



ISSN 1823-626X

Journal of Fundamental Sciences

available online at <http://jfs.ibnusina.utm.my>

Optimal stopping rules for exponential data

Mohammad Saber Fallah Nezhad^{1*}, Mohammad Modarres², Seyed Taghi Akhavan Niaki²

¹Department of Industrial Engineering, Faculty of Engineering, Yazd University, Yazd, Iran

²Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran

Received 25 March 2011, Revised 5 April 2011, Accepted 10 May 2011, Available online 20 May 2011.

ABSTRACT

A problem of sequential sampling from an Exponential Distribution is considered in this research. The problem is formulated in the stochastic dynamic programming framework and the objective is to determine a control policy maximizing the total expected reward. It is assumed that under standard assumptions the control limit policy is optimal. Two types of optimal stopping problems are considered. First one is the problem of sampling without recall that once the decision maker cannot return to that observation at a later time, the second type of optimal stopping problems is sampling with recall where the decision maker can select any observation which he has taken earlier.

|Stochastic dynamic programming | Optimal stopping problem | Exponential distribution |

© 2011 Ibnu Sina Institute. All rights reserved.

1. INTRODUCTION

The problem to be treated in this research is in the category of problems of optimal stopping. In each of these problems, the statistician takes observation sequentially, and at each stage he must decide whether to stop and suffer a specified stopping risk or continue and the next observation at some specified sampling cost.

Two types of optimal stopping problems are considered. First one is the problem of sampling without recall that once the decision maker can not return to that observation at a later time, the second type of optimal stopping problems is sampling with recall where the decision maker can select any observation which he has taken earlier.

The optimal stopping problems have been studied by Macqueen and Miller [1]. Sakaguchi [2] considered problems involve sampling from a distribution with unknown mean. Kramer and Starr [3] characterized the optimal stopping rule and its distribution in a size dependent search. Application of optimal stopping rule in size dependent search has been extended by Bellout [4].

In this research, posterior state of the process after taking each sample has been determined through Bayesian inference.

The Bayesian process control focuses on determining the optimal control policy based on the posterior probability minimizing the total expected cost over a finite horizon, or the long-run expected average cost.

Makis [5] applied Bayesian control through POMDP (partially observable Markov decision process) framework in a multivariate process mean control problem. He developed the optimal stopping rule through control limit policy which minimizes the total expected cost.

The early contributions to Bayesian process control are the models by Girshick and Rubin [6] and Bayesian control design are considered by Tagaras and Nikolaidis [7], Makis and Jiang [8].

One application of this model is Stock-Option Model (Ross [9]). Suppose that you own an option to buy one share of the stock at a fixed price, say c , and you have N days in which to exercise the option. You need never exercise it, but if you do it at a time when the stock's price is t , then your profit is $t-c$. What strategy maximizes your expected profit?

If $V_n(t)$ denote the maximal expected profit when the stock' price is t and the option has n additional days to run, then $V_n(t)$ satisfies the optimality equation

$$V_n(t) = \max \left\{ t - c, \int V_{n-1}(t) f(t) dt \right\}$$

With the boundary condition

$$V_0(t) = \max \{ t - c, 0 \}$$

where $f(t)$ denotes the distribution function of t .

The problem is formulated in the stochastic dynamic programming framework and the objective is to determine a control policy maximizing the total expected reward. It is assumed that under standard assumptions the

Corresponding author at: ¹Department of Industrial Engineering, Faculty of Engineering, Yazd University, Yazd, Iran, E-mail addresses: fallahnezhad@yazduni.ac.ir

control limit policy is optimal. Also the posterior distribution of the state of the system is determined through Bayesian approach.

This paper is organized as follows: Problems of sampling without recall comes in the second section. Properties of optimal stopping rule for the problems of sampling with recall come in section three. The conclusions are presented in the final section.

2. SAMPLING WITHOUT RECALL

2.1 Sampling without Recall with finite stages

We assume that a sequential random sample T_1, T_2, \dots is taken from an exponential distribution for which it is known that the d.f. is F . We suppose that there is constant upper bound n ($n > 1$) on the number of observations which can be taken. At the j th stage of the sampling process ($j=1, 2, \dots, n-1$), after the values $T_1 = t_1, T_2 = t_2, \dots, T_j = t_j$ have been taken, the decision maker either can stop and accept the value t_j as his reward or continue sampling and take T_{j+1} . If he has not stopped earlier, he must stop and accept the final observed value t_n as his reward. This is a problem of sampling without recall, once the decision maker cannot return to that observation at a later time.

Let N denote the random number of observations which have been taken under a stopping rule. The problem is to find an optimal stopping rule that maximizes the expected reward $E(T_N)$.

At any stage of the sampling process, the state of process is denoted by the value of the most recent observation gathered, t , and the optimal stopping rule is $V_{n-i}(t_i)$. Since the decision maker must stop and accept the final observed value t as his reward, it is concluded that $V_0(t) = t$ for $-\infty < t < \infty$.

Since the decision maker is taking samples without recall from a known distribution, the value of expected reward v_j from taking one more observation and then continuing depends on the number of j remained observations in an optimal stopping rule. After one more observation t has been gathered, the expected reward from the optimal continuation over the remaining $j-1$ stages is $V_{j-1}(t)$, hence the expected reward v_j satisfies the following equation,

$$v_j = E(V_{j-1}(t)) = \int_{-\infty}^{\infty} V_{j-1}(t) f(t) dt \tag{1}$$

the expected reward $V_j(t)$ from the optimal stopping rule is maximum of the reward from stopping t , and the expected reward from continuing, v_j , thus following is concluded,

$$V_j(t) = \max\{t, v_j\} \tag{2}$$

Since $V_0(t) = t$ and from equations (1) and (2), the functions V_1, V_2, \dots and the numbers v_1, v_2, \dots are evaluated successively.

Let $v_j, j=0, 1, 2, \dots, n$, denote the maximum expected reward. It can be proven that (Degroot [10]) the optimal stopping rule for problem is to continue sampling whenever an observed value $t_j < v_{n-j}$ and to stop the process elsewhere. Values of v_j can be determined from following equations.

$$\begin{aligned} v_1 &= E(t) \\ v_{j+1} &= v_j + T_F(v_j) \\ T_F(v_j) &= \int_{v_j}^{\infty} (t - v_j) dF(t) \end{aligned} \tag{3}$$

2.2. Sampling without Recall with infinite stages

In this sampling scheme, there is no upper bound on the number of observations which can be taken but there is a fixed cost per observation. Let T_1, T_2, \dots be a sequential random sample from distribution $F(t)$. If the decision maker stops the sampling process after the values $T_1 = t_1, T_2 = t_2, \dots, T_n = t_n$ have been taken, his reward will be $t_n - cn$. Thus problem is to find a stopping rule which maximizes $E(t_N - cN)$. It can be proven that if the variance of the $F(t)$ is finite, then the maximum expected reward v^* among all stopping rules is finite and there is an optimal stopping rule whose expected reward is v^* (Degroot [10]).

After the first observation t has been taken, the decision maker either can stop the sampling process or can continue to take more observations. If he stops the sampling, his reward will be the values of t minus the sampling cost c also if the decision maker continues the sampling, then he is in the same position that he was at the start of the sampling process but he has already spent c for the first observation, thus the expected reward from the optimal continuation is again v^* minus the cost c of the observation gathered.

Thus after the first observation t has been observed, if $T = t < v^*$ then the optimal stopping rule is to continue sampling process and if $T = t \geq v^*$ the optimal stopping rule is to terminate the sampling process. The expected reward from this optimal procedure is $E(\max\{t, v^*\}) - c$, Also, the expected reward from the optimal stopping rule is assumed to be v^* . Thus the following equation is concluded,

$$v^* = E \left(\max \{t, v^*\} \right) - c \quad (4)$$

Assuming $f(t)$ is a exponential distribution with parameter λ

$$f(t) \sim \Gamma(1, \lambda) = \lambda \exp(-\lambda t) \quad (5)$$

The expected reward can be calculated as follows,

$$v^* = E \left(\max \{t, v^*\} \right) - c = \quad (6)$$

$$\int_0^{v^*} v^* f(t) dt + \int_{v^*}^{\infty} t f(t) dt - c$$

Thus,

$$\int_0^{v^*} (t - v^*) f(t) dt = \frac{\exp(-\lambda v^*)}{\lambda} = c \quad (7)$$

$$\Rightarrow v^* = \frac{\ln(\lambda c)}{-\lambda}$$

2.3. Sampling without Recall from an Exponential Distribution with Unknown Mean

Assume that a random sample T_1, T_2, \dots is taken at a cost of c from an exponential distribution with unknown mean (λ is not known). Let the prior distribution of λ is a Gamma distribution. Thus, the posterior distribution at each stage of the sampling process will be a Gamma distribution. The decision maker should find a stopping rule that maximizes the expected reward $E(t_N - cN)$.

It can be proven that if the variance of distribution of each observation is finite then an optimal stopping rule exists and the expected reward of the optimal stopping rule is finite (Degroot [10]). This assumption is satisfied in the problem of sampling from an exponential distribution which has been considered here.

Assume that the prior distribution of λ is a Gamma Distribution, 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).

$$g(\lambda) \sim \Gamma(0, 0) \quad (8)$$

Observations $T_1 = t_1, T_2 = t_2, \dots, T_n = t_n$ are taken, It can be shown by Bayesian Inference that the posterior distribution of λ is (Nair et al. [11])

$$g(\lambda) \sim \Gamma(0, 0) \Rightarrow g(\lambda) \sim \Gamma\left(n, \sum_{i=1}^n t_i\right) \quad (9)$$

Then the marginal distribution of the next observation, t , can be determined as follows:

$$f_T(t) = \int_0^{\infty} f(t|\lambda) g(\lambda) d\lambda =$$

$$\int_0^{\infty} \Gamma\left(n, \sum_{i=1}^n t_i\right) \lambda \exp(-\lambda t) d\lambda = \frac{n \left(\sum_{i=1}^n t_i\right)^n}{\left(\sum_{i=1}^n t_i + t\right)^{n+1}} \quad (10)$$

Hence the mean of random variable T is:

$$\mu_n = \frac{\left(\sum_{i=1}^n t_i\right)}{n} \quad (11)$$

At each stage of the sampling process, the state of the process is characterized by the triple (t_n, μ_n, n) , where t_n is value of the last observation and μ_n is the mean value of the T and n is the number of observations have been taken.

Let $V(t_n, \mu_n, n)$ be the expected reward, without considering the amount spent on gathering samples, when the optimal stopping rule is continued from the state (t_n, μ_n, n) .

At this stage, the decision maker either can stop sampling and accept the reward t_n or can take another observation at a cost c and then continue the sampling process in an optimal procedure. Thus the following functional equation for V is concluded,

$$V(t_n, \mu_n, n) = \max\left(t_n, E\left(V\left(t, \frac{n\mu_n + t}{n+1}, n+1\right)\right) - c\right) \quad (12)$$

Let α_n defined as follows:

$$\alpha_n = E\left(V\left(t, \frac{n\mu_n + t}{n+1}, n+1\right)\right) - c \quad (13)$$

Hence the optimal stopping rule at the state (t_n, μ_n, n) is determined as follows:

If $\alpha_n \leq t_n$ then accept the reward and terminate the sampling process otherwise continue taking samples.

It can be shown that the value of α_{n+1} can be determined from the value of α_n as follows:

Lemma 1: for any value of $n > 0$, then

$$\alpha_n = \frac{\left(\sum_{i=1}^n t_i\right)^n \left(\alpha_{n+1} + \sum_{i=1}^n t_i\right)^{-n+1}}{n-1} - c + \alpha_{n+1} \quad (14)$$

Proof:

According to definition of function V ,

$V\left(t, \frac{n\mu_n + t}{n+1}, n+1\right)$ can be determined as follows:

$$V\left(t, \frac{n\mu_n + t}{n+1}, n+1\right) = \max(t, \alpha_{n+1}) \quad (15)$$

Thus α_{n+1} can be determined as follows:

$$\alpha_n = E\left(\max(t, \alpha_{n+1})\right) - c$$

$$\Rightarrow \int_0^{\alpha_{n+1}} \alpha_{n+1} f(t) dt + \int_{\alpha_{n+1}}^{\infty} t f(t) dt - c = \alpha_n \quad (16)$$

$$\Rightarrow \int_{\alpha_{n+1}}^{\infty} (t - \alpha_{n+1}) f(t) dt + \alpha_{n+1} - c = \alpha_n$$

Using distribution function $f(t)$ derived in equation (10), following is concluded,

$$\int_{\alpha_{n+1}}^{\infty} (t - \alpha_{n+1}) \frac{n \left(\sum_{i=1}^n t_i\right)^n}{\left(\sum_{i=1}^n t_i + t\right)^{n+1}} dt = \alpha_n + c - \alpha_{n+1} \quad (17)$$

$$\Rightarrow \alpha_n = \frac{\left(\sum_{i=1}^n t_i\right)^n \left(\alpha_{n+1} + \sum_{i=1}^n t_i\right)^{-n+1}}{n-1} - c + \alpha_{n+1}$$

Lemma 2: the extreme value of α_n exist and

$\lim_{n \rightarrow \infty} \alpha_n = \mu_t \ln\left(\frac{\mu_t}{c}\right)$ where μ_t is the mean of random variable t .

Proof:

Let α^* be the extreme value of α_n , thus α^* is determined as follows:

$$\alpha^* = \lim_{n \rightarrow \infty} \alpha_n =$$

$$\lim_{n \rightarrow \infty} \left(\frac{\left(\sum_{i=1}^n t_i\right)^n \left(\alpha_{n+1} + \sum_{i=1}^n t_i\right)^{-n+1}}{n-1} - c + \alpha_{n+1} \right)$$

$$= \lim_{n \rightarrow \infty} \frac{\left(\frac{\sum_{i=1}^n t_i}{n-1}\right)^n}{\left(\frac{\alpha_{n+1}}{n} + 1\right)^{n-1}} + \lim_{n \rightarrow \infty} \alpha_{n+1} - c \quad (18)$$

Since all functions in equation (18) are continuous, thus following is concluded:

$$\lim_{n \rightarrow \infty} \frac{\left(\frac{\sum_{i=1}^n t_i}{n-1}\right)^n}{\left(\frac{\alpha_{n+1}}{n} + 1\right)^{n-1}} + \lim_{n \rightarrow \infty} \alpha_{n+1} - c \quad (19)$$

$$= \lim_{n \rightarrow \infty} \frac{\mu_n}{\left(\frac{\alpha_{n+1}}{n\mu_n} + 1\right)^{n-1}} + \lim_{n \rightarrow \infty} \alpha_{n+1} - c$$

$$= \frac{\lim_{n \rightarrow \infty} (\mu_n)}{\left(\frac{\lim_{n \rightarrow \infty} \alpha_{n+1}}{n \lim_{n \rightarrow \infty} (\mu_n)} + 1\right)^{n-1}} + \alpha^* - c$$

Using mathematical methods for convergence of functions and according to the law of large numbers, it is concluded that,

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n t_i}{n-1} = \lim_{n \rightarrow \infty} (\mu_n) = \mu_t$$

Thus following is concluded,

$$\frac{\lim_{n \rightarrow \infty} (\mu_n)}{\lim_{n \rightarrow \infty} \left(\frac{\lim_{n \rightarrow \infty} \alpha_{n+1}}{n \lim_{n \rightarrow \infty} (\mu_n)} + 1 \right)^{n-1}} + \alpha^* - c$$

$$= \frac{\mu_t}{\lim_{n \rightarrow \infty} \left(\frac{\alpha^*}{n \mu_t} + 1 \right)^n} = \frac{\mu_t}{\exp \left(\frac{\alpha^*}{\mu_t} \right)}$$
(20)

From equations (18), (19) and (20), following is concluded,

$$\alpha^* = \mu_t \ln \left(\frac{\mu_t}{c} \right)$$
(21)

If we replace $\mu_t = \frac{1}{\lambda}$ in above equation, it would be similar to equation (7).

3. SAMPLING WITH RECALL FROM AN EXPONENTIAL DISTRIBUTION WITH UNKNOWN MEAN

Now, we consider the problem of sampling with recall. We assume that a random sample T_1, T_2, \dots is taken at a cost of c units per observation from a distribution $F(t)$. Assume that the decision maker can select any observation which he has taken earlier, and accept value of that observation minus the total sampling cost as his reward, thus if decision maker stops taking samples after he has taken the samples $T_1 = t_1, T_2 = t_2, \dots, T_n = t_n$, his reward is $\max \{t_1, t_2, \dots, t_n\} - nc$.

At each stage of the sampling process, the state of the process is denoted by (r, μ_n, n) , where r is the maximum value of the observations and μ_n is the average of the observations t_1, t_2, \dots, t_n and μ_t is the mean of the random variable T and n is the number of observations has been taken.

r denotes the reward to the decision maker, without considering the amount spent on taking samples, if the optimal stopping rule is to continue from this state and the value of the next observation is t , then the state (r, μ_n, n)

is transformed into the new state $\left(r(t), \frac{n\mu_n + t}{n+1}, n+1 \right)$

where $r(t) = \max \{r, t\}$.

Let $V(r, \mu_n, n)$ be the expected reward, Thus V satisfies the following equation:

$$V(r, \mu_n, n) = \max \left(r, E \left(V \left(r(t_{n+1}), \frac{n\mu_n + t_{n+1}}{n+1}, n+1 \right) \right) - c \right)$$
(22)

Where, $r(t_{n+1}) = \max \{r, t_{n+1}\}$ and t_{n+1} is the value of the next observation.

In Lemma 3 and Lemma 4, we derive some properties of optimal solution for this problem.

Lemma 3:

Suppose that $U(r, n) = \frac{\left(\sum_{i=1}^n t_i \right)^n \left(r + \sum_{i=1}^n t_i \right)^{-n+1}}{n-1}$ for given values of r and n . if $U(r, n) > c$ then $V(r, \mu_n, n) > r$.

Proof:

Let $V_1(r, \mu_n, n)$ is defined as follows:

$$V_1(r, \mu_n, n) = E \left(V \left(r(t_{n+1}), \frac{n\mu_n + t_{n+1}}{n+1}, n+1 \right) \right) - c$$
(23)

According to the definition of function V , function $V_1(r, \mu_n, n)$ will be calculated as follows:

$$V_1(r, \mu_n, n) = E(r(t_{n+1})) - c = E(\max \{r, t_{n+1}\}) - c$$
(24)

The value of $E(\max \{r, t_{n+1}\})$ can be calculated by following equation,

$$E(\max \{r, t_{n+1}\}) = r + \int_r^\infty (t_{n+1} - r) f(t_{n+1}) dt_{n+1}$$
(25)

Using the distribution function $f(t_{n+1})$ derive in equation (10), the following equation is concluded,

$$r + \frac{\left(\sum_{i=1}^n t_i \right)^n \left(r + \sum_{i=1}^n t_i \right)^{-n+1}}{n-1} = r + U(r, n)$$
(26)

Thus it is concluded that:

$$V_1(r, \mu_n, n) = r + U(r, n) - c$$
(27)

if $U(r, n) > c$ then it is concluded that $V_1(r, \mu_t, n) > r$ thus $V(r, \mu_n, n) > r$, and the optimal decision is to continue to take samples.

In the next Lemma, we prove that function $U(r, n)$ is strictly decreasing function of r, n . with this property; the optimal stopping rule for this problem can be developed.

Lemma 4: $U(r, n)$ is approximately strictly decreasing function of r, n .

Proof:

By substituting $\sum_{i=1}^n t_i$ with $n\mu_n$ where $\mu_n = \frac{\sum_{i=1}^n t_i}{n}$ in the definition of function $U(r, n)$ in equation (26), following is concluded:

$$U(r, n) = \frac{(n\mu_n)^n (r + n\mu_n)^{-n+1}}{n-1} \tag{28}$$

$$= \left(\frac{r}{n\mu_n} + 1\right)^{-n} \frac{r + n\mu_n}{n-1}$$

Since for sufficient large value of n , $\lim_{n \rightarrow \infty} \mu_n = \mu_t$ the variable μ_n is approximately constant and equal to μ_t .

Since function $\left(\frac{r}{n\mu_n} + 1\right)^{-n}$ is a decreasing function of n

and function $\frac{r + n\mu_n}{n-1}$ is a decreasing function of n , thus

function $U(r, n)$ is strictly decreasing function of n .

The first derivative of function $U(r, n)$ on r is calculated as follows,

$$\frac{\partial U(r, n)}{\partial r} = \frac{-(n\mu_n)^n}{(r + n\mu_n)^n} < 0 \tag{29}$$

Since $\frac{\partial U(r, n)}{\partial r}$ is negative, thus $U(r, n)$ is strictly decreasing function of r, n .

By lemma 3, it is concluded that if $U(r, n) > c$ in one stage then the optimal decision is to stop the sampling and accept r as reward, also by lemma 5, we know that by continuing the sampling, $U(r, n)$ decreases thus remains less than c , as we expected.

In Lemma 5, it is proved that the extreme value of $U(r, n)$ exists and is similar to the optimal stopping rule in equation (21). This means the characteristics of optimal stopping rule in the problems of sampling without recall and sampling without recall are similar.

Lemma 5: the extreme value of $U(r, n)$ exist and

$$\lim_{n \rightarrow \infty} U(r, n) = e^{-\frac{\mu_t}{r}} \mu_t.$$

Proof:

Let U^* be the extreme value of $U(r, n)$, thus U^* is determined as follows:

$$U^* = \lim_{n \rightarrow \infty} U(r, n) = \lim_{n \rightarrow \infty} \left(\left(\frac{r}{n\mu_n} + 1\right)^{-n} \frac{r + n\mu_n}{n-1} \right) \tag{30}$$

$$= \lim_{n \rightarrow \infty} \left(\frac{r}{n \lim_{n \rightarrow \infty} \mu_n} + 1 \right)^{-n} \lim_{n \rightarrow \infty} \left(\frac{r + n \lim_{n \rightarrow \infty} \mu_n}{n-1} \right)$$

$$= e^{-\frac{\mu_t}{r}} \mu_t$$

4. CONCLUSION

In this paper, a category of stopping problems problem has been investigated. In each of these problems, the observations are gathered sequentially, and at each stage he must decide whether to stop and suffer a specified stopping risk or continue and the next observation at some specified sampling cost. The problem has been formulated and analyzed in the stochastic dynamic programming framework. We obtained the optimal stopping rules for different scenarios of sampling. The different characteristics of the value function are shown and an optimal control limit policy is developed.

REFERENCES

- [1] J.B. Macqueen and R. J. Miller, Optimal persistence policies. *Operation Research*, 8(1960), 362-380.
- [2] M. Sakaguchi, Dynamic programming of some sequential sampling designs, *Journal of Mathematical analysis and applications*, 2(1961). 446-466.
- [3] M. Kramer, N. Starr, (1990), Optimal stopping in a size dependent search. *Sequential Analysis*, 9, 59-80.
- [4] D. Bellout, An adaptive rule for stopping a discovery process under time censorship, *Journal of Statistical Planning and Inference*, 137(2007), 2328-2335.
- [5] V. Makis, Multivariate Bayesian process control for a finite production run. *European Journal of Operational Research*, 194(2009), 795-806.
- [6] M.A. Girshick and H. Rubin, A Bayes' approach to a quality control model, *Annals of Mathematical Statistics*, 23 (1952), 114–125.
- [7] Tagaras, G., Nikolaidis, Y., Comparing the effectiveness of various Bayesian \bar{X} control charts. *Operations Research* 50(2002), 878-888.
- [8] V. Makis and X. Jiang, Optimal replacement under partial observations. *Mathematics of Operations Research*, 28(2003), 382-394.
- [9] S. Ross, Introduction to Stochastic Dynamic Programming, Academic Press, New York, 1983
- [10] M.H. Degroot, Optimal Statistical Decisions. Wiley Classics Library, 2004.
- [11] V. N. Nair and B. Tang, and L. Xu, Bayesian inference for some mixture problems in quality and reliability, *Journal of Quality Technology*, 33(2001), 16-28.