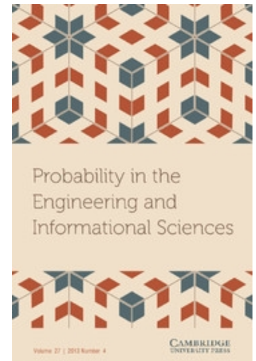

STOCHASTIC SCHEDULING WITH PREEMPTIVE-REPEAT MACHINE BREAKDOWNS TO MINIMIZE THE EXPECTED WEIGHTED FLOW TIME

Xiaoqiang Cai, Xiaoqian Sun and Xian Zhou

Probability in the Engineering and Informational Sciences / Volume 17 / Issue 04 / October 2003, pp 467 - 485
DOI: 10.1017/S0269964803174037, Published online: 24 September 2003

Probability in the Engineering and Informational Sciences, **17**, 2003, 467–485. Printed in the U.S.A.

STOCHASTIC SCHEDULING WITH PREEMPTIVE-REPEAT MACHINE BREAKDOWNS TO MINIMIZE THE EXPECTED WEIGHTED FLOW TIME

Xiaoqiang Cai

Department of Systems Engineering & Engineering Management
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
E-mail: xqcai@se.cuhk.edu.hk

Xiaoqian Sun

Department of Mathematics
Huaiyin Teachers College
Huaian, Jiangsu Province 223001, People's Republic of China
E-mail: xqsun@se.cuhk.edu.hk

Xian Zhou

Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
E-mail: maxzhou@polyu.edu.hk

We study a stochastic scheduling problem with a single machine subject to random breakdowns. We address the *preemptive-repeat* model; that is, if a breakdown occurs during the processing of a job, the work done on this job is completely lost and the job has to be processed from the beginning when the machine resumes its work. The objective is to complete all jobs so that the the expected weighted flow time is minimized. Limited results have been published in the literature on this problem, all with the assumption that the machine uptimes are exponentially distributed. This article generalizes the study to allow that (1) the uptimes and downtimes of the machine follow general probability distributions, (2) the breakdown patterns of the machine may be affected by the job being processed and are thus job dependent; (3) the processing times of the jobs are random variables following arbitrary distributions, and (4) after a breakdown, the processing time of a job may either remain a same but unknown amount, or be resampled according to its probability distribution. We derive the necessary and sufficient condition that ensures the problem

 0269-9648/03 $16.00

with the flow-time criterion to be well posed under the preemptive-repeat breakdown model. We then develop an index policy that is optimal for the problem. Several important situations are further considered and their optimal solutions are obtained.

1. INTRODUCTION

Scheduling problems involving stochastic machine breakdowns have been the subject of extensive studies in the literature for over two decades. Generally, these problems may be categorized into two types, according to the effect of a machine breakdown on the job being processed. One is featured by the so-called *preemptive-resume model*, and the other by the *preemptive-repeat model*, which differ from each other as follows. In the preemptive-resume model, if a machine breakdown occurs during the processing of a job, the work done on the job prior to the breakdown is not lost and the processing of the disrupted job can be resumed at the point where it was interrupted once the machine becomes operable again. In the preemptive-repeat model, however, the work done on this job is lost if the machine breaks down before it is completed, and so its processing will have to restart after the machine resumes its operation.

A significant number of results have been published in the literature on the preemptive-resume model; see Birge et al. [2], Cai and Zhou [3,4], Glazebrook [6,7], Mittenthal and Raghavachari [8], Pinedo [11], Pinedo and Rammouz [10], and Zhou and Cai [15]; to name just a few. In contrast, little progress has been reported on the preemptive-repeat model, although it is equally important in practice. One industrial example of the preemptive-repeat model is in a metal refinery in which the raw material is to be purified by melting it in a very high temperature. If a breakdown (such as power outage) occurs before the metal is purified to the required level, it will quickly cool down and the heating process has to be started again after the breakdown is fixed. Other examples include running a program on a computer, downloading a file from the Internet, performing a reliability test on a facility, and so forth. Generally, if a job must be continuously processed with no interruption until it is totally completed, then the preemptive-repeat formulation should be used to model the processing pattern of the job in the presence of machine breakdowns.

Regarding the preemptive-repeat model, Birge et al. [2] obtained an optimal policy to minimize the expected weighted flow time when the processing time of each job is deterministic. Frostig [5] extended the model of Birge et al. [2] to consider (1) random processing times and (2) different patterns of breakdowns when the machine processes different jobs, under the assumption that after a breakdown, the processing time is resampled and hence independent of the previous time of processing the same job. Adiri, Frostig, and Rinnooy Kan [1] addressed the problem involving a single machine breakdown, to minimize the number of tardy jobs with due dates. It has been widely recognized that when the work done on a job is completely lost and the processing of the job must start over again after a machine breakdown, the problem of deriving an optimal policy to process all jobs is quite complicated and difficult. It is observed that the results in the few works reviewed above have all been derived under the assumption that the uptimes of the machine follow exponential distributions, which substantially simplifies the analysis.

The purpose of this article is to tackle the stochastic scheduling problem with the preemptive-repeat breakdown model. Our study makes several progresses on this model:

1. The problem we consider is built in a general and unified setting, which allows that (i) the uptimes and downtimes of the machine follow general probability distributions, (ii) breakdown patterns of the machine are job dependent, and (iii) processing times of the jobs are random variables following general probability distributions.
2. We consider two different cases for the random processing time after a breakdown: (i) the processing time remains the *same (but unknown) amount* as that *before* the breakdown with respect to the same job and (ii) it is resampled independently after each breakdown. We will refer to case (i) as *without resampling* and case (ii) as *with resampling*.
3. The necessary and sufficient condition to ensure a finite expected time that a job occupies the machine is obtained, a result that is important for the problem with the preemptive-repeat breakdown model to be well posed. The optimal solution that minimizes the expected weighted flow time is derived, which sequences the jobs to be processed by an index policy.
4. Optimal solutions are further induced for a number of specific situations, including some that have interesting practical relevance, under the preemptive-repeat breakdown model.

It is interesting to note that when the processing times are random, there is a distinction between the case with resampling from that without, whereas there is no such distinction with deterministic processing times. In a practical sense, the case without resampling may be used to model the situation in which the randomness of the processing time is *internal* to the job (such as the quality of raw material), which is not influenced by the condition of the machine and so does not vary between machine breakdowns. The case with resampling, on the other hand, considers random factors *external* to the job (such as the condition of the machine), so that the processing time varies independently each time when the same job is repeated, following a specific probability distribution. More discussions on this will be given in Section 5.

The remainder of the article is organized as follows. In Section 2, we formulate the basic model with preemptive-repeat machine breakdowns. Section 3 provides the main results for processing times without resampling, and the case with resampling is considered in Section 4. Comparisons between the cases with and without resampling are further discussed in Section 5. Finally, some concluding remarks are presented in Section 6.

2. MODEL FORMULATION AND ASSUMPTIONS

Suppose that a set of n independent jobs $\{1,2,\ldots,n\}$ are to be processed on a single machine, which are all available at time 0. A *processing time* P_i is required to complete a job i on the machine. Therefore, after job i starts being processed, it will occupy the machine for an amount of time P_i if it is not interrupted before its completion. Nevertheless, all jobs are *preemptive-repeat*, in the sense that the processing of any job will have to be started again if it is interrupted. Let the *occupying time* of job i on the machine be $\mathcal{P}_i$. Then, $\mathcal{P}_i$ may be longer than P_i, if job i is interrupted before it is completed. The machine can process one and only one job at a time. Furthermore, once a job starts being processed by the machine, it cannot be preempted by another job before its completion.

While job i is being processed, the machine may break down, with the breakdown process being characterized by a sequence of finite-valued, positive random vectors $\{Y_{ik}, Z_{ik}\}_{k=1}^{\infty}$, where Y_{ik} and Z_{ik} are the durations of the kth uptime and the kth downtime, respectively, for job i. The distributions of both the uptimes and downtimes Y_{ik} and Z_{ik} are *job dependent*, to reflect the realistic situation when jobs have different levels of impact on the machine during their processing. We assume that the uptimes $Y_{ik}, k = 1,2,\ldots$, are independent and identically distributed (i.i.d.) random variables with an arbitrary distribution function $F_i(t)$ and that the downtimes $Z_{ik}, k = 1,2,\ldots$, are i.i.d. random variables with a distribution function $G_i(t)$. It is assumed that the stochastic processes $\{Y_{ik}, Z_{ik}\}_{k=1}^{\infty}$ for different jobs $i = 1,2,\ldots,n$ are mutually independent.

In the case where P_i remains the same random variable after each breakdown, we assume that $\{P_i\}$, $\{Y_{ik}\}$, and $\{Z_{ik}\}$ are mutually independent with finite means. When P_i is to be resampled after each breakdown (if job i is not completed), let P_{ik} denote the time required to complete job i (without interruption) after the kth breakdown. In such a case, it is assumed that $\{P_{ik}, k = 0,1,2,\ldots\}$ is an i.i.d. sequence of random variables for each i and that $\{P_{ik}\}$, $\{Y_{ik}\}$, and $\{Z_{ik}\}$ are mutually independent with finite means. For the remainder of this section, we consider P_i without resampling. The case with resampling will be considered in Section 4.

Define a counting process $\{N_i(t) : t \geq 0\}$ by

$$N_i(t) = \sup\{k \geq 0 : Y_{i0} < t, Y_{i1} < t, \ldots, Y_{ik} < t\}, \tag{2.1}$$

where $Y_{i0} = 0$, $i = 1,2,\ldots,n$. Then, the time $\mathcal{P}_i$ that job i occupies the machine can be written as

$$\mathcal{P}_i = P_i + \sum_{k=0}^{N_i(P_i)} (Y_{ik} + Z_{ik}), \tag{2.2}$$

where $Z_{i0} = 0$, $i = 1,2,\ldots,n$.

Let $\lambda = (\lambda(1), \lambda(2), \ldots, \lambda(n))$ denote a *sequence* to process the jobs, with $\lambda(k) = i$ representing that job i is the kth to be processed under λ. This is also referred to as a *policy*, or a *solution*, for the problem, which is the decision we seek to determine.

We limit our study to static policies in this article, which is applicable to situations in which altering a decision after it is implemented is very expensive or is even prohibited.

It is easy to see that the completion time of job i under λ can be expressed as

$$C_i(\lambda) = \sum_{j\in\mathcal{B}_i(\lambda)} \mathcal{P}_j = \sum_{j\in\mathcal{B}_i(\lambda)} \left[P_j + \sum_{k=0}^{N_j(P_j)} (Y_{jk} + Z_{jk}) \right], \qquad \textbf{(2.3)}$$

where $\mathcal{B}_i(\lambda)$ denotes the set of jobs sequenced no later than job i under λ.

The problem is to determine an optimal sequence λ^* to minimize the *expected weighted mean flow time*, a criterion that has been widely studied in the scheduling area:

$$\text{EWMF}(\lambda) = E\left[\sum_{i=1}^{n} w_i C_i(\lambda) \right], \qquad \textbf{(2.4)}$$

where $w_i > 0$ is the weight associated with job i. In other words, the objective is to find an optimal sequence λ^* such that

$$\text{EWMF}(\lambda^*) = \min_{\lambda} \text{EWMF}(\lambda). \qquad \textbf{(2.5)}$$

3. MAIN RESULTS FOR THE CASE WITHOUT RESAMPLING

We first give a result on the distribution of the counting process $N_i(t)$ defined by (2.1).

LEMMA 1: *For each $t > 0$, $N_i(t)$ follows a geometric distribution with parameter $1 - F_i(t-)$; that is,*

$$\Pr\{N_i(t) = k\} = F_i^k(t-)[1 - F_i(t-)], \qquad k = 0,1,2,\ldots. \qquad \textbf{(3.1)}$$

As a result,

$$E[N_i(t)] = \frac{F_i(t-)}{1 - F_i(t-)}. \qquad \textbf{(3.2)}$$

PROOF: By the definition of $N_i(t)$ in (2.1) and the assumptions on $\{Y_{ik}\}$, we have

$$\begin{aligned} \Pr\{N_i(t) = k\} &= \Pr\{Y_{i0} < t, Y_{i1} < t, \ldots, Y_{ik} < t, Y_{i,k+1} \geq t\} \\ &= \Pr\{Y_{i1} < t\} \cdots \Pr\{Y_{ik} < t\} \Pr\{Y_{i,k+1} \geq t\} \\ &= F_i^k(t-)[1 - F_i(t-)], \qquad k = 0,1,2,\ldots, \end{aligned}$$

which gives (3.1), and (3.2) then follows immediately. ■

Let $\mu_j = E[Y_{j1}] < \infty$ and $\nu_j = E[Z_{j1}] < \infty$ denote the means of the uptime and downtime, respectively. Theorem 1 gives a result on the expected time $E[\mathcal{P}_j]$ a job occupies the machine, in the situation that the machine is subject to stochastic breakdowns, while the processing of the job follows the preemptive-repeat model after each breakdown. This is an important result for the analysis of problems with the preemptive-repeat breakdown model and is crucial to solving the problem with the mean flow-time criterion.

THEOREM 1: *For the time $\mathcal{P}_j$ that job j occupies the machine,*

$$E[\mathcal{P}_j] = E\left[\frac{1}{1-F_j(P_j-)}\left(\int_0^{P_j}(1-F_j(y))\,dy + \nu_j F_j(P_j-)\right)\right]. \tag{3.3}$$

Moreover, $E[\mathcal{P}_j] < \infty$ if and only if

$$E\left[\frac{1}{1-F_j(P_j-)}\right] < \infty. \tag{3.4}$$

PROOF: Given $t > 0$, as $Y_{j1}, Y_{j2}, \ldots$ are i.i.d.,

$$\begin{aligned} E\left[\sum_{k=0}^{N_j(t)} Y_{jk}\middle| N_j(t) = n\right] &= E\left[\sum_{k=0}^{n} Y_{jk}\middle| Y_{j0} < t, Y_{j1} < t, \ldots, Y_{jn} < t, Y_{j,n+1} \geq t\right] \\ &= \sum_{k=0}^{n} E[Y_{jk}|Y_{j0} < t, Y_{j1} < t, \ldots, Y_{jn} < t, Y_{j,n+1} \geq t] \\ &= \sum_{k=0}^{n} E[Y_{jk}|Y_{jk} < t] = nE[Y_{j1}|Y_{j1} < t] \\ &= \frac{nE[Y_{j1} I_{(Y_{j1}<t)}]}{\Pr(Y_{j1} < t)} = \frac{n}{F_j(t-)}\int_{[0,t)} y\,dF_j(y). \end{aligned} \tag{3.5}$$

Next, noting that the downtimes are independent of the uptimes for every job, we get

$$E\left[\sum_{k=0}^{N_j(t)} Z_{jk}\middle| N_j(t) = n\right] = E\left[\sum_{k=0}^{n} Z_{jk}\right] = nE[Z_{j1}] = n\nu_j. \tag{3.6}$$

Thus, by the law of iterated expectation and (3.5)–(3.6),

$$\begin{aligned} E\left[\sum_{k=0}^{N_j(t)} (Y_{jk} + Z_{jk})\right] &= E\left[E\left[\sum_{k=0}^{N_j(t)} (Y_{jk} + Z_{jk})\middle| N_j(t)\right]\right] \\ &= E\left[N_j(t)\left(\frac{1}{F_j(t-)}\int_{[0,t)} y\,dF_j(y) + \nu_j\right)\right] \\ &= \left(\frac{1}{F_j(t-)}\int_{[0,t)} y\,dF_j(y) + \nu_j\right)\frac{F_j(t-)}{1-F_j(t-)} \\ &= \frac{1}{1-F_j(t-)}\left(\int_{[0,t)} y\,dF_j(y) + \nu_j F_j(t-)\right). \end{aligned}$$

Hence, by (2.3),

$$E[\mathcal{P}_j] = E\left[E\left(P_j + \sum_{k=0}^{N_j(P_j)} (Y_{jk} + Z_{jk}) \middle| P_j\right)\right]$$
$$= E\left[P_j + \frac{1}{1 - F_j(P_j-)}\left(\int_{[0,P_j)} y\, dF_j(y) + \nu_j F_j(P_j-)\right)\right]. \tag{3.7}$$

Furthermore, by Fubini's theorem,

$$\int_{[0,t)} y\, dF_j(y) = \int_{[0,t)} \int_{[0,y)} dx\, dF_j(y)$$
$$= \int_{[0,t)} \int_{(x,t)} dF_j(y)\, dx$$
$$= \int_{[0,t)} [F_j(t-) - F_j(x)]\, dx$$
$$= -t[1 - F_j(t-)] + \int_0^t [1 - F_j(x)]\, dx. \tag{3.8}$$

Substituting (3.8) into (3.7), we get (3.3).

Moreover, as $\int_{[0,P_j)}[1 - F_j(y)]\, dy \le \int_0^\infty [1 - F_j(y)]\, dy = \mu_j$, by (3.3) we see that (3.4) implies

$$E[\mathcal{P}_j] \le (\mu_j + \nu_j) E\left[\frac{1}{1 - F_j(P_j-)}\right] < \infty.$$

Conversely, if $E[\mathcal{P}_j] < \infty$, then by (3.3), we must have $E[F_j(P_j-)/(1 - F_j(P_j-))] < \infty$; hence,

$$E\left[\frac{1}{1 - F_j(P_j-)}\right] = E\left[\frac{F_j(P_j-)}{1 - F_j(P_j-)} + 1\right] < \infty,$$

which completes the proof. ■

Remark 1: Theorem 1 provides (3.4) as the necessary and sufficient condition to ensure a finite expectation for the time that job j occupies the machine in the case without resampling. If Y_{jk} and P_j are exponentially distributed with means $E[Y_{jk}] = 1/\beta_j$ and $E[P_j] = 1/\eta_j$, then condition (3.4) becomes $E[e^{\beta_j P_j}] < \infty$. This holds if and only if $\beta_j < \eta_j$, or $1/\beta_j > 1/\eta_j$; that is, the average length of an uptime for job j must be greater than the average time needed to process that job in order to ensure that the job can be completed within a finite expected time. This is intuitive from a practical point of view when the machine breakdowns are of preemptive-repeat nature and the processing times are not resampled. Furthermore, from Theorem 1, we can easily see that the necessary and sufficient condition for the expected weighted mean flow time (2.4) to be finite is that (3.4) holds for all jobs.

We can now derive the optimal sequence to minimize EWMF.

THEOREM 2: *Suppose that (3.4) holds for $j = 1,2,\ldots,n$, so that the problem with EWMF is well posed under the preemptive-repeat breakdown model. Then, the optimal sequence that minimizes $E[\sum_i w_i C_i(\lambda)]$ is in nondecreasing order of $\{\phi_i/w_i\}$, where*

$$\phi_i = E[\mathcal{P}_i] = E\left[\frac{\int_0^{P_i}(1-F_i(y))\,dy + \nu_i F_i(P_i-)}{1-F_i(P_i-)}\right], \qquad i = 1,\ldots,n. \tag{3.9}$$

PROOF: From Theorem 1 and the first equality in (2.3), it follows immediately that

$$E[C_i(\lambda)] = \sum_{j\in\mathcal{B}_i(\lambda)} E[\mathcal{P}_i] = \sum_{j\in\mathcal{B}_i(\lambda)} \phi_j.$$

Hence, by a straightforward application of the method of adjacent pairwise interchange, we can readily obtain Theorem 2. The details are thus omitted. ■

The following are some special cases and applications of Theorem 2.

Example 1: Exponentially Distributed Uptimes. An important case for the uptime distribution is the exponential distribution, which is often considered in the literature (see, e.g., Frostig [5] and Birge et al. [2]). In this case, let $1/\beta_i$ denote the mean of Y_{ik}, $i = 1,\ldots,n$, $k = 1,2,\ldots$. Then, we have $1 - F_i(t) = e^{-\beta_i t}$, so that

$$\int_{[0,t)}[1-F_j(y)]\,dy = \int_0^t e^{-\beta_j y}\,dy = \frac{1}{\beta_j}(1-e^{-\beta_j t}).$$

Substituting these into (3.3), we obtain

$$\begin{aligned}\phi_j &= E\left[e^{\beta_j P_j}\left(\frac{1-e^{-\beta_j P_j}}{\beta_j} + \nu_j(1-e^{-\beta_j P_j})\right)\right]\\ &= \left(\frac{1}{\beta_j}+\nu_j\right)(E[e^{\beta_j P_j}]-1).\end{aligned} \tag{3.10}$$

Consequently, when the uptimes Y_{ik} are exponentially distributed with mean $1/\beta_i$, and $E[e^{\beta_i P_i}] < \infty$, $i = 1,2,\ldots,n$, the optimal sequence minimizing the expected weighted mean flow time is in nondecreasing order of $\{\phi_i/w_i\}$, with ϕ_i given by (3.10).

If P_j is also exponentially distributed with mean $1/\eta_j$ and $\eta_j > \beta_j$, then

$$E[e^{\beta_j P_j}] - 1 = \frac{\eta_j}{\eta_j-\beta_j} - 1 = \frac{\beta_j}{\eta_j-\beta_j}.$$

Thus, if Y_{ik} and P_i are exponentially distributed with means $1/\beta_i$ and $1/\eta_i$, respectively, and $\eta_i > \beta_i$, $i = 1, 2, \ldots, n$, then the optimal sequence minimizing the expected weighted mean flow time is in nondecreasing order of

$$\left\{ \frac{1 + \beta_i \nu_i}{w_i(\eta_i - \beta_i)}, i = 1, \ldots, n \right\}.$$

Remark 2: Birge et al. [2] obtained a similar result, but only in the case with deterministic processing times and common distributions of uptimes and downtimes across jobs. Theorem 2 extends it to stochastic processing times and job-dependent up/downtime distributions.

Example 2: Uniform Uptimes and Processing Times. Suppose that the uptimes Y_{jk} and the processing times P_j are uniformly distributed over the intervals $[0, u_j]$ and $[0, p_j]$, respectively, with $0 < p_j < u_j$, $j = 1, \ldots, n$. This corresponds to the case where we only know the upper bounds for the uptimes and processing times. In such a case, $F_j(t-) = t/u_j$ for $0 < t < u_j$ and $0 < p_j < u_j$ implies

$$E\left[\frac{1}{1 - F_j(P_j-)}\right] = E\left[\frac{u_j}{u_j - P_j}\right] = \frac{1}{p_j}\int_0^{u_j} \frac{u_j}{u_j - x}\, dx = \frac{u_j}{p_j} \ln\left(\frac{u_j}{u_j - p_j}\right) < \infty.$$

The condition $p_j < u_j$ (i.e., the upper bound of the processing time for job j is less than that of the uptime) is necessary and sufficient for the above expectation to be finite (that ensures the problem is well posed). Assume this basic condition holds. Then, it is easy to calculate

$$\begin{aligned}
E&\left[\frac{1}{1 - F_i(P_i-)} \int_0^{P_i} (1 - F_i(y))\, dy\right] \\
&= E\left[\frac{u_i}{u_i - P_i} \int_0^{P_i} \left(1 - \frac{y}{u_i}\right) dy\right] \\
&= E\left[\frac{u_i}{u_i - P_i}\left(P_i - \frac{P_i^2}{2u_i}\right)\right] \\
&= \frac{1}{2} E\left[P_i - u_i + \frac{u_i^2}{u_i - P_i}\right] \\
&= \frac{1}{2p_i} \int_0^{p_i} \left(x - u_i + \frac{u_i^2}{u_i - x}\right) dx \\
&= \frac{p_i}{4} - \frac{u_i}{2} + \frac{u_i^2}{2p_i} \ln\left(\frac{u_i}{u_i - p_i}\right).
\end{aligned}$$

Furthermore,

$$E\left[\frac{F_i(P_i-)}{1-F_i(P_i-)}\right] = E\left[\frac{u_i}{u_i-P_i}-1\right] = \frac{u_i}{p_i}\ln\left(\frac{u_i}{u_i-p_i}\right)-1.$$

Substituting these expectations into (3.3), we get

$$\phi_i = \frac{p_i}{4} - \frac{u_i}{2} + \frac{u_i^2}{2p_i}\ln\left(\frac{u_i}{u_i-p_i}\right) + \nu_i\left[\frac{u_i}{p_i}\ln\left(\frac{u_i}{u_i-p_i}\right)-1\right]. \qquad \textbf{(3.11)}$$

Consequently, the optimal sequence to minimize the EWMF follows the nondecreasing order of $\{\phi_i/w_i\}$, with ϕ_i given by (3.11).

Example 3: A Problem with Periodical Inspection. This example represents the problem with regular maintenance checkup and repair, which often occurs in practice, and can be described as follows: After starting processing a job, the machine is checked periodically to monitor its condition. The check determines whether the machine needs to be shut down for repair, but the check itself does not interrupt the processing. If a shutdown is necessary, the job will have to start over again after the machine resumes its operation; otherwise, the processing continues without interruption. The probability that a shutdown is necessary, as well as the period between two consecutive checks, are job dependent, due to different impacts/ burdens to the machine created by the job being processed. More specifically, when job j is being processed, the machine undergoes a check every b_j units of time, and there is a probability θ_j $(0 < \theta_j < 1)$ at each check that the machine has to be shut down. Other than these possible shutdowns, the machine works continuously. The problem is to determine the optimal sequence to process the jobs so as to minimize the EWMF.

In this case, a breakdown occurs whenever a check determines to shut down the machine, which is preemptive-repeat, and the repair time represents the downtime. Under the above-described settings, the uptime to process job j is a discrete random variable with masses at mb_j and $\Pr(Y_{jk} = mb_j) = \theta_j(1-\theta_j)^{m-1}$, $m = 1, 2, \ldots$. It follows that $F_j(x) = 0$ for $x < b_j$, and

$$F_j(x) = \sum_{i=1}^{m} \theta_j(1-\theta_j)^{i-1} = \theta_j \frac{1-(1-\theta_j)^m}{1-(1-\theta_j)} = 1-(1-\theta_j)^m \qquad \textbf{(3.12)}$$

for $mb_j \le x < (m+1)b_j$, $m = 1, 2, \ldots$.

Let $m_j = m_j(x)$ satisfy $m_j b_j < x \le (m_j+1)b_j$. Then, by (3.12),

$$1 - F_j(x-) = (1-\theta_j)^{m_j(x)}. \qquad \textbf{(3.13)}$$

Furthermore, given $P_j = x$, let $m = m_j(x)$; then, by (3.12),

$$\begin{aligned}\int_0^x [1 - F_j(y)]\,dy &= \sum_{i=0}^{m-1} (1-\theta_j)^i b_j + (1-\theta_j)^m (x - mb_j) \\ &= \frac{1-(1-\theta_j)^m}{\theta_j} b_j + (1-\theta_j)^m (x - mb_j) \\ &= \frac{b_j}{\theta_j} + (1-\theta_j)^m \left(x - mb_j - \frac{b_j}{\theta_j} \right).\end{aligned}$$

Substituting this and (3.13) into (3.9), we get

$$\begin{aligned}E[\mathcal{P}_j | P_j = x] &= \frac{1}{(1-\theta_j)^m} \left\{ \frac{b_j}{\theta_j} + (1-\theta_j)^m \left(x - mb_j - \frac{b_j}{\theta_j} \right) + \nu_j (1 - (1-\theta_j)^m) \right\} \\ &= \frac{1}{(1-\theta_j)^m} \left(\frac{b_j}{\theta_j} + \nu_j \right) + x - mb_j - \frac{b_j}{\theta_j} - \nu_j \\ &= x - m_j(x) b_j + \frac{b_j + \nu_j \theta_j}{\theta_j} [(1-\theta_j)^{-m_j(x)} - 1]. \end{aligned} \qquad \mathbf{(3.14)}$$

Now, by (3.13) and Theorem 1, $E[\mathcal{P}_j] < \infty$ if and only if

$$E[(1-\theta_j)^{-m_j(P_j)}] < \infty,$$

or, equivalently,

$$E[(1-\theta_j)^{-P_j/b_j}] < \infty. \qquad \mathbf{(3.15)}$$

Assume that condition (3.15) holds for $j = 1, 2, \ldots, n$. Then, by (3.14) and the law of iterated expectation, we get

$$\phi_j = E[\mathcal{P}_j] = E[P_j - m_j(P_j) b_j] + \frac{b_j + \nu_j \theta_j}{\theta_j} E[(1-\theta_j)^{-m_j(P_j)} - 1]. \qquad \mathbf{(3.16)}$$

Consequently, for the above-described maintenance problem, the optimal sequence to minimize the EMWF, by Theorem 2, should follow the nondecreasing order of $\{\phi_j / w_j\}$, where the ϕ_j are given by (3.16). Moreover, because the distribution of $m_j(P_j)$ is given by

$$\Pr(m_j(P_j) = m) = \Pr\{mb_j < P_j \le (m+1) b_j\}, \quad m = 0, 1, 2, \ldots,$$

ϕ_j can also be calculated by

$$\phi_j = E[P_j] + \sum_{m=0}^{\infty} \left[\frac{b_j + \nu_j \theta_j}{\theta_j (1-\theta_j)^m} - mb_j \right] \Pr\{mb_j < P_j \le (m+1) b_j\} - \frac{b_j + \nu_j \theta_j}{\theta_j}.$$

Let us now look at some special cases of Example 3.

Case I: Uniform processing times. Let P_j be uniformly distributed over $(0, Mb_j)$ for some integer $M > 0$. Then,

$$\Pr\{mb_j < P_j \leq (m+1)b_j\} = \frac{b_j}{Mb_j} = \frac{1}{M}, \qquad m = 0,1,2,\ldots,M-1.$$

Hence, by either (3.15) or (3.16),

$$\begin{aligned} E[\mathcal{P}_j] &= \frac{Mb_j}{2} - \sum_{m=1}^{M-1} \frac{mb_j}{M} + \frac{b_j + \nu_j \theta_j}{\theta_j} \left[\sum_{m=0}^{M-1} \frac{(1-\theta_j)^{-m}}{M} - 1 \right] \\ &= \frac{Mb_j}{2} - \frac{M-1}{2} b_j + \frac{b_j + \nu_j \theta_j}{\theta_j} \left[\frac{1}{M} \frac{(1-\theta_j)^{-M} - 1}{(1-\theta_j)^{-1} - 1} - 1 \right] \\ &= \frac{b_j}{2} + \frac{b_j + \nu_j \theta_j}{\theta_j} \left[\frac{1 - (1-\theta_j)^M}{M\theta_j (1-\theta_j)^{M-1}} - 1 \right]. \end{aligned}$$

Case II: Small b_j. If the check is made frequently so that b_j is relatively small, then $m_j(x) b_j \approx x$. Hence, by (3.15), $E[\mathcal{P}_j]$ can be approximated by

$$E[\mathcal{P}_j] \approx \frac{b_j + \nu_j \theta_j}{\theta_j} E[(1-\theta_j)^{-P_j/b_j} - 1]. \tag{3.17}$$

Case III: $b_j \to 0$ but θ_j / b_j remains stable. Note that frequent checks should result in a small chance to shut down the machine at each check. Let $\theta_j = \beta_j b_j$ and $b_j \to 0$, where β_j is a constant. Then, by (3.17),

$$E[\mathcal{P}_j] \approx \left(\frac{b_j}{\beta_j b_j} + \nu_j \right) E[(1-\beta_j b_j)^{-P_j/b_j} - 1] \to \left(\frac{1}{\beta_j} + \nu_j \right) E[e^{\beta P_j} - 1],$$

which is the same as (3.10) with exponential uptimes. Thus, exponential uptimes can be regarded as a limiting case of the maintenance problem in Example 3.

Remark 3: Previous results on preemptive-repeat machine breakdowns are mainly restricted to the case of exponential uptimes. Our results, however, allow a general distribution for the uptimes. This broad coverage allows one to handle a variety of interesting cases, as illustrated by Examples 2 and 3.

4. THE CASE WITH RESAMPLING

We now turn to the case with resampling; that is, each time when a job is repeated, the processing time required is resampled according to its probability distribution. In this case, the counting process defined in (2.1) is no longer applicable and we define, instead, the following random variable:

$$T_i = \sup\{k \geq 0 : Y_{i1} < P_{i0}, Y_{i2} < P_{i1}, \ldots, Y_{i,k+1} < P_{ik}\}. \tag{4.1}$$

Then, the time that job i occupies the machine is given by

$$\mathcal{P}_i = P_{i,T_i} + \sum_{k=0}^{T_i} (Y_{ik} + Z_{ik}), \tag{4.2}$$

and the completion time still has the form $C_i(\lambda) = \sum_{j \in \mathcal{B}_i(\lambda)} \mathcal{P}_j$. Furthermore, the objective function to be minimized remains as the EWMF defined by (2.4).

Let $\mathbf{P}_i = \{P_{ik}\}_{k=0}^{\infty}$ and P_i be a representative of $\{P_{ik}\}$. With similar arguments as in Section 2, we can derive the following results.

LEMMA 2: *Conditional on $\mathbf{P}_i$, the distribution of T_i is given by*

$$\Pr\{T_i = k|\mathbf{P}\} = [1 - F_i(P_{i,k+1}-)] \prod_{j=0}^{k} F_i(P_{ij}-), \qquad k = 0,1,2,\ldots, \tag{4.3}$$

and unconditionally T_i follows a geometric distribution with parameter $1 - E[F_i(P_i-)]$, that is,

$$\Pr\{T_i = k\} = \{1 - E[F_i(P_i-)]\}E^k[F_i(P_i-)]. \tag{4.4}$$

As a result,

$$E[T_i] = \frac{E[F_i(P_i-)]}{1 - E[F_i(P_i-)]}. \tag{4.5}$$

PROOF: By the definition of T_i in (4.1) and the assumptions on $\{Y_{ik}\}$, we have

$$\begin{aligned}\Pr\{T_i = k|\mathbf{P}\} &= \Pr\{Y_{i1} < P_{i1}, \ldots, Y_{ik} < P_{ik}, Y_{i,k+1} \geq P_{i,k+1}|\mathbf{P}\} \\ &= \Pr\{Y_{ik} < P_{ik}|P_{ik}\}\Pr\{Y_{i,k+1} \geq P_{i,k+1}|P_{i,k+1}\} \\ &= [1 - F_i(P_{i,k+1}-)] \prod_{j=0}^{k} F_i(P_{ij}-), \qquad k = 0,1,2,\ldots,\end{aligned}$$

which gives (4.3). Equations (4.4) and (4.5) then follow immediately. ■

THEOREM 3: *The mean of $\mathcal{P}_j$ is given by*

$$E[\mathcal{P}_j] = \frac{1}{1 - E[F_j(P_j-)]}\left(E\left[\int_0^{P_j} (1 - F_j(y))\, dy\right] + \nu_j E[F_j(P_j-)]\right), \tag{4.6}$$

which is finite if and only if $E[F_j(P_j-)] < 1$.

PROOF: The proof is similar to that of Theorem 1, but by conditioning on $\mathbf{P}_j$ and T_j, as in the proof of Lemma 2. Hence, the details are omitted. ■

Theorem 3 leads immediately to the following result on the optimal sequence.

THEOREM 4: *Suppose that $E[F_i(P_i-)] < 1$ for $i = 1, \ldots, n$. Then, the optimal sequence to minimize $E[\sum_i w_i C_i(\lambda)]$ is in nondecreasing order of $\{\phi_i/w_i\}$, where $\phi_i = E[\mathcal{P}_i]$, $i = 1, \ldots, n$, are given by (4.6).*

Remark 4: When Y_{ik} are exponentially distributed with mean $1/\beta_i$, similar to (3.10), we can see that (4.6) reduces to

$$E[\mathcal{P}_i] = \frac{1}{E[e^{-\beta_i P_i}]}\left(\frac{1}{\beta_i} E[1 - e^{-\beta_i P_i}] + \nu_i E[1 - e^{-\beta_i P_i}]\right)$$
$$= \left(\frac{1}{\beta_j} + \nu_j\right)\frac{1 - E[e^{-\beta_i P_i}]}{E[e^{-\beta_i P_i}]}. \tag{4.7}$$

Thus, the optimal sequence minimizing the EWMF is in nondecreasing order of $\{\phi_i/w_i\}$, with $\phi_i = E[\mathcal{P}_i]$ given by (4.7), which coincides with the result of Frostig [5].

From Theorem 4, it is not difficult to obtain $\phi_i = E[\mathcal{P}_i]$ in specific situations. For example, in Example 2 of Section 3, we can show that

$$\phi_i = \frac{2u_i}{2u_i - p_i}\left\{\frac{p_i}{2} - \frac{p_i^2}{6u_i} + \frac{\nu_i p_i}{2u_i}\right\},$$

which is finite as long as $p_i < 2u_i$. In Example 3 of Section 3, let $\rho_i = E[(1 - \theta_i)^{m_i(P_i)}]$. Then,

$$\phi_i = \left(\frac{b_i}{\theta_i} + \nu_i\right)\frac{1 - \rho_i}{\rho_i} + \frac{1}{\rho_i} E[(1 - \theta_j)^{m_i(P_i)}(P_i - m_i(P_i)b_i)],$$

which is finite provided $0 < \theta_i < 1$.

5. COMPARISONS BETWEEN WITH AND WITHOUT RESAMPLING

We have seen in the last two sections that the results differ between the case with resampling and that without. We now attempt to draw some comparisons between these two cases in terms of the expected mean flow time.

First, let us look at an intuitive example. Suppose that the processing time of a job can take any value between 5 and 10 min, say. Assume that a breakdown occurs after the job has been processed continuously for 7 min, but before it is completed. Then in case (i) (without resampling), the job will need at least another 7 min to complete and so the processing time after the machine resumes operation must be between 7 and 10 min. In case (ii) (with resampling), on the other hand, the information from previous experience is lost and the processing time may still take any value between 5 and 10 min. Therefore, in case (i), the work done on a job is lost when a breakdown occurs, but not the information from the previous experience, whereas in case (ii), all is lost once the machine breaks down. Interestingly, this yields the phenomenon that the overall occupying time of a job in case (i) tends to be longer than that in case (ii). This is confirmed, theoretically, in the following proposition.

PROPOSITION 1: *Denote the expected occupying time $E[\mathcal{P}_i]$ of job i by $E_1[\mathcal{P}_i]$ in the case without sampling and by $E_2[\mathcal{P}_i]$ in the case with sampling. If P_i does not degenerate in the support of $F_i(t)$, then $E_1[\mathcal{P}_i] > E_2[\mathcal{P}_i]$.*

PROOF: Define

$$H_i(t) = \int_0^t [1 - F_i(y)]\,dy + \nu_i F_i(t-). \tag{5.1}$$

Then, by Theorems 1 and 3,

$$E_1[\mathcal{P}_i] = E\left[\frac{H_i(P_i)}{1 - F_i(P_i-)}\right], \qquad E_2[\mathcal{P}_i] = \frac{E[H_i(P_i)]}{E[1 - F_i(P_i-)]}. \tag{5.2}$$

It is easy to see that $H_i(t)$ is strictly increasing and $1 - F_i(t-)$ is strictly decreasing, so that $H_i(t)/[1 - F_i(t-)]$ is strictly increasing, in the support of $F_i(t)$. Hence, it follows from Lemma A(b) in the Appendix that

$$E[H_i(P_i)] = E\left[(1 - F_i(P_i-))\,\frac{H_i(P_i)}{1 - F_i(P_i-)}\right] < E[1 - F_i(P_i-)]E\left[\frac{H_i(P_i)}{1 - F_i(P_i-)}\right],$$

which is equivalent to $E_2[\mathcal{P}_i] < E_1[\mathcal{P}_i]$ by (5.2), provided that P_i does not degenerate in the support of $F_i(t)$. ∎

Another interesting difference between the two cases lies in the impact of a breakdown on the remaining occupying time. Let us now compare the expected remaining occupying time of job j conditional on the event that a breakdown occurs before the job is completed (counted from the time that the machine resumes its operation); that is, $E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}]$, with the unconditional expected occupying time $E[\mathcal{P}_j]$. Proposition 2 shows that $E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}]$ is generally greater than $E[\mathcal{P}_j]$ in the case without resampling.

PROPOSITION 2: *In the case without resampling, suppose that condition (3.4) holds; then, $E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}] \geq E[\mathcal{P}_j]$ and the strict inequality holds as long as P_j does not degenerate in the support of $F_j(t)$.*

PROOF: We first modify the proof of Lemma 1 to obtain

$$\Pr\{N_i(t) = k | Y_{i1} < t\} = \frac{\Pr\{Y_{i1} < t, \ldots, Y_{ik} < t, Y_{i,k+1} \geq t\}}{\Pr(Y_{i1} < t)}$$

$$= F_i^{k-1}(t-)[1 - F_i(t-)], \qquad k = 1, 2, \ldots,$$

and so

$$E[N_i(t)|Y_{i1} < t] = \frac{1}{1 - F_i(t-)}. \tag{5.3}$$

Similar to the proof of Theorem 1, we have

$$\begin{aligned} E\left[\sum_{k=2}^{N_j(t)} Y_{jk} \middle| N_j(t) = n, Y_{j1} < t\right] &= \sum_{k=2}^{n} E[Y_{jk}|Y_{jk} < t] = (n-1)E[Y_{j2}|Y_{j2} < t] \\ &= \frac{n-1}{F_j(t-)} \int_{[0,t)} y\, dF_j(y) \end{aligned}$$

and

$$E\left[\sum_{k=2}^{N_j(t)} Z_{jk} \middle| N_j(t) = n, Y_{j1} < t\right] = (n-1)E[Z_{j2}] = (n-1)\nu_j.$$

Thus, by the law of iterated expectation together with (5.3),

$$\begin{aligned} E\left[\sum_{k=2}^{N_j(t)} (Y_{jk} + Z_{jk}) \middle| Y_{j1} < t\right] &= E\left[(N_j(t) - 1)\left(\frac{\int_{[0,t)} y\, dF_j(y)}{F_j(t-)} + \nu_j\right) \middle| Y_{j1} < t\right] \\ &= \left(\frac{1}{F_j(t-)} \int_{[0,t)} y\, dF_j(y) + \nu_j\right)\left(\frac{1}{1 - F_j(t-)} - 1\right) \\ &= \frac{1}{1 - F_j(t-)}\left(\int_{[0,t)} y\, dF_j(y) + \nu_j F_j(t-)\right). \end{aligned} \tag{5.4}$$

It follows from (5.4) and (3.8) that

$$\begin{aligned} &E[(\mathcal{P}_j - Y_{j1} - Z_{j1}) I_{(P_j > Y_{j1})} | P_j = t] \\ &\quad = E[(\mathcal{P}_j - Y_{j1} - Z_{j1}) I_{(t > Y_{j1})} | P_j = t] \\ &\quad = E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j = t > Y_{j1}] \Pr(Y_{j1} < t) \\ &\quad = E\left[t + \sum_{k=2}^{N_j(t)} (Y_{jk} + Z_{jk}) \middle| P_j = t > Y_{j1}\right] F_j(t-) \\ &\quad = F_j(t-)\left\{t + \frac{1}{1 - F_j(t-)}\left(\int_{[0,t)} y\, dF_j(y) + \nu_j F_j(t-)\right)\right\} \\ &\quad = F_j(t-)\left\{\frac{1}{1 - F_j(t-)}\left(\int_0^t [1 - F_j(y)]\, dy + \nu_j F_j(t-)\right)\right\} \\ &\quad = F_j(t-)\tilde{H}_j(t), \end{aligned} \tag{5.5}$$

where $\tilde{H}_j(t) = H_j(t)/(1 - F_j(t-))$, with $H_j(t)$ defined in (5.1). Consequently,

$$E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}] = \frac{E[(\mathcal{P}_j - Y_{j1} - Z_{j1}) I_{(P_j > Y_{j1})}]}{\Pr(P_j > Y_{j1})}$$

$$= \frac{E[F_j(P_j-)\tilde{H}_j(P_j)]}{E[F_j(P_j-)]}. \tag{5.6}$$

On the other hand, by (3.3), we have $E[\mathcal{P}_j] = E[\tilde{H}_j(P_j)]$. Comparing it with (5.6), we can see that $E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}] \geq E[\mathcal{P}_j]$ if and only if

$$E[F_j(P_j-)\tilde{H}_j(P_j)] \geq E[F_j(P_j-)]E[\tilde{H}_j(P_j)]. \tag{5.7}$$

As $F_j(t-)$ and $\tilde{H}_j(t)$ are nondecreasing functions, and strictly increasing in the support of $F_j(t)$, by Lemma A(a) in the Appendix, the inequality in (5.7) is valid for any nonnegative random variable P_j, with the strict inequality holding provided P_j does not degenerate in the support of $F_j(t)$. ■

For the case with resampling, arguments similar to those above (by conditioning on $\mathbf{P}_j$ and T_j) show that $E[\mathcal{P}_j - Y_{j1} - Z_{j1} | P_j > Y_{j1}] = E[\mathcal{P}_j]$ (see also Frostig [5]). Therefore, the expected remaining occupying time of a job increases after a breakdown in the case without resampling, but remains the same in the case with resampling.

6. CONCLUDING REMARKS

Stochastic scheduling subject to preemptive-repeat machine breakdowns is an important and challenging problem. However, unlike the problem with preemptive-resume breakdowns, progress achieved up to date on this problem is very limited. In this article, we have studied the problem in a fairly general and unified setting, which allows the uptimes and downtimes of machine breakdowns to be job dependent, the uptimes and downtimes to follow any general probability distributions, and the processing times to follow any general probability distributions. We have further considered two possible situations on the realization of a random processing time after a machine breakdown, the case without resampling and the case with resampling, and revealed some interesting phenomena on the differences between these cases. We have investigated the optimal solutions under the criterion to minimize the expected weighted flow time. We show that the optimal solutions can be constructed under an index rule comprising the parameters of the model. Results for some cases of important practical relevance have also been developed.

The investigations on the problem with the preemptive-repeat model are, nevertheless, far from being complete, and there are many important and interesting questions to be further studied. An interesting but difficult problem is to consider multiple machines, configured in parallel or as a flowshop or jobshop. Optimal solutions with respect to other performance measures (such as those involving due dates for the jobs) are also interesting topics for further investigation.

Acknowledgments

The authors are grateful to an anonymous referee whose valuable suggestions and comments have helped improve the article. This research was partially supported by the Research Grants Council of Hong Kong under Earmarked Grants CUHK 4418/99E and PolyU 5142/01E.

References

1. Adiri, I., Frostig, E., & Rinnooy Kan, A.H.G. (1991). Scheduling on a single machine with a single breakdown to minimize stochastically the number of tardy jobs. *Naval Research Logistics* 38: 261–271.
2. Birge, J., Frenk, J.B.G., Mittenthal, J., & Rinnooy Kan, A.H.G. (1990). Single-machine scheduling subject to stochastic breakdown. *Naval Research Logistics* 37: 661–677.
3. Cai, X. & Zhou, X. (1999). Stochastic scheduling on parallel machine subject to random breakdowns to minimize expected costs for earliness and tardy cost. *Operations Research* 47: 422–437.
4. Cai, X. & Zhou, X. (2000). Asymmetric earliness-tardiness scheduling with exponential processing times on an unreliable machine. *Annals of Operations Research* 98: 313–331.
5. Frostig, E. (1991). A note on stochastic scheduling on a single machine subject to breakdown—the preemptive repeat model. *Probability in the Engineering and Informational Sciences* 5: 349–354.
6. Glazebrook, K.D. (1984). Scheduling stochastic jobs on a single machine subject to breakdowns. *Naval Research Logistics Quarterly* 31: 251–264.
7. Glazebrook, K.D. (1987). Evaluating the effects of machine breakdowns in scheduling stochastic problems. *Naval Research Logistics* 34: 319–335.
8. Mittenthal, J. & Raghavachari, M. (1993). Stochastic single machine scheduling with quadratic early-tardy penalties. *Operations Research* 41: 786–796.
9. Pinedo, M. (1995). *Scheduling: Theory, algorithms, and systems*. Englewood Cliffs, NJ: Prentice-Hall.
10. Pinedo, M. & Rammouz, E. (1988). A note on stochastic scheduling on a single machine subject to breakdown and repair. *Probability in the Engineering and Informational Sciences* 2: 41–49.
11. Zhou, X. & Cai, X. (1997). General stochastic single-machine scheduling with regular cost functions. *Mathematical and Computer Modelling* 26: 95–108.

APPENDIX

LEMMA A: *Let A be a subset of* $[0,\infty)$ *and X be a nonnegative random variable which is not degenerate in A.*

(a) If $f(x)$ *and* $g(x)$ *are two nondecreasing functions on* $[0,\infty)$, *then*

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)] \tag{A.1}$$

and the strict inequality holds if $f(x)$ *and* $g(x)$ *are strictly increasing in A.*

(b) If $f(x)$ *is strictly increasing and* $g(x)$ *is strictly decreasing in A, then*

$$E[f(X)g(X)] < E[f(X)]E[g(X)].$$

PROOF: Assume that $f(x)$ and $g(x)$ are two nondecreasing functions on $[0,\infty)$. Define $\mu_f = \mathrm{E}[f(X)]$ and $a = \inf\{x : f(x) > \mu_f\}$. Then, $f(x) \geq \mu_f$ for $x > a$ and $f(x) \leq \mu_f$ for $x < a$. Hence,

$$[f(x) - \mu_f]g(x) \geq [f(x) - \mu_f]g(a) \qquad \text{for } x > a$$

$$[\mu_f - f(x)]g(x) \leq [\mu_f - f(x)]g(a) \qquad \text{for } x < a.$$

As a result,

$$[f(x) - \mu_f]g(x) \geq [f(x) - \mu_f]g(a) \qquad \text{for all } x \geq 0. \tag{A.2}$$

It follows that

$$\begin{aligned} E[f(X)g(X)] - E[f(X)]E[g(X)] &= E[(f(X) - \mu_f)g(X)] \\ &\geq E[(f(X) - \mu_f)g(a)] \\ &= E[f(X) - \mu_f]g(a) \\ &= (\mu_f - \mu_f)g(a) \\ &= 0. \end{aligned} \tag{A.3}$$

Thus, (A.1) holds. Furthermore, the equality in (A.3) can only hold if the equality in (A.2) holds for all x in the support of the distribution of X, which is not possible if $f(x)$ and $g(x)$ are strictly increasing in A. This proves part (a); part (b) then follows immediately by applying part (a) to $f(x)$ and $-g(x)$. ■