Research Notes for Chapter 11*

Theorem 11.1 is due to Makino (1965). Talwar (1967) proved Talwar's Rule for up to 4 jobs and conjectured it holds for any n, with respect to the stochastic counterpart problem. The conjecture—Theorem 11.2—was subsequently proved by Cunningham and Dutta (1973). Ku and Niu (1986) developed a sufficient condition for stochastic optimality and showed that Talwar's Rule satisfies that condition. Kamburowski (1999) weakened (and thus improved) that sufficient condition, but no new cases that satisfy the improved condition were identified. However, Kalczynski and Kamburowski (2006) generalized Talwar's rule for a Weibull distribution with constant coefficient of variation and showed by simulation that it is close to optimal for several distributions, including the Weibull itself, subject to the constant coefficient of variation assumption. (Whereas that assumption is quite restrictive, in our Research Notes for Chapter 19 we discuss a recent empirical field study where we could not reject the hypothesis that activity times have lognormal distributions with consistent coefficients of variation. The heuristic of Kalczynski and Kamburowski might work very well in such cases.) Safe scheduling results for flow shops were developed by Portougal and Trietsch in a series of papers (1998, 2001, 2006). Table 11.1 is similar to one that appeared in Portougal and Trietsch (2006), where the calculations for the normal API case are based on a result originally published by Clark (1961). However, we discovered a programming error in the spreadsheet they used, so their example was erroneous. Accordingly, we developed a new counterexample to show that the normal distribution is not guaranteed to yield a unique API-stable sequence. We can report, however, that we had to generate a very large number of examples to obtain that counterexample. In other words, in most practical cases we may expect the API heuristic to yield a unique sequence. Dodin (1996) observed that the makespan of a flow shop with independent processing times tends to be approximately normal when the number of jobs is large. The experiments we reported in the chapter confirmed that observation for as few as seven jobs. In experiments conducted by Portougal and Trietsch (2006) with 2000 normally distributed jobs with constant expectation, the makespan distribution could indeed be approximated by the normal in general; however, deviations from normality could still be detected at the tails. That is, the distribution was approximately normal but perceptibly skewed. That may not be surprising, as there are maximum operators at play in the calculations which tend to create skewed distributions.

Theorem 11.3 is also due to Portougal and Trietsch (2006). Because Theorems 11.4 and 11.5 rely on the same proof as that of Theorem 11.3, we repeat it here. For completeness, we start by listing the two regularity conditions, followed by two lemmas.

R1. $(1/u)\sum_{j=1,\ldots,u}(\mathrm{E}(A_j) + \mathrm{E}(B_j)) \geq 2\delta$ where $\delta > 0$; $u = 1, 2,\ldots$

R2. $(1/u)\sum_{j=1,\ldots,u}(\mathrm{V}(A_j) + \mathrm{V}(B_j)) \leq \gamma^2$ where γ is finite; $u = 1, 2,\ldots$

For notational convenience, we assume that jobs are sequenced in index order. Recalling the definition of y_j from Chapter 10, we define a random variable Y_j,

$$Y_j = \sum_{k=1}^{j}[A_k - B_{k-1}]$$

Assuming that jobs are processed on machine 1 without inserted idle time, Y_j represents the difference between the time required to process the first j jobs on machine 1 and the time required to process the first $(j-1)$ jobs on machine 2. Before job j starts on machine 2, there must have been at least this much idle time on machine 2, so to minimize the makespan it is sufficient to minimize $\max\{Y_j\}$. To solve the stochastic counterpart problem, our task is to minimize $\mathrm{E}[\max\{Y_j\}]$. Denote $\mathrm{E}(Y_j)$ by y_j and recall that we denote the expected processing times by a_k and b_k. Because Y_j is a simple sum, its expected value is given by

$$\mathrm{E}(Y_j) = y_j = \sum_{k=1}^{j}[a_k - b_{k-1}]$$

That is, y_j is the value we defined in Chapter 10, but for the deterministic counterpart problem. By Jensen's inequality $\mathrm{E}[\max_j\{Y_j\}] \geq \max_j\{\mathrm{E}(Y_j)\} = \max_j\{y_j\}$. Hence, the optimal makespan of the deterministic counterpart, denoted DCM, provides a lower bound for the optimal expected makespan, $\mathrm{E}(M(s^*))$. But by condition 1, DCM $> \delta n$. Thus we obtain a lower bound, $\mathrm{E}(M(s^*)) \geq \mathrm{DCM} > \delta n$. The following lemmas provide an upper bound.

Lemma RN11.1. For any given set of processing time realizations,

$$\max_j\{Y_j\} \leq \max_j\{y_j\} + \max_k\left\{\sum_{i=1}^{k}[(A_i - a_i) - (B_{i-1} - b_{i-1})], 0\right\};\ 1 \leq k \leq j \leq n$$

Proof.

»» By construction, for $1 \leq k, m \leq j \leq n$

$$Y_j = y_j + \sum_{i=1}^{j}[(A_i - a_i) - (B_{i-1} - b_{i-1})] \quad \leq$$

$$\max_m\{y_m\} + \max_k\left\{\sum_{i=1}^{k}[(A_i - a_i) - (B_{i-1} - b_{i-1})], 0\right\}$$

If $j = m = k$ and if the maximum by k is positive, the inequality is satisfied weakly. Otherwise, the sum of the two maxima must strictly exceed y_j for any j (including the index that maximizes Y_j). ««

Lemma RN11.2. Under conditions 1 and 2, $E[M(s_J) - M(s^*)] \leq 2\gamma\sqrt{n}$, with probability 1 (*w.p.*1).

Proof.

»» The proof relies on a classical Martingale result that we do not develop here. Instead, we refer to Kushner (1984, p. 4). Martingales model the cumulative deviation of a continuous stochastic process from its mean. Martingale theory can also be used to model discrete processes that can be embedded in a continuous process. The result we use is that the expected second moment of the deviation of a Martingale is bounded by four times the sum of expected second moments of the elements, *w.p.*1. Consider the discrete stochastic process given by

$$S_k = \sum_{i=1}^{k}[(A_i - a_i) - (B_{i-1} - b_{i-1})]$$

By construction, as shown in the proof of Lemma RN11.1, $S_k = Y_k - y_k$. For each element in the sum, we have an expected value of 0 and by condition 2, a finite variance given by $V(A_i) + V(B_{i-1})$. Because $E(S_i) = 0$, $V(A_i) + V(B_{i-1})$ is also the second moment of element i. Therefore, it can be shown that S_k is embedded in a Martingale. By the Martingale result we cited above, $E[\sup_k\{S_k^2\}] \leq_{as} 4\Sigma_{\forall i}[V(A_i)+V(B_{i-1})]$ (where $\leq_{as}$ denotes $\leq$ *almost surely*, i.e., *w.p.*1). By condition 2, it follows that $E[\sup_k\{S_k^2\}] \leq_{as} 4n\gamma^2$. Because $E(X^2) = E(|X|^2) \geq (E(|X|))^2$, we thus obtain $E(\sup_k\{|S_k|\}) \leq_{as} 2\gamma\sqrt{n}$. To complete the proof recall that $DCM \leq E(M(s^*)$. ««

Theorem 11.3. Consider a stochastic two-machine flow shop with independent processing time distributions subject to regularity conditions R1 and R2. Let s_J denote the deterministic counterpart sequence (from Johnson's Rule), s^* the optimal sequence, and $M(s)$ the makespan

associated with the sequence s. Then, as $n \to \infty$, $E[M(s_J) - M(s^*)]/E[M(s^*)] \to 0$, *w.p.*1.

Proof.

»» $E(M(s^*)) \geq DCM > \delta n$, so it is sufficient to show that under conditions 1 and 2, for any [small] $\varepsilon > 0$ there exists a number n_ε such that for $n > n_\varepsilon$, $E(M(s_J)) - E(M(s^*))/\delta n \leq_{as} \varepsilon$. By Lemma 2, $n_\varepsilon = 4\gamma^2/(\delta\varepsilon)^2$ satisfies the theorem. ««

Essentially, the proof shows that the Jensen gap is bounded from below by 0 and from above by a function that is proportional to the square root of the total variance in the system, $\sum_{i=1,\ldots,n}[V(A_i) + V(B_i)]$. By the same token, the squared Jensen gap is also bounded by a function that is proportional to the total variance in the system. If we imagine a large stored sample where for each row of realizations we calculate the Jensen gap, essentially we are treating the Jensen gap as a random variable (whereas the true Jensen gap is defined as the expectation of such results). But the makespan of each row is given by DCM plus the Jensen gap realization. The limit on the expected squared Jensen gap realization then implies that the variance of the makespan is similarly bounded. Therefore, the ratio of the standard deviation of the makespan to the expected makespan tends to zero as n grows large; i.e., $cv \to 0$ as $n \to \infty$. This argument is sufficient to also prove Theorems 11.4 and 11.5. They hold because as $cv \to 0$ the optimal safety time becomes negligible relative to the expected makespan.

However, it is implausible that the true coefficient of variation tends to zero as $n \to \infty$. Including linear association in a model is one way to improve its practicality. This observation indicates that the practical value of Theorems 11.3, 11.4 and 11.5 is limited. Would those results hold with linear association? The answer is affirmative because as far as the initial values are concerned, in the limit as $n \to \infty$, $cv \to 0$. Therefore, the distribution of the adjusted makespan converges to that of Q scaled by the mean. Thus, in the limit, the deterministic counterpart means are sufficient to compare two sequences. That proves the following three theorems:

Theorem RN11.1. Consider a stochastic, n-job, two-machine flow shop with linearly associated processing time distributions, with the objective of minimizing the expected makespan. Suppose that the initial (independent) processing times are subject to regularity conditions R1 and R2. Let s_J denote the Johnson deterministic counterpart sequence and s^* the [unknown] optimal sequence. Then, as $n \to \infty$, $[M(s_J) - M(s^*)]/M(s^*) \to 0$, *w.p.*1.

Theorem RN11.2. Consider a stochastic, n-job, two-machine flow shop with linearly associated processing time distributions, with the objective of

minimizing $Z(s) = d + \gamma E(T)$. Suppose that the initial (independent) processing times are subject to regularity conditions R1 and R2. Let s_J denote the Johnson deterministic counterpart sequence, s^* the [unknown] optimal sequence and $Z^*(s)$ the optimal value associated with the sequence s (where the adjusted distribution is used to set d). Then, as $n \to \infty$, $[Z^*(s_J) - Z^*(s^*)]/Z^*(s^*) \to 0$, *w.p.*1.

Theorem RN11.3. Consider a stochastic, n-job, two-machine flow shop with linearly associated processing time distributions, with the objective of minimizing $d(s)$ subject to to a service level constraint $SL \geq b$. Suppose that the initial (independent) processing times are subject to regularity conditions R1 and R2. Let s_J denote the Johnson deterministic counterpart sequence, s^* the [unknown] optimal sequence, and $d^*(s)$ the optimal value associated with the sequence s (where the adjusted distribution is used to set d^*). Then, as $n \to \infty$, $[d^*(s_J) - d^*(s^*)]/d^*(s^*) \to 0$, *w.p.*1.

Next, we present a minor unpublished result for two machines and two jobs with independent processing times.

Theorem RN11.4. For a two-machine, two-job flow shop with stochastically independent operation processing times, the variance of the makespan is bounded from above by the sum of variances of all operations.

Proof.

»» Without loss of generality, assume the jobs are sequenced in index order. The makespan is given by $A_1 + \max\{A_2, B_1\} + B_2$. By independence, the variance is equal to $V(A_1) + V(\max\{A_2, B_1\}) + V(B_2)$. But by Theorem RN9.4, $V(\max\{A_2, B_1\}) \leq V(A_2) + V(B_1)$. ««

Again, by itself, this result is minor. However, extensive numerical simulation suggests that the variance of the makespan cannot exceed the total variance of all jobs for m machines and n jobs, either. We conjecture that the same property applies to the makespan in job shops and projects as well.

Some Empirical Results for the Stochastic Counterpart

In Baker and Trietsch (2009), we experimented with the three main heuristics presented in the chapter: Johnson's heuristic, Talwar's heuristic, and the API heuristic. As might be expected, our results show that Johnson's heuristic is quite robust, API tends to improve

upon the Johnson heuristic, but at the expense of slightly higher variance, and the Talwar heuristic works very well when coefficients of variation are high.

The m-Machine Model

Not too many results exist for the stochastic m-machine model, and much of the existing results rely on very strong assumptions. For example, Pinedo (1982) addresses the stochastic m-machine flow shop with the expected makespan objective, but he assumes that each job has a single distribution that applies to all operations. (In other words, the processing times of job j constitute a sample of m draws from the same distribution). He also assumes that the n distributions are non-overlapping (in other words, we can sort the jobs by expected processing time and the realization will be in the same order on all machines). In the deterministic counterpart of this model, the sequence does not matter. Pinedo shows that any SEPT-LEPT sequence—defined such that the first few jobs are ordered in increasing expected processing time (smallest EPT first) and the remainder in decreasing EPT—minimizes the makespan by expectation. Any pyramid sequence is therefore optimal. Under these conditions, the shop is ordered for any possible realization, but recall from Chapter 10 that in general it *does* matter which pyramid sequence is selected in an ordered shop, whereas here any one will do. Ku and Niu (1986) observe that Pinedo's proof can be extended to show that any SEPT-LEPT sequence is stochastically minimal. Pinedo (1982) also addressed a case where only two jobs are stochastic, and not necessarily non-overlapping. Here, the makespan obtained by arbitrarily selecting one of the stochastic jobs to go first, the other one last, and the deterministic jobs at any order in between, is stochastically minimal.

The m-machine model with blocking can be generalized directly for the stochastic counterpart, but recall from Chapter 10 that even in the deterministic case no efficient solution is known for more than for two machines. For two machines, this case retains the TSP structure of its deterministic counterpart except that the distance matrix contains elements of the stochastic form $D_{ij} = \max\{A_j, B_i\}$ instead of the deterministic form $D_{ij} = \max\{a_j, b_i\}$. Therefore, the polynomial algorithm of Gilmore and Gomory does not apply. When processing times are independent, we can always compute the cdf of the maximum of two or more random variables as the product of their cdfs. For instance, for $\max\{X, Y\}$ we have $F_{max}(t) = F_X(t)F_Y(t)$. Thus, we can compute the expected makespan for any sequence and any m. Therefore, as in the deterministic case, we can still use neighborhood search heuristics in a straightforward manner. In contrast, the no-wait case does not have a proper deterministic counterpart because we cannot guarantee no waiting unless we postpone the start of job $[j]$ on machine 1 until the actual completion time of job $[j-1]$ on machine m. Obviously, this is a very conservative solution, and a safe scheduling approach would be more suitable. However, to our knowledge, this research area has not been explored yet.

References

Baker, K.R. and D. Trietsch (2009) "Three Heuristic Procedures for the Stochastic Flow Shop Problem," *Proceedings*, *Multidisciplinary International Scheduling Conference: Theory and Application* (*MISTA*), Dublin (2009).

Clark, C.E. (1961) "The Greatest of a Finite Set of Random Variables," *Operations Research* 9, 145-162.

Cunningham, A.A. and S.K. Dutta (1973) "Scheduling Jobs with Exponentially Distributed Processing Times on Two Machines of a Flow Shop," *Naval Research Logistics Quarterly* 16, 69-81.

Dodin, B. (1996) "Determining the Optimal Sequences and the Distributional Properties of their Completion Times in Stochastic Flow Shops," *Computers & Operations Research* 23, 829-843.

Johnson, S.M. (1954) "Optimal Two- and Three-Stage Production Schedules with Setup Times Included," *Naval Research Logistics Quarterly* 1, 61-68.

Kalczynski, J.P. and J. Kamburowski (2006) "A Heuristic for Minimizing the Expected Makespan in Two-Machine Flow Shops with Consistent Coefficients of Variation," *European Journal of Operational Research* 169, 742-750.

Kamburowski, J. (1999) "Stochastically minimizing the makespan in two-machine flow shops without blocking," *European Journal of Operational Research* 112, 304-309.

Ku, P.-S. and S.-C. Niu (1986) "On Johnson's Two-Machine Flow Shop with Random Processing Times," *Operations Research* 34, 130-136.

Kushner, H.J. (1984) *Approximation and Weak Convergence Methods for Random Processes*, *with Applications to Random Systems Theory*, The MIT Press.

Makino, T. (1965) "On a Scheduling Problem," *Journal of the Operations Research Society Japan* 8, 32-44.

Pinedo, M. (1982) "Minimizing the Expected Makespan in Stochastic Flow Shops," *Operations Research* 30, 148-162.

Portougal, V. and D. Trietsch (1998) "Makespan-Related Criteria for Comparing Schedules in Stochastic Environments," *Journal of the Operational Research Society* 49, 1188-1195.

Portougal, V. and D. Trietsch (2001) "Stochastic Scheduling with Optimal Customer Service," *Journal of the Operational Research Society* 52, 226-233.

Portougal, V. and D. Trietsch (2006) "Johnson's Problem with Stochastic Processing Times and Optimal Service Level," *European Journal of Operational Research* 169, 751-760.

Talwar, P.P. (1967) "A Note on Sequencing Problems with Uncertain Job Times," *Journal of the Operations Research Society Japan* 9, 93-97.