CHAPTER 3

BASIC THEORY

3.1. Scheduling

3.1.1. Scheduling Definitions and Benefits

Scheduling is the process of organizing, choosing, and timing resource usage to carry out all the activities necessary to produce the desired outputs at the desired times, while satisfying a large number of time and relationship constraints among the activities and the resources (Morton and Pentico in Sipper and Bulfin, 1997). In other words, schedule specifies the time each job starts and completes on each machine as well as any additional resources needed.

Dervitsiotis (1981) defines the importance of scheduling from two different considerations. They are:

1. Inefficient scheduling results in poor utilization or available resources, (the idleness of facilities, human resources, equipment waiting for orders to be processed) as a result production cost increase, and this reduces the competitiveness or effectiveness of the organization.
2. Poor scheduling frequently create delays in the flow of some order system. This calls for expediting measures that again increase cost, upset previous plans, and delays other orders, whose late delivery results in unhappy customers.

While Nahmias (2005) states there are several scheduling objectives:

1. Meets due dates

2. Minimize Work In Process (WIP)inventory
3. Minimize the average flow time average through the system
4. Provide for high machine or worker time utilization (minimize machine or worker idle time)
5. Provide for accurate job status information
6. Reduce setup time
7. Minimize production and worker cost.

3.1.2. Scheduling Classifications

One of characteristic that distinguishes scheduling systems is whether the schedule is generated forward or backward in time (Chase, 2004).

a. Forward Scheduling

Forward scheduling refers to the situation in which the system takes an order then schedules each operation that must be completed forward in time. A system that forward schedule can tell the earliest date that an order can be completed.

b. Backward Scheduling

Backward Scheduling starts from some date in the future (possibly a due date) and schedule required operations is reverse sequencing. The backward schedule tells when an order must be started in order to be done by a specific date.

Based on manufacturing facilities, scheduling is divided into;

a. Flowshop Scheduling

The system of manufacturing facilities is called flowshop when batches (is referred to as jobs) visit the same sequence of workstation (Askin, 1993). Dudek et al in Elsayed (1994) states that a large portion of sequencing research deals with flowshop problems utilizing makespan as the objectives criterion. Scheduling problem in this classification usually deals with single or *n* machines scheduling.

b. Jobshop Scheduling

The system of manufacturing facilities is called jobshop when each part type has its own route (Askin, 1993). Askin (1993) also states that simple dispatching rules are often used in jobshop scheduling. They are:

1. SPT (Shortest Processing Time)
   Select a job with minimum processing time.
2. EDD (Earliest Due Date)
   Select a job due first.
3. FCFS (First Come First Served)
   Select a job that has been on the workstation's queue the longest.
4. FISFS (First in System, First Served)
   Select a job that has been on the shop floor the longest.
5. S/RO (Slack per Remaining Operation)
   Select a job with the smallest ratio of slack to operations remaining to be performed. Slack defined as the time until the due date minus remaining processing time.
6. Covert
   Orders jobs based on ratio slack-based on priority to processing time.

7. LTWK (least Total Work)
   Select a job with the smallest total processing time.
8. LWKR (Least Work Remaining)
   Select a job with the smallest total processing time for unfinished operations.
9. MOPNR (Most Operation Remaining)
   Select a job with the most operations remaining in its processing sequences.
10. MWKR (Most Work Remaining)
    Select a job with the most total processing time remaining.
11. RANDOM
    Select a job at random.
12. WINQ (Work in Next Queue)
    Select a job whose subsequent machine currently has the shortest queue.

Based on the manner in which jobs arrive, jobshop scheduling is also classified into static and dynamic scheduling (Elsayed, 1994).

In the static pattern, there are n jobs each of which must be processed by a set of machine. All the n jobs are available for processing at the initiation of the scheduling period, and no new jobs arrive during the period (Elsayed, 1994). LTWK and EDD (assuming due dates are fixed are static rules (Askin,1993).

While the dynamic rules change with time and queue characteristic. LWKR is dynamic, since the remaining processing time decreases as the job progresses through the shop, which is through time (Askin,1993).

3.1.3. Scheduling Terminology

List of terms below is often used in operation scheduling:

1. Flow time. The flow time of job I is the time that elapses from the initiation of the first job on the first machine to the completion of job I (Nahmias, 2005)
2. Average flow time. Is a common measure of system performance, is the arithmetic average of the flow times for all n jobs.
3. Lateness. Appropriateness between completion time of a job and due date. It is divided into:
   a. Tardiness. It is a condition when job completed after the due date.
   b. Earliness. It is a condition when job completed before the due date
4. Setup time. Time required to prepare the work centre for operation (Arnold, 2001)
5. Run time. Time needed to run the order through the operation (Arnold, 2001).
6. Work center. Work centers are the machine, persons, tools, or vendors used to perform the operation.

3.2. Makespan

Makespan is the time required to complete all *n* jobs (Nahmias, 2005). Minimizing makespan will directly affect the total production cost and labor cost. Therefore, typical objectives (in scheduling) are minimizing average flowtime, minimizing makespan, minimizing average tardiness, minimizing maximum tardiness, and minimizing the number of tardy jobs (Askin, 1993).

3.3. Lot Splitting

Lot splitting is defined as the division of the lot into *nb* sub-batches that can be transferred to the next operation as soon as the former operation has been performed for all parts in the sub-batch (int. j. prod. res., 15 June 2004, vol. 42, no. 12, 2325-2338, Time Bucket Length and Lot-Splitting Approach, J. RIEZEBOS in http://209.85.175.104/search?q=cache:t0PL8dL5WnkJ:www.bdk.rug.nl/medewerkers/j.riezebos/PDF/IJPR_Time%2520Bucket%2520Length%2520and%2520lot-splitting%2520approach.pdf+lot+splitting&hl=id&ct=cln&cd=1&gl=id). This journal also states that lot splitting may reduce the throughput time of the complete lot if more operations are involved.

While in http://goliath.ecnext.com/coms2/gi_0199-661999/Multi-objective-lot-splitting-for.html states that the lot-splitting/lot-streaming literature is mostly associated with minimizing makespan and mean flowtime since these objectives are directly improved by applying this technique.

3.4. Bill of Material (BOM)

Bill of material is a master list of component, purchased parts, and subassemblies that is required to produce a complete product (Elsayed, 1994). Bill of material is also called product structure. The product structure diagram details the parent-child relationship of components and end parts at each level, the number of periods required for production of each components, and the number of component required at the child level to produce one unit of the parent level (Nahmias, 2005). Lulloff (Business Management 361 Section 1, Dr. Foster, Marriott School, Brigham Young University) in his presentation titled Bill of Material defines the four types of BOM (www.freequality.org/sites/www_freequality_org/Documents/Training/bill_of_materials.ppt -). They are:

1. Static (fixed) bill

   It is a bill of material for a part that is normally made from the same components, labor and raw materials. It is used for standard assemblies, components, and engineer-to-order customer orders. Example: A bill of materials for a standard chair

2. Dynamic (parametric) bill

   It is a bill of material for a product or part for which size, color, laminate, and other options can be selected. An example of this type of BOM is for a Dell computer. For the most part the components are standard, but the customer can choose different features and so those parts will vary from project to project.

3. Single level bill of material

   It is a bill of material that lists the materials, parts and labor required to make another part.

4. Multilevel bill of material

It is a bill of material that lists the components, assemblies, and materials required to make a part, the components, assemblies, and materials required to make each component and assembly of the part, and so forth. A multilevel BOM not only lists the parts of a product, but the parts of the components in the product. For example the BOM not only lists the parts in a Dell computer, but also the parts that are in the battery in the Dell computer (http://www.feldmanengineering.com/BoM_Glossary.htm) .

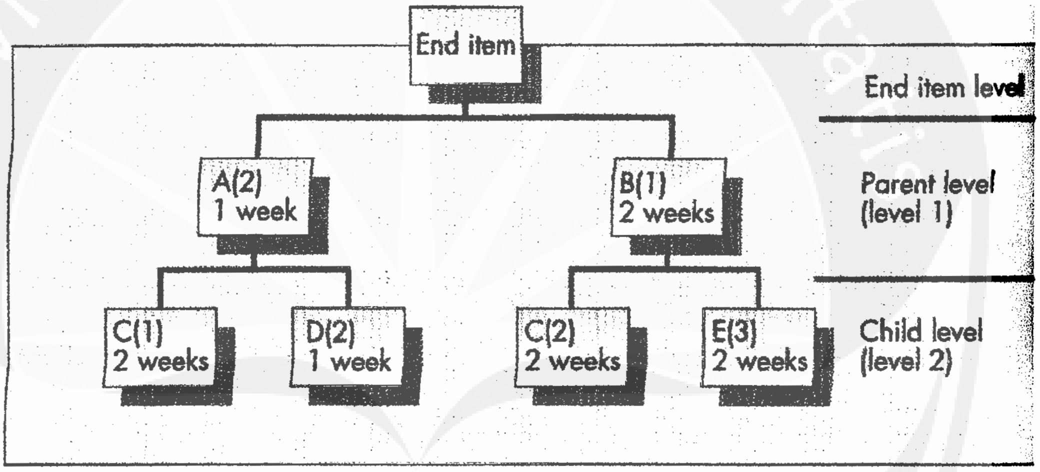

Figure 3.1. The Example of Bill of Material (source: Nahmias, 2001)

3.5. Routing File

There is a record in this routing file for each part manufactured. The routing consists of a series of operations required to make the part. For each product, this file contains a step by step set of instruments describing how the product is made (Arnold, 2001).

In other words, route are the sequence and length of time it takes operation to produce a bill of material part (http://msdn.microsoft.com/en-us/library/aa550330(AX.10).aspx)

3.6. Gantt Chart

Gantt chart is pioneered by Henry Gantt in the early of 1900 which aimed to increase the productivity through better scheduling.

Gantt Chart is a type of bar chart that plots tasks against time (Chase, 2004). Chase also states that Gantt Chart are used for project planning as well as to coordinate a number of scheduled activity

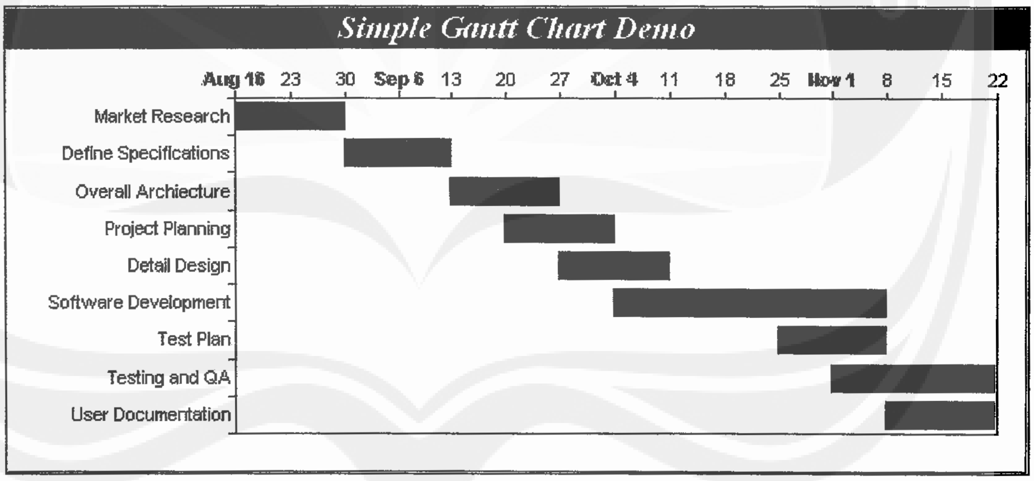

Figure 3.2. Simple Gantt Chart Demo (source: http://www.barchart.be/ganttchart.jsp)

3.7. Hypothesis testing

Hypothesis testing begins with an assumption called a hypothesis. Hypothesis is an unproved theory preposition, supposition, etc tentatively accepted to explain certain facts or to provide a basis for further investigation, arguments, etc (Webster's new world dictionary in Supranto, 1986).

Cohen and Nagel in Supranto (1986) define the function of a hypothesis is to direct our search for order among facts.

Hypothesis, judgment, or opinion is often used to make decision. Therefore, hypothesis must be tested based on empirical data that is from the result of sample research, and so the decision making will be accept or reject hypothesis can also be wrong (Supranto, 1986).

3.7.1. Statistical Hypothesis

Statistical Hypothesis is a statement about a certain population that can be tested through a random sampling (Wonnacot, 1989).

Supranto (1986) defines Statistical Hypothesis as the statement of parameter value. While parameter is an actual value that is calculated based on research in certain population.

3.7.2. Null Hypothesis (H_0) and Alternative Hypothesis (H_1)

Null Hypothesis (H_0) is the assumption we wish to test (Levin, 1981). And alternative hypothesis (H_1) is the conclusion we accept when the data fail to support the null hypothesis (H_0). Whenever we reject the null hypothesis, the conclusion we do accept the alternative hypothesis (Levin, 1981).

Null Hypothesis (H_0) is usually noted by "=". While the alternative hypothesis (H_1) uses ">", "<", or "≠". For example:

$H_0 : \mu_1 = \mu_2 = 100$

(The null hypothesis is that the first population mean is equal to the second population mean is equal to 100)

$H_1 : \mu_1 \neq \mu_2 \neq 100$

(The alternative hypothesis is that the first population mean is not equal to the second population mean is not equal to 100)

3.7.3. Significance Level

Significance Level is a value indicating the percentage of sample values that is outside certain limits, assuming the null hypothesis is correct, i.e., the probability of rejecting the null hypothesis when it is true (Levin, 1981).

Sari (2008) defines that there are some definitions in hypothesis testing, which are:

1. Significance level (α)

   Significance level is an area under the distribution normal curve which refers to the rejected null hypothesis area
2. Confidence level ($1-\alpha$)

   Confidence level is an area under the distribution normal curve which refers to the accepted null hypothesis area

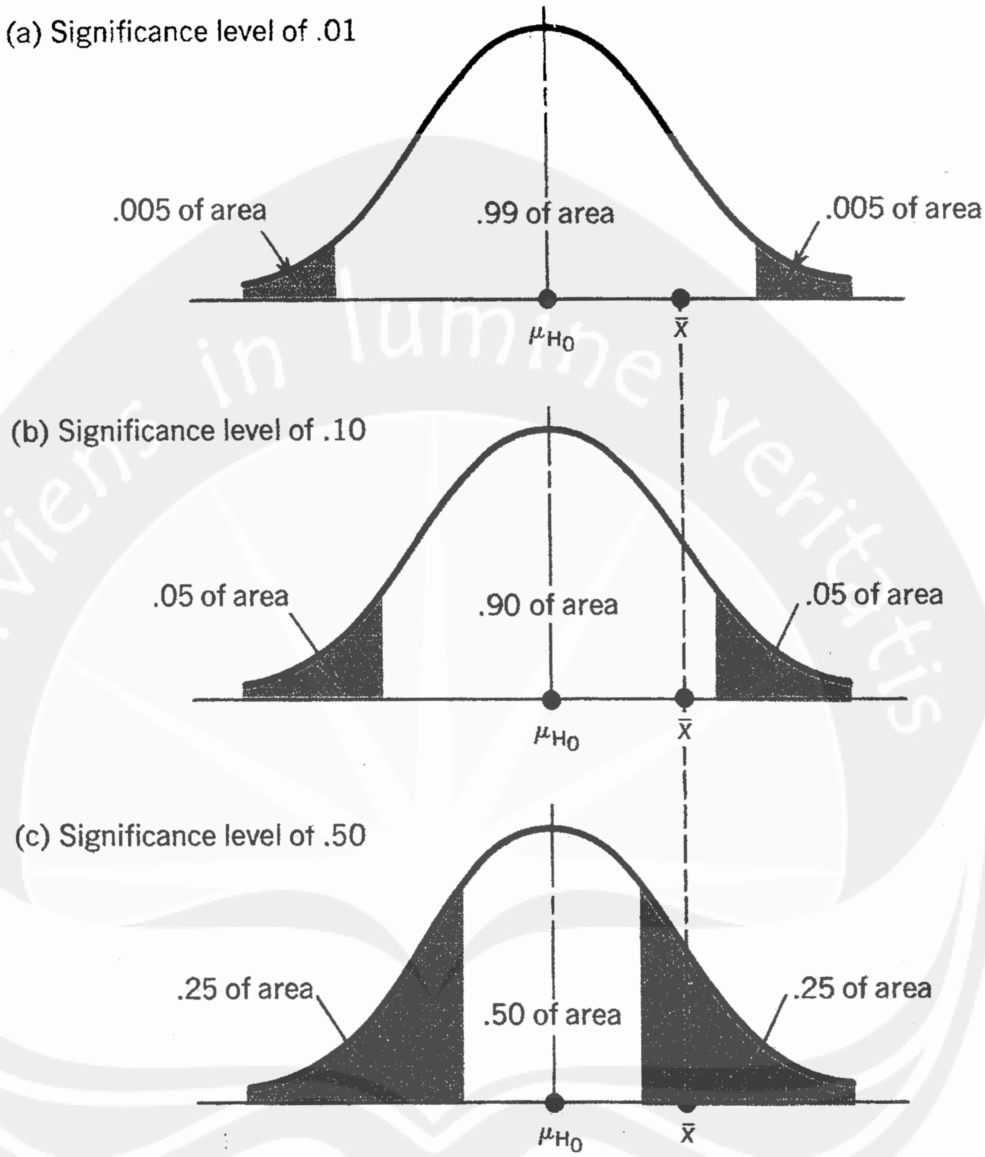

Figure 3.3. Hypothesis Testing at Three Different Significance Levels (source: Levin, 1981)

The Figure 3.3 above ilustrates a hypothesis test at three different significance levels. They are 0.01 which means 0.99 of confidence level, 0.10 which means 0.90 of confidence level and 0.50 which means 0.50 of confidence level. The shaded area is where the null hypothesis is rejected.

The higher significance level we use for testing a hypothesis, the higher probability of rejecting the null hypothesis when it is true (Levin, 1981).

3.8. Analysis of Variance (ANOVA)

Analysis of variance is a statistical technique used to test the equality of 3 or more samples means and thus make interferences as to whether the samples come from populations having the same mean (Levin, 1981). There are two types of ANOVA. They are one way ANOVA and two ways ANOVA.

This research uses one way anova. A One-Way Analysis of Variance is a way to test the equality of three or more means at one time by using variances (http://people.richland.edu/james/lecture/m170/ch13-1wy.html).

Generally, there are several steps to build ANOVA. They are:

1. Make Hypothesis statement
2. Take F test to find F statistic and critical value in F table
3. Compare F statistic and critical value in F table

3.8.1. Hypothesis statements

Hypothesis statements are about null hypothesis (H_0) and alternative hypothesis or also called research hypothesis (H_1). The common hypothesis statements used in ANOVA is following:

Null hypothesis:

$H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_n$

Alternative hypothesis:

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots \neq \mu_n$

The null hypothesis claims that the populations are identical in distribution to each other. And the alternative hypothesis claims that they differ each others.

3.8.2. F Test

ANOVA uses F test that is designed to test if two population variances are equal. It does this by comparing the ratio of two variances (http://people.richland.edu/james/lecture/m170/ch13-f.html).

3.8.2.1. F Statistic

F Statistic for ANOVA is the ratio of variability for the two sources of variation: the between sample variability divided by the within sample variability (Siegel, 2000).

Basically, there are 4 steps to find the value of F statistic. They are:

1. Grand Mean / Grand Average

   Grand Mean is the mean for entire group of subjects from all the samples in the experiment (Levin, 1981). In other words, grand mean of a set of samples is the total of all the data values divided by the total sample size.

$$\bar{\bar{X}} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \ldots + n_k \bar{X}_k}{n} = \frac{\sum_{i=1}^{k} n_i \bar{X}_i}{n} \qquad \ldots\ldots(3.1)$$

Where:

$\bar{\bar{X}}$ = Grand mean or grand average

n_i = i^{th} Sample size

$\bar{X}_i$ = Mean or average of i^{th} sample

2. Between samples variability

Between samples variability (or between column variance) is an estimate of the population variance derived from the variance among the sample means (Levin, 1981). It is donated as MS(B) for Mean Square Between groups. In other words, This is the between group variation (SS(B) for Sum of Squares Between groups) divided by its degrees of freedom (http://people.richland.edu/james/lecture/m170/ch13-def.html).

$$MS(B) = \frac{\sum_{i=1}^{k} n_i (\bar{X} - \bar{\bar{X}})^2}{k-1} \qquad \ldots\ldots\ldots\ldots\ldots (3.2)$$

Where:

$\bar{\bar{X}}$ = Grand mean or grand average

n_i = i^{th} Sample size

$\bar{X}_i$ = Mean or average of i^{th} sample

k = number of sample group

$k - 1$ = degrees of freedom

The degrees of freedom number expresses variability measurement of k averages.

3. Within sample variability

Within sample variability or within column variance is an estimate of the population variance based on variances within samples (Levin, 1981). It is denoted as MS(W) for Mean Square Within groups. Or this is the within group variation (This is the within group variation divided by its degrees of freedom) divided by its degrees of freedom (http://people.richland.edu/james/lecture/m170/ch13-def.html).

$$MS(W) = \frac{\sum_{i=1}^{k}(n_i - 1)(S_i)^2}{n-k} \quad (3.3)$$

Where:

n_i = i^{th} Sample size

S_i = Standard deviation of i^{th} sample

k = number of sample group

n = number of total sample

$n - k$ = degrees of freedom

4. F Statistic

$$F\ Stat = \frac{MS(B)}{MS(W)} \quad (3.4)$$

The degrees of freedom numbers expresses variability measurement of all n data values about their sample averages but have lost k degrees of freedom because k different sample averages were estimated.

All of these formulas above can be tabled as table 3.1. below:

Table 3.1. ANOVA Table

	Degrees of Freedom	Sum Squares	Mean Squares	F statistic
Between Groups	$k-1$	SSB $=\sum_{i=1}^{k} n_i (X-X)^2$	MSB $=\frac{\sum_{i=1}^{k} n_i (X-X)^2}{k-1}$	F Stat $=\frac{MS(B)}{MS(W)}$
Within Groups	$n-k$	SSW$=\sum_{i=1}^{k}(n_i-1)(S_i)^2$	MSW $=\frac{\sum_{i=1}^{k}(n_i-1)(S_i)^2}{n-k}$	
Total	$n-1$	SST		

3.8.2.2. F Table

F table is a list of critical value for the distribution of the F statistic (Siegel, 2000). To find critical value, the numbers of degrees of freedom are used to find the row and column in F table, depending on the significance level tested.

The denominator (column in F table) is taken from degrees of freedom *n-k* for within sample variability. While the numerator (row in F table) is taken from degrees of freedom *k-1* for between sample variability. Sometimes interpolation is needed to give an approximately value between two values.

3.8.3. F statistic versus critical value in F table

The last step is comparing the value between F statistic versus critical value in F table and taking conclusion based on it.

Siegel (2000) states that if F statistic is smaller than the critical value from F table:

- Accept the null hypothesis, H_0 as a reasonable possibility
- Do not accept the research hypothesis, H_1
- The sample averages are not significantly different from each other
- The observed difference among the sample averages could reasonably be due to random chance alone
- The result is not statistically different

(All of the above statements are equivalent to one another)

When F statistic is greater than the critical value from F table:

- Accept the research hypothesis, H_1
- Reject null hypothesis, H_0
- The sample averages are significantly different from each other
- The observed difference among the sample averages could not reasonably be due to random chance alone
- The result is statistically different

(All of the above statements are equivalent to one another)

P-value also plays a role in concluding hypothesis. P-value stands up for Probability Value. Wonnacott (1989) states that P-value is H_0's credibility measurement. If p-value is less than significance level (α) reject the null hypothesis (H_0).

Another benefit of the p-value is that the statistician immediately knows at what level the testing becomes significant. That is, a p-value of 0.06 would be rejected at an 0.10 level of significance, but it would fail to reject at an 0.05 level of significance (http://people.richland.edu/james/lecture/m170/ch09-pvl.html). In practice, many computer software or statistical package can give the exact p-value.