these cases and pick the one with the lowest value. Then, we compare the mean hazard rate,

$$E(\lambda) = \frac{m}{\sum_{i=1}^{m} t_i},$$

with the optimum boundary points of the objective function and make the decision based upon the framework given in (11).

In the decision-making framework given in (11), we note that whenever in a given stage of the sampling process, the expected hazard rate is either less than th_1 or greater than th_2 we continue sampling in the next stage. Since this event occurs with a low probability, the probability of not making the decision in stage n, as n becomes large goes to zero. In other words, the proposed method eventually converges to make a decision.

$$1)\ \frac{\partial H_n(\lambda)}{\partial d_n} \le 0, \frac{\partial H_n(\lambda)}{\partial d'_n} \le 0 \Rightarrow d_n = \infty, d'_n = th_2 \Rightarrow \begin{cases} \text{a) if } \frac{m}{\sum_{i=1}^{m} t_i} \le th_2 \Rightarrow \text{continue the production process} \\ \text{b) else, go to the next stage} \end{cases}$$

$$2)\ \frac{\partial H_n(\lambda)}{\partial d_n} \le 0, \frac{\partial H_n(\lambda)}{\partial d'_n} \ge 0 \Rightarrow d_n = \infty, d'_n = 0 \Rightarrow \text{go to the next stage} \tag{11}$$

$$3)\ \frac{\partial H_n(\lambda)}{\partial d_n} \ge 0, \frac{\partial H_n(\lambda)}{\partial d'_n} \le 0 \Rightarrow d_n = th_1, d'_n = th_2 \Rightarrow \begin{cases} \text{a) if } th_1 \le \frac{m}{\sum_{i=1}^{m} t_i} \le th_2 \rightarrow \text{Since } d_n \ge d'_n\text{, this case} \\ \text{is not feasible and should not be considered.} \\ \text{b) if } th_2 \le \frac{m}{\sum_{i=1}^{m} t_i} \le th_1 \rightarrow \text{go to the next stage} \\ \text{c) if } th_1, th_2 \le \frac{m}{\sum_{i=1}^{m} t_i} \rightarrow \text{halt the production process} \\ \text{d) if } \frac{m}{\sum_{i=1}^{m} t_i} \le th_2, th_1 \rightarrow \text{continue the production process} \end{cases}$$

$$4)\ \frac{\partial H_n(\lambda)}{\partial d_n} \ge 0, \frac{\partial H_n(\lambda)}{\partial d'_n} \ge 0 \Rightarrow d_n = th_1, d'_n = 0 \Rightarrow \begin{cases} \text{a) if } \frac{m}{\sum_{i=1}^{m} t_i} \ge th_1 \rightarrow \text{halt the production process} \\ \text{b) else, go to the next stage} \end{cases}$$

In summary, we propose the following algorithm to solve the problem at hand:

THE SOLUTION ALGORITHM

According to what we derived in earlier section, the steps involved in the solution algorithm are:

- Based on the given values of the parameters α, β, R, H, C, D, ε_1, ε_2, δ_1 and δ_2, in the first stage, n = 1, we define $H_n(\lambda)$ using Eq. 6.
- Using Eq. 9 and 10 and by numerical integrations, next we determine th_1 and th_2 as the thresholds of d_1 and d_2, respectively.
- Knowing that the optimal value of $H_n(\lambda)$can only happen in one of the four cases ($d_1 = \infty$, $d'_1 = 0$), ($d_1 = \infty$, $d'_1 = th_2$), ($d_1 = th_1$, $d'_1 = 0$) and ($d_1 = th_1$, $d'_1 = th_2$), we evaluate $H_n(\lambda)$ at these points and pick the point with the minimum value of $H_n(\lambda)$.
- We employ the framework given in (11) to make the decision at the stage $H_{n-1}(\lambda)$. If the optimal decision is to go to the next stage, then we go to step 5. Else, we stop the decision making process.
- Set n = n + 1 and determine the optimal value of $H_{n-1}(\lambda)$. Then, go to step 1.

We note that in order to evaluate the optimal value of $H_{n-1}(\lambda)$ in step 5, we need to calculate the optimal values of $H_{n-2}(\lambda)$, $H_{n-3}(\lambda)$,... and $H_1(\lambda)$.

The flowchart given in Fig. 1 summarizes the steps involved in the proposed algorithm.

Numerical example 1: In this example, the parameters are set such that the optimal decision is made in stage 2 of the decision-making framework. Suppose $\alpha = 5$, $\beta = 80$, R = 100, H = 1000, C = 1, D = 1000, $\varepsilon_1 = 0.05$, $\varepsilon_2 = 0.1$, $\delta_1 = 0.04$ and $\delta_2 = 0.1$. Knowing that $\lambda \sim \Gamma(\alpha = 5, \beta = 80)$, in the first step of the algorithm we define:

$$H_1(\lambda) = \frac{100\,P(\lambda \ge d_1) + 62.5\,P(\lambda \le d_1') + \left(1 - P(\lambda \ge d_1) - P(\lambda \le d_1')\right)\alpha'\,V_0(\lambda)}{0.099P(\lambda \ge d_1) + 0.37P(\lambda \le d_1') + \left(1 - P(\lambda \ge d_1) - P(\lambda \le d_1')\right)\alpha'\,W_0(\lambda)}$$

In the second step, using Eq. 9 and 10 for the production process to be in good and bad states, respectively, we have:

$$f(\lambda) \sim \text{gamma}\left(5, \frac{5}{0.04} = 125\right) \text{ and } d_n \ge F^{-1}(1 - \varepsilon_1) = th_1 \tag{12}$$

$$f(\lambda) \sim \text{gamma}\left(5, \frac{5}{0.1} = 50\right) \text{ and} \Rightarrow d'_n \le F^{-1}(\varepsilon_2) = th_2 \tag{13}$$

which are numerically evaluated for $th_1 = 0.091$ and $th_2 = 0.048$.

Then in the third step of the algorithm, we evaluate the objective function for different possible boundary values of d_1,d'_1 and then choose d_1 and d'_1 that minimizes the objective function, i.e:

$$\begin{aligned}
&d_1 = 0.091, d'_1 = 0 \Rightarrow H_1(0.0625) = 1003.69\\
&d_1 = \infty, d'_1 = 0.048 \Rightarrow H_1(0.0625) = 193.98\\
&d_1 = \infty, d'_1 = 0 \Rightarrow H_1(0.0625) \Rightarrow \text{no answer}\\
&d_1 = 0.091, d'_1 = 0.048 \Rightarrow H_1(0.0625) = 253.11
\end{aligned}$$

Hence, the optimum values for d_1 and d'_1 are $d_1 = \infty$, $d'_1 = 0.048$.

In the fourth step of the solution algorithm, since the expected hazard rate is equal to the mean of the Gamma distribution with parameters $\alpha = 5$, $\beta = 80$, that is 0.0625, we are in state (1-a) of the decision tree in Eq. 11 and should continue to the next stage.

In stage n = 2, assume $\alpha = 1$ and $\beta = 280$. According to the solution algorithm, first we should determine

$$V_1\left(\lambda = \frac{10}{220} = 0.0454\right) \text{ and } W_1(0.0454).$$

In the second step, using Eq. 9 and 10 for good and bad states we have

$$\begin{aligned}
&f(\lambda) \sim \text{Gamma}\left(10, \frac{10}{0.04} = 250\right) \text{ and}\\
&d_n \ge F^{-1}(1 - \varepsilon_1) = th_1 \Rightarrow th_1 = 0.0628
\end{aligned} \tag{14}$$

$$\begin{aligned}
&f(\lambda) \sim \text{Gamma}\left(10, \frac{10}{0.1} = 100\right) \text{ and}\\
&\Rightarrow d'_n \le F^{-1}(\varepsilon_2) = th_2 \Rightarrow th_2 = 0.0622
\end{aligned} \tag{15}$$

Then in the third step of the algorithm, we evaluate the objective function for different possible boundary values of d_1,d'_1 and then choose d_1 and d'_1 that minimizes the objective function, i.e.:

$$\begin{aligned}
&d_1 = 0.036, d'_1 = 0 \Rightarrow H_1(0.0454) = 66448.18\\
&d_1 = \infty, d'_1 = 0.024 \Rightarrow H_1(0.0454) = 117.67\\
&d_1 = \infty, d'_1 = 0 \Rightarrow H_1(0.0454) \Rightarrow \text{no answer}\\
&d_1 = 0.036, d'_1 = 0.024 \Rightarrow H_1(0.0454) = 152.56
\end{aligned}$$

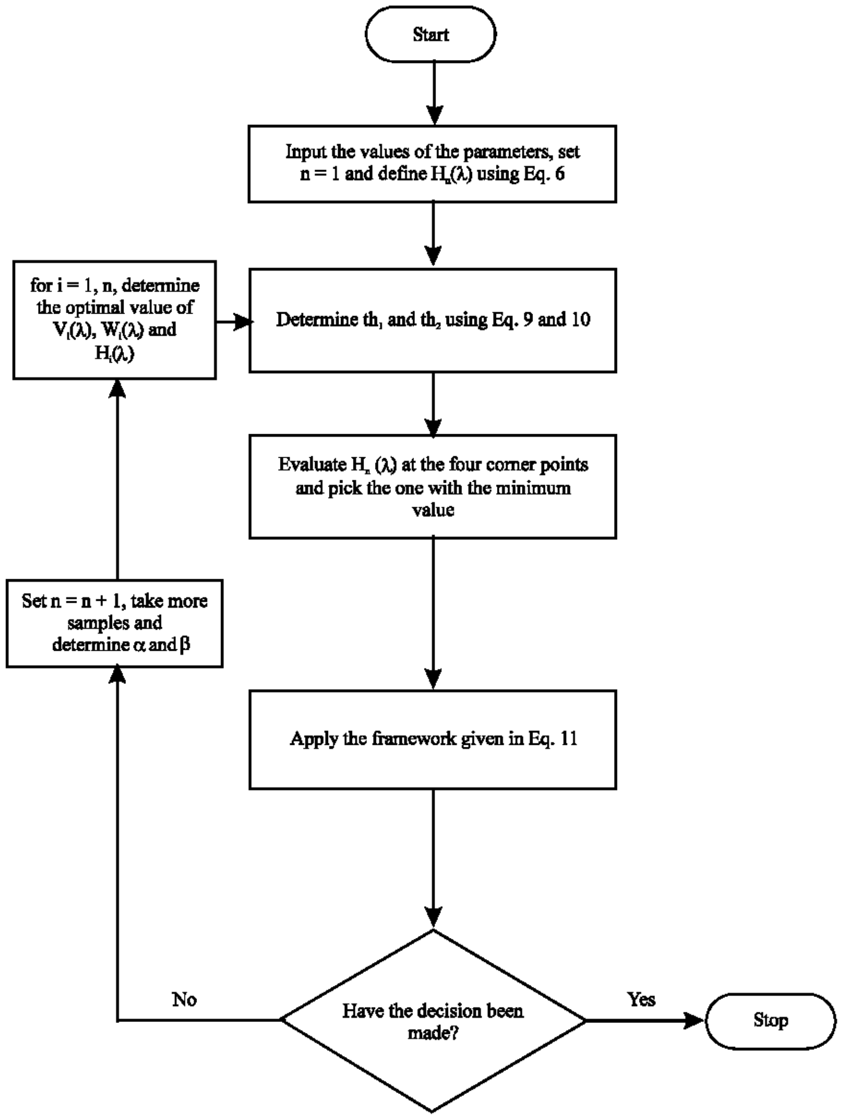

Fig. 1: The flowchart of the proposed algorithm

Hence, the optimum values for d_1 and d'_1 are $d_1 = \infty$, $d'_1 = 0.0622$, $V_1(0.05454) = 39.76$ and $W_1(0.0454) = 0.337$, which enables us to calculate the values of $V_2(0.0454)$, $W_2(0.0454)$.

Then in the fourth step of the algorithm, we evaluate the objective function for different possible boundary values of d_2, d'_2 and then choose d_2 and d'_2 that minimizes the objective function, i.e.:

$$d_2 = 0.0628, d'_2 = 0 \Rightarrow H_2(0.0454) = 156.85$$
$$d_2 = \infty, d'_2 = 0.0622 \Rightarrow H_2(0.0454) = 117.82$$
$$d_2 = \infty, d'_2 = 0 \Rightarrow H_2(0.0454) = 117.98$$
$$d_2 = 0.0628, d'_2 = 0.0622 \Rightarrow H_2(0.0454) = 136.21$$

In the fifth step of the solution algorithm, since the expected hazard rate is equal to the mean of the Gamma distribution with parameters $\alpha = 10$, that is 0.0454, we are in state (1-a) of the decision tree in Eq. 11 and should continue the production process.

Numerical example 2: In this example, the optimal solution is to halt the production process at the first stage of the sampling process. Suppose $\alpha = 5$, $\beta = 40$ and other parameters are the same as the numerical example 1.

In the second step, using Eq. 9 and 10 for a good and a bad process, respectively, we obtain

$$f(\lambda) \sim \text{Gamma}\left(5, \frac{5}{0.04} = 125\right) \text{ and} \\ d_n \geq F^{-1}(1-\varepsilon_1) = th_1 \Rightarrow th_1 = 0.091 \quad (16)$$

$$f(\lambda) \sim \text{Gamma}\left(5, \frac{5}{0.1} = 50\right) \text{ and} \\ \Rightarrow d'_n \leq F^{-1}(\varepsilon_2) = th_2 \Rightarrow th_2 = 0.048 \quad (17)$$

Then, in the third step of the algorithm, we evaluate the objective function for different possible boundary values of d_1, d'_1 and then choose d_1 and d'_1 that minimizes the objective function, i.e:

$$\begin{aligned} &d_1 = 0.091, d'_1 = 0 \Rightarrow H_1(0.125) = 159.02 \\ &d_1 = \infty, d'_1 = 0.048 \Rightarrow H_1(0.125) = 5278.2 \\ &d_1 = \infty, d'_1 = 0 \Rightarrow H_1(0.125) \Rightarrow \text{no answer} \\ &d_1 = 0.091, d'_1 = 0.048 \Rightarrow H_1(0.125) = 168.59 \end{aligned}$$

Hence, the optimum values for d_1 and d'_1 are $d_1 = 0.091$, $d'_1 = 0$.

In the fourth step of the solution algorithm, since the expected hazard rate is equal to the mean of the Gamma distribution with parameters $\alpha = 5$, $\beta = 40$, that is 0.125, we are in state (4-a) of the decision tree in Eq. 11 and should halt the production process and do maintenance action.

An error study: Here, we investigate the performance of the proposed method in terms of type-one and type-two error. To do this, let us consider the simplest case where we have only one stage for the decision-making process (n = 1). The objective function for this stage is:

$$H_1(\lambda) = \frac{R\,P(\lambda \geq d_n) + CH\lambda\,P(\lambda \leq d'_n)}{P(\lambda \geq d_n)\int_{\lambda_2}^{\infty} f(\lambda)d\lambda + P(\lambda \leq d'_n)\int_0^{\lambda_1} f(\lambda)d\lambda}$$

It can be easily shown that to minimize $H_1(\lambda)$, either $P(\lambda \geq d_n)$ or $P(\lambda \leq d'_n)$ should be equal to zero. To prove it, assume

$$\frac{R}{\int_{\lambda_2}^{\infty} f(\lambda)d\lambda} \leq \frac{CH\lambda}{\int_0^{\lambda_1} f(\lambda)d\lambda}.$$

If in the minimum value of $\underset{d_n, d'_n}{\text{Min}}\{H_1(\lambda)\}$, both $P(\lambda \geq d_n)$ and $P(\lambda \geq d'_n)$ are more than zero, then we have

$$\begin{aligned} &\underset{d_n, d'_n}{\text{Min}}\{H_1(\lambda)\} \leq \underset{d_n, 0}{\text{Min}}\{H_1(\lambda)\} \Rightarrow \\ &\frac{R\,P(\lambda \geq d_n) + CH\lambda\,P(\lambda \leq d'_n)}{P(\lambda \geq d_n)\int_{\lambda_2}^{\infty} f(\lambda)d\lambda + P(\lambda \leq d'_n)\int_0^{\lambda_1} f(\lambda)d\lambda} \\ &\leq \frac{R}{\int_{\delta_2}^{1} f(\lambda)d\lambda} \Rightarrow \frac{R}{\int_{\lambda_2}^{\infty} f(\lambda)d\lambda} \geq \frac{CH\lambda}{\int_0^{\lambda_1} f(\lambda)d\lambda} \end{aligned}$$

which is contradiction. Hence, we should only consider two cases:

Case 1: $d_1 \geq th_1$, $d'_1 = 0$

Case 2: $d'_1 \leq th_2$, $d_1 = \infty$

For given values of R = 100, C = 1, D = 1000, a defective production rate of 0.05, type-two error = 0.1, $\lambda_1 = 0.05$ and $\lambda_2 = 0.1$, based on the information from some simulated samples of different sizes, suppose we want to estimate the size of type-one error of the proposed method. Table 1 shows the results of this estimation.

The first column of Table 1 shows samples with different sizes taken for evaluation. The second column is the threshold value of the hazard rate of the production process. It means that if the hazard rate (defective product rate) is less than this threshold, we continue producing items. In column three, the probability of the right decision (continuing the production process) has been given. This probability has been calculated by Gamma distribution. The results of Table 1 show that, as expected, the type-one error associated with the performance of the proposed methodology decreases as the sample size increases. However, as the cost increases with an increase in the sample size, we need to determine the optimum value of the sample size.

Based on the information given in Table 1, we may estimate the probability of making correct decision (accepting the batch and continue the process) as a function of sample size. The regression function using excel software is shown in Fig. 2, in which the coefficient of determination is 0.8356.

If C_1 denotes the fixed cost associated with a sample of size one, R denotes the cost of halting the process, then using the regression function $y = 0.0243\ln(x) + 0.8989$ and y denotes the probability of correct decision, hence the cost function of taking a sample of size x is:

$$g(x) = C_1 x + (1-(0.0243\ln(x) + 0.8989))\,R$$

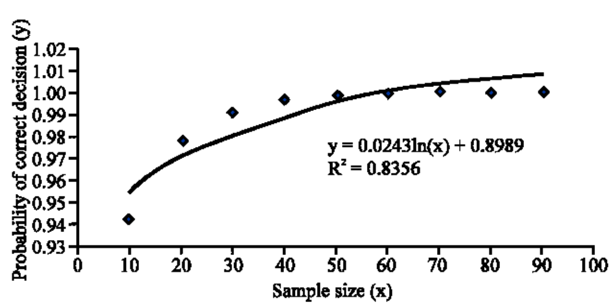

Fig. 2: Regression function for the probability of correct decision

Table 1: Type-one error estimation of the proposed method for different samples

Sample size	Rejection threshold	Probability of accepting the batch	Type-one error
10	0.077	0.942113	0.057887
20	0.075	0.978127	0.021873
30	0.074	0.990765	0.009235
40	0.074	0.996570	0.003430
50	0.074	0.998699	0.001301
60	0.074	0.999500	0.000500
70	0.073	0.999684	0.000316
80	0.073	0.999868	0.000132
90	0.073	0.999945	5.54E-05

Table 2: Type-two error estimation of the proposed method for different samples

Sample size	Probability of accepting the batch (type two error)
10	0.246920
20	0.124781
30	0.065763
40	0.039253
50	0.023899
60	0.014745
70	0.006926
80	0.004189
90	0.002548

which has its minimum value at

$$x^* = \frac{C_1}{0.0243R}.$$

In order to estimate the type-two error of the proposed method, let a process to have a defective product rate of 0.1. Then, the probability of making a wrong decision (continuing the production process) has been calculated based upon different sample sizes and is given in Table 2. The results of Table 2 indicate that as the sample size increases the probability of accepting a batch with a wrong defective rate decreases; implying a similar trade-off between the costs of sampling and the probability of correct decision.

The evaluation study for the cases in which the decision is made in stage $n>1$ can be performed in a similar way.

CONCLUSIONS

In this research, we applied Bayesian inference and stochastic dynamic programming to model a decision-making problem in production environments in which we observe the time between breakdowns based on the produced defective items. Assuming the time between breakdowns follow an exponential distribution with parameter λ, to estimate λ at any stage of the sampling process, we proposed a sequential decision-making framework such that not only the total costs of unsatisfied customers, defective products and the corrective maintenance will be minimized, but also the probability of making correct decision is maximized. In order to demonstrate the application of the proposed framework, we provided two numerical examples.

For further research, we propose either to consider some other objective functions or to employ other functions to model the state of the system. Moreover, we can employ other functions to model the probability of correct choice when the production process is accepted or rejected.

REFERENCES

Albright, S.C., 1979. Structural results for partially observable Markov decision processes. Operat. Res., 27: 1041-1053.

Bertsekas, D., 1976. Dynamic Programming and Stochastic Control. Academic Press, New York.

Eshragh, J.A. and M. Modarres, 2001. A new approach to distribution fitting: Decision on beliefs. Proceedings of the 53rd ISI Session, Seoul, Korea.

Eshragh, J.A. and S.T.A. Niaki, 2006. The application of decision on beliefs in response surface methodology. Proceedings of the 4th International Industrial Engineering Conference, Tehran, Iran.

Grosfeld-Nir, A., 2007. Control limits for two-state partially observable Markov decision processes. Eur. J. Operat. Res., 182: 300-304.

Kuo, Y., 2006. Optimal adaptive control policy for joint machine maintenance and product quality control. Eur. J. Operat. Res., 171: 586-597.

Kyriakidis, E.G. and T.D. Dimitrakos, 2006. Optimal preventive maintenance of a production system with an intermediate buffer. Eur. J. Operat. Res., 171: 86-99.

Lovejoy, W., 1987. Ordered solution for dynamic programs. Math. Operat. Res., 12: 269-276.

Monahan, G.E., 1982. A survey of partially observable Markov decision processes: Theory, models and algorithms. Manage. Sci., 28: 1-16.

Nair, V.N., B. Tang and L. Xu, 2001. Bayesian inference for some mixture problems in quality and reliability. J. Qual. Technol., 33: 16-28.

Ross, S.M., 1983. Introduction to Stochastic Dynamic Programming. Academic Press, London, UK.

Scarf, P.A., 1997. On the application of mathematical models in maintenance. Eur. J. Operat. Res., 99: 493-506.

Sinuany-Stern, Z.S., I. David and S. Biran, 1997. An efficient heuristic for a partially observable Markov decision process of machine replacement. J. Comput. Operat. Res., 24: 117-126.

Tagaras, G., 1988. Integrated cost model for the joint optimization of process control and maintenance. J. Operat. Res. Soc., 39: 757-766.

Valdez-Flores, C. and R.M. Feldman, 1989. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. Naval Res. Logistics, 36: 419-446.

Wang, H., 2002. A survey of maintenance policies of deteriorating systems. Eur. J. Operat. Res., 139: 469-489.

White, C., 1979. Optimal control-limit strategies for a partially observed replacement problem. Int. J. Syst. Sci., 10: 321-331.

White, D.J., 1988. Further real applications of Markov decision processes. Interfaces, 18: 55-61.