J Comb Optim (2009) 17: 400–416
DOI 10.1007/s10878-007-9122-0

Analytic evaluation of the expectation and variance of different performance measures of a schedule on a single machine under processing time variability

Subhash C. Sarin · Balaji Nagarajan · Sanjay Jain · Lingrui Liao

Published online: 6 December 2007

Abstract In this paper, we present closed-form expressions, wherever possible, or devise algorithms otherwise, to determine the expectation and variance of a given schedule on a single machine. We consider a variety of completion time and due date-based objectives. The randomness in the scheduling process is due to variable processing times with known means and variances of jobs and, in some cases, a known underlying processing time distribution. The results that we present in this paper can enable evaluation of a schedule in terms of both the expectation and variance of a performance measure considered, and thereby, aid in obtaining a stable schedule. Additionally, the expressions and algorithms that are presented, can be incorporated in existing scheduling algorithms in order to determine expectation-variance efficient schedules.

Keywords Single machine scheduling · Processing time variability · Various performance measures · Expectation-variance analysis

1 Introduction and problem statement

As true as it would be with any other field within the manufacturing domain, the issue of uncertainty is of considerable importance in production scheduling. The randomness in the scheduling system could be due to varying processing times, machine breakdowns, incomplete information about customer due dates, among others. If the uncertainty in the scheduling parameters is not adequately addressed or accounted for, then some of the ill-effects would include system instability, excess inventory, customer dissatisfaction by not meeting the due dates, and more importantly, loss of

S.C. Sarin (✉) · B. Nagarajan · S. Jain · L. Liao
Grado Department of Industrial and Systems Engineering, Virginia Tech., Blacksburg, VA 24061, USA
e-mail: sarins@vt.edu

revenue. Hence, in the world of agile and lean manufacturing, effective scheduling under uncertainty has become a survival necessity for the companies to meet committed shipping dates and effective utilization of the available resources. Consequently, it is imperative to devise the right scheduling strategies to employ in practice.

The ultimate goal of any stochastic analysis in scheduling is to find a schedule that has the 'best' statistical distribution with respect to the performance measure of interest. Knowing such a distribution will enable the management to plan for capacity and quote delivery dates in a manner that achieves set target service levels and higher customer satisfaction. However, finding the distribution of a scheduling criterion can be extremely complex and, most often, not a viable exercise. Hence, the effort is typically resorted to optimizing some function of the performance measure, which almost always is the expectation function. However, by focusing only on the expected value and ignoring the variance of a schedule, the scheduling problem becomes purely deterministic and the significant ramifications of schedule variability are neglected. In many a practical cases, a scheduler would prefer to have a stable schedule with minimum variance than a schedule that has lower expected value but unknown (and possibly higher) variance. To achieve this goal, it is essential to include the variance along with the expectation of a performance measure in schedule selection.

In light of the above, in this paper, we address the issue of determining closed-form expressions or methodologies to help determine the expectation and variance of a given schedule on a single machine. We consider a variety of completion time and due date-based performance measures. The randomness in the scheduling process is assumed to be due to variable job processing time with known mean and variance, and, in some cases, a known processing time distribution. Having knowledge about the variance of a suitable performance measure of a schedule would help the decision maker in selecting an appropriate schedule from those available. Additionally, these expressions and methodologies can be incorporated in existing scheduling algorithms in order to determine effective expectation—variance efficient schedules.

Variation in the processing time of a job is a major factor or cause of uncertainty in scheduling and its impact on the efficiency of a schedule has been a subject of discussion in the literature. McKay et al. (1988) points out that the primary reason of poor applicability of scheduling theory, in practice, is in its inability to properly account for extreme variations in processing times. This is primarily due to the ubiquitous attempt to use deterministic models in the practical situations that are highly stochastic. Dodin (1996) contends that the pseudo-deterministic sequence that is obtained by sequencing the tasks when all activities are assumed to take their expected times deterministically, does not fully reflect the goals of stochastic analysis of a schedule. He further suggests the use of an alternative sequence determined based on a ranking system of *Optimality Indices* (*OI*), defined as sequences' respective probabilities of being the best. Along the same lines, Portougal and Trietsch (1998) suggest that a variance reduction objective should be considered explicitly to attain optimal service levels while retaining the expected completion time as well. Ayhan and Olsen (2000) propose two heuristic procedures to schedule a multi-class single server to minimize the throughput time variance and the outer percentiles of the throughput time.

There has also been some notable work reported in the field of stochastic scheduling under processing time uncertainty that has attempted to consider both the expectation and variance of a performance measure. The notion of "nondominating schedules" has been used to identify a set of schedules when dealing with multi-objective optimization. This set contains the solutions that are reasonably good in terms of one criterion or the other which, in our case, will be either the expectation or the variance. Various efficient heuristics for generating this nondominated set for multi-objective scheduling problems, dealing with the expectation and variance of a single objective, have been developed in the literature. Jung et al. (1990) proposed a heuristic algorithm based on a pairwise-job-interchange method for a stochastic multi-objective flow time scheduling problem. Similar efforts have also been made by Morizawa et al. (1993), Murata et al. (1996) and Rajendran (1995). However, this set of nondominating schedules might contain hundreds of such schedules. Hence, a major task is to determine a methodology of finding a preferred schedule from among a set of nondominated schedules, based on certain user inputs. Shing and Nagasawa (1996) and Nagasawa and Shing (1998) have addressed this problem and have developed an "interactive stochastic multi-objective scheduling system (*ISMSS*)" for selecting a preferred schedule from among a set of nondominating schedules (N) for a single machine problem. Nagasawa and Shing (1997) further extended this analysis to the scheduling of jobs on parallel machines. De et al. (1992) define "expectation-variance efficient sequences" when identifying schedules that are efficient in terms of the mean and variance of a performance measure. They have proposed two different approaches, one based on dynamic programming and the other involving a linear assignment problem subject to a single side constraint (*bi-criteria assignment problem*), to determine the expectation-variance efficient sequences for a single machine flow time problem.

The concept of robust scheduling has also been employed to counter the effects of processing time variance on a performance measure. The notion of schedule robustness as presented by Daniels and Kouvelis (1995) is based on determining a schedule that minimizes the worst-case deviation from optimality for the performance criterion of interest, in the face of all possible scenarios. They formulated an Absolute Deviation Robust Scheduling Problem (*ADRSP*) with total flow time as the optimizing measure. Kouvelis et al. (2000) have extended this analysis to the two-machine flow shop problem. Daniels and Carrillo (1997) formulated a β-Robust scheduling problem in a similar fashion, where each possible scenario has an associated probability and a preferred target level 'T' of the performance measure. The objective is to determine a sequence that maximizes the likelihood of achieving system performance no worse than the target level 'T'. The single machine robust scheduling problem with finite scenarios is *NP*-complete and can be exactly solved using a dynamic programming approach (see Yang and Gang 2002).

Another idea, albeit for expectation function only, that has been explored in the literature is to study the asymptotic behavior of the optimal value in the presence of uncertainty. In particular, this behavior has been studied for multidimensional assignment problems (which can be viewed as a special type of scheduling problems), as the number of dimensions as well as the number of entities in each dimension grow to infinity, where each assignment cost is assumed to be a random variable. For work in this regard, and related work, please see Krokhmal et al. (2007).

2 Analytic evaluation of expectation and variance of performance measures

As alluded to earlier, our analysis, in this paper, is contingent upon the fact that the schedule (or sequence) is known *a priori*, and hence, the position of each job is known with certainty. The goal is then to devise appropriate methodologies and develop analytic expressions (wherever possible) to determine the expectation and variance of the objective function under consideration. The randomness in the scheduling process is attributed only to variable processing times with known means and variances. All other parameters like the job due dates and weights, which represent the importance factors of jobs, are assumed to be deterministic. Also, all jobs are assumed to be available at time zero.

Before we proceed with the analysis, we first present the notation that we use in the sequel. Let n be the total number of jobs to be scheduled on the machine. Also, let $P_{[j]}$ be the processing time (a random variable), $\mu_{[j]} = E[P_{[j]}]$ and $\sigma^2_{[j]} = \text{var}[P_{[j]}]$, $C_{[j]}$ be the completion time ($=\sum_{i=1}^{j} P_{[i]}$, a random variable), $w_{[j]}$ be the weight or importance factor, $d_{[j]}$ be the due date, $L_{[j]}$ be the lateness ($=C_{[j]} - d_{[j]}$), $T_{[j]}$ be the tardiness ($=\max(0, L_{[j]})$, and $U_{[j]}$ be the unit penalty ($=1$, if $C_{[j]} > d_{[j]}$, and $=0$, otherwise), all respectively, associated with the job in the jth position of the given sequence.

2.1 Completion time-based measures

We consider the following completion time-based measures in our analysis: (1) Total Completion Time; (2) Total Weighted Completion Time; and (3) Total Weighted Discounted Completion Time.

2.1.1 Total completion time and weighted completion time

The calculation of the expectation and variance for these measures is rather straightforward and the corresponding expressions are given in Table 1, where $G_{[j]} = \sum_{i=j}^{n} w_{[i]}$.

Example 1 As an illustration of implementing the expressions given in Table 1, consider an instance of four jobs with the mean and variance of their processing times given in Table 2. There are 24 ways of sequencing four jobs on a machine. The sequence 2-4-1-3, obtained by sequencing the jobs in *SEPT* (*Shortest Expected Processing Time*) order, has an expectation and variance of 295 and 365, respectively. Apparently, this sequence is optimal with respect to the expected value of the

Table 1 Expectation and variance of total completion time and total weighted completion time

Measure	Expectation	Variance
$\sum_{j=1}^{n} C_{[j]}$	$\sum_{j=1}^{n} (n+1-j)\mu_{[j]}$	$\sum_{j=1}^{n} (n+1-j)^2\sigma^2_{[j]}$
$\sum_{j=1}^{n} w_{[j]} \sum_{i=1}^{j} P_{[j]}$	$\sum_{j=1}^{n} G_{[j]}\mu_{[j]}$	$\sum_{j=1}^{n} G^2_{[j]}\sigma^2_{[j]}$

Table 2 Data set for total completion time problem

Job Index	1	2	3	4
Mean	40	20	60	25
Variance	15	15	20	5

total completion time. On the other hand, consider two other sequences that are optimal with respect to the variance of the completion time; namely, sequences 4-1-2-3 and 4-2-1-3. These are obtained by arranging the jobs in *SPTV* (*Shortest Processing Time Variance*) order and their respective means and variances are $(320, 295)$ and $(300, 295)$. Note that the sequence 4-2-1-3 deviates by less than 2% from the optimal expected completion time value while it reduces variance by almost 15%. It indicates that by sacrificing a little on the expected value, significant gains can be achieved on the variance front. Such a schedule would be preferable as it reduces the risk of not meeting a desired completion time of the jobs.

2.1.2 Total weighted discounted completion time

Assume that the costs are now discounted at a rate of r, $0 < r < 1$ per unit time, and if job $[j]$ is not completed by time t, an additional cost, $w_j e^{-rt} dt$, is incurred over the interval $[t, t + dt]$. Consequently, for a job $[j]$ that is completed at time t, the total cost incurred over the interval $[0, t]$ is $w_j(1 - e^{-rt})$.

For a given schedule, the sum of weighted discounted completion times of the jobs is given by

$$\sum_{j=1}^{n} w_{[j]}(1 - e^{-rC_{[j]}}) = W - \sum_{j=1}^{n} w_{[j]} e^{-rC_{[j]}}, \quad \text{where } W = \sum_{j=1}^{n} w_{[j]}. \tag{2.1}$$

We have,

$$E\left[W - \sum_{j=1}^{n} w_{[j]} e^{-rC_{[j]}}\right] = W - \sum_{j=1}^{n} w_{[j]} \mu_{C_{[j]}}, \quad \text{where } \mu_{C_{[j]}} = E[e^{-rC_{[j]}}]. \tag{2.2}$$

Also,

$$\begin{aligned} \mathrm{Var}\left[W - \sum_{j=1}^{n} w_{[j]} e^{-rC_{[j]}}\right] &= \mathrm{Var}\left[\sum_{j=1}^{n} [w_{[j]} e^{-rC_{[j]}}]\right], \quad \text{as } \mathrm{Var}[W] = 0, \\ &= \sum_{j=1}^{n} w_{[j]}^2 \sigma_{C_{[j]}}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{[i]} w_{[j]} \sigma_{C_{[ij]}}, \end{aligned} \tag{2.3}$$

where $\sigma_{C_{[j]}}^2 = \mathrm{Var}[e^{-rC_{[j]}}]$, and $\sigma_{C_{[ij]}} = \mathrm{Cov}[e^{-rC_{[i]}}, e^{-rC_{[j]}}]$.

First consider expression (2.2). Note that $E[e^{-rC_{[j]}}]$ is the moment-generating function of the random variable $C_{[j]}$ with $t=-r$, $0<r<1$. Since $C_{[j]}$ is the sum of j independent random variables $P_{[i]}$, $i=1,\ldots,j$, the moment-generating function, $E[e^{-rC_{[j]}}]$, can be determined as a product of the moment generating functions of the individual processing times. That is,

$$M_{C_{[j]}}(-r)=\prod_{i=1}^{j}M_{p_{[i]}}(-r). \tag{2.4}$$

Next, consider expression (2.3). In order to determine $\sigma^2_{C_{[j]}}$, note that,

$$\sigma^2_{C_{[j]}}=\mathrm{Var}[e^{-rC_{[j]}}]=E[(e^{-rC_{[j]}})^2]-(E[e^{-rC_{[j]}}])^2=E[e^{-2rC_{[j]}}]-(E[e^{-rC_{[j]}}])^2.$$

Moreover,

$$\sigma_{C_{[ij]}}=\mathrm{Cov}[e^{-rC_{[i]}},e^{-rC_{[j]}}]=E[e^{-rC_{[i]}}.e^{-rC_{[j]}}]-(E[e^{-rC_{[i]}}].E[e^{-rC_{[j]}}]).$$

But, $E[e^{-rC_{[i]}}.e^{-rC_{[j]}}]=E[e^{-r(C_{[i]}+C_{[j]})}]$, and is the moment-generating function of the random variable $C_{[i]}+C_{[j]}$ which, in turn, is the sum of independent random variables. That is

$$C_{[i]}+C_{[j]}=2P_{[1]}+2P_{[2]}+\cdots+2P_{[i]}+P_{[i+1]}+\cdots+P_{[j]},\quad \text{where } i<j.$$

Therefore, by expression (2.4),

$$M_{C_{[i]}+C_{[j]}}(-r)=M_{p_{[1]}}(-2r).M_{p_{[2]}}(-2r)\cdots M_{p_{[i]}}(-2r).M_{p_{[i+1]}}(-r)\cdots M_{p_{[j]}}(-r).$$

The moment generating functions can be determined after having known the probability density functions of the random job processing times. Standard formulas for moment generating functions exist for all commonly used continuous distributions like exponential and normal, among others.

As an illustration for evaluating the expectation and variance for weighted discounted completion time, consider exponentially distributed job processing times with rate $\lambda_{[j]}$, for $j=1,\ldots,n$. The moment generating function for the processing time of the job in the jth position of the sequence is given by

$$M_{p_{[j]}}(-r)=E[e^{-rp_{[j]}}]=\frac{\lambda_{[j]}}{\lambda_{[j]}+r}.$$

Substituting this expression in (2.4), we have

$$M_{C_{[j]}}(-r)=E[e^{-rC_{[j]}}]=\prod_{i=1}^{j}\left(\frac{\lambda_{[i]}}{\lambda_{[i]}+r}\right). \tag{2.5}$$

In view of expression (2.5), the expectation of the total weighted discounted completion time for exponentially distributed job processing times is given by (see expressions (2.1) and (2.2)),

$$E\left[\sum_{j=1}^{n} w_{[j]}(1-e^{-rC_{[j]}})\right] = W - \sum_{j=1}^{n}\left(w_{[j]}\prod_{i=1}^{j}\left(\frac{\lambda_{[i]}}{\lambda_{[i]}+r}\right)\right).$$

Similarly, variance for the job in position j,

$$\sigma^2_{C_{[j]}} = E[e^{-2rC_{[j]}}] - (E[e^{-rC_{[j]}}])^2 = \prod_{i=1}^{j}\left(\frac{\lambda_{[i]}}{\lambda_{[i]}+2r}\right) - \prod_{i=1}^{j}\left(\frac{\lambda^2_{[i]}}{(\lambda_{[i]}+r)^2}\right). \tag{2.6}$$

And, for the covariance,

$$\sigma_{C_{[ij]}} = E[e^{-rC_{[i]}}.e^{-rC_{[j]}}] - (E[e^{-rC_{[i]}}].E[e^{-rC_{[j]}}]).$$

Since

$$E[e^{-rC_{[i]}}.e^{-rC_{[j]}}] = \prod_{k=1}^{i}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+2r}\right)\prod_{k=i+1}^{j}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+r}\right),$$

we have

$$\sigma_{C_{[ij]}} = \prod_{k=1}^{i}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+2r}\right)\prod_{k=i+1}^{j}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+r}\right) - \prod_{k=1}^{i}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+r}\right)\prod_{k=1}^{j}\left(\frac{\lambda_{[k]}}{\lambda_{[k]}+r}\right). \tag{2.7}$$

Expressions (2.6) and (2.7) can be substituted in (2.3) to obtain the variance of the total weighted discounted completion time for exponentially distributed job processing times.

2.2 Tardiness-based objectives

The tardiness-based measures considered in our analysis are: (1) Total Tardiness; (2) Total Weighted Tardiness; (3) Total Number of Tardy Jobs; (4) Total Weighted Number of Tardy Jobs; (5) Mean Lateness; and (6) Maximum Lateness.

2.2.1 Total tardiness

The expectation of total tardiness is given by,

$$E\left[\sum_{j=1}^{n} T_{[j]}\right] = \sum_{j=1}^{n} E[T_{[j]}] = \sum_{j=1}^{n} \mu_{T_{[j]}}, \quad \text{where } \mu_{T_{[j]}} = E[T_{[j]}]. \tag{2.8}$$

The variance of total tardiness is similarly given by,

$$\begin{aligned}\mathrm{Var}\left[\sum_{j=1}^{n} T_{[j]}\right] &= \sum_{j=1}^{n}\mathrm{Var}[T_{[j]}] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{Cov}[T_{[i]},T_{[j]}] \\ &= \sum_{j=1}^{n}\sigma^2_{T_{[j]}} + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sigma_{T_{[ij]}},\end{aligned} \tag{2.9}$$

where $\sigma^2_{T_{[j]}} = \text{Var}[T_{[j]}]$ and $\sigma_{T_{[ij]}} = \text{Cov}[T_{[i]}, T_{[j]}]$. Note that the tardiness random variables ($T_{[j]}$'s) are not independent (as they are correlated by their individual completion times), and hence, it is necessary to compute the covariance between every pairs of $T_{[j]}$'s. We have

$$\begin{aligned}\text{Cov}[T_{[i]}, T_{[j]}] = \sigma_{T_{[ij]}} &= E[T_{[i]}.T_{[j]}] - E[T_{[i]}]E[T_{[j]}] \\ &= E[T_{[i]}.T_{[j]}] - \mu_{T_{[i]}}\mu_{T_{[i]}}.\end{aligned} \tag{2.10}$$

Recall that $T_{[j]} = \max(C_{[j]} - d_{[j]}, 0)$. As a result, $\mu_{T_{[j]}}$, $\sigma^2_{T_{[j]}}$ and $\sigma^2_{T_{[ij]}}$ in expressions (2.8), (2.9) and (2.10) are difficult to compute. However, in a pioneering work, Clark (1961) developed a method to recursively estimate the expectation and variance of the greatest of a finite set of random variables that are normally distributed. Accurate results can be obtained for maximum functions with two arguments while for higher number of arguments, close-to-accurate approximations can be obtained. This method has been further extended to the case of lognormal distribution by Wilhelm (1986). Hence, assuming normally distributed processing times for all the jobs, Clark's method could be applied to find the desired expectation and variance expressions of $\sum_{j=1}^{n} T_j$.

$C_{[i]}$ and $C_{[j]}$ are linear sums of the job processing times that are assumed to be normally distributed, and hence, are normally distributed by the reproductive property of normal random variables. Once again, recall that $T_{[j]} = \max(L_{[j]}, 0) = \max(C_{[j]} - d_{[j]}, 0)$. The second argument inside the maximum function viz., 0 can be assumed to be a random variable with mean, $\mu = 0$ and variance, $\sigma^2 = 0$. The mean and variance of the first argument, $C_{[j]} - d_{[j]}$, are as follows:

$$\mu_{[j]} = E[C_{[j]} - d_{[j]}] = E[C_{[j]}] - d_{[j]} = \sum_{k=1}^{j} \mu_{[k]} - d_{[j]},$$

$$\sigma^2_{[j]} = \text{Var}[C_{[j]} - d_{[j]}] = \text{Var}[C_{[j]}] = \sum_{k=1}^{j} \sigma^2_{[k]}.$$

The coefficient of linear correlation between the two arguments inside this maximum function is zero. Clark's method can now be applied to accurately determine the expectation and variance of $\max(C_{[j]} - d_{[j]}, 0)$. $\mu_{T_{[j]}}$ is given by the first moment, $\nu_{1_{[j]}}$, of the random variable $\max(C_{[j]} - d_{[j]}, 0)$:

$$\nu_{1_{[j]}} = \mu_{T_{[j]}} = \mu_{[J]}\Phi(\alpha_{[j]}) + a_{[j]}\varphi(\alpha_{[j]}). \tag{2.11}$$

The second moment is given by

$$\begin{aligned}\nu_{2_{[j]}} &= (\mu^2_{[J]} + \sigma^2_{[J]})\Phi(\alpha_{[j]}) + \mu_{[J]}a_{[j]}\varphi(\alpha_{[j]}), \\ \sigma^2_{T_{[j]}} &= \nu_{2_{[j]}} - \nu^2_{1_{[j]}} \\ &= [(\mu^2_{[j]} + \sigma^2_{[j]})\Phi(\alpha_{[j]}) + \mu_{[j]}a_{[j]}\varphi(\alpha_{[j]})] - [\mu_{[j]}\Phi(\alpha_{[j]}) + a_{[j]}\varphi(\alpha_{[j]})]^2\end{aligned} \tag{2.12}$$

where, $a^2_{[j]} = \sigma^2_{[j]}, \alpha_{[j]} = \mu_{[j]}/a_{[j]}$, $\varphi(z) = \frac{1}{\sqrt{2\pi}}\exp(-z^2/2)$, and $\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt$.

We still need to evaluate $E[T_{[i]}T_{[j]}]$, where without loss of generality job j is assumed to succeed job i. To that end, let X and Y to be two random variables such that X is the sum of the processing times of the first 'i' jobs in the sequence and Y is the sum of the processing times of the jobs from position '$i+1$' to 'j' in the given sequence. In other words, X and Y are $C_{[i]}$ and $C_{[j]} - C_{[i]}$, respectively. Similar to the earlier discussion, X and Y are linear sums of the job processing times that are assumed to be normally distributed. By the reproductive property of normal random variables, X and Y are also normally distributed with mean and variance of (μ_x, σ_x^2) and (μ_y, σ_y^2), respectively. By definition, X and Y are independent random variables, and therefore, the joint probability density function for X and Y, $f_{XY}(x, y) = f_X(x).f_Y(y)$, i.e.,

$$\begin{aligned} f_{XY}(x,y) &= \frac{1}{\sigma_X\sqrt{2\pi}}\exp\left(\frac{-(x-\mu_X)^2}{2\sigma_X^2}\right)\frac{1}{\sigma_Y\sqrt{2\pi}}\exp\left(\frac{-(y-\mu_Y)^2}{2\sigma_Y^2}\right) \\ &= \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left(-\frac{1}{2\sigma_X^2\sigma_Y^2}(\sigma_Y^2(x-\mu_X)^2+\sigma_X^2(y-\mu_Y)^2)\right). \end{aligned}$$

Now,

$$\begin{aligned} T_{[i]}T_{[j]} &= \max(C_{[i]} - d_{[i]}, 0).\max(C_{[j]} - d_{[j]}, 0) \\ &= \max(C_{[i]} - d_{[i]}, 0).\max((C_{[j]} - C_{[i]}) + C_{[i]} - d_{[j]}, 0) \\ &= \max(X - d_{[i]}, 0).\max(Y + X - d_{[j]}, 0). \end{aligned} \tag{2.13}$$

Note that $T_{[i]}T_{[j]}$ is a function of two independent random variables X and Y, and therefore,

$$E[T_{[i]}T_{[j]}] = \int_{y=0}^{\infty}\int_{x=0}^{\infty}\max(x-d_{[i]},0).\max(y+x-d_{[j]},0)f_{XY}(x,y)dxdy.$$

Since the integrand will be zero for any value of x and y such that $x < d_{[i]}$ or $y < d_{[j]} - x$, the expression can be simplified to,

$$\begin{aligned} &E[T_{[i]}T_{[j]}] \\ &\quad= \int_{x=d_{[i]}}^{\infty}\int_{y=d_{[j]}-x}^{\infty}(x-d_{[i]})(y+x-d_{[j]})f_{XY}(x,y)dydx \\ &\quad= \int_{x=d_{[i]}}^{\infty}\int_{y=d_{[j]}-x}^{\infty}(x-d_{[i]})[(y-\mu_Y)+(x-d_{[j]}+\mu_Y)]f_X(x).f_Y(y)dydx \\ &\quad= \int_{x=d_{[i]}}^{\infty}(x-d_{[i]})(x-d_{[j]}+\mu_Y)\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right) \\ &\qquad\times\int_{y=d_{[j]}-x}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)dydx \\ &\qquad+\int_{x=d_{[i]}}^{\infty}(x-d_{[i]})\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right) \end{aligned}$$

$$
\begin{aligned}
&\times \int_{y=d_{[j]}-x}^{\infty} .(y-\mu_Y)\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)dydx \\
&= \int_{x=d_{[i]}}^{\infty}(x-d_{[i]})(x-d_{[j]}+\mu_Y)\frac{1}{\sigma_X}\varphi\left(\frac{x-\mu_X}{\sigma_X}\right)\left[1-\Phi\left(\frac{d_{[j]}-x-\mu_Y}{\sigma_Y}\right)\right]dx \\
&\quad + \int_{x=d_{[i]}}^{\infty}(x-d_{[i]})\frac{1}{\sigma_X}\varphi\left(\frac{x-\mu_X}{\sigma_X}\right)\left[-\frac{\sigma_Y}{\sqrt{2\pi}}\exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)\Big|_{d_{[j]}-x}^{\infty}\right]dx \\
&= \int_{x=d_{[i]}}^{\infty}(x-d_{[i]})(x-d_{[j]}+\mu_Y)\frac{1}{\sigma_X}\varphi\left(\frac{x-\mu_X}{\sigma_X}\right)\Phi\left(\frac{x-d_{[j]}+\mu_Y}{\sigma_Y}\right)dx \\
&\quad + \int_{x=d_{[i]}}^{\infty}(x-d_{[i]})\frac{1}{\sigma_X}\varphi\left(\frac{x-\mu_X}{\sigma_X}\right)\sigma_Y\varphi\left(\frac{x-d_{[j]}+\mu_Y}{\sigma_Y}\right)dx. \qquad (2.14)
\end{aligned}
$$

It does not seem likely to simplify expression (2.14) any further. However, it is easy to evaluate it numerically. Note that we can first apply the substitution $s=(x-d_{[i]})/x$, so that the region of integration becomes a finite interval.

Next, we provide a numerical example to illustrate the use of the expressions (2.8) and (2.9) and their subsequent expositions, to calculate the mean and variance of total tardiness of a schedule.

Example 2 Consider the problem presented in Table 2. In addition, we assume that the job due dates are as shown in Table 3.

Consider the sequence 4-2-3-1. We start with the calculation of $\mu_{T_{[j]}}$ and $\sigma^2_{T_{[j]}}$. Take $T_{[3]}$ for example. We have

$$\mu_{[3]} = E[C_{[3]}-d_{[3]}] = \sum_{k=1}^{3}\mu_{[k]} - d_{[3]} = (25+20+60)-90 = 15,$$

$$\sigma^2_{[3]} = \mathrm{Var}[C_{[3]}-d_{[3]}] = \sum_{k=1}^{3}\sigma^2_{[k]} = 5+15+20 = 40.$$

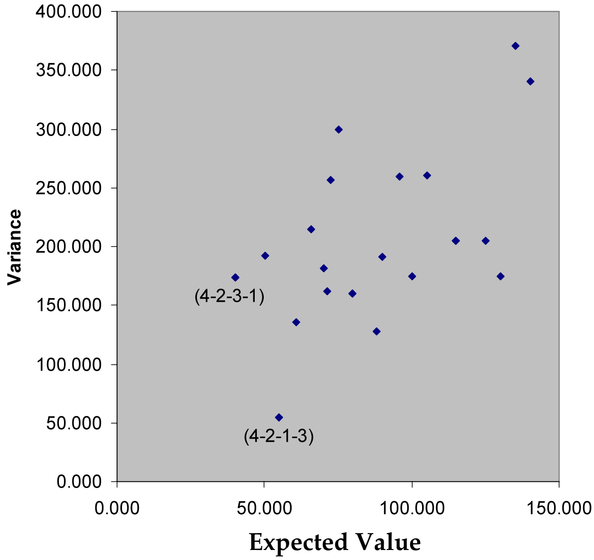
By expressions (2.11) and (2.12), respectively, we can obtain $\mu_{T_{[j]}}$ and $\sigma^2_{T_{[j]}}$. For the calculation of $\sigma_{T_{[ij]}}$, consider $\sigma_{T_{[23]}}$ as an example. We have

$$
\begin{aligned}
&\mu_X = E[C_{[2]}] = 25+20 = 35; \qquad \sigma_X^2 = \mathrm{Var}[C_{[2]}] = 5+15 = 20,\\
&\mu_Y = E[C_{[3]}-C_{[2]}] = 60; \qquad \sigma_Y^2 = \mathrm{Var}[C_{[3]}-C_{[2]}] = 20.
\end{aligned}
$$

Table 3 Data set for total tardiness problem

Job Index	1	2	3	4
Mean	40	20	60	25
Variance	15	15	20	5
Due Date	120	80	90	60

Fig. 1 Expected and variance values of total tardiness for various sequences



By expression (2.14) and subsequently (2.10), we obtain $\sigma_{T_{[23]}}$. Proceeding in this way, and eventually using expressions (2.8) and (2.9), we obtain the mean and variance of the total tardiness of sequence 4, 3, 2, 1 to be,

$$E\left[\sum_{j=1}^{4} T_{[j]}\right] \approx 40.019; \qquad \mathrm{Var}\left[\sum_{j=1}^{4} T_{[j]}\right] = 173.596.$$

We applied the foregoing process to all possible sequences. The results are presented in Fig. 1. Note that the sequence 4-2-3-1 gives the minimum expected value of total tardiness (40.019), while its total variance is 173.596. There is another sequence, 4-2-1-3, for which the expected and variance values are 55 and 55, respectively. Therefore, with a slight increment in the expected tardiness value, the sequence 4-2-1-3 results in a significantly reduced value of its variance over that for the sequence 4-2-3-1, and may be preferred depending on the relative importance given to the two objectives.

2.2.2 *Total weighted tardiness*

The expectation and variance for the total weighted tardiness measure are as shown below.

$$E\left[\sum_{j=1}^{n} w_{[j]}T_{[j]}\right] = \sum_{j=1}^{n} E[w_{[j]}T_{[j]}] = \sum_{j=1}^{n} w_{[j]}\mu_{T_{[j]}},$$

$$\mathrm{Var}\left[\sum_{j=1}^{n} w_{[i]}T_{[j]}\right] = \sum_{i=1}^{n} w_{[j]}^2 \mathrm{Var}[T_{[j]}] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{[i]}w_{[j]}\mathrm{Cov}[T_{[i]}, T_{[j]}]$$

$$= \sum_{j=1}^{n} w_{[j]}^2 \sigma_{T_{[j]}}^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{[i]} w_{[j]} \sigma_{T_{[ij]}}.$$

For the case of normally distributed processing times, these can be determined by using the expressions for $\mu_{T_{[j]}}, \sigma_{T_{[j]}}^2$ and $\sigma_{T_{[ij]}}$ developed in Sect. 2.2.1.

2.2.3 Number of tardy jobs

The expectation and variance of the total number of tardy jobs for a given sequence are given as follows:

$$E\left[\sum_{j=1}^{n} U_{[j]}\right] = \sum_{j=1}^{n} E[U_{[j]}] = \sum_{j=1}^{n} \mu_{U_{[j]}}, \tag{2.15}$$

$$\begin{aligned} \mathrm{Var}\left[\sum_{j=1}^{n} U_{[j]}\right] &= \sum_{j=1}^{n} \mathrm{Var}[U_{[j]}] + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathrm{Cov}[U_{[i]}, U_{[j]}] \\ &= \sum_{j=1}^{n} \sigma_{U_{[j]}}^2 + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sigma_{U_{[ij]}} \end{aligned} \tag{2.16}$$

where $\sigma_{U_{[ij]}} = E[U_{[i]} U_{[j]}] - \mu_{U_{[i]}} \cdot \mu_{U_{[j]}}$.

Note that the random variable, $U_{[j]}$, involves two outcomes, one for a job being early and another for it being late, and hence, can be modeled by using the Bernoulli distribution. Therefore,

$$\begin{aligned} \mu_{U_{[j]}} &= (1.\Pr[C_{[j]} > d_{[j]}] + 0.\Pr[C_{[j]} \le d_{[j]}]) \\ &= 1 - \Pr[C_{[j]} \le d_{[j]}] = 1 - F_{C_{[j]}}[d_{[j]}], \end{aligned} \tag{2.17}$$

$$\sigma_{U_{[j]}}^2 = \mu_{U_{[j]}}(1 - \mu_{U_{[j]}}). \tag{2.18}$$

By substituting expressions (2.17) and (2.18) in expressions (2.15) and (2.16), we have

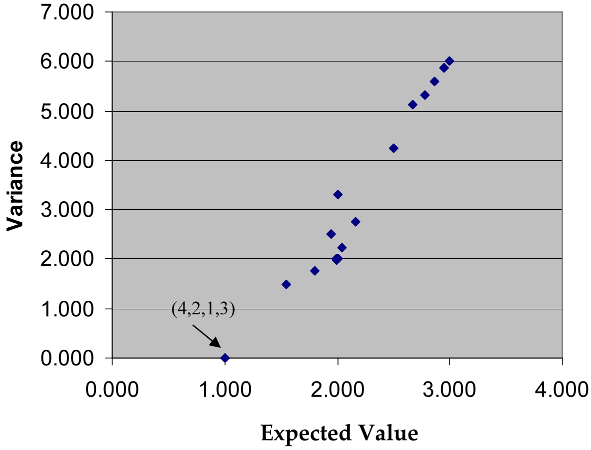
$$E\left[\sum_{j=1}^{n} U_{[j]}\right] = \sum_{j=1}^{n} (1 - F_{C_{[j]}}[d_{[j]}]),$$

$$\begin{aligned} \mathrm{Var}\left[\sum_{j=1}^{n} U_{[j]}\right] = &\sum_{j=1}^{n} [(1 - F_{C_{[j]}}[d_{[j]}]).F_{C_{[j]}}[d_{[j]}]] \\ &+ 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} [E[U_{[i]} U_{[j]}] - (1 - F_{C_{[i]}}[d_{[i]}]) \cdot (1 - F_{C_{[j]}}[d_{[j]}])]. \end{aligned}$$

Similar to the calculation of $E[T_{[i]} T_{[j]}]$, we have

$$E[U_{[i]} U_{[j]}] = \int_{x=d_{[i]}}^{\infty} \int_{y=d_{[j]}-x}^{\infty} f_X(x) f_Y(y) dy dx.$$

Fig. 2 Expected and variance values of the total number of tardy jobs for various sequences



We illustrate the use of the above expressions by assuming normally distributed job processing times. As a result, note that $C_{[i]}$ and $C_{[j]}$ are also normally distributed. Hence,

$$F_{C_{[j]}}(d_{[j]}) = \Phi\left(\frac{d_{[j]} - \mu_{C_{[j]}}}{\sigma_{C_{[j]}}}\right).$$

Furthermore,

$$\begin{aligned} E[U_{[i]}U_{[j]}] &= \int_{x=d_{[i]}}^{\infty}\int_{y=d_{[j]}-x}^{\infty} f_X(x) f_Y(y) dy dx \\ &= \int_{x=d_{[i]}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right) \\ &\quad \times \int_{y=d_{[j]}-x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right) dy dx \\ &= \int_{x=d_{[i]}}^{\infty} \frac{1}{\sigma_X}\varphi\left(\frac{x-\mu_X}{\sigma_X}\right)\Phi\left(\frac{x-d_{[j]}+\mu_Y}{\sigma_Y}\right) dx. \end{aligned}$$

Again, we can use numerical integration to calculate $E[U_{[i]}U_{[j]}]$.

Consider Example 2. We evaluated $E[\sum_{j=1}^{n} U_{[j]}]$ and $\mathrm{Var}[\sum_{j=1}^{n} U_{[j]}]$ for all possible sequences. The results are plotted in Fig. 2. Note that, for this example, 4-2-1-3 is the only non-dominated sequence.

2.2.4 Total weighted number of tardy jobs

The expectation and variance for the total weighted number of tardy jobs for a given sequence are as shown below.

$$E\left[\sum_{j=1}^{n} w_{[j]}U_{[j]}\right] = \sum_{j=1}^{n} w_{[j]}E[U_{[j]}] = \sum_{j=1}^{n} w_{[j]}\mu_{U_{[j]}},$$

$$\begin{aligned}\mathrm{Var}\left[\sum_{j=1}^{n} w_{[i]}U_{[j]}\right] &= \sum_{i=1}^{n} w_{[j]}^2 \mathrm{Var}[U_{[j]}] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{[i]}w_{[j]}\mathrm{Cov}[U_{[i]}, U_{[j]}]\\ &= \sum_{j=1}^{n} w_{[j]}^2 \sigma_{U_{[j]}}^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{[i]}w_{[j]}\sigma_{U_{[ij]}}.\end{aligned}$$

It is easy to see that an analysis similar to the one presented in Sect. 2.2.3 for the total number of tardy jobs, can be used to determine the expectation and variance expressions by using the Bernoulli distribution.

2.2.5 Mean lateness

The mean lateness for a given schedule is given by,

$$\bar{L} = \frac{1}{n}\left[\sum_{j=1}^{n} L_{[j]}\right] = \frac{1}{n}\left[\sum_{j=1}^{n}(C_{[j]} - d_{[j]})\right] = \frac{1}{n}\left[\sum_{j=1}^{n} C_{[j]} - \sum_{j=1}^{n} d_{[j]}\right].$$

Hence, similar to the development in Sect. 2.1.1, we have

$$\begin{aligned}E[\,\bar{L}\,] &= E\left[\frac{1}{n}\left[\sum_{j=1}^{n} C_{[j]} - \sum_{j=1}^{n} d_{[j]}\right]\right] = \frac{1}{n}\left(E\left[\sum_{j=1}^{n} C_{[j]}\right] - \sum_{j=1}^{n} d_{[j]}\right)\\ &= \frac{1}{n}\left[\sum_{j=1}^{n}(n+1-j)\mu_{[j]} - \sum_{j=1}^{n} d_{[j]}\right],\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}[\,\bar{L}\,] &= \mathrm{Var}\left[\frac{1}{n}\left(\sum_{j=1}^{n} C_{[j]} - \sum_{j=1}^{n} d_{[j]}\right)\right] = \frac{1}{n^2}\left(\mathrm{Var}\left[\sum_{j=1}^{n} C_{[j]}\right]\right)\\ &= \frac{1}{n^2}\left[\sum_{j=1}^{n}(n+1-j)^2\sigma_{[j]}^2\right].\end{aligned}$$

2.2.6 Maximum lateness

$$\begin{aligned}L_{\max} &= \max(L_{[1]}, L_{[2]}, \ldots, L_{[n-1]}, L_{[n]})\\ &= \max(\max(\max(\ldots\max(\max(\max(L_{[1]}, L_{[2]}), L_{[3]}), L_{[4]}), \ldots, L_{[n-2]}),\\ &\qquad L_{[n-1]}), L_{[n]}).\end{aligned} \tag{2.19}$$

The mean and variance of the lateness of a job j, $L_{[j]}$, are given by

$$\mu_{L_{[J]}} = E[C_{[j]} - d_{[j]}] = E[C_{[j]}] - d_{[j]} = \sum_{k=1}^{j}\mu_{[k]} - d_{[j]},$$

$$\sigma^2_{L_{[J]}} = \mathrm{Var}[C_{[j]} - d_{[j]}] = \mathrm{Var}[C_{[j]}] = \sum_{k=1}^{j} \sigma^2_{[k]}.$$

Given that the processing times of the jobs are normally distributed, the random variable $L_{[j]}$ is also normally distributed with mean $\mu_{L_{[J]}}$ and variance $\sigma^2_{L_{[J]}}$.

Clearly, by referring to expression (2.19), we can see that the first and second moments of $L_{\max}$ can be determined by recursively applying Clark's equations. To begin with, it requires determination of covariance between $L_{[i]}$ and $L_{[j]}$ of two jobs i and j. We have

$$\mathrm{Cov}(L_{[i]}L_{[j]}) = E[L_{[i]}L_{[j]}] - \mu_{L_{[i]}}\mu_{L_{[j]}}.$$

In order to determine $E(L_{[i]}L_{[j]})$, as before let $L_{[i]} = C_{[i]} - d_{[i]} = X - d_{[i]}$ and $L_{[j]} = C_{[j]} - d_{[j]} = X + Y - d_{[j]}$. We have

$$E[L_{[i]}L_{[j]}] = \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} (x - d_{[i]})(y + x - d_{[j]}) f_X(x) f_Y(y) dy dx.$$

When processing times have symmetric distributions,

$$\begin{aligned}
E[L_{[i]}L_{[j]}] &= \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} (x - d_{[i]})[(y - \mu_Y) + (x - d_{[j]} + \mu_Y)] \\
&\quad \times f_X(x).f_Y(y)dydx \\
&= \int_{x=-\infty}^{\infty} (x - d_{[i]})(x - d_{[j]} + \mu_Y) f_X(x) \int_{y=-\infty}^{\infty} f_Y(y)dydx \\
&\quad + \int_{x=d_{[i]}}^{\infty} (x - d_{[i]}) f_X(x) \int_{y=-\infty}^{\infty} .(y - \mu_Y) f_Y(y)dydx \\
&= \int_{x=-\infty}^{\infty} (x - d_{[i]})(x - d_{[j]} + \mu_Y) f_X(x)dx \\
&= \int_{x=-\infty}^{\infty} [x^2 + (\mu_Y - d_{[i]} - d_{[j]})x + d_{[i]}d_{[j]} - d_{[i]}\mu_Y] f_X(x)dx.
\end{aligned}$$

Now, since $\int_{x=-\infty}^{\infty} x^2 f_X(x)dx = \sigma_X^2 + \mu_X^2$, $\int_{x=-\infty}^{\infty} x f_X(x)dx = \mu_X$ and $\int_{x=-\infty}^{\infty} f_X(x)dx = 1$, we have

$$E[L_{[i]}L_{[j]}] = \sigma_X^2 + \mu_X^2 + (\mu_Y - d_{[i]} - d_{[j]})\mu_X + d_{[i]}d_{[j]} - d_{[i]}\mu_Y.$$

Therefore,

$$\begin{aligned}
\mathrm{Cov}(L_{[i]}L_{[j]}) &= E[L_{[i]}L_{[j]}] - \mu_{L_{[i]}}\mu_{L_{[j]}} \\
&= \sigma_X^2 + \mu_X^2 + (\mu_Y - d_{[i]} - d_{[j]})\mu_X + d_{[i]}d_{[j]} - d_{[i]}\mu_Y \\
&\quad - (\mu_X - d_{[i]})(\mu_X + \mu_Y - d_{[j]}) \\
&= \sigma_X^2 = \sigma^2_{L_{[i]}}.
\end{aligned}$$

Referring to expression (2.19), determination of such covariances will get rather cumbersome with recursion, and one can resort to approximation techniques for their computation (see Wilhelm 1986).

3 Concluding remarks

In this paper, we have considered processing time variability in evaluating the performance of a schedule. Variability in scheduling has been modeled in the literature using random variables, and the majority of work undertaken in this field has considered optimizing the expected value of a performance measure of interest. This paper has addressed the problem of devising methodologies and developing closed-form expressions (wherever possible) to determine the expectation and variance of a given schedule for different performance measures, which can help in selecting schedules that are both expectation and variance efficient. The significance of our work lies in the fact that it enables evaluation of a schedule in terms of both the expectation and variance of a performance measure considered, and thereby, can aid in obtaining a more stable schedule. Our results can be easily incorporated in existing scheduling algorithms to help determine expectation-variance efficient schedules.

Acknowledgements This research has been supported partly by the Center for High Performance Manufacturing (CHPM) at Virginia Tech under CTRF Grant No. SE2002-2003.

References

Ayhan H, Olsen TL (2000) Scheduling of multi-class single server queves under nontraditional performance measures. Oper Res 48(3):482–489

Clark CE (1961) The greatest of a finite set of random variables. Oper Res 9:145–162

Daniels RL, Carrillo JE (1997) β-robust scheduling for single-machine systems with uncertain processing times. IIE Trans 29:977–985

Daniels RL, Kouvelis P (1995) Robust scheduling to hedge against processing time uncertainty in single-stage production. Manag Sci 41(2):363–376

De P, Ghosh JB, Wells CE (1992) Expectation-variance analysis of job sequences under processing time uncertainty. Int J Prod Econ 28:289–297

Dodin B (1996) Determining the optimal sequences and the distributional properties of the completion times in stochastic flow shops. Comput Oper Res 23:829–843

Jung YS, Nagasawa H, Nishiyama N (1990) Bicriterion single-stage scheduling to minimize both the expected value and the variance of the total flow time. J Jpn Ind Man Assoc 39:76–82 (in Japanese)

Kouvelis P, Daniels RL, Vairaktarakis G (2000) Robust scheduling of a two-machine flow shop with uncertain processing times. IIE Trans 32:421–432

Krokhmal P, Grundell D, Pardalos PM (2007) Asymptotic behavior of the optimal value of multidimensional assignment problems. Math Program 109(2–3):525–551

McKay KN, Safayeni FR, Buzacott JA (1988) Job-shop scheduling theory: What is relevant? Interfaces 18:84–90

Morizawa K, Ono T, Nagasawa H, Nishiyama N (1993) An interactive approach for searching a preferred schedule. J Jpn Ind Man Assoc 39:76–82

Murata T, Ishibuchi H, Tanaka H (1996) Multi-objective genetic algorithm and its applications to flowshop scheduling. Comput Ind Eng 30:957–968

Nagasawa H, Shing C (1997) Interactive decision system in parallel-machine stochastic multi-objective scheduling. In: Proceedings of the 1st international conference on engineering design and automation, Bangkok, Thailand, pp 421–424

Nagasawa H, Shing C (1998) Interactive decision system in stochastic multi-objective scheduling to minimize the expected value and variance of total flow time. J Oper Res Soc Jpn 41(2):261–278

Portougal V, Trietsch D (1998) Makespan related criteria for comparing schedules in stochastic environments. J Oper Res Soc 49(11):1188–1195

Rajendran C (1995) Heuristics for scheduling in flowshop with multiple objectives. Eur J Oper Res 82:540–555

Shing C, Nagasawa H (1996) Interactive decision support system in stochastic multi-objective scheduling. Bull Osaka Prefect Univ Ser A 45(2):133–141

Wilhelm WE (1986) The application of lognormal models of transient operations in the flexible manufacturing environment. J Manuf Syst 5(4):253–266

Yang J, Gang Y (2002) On the robust single machine scheduling problem. J Comb Optim 6:17–33