Base Stock Level Determination of “Insurance Type” Spares

JOHN WALKER

Nanyang Business School, Nanyang Technological University,

Singapore 639798

Abstract

Expensive renewable spares known as “insurance type” spares are often a major concern in the design and setting up of industrial, commercial and military systems. These spares, though low in demand, are critical to the system’s operation and their unavailability can lead to excessive downtime costs. Due to their nature, the (S-1,S) inventory control model provides an appropriate replenishment policy for this class of items, where S is the maximum number of spares in inventory. A (S-1,S) model with Exponential distribution of failure-free operating time at each of a *finite* number of machines and Exponential distribution of re-supply lead-time is developed. A graphical aid is presented which, for a given number of machines, indicates the range of the ratio {mean lead-time/mean failure-free operating time} for which a minimum S is required in order to satisfy a service level constraint on the service measure Pr[a spare is available at a machine stoppage due to part failure].

Key words: inventory, spares, heuristics

INTRODUCTION

Suppose that a new industrial, commercial or military system is to be purchased/constructed. The system contains several identical “machines” that are to operate within the same economic and operational environment. A critical expensive part, *unique* to the machines and also *essential* to their operation, has been identified. The decision to stock at least one spare part has been made and the (S-1,S) inventory control model is to be implemented as the replenishment policy for the part. At the time of purchase/construction of the system, S spares for the part are to be purchased from a supplier. Such critical expensive spares, which have a high probability that they will not be needed during the system lifetime, are called “insurance type” spares. In the operation of the system, for the relatively rare event that a machine may stop operating due to failure of the part, an order is placed with the supplier to obtain a new part. Upon receipt of the part, it becomes a spare (unless a machine is down in which case the part goes immediately into operation). The cost per spare at the initial purchase, although high, is usually much lower than that at subsequent orders where the supplier will, at a premium price, set up a special production run in order to expedite the order. In those cases where there is a machine stoppage due to part failure and a spare is not available from stock, the time involved in obtaining the part from the supplier causes *excessive losses*, since the machine is down during that time. In those cases where a spare is not required during the lifetime of the system, its net salvage value is negligible.

Consider a commercial flight line with a number of planes supported by a regional repair facility. At random times, a unit malfunction is discovered and the unit is removed from the plane for repair. If a spare unit is available, it is used to replace the failure and the plane is returned to service. If no spares are available, the plane is grounded until a serviceable unit is provided by the repair facility. An industrial example is given by Dhakar, Schmidt and Miller[1] in considering how many machine tools should be carried as spares in the operation of a large paper mill. Further examples of insurance type spares are given by Mitchell[2], Geurts and Moonen[3] and Burton and Jaquette[4]. A comprehensive review of the literature when spares are "re-supplied" by means of repair is given by Nahmias[5].

The decision on the initial number of spares to purchase can be developed in one of two ways. First, one can assign a carrying cost to stocks of spares and a downtime cost to shortages of spares. The number of spares can then be selected so as to minimise the expected total cost of carrying and downtime costs per period. However, estimating downtime costs is a difficult task. Therefore, in this paper, a second approach is adopted, namely to specify a desired service measure and then to select the minimum number of spares so as to satisfy the associated service level constraint. This paper uses a relatively easily understood service measure, namely the probability that at a machine stoppage due to part failure a spare is available. Moreover, because of the high downtime cost, probability values of 0.90, 0.95, and 0.99 are explicitly considered for the service level constraint in a simple graphical aid for determining the initial number of spares to purchase.

NOTATION AND ASSUMPTIONS

The notation to be used in this paper includes the following.

M = the number of machines representing sources of part failure.

S = the maximum number of spares.

$1/\lambda$ = the mean failure-free operating time of a part.

$1/\mu$ = the mean lead-time for supply of a part.

ν = λ/μ

$p_j(S,M,\nu)$ = Pr[at any arbitrary point in time j parts are on order].

$q_j(S,M,\nu)$ = Pr[at a machine stoppage due to part failure j parts are on order].

$r(S,M,\nu)$ = Pr[a spare is available at a machine stoppage due to part failure].

α = the required service level constraint i.e., $r(S,M,\nu) \geq \alpha$.

The distribution p_j is said to represent the viewpoint of an *outside observer* because it describes the distribution of states that an outside observer would see if he (or she) were to observe the system over all time. In contrast, q_j considers the viewpoint of a *failing part,* which, by definition, observes the system at its failure epoch.

The assumptions to be used include the following.

1. Each of the M machines contains the critical part with Exponential distribution of failure-free operating time and with a mean of $1/\lambda$ periods.
2. The supplier has either:
    (a) limited re-supply capacity comprising of a single “re-supply channel”; or
    (b) “ample” re-supply capacity comprising of S+M “re-supply channels."
    Each re-supply channel has Exponential distribution of lead-time with a mean of $1/\mu$ periods.
3. Over the lifetime of the process, the failure-free-operating time and lead-time distributions are stationary. At the end of the process lifetime, any spares in stock have negligible salvage value.

Much of the literature on insurance type spares e.g., Mitchell[2], Geurts and Moonen[3] deals with systems where the value of M is sufficiently large, so the source population of part failures can be treated as infinite. In such cases the mean failure rate can be treated as independent of the number of machines in operation. In this paper, finite values of M are considered and, therefore, the mean failure rate depends upon the number of machines in operation. The assumption of Exponential failure-free operating time corresponds to "accidental" part failure as a Poisson process with rate λ and is commonly used in the literature. An estimate of the value of λ may be determined from a mixture of the supplier's quoted "mean time before failure" (if available) and managerial/engineering judgement on the future economic and operational environment. It is clear that in many cases the value of λ should be regarded as a rough estimate only.

Much of the literature on insurance type spares e.g., Mitchell[2], Geurts and Moonen[3] deals with systems where there exists "ample" re-supply capacity, i.e., in the context of this paper, S+M re-supply channels are available and no order has to queue. The assumption of a single re-supply channel applies where, in the event that an order arrives when prior orders are waiting for processing and/or when the supplier is manufacturing a part, the new order joins the queue. The assumption of Exponential lead-time applies when the majority of orders are supplied according to the supplier's "relatively short" lead-time but occasionally, perhaps due to other contractual commitments, breakdowns, etc., much longer times are required. It is clear that in many cases the value of μ should be regarded as a rough estimate only.

The stationary assumption on λ and μ is clearly very strong and almost certainly erroneous. Changes in the economy causing changes in the process usage, changes in production technology causing changes in operating practice and/or the maintenance regime provide instances where the stationary assumption may be questioned. It may seem that the most appropriate approach would be to build a model of the dynamic decision process and to derive an “optimal strategy” from it. However, as stated by Geurts and Moonen[3], “...where are the data to come from to estimate the values of the parameters of even the simplest models?" The approach adopted in this paper is to provide a graphical aid indicating, for a given M and α, the range of $\nu = \lambda/\mu$ for which a minimum S is required in order to satisfy the service level constraint that the Pr[a spare is available at a machine stoppage due to part failure] $\geq \alpha$. The graphs indicate the sensitivity of the value of S to values of ν. The graphs could also provide some aid in the dynamic decision context. If the economic and/or operational environment indicates that extra machines are required and/or values of λ (μ) are to increase (decrease) then, using what evidence was available for the values of λ and μ, the graphs would indicate if the current value of S needed to be increased in order to satisfy the service level constraint.

(S-1,S) MODEL

Under assumptions 1-3, let the states j = 0, 1, ..., S+M represent the number of parts on order. The mean failure rate λ_j is as follows:

$$\lambda_j = \begin{cases} M\lambda & , 0 \leq j \leq S \\ (S+M-j)\lambda & , S+1 \leq j \leq S+M. \end{cases}$$

The mean re-supply rate is as follows:

$$\mu_j = \begin{cases} \mu & \text{for assumption 2(a)} \\ j\mu & \text{for assumption 2(b)} \end{cases}$$

Application of the standard balance flow procedure to the underlying birth-death process results in the following steady-state equations for $p_j(S,M,\nu)$.

$$\begin{aligned}
\mu_1 p_1(S,M,\nu) &= \lambda_0 p_0(S,M,\nu) \\
\mu_{j+1} p_{j+1}(S,M,\nu) &= (\lambda_j + \mu_j) p_j(S,M,\nu) - \lambda_{j-1} p_{j-1}(S,M,\nu) \qquad , 1 \le j \le S+M-1 \qquad (1) \\
\mu_{S+M} p_{S+M}(S,M,\nu) &= \lambda_{S+M-1} p_{S+M-1}(S,M,\nu)
\end{aligned}$$

Let
$$\omega_j = \begin{cases} M\nu & , 0 \le j \le S \\ (S+M-j)\nu & , S+1 \le j \le S+M. \end{cases} \qquad \text{for assumption 2(a)}$$

and
$$\omega_j = \begin{cases} [M/(j+1)]\nu & , 0 \le j \le S \\ [(S+M-j)/(j+1)]\nu & , S+1 \le j \le S+M. \end{cases} \qquad \text{for assumption 2(b)}$$

then, using (1) together with the normalising condition $\sum_{j=0}^{S+M} p_j = 1$, convenient recursive computational forms for $p_j(S,M,\nu)$ are obtained:

$$\begin{aligned}
p_0(S,M,\nu) &= 1/[1+\sum_{j=1}^{S+M} \prod_{k=0}^{j-1} \omega_k], \\
p_j(S,M,\nu) &= \omega_{j-1} p_{j-1}(S,M,\nu) \qquad , 1 \le j \le S+M.
\end{aligned}$$

Having determined $p_j(S,M,\nu)$, $q_j(S,M,\nu)$ can be obtained using the following relationship (see for example, Lavenberg and Reiser[6]):

$$q_j(S,M,\nu) = p_j(S-1,M,\nu) \qquad , 0 \le j \le S+M$$

Then $$r(S,M,\nu) = \sum_{j=0}^{S-1} q_j(S,M,\nu)$$

From the tabulated values of $r(S,M,\nu)$, it is relatively simple to determine and graph, for given M and α, the range of ν for which a given S is the minimum number of spares which satisfies the service level constraint $r(S,M,\nu) \ge \alpha$. Examples of such graphs for the limited re-supply capacity of one re-supply channel are illustrated in Figures 1-3. To use the graphs, the manager/engineer determines M, estimates λ and μ and the required value of α. For $\nu = \lambda/\mu$, plotting the point (ν,M) determines the minimum value of S which satisfies the required service level constraint and also the sensitivity of S to the value of ν. The production of graphs for the ample re-supply capacity of S+M re-supply channels is left to the interested reader.

Figure 1. Number Of Spares For A Service Level Of 90%

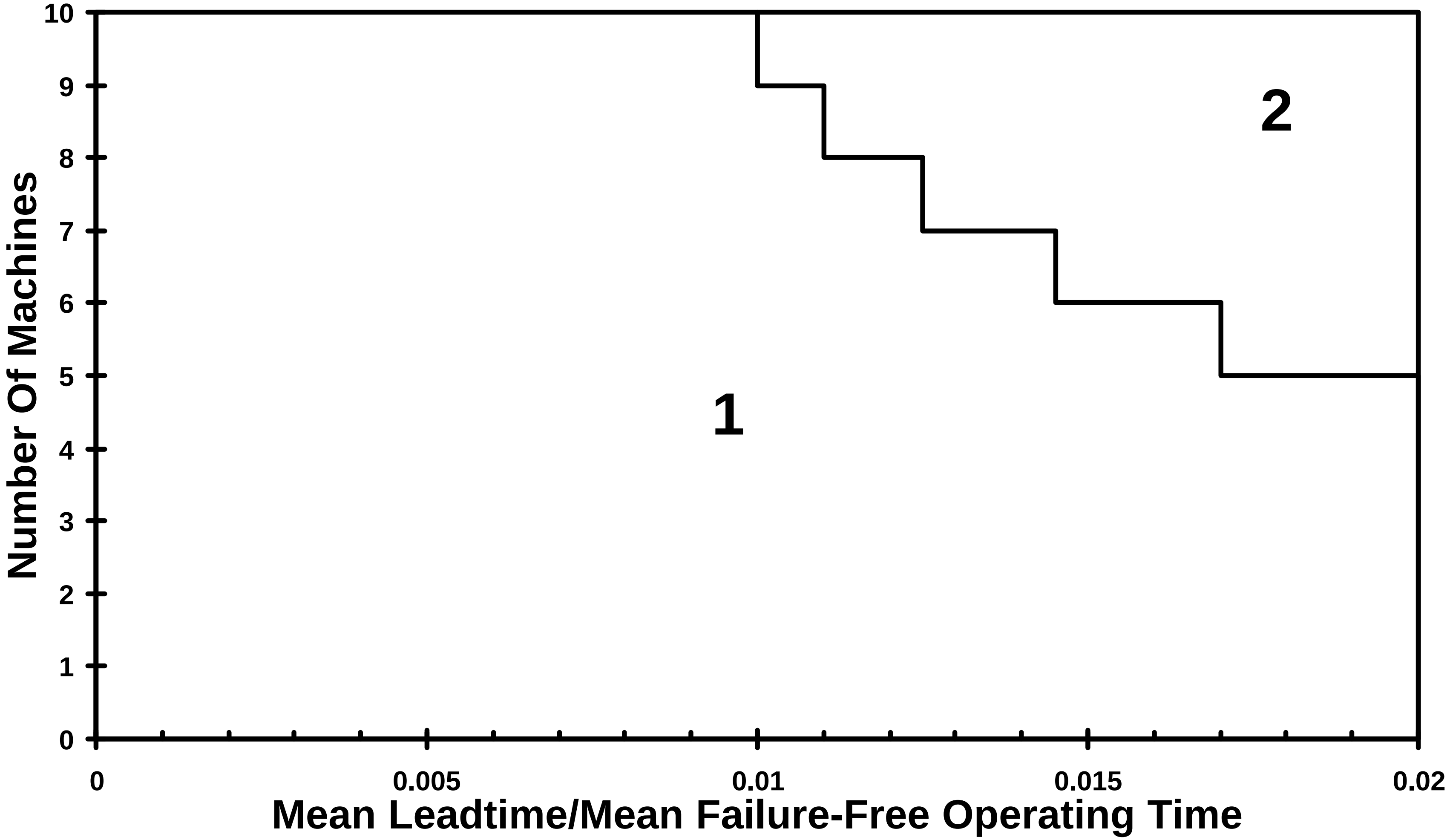


Figure 2. Number Of Spares For A Service Level Of 95%

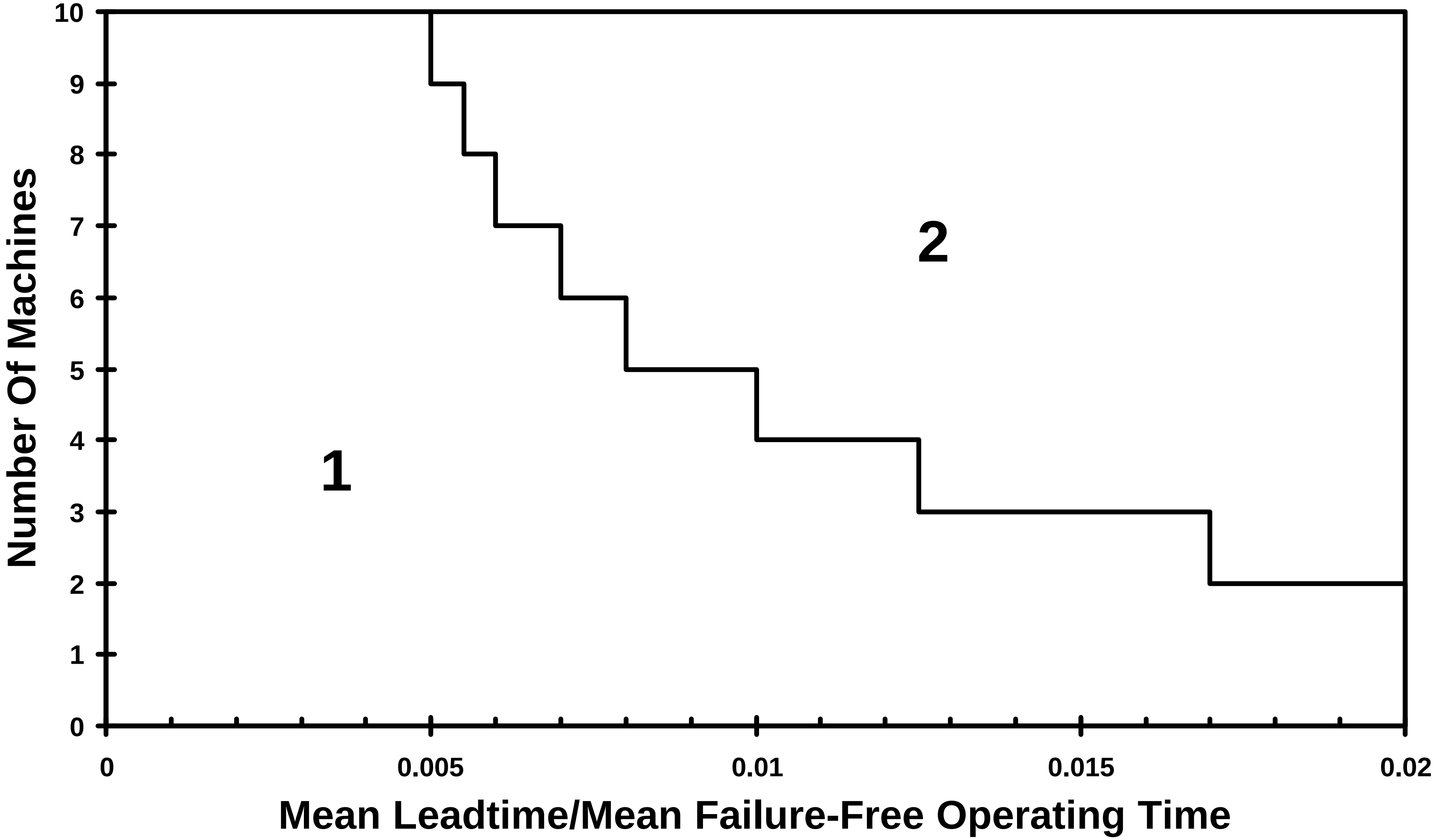
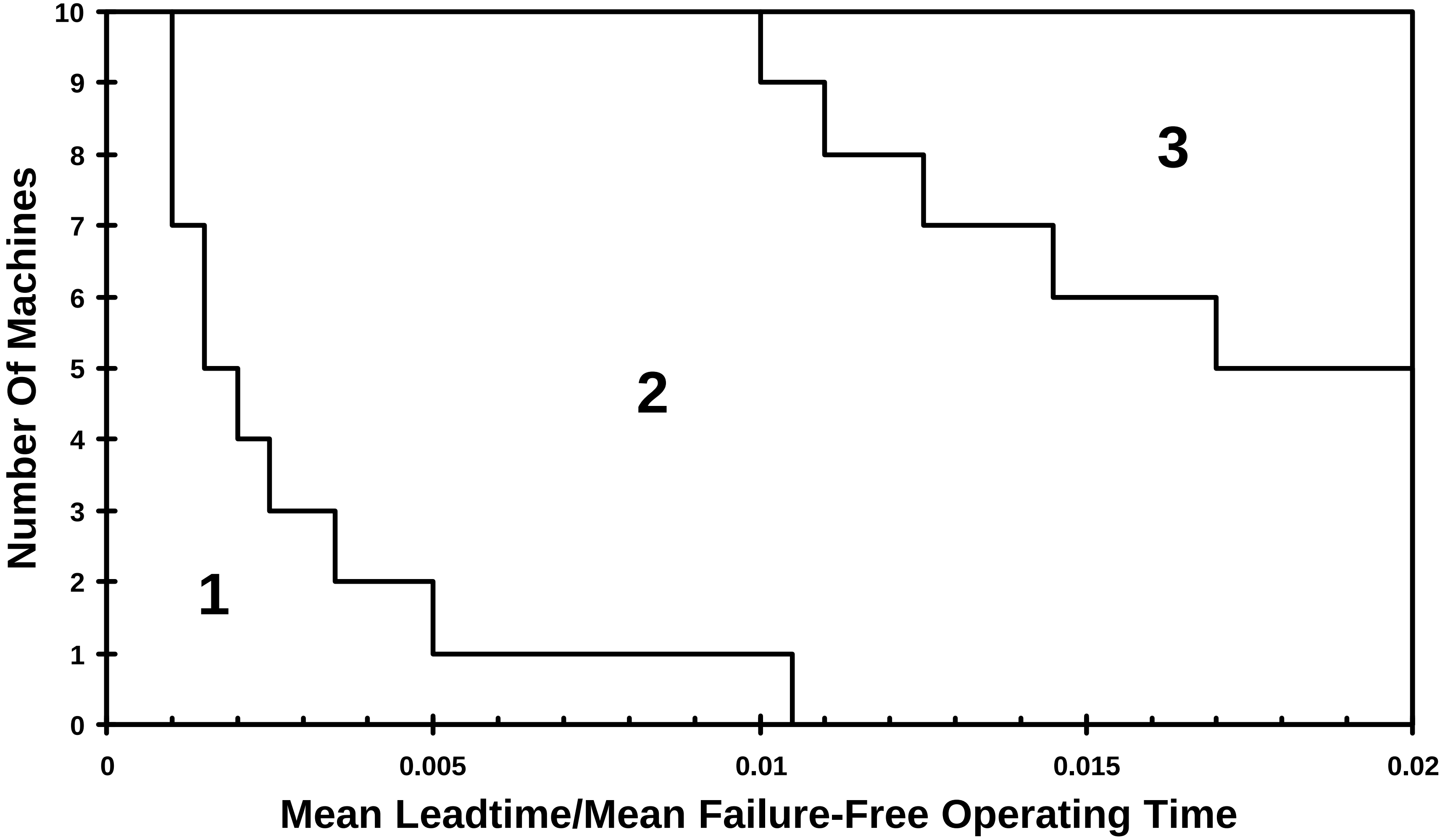


Figure 3.   Number Of Spares For A Service Level Of 99%



CONCLUSION

In this paper, a very simple graphical implementation aid has been developed for choosing the initial number of insurance type spares to purchase for systems having a finite population source of part failures. The simplicity of the method is appropriate for the poor quality of available data and indicates the sensitivity of the decision to the ratio {mean lead-time/mean failure-free operating time}.

REFERENCES

1. T. S. DHAKAR, C. P. SCHMIDT and D. M. MILLER (1994) Basestock level determination for high cost low demand critical repairable spares. *Computer Ops Res*. **21**, 411-420.
2. G. H. MITCHELL (1962) Problems of controlling slow moving engineering spares. *Opl. Res. Q.* **13**, 23-39.
3. J. H. J. GEURTS and J. M. C. MOONEN (1992) On the robustness of ‘insurance type’ spares provisioning strategies. *J. Opl. Res. Soc.* **43**, 43-51.
4. R. W. BURTON AND S. C. JAQUETTE (1973) The initial provisioning decision for insurance type items. *Naval Res. Logist. Q.* **20**, 123-146.
5. S. NAHMIAS (1981) Managing repairable item inventory systems: a review. In *Multi Level Production/Inventory Control Systems: Theory and Practice* (L. B. Schwarz, Ed.), pp. 253-277. Vol. 16, Studies in the Management Sciences, North Holland, Amsterdam
6. S. S. LAVENBERG and M. REISER (1979) Stationary state probabilities at arrival instants for closed queueing networks with multiple types of customers. *IBM Research Report RC 7592*.