Summaries modules 1 and 2

Industrial Engineering and Management

Study Association Stress
RAVELIJN BUILDING | RA1336

Table of Contents

Industrial Engineering and Management 0
Stress 2
Education 2
HELP! 3
Become active at Stress! 4
Member Initiative 4
More information about Stress 4
Module 1: Introduction to IEM 5
Summary 1: Mathematics 6
Summary 2: Probability 4
Summary 3: Research Methods 14
Summary 4: Excel & VBA Manual Summary 22
Module 2: Operations management 30
Summary 1: Operations Strategy 1 31
Summary 2: Operations Strategy 2 60
Summary 3: Operations Research 1 87
Summary 4: Operations Research 2 93

Stress

Stress is the study association for Business Administration (BA), International Business Administration (IBA) and Industrial Engineering and Management (IEM) of the University of Twente. Stress was founded on May 21st, 1974. Currently, Stress has over 2100 members and is the largest UT (study) association. Stress organizes various activities to support, expand and complement all of its studies. Stress has five principles, which will greatly enhance your time as a student: Study, Meet, Practice, Develop and International. Following these, several activities are organized by the roughly 120 active members, and are partially made possible thanks to the sponsorship and participation of various companies. Moreover, we have regular contact with other business and management study associations across the Netherlands.

Education

As a study association, Stress represents its members towards the faculty. Therefore we have a Commissioner of Educational Affairs who, with help of the Education Committee, deals with everything concerning your education. From collecting summaries to handling complaints to organizing educational events. If you have any questions concerning your study, the teachers, the faculty or anything else, feel free to ask the Commissioner of Educational affairs!

Education Committee

The committee consists of representatives of every cohort. The representative of the freshmen will be introduced during the first module. This committee is aimed at forming the bridge between teachers and students and therefore, improving the quality of the study. They keep the summary database on the Stress site up to date, organise study evenings and try to keep an eye out for the quality of our education. When there are complaints from the students concerning education, about one of the courses of one of our studies for example, they will be handled by this committee.

Summaries

Next to handling complaints, we also collect and check summaries of all the courses you follow! So it does not matter if the course is in the first, fifth or eighth module, you can send your summary in and we will check if your summary will make a good addition to our collection. To hand in your summary, simply send the file to ec@stress.utwente.nl. The Education Committee will then check the summary and if it is found to be sufficient, you will be compensated for your efforts. If your summary is the first one of a course, you will receive €15. If it is the second one, €10, and for the third one you will get €5. If you think you have made a better summary than the ones online, you can also send yours in and earn €5,-. Our summary collection can be found at the bottom of the 'study' page on www.stress.utwente.nl.

Panel Meetings

The panel meetings are organised in every module to improve the education. Here, the teachers of the module together with some students discuss the module. The students are able to give their opinion about the module and what they would like to see improved. The teachers can also ask questions about the opinions of the students. This way teachers know what went right in a module and what went wrong so they can improve the module for next year. The panel meetings are for all the students which are taking the module. You can also join the feedback panel. This means that you join a group of student who attends the panel meeting each module and gives valuable feedback to the programme.

Study sessions

For some courses, Stress organises study sessions. During these afternoons or evenings one or two student assistants of the course will be present. The study sessions are free to attend and coffee, tea and snacks are provided for you by Stress. If you think a study session will be valuable for a course you are following, please contact the Commissioner of Educational Affairs or the Education Committee. They will check if there is more demand for a study evening for this course and act accordingly.

Complaints

If you have a complaint, you can submit it at the 'Study' page on www.stress.utwente.nl or talk to someone of the Education Committee. However, if you feel it is a really important complaint or you want to explain it personally, you can come to the Stress room and talk to the Commissioner of Educational Affairs or send an email. We will then contact the programme management team and discuss what actions can be taken. They value bundled complaints greatly because it tells them a lot more when multiple people have the same complaint, this is the most important reason to always voice your opinion.

Ordering Books

For the first module, you can order your books during the Kick-In. For the following modules, you will have to order them by yourself. You can do this online, at our website. The only requirement to order the books is that you are a member of Stress.

To order books online you have to go to the 'Study' page on www.stress.utwente.nl. On the left of the screen you find a header: 'BOOKSALE', and below the option: 'Order your books'; select this option. Next, you can use the dropdown menus to select your study and module. Once you have chosen the correct options, press 'To booklist'. After this, you can select all the books you would like to order, and then proceed to 'Checkout. After paying, the books will be shipped to the address you enter.

For any questions about the books you need, the ordering of the books or anything else book-related, you can send an email to books@stress.utwente.nl.

Tutor platform

If you are having trouble studying for a course, we have the tutor platform to provide you with the right student for your struggles. We have a wide variety of students who have gone before you and who are willing to help you out for a small compensation. Send an email to tutor@stress.utwente.nl and mention your study, study-year, course you need help with, how many hours you need and any requirements you might have for the tutor. The payment of the tutor can be negotiated but keep in mind that students get paid €10 to €15 by the university when working as a teaching assistant, and you have to pay for it yourself.

The other way around, we are always looking for new tutors. If you are interested in joining our tutor pool let us know. We will add you to the WhatsApp group and you can reply to students asking for tutoring.

HELP!

Often, students do not know where to go with any problems, either study-related or personal. Here you find some information about the most common places to find help.

Study advisor

The study advisor is not only there to answer all your questions about your study, but also there to help you with any personal conditions or other issues that might affect you or your study progress. If you have any problem at all, go see your study advisor. Even if they are not the person who can help you, they can send you to someone who can. Every talk with study advisors is confidential and they will always do their best to help you. You can make an appointment with the study advisors on www.bms.planner.utwente.nl. The study advisors for IEM are Cornelis ten Napel and Ellen van Zeijts. The office of Cornelis is RA3246 and the office of Ellen is RA3256. Their emails are c.tennapel@utwente.nl and e.w.g.vnzeijts@utwente.nl.

Red desk / Student Affairs Coaching & Counselling

If your study and personal life are all on track, this bit of information might not be really relevant for you. But if it is not the case, when your study is completely going the wrong way, or you find it hard to adapt to living away from your parents or you have a difficult situation back at home, the Student Affairs Coaching & Counselling, also called the 'Red Desk', is the place where they can help you. Every possible question about study or personal issues will be answered here, or you will be forwarded to a trained professional. The Red Desk can be contacted at sacc@utwente.nl and is located in the Vrijhof (building 47), third floor, room 311.

Become active at Stress!

Next to your study, you can become an active member of our association! Stress offers many different committees which have organisational tasks or supporting tasks. On our website, you can check out all the committees from Stress. To find out which committee suits you best, email the Commissioner of Internal Affairs at internal@stress.utwente.nl.

Member Initiative

Have you always wanted to organize something big, but never had the resources? We appreciate initiatives from our members! So, if you have a clever idea for something within Stress or the committees, please contact us and we can see what is possible.

More information about Stress

Do you want to know more about Stress? Or do you want to check out our website and social media? Make sure to scan the QR code:

Module 1: Introduction to IEM

Contact: intro-iem@utwente.nl

! Disclaimer: always check if what you need to study corresponds with the content of the summaries, courses can be changed which could cause changes in study material for your exams

This module consists out of several courses, a wide introduction to multiple subjects and a project. Below you find information about which courses you have in this module, and about the summaries for this module. If you made a summary for a course this module you can send them to education@stress.utwente.nl and depending on how many summaries we have for this course you will receive compensation for your work.

Courses

- Introduction to Mathematics & Calculus 1A
- Probability
- VBA Programming
- Professional Skills
- Research Methods
- Core topics: information, production, supply chain and financial management
- Project

Summary 1

Course: Introduction to Mathematics
Book: Dictate
Chapters: 1, 2, 3
Year the summary was received: 2018/2019

Summary 2

Course: Probability
Book: Dictate
Chapters: 1, 2, 3, 4, 6 and 8.1
Year the summary was received: 2018/2019

Summary 3

Course: Research Methods
Book: Heerkens, H. & Van Winden, A. (2017) *Solving managerial problems systematically,* Wolters-Noordhoff B.V.
Chapters: 1-11
Year the summary was received: 2018

Summary 4

Course: Excel & VBA Manual summary
Book: Manuals
Chapters:
Year the summary was received: 2023

Summary 1: Mathematics

"Is math related to science" & "Mathematics is the language in which God has written the universe."

Three logicians walk into a bar. Waitress: "Would you all like a beer?"

Logician 1: "I don't know"
Logician 2: "I don't know"
Logician 3: "Yes"

Chapter 1 Basic set theory and logic

Definition Set: A set is a well-defined unordered collection of distinct elements.

- Unordered: $\{1,3,6\} = \{1,6,3\}$
- Distinct elements: $\{1,3,6\} = \{1,6,3,6\}$
- Well-defined (not ambiguous): 'All friendly people' is not a set.

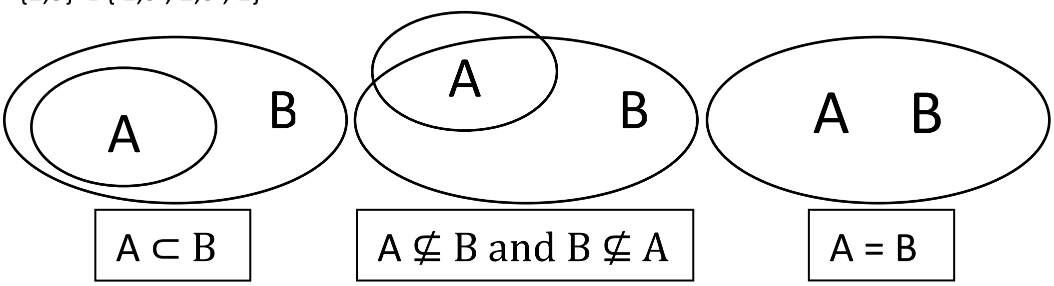
Examples:

$\{1,2,3,4,5,6,7,8,9,10\} = \{1,2,3,\ldots,10\}=\{n \mid n \text{ is an integer with } 1 \leq n \geq 10\}$
This means $4 \in \{1,2,3,\ldots,10\}$

Definition subsets $A \subseteq B$: every element $a \in A$ is also an element of B (if $a \in A$, then $a \in B$)
Definition proper subset $A \subset B$: $A \subseteq B$ and $A \neq B$
Definition equal sets $A=B$: $A \subseteq B$ and $B \subseteq A$ (A and B contain exactly the same elements)
Definition empty set $\emptyset$: $\emptyset$ is a set without any elements ($x \notin \emptyset$ for all x)

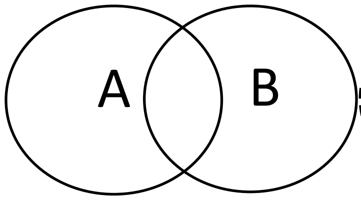
There exist only one empty set , the one with no elements
$\emptyset \subseteq \emptyset'$ and $\emptyset' \subseteq \emptyset$, so that means $\emptyset = \emptyset'$

- $\{1,5\} \subseteq \{1,2,3,\ldots\}$
- $\{1,5\} \subset \{1,2,3,\ldots\}$
- $\{1,5\} \subseteq \{1,5\}$
- $\{1,5\} \not\subset \{1,5\}$
- $\{1,5\} \nsubseteq \{ 1,5 , 2,5 , 1\}$
- $\{1,5\} \in \{ 1,5 , 2,5 , 1\}$

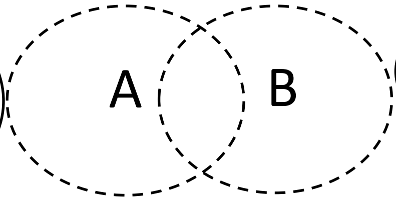


Definition intersection $\cap$ and union $\cup$:

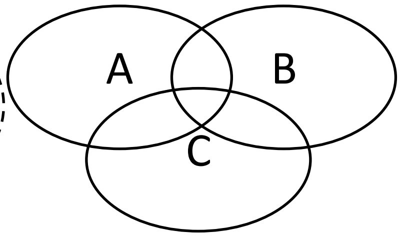
$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$
$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$



A ∩ B



A ∪ B



A ∩ B ∩ C

Which of the following is correct, given $A \cap B \subseteq B$ and $A \cup B \subseteq$ A?:

- A⊆B
- B⊆A
- B=A
- B=∅

From $A \cap B \subseteq B$ we can't conclude anything, since $A \cap B$ is always a subset of B.

$A \cup B \subseteq A$ means that every element of $A \cup B$ is also in A.
In particular, $A \cup B$ contains all elements of B, so every element of B is also in A. **Therefore A ⊆ B.**

Examples:

A={7,8} & D={7} & E={7,9} & ∅={ }

D⊆A because all ∈ off D are in A
E⊈A because not all ∈ off E are in A
∅⊆A because all ∈ off ∅ are in A (∅ is always a subset of an other set)

If a then b means a→b
a sufficient for b means a→b
a necessary for b means b→a

Natural numbers $\mathbb{N} = \{ 1,2,\ldots\}$
Integers $\mathbb{Z} = \{0,-1,1,-2,2,-3,3,\ldots\}$
Rational numbers $\mathbb{Q} = \left\{\frac{a}{b} \middle| a, b \in \mathbb{Z}, \mathrm{b} \neq 0\right\}$
Real numbers $\mathbb{R}$ 'all numbers on the real number line'

$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subseteq \mathbb{R}$

Definition proposition: a proposition is a statement whose truth can be expressed by the values 'true' and 'false'.

Examples:
Proposition = 'it rains' and 'if it rains, I will stay inside'
Conclusion = 'I will stay inside'

Proposition = 'if it rains, I will stay inside' and 'I will stay inside'
Conclusion = '???'

Symbol	Name	Proposition	... is true if and only if
∧	And-operator	P ∧ Q	P is true and q is true
∨	Or-operator	P ∨ Q	P is true and/or q is true
¬	Not-operator	¬P	P is false
→	Implication-operator	P → Q	If p is true, then q is also true

↔	Equivalence- operator	P ↔ Q	(P → Q) ∧ (Q →P)

Truth tables

P	Q	¬P	P ∧ Q	P ∨ Q	P → Q	P ↔ Q
0	0	1	0	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	0	0
1	1	0	1	1	1	1

Definition predicate: a predicate is a statement that can contain variables that influence the truth value of the statement

- Predicates
  - 1+n > 5
  - If it rains x will stay inside
  - N^2 < 0
- Propositions
  - 1+8 > 5
  - If it rains, all persons stay inside
  - There exists an integer n ∈ ℤ such that n^2 < 0

Predicates can be turned into propositions by choosing a value.

∀ means 'universal quantifier' ➔ For all
∃ means 'existential quantifier ➔ There exists a …

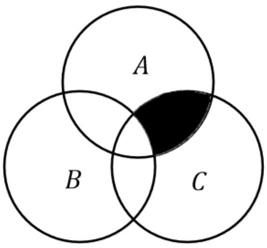
<u>Introduction to mathematics</u>

What does ∃ n ∈ ℤ(m=2n) mean?

a. **M is even**
b. Even numbers exist
c. All numbers are even
d. There exist numbers that are twice as large as other numbers

The black area represents?

a. A∪C
b. A∩C
c. (A∪C)-B
d. **(A∩C)-B**



Chapter 2 Proof techniques

Different proof techniques:

- (counter)example

Proposition: Cows don't exist
Counterexample → you see a cow
Proposition: Every person has the same gender
Counterexample → one male and one female

Theorem: ∃ m ∈ ℤ(∃ n ∈ ℤ(m=2n))

- Direct proof

Theorem: There exist at least two Facebook users with the same number of Facebook friends

Proof:

There are billion pf users on Facebook. Denote the set of users by U
Each user has a number of friends from the set F={0,1,2,…,5000}
Since |U|>|F|, at least two users have the same number of friends

|U|>|F| here they use the pigeonhole principle: When n items are put into m containers, where n > m, then at least one container contains more than one item.

- Case distinction

Theorem: Even without a limit to the number of friends, there exist at least two Facebook users with the same number of Facebook friends.

Proof:

Denote the number of Facebook users by n.
Each user had a number of friends from the set F = {0,1,2,…,5000}

Case 1: suppose no user has 0 friends. We have |{1,2,…,n-1}| = n –1 < n. By the pigeonhole principle, the theorem holds.

Case 2: suppose some user u has 0 friends. Now no user had n-1 friends, since u is not her friend. Therefore |{1,2,…,n-2}| = n –1 < n. By the pigeonhole principle, the theorem holds

Example:

1. *If there are at least two students, there are always at least two students with the same number of students in their house.*
2. *If there are at least two students, there are always at least two students helping the same number of students with their study.*
    a. *Both statements are false*
    b. *Statement 1 is false, and statement 2 is true*
    c. *Statement 1 is true, and statement 2 is false*
    d. *Both statements are true*

Solution:

1. *If you define two students who live together as 'Facebook friends'*
   By the theorem on Facebook friends, this statement holds
2. *Consider two students {1,2}. 1 helps 2 doesn't mean 2 helps 1*

Theorem: Even without a limit of the number of friends, there exist at least two Facebook friends with the same number of friends.

Proof:

Number of Facebook users = n. Each user has a number of friends F={1,2,…,n-1}

Case 1: suppose no user had 0 friends |{1,2,…,n-1}| = n –1 < n.
Case 2: suppose some user has 0 friends. Now no users has n-1 friends, since u is not her friend. Therefor |{0,1,2,…,n-2}| = n –1 < n. by the pigeonhole principle, the theorem holds.

- Proof by contradiction

Theorem: $\sqrt{2} \notin \mathbb{Q}$

Proof:

Suppose $\sqrt{2} \in \mathbb{Q}$ i.e. (that is) $\exists p,q \in \mathbb{Z}(\sqrt{2} = \frac{p}{q})$. Assume w.l.o.g. (with loss of generality) that p and q have no common divisor. (If they do, divide both by this divisor.)

$\sqrt{2} = \frac{p}{q}$ ➔ $2 = \frac{p^2}{q^2}$ ➔ $2q^2=p^2$

➔ p^2 is even ➔ p is even (prove yourself…)

➔ $\exists\ n \in \mathbb{Z}(p=2n)$

➔ $2q^2 = p^2 = (2n)^2 = 4n^2$ ➔ $q^2=2n^2$

➔ q^2 is even ➔ q is even

➔ 2 divides both p and q ➔ CONTRADICTION!

Try to prove: $\sqrt{\sqrt{2}} \notin \mathbb{Q}$ and $\sqrt{3} \notin \mathbb{Q}$

- Mathematical induction

Given is the first domino falls. If the first domino falls, the next domino will also fall.

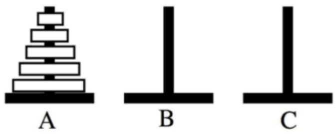
Definition (mathematical induction): Consider a predicate S(n), $n \in \mathbb{N}$.

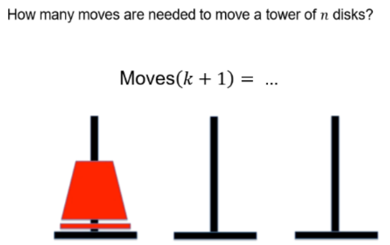
Suppose that:
1. S(1) is true
2. $\forall\ k \in \mathbb{N}\ (S(k) \rightarrow S(k+1))$

Then S(n) is true for all $n \in \mathbb{N}$

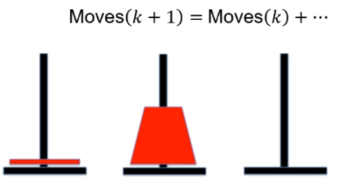
Goal: Move all disks from rod A to rod C

Rules: Move one disk at the time ; No disk can be places on top of a smaller disk

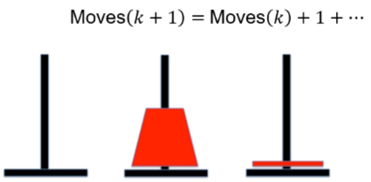




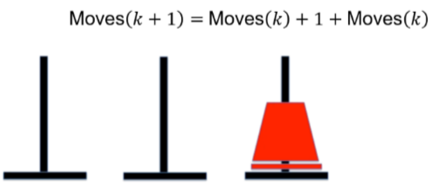
1.



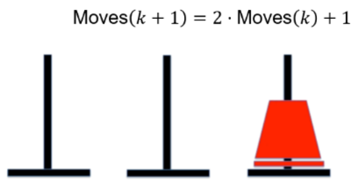
2.



3.



4.



5.

Lemma: $\forall k \in \mathbb{N}$(Moves k + 1) = 2 · Moves (k) + 1
Theorem: $\forall n \in \mathbb{N}$(Moves (n) = $2n - 1$)
Mathematical induction: Consider a predicate $S\ n\ , n \in \mathbb{N}$
Suppose that: 1. $S(1)$ (is true)
2. $\forall k \in \mathbb{N}\ (S\ (k) \rightarrow S\ (k + 1))$
Then $S\ (n)$ is true for all $n \in \mathbb{N}$

Proof:

S(n): moving a tower of n disks takes $2^n - 1$ moves
Step 1: prove S(1) → moves (1)=$2^1 - 1 = 1$, the number of moves required to move 1 disk

Step 2: Prove $\forall k \in \mathbb{N}\ (S(k) \rightarrow S(k + 1))$.
$S\ (k)$:Moving k disks takes $2k - 1$ moves
$S\ (k + 1)$:Moving $k + 1$ disks takes $2k$+1 − 1 moves
Moves (k + 1) = 2 · Moves k + 1 → (Lemma)
= 2 · $2k$ − 1 + 1 → (By Induction Hypothesis)
= $2k$+1 − 2 + 1 = $2k$+1 − 1

<u>Introduction to mathematics</u>

Where does the proof that every person has a common gender fail?

a. The proof does not consider the group of 0 people
b. Groups can be bigger that the number of people on earth
c. **The induction step does not hold for all k**
d. The basis step should not start with n=1

Fake theorem: every person has the same gender

Proof:

S(n): for every group of size n, all persons in that group have the same gender
Step 1: each person has the same gender as him/herself
Step 2: prove $\forall$ k $\in \mathbb{N}$(S(k)→S(k+1))

Suppose for some k, S(k): in every group of size k, each person has the same gender.
We now need to show S(k+1): in each group {1,2,…,k+1} of size k+1 each person has the same gender
By the induction hypothesis, both {1,2,…,k+1} and {2,3,…,k+1} have the same gender. This is the same gender for all {1,2,..,k+1}, since it is shares by {2,3,..,k+1}

Counterexample:

Suppose S(1): in every group of size 1, each person has the same gender
We now need to show S(2): in each group {1,2} of size 2 each person has the same gender.
By the induction hypothesis, both {1} and {2} have a common gender within their group. However, the groups don't have any person in common.

Chapter 3 Counting

Theorem (sum rule): If A and B are finite disjoint sets ($A \cap B = \emptyset$), then $|A \cup B| = |A| + |B|$

Example: I can pick my lunch by selecting from a set of 3 fruits A={apple, banana, cherry} or a set of 3 vegetables B = {eggplant, broccoli}. Now I have $|A|+|B|$=3+2=5 choices for lunch

Theorem (product rule): If A and B are finite sets, then there are $|A| \cdot |B|$ different pairs (a, b) with $a \in A$, $b \in B$

Example: My lunch consists of 1 fruit and 1 vegetable from a set of 3 fruits A={apple, banana, cherry} and a set of 2 vegetable B={eggplant, broccoli}.
Now I have $|A| \cdot |B| = 3 \cdot 2 = 6$ choices of lunch

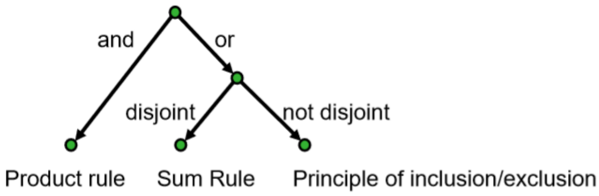
My lunch consists of 1 fruit and 1 vegetable from a set of 3 fruits A = {Apple, Banana, Tomato} and a set of 2 vegetables B = {Eggplant, Tomato}.
Now I have $|A| \cdot |B| = 3 \cdot 2 = 6$ choices for lunch.

Theorem (sum rule revised): If A and B are finite disjoint sets ($A \cap B = \emptyset$), then $|A \cup B| = |A| + |B|$

Example: I can pick my lunch by selecting from a set of 3 fruits A = {Apple, Banana, Tomato} or a set of 2 vegetables B = {Eggplant, Tomato} . Now I have only $4 \neq |A| + |B|$ choices for lunch.

Theorem (principle of inclusion/exclusion): If A and B are finite sets, then $|A \cup B| = |A| + |B| - |A \cap B|$.

Example: I can pick my lunch by selecting from a set of 3 fruits A = {Apple, Banana, Tomato} or a set of 2 vegetables B = {Eggplant, Tomato} . Now I have $|A| + |B| - |A \cap B| = 3 + 2 - 1 = 4$ choices for lunch.



Definition permutations:

a permutation of n distinct elements is an ordering of the n elements. There are $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ possible permutations of n elements.

Definition r-permutations:

an r-permutation of n distinct elements is an ordering of r out of n elements, where $0 \leq r \leq n$. There are $P(n, r) = n \cdot (n-1) \cdot \ldots \cdot (n-r+1) = \frac{n!}{(n-r)!}$ Possible r-permutations of n elements

Example: Suppose we want to order four out of six possible foods {apple, banana, cherry, broccoli, eggplant and tomato}. For example: (tomato, apple, eggplant, broccoli).
Number of possible orderings: $6 \cdot 5 \cdot 4 \cdot 3 = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = \frac{6!}{2!}$, with $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ is n and $6 \cdot 5 \cdot 4 \cdot 3$ is r.

Definition combinations:

An r-combination of n distinct elements is an unordered subset of r out of n elements, where $0 \leq r \leq n$ There are $C(n,r) = \binom{n}{r} = \frac{P(n,r)}{r!} = \frac{n!}{r! \cdot (n-r)!}$ possible r-combinations of n elements.

Permutations	Combinations
P(n,r)	C(n,r) = $\binom{n}{r}$
Order matters	Order doesn't matter
$\frac{n!}{(n-r)!}$	$\frac{n!}{r!\cdot(n-r)!}$

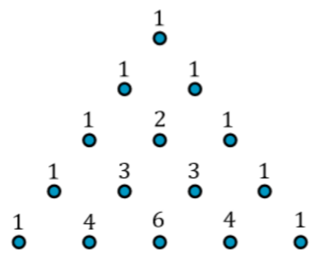
Theorem (birthday paradox): in a group of 23 people, the probability that at least 2 share their birthday is approximately 50%

Solution:

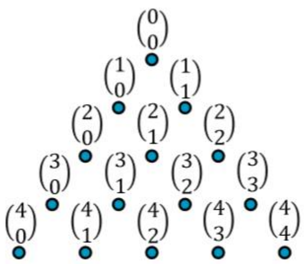
- A scenario is an allocation of birthdays to the 23 people, e.g. (March 14th, January 5th , …)
- For simplicity assume that all scenarios are equally likely.
- Total number of scenario's: 365^{23} (product rule)
- Number of scenario's for which no pair shares a birthday: $365 \cdot 364 \cdot \ldots \cdot (365 - 23 + 1) = \frac{365!}{342!}$
- Number of scenario's for which at least one pair shares a birthday: $365^{23} - \frac{365!}{342!}$
- Probability at least one pair shares birthday: $\frac{365^{23} - \frac{365!}{342!}}{365^{23}} \approx 50\%$

<u>For the test, you don't need to know probabilities, but you can compute the number of scenario's.</u>

Pascal's triangle (Newton's binomial theorem)



Numbers denote number of ways to arrive at a position



Numbers denote number of ways to arrive at a position.

Going right r out of n times: $\binom{n}{r}$

$(x + y)^2 = x^2 + 2xy + y^2$: represents the third row from the top of the pyramid
$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$: represents the fourth row from the top of the pyramid

Binomial Theorem: $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$

Summary:

- *And means: product; or means: sum of disjoint; otherwise you use principle of inclusion/exclusion*
- *Permutations: Order matters*
- *Combinations: Order does not matter*

Summary 2: Probability

Chapter 1 Experiment, sample space and probability

An experiment is **probabilistic** of **stochastic** if the experiment does not necessarily lead to the same outcome when it is repeated under equal conditions.

The **sample space S** of an experiment is the set of all possible outcomes.
An **event** is a subset of the sample space.
The **empty set Ø** is called the **impossible event**. The set S is called the **certain event** since every outcome is in S and thus this event will always occur. An event which consist of a single outcome s is called an **elementary event**: {s}. both s and {s} are sometimes called a **'sample point'**.

If A and B are events, then:

* $\boldsymbol{\bar{A}}$ is the **complement of A** or the **complementary event** of A (i.e. the event which occurs if A does not occur).
* $\boldsymbol{A} \cup \boldsymbol{B}$, the **union** of A and B -> at least one or both of the events A and B occurs.
* $\boldsymbol{A} \cap \boldsymbol{B}$ or **AB**, the **intersection** of A and B -> both A and B occur.
* $\boldsymbol{A} \subset \boldsymbol{B}$ (A is a **subset** of B): A implies B -> if A occurs, then also does B.

A and B are **mutually exclusive** (or **disjoint)** events if **AB = Ø**.
The events $A_1, A_2, \ldots, A_n$ or $A_1, A_2, \ldots$ are called **mutually exclusive** if $A_iA_j = \emptyset$ for every possible combination (i, j) for which $i \neq j$.
The sequence of events $\{A_i\}$ is a **partition** of the event B if the events A_i are mutually exclusive and $B = \bigcup_i A_i$.

Properties of events:

* $A \cap (B \cup C) = (A \cap B) \cup (A \cap B)\ and\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
* $A \cup B = A \cup (\bar{A}B)\ and\ B = (AB) \cup (\bar{A}B)$
* **De Morgan's laws**: $\overline{A \cup B} = \bar{A} \cap \bar{B}\ and\ \overline{A \cap B} = \bar{A} \cup \bar{B}$

The **probability of event A** of S is **P(A)**. The **pair (S, P)** a **probability space**.
Laplace's definition: When the sample space S of an experiment N(S) contains equally likely outcomes and the event A contains of N(A) outcomes, then the **probability of an event A**, denoted by P(A) equal: $\boldsymbol{P(A) = \frac{N(A)}{N(S)} = \frac{favorable\ number}{total\ number}}$.

If S is a finite sample space of an experiment and the probabilities P(A) of events A are defined according to Laplace's definition, the pair (S, P) is called a **symmetric probability space.**

Properties of a symmetric probability space:

* $P(A) \geq 0$ for every event A,
* $P(S) = 1$,
* If $A \subset B, then\ P(A) \leq P(B)$,
* $\boldsymbol{P(\bar{A}) = 1 - P(A)}$
* If $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

Assume that we have an experiment with sample space S which we can repeat arbitrarily often. If the event A occurred n(A) times in total with n repetitions, then we define

$$\boldsymbol{f_n(A) = \frac{n(A)}{n}}$$

as the **relative frequency of A** in n repetitions.
Experimentally it appears that $f_n(A)$ for increasing n 'converges' to a constant, the probability of A. This phenomenon is called **empirical lax of large number**s.

Axioms of Kolmogorov:

Consider an experiment with a random non-empty sample space S. A function P which assigns a real number P(A) to every event $A \subset S$, is called a **probability** or **probability measure** on S if:

* $\boldsymbol{P(A) \geq 0}$ for every event A,
* $\boldsymbol{P(S) = 1}$ and
* For every countable sequence of mutually exclusive events $A_1, A_2, \ldots, A_n$ or $A_1, A_2, \ldots$:
$\boldsymbol{P(\bigcup_i A_i) = \sum_i P(A_i)}$

When S is a sample space and P is the probability on S then we call the pair (S, P) a probability space.

* $\boldsymbol{P(\emptyset) = \emptyset}$
* **Complement rule:** $\boldsymbol{P(\overline{A}) = 1 - P(A)}$
* For two events A and B with $\boldsymbol{A \subset B}$ we have: $\boldsymbol{P(A) \leq P(B)}$**.**
* The **general addition rule**: for two events A and B (which are not necessarily mutually exclusive): $\boldsymbol{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$
* **The addition rule for mutually exclusive events:**
If A and B are mutually exclusive, then: $\boldsymbol{P(A \cup B) = P(A) + P(B)}$**.**

Chapter 2 Combinatorial Probability

Combinatorics is the 'art of counting'.

The product rule: When an experiment consists of performing k partial experiments and the i^{th} partial experiment has n_i possible outcomes (i = 1, …k), no matter what the results of the partial experiments are, then $\boldsymbol{n_1 * n_2 * \ldots * n_k}$ outcomes of the total experiment are possible.

* **5!** is pronounced as '5 **factorial**'. $5! = 5 * 4 * 3 * 2 * 1$

The permutation rule: The number of orders or **permutations** in which k different things can be arranged in **k!**.

When $A_1, A_2, \ldots, A_k$ are **mutually exclusive** events, then: $\boldsymbol{N(\bigcup_{i=1}^{k} A_i) = \sum_{i=1}^{k} N(A_i)}$**.**

* With **permutations** is the **order important** and it is **without replacement**. Think of an vase model.
* With **combinations** is the **order not important** and it is **without replacement**.

		Method of drawing	
		Without replacement	With replacement
Kind of outcome	Ordered	**Permutations** of 3 out of 10: S = {numbers with 3 different digits} $\boldsymbol{N(S) = 10 * 9 * 8 = \frac{10!}{7!}}$ **nPr-buuton: 10nPr3**	S = {numbers with, possibly the same, 3 digits} $\boldsymbol{N(S) = 10^3}$
	Unordered	**Combinations** of 3 out of 10: S = {subsets with 3 different digits} $\boldsymbol{N(S) = \binom{10}{3}}$ **nCr-button: 10nCr3** **'10 choose 3'** $\binom{10}{3}$ is called the **binomial coefficient.**	S = {Unordered threesomes with repetitions} $(N(S) = \binom{12}{3})$ (S, P) is **not symmetric**: refine the sample space to the corresponding symmetric probability space of ordered threesomes.

We randomly draw k times from a set of n different elements, then in the following three cases the probability space is **symmetric**:

a. Draw with replacement, ordered outcomes: $\boldsymbol{N(S) = n^k}$

b. Draw without replacement, ordered outcomes: (**variations/permutations of k out of n)**:
$N(S) = n * (n-1) * \dots * (n-k+1) = \frac{\boldsymbol{n}!}{(\boldsymbol{n}-\boldsymbol{k})!}$

c. Draw without replacement, unordered outcomes: (**combinations of k out of n):**
$\boldsymbol{N(S)} = \binom{\boldsymbol{n}}{\boldsymbol{k}}$

d. Drawing with replacement, unordered outcomes, the probability space is **non-symmetric**. We can redefine a symmetric probability space by considering the corresponding ordered outcomes.

* $\mathbf{0! = 1}$
* **# = number**

Hypergeometric formula: If we draw n times, at random and without replacement, from a set of N balls, consisting of R red and N-R white balls, the probability of event A_k, that we draw k red (and n-k white) balls, is given by:

$$\boldsymbol{P(A_k)} = \frac{\binom{R}{k} * \binom{N-R}{n-k}}{\binom{N}{n}}$$

Using this formula we can compute probabilities of a specific number of k red balls. Such a numerical variable is called a **random variable X**, a quantitative variable in a stochastic experiment. If we determine for X all possible probabilities P(X = k), using the hypergeometric formula, this list of probabilities is called **the hypergeometric distribution of X.**

Chapter 3 Conditional probability and independence

Conditional probability is the probability of a subpopulation in a population.
When A and B are events and P(B) > 0, then we define

$$\boldsymbol{P(A|B) = \frac{P(AB)}{P(B)}}$$

as **the (conditional) probability of A under condition of B** (or **the (conditional) probability of A given B).** -> P(B) > 0

* $P(\bar{A}|B) = 1 - P(A|B)$

The general product rule: $\boldsymbol{P(AB) = P(A) * P(B|A)}$.

For n events $A_1, A_2, \dots, A_n$ with $n \geq 2$ and $P(A_1 A_2 \dots A_{n-1}) > 0$ we have:
$\boldsymbol{P(A_1 A_2 \dots A_n) = P(A_1) * P(A_2|A_1) * \dots * P(A_n|A_1 A_2 \dots A_{n-1})}$

The law of total probability: $\{S_i\}$ is a partition of S such that $P(S_i) > 0$ for all i, then for each event A we have: $\boldsymbol{P(A) = P(A|S_1) * P(S_1) + P(A|S_2) * P(S_2) + \dots = \sum_i P(A|S_i) * P(S_i)}$

* $P(A) = P(A|B) * P(B) + P(A|\bar{B}) * P(\bar{B})$
* $\boldsymbol{P(ABC) = P(A) * P(B|A) * P(C|AB)}$

The complement rule: $\boldsymbol{P(\bar{A}) = 1 - P(A)}$

Bayes' rule: If $\{S_i\}$ is a partition of S with $P(S_i) > 0$ for each i, then for each event A with P(A) > 0 we have:

$$\boldsymbol{P(S_k|A) = \frac{P(AS_k)}{P(A)} = \frac{P(A|S_k)P(S_k)}{\sum_i P(A|S_i) * P(S_i)}}$$

The events A and B are **independent** when $\boldsymbol{P(AB) = P(A) * P(B)}$.
When we say that two **experiments** are **independent**, we mean that every pair of events A and B, where A only relates to the first experiment and B only to the second, can be assumed independent.
We will call the events $\{A_i\}$ **pairwise independent** when each pair in this sequence of events is independent. Pairwise independence does not rule out that there is a certain dependence between events.

1. $\boldsymbol{P(ABc) = P(A) * P(B) * P(C)}$
2. $P(AB) = P(A) * P(B)$
3. $P(AC) = P(A) * P(C)$
4. $P(BC) = P(B) * P(C)$

The events A_1, A_2 A_3, … are independent if for each subsequence A_{i1}, A_{i2}, …, A_{ik} with $k \geq 2$, it is true that: $\boldsymbol{P(A_{i1}A_{i2} \dots A_{ik}) = P(A_{i1}) * P(A_{i2}) * \dots * P(A_{ik})}$

A series of experiments is called **Bernoulli experiments/trails** if:

1. Each experiment/trail has two possible outcomes, often denoted with 'Success' and 'Failure',
2. The experiments are independent and
3. The probability of success is the same for each experiment.

The **success probability** is usually denoted by **p** and the **probability of failure** with **1 – p**.

The **binomial formula**: If X is the number of successes in n Bernoulli experiments/trails with success probability p, then:
$\boldsymbol{P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}}$, where k = 0, 1, 2, …, n.
X is said to have a **binomial distribution** with parameters n and p.

> $p^k(1-p)^{n-k}$ is the probability that the first k trails are successful and the last n – k are failures and $\binom{n}{k}$ the number of possible orders of k successes and n – k failures.

The **geometric formula**: If we conduct Bernoulli trails with success probability p until a success occurs and X is the number of required trails, then: $\boldsymbol{P(X = k) = (1-p)^{k-1} p}$, where k = 1, 2, 3,…
$\boldsymbol{E(X) = \frac{1}{p}}$ **is the expected value of X.**

X is said to have a **geometric distribution** with parameter p.

Chapter 4 Discrete random variables

If S is the sample space of an experiment, then a real function $X: S \to \mathbb{R}$, which assigns a real number X (s) to each outcome $s \in S$, is a **random variable**.

* $\mathbb{R}$ is the set of all real numbers.

The **range S_X** of a random variable X, defined on a sample space is the set of all possible realizations X (s). So $S_X = \{X(s) | s \in S\}$.
The range of a variable can be

- **Finite:** -> $S_X = \{1, 2, 3, 4, 5, 6\}$
- **Countably finite:** 'countable' using the natural numbers 1, 2, 3,…
  -> $S_X = \{1, 2, 3, 4, \dots\}$
- **Not countably infinite**: -> $S_X = [-, \infty)$

A **random variable** X is **discrete** if the range S_X is denumerable.
If X is discrete, S_X has the shape $\{x_1, x_2, \dots, x_n\}$ or $\{x_1, x_2, x_3, \dots\}$.

If X is a discrete random variable, then we will call the function that assigns a probability $P(X = x)$ to each $x \in S_x$ **the probability function of X**. The **probability function of X** is usually graphed using a so called bar graph of probabilities, with x in S_X on the X-axis and the probabilities P(X = x) on the Y-axis. Note that the total length of the bars is the total probability 1.

For the probability function of a discrete random variable X we have:

1. $\boldsymbol{P(X = x) \geq 0 \; for \; x \in S_X \; and}$
2. $\boldsymbol{\sum_{X \in S_X} P(X = x) = 1}$

The probabilities $P(X \in B)$ for each $B \subset S_X$ are, all together, called the **(probability) distribution** of the random variable X.
When all probabilities are equal we say that X has a **homogeneous distribution** on S_X.

The **expectation** of **expected value E(X)** of a discrete random variable X is given by:

$$\boldsymbol{E(X) = \sum_{x \in S_X} xP(X = x)}$$

Instead of the symbol E(X), **EX, μ or μ_X** is also used.
Since the variable X with range S_X and probability function P(X = x) form a probability model for a population, E(X) is often referred to as the **population mean**.
The **sample mean $\bar{x}$** is the **average value of observations in a sample**.

If the probability function is symmetric with respect to x = c, then **E(X) = c**.

E(X) is sometimes confused with the so called **median**: that is the value M, such that $P(X \leq M) = P(X \geq M) \geq 50\%$ or $\boldsymbol{P(X \leq M) = P(X \geq M) = \frac{1}{2}}$

If X is a discrete random variable and g a (real) function, then:

$$E\big(g(X)\big) = \sum_{x \in S_X} g(x)P(X = x)$$

If Y is a **linear function of X**, that is Y = aX + b for any real constants $a, b \in \mathbb{R}$, then

$$\boldsymbol{E(aX + b) = a * E(X) + b}$$

If X is a discrete random variable and g and h are real functions, then for real constants $a, b \in \mathbb{R}$ we have:

1. $\boldsymbol{E(aX + b) = a * E(X) + b}$
2. $\boldsymbol{E[ag(X) + bh(X)] = aEg(X) + bEh(X)}$

$$E\big(X^k\big) = \sum_x x^k P(X = x)$$

$\boldsymbol{E\big(X^k\big)}$ is the $\boldsymbol{k^{th}}$ **moment** of the random variable X (k = 1, 2, 3, …)
$E(X^k)$ is a weighted average that can be considered as a **measure for the center** of the distribution of X. However E(X) does not tell anything about the **magnitude of the differences** in the values of X. The thing that tells us that is the concept of **measure of variation**.

The **variance** of X ($\boldsymbol{var(X) \; or \; \sigma_X^2}$) is defined as

$$\boldsymbol{var(X) = E(X - \mu_X)^2 = E\big(X^2\big) - \big(E(X)\big)^2}$$

The **standard deviation of X** ($\boldsymbol{\sigma_X}$) is the square root of the variance: $\boldsymbol{\sigma_X = \sqrt{var(X)}}$

Properties of variance and standard deviation:

1. $\boldsymbol{var(X) \geq 0 \; and \; \sigma_X \geq 0}$
2. $\boldsymbol{var(X) = E\big(X^2\big) - \mu_X^2}$
3. $\boldsymbol{if \; var(X) > 0, so \; if \; X \; is \; not \; degenerate, we \; have \; E(X^2) > (EX)^2}$
4. $\boldsymbol{var(aX + b) = a^2 * var(X) \; and \; \sigma_{aX+b} = |a| * \sigma_X}$

Chebyshev's inequality: For any real number $c > 0$, we have: $\boldsymbol{P(|X - \mu_X| \geq c) \leq \frac{var(X)}{c^2}}$
The inequality is valid for any random variable X and gives us an **upper bound of the probability** of values **outside the interval ($\boldsymbol{\mu_X - c, \mu_X + c}$),** so deviating more than c from µ.
Chebyshev's rule is valid for any distribution.

Empirical rule: If the graph of the distribution of X shows a bell shape, then the approximately probabilities for X having a value within the interval

* $\boldsymbol{(\mu - \sigma, \mu + \sigma)\ is\ 68\%}$
* $\boldsymbol{(\mu - 2\sigma, \mu + 2\sigma)\ is\ 95\%}$
* $\boldsymbol{(\mu - 3\sigma, \mu + 3\sigma)\ is\ 99,7\%}$

Alternative distribution:

X = 1 -> success
X = 0 -> failure

$\boldsymbol{X \sim B(1, p)}$

* $\boldsymbol{P(X = 1) = p}$
* $\boldsymbol{E(X) = p}$
* $\boldsymbol{E(X^2) = p}$
* $\boldsymbol{var(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p)}$

When for example to use: 'dice results is 6 (X = 1) of not (X = 0).

* **~** means 'is having'.

Binomial distribution:
X is **binomially distributed (X ~ B(n, p)) with parameters n and p**, for all n = 1, 2, … and $p \in [0, 1]$, if the probability function of X is given by:

* $\boldsymbol{P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}}$**, where k = 0, 1, 2, …, n**
* $\boldsymbol{E(X) = np}$
* $\boldsymbol{var(X) = np(1 - p)}$

Use when we have **random draws** from a population the independence is only secured if we draw **with replacement**. Example: 'number of sixes is 30 rolls of a dice'

* If **p = 1**, then X has a **degenerate distribution**.
* If **n = 1**, then X has an alternative distribution.

Hypergeometric distribution:
X is **hypergeometrically distributed** (with parameters N, R and n) if

* $\boldsymbol{P(X = k) = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}}$**, where k = 0, 1, 2, …, n**
* $\boldsymbol{E(X) = np}$
* $\boldsymbol{var(X) = np(1 - p) * \frac{N-n}{N-1}}$**, where** $\boldsymbol{p = \frac{R}{N}}$

For relatively large R and N-R and relatively small n the hypergeometric distribution with parameters N, R and n can be approximated by a $\boldsymbol{B\left(n, \frac{R}{N}\right)}$**-distribution**. A rule of thumb for applying the binomial distribution from above is $\boldsymbol{N > 5n^2}$.

We apply this distribution whenever we consider a number of **random draws without replacement** from a so called **dichotomous population**: consisting of elements which do or do not have a specific property. Example: 'number of girls if we choose 5 from a group of 10 boys and 12 girls'.

Geometric distribution:
X had a **geometric distribution** with parameter $p \in (0, 1]$, if

* $\boldsymbol{P(X = k) = (1 - p)^{k-1} p}$, where k = 1, 2, …
* $E(X) = \frac{1}{p}$
* $\boldsymbol{var(X) = \frac{1-p}{p^2}}$
* $\boldsymbol{P(X > k) = (1 - p)^k}$ -> we need more that k trails to score a success equals the probability that we are not successful in the first k trails.

Example use: 'number of rolls of a dice until 6 occurs'.
Poisson distribution:
X has a **Poisson distribution** with parameter **μ > 0**, if

* $\boldsymbol{P(X = k) = \frac{\mu^k e^{-\mu}}{k!}}$, **for k = 0, 1, 2, …**
* $\boldsymbol{E(X) = \mu}$
* $\boldsymbol{var(X) = \mu}$

If X has a B(n, p)-distribution with "**large n and small p"**, then X has approximately a Poisson distribution with parameter $\boldsymbol{\mu = np}$**.**
A rule of thumb for the above is $\boldsymbol{n > 25 \text{ and } np < 10 \text{ or } n(1 - p) < 10}$.
This distribution is used with **numbers of rare events**. Usually **the area and the time interval** for the occurrence of these events are restricted. Example 'number of clients that enter an office in 10 minutes'.
Homogeneous distribution:

* $\boldsymbol{P(X = x) = \frac{1}{N}}$, **x = 1, 2, …, N**
* $\boldsymbol{E(X) = \frac{N+1}{2}}$

Example 'result of one roll of a dice'

Chapter 6 Continuous random variables

For **continuous random** variable, in general, we have $\boldsymbol{P(X = x) = 0}$ for each real value of x.
For continuous variables we cannot define the probability model by a probability function: we need another kind of model where probabilities that the variable X attains values in an interval of real values lead to positive values. Such a model is given by the **(probability) density function (pdf)** of a continuous random variable.

The **density function** of a continuous random variable X is a non-negative function f, such that
$$\boldsymbol{P(a \le X \le b) = \int_a^b f(x)dx = F(x)\Big|_{x=a}^{x=b} = F(b) - F(a)}$$
The total area under the graph of f should be1 (the total probability is 100%)

f is a density function if:

a. $\boldsymbol{f(x) \ge 0 \text{ and}}$
b. $\boldsymbol{\int_{-\infty}^{\infty} f(x)dx = 1}$

The **expectation (expected value)** of a continuous random variable X is

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

For every real valued function g we have: $Eg(X) = \int_{-\infty}^{\infty} g(x)f(x)dx$

* $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$
* $var(X) = E(X^2) - \left(E(X)\right)^2$

The function F, defined $F(x) = P(X \leq x)$ with $x \in \mathbb{R}$, is the (cumulative) **distribution function (cdf)** of the random variable X.
$F_X(x)\ and\ F_Y(y)$ are the **marginal distribution functions** of the random variables X and Y.

For any distribution function F(x) of a random variable X, we have:

a. **F is non-decreasing (if $x_2 > x_1$, then $F(x_2) > F(x_1)$).**
b. $\lim_{x\to\infty} F(x) = 1\ and\ \lim_{x\to-\infty} F(x) = 0$
c. **F is continuous from right ($\lim_{h\to0^+} F(x+h) = F(x)$)**

A random variable X is **continuous** if the distribution function F of X is a continuous function.
The relation between density and distribution function is: $F(x) = P(X \leq x) = \int_{-\infty}^{\infty} f(u)du$
F(x) is a specific anti-derivative -> $F'(x) = f(x)$.
For a continuous random variable X with density function f and (cumulative) distribution function F we have:

a. $P(X = x) = 0,\ for\ x \in \mathbb{R}$
b. $P(X \in [a, b]) = \int_a^b f(x)dx = F(b) - F(a)$
c. $F(x) = \int_{-\infty}^{x} f(u)du$
d. $f(x) = \frac{d}{dx}F(x)$
e. **If the density function f(x) of X is symmetric about x = c, then $E(X) = c$**

The uniform distribution on the interval [a, b]:
The random variable X has a **uniform distribution on the interval [a, b]**, if

$$f(x) = \begin{cases} \frac{1}{b-a}, for\ x \in [a, b] \\ o,\quad for\ x \notin [a, b] \end{cases}$$

X ~ U(a, b)

The expectation and variance of the uniform distribution on [a, b] are:

* $E(X) = \frac{a+b}{2}$
* $var(X) = \frac{(b-a)^2}{12}$

The exponential distribution with parameter λ:
The random variable X has an **exponential distribution parameter λ** (> 0) if:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}\ for\ x \geq 0 \\ 0 \qquad for\ x < 0 \end{cases}$$

X ~ Exp(λ)

If X is an exponential distributed variable, with parameter λ, then:

a. $P(X > x) = e^{-\lambda x}$, **for** $x \geq 0$
b. $F(x) = \begin{cases} 1 - e^{-\lambda x}\ for\ x \geq 0 \\ 0 \qquad for\ x < 0 \end{cases}$

c. $\boldsymbol{E(X) = \frac{1}{\lambda}}$
d. $\boldsymbol{var(X) = \frac{1}{\lambda^2}}$

The standard normal distribution:
The continuous random variable Z has a **standard normal distribution** if
$\boldsymbol{\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$**, where** $\boldsymbol{z} \in \mathbb{R}$
* $\phi(z) = P(Z \leq z)\ or\ F_Z(z)$
* $\phi(-z) = 1 - \phi(z)$

If the continuous random variable X has a density function f_X, then for Y = aX + b, with $a \neq 0$:
$\boldsymbol{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)}$

Examples:
$Y = \sqrt{X} \rightarrow X \sim Exp(\lambda = 2)$
1. $F_y(Y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F_X(y^2)$
2. $f_Y(y) = \frac{d}{dy}[F_X(y^2)] = 2y * f_x(y^2)$
3. $f_Y(y) = 2y * 2e^{-2y^2}\ for\ y \geq 0$
   $f_Y(y) = 0\ if\ y < 0$

$X \sim U(0,1) \rightarrow Y = -\frac{\ln(x)}{2} \sim Exp(\lambda = 2) \rightarrow Y = -\frac{\ln x}{\lambda} \sim Exp(\lambda)$
1. $F_Y(y) = P\left(-\frac{\ln(x)}{2} \leq y\right) = P(\ln(x) \geq -2y) = P(X \geq e^{-2y}) = 1 - P(X \leq e^{-2y}) = 1 - F_X(e^{-2y})$
2. $f_Y(y) = \frac{d}{dy}[1 - F_X(e^{-2y}) = 2e^{-2y} * f_X(e^{-2y})$
3. $f_Y(y) = 2e^{-2y} * 1 \rightarrow 0 \leq e^{-2y} \leq 1 = y \geq 0$
   $Y \sim Exp(\lambda = 2)$

1. **First express $\boldsymbol{F_Y(y)}$ in $F_X \rightarrow$ $\boldsymbol{F_Y(y) = P(Y \leq y) = \cdots}$**
2. **Then compute the derivative to express $\boldsymbol{f_Y(y)\ in\ f_X \rightarrow\ f_Y(y) = \frac{d}{dy}F_Y(y)}$**
3. **Finally, use the specified density function f_X to determine the formula for $\boldsymbol{f_Y(y)}$**

If X has a uniform distribution on (0, 1), then $\boldsymbol{Y = -\frac{\ln(X)}{\lambda}}$ has **exponential** distribution with parameter λ (> 0)
If random variable X has a **normal distribution with parameters $\boldsymbol{\mu\ and\ \sigma^2}$** if the density function of x is defined by $\boldsymbol{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$, with $x \in \mathbb{R}$
$\boldsymbol{X \sim N(\mu, \sigma^2)}$
The two **inflection points** of the graph of f are $\boldsymbol{x = \mu - \sigma\ and\ x = \mu + \sigma}$

If $X \sim N(\mu, \sigma^2)\ and\ Z \sim N(0, 1)$, then:
a. $\boldsymbol{\sigma Z + \mu \sim N(\mu, \sigma^2)}$
b. **The z-score** $\boldsymbol{\left(\frac{X-\mu}{\sigma}\right) \sim N(0, 1)}$
c. $\boldsymbol{E(X) = \mu\ and\ var(X) = \sigma^2}$

The empirical rule, z-scores and percentiles:
When introducing the variance σ^2 and the standard deviation σ there was given an interpretation of these measures of variation, using the **empirical rule** for mount shaped distributions.

$\boldsymbol{P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \phi\left(\frac{x-\mu}{\sigma}\right)}$, where the real value $\frac{x-\mu}{\sigma}$ is called **z-score** (z-value) of the bound x.
The $\boldsymbol{k^{th}}$ percentile is such a value c that $\boldsymbol{P(X \leq c) = k\%}$

For a $N(\mu, \sigma^2) - distributed$ random variable X we have:
$\boldsymbol{Y = aX + b \text{ is } N(a\mu + b, a^2\sigma^2) - distributed \; (for\ all\ a \neq 0\ and\ b} \in \mathbb{R})$

Chapter 8.1 Waiting times

The distribution of a random variable X has the **lack of memory property** on its range S_X, if for all $t, s \in S_X$
$\boldsymbol{P(X > t + s | X > s) = P(X > t)}$

$$P(Y > m + n | Y > n) = \frac{P(Y > m + n)}{P(Y > n)} = \frac{(1-p)^{m+n}}{(1-p)^n} = (1-p)^m = P(Y > m)$$

For a continuous random variable X with range $S_x = [0, \infty)$ the following statements are equivalent:

* **X is exponentially distributed for parameter λ.**
* $\boldsymbol{P(X > t) = e^{-\lambda t},\ for\ t \geq 0}$
* **The distribution of X has the lack of memory property on S_X and** $\boldsymbol{E(X) = \frac{1}{\lambda}}$

For a discrete random variable X with range $S_X = \{1, 2, \ldots\}$ the following statements are equivalent:

* **X is geometrically distributed with parameter p.**
* $\boldsymbol{P(X > n) = (1-p)^n, for\ n = 0, 1, 2, \ldots}$.
* **The distribution of X has the lack of memory property on S_X and** $\boldsymbol{p = P(X = 1)}$

Summary 3: Research Methods

Chapter 1: A Framework for The Best Solution

MPSM: Managerial Problem-Solving Method. The MPSM allows you to embark on necessary creative jaunts within the systematic approach to a problem.
The MPSM is one of several systematic problem-solving approaches. The MPSM consists of the following **seven phases**:

1. Defining the problem.
2. Formulating the approach.
3. Analysing the problem.
4. Formulating (alternative) solutions.
5. Choosing a solution.
6. Implementing the solution.
7. Evaluating the solution.

The MPSM takes into account that no problem is an isolated issue. A problem is embedded in the context of an organisation, and needs a fitting solution.
When using the MPSM, you express a problem in terms of variables.

* **Language of Variables:** Expressing problems in terms of variables. You untangle an issue and define it as a series of variables, measurable attributes. Some variables can only be measured using indicators.

Chapter 2 Untroubled Problems

Discrepancy: a deviation, a situation where two things do not match. Problem-solving involves dealing with a discrepancy, a deviation, between the norm and reality.
Action problem: a discrepancy between the norm and reality, as perceived by the problem owner. An action problem can be resolved by using the MPSM.
Problem-owner: the person(s), group(s), or organisation(s) who are (or perceive themselves to be) responsible for a solution to a problem.
Research population: The person(s), group(s), or organisation(s) that are the subject of the investigation you are performing.
Knowledge problem: a description of the research population, the variables and, if necessary, the relationships that need to be investigated.
A knowledge problem is resolved using the **research cycle**. It consists of eight phases:

1. Formulating the research goal.
2. Formulating the problem statement.
3. Formulating the research questions.
4. Formulating the research design.
5. Performing the operationalisation.
6. Performing the measurements (gathering date).
7. Processing data.
8. Drawing conclusions (reviewing the problem statement).

Description	Action problem	Knowledge problem
Appearance	Thing are not as you want them to be.	You require knowledge.
Formulation	Problem identification.	Problem statement.
Problem description	Perception	Question
Norm	Specific	General

Problem-solving approach	MPSM	Research cycle.

When dealing with an action problem, you always need a certain amount of knowledge. As soon as you find that there is information you need and do not already have, you move from the MPSM, and enter the research cycle. Once you have completed the cycle, you re-enter the MPSM at the phase from which you left off.

Chapter 3 A Trouble-Shooter's Log

MPSM phase	**Activities**	**Results**
Phase 1: **Defining the problem**	Drafting inventory of problems. Make problem cluster. Select core problem. Express problem in variables.	Global problem identification.
Phase 2: **Formulating the problem-solving approach**	Draft problem-solving approach. Use D3 to describe the activities and knowledge required, and choose from the various options available within your approach.	Plan for the problem-solving process.
Phase 3: **Analysing the problem**	Re-examine problem identification and problem cluster, and fill in missing details. Look for causes. Review why any earlier solutions have not worked. Use research cycle to solve knowledge problems. Use a model to document relationships between problem and cause.	Definition of analysis both problem and problem identification.
Phase 4: **Formulating (alternative) solutions**	Describe a solution. Establish decision-making process. Draft, scale and weight list of criteria. Invent alternatives or use existing possibilities. And evaluate alternatives.	Report on alternatives and their desirability.
Phase 5: **Choosing a solution**	Select one of the possible solutions.	Solution to the problem.
Phase 6: **Implementing the solution**	Draft an implementation plan. Provide brief and step-by-step description of activities. Detail approach to possible resistance.	Change
Phase 7: **Evaluating the solution**	Convince participants of the need for evaluation. Evaluate all phases of the MPSM. Preform structured evaluation.	Comparison of effected situation to desired situation.

Chapter 4 In Search of the Core Problem

Problem context: the full scope of existing problems, which may or may not be related.

When identifying a problem in an organization, you proceed step by step:

* You make an inventory of existing problems. Your work should be quick and dirty. You do not perform any real research. You inventory existing data and perspectives. You use different sources for your inventory, such as available documentation and information from interviews.

Quick and dirty (identification): a rapid, uncomplicated inventory of existing inventory of existing data and perspectives made without performing actual research.

* Next, you look for causes and effects. You draft a problem cluster in which you map the different problems and their mutual relationships. Doing so structures the problem context, allowing you to identify the core problem. You use the following rules of thumb for this process:
    - You only include a problem in the cluster if you are sufficiently sure it is an actual problem. If you are not, you do not add it to the cluster.
    - Proceed down the chain of causes and effect until you reach the cause which is furthest removed from your initial problem. That cause is no longer effect of another cause.
    - If you cannot influence it, it cannot be labelled a core problem.

Problem identification: systematically establishing a quantifiable core problem.
Problem cluster: a model used to map different problems and their mutual relationships. A problem cluster serves as a mean of structuring the problem context, and is used to identify the core problem.

* Then, you choose which core problem to deal with. That will be the most important problem in the problem cluster. At this time, you can choose to first perform a global cost-benefit analysis for the approaches to the different core problems. If you do, pick the most appropriate problem – lowest costs, highest reward.
* The last step is to make the problem manageable. You express the norm and the reality of your problem in a concrete, measurable variable. You may need to use indicators in order to actually measure anything. When using indicators, consider the following tips:
    - Use more than one, but no too many, indicators, preferably three of four.
    - Consider all circumstances when choosing you indicators.
    - Preferably use only direct consequences when using indicators to measure effects.
    - Never use a cause as an indicator.

Concretising variables: a method of quantifying variables. Using indicators, you highlight one of more aspects of your variable. You may choose to use more than one indicator for the same variable. Together, these indicators may or may not exactly encompass the variable.
Decomposing variables: a method of quantifying variables. Variables can sometimes be split up into separate parts. Combined, these parts add up to your variable exactly. The variable for profit, for example, is made up of the indicators of revenue and costs.

Chapter 5 D3: Do, Discover, Decide

When formulating your problem-solving approach, you describe your planned actions as precisely as possible. You review all actions needed to successfully resolve your approach. You need **D3 (Do, Discover, Decide). Do** is everything you need to do for your approach to work. **Discover** is everything the parties solving the problem need to know. **Decide** is selecting the appropriate options from the multitude of aspects of an often complex problem.

The next step is systematically mapping your problem. This starts with asking the questions you need answered, including at least the following:

* Which parties have to deal with the problem? Who are the owners and victims of the problem?
* Which parties are in charge of the required resources?
* What information do you require?
* Whose help will you need during which phases of your approach? Who are you helpers?
* What is the division of labour in you project team?
* What are the possible of existing limitations?

Once you have the answers to your questions, you can begin formulating your approach. This requires precision. When formulating an approach, consider the following pointers:

* Be accurate in describing your activities. You should be able to rely on your descriptions later.
* Be selective! Only write down what you need for your problem-solving approach. Do not overstate or embellish matters.
* Demonstrate that you are able to solve the problem.

Research cycle: a systematic approach to research performed in order to obtain required information. Allows the researcher to arrive at a conclusion through a stepwise process, thereby answering a previously formulated problem definition.

Chapter 6 In Search of the Unknown

During the problem analysis phase, you conduct research because you need more information. You review your problem identification and problem cluster. You fill in the missing details. You look for as yet unknown explanations for your problem. You re-examine previously suggested solutions.
Problem analysis: researching a problem to find out its causes, and establishing interrelationships between causes and effects.

You use the research cycle to solve your knowledge-based problem. This takes you to the answer to your problem definition step by step.

Measurement levels: levels at which variables can be measured. There are four:

* Nominal variables: values are named (not quantitatively related).
* Ordinal variables: values are ordered (the exact value is unknow).
* Interval variables: distance between variables are known.
* Ratio level variables: ratio’s between variables are known.

A relationship between problems can be defined as a **causality**, a cause-and-effect type relationship, if it meets the following three criteria:

1. There is a **sequential order:** a problem is always preceded by a cause.
2. There is a **statistical order:** a numerical relationship between a cause and a problem. Involves variables whose values change together.
3. There are **no alternative explanations** for this part of the problem.

During your research, you keep the following in mind:

* Make sure you quantify your variables using indicators.
* Formulate your variables at the highest possible measurement level.
* Limit your research to important causes that will contribute to the solution.
* Look for causal relations and not just statistical ones.

Chapter 7 An Overview of Options

Seven steps to a possible solution:

1. Defining the decision: you decide what exactly needs to be decided on.
2. Defining the decision-making criteria:
    - Establish who the decision-makers are.
    - Determine how the decision is constructed.
    - Agree on rules for the decision-making process.
    - To weight or not to weight. Weight is what indicates how important different criteria are. The higher the weight, the more important the criterion.
    - Be mindful of subjective elements. Subjective elements are issues which, considering the goal of the decision, should not actually be of influence to the decision.
    - Ensure the decision-making process is clearly established.
3. Establishing criteria: ask the client what they require of a solution.
4. Scaling criteria: you assign points on a scale to the criteria, for example between 1 and 5.
5. Weighting criteria: you weight the criteria. There are several methods for weighting criteria which can involve users of the final solution of product.
6. Coming up with optional solutions or using existing possibilities: you come up with options or use existing possibilities. First, think of as many possible solutions as you can. Then, involve others in your search for alternative options, for example by conducting interviews or using creative techniques, such as brainstorming, synectics, or morphology. Be sure to use information or solutions other researchers have already come up with, if possible.
7. Evaluating options: you indicate which solution is best suited to solving the problem.

Attributes: characteristics of solutions.
Conjunction criteria: demands that a potential solution has to meet by definition. Scoring too high or too low in these criteria cannot be redeemed or compensated by scoring higher elsewhere. Conjunction criteria are also known as **cut-off criteria, non-compensatory criteria of killer requirements**.
Consensus: a state where all parties participating in a meeting are in agreement.
Qualified majority: an agreement which sets additional demands with regards to decision-making.
Lexicographical method: a method of deciding using a set of ranked criteria. The options are first compared based on the most important criterion. Whichever option scores the highest, is selected. In case of a tie, the next-most important criterion is used until there is a winner.
Majority rule: a rule that favours the option which scores best in the largest number of criteria. Here, all criteria have equal importance, and non are weighted over others.
Morphology: a creative technique used to come up with solutions.
MoSCoW rules: a method of prioritising requirements and demands that are part of a solution. the consonants in Moscow represent *must have, should have, could have, want to have but will not have this time round.*
Sub-attributes: part of a criterion, a characteristic of an option. Splitting a criterion or attribute into component parts yields sub-attributes.
Synectics: a creative technique for coming up with solutions based on looking at a problem in a different realm, time or place.

Chapter 8 The Clients Move

The final decision is the client's prerogative. Bases on other considerations, a client may deviate from professional and founded advice. As such, advisors and decision-makers can arrive at different choices. These may often come across as politically charged in the eyes of the advisors. That need not be the case – it need not even be a bad thing.

The trouble-shooter compares the options and assesses what the best possibility would be. The decision-maker reviews the submitted selection, but their decision is an independent one. The decision-maker places the choice for a particular option in a boarder of context.

When assessing alternatives, an advisor reviews scores, weights, and prerequisites. The decision-maker may deal differently with the weighting, for example by choosing to apply the majority rule. This leads to whichever option conforms to the largest number of attributes.
Decision-makers explicitly take all risks into account in their deliberations. They view the risks in a boarder aspect than the advisor. This may lead them to make different choices.

Threshold value: minimum amount (of products needed for profitable production).
Probability theory: a method used to calculate the risks involved in an option or event. Risk is defined as *probability times consequences*. The probability is the average number of times a certain thing occurs, usually expressed as a percentage. The consequence is the loss of value incurred when the risk occurs. Based on the *probability times consequences* formula, risks are assessed as acceptable of inacceptable.
Scenario: a construction of possible future state of an aspect of the world. Scenario planning is a way of using various scenarios as the basis for a decision, since it provides an insight into how certain options would work in practice.

Chapter 9 A Systematic Approach

Your **implementation plan** is concerned with precisely detailing the execution of the solution, as well as with the necessary steps to ensure everyone is on board.
Implementation plan: a plan of attack which details what need to happen in order to implement adopted advice in an organisation.

Bottom-up: a direction of communications in an organisation. Bottom-up means workplace input is possible.
BCH model: a model used for an internal analysis of an organisation. The letters represent Balance, Cohesion and Heterogeneity.
Top-down: a direction of communication in an organisation. Top-down means all communications originate from management and move along any existing intermediate layer until they arrive at the workplace.
Change management: supervise, organise, and (have others) execute change within an organisation.
Yammer: yammer is like twitter for companies. One of various social media platforms which work by submitting short messages.

Chapter 10 Keep Improving

* The evaluation of an implemented solution should be handled in a structured manner.
* People are often not keen on an evaluation. Convince them of its usefulness.
* Use your findings from the different phase of the MPSM if the desired effects of a change are not encountered strongly enough. You may find the causes in the analysis, or may be able to use solutions you had discounted earlier.
* Handle the evaluation professionally. Among others, this involves: establishing the goal of the evaluation, making an inventory of people experiencing the problem, formulating your goals as to be able to evaluate them after, determine the indicators of the problem, assembling the evaluation team and, above all, involving others in the evaluation process.

Formative evaluation: a way of evaluating whereby you review the causes of the effects encountered. Reviewing whether the implementation has gone according to plan.
Summative evaluation: a way of evaluating which focusses on establishing whether your goals have been met. You measure whether the size of severity of the problem has decreased without looking at its causes.

Chapter 11 In Search of Knowledge

The **research cycle** consists of the following eight phases:

1. Formulating the research goal: state the reason for wanting to solve the knowledge problem.
2. Formulating the problem statement: phrase it in the form of a question. Consider the type of research involved: descriptive or explanatory. Define your variables and indicate the relationships to be researched. Establish your research population. Draft a model. Use a theoretical perspective if necessary.
3. Formulating the research questions: these are questions used to split up a problem statement. Formulate the research questions systematically.
4. Formulating the research design: the intended strategy of your research. Answer the following questions:
    - Do you want to give your subjects stimulus?
    - Do you want or need to be in contact with your research population?
    - Do you conduct your research in laboratory conditions?
    - Do you want your research to be broad or in-depth?
    - Do you want to conduct cross-sectional of longitudinal research?
5. Performing the operationalisation: operationalise your variables. You can choose to concretise: use indicators to highlight several aspects of your variable which need not necessarily amount to the variable exactly. Or you can choose to decompose: split up your variables into separate components which, when combined, add up to your variable exactly.
6. Performing the measurements (gathering date).
7. Processing data: indicate how you processed the data. Clarify how you gathered your date. Choose as uncomplicated a method of analysis as possible. Indicate any potential influences on their validity and reliability.
8. Drawing conclusions (reviewing the problem statement): end by drawing conclusions and revisiting the problem statement. Consider whether you can fully answer the problem statement. If so, use your newfound knowledge to solve your action problem. If your conclusions can only apply to certain aspects or situations, you will need to conduct additional research.

Cross-sectional research: a method of research which involves performing all measurements virtually simultaneously.
Longitudinal research: a method of research which involves performing measurements at different points in time. These measurements are then compared for differences.

Operationalisation: using indicators to make variables quantifiable.
Validity: refers to the quality of research results. Are you measuring what you had intended to measure? There are several types of validity, such as **internal validity** (concerned with properly formulating and constructing research designs and measuring instruments), **external validity** (concerned with the extent to which the research can be applies to groups other that the research population), and **construct validity** (concerned with whether abstract concepts are properly operationalised, logically coherent and, where possible, based on available scientific knowledge).

Summary 4: Excel & VBA Manual Summary

Excel tricks

- CTRL+ARROW DOWN (to jump to the last cell in the column)
- CTRL+ARROW RIGHT (to jump to the last cell in the row)
- CTRL+HOME (to jump back to cell A1)
- CTRL+SHIFT+ARROW DOWN (to select the entire column)
- CTRL+SHIFT+ARROW RIGHT (to select all data cells)
- Hold CTRL while clicking on column A and then column D. Now you have selected two noncontiguous columns. Of course this also works on rows
- Click on column B and Hold SHIFT while clicking on column E. Now you have selected all (contiguous) columns from B to E.
- You can give a certain range a name in the space above A1
- =FORMULA(f3 to see all the known options/ranges
- F4: to freeze cells
- Alt + F1: make a chart
- F2 = change something in that specific cell
- CTRL shift 4 = currency format - F9: refreshes entire workbook
- Shift F9: refreshes current worksheet
- CTRL + 1 = format cells
- Format > custom > 0000 = always fill the number up with zeros up to 4 digits
- To start every number with 00 → 00### or with a letter or in parentheses or in a color → [blue]
- CTRL + T: make a table o To sort without sorting top row: sort - custom sort → my data has readers
- Paste special: first copy, then paste special → only values, to only keep the values without formula; transpose, to paste it horizontally/vertically
- Pivot table: pivot table design: report layout: in tabular form, so it doesn't say column and row but the right names o Right click in value → value field settings for formatting
    - You can also right click and then number format, **not** format cells, because only number format will stay when the pivot table changes

  o You can drag the same variable into the values box multiple times and by using value field settings you can change the type of calculation

    - You can also right click and then summarize values by … o Stddevp = standard deviation for population → if you have the data of the whole population

  o DON'T FORGET TO REFRESH THE PIVOT TABLE IF YOU CHANGE SOMETHING IN THE DATA → analyse → refresh
- Vertical lookup: VLOOKUP(the thing you want to know a value from; range where you want to look in; number of column you want to find the value in (column 1 is the for the thing you want to know the value from); 0 (false = exact match) or 1 (true = approximate match)) o Don't forget to freeze the range when you copy/pull down the formula o Default is approximate match

- XLOOKUP = simplified version of VLOOKUP where the return column or row may be on either side of the lookup-column and you don't need to count the columns, you only say in which column you want to look up and in which column you want to find a value o Default is exact match
- HLOOKUP = horizontal look up, it will return a value from a different row in the same column
- LOOKUP = combination of H and VLOOKUP → older version
- FIND = to find a position of a certain character within a text → FIND("a";A2) with A2 = apple will return 1 o FIND("";A2) will return the number of the first character in cell A2 o FIND(" ";A2) will return the position of the first space in cell A2
    - o SEARCH is the same but not case sensitive
- LEFT returns the first character(s) from the left, given the number of characters you want
- RIGHT same but from the right - LEN returns how long a text is - How to split names:
    - o To find the first name: =LEFT(cell, FIND(" ", cell)-1) → you're finding the space between first and last name and then take that position minus 1 to get the characters from the left up till the position of the last letter of the first name
    - o To find the last name: =RIGHT(cell, LEN(cell)-FIND(" ";cell))
- AND to say what return value you want when 2 statements are true
- OR ^ when either statement is true
- NOT reverses FALSE into TRUE and vice versa. E.g. NOT(2>1) gives FALSE
- In a formula, when you want to say <10, you have to put "<10"
- Formulas → formula auditing → trace dependents, shows what cells are used in the formula - Formulas → formula auditing → trace precedents, shows what cells are using this cell - Data → forecast → datatable:
    - o 1 variable: choose the variable you want to test with different values (e.g. weight) and choose the values (e.g. 60-90), choose the formula which you want to use to test the values on (e.g. BMI score) and put a reference to this formula in the column next to the values. Then datatable → column input is the cell with the formula (freeze it)
    - o 2 variables: put the reference to a formula (e.g. BMI score) in a cell and put values for the first variable (e.g. weight) in that column and values for the second variable (e.g. length) in that row. The row input is the original value for the second variable and the column input is the original value for the first variable.
- Data → forecast → goal seek: to calculate what value a certain variable should have in order to retrieve a certain value from a formula (e.g. what weight you should have, to have a BMI score of 25)
- & to put values of 2 cells in 1 cell
- CONCATENATE: same thing as & but with multiple cells
- INDEX: to find a value in an array o =INDEX(Array; row number; column number) → will return the value of that cell
- OFFSET: to move the cursor a number of rows and/or columns o E.g.: =average(A4:offset(A4;5;0)) → will return the average of A4-A9
- Data → data tools → data validation → list → select a range, then use vlookup to get the value you want o Select a column → name it, then in data validation when you choose list, you can do f3 and choose the column you named

- Select multiple columns and name it so in vlookup you can choose the range you already named using f3

VBA manual

Always make codes in a module!

- Tools (Extra in Dutch) row→ Options → Editor → Require Variable Declaration: to put Option Explicit at the top so VBA cannot declare variables automatically, which avoids errors
- Record macros by pressing record and then do what you want the macro to do and then stop recording → insert a button and select the cells you want to apply the macro to and click the button
- Run a code line by line by pressing f8
- Create a function by inserting a module → "Function *name function* ()" → put the command between function and end function o E.g. function with variables:

  Function AddTwo (Arg1, Arg2)

  AddTwo = Arg1 + Arg2

  End Function → then if you do =addtwo in the excel sheet and add 2 values they will be added
- Subs (=macro) are used to execute several actions and not only return a value
- You can denote a range in the following ways: "Range("A1"), Range("B3:z125"), Range("*name range*"), Range(Cells(1,1), Cells(23,54)) → Cells(Row number, column number) o E.g.:

  To change the content of a cell into text:

  Sub (name) Range("A1") =

  "text"
- Subs will run in the active worksheet. To run in specific worksheet, put 'worksheets("Sheet1".cells()

`Range("B1").OffSet(1,2)`	gives a new range: `Range("D2")`. The 1,2 is the shift by Row (+1) and Column (+2) of the indicated address B1
`Range("B1").Value`	contains the cell value
`Range("B1").Text`	contains the formatted text as in the cell
`Range("B1:B10").Count`	is the number of cells
`Range("B1").Columns.Count` or `.Rows.Count`	the number of columns or rows in the range
`Range("B1").Column` or `.Row`	column or row number
`Range("B1").Address`	text with range address, here: B1
`Range("B1").HasFormula`	true/false
`Range("B1").Formula`	text with cell formula, f.e. "=SUM(A1:C1)"
`Range("A1:E10").Clear`	clears all the data *and formatting* of cells in the range
`Range("A:A").ClearContents`	clears all the data of cells in the range

- To create a variable: "dim *name* as …." (see data types next page) If you don't declare what data type you are using, VBA will declare it automatically as a Variant. It is better to declare it yourself and therefore turn on the option "Require Variable Declaration", so that VBA will never do anything automatically and therefore you avoid making mistakes (VBA tools → options → editor/Dutch: VBA extra → opties → variabelen declareren vereist)

- If you define a variable *within* a sub, you can only use it within this sub, and it is not known outside of this sub. Its value is removed from the computer's memory when the sub ends. To maintain the value, type "Static" instead of Dim
- Global variable: a variable created outside of a sub → so you can use a variable outside a sub
- Local variable: a variable created inside a sub
- Public (instead of Dim): to use the variable also in other modules

Table 7-1 VBA's Built-In Data Types

Data Type	*Bytes Used*	*Range of Values*
Boolean	2	True or False
Integer	2	–32,768 to 32,767
Long	4	–2,147,483,648 to 2,147,483,647
Single	4	–3.402823E38 to 1.401298E45
Double (negative)	8	–1.79769313486232E308 to –4.94065645841247E-324
Double (positive)	8	4.94065645841247E–324 to 1.79769313486232E308
Currency	8	–922,337,203,685,477.5808 to 922,337,203,685,477.5807
Date	8	1/1/100 to 12/31/9999
String	1 per char	Varies
Object	4	Any defined object
Variant	Varies	Any data type
User defined	Varies	Varies

- Variables can be passed to subs and functions as parameters; as input or as output. As input, their values are used within the sub/function, as output, they may receive new values by the sub/function.
- ByVal = input variable → the value is not changed but used as input
- ByRef = output variable → vba takes the value of this variable and uses it to create a new value
- Optional variable: you can say what the default value is and if you want you change that value when calling the sub
- 

  Here you can customize what buttons you have
- Step into, step over, step out - CTRL+space = fills in a word for you - How to make a constant:

  e.g.:

  Const ActLevel_Sedentary = 1.2

  Const ActLevel_LightlyActive = 1.373
- Foutopsporing → VBAproject compileren: to check if there are bugs
- If … mod 2 = 1 → if you divide something by 2 and there is 1 left (e.g. 21/2 = 10 + 1)

1. Running a sub

- A button cannot be used to call a function or a sub when the sub contains parameters
- "Call": to call a sub
- Const: to create a "variable" with a constant value
- MsgBox: to show a message box
  - E.g. MsgBox("Product is " & Cells(1, 1) * Cells(1, 2))
  - A message box is defined as MsgBox(the message [, type of buttons/icon style/default button] [, title])
- ISEMPTY = to check if a cell is empty
- Isnumeric(arg1) → to check if arg1 is a number
- Int(arg1) → to check if arg1 is integer
- Select Case: same as If…Then…Else… statement, but with multiple Else and easier o

Select Case grade (this should refer to a variable) Case 1 (or e.g. 1 to 3) grade = "good"
Case 2

grade = "satisfactory" etc.

End Select

- InputBox("Enter a number"): to make a user enter a number
- <>: not equal to
- End an if…then statement with end if

Group	Constant	Value	Description
1	vbOKOnly	0	Display OK button only
1	vbOKCancel	1	Display OK and Cancel buttons
1	vbAbortRetryIgnore	2	Display Abort, Retry, and Ignore buttons
1	vbYesNoCancel	3	Display Yes, No, and Cancel buttons
1	vbYesNo	4	Display Yes and No buttons
1	vbRetryCancel	5	Display Retry and Cancel buttons
2	vbCritical	16	Display Critical Message icon
2	vbQuestion	32	Display Warning Query icon
2	vbExclamation	48	Display Warning Message icon
2	vbInformation	64	Display Information Message icon
3	vbDefaultButton1	0	First button is default
3	vbDefaultButton2	256	Second button is default
3	vbDefaultButton3	512	Third button is default
3	vbDefaultButton4	768	Fourth button is default
4	vbApplicationModal	0	Application modal; the user must respond to the message box before continuing work in the current application
4	vbSystemModal	4096	System modal; all applications are suspended until the user responds to the message box
4	vbMsgBoxHelpButton	16384	Adds Help button to the message box
4	VbMsgBoxSetForeground	65536	Specifies the message box window as the foreground window
4	vbMsgBoxRight	524288	Text is right-aligned
4	vbMsgBoxRtlReading	1048576	Specifies text should appear as right-to-left reading on Hebrew and Arabic systems

2. Loops

- For…Next loop: e.g.:

o Dim i as long

Dim total as long

For i = 1 to 10

Total = total + 10

Next i o Adds 10 to the variable total
10 times

- For i = 10 to 1 step -1 → to loop backwards
- Within a for…loop you can put: "if …. Then exit for" to end the For loop at a specific value
- MsgBox ("i stopped in loop: " & i) after a for next loop to state how many times the function looped
- Nested loops: if you put a loop in a loop
  o E.g.:

```
Dim x As Long, y As Long 'two loop variables
For x = 1 To 10
For y = 1 To 10
Cells(y, x) = 1000
Next y
Next x
```

- Do loop: when you don't want to repeat your statements a fixed number of times o Do Until: repeats loop until the condition is true o Do while: repeats loop until condition is false o E.g.:

```
X = 0
Do While x < 2
x = x + 1
Loop o To exit Do loop, write
Exit Do
```

- For each loop: another way to loop through a range o E.g.:

```
Dim c as Range
For each c in Range("K1:K10")
    c.value = "XYZ" next c
```

- To make a variable empty after one loop, put variable = 0 at the beginning of the loop

- ARange.Rows.Count: to find the amount of rows in a range
- Or: Range("A1:A" & rows.count)

3. Arrays

- Static array:

E.g.:

```
Dim Students(1 to 5) as string
Students(1) = "Name"
Students(2) = "Name"
Etc.
MsgBox Students(4)
```

- Or: Dim AnArray (10) as long

The array is now from 0 to 10

If you want 1 to 10, you put this at the top:

Option base 1 -

Dynamic array:

Dim AnArray() as long

- To resize an already existing array (1 to 5) **and** keep the data → redim preserve anarray(1 to 6)
- Arrays with more dimensions:

E.g.:

Dim sudoku(1 to 9, 1 to 9) as long

- LBound and UBound to find the lower and upperbound of (dimensions of) an array

E.g.:

```
Dim Matrix(1 to 10, 1 To 50) As Long
Dim i As Long, j As Long
    For i = LBound(Matrix, 1) To UBound(Matrix, 1) → 1 = the first dimension of the array
        For j = LBound(Matrix, 2) To UBound(Matrix, 2)
            Matrix(i,j) = 15 'some code goes here
        Next j
    Next i
```

- Bubblesort: to sort values from e.g. small to large e.g.

```
Sub BubbleSort(AnArray())
Dim FirstItem As Long, LastItem As Long, i As Long, j As Long, Tmp As Long
FirstItem = LBound(AnArray)
LastItem = UBound(AnArray)
For i = FirstItem To LastItem - 1
    For j = i + 1 To LastItem
        If AnArray(i) > AnArray(j) Then
            Tmp = AnArray(j) AnArray(j) = AnArray(i) AnArray(i)
            = Tmp
        End If
    Next j
  Next i
End Sub
```

- Erase [DynamicArray] → removes all elements of the array - Erase [StaticArray] → sets values of all elements to 0

4. Random tricks

- Range(…).clearcontents → to clear all cells
- Calculating random non-integer numbers between 0 and 1: RND (like the excel function rand())
- Calculating random non-integer numbers between a and b: RND*(a-b)+a

Examples

1. Find the 2 largest numbers

```
Dim LargestNumber As Double, SecondLargestNumber As Double
  Dim r As Long, c As Long
  LargestNumber = ARange.Cells(1, 1)
  If ARange.Cells(1, 2) > LargestNumber Then
    SecondLargestNumber = LargestNumber
```

```
      LargestNumber = ARange.Cells(1, 2)
    Else
      SecondLargestNumber = ARange.Cells(1, 2)
    End If

    For r = 1 To ARange.Rows.Count
      For c = 1 To ARange.Columns.Count
        If r > 1 Or c > 2 Then
          If ARange.Cells(r, c) > LargestNumber Then
              ARange.Cells(r, c).Interior.Color = 65535
              SecondLargestNumber = LargestNumber
              LargestNumber = ARange.Cells(r, c)
          ElseIf ARange.Cells(r, c) > SecondLargestNumber Then
            ARange.Cells(r, c).Interior.Color = 65535
            SecondLargestNumber = ARange.Cells(r, c)
          End If
        End If

      Next c
    Next r
```

2. Check if an argument is a number and integer within a function

```
  Function Faculty(arg1 As Long)
  Dim i As Long
  Faculty = 1
  If IsNumeric(arg1) Then
    If Int(arg1) Then
      For i = arg1 To 2 Step -1
        Faculty = Faculty * i
      Next i
    Else
    Faculty = "This number is not integer"
    Exit Function
    End If
  Else
  Faculty = "This is not a number"
  Exit Function
  End If

  End Function
```

Module 2: Operations management

! Disclaimer: always check what you need to study corresponds with the content of the summaries, courses can be changed which could cause changes in study material for your exams

This module consists out of four courses, and a four part project of which the last one is based on Diversey Enschede. Below you find information about which courses you have this module, and about the summaries for this module. If you made a summary for a course this module you can sent them to education@stress.utwente.nl and depending on how many summaries we have for this course you will receive compensation for your work.

Courses

- Math B2
- Operations Strategy
- Operations Research
- Professional Skills
- Project

Summary 1

Course: Operations Strategy 1*
Book: Slack, N., Chambers. S. & Johnston, R. (2016) *Operations Management,* Pearson Education
Chapters: 1 - 9
Year the summary was received: 2019/2020

Summary 2

Course: Operations Strategy 2*
Book: Slack, N., Chambers. S. & Johnston, R. (2016) *Operations Management,* Pearson Education
Chapters: 10, 11, 14, 15, 16, 17, 18
Year the summary was received: 2019/2020

Summary 3

Course: Operations Research 1
Book: Winston, W. (2003). Operations Research: Applications and Algorithms (4th ed.). Boston, United States of America: Duxbury Press
Chapters: 1, 3 , 9.1/9.2 and 18.1 – 18.6
Year the summary was received: 2018/2019

Summary 4

Course: Operations Research 2
Book: Winston, W. (2003). Operations Research: Applications and Algorithms (4th ed.). Boston, United States of America: Duxbury Press
Chapters: 15, 16.1/2/3/6, 18.7, 20.1/2/4/5/6/9
Year the summary was received: 2018/2019

* There is another summary available on www.stress.utwente.nl

Summary 1: Operations Strategy 1

Chapter 1

Operations management (OM): activity of managing the resources that create and deliver services/products. All organizations have 'operations' that produce some mix of services/products, operations manager responsible. Three core functions in any organization:

- **Marketing** (incl sales) -> communicating organization's services and products to its markets to generate customer requests
- **Product-service development** -> coming up with new and modified services/products to generate future customer requests
- **Operations** -> creation and delivery of services and products based on customer requests

Always some additional support functions, can depend on type of organization. In this book: **operations function**: all activities necessary for day-to-day fulfilment of customer requests within constraints of environmental and social sustainability.

Fundamental that functional boundaries should not hinder efficient internal processes. All organizations have 1 thing in common: make a profit from creating and delivering their products or services. Non-profit -> to serve society in some way. OM uses 'resources to appropriately create outputs that fulfil defined market requirements'. OM is at the forefront of coping with, and exploiting, developments in business and technology.

All operations create and deliver service/products by changing inputs into outputs using an '**input-transformation-output**' process. One set to any operations processes is transformed resources: treated, transformed or converted in the process. Usually mix of: **materials** (transform physical properties, location, possession, store), **information** (transform informational properties, possession, store, location) and **customers** (change physical properties, accommodate, location, physiological state). When customers play role in operations: 'co-production'.

Other input: transforming resources: **facilities** (building, equipment etc) and **staff**.

Facilitating services: e.g. aluminum smelters also delivering technical advice. Most operations produce a mixture of tangible products and intangible services. Whether an operation produces tangible products or intangible services is becoming increasingly irrelevant. In a sense all operations produce service for their customers.

It is critical that operations managers are aware of customers' needs, both current and potential.

All operations consist of a collection of processes interconnecting with each other. **Process**: arrangement of resources and activities that transform inputs into outputs that satisfy customer needs. Each process is an internal supplier and an internal customer for other processes. Any process is made up of a network of resources. Business or operations itself is also part of a greater process: supply network. Process principle can be used at three levels 'levels of analysis': level of operations itself, level of supply network and level of individual processes.

Not just the operations function that manages processes; each function will have its 'technical' knowledge. OM is relevant for all functions, and all managers should have something to learn from the principles, concepts, approaches and techniques of OM -> distinguish between: 1. 'Operations' as a function: part of the organization which creates and delivers services and products for the organizations external customers. 2. 'Operations' as an activity: management of the processes within any of the organization's functions.

Processes are defined by how the organization chooses to draw process boundaries. Operations processes differ in 4 V's:

- **Volume** of their output

High (e.g. McDonalds): high repeatability, specialization, capital intensive, low unit costs. Low (e.g. local cafeteria): low repetition, each staff member performs more of each task (might be more rewarding), less systemization, high unit costs, less feasible to invest in specialized equipment.

- **Variety** of their output

High (e.g. taxi): flexible, complex, match customer needs, high unit costs. Low (e.g. bus service): well defined, routine, standardized, regular, low unit costs.

- **Variation** in the demand for their output

High (e.g. summer holiday resort hotel): changing capacity, anticipation, flexibility, in touch with demand, high unit costs. Low (e.g. formule 1 hotel): stable, routine, predictable, high utilization, low unit costs.

- Degree of **visibility** which customers have of the creation of their output

High (e.g. 'bricks and mortar'): short waiting tolerance, satisfaction governed by customer perception, customer contact skills needed, received variety is high, high unit costs. Low (e.g. web-based): time lag between production and consumption, standardization, low contact skills, high staff utilization, centralization, low unit costs.

Some mixed high- and low-visibility: e.g. airports.

Operations and processes can reduce their costs by increasing volume, reducing variety, reducing variation and reducing visibility.

OM activities:

- **Directing** overall strategy of the operation. General understanding of operations and processes and their strategic purpose and performance, together with appreciation of how strategic purpose is translated into reality.
- **Designing** the operation's resources and processes. Determining the physical form, shape and composition of operations and processes in line with services/products that they create.
- Planning and control process **delivery**. After being designed, the delivery of services/products from suppliers and through the total operation to customers must be planned and controlled.
- **Developing** capabilities of their processes to improve process performance.

Environmental sustainability: extent to which business activity negatively impacts the natural environment. Decisions -> affect utilization of materials both in short and long term recyclability. Design -> proportion energy, labor, materials wasted. Planning and control -> energy, labor, materials wastage. Reducing waste also reduces costs.

Chapter 2

Operations management being able to make or break any business. Understand importance of operations management -> first understand why things can go wrong in operations and their impact. 'Make' the organization in several ways: 1. OM concerned with doing things better, improvement can potentially make operations the driver of improvement for the whole organization. 2. Through continual learning that can come from its improvement activities, OM can build 'difficult to imitate' capabilities -> strategic impact. 3. OM concerned with process, good OM -> best way to produce good services/products.

Operations managers make decisions on economic, social, political and technological environment changes.

How operations can judge its performance at 3 levels:

- Broad, societal level, 'triple bottom line'
- Strategic level of how an operation can contribute to the organization's overall strategy
- Operational level, 5 'performance objectives'

Societal level – Decisions will affect whole variety of 'stakeholders'. Stakeholders: people and groups who have a legitimate interest in the operation's activities, internal or external. Responsibility of the operations function to understand the objectives of its stakeholders. This idea often termed 'corporate social responsibility (CSR)'. Of increasing importance, both ethical and commercial point of view.

Triple bottom line; 'people, plant and profit': organizations should measure not only on traditional economic profit, also on impact their operations have on society and ecological impact on environment -> sustainability. Balance economic, environmental and societal interests. Social bottom line performance: businesses should accept that they bear some responsibility for impact they have on society and balance the external consequences with more direct internal.

Environmental bottom line (=planet): sustainability: extent to which business activity negatively impacts the natural environment.

Economic bottom line (=profit): operations managers must use operation's resources effectively.

How is operations performance judged at a strategic level? 5 aspects are contributing to the 'economic' aspect of the triple bottom line. How operations affect profit.

- OM affects **costs** -> most important aspect how they judge their performance. Efficiency with which operation purchases its transformed, transforming resources , converts those resources will determine costs.
- OM affects **revenue** -> judges on organization's profitability.
- OM affects the **required level of investment** -> How the transforming resources that are necessary to produce the required type and quantity of its products and services.
- OM affects the **risk of operational failure** -> reduce risks and recover quickly with less disruption (resilience)
- OM affects **ability to build the capabilities on which future innovation** is based -> learn from experience, improve over time.

Operations performance objectives:

Quality – consistent conformance to customers' expectations. Major influence on customer satisfaction. Makes internal easier as well -> quality reduces costs (fewer mistakes -> less time needed to correct, less confusion). Quality increases dependability (can give potential for better services and products.

Speed – elapsed time between customers requesting products/services and their receiving them. The faster -> more likely to buy it, more pay, greater benefit they receive. Speed reduces inventories. Speed reduces risks (forecasting). Speed can give the potential for faster delivery of services and products, and save costs.

Dependability – doing things in time for customers to receive products or services exactly when they are needed/promised. Dependability saves time, saves money (ineffective use of time), gives stability ('quality' of time and will be trust in operation). Dependability can give the potential for more reliable delivery of services and products, and save costs.

Flexibility – being able to change the operation in some way.

- **Product/service flexibility**: ability to introduce new/modified products/services

- **Mix flexibility**: ability to produce a wide range or mix of products/services
- **Volume flexibility**: ability to change its level of output/activity to produce different quantities/volumes of products and services over time
- **Delivery flexibility**: ability to change the timing of delivery of its services/products

Some companies have developed their flexibility in such a way that products and services are customized for each individual customer, yet manage to produce in high volume, mass production -> keeps costs down: **mass customization**.

Agility -> responding to market requirements by producing new and existing products and services fast and flexibly.

Flexibility speeds up response, saves time, maintains dependability. Flexibility can give the potential to create new, wider variety, differing volumes and differing delivery dates of services and products, and save costs.

Cost – important objective for OM, even if the organization does not compete directly on price. Productivity: ratio of what is produced (output) to what is required (input). Often partial measures of input/output used so that comparisons can be made: single-factor measure. $P_s = \frac{O}{I_i}$, $P_m = \frac{O}{\sum_i I_i}$. O = output; I = input; i = factor considered.

Improving productivity is to reduce costs. E.g. reducing costs some or all transformed and transforming resource inputs. Or making better use of inputs. All performance objectives affect cost, so an important way to reduce costs is to improve all the other objectives.

Polar representation of performance objectives good way to represent them all.

Performance measurement: business will need to measure how well, or not, it is doing. -> process of quantifying action where measurement means process of quantification and the performance of the operation is assumed to derive from actions taken by its management.

What factors to include as performance measures? Already discussed 3 levels: social, strategic and operational level. 1. Sometimes to these measures are aggregated into composite measures that combine several measures, help to present a picture of the overall performance. 2. All these factors can be broken down to more detailed ones.

Which are the most important performance measures? Often a compromise reached by making sure there is a clear link between the operation's overall strategy, most important KPI's that reflect strategic objectives, and the bundle of detailed measures used to 'flesh out' each KPI.

Multi-dimensioned performance measurement approaches, such as **balanced scorecard**, give a broader indication of overall performance.

Improving the performance of one objective, but cause others to sacrifice. Two views: 1. Emphasizing 'repositioning' performance objectives 2. Emphasizes increasing the 'effectiveness' by overcoming **trade-offs**. Presumably all the operations would ideally like to be able to offer very high variety while still having very high levels of cost efficiency. Operations that lie on the '**efficient frontier**' have performance levels that dominate those which do not: In these circumstances they may choose to reposition themselves at some other point along the efficient frontier. Those with a position on the efficient frontier will generally also want to improve their operations effectiveness by overcoming the trade-off that is implicit in the efficient frontier curve. An operation's improvement path can be described in terms of repositioning and/or overcoming its performance trade-offs. The distinction between positioning on the efficient frontier and increasing operations effectiveness by extending the frontier is an important one. Any business must make clear the extent to which it is expecting the operation to reposition itself in terms of its performance objectives and the extent to which it is expecting the operation to improve its effectiveness in several ways simultaneously.

Chapter 3

Strategic decisions: are widespread in their effect on the organization to which the strategy refers; define the position of the organization relative to its environment; and move the organization closer to its long-term goals. **Strategy**: total pattern of the decisions and actions that influence the long-term direction of the business. 'Operations' is not the same as 'operational'; it does not have a strategic role. **Content** of operations strategy: specific decisions and actions that set the operations role, objectives and activities. **Process**: method used to make specific content decisions.

3 roles of operations:

- **Implementing** business strategy: without effective implementation even the most brilliant strategy will be rendered totally ineffective
- **Supporting** business strategy: beyond implementing, developing the capabilities allow the organization to improve and refine its strategic goals.
- **Driving** business strategy: by giving it a unique and long-term advantage. The operation drives the company's strategy.

Professors Hayes and Wheelright developed 4 stage model, used to evaluate the role and contribution of the operations function.

1. **Stage 1: Internal neutrality**: very poorest level of contribution by operations function. Holding company back from competing, very little positive contribute towards competitive success, its goal 'to be ignored', improve by 'avoiding making mistakes'
2. **Stage 2: External neutrality**: begin comparing itself with similar companies in outside market
3. **Stage 3: Internally supportive**: among the best in their market. They still aspire to be clearly and unambiguously the very best in the market -> achieve by gaining a clear view of the company's competitive or strategic goals and supporting it by developing appropriate operations resources.
4. **Stage 4: Externally supportive**: operations function as foundation for its competitive success. Look at long-term, innovative, creative and proactive '1 step ahead'.

From 1 to 2: implementing strategy. 2 to 3: supporting strategy. 3 to 4: driving strategy.

Four perspectives of operations strategy:

- **Top-down** reflection of what whole group/business wants to do
- **Bottom-up** activity where operations improvements cumulatively build strategy
- **Market requirements perspective**: involves translating market requirements into operations decisions ('outside-in')
- **Operations resource perspective**: involves exploiting the capabilities of operations resources in chosen markets ('inside-out')

Corporate strategy: strategic positioning of a corporation and the businesses with it

Top-down strategies: operations strategies should take its place in hierarchy, main influence whatever the business sees as its strategic direction.

Business strategies: strategic positioning of a business in relation to its customers, markets and competitors, a subset of corporate strategy.

Functional strategy: overall direction and role of a function within the business, subset of business strategy.

Operations strategies should reflect top-down corporate and/or business objectives. Operations strategy should reflect bottom-up experience of operational reality.

Bottom-up strategies - Top-down perspective provides orthodox view of how functional strategies should be put together. However, more complex -> group reviewing its corporate strategy, also take circumstances, experiences, capabilities of businesses into account. Concept of **emergent strategies**: strategy gradually shaped over time and based on experience instead of theoretical positioning. Key virtues required shaping strategy from bottom-up -> ability learn from experience and philosophy of continual and incremental improvement.

Bottom-up and top-down can reinforce each other -> top-down to implement strategy -> bottom-up to exploit capabilities strategically.

Market-requirements-based strategies – Impossible to ensure that operation is achieving right priority between performance objectives without understanding of what market requires. Operations satisfy customers through developing their 5 performance objectives. Key point is whatever competitive factors are important should influence the priority of each other.

Competitive factors: order winners or qualifiers. **Order-winning factors**: directly and significantly contribute to winning business, key reasons to buy product/service, raising performance -> (improve chances) more business. **Qualifying factors**: performance above particular level just to be considered by the customer. Above will not necessarily gain more business. **Less important factors**: neither one of them.

If operation produces for multiple customer groups -> determine competitive factors for each group.

Way of generalizing behavior of customers and competitors: link to **life cycle** of products/services operation is producing. 4 stages:

1. **Introduction stage**

Product/service first introduced, likely to offer new in terms of design/performance, few competitors with same product/service. Need customers unlikely to be well understood, OM need flexibility to cope with changes.

2. **Growth stage**

Competitors may enter growing market. Keeping up with demand could be main operation preoccupation. Quality level ensure company keeps share of the market.

3. **Maturity stage**

Demand starts to level off. Some early competitors may have left, industry probably dominated by few larger companies. Operations expected: costs down (maintain profits, allow price cutting). Cost, productivity issues, dependable supply main concerns.

4. **Decline stage**

Sales will decline with more competitors dropping out, might be residual market, unless shortage of capacity develops, market will continue to be dominated by price competition.

Operations resources perspective – based on resource-based view (RBV). Firms with above average strategic performance likely to have gained their sustainable competitive advantage because of their core competences of their resources -> way organization acquires resources will, over long term, have significant impact on strategic success. Also, understanding and developing capabilities particularly important. This perspective may identify constraints to satisfy some markets but can also identify capabilities which can be exploited in other markets. Very important to be aware of the **intangible resources** such as relationship with suppliers, reputation with customers, knowledge of process technologies, how staff can work together etc. Central issue for OM: ensure its pattern of strategic decisions really does develop appropriate capabilities within its resources and processes. Resources can have particularly influential impact on strategic success if exhibit properties:

- **Scarce**: unequal access to resources, so that not all competing firms have scarce resources can strengthen competitive advantage
- **Not very mobile**: e.g. specialized staff that does not want to move -> process's resources more likely to be retained over time
- **Difficult to imitate** or substitute for: the less tangible and more connected with tacit knowledge embedded within the organization, the more difficult to understand and copy

Distinction operation's structure and infrastructural decisions -> **structural**: classed as primarily influencing design activities. **Infrastructural**: influence workforce organization, planning, control, improvement. Compare with hardware/software. The best and most costly facilities and technology only effective if operation also has appropriate infrastructure which governs day-to-day basis.

How can operations strategy form the basis for operations improvement? 2 models use market requirements and operations capabilities perspectives.

1. '**Line of fit**' between market requirements and operations capabilities

There should be a fit between what an operation is trying to achieve in its markets (market requirements) and what it is good at doing (operations capabilities). In terms of the framework improvement means:

i) **Achieving 'alignment'** – achieving approximate balance between required market performance and actual operations performance.
ii) **Achieving 'sustainable' alignment** – also important whether operations processes could adapt to the new market conditions.
iii) **Improving overall performance** - the more demanding the level of market requirements, the greater the level of operations capabilities will have to be.

Sometimes deviating from the line brings risks, sometimes improvements.

2. **Importance-performance matrix**

More focused -> gain understanding of relative importance customers of various competitive factors. Competitors: points of comparison against which operation can judge its performance. Competitive viewpoint: operations improve their performance, the improvement which matters most is that which takes the operation past the performance levels achieved by competitors.

Both importance and performance need to be viewed together to judge the prioritization of objectives:

- **Judging importance to customers** – e.g. order-winning, qualifying and less important factors, within -> strong, medium, weak
- **Judging performance against competitors** – also in point scale

All factors can be put in a matrix. Different zones:

- **'Appropriate' zone** – satisfactory
- **'Improve' zone** – Below lower bound of acceptability
- **'Urgent-action' zone** – Important to customers but performance below that of competitors
- **'Excess?' zone** – High performance but not important for customers

'Process' of strategy -> 'how' strategies are put together. 4 stages:

Formulation – of operations strategy: process of clarifying the various objectives and decisions that make up the strategy, and the links between them. Infrequent activity.

What should the formulation process be trying to achieve?

- Is operations strategy **comprehensive**? – Does it include all important issues?
- Is operations strategy **coherent**? – coherence = choices made in each decision area all direct the operation in the same strategic direction, complementing and reinforcing each other.
- Does operations strategy have **correspondence**? – Decision pursued in each part of the strategy should correspond to the true priority of each performance objective.
- Does operations strategy identify **critical issues**? The more critical, the more attention deserves

Implementation – strategies operationalized or executed. 3 issue important:

- **Clarity of strategic decisions** - crucial for formulation: clarity. The clearer, the easier to implement the intention.
- **Motivational leadership** – to bring sense and meaning to strategic aspirations, maintain sense of purpose, modify plan in light of experience
- **Project management** – breaking up complex plan into set of distinct activities

Monitoring – Track performance make sure changes are proceeding as planned.

Control – involves evaluation of results from monitoring the implementation. Difficult at strategic level because strategic objectives not always clear. Many strategies too complex for that. So, rather than adhering dogmatically to a predetermined plan, it may be better to adapt as circumstances change.

Chapter 4

Innovation: doing something new. **Invention**: novel or unique but does not necessarily imply that the novel device/method has potential to be practical, economic etc. Innovation goes further than invention. Innovation not just the idea, also process of transforming into something that provides a return for organization, customer etc. **Creativity**: ability to move beyond conventional ideas, rules or assumptions, in order to generate new ideas. Innovation creates the novel idea; **design** makes it work in practice.

New ideas usually follow **S-shaped progression**. Begin: often large amounts resources, time, effort needed, relatively small performance improvements. With time, experience and knowledge grow -> performance increases. Idea established, extending performance further becomes increasingly difficult. Again improved -> many s's connected.

Obvious difference between how pattern of new ideas emerges in different operations is rate and scale of innovation. **Radical innovation**: discontinuous, 'breakthrough' change, and **incremental innovation**: smaller, continuous changes. New entrants to market often more radical.

The Henderson-Clark model – why some established companies sometimes fail to exploit seemingly obvious incremental innovations. **Technological knowledge** into 'knowledge of the components of knowledge' and 'knowledge of how the components of knowledge link together' -> '**architectural knowledge**'. **Incremental innovation**: upon existing component and architectural knowledge. **Radical innovation**: changes both component and architectural knowledge. **Modular innovation**: existing architectural knowledge, requires new knowledge for components. **Architectural innovation**: great impact upon linkage of components (architecture), knowledge components same.

Matters -> ability to identify innovations and manage their transformation into effective designs so that they can sustain competitive advantage and/or generate social payback.

Product designers: try achieve pleasing designs meet/exceed customers' expectations, performs well, reliable, manufactured easily and quickly. Service designers: also customers' expectations, but within capabilities and delivered at reasonable cost. Design can add value to any organizations, it can:

- Drive and operationalize innovation
- Differentiate products/services
- Strengthen branding
- Reduce overall costs associated with innovation

Design process as input-transformation-output. The performance of the design process can be assed in much the same performance objectives, now also include sustainability as design objective.

Quality for design process – distinguish high- and low-quality designs in terms of ability to meet market requirements. Distinction between specification quality and conformance quality of designs important. 'Conformance' failures in the design process -> product recalls. 'Specification' quality -> degree of functionality/experience etc.

Speed for design process – depends on industry. Advantages of fast design:

- Early market launch
- Starting design late
- Frequent market stimulation

Dependability for design process – dependable -> minimize design uncertainty. External disturbances to innovation process will remain. Flexibility one of most important ways to ensure dependability.

Flexibility for design process – ability to cope with external/internal change. External when requirements customers change, especially necessary in fast moving markets. Internal -> increasing complexity and interconnectedness of product/service. One way measuring innovation flexibility: compare cost of modifying design with consequences profitability if no changes made. The lower the cost, the higher the flexibility.

Cost for design process – usually in 3 categories: cost buying inputs to process, cost providing labor, other general overhead costs of running process. Delayed completion design -> more expenditure and delayed revenue.

Sustainability for design process – extent to which benefits the triple bottom line (people, planet, profit). Design innovation process important in impacting ethical, environmental and economic well-being of stakeholders. Lots of examples in the book.

Design process involve a number of stages that move an innovation from a concept to a fully specified state. Stages:

1) <u>Concept generation</u>

Innovation can come from sources:

- Ideas from customers
- Listening to customers
- Ideas from competitor activity ('reverse engineering')
- Ideas from staff
- Ideas from R&D

Open-sourcing: products developed by open community, incl people who use it. E.g. Google, Wikipedia. Concept: large communities of people around the world come together and produce software product. Regularly updated and free. **Crowdsourcing**: process of getting work/funding/ideas from a crowd of people. **Parallel-path approach**: utilizing variety of different sources and approaches to generating ideas. **Ideas management**: type of enterprise software that can help operations to collect ideas from employees. Often used to focus ideas of specific organizational targets.

2) Concept screening

Evaluate concepts by assessing the worth or value of design options, assessing against number of design criteria. 3 broad categories:

1. **Feasibility** of design option – can we do it?
Do we have skills? Do we have organizational capacity? Have financial resources?

2. **Acceptability** of the design option – do we want to do it?
Satisfy performance criteria? Will our customers want it? Give enough financial return?

3. **Vulnerability** – do we want to take the risk?
Do we understand full consequences? What could go wrong? 'Downside risk' -> consequences if everything goes wrong.

Uncertainty surrounding the design reduces as the number of alternative designs considered decreases -> **design funnel**. As the design progresses, changes become more expensive.

3) Preliminary design

Objective this stage -> have a first attempt at specifying the individual components of products/services and relationship between them -> constitute final offering. First task: define exactly what will go into product/service -> require information about constituent component parts so package and component structure.

3 approaches to reduce design complexity (and reduce costs): **Standardization** – restrict variety to which has real value for customer (e.g. fast-food restaurants, clothes (sizes)). **Commonality** – using common elements within a product/service simplifies. Likewise, standardizing format of information inputs to a process can be achieved by using appropriately designed forms or screen formats. **Modularization** – designing standardized 'sub-components' which can be put together in different ways. Possible to create wide choice through fully interchangeable assembly of various combinations of smaller number of standard sub-assemblies.

4) Evaluation and improvement

To take preliminary design and subject it to series of evaluations to see if can be improved before tested in market. Several techniques, best known: **quality function deployment** (QFD): to try ensure eventual innovation actually meets needs of customers. QFD matrix: formal articulation of how company sees relationship between requirements of customers (whats) and design characteristics of new product (hows):

- Whats: list of competitive factors which customers find significant, relative importance is scored
- Competitive scores indicate relative performance
- Hows: various dimensions of design which will operationalize customer requirements
- Central matrix (relationship matrix): view of interrelationship betweens whats and hows.
- Technical assessment: contains absolute importance of each design characteristic
- 'Roof' captures information team has about correlations between various design characteristics.

5) Prototyping and final design

Turn improved design into prototype -> tested. Trials can include simulations, but also implementation. **Computer-aided design (CAD)** to create and modify product drawings. Store and retrieve design data quickly, also manipulate design details. Enhance flexibility of design activity, use of standardized libraries of shapes can reduce possibility of errors. **Alpha and beta testing** -> Alpha testing: internal process where developers examine product for errors, private process, in simulated environment. After that -> Beta testing: product released for testing by selected customers, external pilot test in real world, 'field testing' or customer validation.

Benefits of interactive product and service innovation? – Generally considered mistake to separate product and service design from process design. Merging stages of design innovation: 'interactive design'. Main benefit -> reduction in elapsed time for whole design innovation activity -> 'time to market' (TTM). Gives increased competitive advantage, more opportunities to improve performance. 3 factors significantly reduce time:

- Simultaneous development of various stages in overall process

Advantages of stages step-by-step: easy to manage, each stage can focus on its skills and expertise. Main problem: time consuming and costly. Often little need to until the absolute finalization of one stage. Often: 'simultaneous/concurrent engineering'.

- Early resolution of design conflict and uncertainty

A decision, once made, need not totally and utterly commit the organization. Manage to resolve conflict early in design activity, reduce degree of uncertainty within project and reduce extra costs, time. Design process requires strategic attention early, when there is most potential to affect design decisions.

- Organizational structure which reflects the development project

Total process of developing concepts through to market will almost certainly involve personnel from several different areas of the organization. Organizational question which of these 2 ideas – various organizational functions which contribute or design project itself – should dominate the way in which design activity is managed?

- Pure functional organization: all staff associated with design project based unambiguously in their functional groups, no project-based group at all
- Project forms: all individual members of staff who are involved could be moved out of their function to a task force dedicated solely to the project

In between various types with varying emphasis on two aspects of the organization. **Skunkworks**: organizational structure claimed to release design and development creativity of a group. Small team taken out of normal work and granted freedom from their normal management.

Chapter 5

Structure of operation's supply network: shape and form of the network.

Scope of operation's supply network: extent that operation decides to do activities performed by network itself, as opposed to requesting a supplier to do them.

Supply network: setting an operation in the context of all the other operations with which it interacts (suppliers/customers).

On supply side -> operation has suppliers of parts, information, services. These suppliers have their own suppliers etc.:

- First-tier suppliers: group of operations that directly supply the operation
- Second-tier suppliers: supply the first-tier suppliers

On demand side -> operation has customers, might have own set of customers

- First-tier customers: main customers group of operation
- These supply second-tier customers etc.

Its **immediate supply network**: suppliers and customers who have direct contact with an operation. All the operation which form network of suppliers and customers: **total supply network**.

Why look at whole supply network? 3 reasons:

- It helps an **understanding of competitiveness** – immediate customers and immediate suppliers main concern but looking beyond these to understand why they act like they do
- It helps **identify significant links** in the network – some operations contribute more to the performance objectives which valued by customers, so analysis of network needs to understand the downstream and upstream operations which contribute most to end customers service
- It helps **focus on long-term issues** – times when circumstances render parts of a supply network weaker than its adjacent links. Long-term supply network view would involve constantly examining technology and market changes to see how each operation might be affected.

Decisions relating to structure and scope often interrelated. The structure of operations supply network determined by 3 sets of decisions:

1. <u>How should the network be configured?</u>

'Configuring' a network: determining overall pattern. Involves attempting to manage network behavior by reconfiguring the network so as change the nature of the relationships between them, sometimes parts merging. Another trend: 'disintermediation' -> cutting out intermediaries, making direct contact with customers' customers etc.

Co-opetition – business surrounded by suppliers, customers, competitors and complementors. Complementors: your products more valued when also products of complementor. Competitors: your product less valued when also products of competitor. Can be both. Customers and suppliers should have 'symmetric' roles. Harnessing the value of suppliers, just as important as listening to needs of customers. Co-opetition when all parts can be enemies and friends.

Business ecosystem – economic community supported by a foundation of interacting organizations and individuals. Economic community produces goods/services of value to customers, who members of ecosystem. Difference with co-opetition: inclusion in idea of ecosystem of business no or little direct relationship with main supply network, exist because of network.

Dyadic interaction: focus on individual interaction between 2 specific operations to understand them better. Recently, more triads that are basic elements of supply network, complex can be broken down in **triads**. Thinking in triads strategically important: emphasizes dependence organizations are placing on their suppliers' performance when outsource service delivery. 2nd: control that buyer has over service delivery to its customer is diminished. Products/service bypass the buying organization and go directly from provider to customer. 3rd: direct link between service provider and customer can result in power gradually transferring over time from buying organization to supplier that provides service. 4th: becomes increasingly difficult for buying organization to understand what is happening between supplier and customer at day-to-day level. 5th: closeness between supplier and customer could prevent buyer from building important knowledge.

2. <u>What physical capacity should each part of the network have? (long-term capacity decision)</u>

Low occupancy -> high cost per customer. Operating high level of capacity -> longer waiting times, reduced customer service, overtime -> reduce productivity. As nominal capacity increases, lowest cost point at first reduces -> fixed cost and capital costs of constructing do not increase proportionally. These 2 factors: **economies of scale**. Do not go one forever -> **diseconomies of scale** -> complexity costs increase as size increases (communication, coordination) and larger center more likely to be partially underutilized (transport cost).

Advantages small scale in 4 areas:

- Allow businesses to locate near 'hot spots'
- Can respond rapidly to regional customer needs by basing more and smaller units of capacity close to local markets
- Take advantage of potential for HR development by allowing staff greater degree of local autonomy.
- Explore radically new technologies by acting in same way as smaller, more entrepreneurial rival.

In deciding when new capacity to be introduced company can mix 3 strategies:

- Capacity introduced generally to lead demand – always sufficient capacity to meet forecast demand
- Capacity introduced generally to lag demand – demand always ≥ capacity
- Capacity introduced to sometimes lead and sometimes lag demand – **smoothing with inventory**: inventory built up during lead times used to meet demand during lag times

Each advantages and disadvantages

Alternative view -> examining cost implications of adding increments of capacity on break-even basis. Each additional unit capacity -> fixed-cost break. Operation unlikely profitable at low out-puts. Assuming prices > marginal costs -> revenue exceed total costs. However, output equal to capacity may not be sufficient and thus unprofitable.

3. <u>Where should each part of the network be located? (location decision)</u>

Location decisions have effect on operation's costs, ability to serve customers, difficult to undo. Reasons for location decisions:

- **Changes in demand** – customer demand shift, changes volume demand (expand on existing site, larger site, additional location)
- **Changes in supply** – changes in cost, availability (mining, India -> cheap)

Aim location decision to achieve appropriate balance between (different for profit and non-profit organization):

- Spatially variable costs of operation (spatially variable = something changes with geographical location)
- Service the operation is able to provide to its customers
- Revenue potential

Location decision for any operation determined by relative strength of number of factors:

- Labour costs – simple wage costs can be misleading, must take effects productivity difference and different exchange rate in account.
- Labour skills availability – e.g. science parks close to university
- Land costs – trade-off between hotspot and generate enough revenue

- Energy costs – availability of relatively inexpensive energy
- Transportation costs – include transporting inputs form their source to operation and cost of transporting outputs to customers. Proximity to sources of supply dominates decision, also transportation to customers.
- Community factors – derive from social, political and economic environment (tax, corruption, language etc.)
- Suitability of the site itself – intrinsic characteristics can affect ability to serve customers (hotel near beach).
- Image of the location – e.g. fashion design in Milan
- Convenience for customers – important for service to customers, e.g. hospitals close to centers of demand

Scope of operation's activities within network determined by 2 decisions:

1. <u>Extent and nature of the operation's vertical integration</u>

Vertical integration: extent to which an organization owns network of which it is a part. Usually involves organization assessing wisdom of acquiring suppliers/customers. OEMs: original equipment manufacturers. Organization's vertical integration strategy can be defined in following terms:

- Direction of integration – strategy of expanding on supply side of network: **backward** or 'upstream' vertical integration. Expanding on demand side: **forward** or 'upstream' vertical integration. Backward -> to gain cost advantages or prevent competitors gaining control of important suppliers. Forward -> organization closer to its markets, allows more freedom make direct contact customers.
- Extent of process span of integration – some organization deliberately choose not to integrate far, some choose very vertically integrated.
- Balance among vertically integrated stages – amount of capacity at each stage devoted to supplying next stage. Totally balanced -> one stage produces only for next and fully satisfied.

Perceived advantages of vertical integration:

- Secures dependable access to supply or markets – more secure supply or bring business closer to customers. Downstream can give firm greater control over its market positioning.
- May reduce costs – take start-up and learning costs in account. Reduces transportation, loading costs
- May help to improve product/service quality – for specialist or technological advantage, preventing getting in hand of competitors.
- Helps in understanding other activities in the supply network

Perceived disadvantages of vertical integration:

- Creates an internal monopoly
- You cannot exploit economies of scale – specialist suppliers who can serve more than one customer likely to have volumes larger
- Results in loss of flexibility – high proportion of costs -> fixed -> reduction total volume easily below break-even point
- Cuts you off from innovation
- Distracts you from core activities (loss of focus) – vertical integration -> doing more things which can distract from few particularly important things.

2. <u>Nature and degree of outsourcing it engages in</u>

Vertical integration and **outsourcing** are the same thing. No business does everything that is required. Outsourcing -> ‘do-or-buy’ decision. Many indirect and administrative processes outsourced: business process outsourcing (BPO). Reason: reduce cost. Sometimes gains in quality and flexibility. <u>Difference</u>: vertical integration usually applied to whole operations. Outsourcing: usually smaller sets of activities previously performed in-house.

Assessing the advisability of outsourcing should include how it impacts relevant performance objectives. Also include consideration of the strategic importance of the activity and operation’s relative performance.

Outsourcing: deciding to buy in products/services rather than perform the activities in-house.
Offshoring: obtaining products/services form operations that are based outside one’s own country. One can do both, very closely related, motives may be similar.

Chapter 6

Process design -> at start important to understand design objectives, when overall shape and nature of process being decided. To do by positioning according to its volume and variety characteristics. Eventually, details analyzed to ensure that fulfils objectives effectively. Design of processes cannot be done independently of the products/services that they are creating. Point of process design -> make sure that the performance of the process is appropriate for whatever it is trying to achieve. Again, include sustainability as operational objective. Design of process judged on quality, speed, dependability, flexibility, costs and sustainability.

‘Micro’ performance flow objectives used to describe process flow performance:

- **Throughput rate** (flow rate): rate at which items emerge form process, number of items passing through process per unit of time
- **Cycle time** (takt time): reciprocal of throughput rate; time between items emerging from process.
- **Throughput time**: average elapsed time taken for imputs to move through process and become outputs
- **Number of items in process** (‘work-in-progress’/in-process inventory) as an average over period of time
- **Utilization of process resources**: proportion of available time that resources within process are performing useful work.

Throughput time = work-in-progress x cycle time (e.g. 10-minute wait = 10 people in system x 1 min per person) → **Little’s law**. -> average number of things in system is product of average rate at which things leave system and average time one spends in system.

Throughput efficiency: % of time item being worked on during item being processed. This assumes all ‘work content’ is needed, be that individual elements may not be considered ‘value-added’. -> value-added throughput efficiency. Throughput efficiency = work content/throughput time x 100%.

When transformed resource is information, when information technology used to move, process design: ‘**workflow’**.

Bottleneck in process: congestion because workload greater than capacity. Will dictate rate at which process can operate. Bottlenecks reduce efficiency, other stages underloaded. **Balancing**: trying to allocate work equally between stages, wasted time as percentage: **balancing loss**. Process design must respect task precedence, precedence diagram: representation of ordering of elements.

Advantages of long thin arrangement of stages:

- Controlled flow of items, easy to manage
- Simple handling, when heavy and difficult to move
- Lower capital requirements, specialist piece of equipment only bought once
- More efficient operation, one person small part of the job

Advantages short fat arrangement:

- Higher mix flexibility, each stage could specialize in different types
- Higher volume flexibility, when volume varies each stage can be closed or opened
- Higher robustness, if one breaks down, others can continue
- Less monotonous work

Important design objectives: to which extent process designs should be standardized. Doing things differently gives degree of autonomy and freedom but can cause confusion, misunderstanding, inefficiency. Practical dilemma: how draw line between processes required to be standardized and those allowed to be different.

Fundamental issues in regard to sustainability:

- Source of inputs
- Quantities and sources of energy consumed in the process
- Amounts and type of waste material, created in manufacturing processes
- Life of product itself
- End of life of the product

Life cycle analysis -> analyses all production inputs, life cycle use of product and final disposal (total energy used and emitted waste).

The design of any process should be governed by the volume and variety it is required to produce. General approaches to designing and managing processes: **process types**.

1. **Project processes** – discrete, usually highly customized products, relatively long timescale to complete. Low volume and high variety. Transforming resources may have to be organized especially for each item (movie production, construction)
2. **Jobbing processes** – high variety and low volumes, each product has to share operation's resources with many others, can be complex, fewer unpredictable circumstances (made-to-measure tailors, furniture restorers)
3. **Batch processes** – wide range volume and variety (special gourmet frozen foods, components which go into mass automobiles)
4. **Mass processes** – high volume and relatively narrow variety, repetitive and largely predictable (frozen food production, television factories)
5. **Continuous processes** – even higher volume and usually lower variety than mass, operate for longer periods of time, relatively inflexible, capital-intensive technologies, smooth flow from part to part (water processing, steel making)

1. **Professional services** – high-contact processes, spend considerable time, high level of customization, people based (management consultants, lawyers, architects)
2. **Service shops** – levels of volume and variety between the extremes (banks, schools, hotels, travel agents)

3. **Mass services** – limited contact, many customer transactions, staff follow set procedures (supermarkets, airport, library, call center). **Script**: using carefully designed enquiry process to cope with high volume.

Common to show volume-variety position in 'product-process' matrix. Many important elements of process design strongly related to volume-variety position of process. Most processes should lie close to diagonal of matrix -> 'natural' diagonal, 'line of fit'. On line -> normally lower operating costs. On the right: lower volumes and higher variety, more flexible than seems -> not taking advantage of ability to standardize activities -> higher costs. Left: position normally for higher volume and lower variety, now 'over-standardized' and too inflexible for their position -> higher costs.

Detailed design of process involves identifying all individual activities and deciding sequence. Can be done using **process mapping** – describing processes in terms of how activities within the process relate to each other. Many techniques, all identify different types of activity that take place during process and show flow materials/people/information through process. Symbols used to classify different types of activities. See symbols in book!!!

Sometimes too complex -> mapped at more aggregated level: **high-level process mapping**. More detailed level: **outline process map**, in a general way. All activities: **detailed process map**. **Micro-detailed process map**: specify every single motion involved in each activity.

Sometimes useful to map such processes in a way that makes the degree of visibility of each part of the process obvious -> allows those parts of the process with high visibility to be designed so that they enhance the customer's perception of the process. Line of visibility': boundary between activities that the customer could see and those they couldn't. The highest level of visibility: above 'line of interaction' are those activities that involve direct interaction between staff and customer.

Effects of process variability – 2 fundamental types of variability:

- Variability in demand for processing at an individual stage within the process, variation in the inter-arrival times of items to be processed
- Variation in time taken to perform activities at each stage

When arrival and process times are variable, sometimes the process will have items waiting to be processed, sometimes process will be idle. Process will have a 'non-zero' average queue. The greater the variability in the process, the more the waiting time utilization deviates from the simple rectangular function. 3 options to process designers wishing to improve waiting time or utilization performance:

- Accept long average waiting times and achieve high utilization
- Accept low utilization and achieve short average waiting times
- Reduce variability in arrival times, activity times, or both, and achieve higher utilization and short waiting times.

Chapter 7

'Layout' of operation or process: how its transforming resources are positioned relative to each other, how its various tasks are allocated to these transforming resources and general appearance of the transforming resources. Layout must start with full appreciation of objectives, will largely depend on strategic objectives, some general:

- Inherent safety – e.g. fire exits clearly marked, pathways clearly defined
- Security – anyone with malicious intent cannot gain access
- Length of flow – flow of materials/information/customers channeled to be appropriate for objectives, e.g. minimizing distance travelled, in supermarkets -> maximize
- Minimize delays – caused by over-long routes, inconvenient placing facilities, bottlenecks
- Reduce work-in-progress – excessive caused by bottlenecks, layout can limit
- Clarity of flow – all flow of materials/customers well signposted, clear to staff/customers
- Staff conditions – layout should provide pleasant working environment
- Communication – layout designed to promote meetings etc.
- Management co-ordination – supervision and communication should be assisted by relative location of staff, use of communication devices and information plants
- Accessibility – all machines, plant or equipment should be accessible so sufficient for inspection, cleaning etc.
- Use of space – all layouts achieve appropriate use of total space (maximize or minimize)
- Use of capital – minimized when finalizing layout
- Long-term flexibility – layouts changed periodically as needs of operation change. Good layout devised with possible future needs in mind.
- Image – layout can help to shape image, appearance of layout can be used as deliberate attempt to establish company's brand.

Most practical layouts are derived from 4 basic types:

1. **Fixed-position layout**

Transformed resources do not move between transforming resources. Instead of materials/information/customers flowing through operation, recipient of the processing is stationary and equipment move. E.g.: motorway construction, open-heart surgery, shipbuilding. Loosely related to project process type.

<u>Advantages</u>: very high product and mix flexibility, product/customer not moved/high variety of tasks for staff

<u>Disadvantages</u>: very high unit costs, scheduling space and activities can be difficult

<u>Objective</u>: maximize transforming resources' contribution to the transformation process

2. **Functional layout**

Resources or processes located together. When products/information/customers flow through operation, will take route from activity to activity according to their needs. Flow pattern complex. E.g.: hospital, machining the parts which go into aircrafts engines, supermarket, library (see book for extensive example). Loosely related to jobbing and batch process.

<u>Advantages</u>: high product and mix flexibility, relatively robust in the case of disruptions, easy to supervise

<u>Disadvantages</u>: low utilization, can have very high WIP, complex flow, with much transportation

Objective: Minimize distance travelled

3. **Cell layout**

Where transformed resources entering the operation are pre-selected to move to one part of the operation in which all the transforming resources, to meet their immediate processing needs, are located. Cell itself may be functional or line layout. Cell layout to bring order. E.g.: computer component manufacture, 'lunch' products area in supermarket, maternity unit in hospital. Service -> department sore with units of particular class. Loosely related to mass process

Advantages: can give good compromise, fast throughput, group work can result in good motivation

Disadvantages: can be costly to rearrange existing layout, can need more plant

Objective: create efficient flow for different component/product families

4. **Line ('product') layout**

Involves locating transforming resources entirely for convenience of transformed resources. Each product/information/customer follows prearranged route in which sequence activities required matches sequence in which processes have been located. 'Flow or product layout'. E.g.: automobile assembly, self-service cafeteria. Still improvements are made to the line layout.

Advantages: low unit costs for high volume, opportunities for specialization of equipment

Disadvantages: can have low mix flexibility, not very robust in case of disruptions, work can be very repetitive

Objective: balancing flow

Many operations combine elements of some or all basic layouts, e.g. hospitals. Restaurant -> kitchen functional layout with various processes grouped together. Traditional service restaurant arranged in fixed-position layout. Buffet restaurant cell-type layout. Cafeteria restaurant, line layout.

Importance of flow will depend on volume and variety. Volume very low and variety high -> fixed-position. Resources in low-volume-high-variety processes should be arranged to cope with irregular flow.

For any particular product/service, the fixed costs of physically constructing a fixed-position layout are relatively small compared with any other way of product or service. The variable costs of producing each individual product/service are relatively high compared to the alternative layout types. The total costs for each layout type will depend on the volume of products or services produced. There is uncertainty about the exact fixed and variable costs of each layout -> decision can rarely be made on cost alone. Different layout types have different fixed and variable which determine the appropriateness of layout for varying volume-variety characteristics.

Look and feel very important, especially in high-visibility operations. Factors deal with physiological aspects of working. Aesthetics reflect culture of organization. It can encourage desired behaviors. Allen curve -> shows powerful negative correlation between physical distance and frequency of communication. Even when sitting closely, more likely to email more often. Look and feel of environment within operation from customer's perspective: '**servicescape'**. Individual factors that influence experience will lead to certain responses: 1. Cognitive (what people think). 2. Emotional (what they feel). 3. Physiological (what their body experiences). Also contain subjective stimuli.

Detailed design in fixed-position layout – location of resources determined on convenience of transforming resources.

Detailed design in functional layout – complex due to different options of flow. For N centers there are N! different ways of arranging. Designer needs essential pieces of information:

- Area required by each work center
- Constraints on shape of the area allocated to each work center
- Degree and direction of flow between each work center -> usually shown on **flow record chart**
- Desirability of work centers being close together or close to some fixed point in the layout

Prime objective usually to minimize costs -> minimizing total distance. **Effectiveness of layout** = $\sum F_{ij} D_{ij}$. F_{ij} = flows in loads per period of time from work center i to work center j. D_{ij} = distance between work center i and work centre j. The lower the score, the better.

Due to complexity now several heuristic procedures to derive a good sub-optimal solution. One of them: **CRAFT** -> when N is large, it is feasible to start with initial layout and then evaluate all different ways of exchanging two work centers. There are $\frac{N!}{2!(N-2)!}$ Possible ways of exchanging 2 out of N work centers. 3 inputs required: matrix of flow between departments; matrix of cost associated with transportation between each department; spatial array initial layout. From these: location of centroids of each department calculated; flow matrix weighted by cost matrix, weighted flow matrix is multiplied by distances between departments -> total transportation costs of initial layout; then calculates cost consequence of exchanging every possible pair of departments.

Detailed design in cell layout – In dividing layout into 4 cells, OM implicitly taken two interrelated decisions regarding: 1. Extent and nature of the cells it has chosen to adopt; 2. Which resources to allocate to which cells. Detailed design of cellular layouts difficult, partly because idea of cell itself is compromise between process and product layout. Simplify -> concentrate either process or product aspects of cell layout.

Concentrate on processes -> could use cluster analysis to find which processes group naturally together.

PFA production flow analysis = one approach to allocate tasks and machines to cells, examines both product and requirements and process grouping simultaneously.

Detailed design in line layout – instead of 'where to place what' it is 'what to place where'. Work tasks allocated to each location. Layout activity very similar to aspects of process design -> see chapter 6! Main decisions:

- What cycle time is needed?
- How many states are needed?
- How should the task-time variation be dealt with?
- How should the layout be balanced? (bottlenecks reduced)
- How should the stages be arranged? ('long thin' or 'short fat')

Chapter 8

How operations managers deal with process technology is now one of the most important decisions that shape the capabilities of operations. High-volume services have for years understood the value of process technologies. Operations managers do not need to be technologists. Yet they should be able to do three things:

1. Understand the process technology to the extent that they are able to articulate what it should be able to do: What do operations managers need to know about process technology?
2. Evaluate the process technology – should be able to evaluate alternative technologies and share in the decisions of which technology to choose: How does the process technology affect the operation?
3. Must implement the process technology so that it can reach its full potential in contributing to the performance of the operation as a whole: How can operations managers introduce new process technology smoothly?

Process technology: machines/equipment/devices that create/deliver products/services. This is pervasive in all types of operations. Without it many of the products and services we all purchase would be less reliable, take longer to arrive and arrive unexpectedly, only be available in a limited variety, and be more expensive. Process technology has a very significant effect on quality, speed, dependability, flexibility and cost. Facilitating direct inputs: indirect process technology, increasingly important.

Distinguish different types of process technologies by that they process:

Material-processing technologies: include any technology that shapes, transports, stores, or in any way changes physical objects. E.g.: machines and equipment found in manufacturing operations, trucks, warehouse systems, etc.

Information-processing technology: Most common single type of technology within operations, and includes any device which collects, manipulates, stores or distributes information. The use of internet-based technology increases both reach and richness.

Customer-processing technology: Very much in evidence in many services. Human element of service is being reduced with customer-processing technology being used to give an acceptable level of service while significantly reducing costs. Three types of customer-processing technologies:

1. **Active interaction technology** such as automobiles, telephones and ATMs. In all of these,

customers themselves are using the technology to create the service.

2. **Passive interactive technologies**: they 'process' and control customers by constraining their

actions in some way. E.g. aircraft, cinemas and theme parks.

3. Technologies that are 'aware' of customers but not the other way round. E.g.: security monitoring technologies in shopping malls. The objective of these 'hidden technologies' is to track customers' movement or transactions in an unobtrusive way.

Some technologies process more than one type of resource. -> **Integrating technologies** = the

technologies that process combinations of material, people and customers.

Understanding process technology: knowing about the principles behind the technology to be comfortable in evaluating some technical information, capable of dealing with experts in the technology and confident enough to ask relevant questions. In particular the following 4 key questions can help operations to grasp the essentials of the technology:

What does technology do (that is different from other similar technologies)?

How does it do it? = what particular characteristics of the technology are used to perform its function?

What benefits does it give (using the technology give to the operation)?

What constraints or risks does it impose (using the technology place on the operation?

Operations managers should understand enough about process technology to evaluate alternatives.

Implications: natural consequence for operation of adopting the technology, effects. Emerging technologies can have a potentially significant impact on how operations are managed.

Technology-related decision -> choice between alternative technologies. Process technologies can be evaluated in terms of their fit with process tasks, their effect of performance and their financial impact.

<u>Does the processing technology fit the processing task?</u> – High-variety-low-volume require general purpose, high-volume-low-variety more dedicated. In between 3 dimensions:

- **The degree of automation**

The ratio of technological to human effort it employs: the capital intensity of the process technology. Generally processes that have high variety and low volume will employ process technology with lower degrees of automation than those with higher volume and lower variety

- **Scale/scalability**, capacity of technology to process work

The virtues of smaller-scale technology are often the nimbleness and flexibility that is suited to high-variety, lower-volume processing. The advantages of large-scale technologies are similar to those of large-capacity increments (discussed in chapter 6). *Scalability*: ability to shift to a different level of useful capacity quickly and cost-effectively.

- **Coupling/connectivity**, extent to which it is integrated with other technologies

Coupling: linking together of separate activities within a single piece of process technology to form an interconnected processing system. Tight coupling usually gives fast process throughput. Closely coupled technology can be both expensive and vulnerable. Process technology in high-volume, low-variety processes is relatively automated, large-scale and closely coupled when compared to that in low-volume, high-variety processes. Coupling generally more suited to relatively low-variety-high-volume. Higher-variety processing generally requires a more open and unconstrained level of coupling -> different products and services will require a wide range of processing activities.

<u>How does technology improve operation's performance?</u>

A sensible approach to evaluating the impact of any process technology on an operation is to assess how it affects the quality, speed, dependability, flexibility and cost performance of the operation. (Flexibility usually only volume and delivery improve)

<u>Does the technology give an acceptable financial return?</u>

While the benefits of investing in new technology can be spread over many years into the future, the costs associated with investing in the technology usually occur up front. The rate of interest assumed: discount rate. More generally, the present value of €x in n years' time, at a discount rate of r per cent is: €$\frac{x}{(1+\frac{r}{100})^n}$.

Implementing process technology: organizing all the activities involved in making the technology

work as intended. No matter how potentially beneficial technology, remains only prospective until it has been implemented successfully. Four particularly important issues that affect technology implementation:

1. Technology planning in the long-term – technology road mapping

Technology roadmap: approach that provides structure that attempts to assure the alignment of developments in technology. TRM process supports technology development by facilitating collaboration between various activities and contribute technology strategy. Devine technological evolution in advance by planning timing and relationships between elements involved. Benefits: way bring together stakeholders and various perspectives they have.

2. Resource and process 'distance'

The degree of difficulty in the implementation of process technology will depend on the degree of novelty of the new technology resources and the changes required in the operation's processes. The less that the new technology resources are understood, the greater their 'distance' from the current technology resource base of the operation. The greater the resource and process distance, the more difficult any implementation is likely to be. The more distance -> more difficult to know what has (not) worked and why

3. Customer acceptability

If customers are to have direct contact with technology, they must have some idea of how to operate it. Where customers have an active interaction with technology, the limitations of their

understanding of the technology can be the main constraint on its use. The ability of the operation to train its customers in the use of its technology depends on 3 factors: complexity, repetition and variety of tasks performed by customers. If services are complex, higher levels of 'training' may be needed. Frequency of use -> payback for the 'investment' in training will be greater if customer uses the technology frequently. Training easier if customer is presented with a low variety of tasks.

4. Anticipating implementation problems

The implementation of any process technology will need to account for the 'adjustment' issues that almost always occur when making any organizational change. Adjustment issues: losses that could be incurred before the improvement is functioning as intended. But estimating the nature and extent of any implementation issues is notoriously difficult. Murphy's Law seems to prevail, this law: "if anything can go wrong, it will" -> effect has been identified empirically in range of operations, especially when new types of process technology are involved. New technology rarely behaves as planned and as changes are made their impact ripples throughout the organization. It is recognized that implementation may take some time; therefore allowances are made for the length and cost of a 'ramp-up' period. As the operation prepares for the implementation, the distraction causes performance actually to deteriorate. The area of the dip indicates the magnitude of the adjustment costs, and therefore the level of vulnerability faced by the operation.

Chapter 9

Human resources aspects are especially important in the operations function, where most 'human resources' are to be found. People in organization: contribute to human resource strategy; understand organization design; design the working environment; design individuals' and groups' jobs; allocate work times.

'Organizational' culture can apply to a single function. Overcoming cultural differences between different functions -> cultural fragmentation, own subcultures. 3 elements of operations culture:

- Believe – what the people within the operations function accept as self-evident
- Know – understand underlying principles that govern how operations and processes work, only with thorough understand contribute fully to success
- Behave – not fundamentally different from any effective manager

Human resources strategy: overall long-term approach to ensuring that an organization's human resources provide a strategic advantage. It involves 2 interrelated activities:

1. Identifying the number and type of people that are needed to manage, run and develop the organization so that it meets its strategic business objectives.

2. Putting in place the programmes and initiatives that attract, develop and retain appropriate staff.

Dave Ulrich: traditional HR departments often inadequate at fulfilling meaningful strategic role. 4 elements to HR activity:

1. Being a 'strategic partner' to the business
2. Administering HR procedures and processes
3. Being an 'employee champion'
4. Being a 'change agent'

People issues are inter-reliant -> strategic perspective aimed at identifying the relationship between all 4 roles is necessary. 1st step in developing HR strategy to understand overall strategy.

Stress can undermine quality of people's working lives -> effectiveness. Stress: adverse reaction people have to excessive pressures or other types of demand placed on them. Business-related benefits:

- Staff feeling happier, perform better
- Introducing improvements easier when stress is managed effectively
- Employment relations: problems resolved more quickly
- Attendance levels increase and sickness reduces

Organization structure: the way in which tasks and responsibilities are divided into distinct groupings, and how responsibility and co-ordination relationships between the groupings are defined.

There are many valid approaches to describing organizations. The process perspective is a particularly valuable one.

Organizations are machines - resources within organizations can be seen as 'components' in a mechanism whose purpose is clearly understood. Where it is important to impose clarity such a metaphor can be useful, and is the basis of the 'process approach'.

Organizations are organisms: Organizations are living entities. Their behaviour is dictated by the behaviour of the individual humans within them. Useful if parts of the environment change radically. The survival of the organization depends on its ability to exhibit enough flexibility to respond to its environment.

Organizations are brains: like brains, organizations process information and make decisions they balance conflicting criteria, weigh up risks and decide when an outcome is acceptable.

Organizations are cultures: its shared values, ideology, pattern of thinking and day-to-day ritual. A major strength of seeing organizations as cultures is that it draws attention to their shared 'enactment of reality.

Organizations are political systems: Organization, like communities, are governed. Individuals and groups seek to pursue their aims through the detailed politics of the organization. Such a view is useful in helping organizations to legitimize politics as an inevitable aspect of organizational life.

Most organizational designs attempt to divide an organization into discrete parts that are given some degree of authority to make decisions; it allows specialization. 3 approaches to grouping:

1. According to their functional purpose
2. By the characteristics of the resources themselves (similar technologies, similar skills, resources for particular products)
3. By markets which the resources are intended to serve (location, type of customers)

Some pure types of organization:

- **U-form organization** - Clusters resources primarily by functional purpose. Pyramid management structure. Can emphasize process efficiency above customer service and ability to adapt to changing markets. Keeps together expertise an can promote creation and sharing of technological knowledge. Problem: flexibility of deployment
- **M-form organization** – groups either resources needed for each product/service group or those needed to serve a particular geographical market. Separate functions may be distributed throughout different divisions can reduce economies of scale and operating efficiency. Allows each individual division to focus on specific needs of market
- **Matrix forms** – hybrid, combining U and M-form. 2 different structures. Each resource cluster has at least 2 lines of authority. Ensures representation of all interests within the company, can be complex and confusing.
- **N-form organization** – n = network. Resources are clustered into groups as in other organizational forms but with more delegation of responsibility for strategic management of those resources. Little hierarchical reporting. Relative strength of relationships between clusters.

4 types of operations developer role

- **Top down or bottom up** – Top down: programmatic approach, emphasizing implementation of overall strategy. Bottom up: emergent model where individual business operations together contribute to overall building of operations expertise.
- **Market requirements or operations resource focus** – Market requirements: focus on explicit performance achieved by each part of operations function. Operations resource focus: emphasizes way in which each part of operation function develops its competences and successfully deploys them in marketplaces.

How they could work:

- Operations developers as **governors** – sets clear goals for each part, judges performance and wants to know reason why
- Operations developers as **curators** – collecting performance information, disseminating information, adopt best practice elsewhere

- Operations developers as **trainers** – develop clear objectives, devise improvement methodologies that can be customized
- Operations developers as **facilitators** – advise, support and aid development through process of mentoring, relatively long-term approach

Job design involves number of elements:

- What tasks are to be allocated to each person in the operation?
- What is the best method of performing each job?
- How long will it take and how many people will be needed?
- How do we maintain commitment?
- What technology is available and how will it be used?
- What are the environmental conditions of the workplace?

Division of labour = dividing the total task down into smaller parts, each of which is accomplished by a single person or team. Some real advantages in division-of-labour principles:

1. **Promotes faster learning** - easier to learn how to do a relatively short and simple task than a long and complex one.

2. **Automation becomes easier** - dividing a total task into small parts raises the possibility of automating some of those small tasks.

3. **Reduced non-productive work** - In large, complex tasks the proportion of time spent picking up tools and materials, putting them down again and generally finding, positioning and searching can be very high indeed.

There are also serious drawbacks to highly divided jobs:

1. **Monotony**: The shorter the task, the more often operators will need to repeat it.

2. **Physical injury**: The continued repetition of a very narrow range of movements can, in extreme cases, lead to physical injury. The over-use of some parts of the body can result in pain and a reduction in physical capability = repetitive strain injury (RSI).

3. **Low flexibility**: Dividing a task up into many small parts often gives the job design a rigidity which is difficult to alter under changing circumstances.

4. **Poor robustness**: Highly divided jobs imply materials passing between several stages. If one of these stages is not working correctly, the whole operation is affected. On the other hand, if each person is performing the whole of the job, any problems will only affect that one person's output.

<u>Designing job methods – scientific management</u>

Frederick Taylor identified what he saw as the basic tenets of scientific management:

- All aspects of work should be investigated on a scientific basis to establish the laws, rules and formulae governing the best methods of working.
- Such an investigative approach to the study of work is necessary to establish what constitutes a 'fair day's work'.
- Workers should be selected, trained and developed methodically to perform their tasks.
- Managers should act as the planners of the work, while workers should be responsible for carrying out the jobs to the standards laid down.
- Co-operation should be achieved between management and workers based on the 'maximum prosperity' of both.

Scientific management -> 2 fields of study: method study, which determines the methods and activities to be included in jobs; and work measurement, which is concerned with measuring the time that should be taken for performing jobs. Together, these 2 fields are often referred to as work study.

Physiology -> way the body functions and involves two aspects: how a person interfaces with his/her immediate working sea and how people react to environmental conditions. **Ergonomics** sometimes referred to as human factors engineering or just 'human factors'. Both aspects are linked by 2 common ideas:

- There must be a fit between people and the jobs they do. To achieve this fit there are only 2 alternatives. Either the job can be made to fit the people who are doing it, or the people can be made to fit the job. Ergonomics addresses the former alternative.
- It is important to take a 'scientific' approach to job design.

Many ergonomic improvements are primarily concerned with the **anthropometric aspects of jobs**: aspects related to people's size, shape and other physical abilities. Anthropometric data: data which ergonomists use when doing this. Because we all vary in our size and capabilities, ergonomists are particularly interested in our range of capabilities, which is why anthropometric data is usually expressed in percentile terms.

Job design should also take into account the desire of individuals to fulfil their needs for self-esteem and personal development -> motivation theory and contribution to behavioral approach important. Achieves 2 important objectives in job design: 1. Provides jobs which have an intrinsically higher quality of working life and ethically desirable end in itself. 2. More motivation -> better performance, quality and quantity. This involves 2 conceptual steps: 1. Exploring how various characteristics of job affect people's motivation; 2. Exploring how individual motivation towards the job affects performance.

Job rotation: If increasing the number of related tasks in the job is constrained in some way, e.g. by technology, one approach may be to encourage job rotation = moving individuals periodically between different sets of tasks to provide some variety in their activities. When successful, job rotation can increase skill flexibility and make a small contribution to reducing monotony.

Job enlargement = allocation a larger number of tasks to individuals which are as the same type as those in the original job. It may provide a more complete and therefore slightly more meaningful job. People performing an enlarged job will not repeat themselves as often -> marginally less monotonous. Operators repeat themselves less frequently and presumably the variety of tasks is greater, although no further responsibility or autonomy is necessarily given to each operator.

Job enrichment = increasing the number of tasks and allocating extra tasks which involve more decision making, greater autonomy and greater control over the job. The effect is both to reduce repetition in the job and to increase autonomy and personal development. **Horizontal changes** = those which extend the variety of similar tasks assigned to a particular job; **Vertical changes** = those which add responsibilities, decision making or autonomy to the job.

Job enlargement -> movement only in the horizontal scale, job enrichment certainly implies movement on the vertical scale and perhaps on both scales.

Empowerment = extension of the autonomy job characteristic prominent in the behavioural approach to job design. It is usually taken to mean more than **autonomy** = giving staff the ability to change how they do their jobs. The benefits: providing fast responses to customer needs, employees who feel better about their jobs and who will interact with customers with more enthusiasm, promoting 'word-of-mouth' advertising and customer retention. There are costs associated, including higher selection and trainings costs, perceived inequity of service and the possibility of poor decisions begin made by employees.

Teamworking (closely linked to empowerment) = where staff, often with overlapping skills, collectively perform a defined task and have a high degree of discretion over how they actually perform their task. The concept of teamwork, is more prescriptive and assumes a shared set of objectives and responsibilities. The benefits of teamwork can be summarized as:

- Improving productivity through enhanced motivation and flexibility
- Improving quality and encouraging innovation
- Increasing satisfaction by allowing individuals to contribute more effectively
- Making easier to implement technological changes in workplace because teams are willing to share challenges this brings

From an operations management perspective, 3 aspects of flexible working are significant:

Skills flexibility: flexible workforce that can move across several different jobs could be deployed in whatever activity is in demand at the time. Greater emphasis must be placed on training, learning and knowledge management. Defining what knowledge and experience are required to perform particular tasks and translating into training activities are clearly prerequisites for effective multi-skilling.

Time flexibility: Many people, only want to work for part of their time, sometimes during specific parts of the day/week. Bringing both the supply of staff and the demand for their work together is the objective of 'flexible time' or flexi-time working systems.

Location flexibility: The number of jobs which are not location-specific has increased. Location-specific: a job must take place in one fixed location. Many jobs could be performed at any location where there are communication links to the rest of the organization.

<u>Designing the working environment – ergonomics</u>

One aspect of ergonomics is concerned with how a person interfaces with the physical aspects of his or her immediate working area. The immediate environment in which jobs take place will influence the way they are performed. Working conditions:

<u>Working temperature</u>: individuals vary in the way their performance and comfort vary with temperature. Some points regarding working temperatures provide guidance to job designers:

- Comfortable temperature range will depend on type of work
- Effectiveness at performing vigilance tasks reduces at temperatures too warm
- Chances of accidents occurring when temperature too high or too low

<u>Illumination levels</u>: intensity of lighting required to perform any job satisfactorily will depend on the nature of the job.

<u>Noise levels</u>: Noise-induced hearing loss is a well-documented consequence of working environments where noise is not kept below safe limits. In addition to the damaging effects of high levels of noise, intermittent and high-frequency noise can also affect work performance at far lower levels.

<u>Ergonomics in the office</u>: As the number of people working in offices has increased, ergonomic principles have been applied increasingly to this type of work. At the same time, legislation has been moving to cover office technology such as computer screens and key-boards.

At best, any 'measurement' of how long a task will, or should, take will be an estimate. Why this process of estimating work times: '**work time allocation**' -> allocating a time for completing a task because we need to do so for many important operations management decisions. The advantage of structured and systematic work measurement -> gives common currency for the evaluation and comparison of all types of work.

Work measurement: process of establishing the time for a qualified worker (=one who is accepted as having the necessary physical attributes, intelligence, skill, education and knowledge to perform the task to satisfactory standards of safety quality and quantity), at a defined level of performance, to carry out a specified job (=one for which specifications have been established to define most aspects of the job).

<u>Work measurement techniques</u>:

1. **Synthesis from elemental data**: totalling element times obtained previously from studies in other jobs containing the elements concerned or from synthetic data.

2. **Predetermined motion-time systems** (PMTS): technique whereby times established for basic human motions are used to build up the time for a job at a defined level of performance.

3. **Analytical estimating**: technique which is a development of estimating, estimated from knowledge and experience of the elements concerned.

4. **Activity sampling**: a technique in which a large number of instantaneous observations are made over a period of time of a group of machines, processes or workers.

Summary 2: Operations Strategy 2

Chapter 10

<u>Planning and control</u>: activities that attempt to reconcile the demands of the market and ability of the operation's resources to deliver. It provides the systems, procedures and decisions which bring different aspects of supply and demand together. Customers' perceptions of an operation will partially be shaped by the customer interface of its planning and control system. **Planning** -> formalization of what is intended to happen at some time in the future, no guarantee, more statement of intention. **Control** -> process of coping types of change, e.g. plans redrawn, make adjustments which allow the operation to achieve the objectives that the plans has set, even when assumptions on which the plan was based do not hold true. Planning and control are separate but closely related.

Very long term -> plans concerning what they intend to do, resources needed, objectives hope to achieve, emphasis on planning, demand described in aggregated terms, mainly focus on volume and financial targets.

Medium-term planning -> more detailed, overall demand in partially disaggregated manner.

Short-term planning -> difficult to make large changes, totally disaggregated.

Low-volume/high-variety: slow customer responsiveness, short planning horizon, major planning decision: timing, detailed control decisions, high robustness.

High-volume/low-variety: fast customer responsiveness, long planning horizon, major planning decision: volume, aggregated control decisions, low robustness.

Plan and control depend on both nature of demand and supply:

Uncertainty in supply and demand – makes planning and control more difficult. Systems should be able to cope with uncertainty in demand.

Dependent and independent demand – **Dependent demand**: predict demand with relative certainty because demand dependent upon some other factor which is known, e.g. car tires in an automobile factory, custom-made dressmaker, high-class restaurant. **Independent demand**: do not have firm 'forward visibility' of customer orders. E.g. tire fitter, governed by type of car arriving, fluctuations.

Responding to demand – 'Design-resource, create and deliver to order': e.g. advertising agency. 'Design, create and deliver to order': e.g. website designer. 'Create and deliver to order': e.g. house builder who has standard designs, or 'make to order'. 'Partially create and deliver to order': e.g. 'assemble to order' computers. 'Create to stock': e.g. preserved food production. 'Collect/choose from stock': e.g. IKEA. See the connecting for these types of operations and volume-variety. Planning and control activity will vary depending on how much work is done before demand is known.

P:D ratios – D: demand time: total length of time customers have to wait between asking for service/product and receiving it. P: total throughput time: how long operation takes to design the service/product, obtain resources, create and deliver it. The larger the P:D ratio, the more speculative the operation's planning and control activities will be.

Planning and control include 4 overlapping activities:

- **Loading** = amount of work allocated to a work center. Finite loading: approach which only allocates work to a work center up to a set limit, estimate of capacity for the work center. Especially relevant for operations where:
    a. It is possible to limit the load (appointment system for hairdresser, general medical practice)
    b. It is necessary to limit the load (safety reasons -> amount of people/luggage in aircraft)
    c. The cost of limiting the load is not prohibitive (maintain a finite order book at specialist sports car manufacturer does not adversely affect demand)

Infinite loading: approach loading where no limit, tries to cope with it. Relevant for operations where:

  a. It is not possible to limit the load (emergency department hospital)
  b. It is not necessary to limit the load (fast food outlets designed to flex capacity)
  c. The cost of limiting the load is prohibitive (retail bank turned away customers because of a set number of customers -> customers unhappy)

- **Sequencing** = decisions taken on the order in which the work will be tackled.

The **physical nature** of the inputs being processed may determine the priority of work e.g. operation using paints will use lighter shades before darker ones. Sometimes the mix of work arriving may determine the priority, e.g. jobs that physically fit together may be scheduled together to reduce waste.

Sometimes there is **customer priority** sequencing e.g. in emergency room of hospital.

DD: prioritizing by due date: work sequenced according to when it is 'due' for delivery, irrespective of the size/importance, e.g. printing unit when copies are needed. Improve: delivery dependability and average delivery speed, may not provide optimal productivity, can be flexible.

LIFO: last in, first out, e.g. unloading an elevator, more practical.

FIFO: first in, first out, e.g. queues in theme parks.

LOT: longest operation time. Advantage of occupying work centers for long periods, relatively small jobs progressing through the operation will take up time at each work center. This only takes high utilization into account.

SOT: shortest operation time first. When operation become cash constrained. Improving delivery performance, however may adversely affect total productivity and can damage service to larger customers.

Performance objectives often used: meeting 'due date' (dependability), minimizing flow time (speed), minimizing work-in-progress inventory (element of cost), minimizing idle time (element of cost)

- **Scheduling** = detailed timetable showing at what time or date jobs should start and when they should end e.g. bus schedules. Rapid-response services cannot plan when customers arrive.

Scheduling one of most complex tasks. E.g. Number of possible schedules = $(n!)^m$, n: number of jobs, m: number of machines. So scheduling often not optimal solution but an acceptable one.

Forward scheduling: starting work as soon as it arrives. **Backward scheduling**: starting jobs at the last possible moment. Choice depends largely upon circumstances.

Gantt charts – advantages: provide simple visual representation both of what should be happening and of what is actually happening. Used to 'test out' alternatives. They are not an optimizing tool; merely facilitate by communicating alternative schedules effectively.

Staff rostering: where dominant resource is staff, schedule of work times effectively determines the capacity of the operation itself, main task sufficient number of people working at point in time. E.g. retail shops schedule with demand in mind -> high visibility, have to respond directly to customer demand. Very complex because of vacation, fluctuations in demand, skill etc.

- **Monitoring and control** = ensure that planned activities really happen. Deviations from plan -> re-planning.

Push system of control -> activities scheduled by means of a central system and completed in line with central instructions. Each center pushes work without considering whether the next can make use of it. Often characterized by idle time, inventory and queues.
Pull system of control -> the place and specification of what is done are set by the customer workstation, which pulls work from previous. Demand is transmitted back through the stages from the original point of demand by original customer. Pull less likely for inventory build-up -> favored by lean operations.

Drum, buffer, rope concept from theory of constraints and optimized production technology concept -> helps to decide exactly where in a process control should occur. Most do not have same amount of work loaded -> likely to be a part of the process which acting as bottleneck -> should be control point: *drum*. Bottleneck working all the time (not enough capacity) -> sensible to keep *buffer* in front. Communication between bottleneck and input -> input not overproduce: *rope*.

Simple monitoring control model often simplification. Critical questions:

Is there consensus over what the operations objectives should be? – because individual managers have different and conflicting interests. Objectives ambiguous -> strategy has to cope with unpredictable changes in environment.

Are the effects of interventions into the operation predictable? - Relationships between intervention and resulting consequence within process are predictable, which assumes that degree of process knowledge is high.

Are the operation's activities largely repetitive? – non-repetitive -> little opportunity for learning.

These questions can form a decision tree. Any divergence from the conditions necessary for routine control implies a different type of control.

- **Expert control** -> If objectives unambiguous, effects well understood but activity not repetitive. Need to 'network' -> acquiring expertise and integrating expertise into organization.
- **Trial-and-error control** -> if strategic object relatively unambiguous, effects intervention unknown, activity repetitive -> learn through own failures.
- **Intuitive control** -> if objectives relatively unambiguous, effects and control interventions unknown, neither repetitive, learning by trial-error not possible -> intuition.
- **Negotiated control** -> objectives ambiguous -> negotiation to make unambiguous objectives. To some extent dependent on power structures.

Chapter 11

Capacity: maximum level of value-added activity over a period of time that the process/operation can achieve under normal circumstances. -> **scale of capacity**. **Processing capabilities**: more important, level of output.
Capacity management: activity of understanding the nature of demand for products/services, and effectively planning and controlling capacity in the short, medium and long term.

Established **long-term capacity** -> decide on **medium-term**: assessment of demand forecasts (2-18 months), during which time planned output can be varied. **Short-term** -> demand varies and how to respond.

Setting capacity levels over medium/short term in **aggregated terms** -> making overall broad capacity decisions, not concerned with all of the detail of the individual products/services offered. Aggregated: different products/services bundled together in order to get a broad view of demand and capacity, may mean some degree of approximation. Ultimate aggregation measure is money.

Some parts of operation at their capacity 'ceiling' -> capacity constraint for whole operation.

Decision in devising their capacity plans will affect aspects of performance:

- Costs – balance between capacity and demand, under-utilization -> high unit cost
- Revenues – balance capacity and demand, capacity ≥ demand -> no revenue lost
- Working capital – affected if build up finished goods prior to demand, demand met but inventory costs
- Quality – affected by capacity plan that involves large fluctuations in capacity levels
- Speed – of response enhanced by build-up of inventories or deliberate provision of surplus capacity to avoid queuing
- Dependability – affected by how close demand levels are to capacity
- Flexibility – especially volume flexibility enhanced by surplus capacity, if demand and capacity balanced then not able to respond to unexpected increase in demand

1st step in managing capacity: measure the aggregate demand and capacity levels and understand changes in these levels for the planning period. 2nd step: determine the operation's base level of capacity from which adjustments up or down will be made -> largely determined by performance objectives, perishability of outputs and degree of variability in demand and supply. 3rd step: identify and select methods of coping with mismatches between demand and capacity. 4th: understand the consequence of different capacity decisions.

1. <u>How are demand and capacity measured?</u>

Understanding demand for products and services – Demand forecasting very important input into the capacity management decisions. 3 requirements from demand forecast, concerned with capacity management:

- Expressed in terms that are useful for capacity management, same units such as machine hours per year or space
- As accurate as possible. To meet demand, output must be decided from a forecast.
- Gives an indication of relative uncertainty. Decisions to operate extra hours e.g. usually based on forecast levels of demand, which could differ from relative.

Understanding changes in demand – Most markets are influenced by some kind of seasonality, e.g. climatic (holidays), festive (gifts). **Seasonality**: describe changes to demand over a period of a year. Can also be weekly/daily. Extent to which an operation will have to cope with very short-term fluctuations is partly determined by how long customers are prepared to wait. E.g. emergency services cannot let their customers wait long.

Better forecasting or better operations responsiveness? – Most organizations find a balance between having better forecasts and being able to cope without perfect forecasts. Capacity management requires combining attempts to increase market knowledge with attempts to increase operations flexibility.

Understanding capacity – Only one type of output, the **output capacity measure** is pretty simple. With a much wider range, output places varying demands on the process so output measures capacity are less useful -> **input capacity measures** frequently used. See table in book for examples.

- *Capacity depends on activity mix* - Capacity is a function of service/product mix, duration and product service specification, e.g. in a hospital it is very hard to predict the output. Some problems caused by variation mix can be partially overcome by using **aggregated capacity measures**.
- *Capacity depends on the duration over which output is required* – Capacity is the output that can be delivered in a defined unit of time. The level of activity and output that may be achievable over a short period is not the same as the capacity that is sustainable on a regular basis. 3 different measures of capacity:
    a. **Design capacity** – theoretical capacity is what technical designers had in mind.
    b. **Effective capacity** – the capacity after planned losses are accounted for.
    c. **Actual output** – capacity after both planned and unplanned losses are accounted for.

    Utilization = $\frac{Actual\ output}{Design\ capacity}$. Efficiency = $\frac{Actual\ output}{Effective\ capacity}$.
- *Capacity depends on the specification of output* – Some operation can increase their output by changing the specification of the product/service. Important to distinguish the most important tasks to increase capacity in short term.
- *Capacity 'leakage'* - = reduction in capacity due to quality problems, delays in delivery, machine breakdown etc. Popular method of assessing this leakage is the overall equipment effectiveness (OEE) measure, calculated: OEE = a x p x q. a: availability of a process, p: performance or speed of a process, q: quality of product/service that process creates. To process effectively all 3 need high levels of performance. OEE represents valuable operating time as a percentage of the capacity something was designed to have.

Understanding changes in capacity – capacity management decisions should reflects both predictable and unpredictable variations in capacity and demand.

2. <u>How should the operation's base capacity be set?</u>

The higher the **base level** of capacity, the less capacity fluctuations is needed to satisfy demand.

Setting demand – Base level of capacity should be related to 3 factors:

1. *Relative importance of the operation's performance objectives*

E.g. high base level -> high cost (lots of underutilization) but also ability to flex output. Trade-off between fixed and working capital -> low base level, low investment but increasing working capital to meet future demand.

2. *The perishability of the operation's outputs*

Demand or supply perishable -> base capacity relatively high level, because cannot be stored.

3. *The degree of variability in demand or supply*

Variability in either demand or capacity will reduce ability to process its inputs -> reduce effective capacity. -> long throughput times, queue build up so affects inventory levels. So when high level of variability -> relatively high base level of capacity to provide extra capacity.

3. What are the ways of coping with mismatches between demand and capacity?

The nature of capacity management depends on the mixture of predictable and unpredictable demand and capacity variation. 3 options for coping with variation, usually organizations use a combination.

- **Level capacity plan** – capacity is fixed regardless of fluctuations in forecast demand. Can achieve: stable employment patterns, high process utilization, high productivity, low unit costs. But also: create considerable inventory, decisions taken for what immediately sale and what inventory? For e.g. hotel and supermarket a high level of capacity would be necessary at this plan. Low utilization can be at e.g. expensive jewelry and real-estate agents. May let customers wait long.
- **Chase demand plan** – attempts to match capacity closely to varying levels of forecast demand, opposite and more hard than level plan. Usual by operations that not able to store their output such as perishable products or certain services. Where output can be stored, used to minimize inventory. Adjusting capacity by e.g. overtime, part-time staff, outsourcing, hiring, firing, subcontracting.
- **Demand management** – change pattern of demand to bring it closer to available capacity, some methods:
  a. *Constraining customer access* – customers only allowed to access products/services at certain times (reservation systems)
  b. *Price differentials* – adjusting price to reflect demand (vacation in high-season more expensive)
  c. *Scheduling promotion* – varying degree of market stimulation through promotion to encourage demand during low periods (promoting turkeys outside thanksgiving)
  d. *Service differentials* – allowing service levels to reflect demand, if explicitly then customers educated to move to periods of lower demand.

  More radical approach is to create alternative products/services to fill capacity in quiet periods. New products/services should meet criteria: (1) can be produced on same processes, (2) have different demand patterns of existing offerings, (3) they are sold through similar marketing channels. E.g. universities fill lecture rooms with conferences during vacation. Operation must be fully capable of serving both markets, weigh risks against benefits.

Yield management: collection of methods used to maximize operation's potential to generate profit. For operations which have relatively fixed capacities like airlines and hotels. E.g. from data airlines can balance risks of over-booking and under-booking. Hotels can give discounts outside holidays.

Mixed plans – most organizations use mix of approaches. E.g. peak demand been brought forward by offering discounts (manage demand). Capacity adjusted to 2 points in the year to reflect broad changes in demand (chase demand). Adjustment in capacity not sufficient to avoid totally build-up of inventories (level capacity).

4. How can operations understand the consequences of their capacity decisions?

Before operation can choose a capacity plan, it must understand consequences by adopting plans:

- *Consider capacity decisions using cumulative representations*

Assessing capacity plans to first plot demand on a **cumulative basis**. Cumulative representation of demand shows although peak demand in September, due to number of productive days the production starts in August. Also, fluctuation in demand is even greater than seemed. Demand per productive day is more relevant, because productive days represent the time element of capacity. Most useful: plotting demand and capacity on same graph, feasibility and consequences of capacity plan can be assessed. Vertical distance between cumulative demand and production shows surplus or shortage. Meet demand as it occurs -> cumulative production line always above demand.

Chase demand also illustrated on a cumulative representation. But production line have varying gradient, match the demand line. Inventory, carrying costs now zero/low, but now costs with changing capacity levels.

- *Considering capacity decisions using queuing principles*

Cumulative useful if able to story goods as inventory. Not possible -> waiting or **queueing theory**. Accepts that sometimes customers have to wait. Especially when arrival difficult to predict, time to make product/service uncertain, or both. Common set of elements that define queueing behavior:

a. **Source of customers** – 'population', can be finite or infinite. Finite -> probability of customers arriving depends on number of customers already being served.
b. **Arrival rate** – rate at which customers arrive. Some variability so describe in terms of probability distributions.
c. **The queue** – sometimes finite, sometimes infinite
d. **Rejecting** – if customers in system is already maximum -> new customers could be rejected
e. **Baulking** – customer refuses to wait, queue too long
f. **Reneging** – leaves queue before being served
g. **Queue discipline** – determine which order being served, e.g. fifo, lifo.
h. **Servers** – can be parallel, or in series. Human servers will vary in time it takes to serve, so processing time is usually described by a probability distribution.

Dilemma -> how many servers available? Because of variability in processing capability and arrival rate, there will always be queues and idle time. Trade-off between customers waiting and system utilization. Customer reactions to having to queue will be influenced by set of principles: unoccupied time feels longer than occupied,

a. Unoccupied time feels longer than occupied time, Pre-process waits feel longer than in-process waits
b. Anxiety makes wait seem longer, uncertain waits feel longer, unexpected waits feel longer, unfair waits feel longer, solo waiting feels longer, uncomfortable waits feel longer, new or infrequent users feel they wait longer
c. The more valuable the service, the longer customers will 'happily' wait

- *Considering capacity decisions over time*

Capacity management far more dynamic process, involves controlling and reacting to actual demand and actual capacity. At beginning of each period, same type of decisions must be made in the light of the new circumstances. Considerations of capacity strategy usually is the difference between long- and short-term outlook for volume and demand.

Chapter 14

Planning and control systems: information-processing, decision support and execution mechanisms that support operations planning and control activity. Common elements:

- Customer (demand) interface that forms two-way information link between the operation's activities and its customers;

Set of activities that interface with both individual customers and market more broadly. Defines nature of the customer experience. Quality of the service defined by the gap between customers' expectations and their perceptions of they service they receive. Managing of customer expectations is particularly important in the early stages of the experience. It should reflect the operation's objectives. It acts as a trigger function, what is triggered will depend on nature of business.

- Supply interface does same for suppliers;

Provides link between activities of operation and its suppliers. Has long and short term function. Important to manage supplier expectations.

- Set of overlapping 'core' mechanisms that perform basis tasks;

Recognizes difficulties and tries to bring order to complexity by dividing up many interrelated planning and control decisions into sub-problems to reflect the organizational hierarchy, separate different kinds of decisions at different levels in the organization and over different time periods. Gives certain amount of independence to the planners at different levels. Effectiveness -> depend largely on how effective and consistent boundaries between levels of hierarchy are managed.

- Decisions mechanism involving operations staff and information systems that makes or confirms planning and control decisions.

Does systems integrate human with 'automated' decision making? Computer can cope with immense complexity. Humans 'soft' qualitative tasks:

a. Flexibility, adaptability and learning -> deal with ambiguous, incomplete, unstable goals and constraints
b. Communication and negotiation
c. Intuition

This should provide a clue as to what should be appropriate degree of automation built into decision making.

Enterprise resource planning (**ERP**) - series of interrelated decisions about volume (quantity) and timing of materials needed, basis for ERP, called: **materials requirements planning** (**MRP**). ERP systems automates and integrates core business processes. Helps company to 'forward-plan' and understand all implications of any changes to the plan.

MRP popular during 1970's because availability of computer power to drive basic planning and control mathematics. MRP uses bill of materials (**BOM**), with demand information -> form master production schedule (**MPS**). **MRP II** during 1980's -> higher degree of processing power and communication, modelling 'what-if' scenarios. Strength of both: explore consequences of any changes.

ERP -> consequences of decisions in one part reflected in planning and control systems of rest of the organization. ERP fully exploited when a web-integrated ERP widely implemented (allow different ERP systems to communicate).

Benefits of ERP:

- Absolute visibility of what is happening in all parts of the business
- Discipline of forcing business-process-based changes effective mechanism for making all parts of the business more efficient
- Better 'sense of control' of operations that will form the basis for continuous improvement
- Enables for more sophisticated communication with customers, suppliers and other business partners -> more accurate and timely information
- Capable of integrating whole supply chains including suppliers' suppliers and customers' customers

ERP systems are only fully effective if the way a business organizes its processes is aligned with the underlying assumptions of its ERP system. Usually companies invest in ERP because the prospect of such organizational efficiency is attractive, or because all competitors do so. Effectiveness depends partly on suppliers' and customers' ERP systems.

High failure rate for IT projects -> often managerial, implementation or organizational factors -> one of most important: degree of alignment and integration between firm's overall IT strategy and general strategy of the firm. Impact of some IT is limited: **function IT** -> function IT can be adopted with or without any changes to other organizational structures. **Enterprise IT**: extends across much of or even the entire organization, most common: ERP. **Network IT**: facilitates exchanges between people and groups inside/outside the organization, challenge: learn how to exploit its emergent potential.

Key issues in ERP implementation -> what **critical success factors** (CSFs) should be managed. CSFs: get right in order for ERP system to work effectively. Likely problems in ERP implementation:

- Total cost of underestimated
- Time and effort often underestimated
- Resourcing from business and IT function often higher than expected
- Level of outside expertise required will be more than expected
- Changes required to business processes greater than expected
- Controlling the scope of the project more difficult than expected
- Will never be enough training
- Need for change management often not recognized until it is too late, changes required to corporate culture often underestimated (single biggest failure point)

Supplement to chapter 14

Materials requirements planning (MRP): approach to calculating how many parts/materials are required and what times. This requires data files that can be checked and updated. Calculations based on customer orders and forecast demand.

MPS main input to MRP, contains statement of volume and timing of end products to be made. It is the basis for the planning and utilization of labor and equipment, it determines the provisioning of materials and cash. Should include all sources of demand.

MPS contain statement of demand and currently available stock -> available inventory is projected ahead in time. MPS is '**chasing demand**': MPS increasing as demand increases, aims to keep inventory zero. **Levelled MPS**: averaging the amount required, smooth out peaks, more inventory.

Sales function can load known sales orders against the MPS and keep track of what is available to promise (ATP), ATP line shows maximum that is still available in any one week, against which sales can be loaded.

Bill of materials (BOM): information of what parts are required for each product. Different levels of assembly are shown, finished: level 0. Intended bill of materials: indentation of level of assembly, know required number of each part.

Calculate 'net' requirements (extra requirements so demand is met) -> requires: item master file (contains unique standard identification code), transaction file (record of receipts into stock, from stock and running balance) and location file (identifies where inventory is located).

Most important element of MRP netting process: systematic process of taking planning information and calculating the volume and timing requirements which will satisfy demand. MRP check how many of required parts already available in stock, then generates 'works orders' for net requirements, through BOM next level down etc.

MRP also considers when each of these parts is required, by process: back-scheduling, takes lead time into account.

3 planning routines to check production plans against operations resources:

- Resource requirements plans (RRPs): looking forward in long term to predict the requirements for large structural parts of the operation
- Rough-cut capacity plans (RCCPs): medium to short term, check the MPSs against known capacity bottlenecks, in case capacity constraints are broken, checks MPS and key resources only.
- Capacity requirements plans (CRPs) look at daily effect of works orders issued from MRP on the loading individual process stages.

Chapter 15

Focus of **lean**: achieve a flow of materials, information or customers that delivers exactly what customers want, in exact quantities, exactly when needed, exactly where required, at the lowest cost.

3 perspectives of lean:

- Lean is a **philosophy** of how to run operations – the involvement of staff in the operation, the drive for continuous improvement, the elimination of waste.
- Lean is a **method of planning and controlling** operations – many ideas are concerned with how many items flow through the operation and how managers can manage this flow. Uses 'pull' in contrast to MRP which uses 'push'.
- Lean is a **set of tools that improve operations** performance – the 'engine' room of the lean philosophy is a collection of improvement tools and techniques that are the means for cutting out waste.
-

How lean operations consider…

…<u>Flow</u> – Traditional: buffer between stages, makes stages insulated -> paid in terms of inventoy, queues, slow throughput times. Also, when problem in a stage, it is up to that stage to fix it and will not directly be noticed by other stages. In lean: items directed passed to next stage, problem will be seen quickly by entire system so chances to fix it increase. 1. The downstream 'customer' stage signals the need for action, customer 'pulls' item through process. 2. Flow in synchronized matter instead of dwelling in inventory. 3. This affects the motivation to improve because stages are no longer decoupled. 4. This improves exposes waste and encourages its elimination.

…<u>Inventory</u> – Gradually reducing the inventory exposes the worst of the problems which can be resolved, and so on.

…<u>Capacity utilization</u> – in lean process, any stoppage will affect the whole process -> lower capacity utilization, short term. May sound bad but producing just to keep high utilization is pointless.

…<u>The role of people</u> – Lean encourages/requires team-based problem solving, job enrichment, job rotation and multi-skilling. 'Basic working practices' used to implement the 'involvement of everyone' principle:

- **Discipline** – critical for safety of staff, environment and quality, by everyone all the time
- **Flexibility** – should be possible to expand responsibilities to the extent of people's capabilities, barriers (grading systems e.g.) removed
- **Equality** – through uniforms, consistent pay structures, despite position
- **Autonomy** – Delegate responsibility to people involved in direct activities so that management's task becomes one of supporting processes.
- **Development of personnel** – create more members who can support rigors of being competitive
- **Quality of working life (QWL)** – e.g. enjoyment, working area facilities
- **Creativity** – indispensable element of motivation. Not just doing the job but improving how its done, building the improvement in the business
- **Total people involvement** – staff take on more responsibility to use their abilities to the benefit of the company as a whole.

…<u>Improvement</u> – get closer to ideals over time. **Continuous improvement** -> if its aims are set in terms of ideals which individual organizations may never fully achieve, then the emphasis must be on the way in which an organization moves closer to the ideal state.

Most significant: focus on elimination of all forms of waste. **Waste**: any activity that does not add value. Terms conveying 3 causes of waste that should be reduced or eliminated:

- **Muda** - = activities in a process that are wasteful, do not add value to operation/customer. Causes likely to be poorly communicated objectives, inefficient use of resources.
- **Mura** – 'lack of consistency' -> results inn periodic overloading of staff/equipment.
- **Muri** - =absurd/unreasonable. Unreasonable requirements put on a process will result in poor outcomes. Waste can be caused by failing to carry out basic operations planning tasks.

Types of waste:

- *Waste from irregular flow* – barriers that prevent streamlined flow include:
  - Waiting time
  - Transport
  - Process inefficiencies
  - Inventory
  - Wasted motion
- *Waste from inexact supply* – barriers to achieving an exact match between supply and demand include:
  - Over/under-production
  - Early/late delivery
  - Inventory
- *Waste from inflexible response* – symptoms of inadequate flexibility include:
  - Large batches

- Delays between activities
- More variation in activity mix than in customer demand

- *Waste from variability* – symptoms of poor variability include:
  - Poor reliability of equipment
  - Defective products/services

Looking for waste - Gemba ('the actual place') term used -> if you really want to understand something, you go to where it takes place. Lean uses 'the Gemba walk' to make problems visible.

Eliminating waste through…

… streamlined flow – Primarily, reconfiguring the layout the layout of a process to aid lean synchronization -> moving from functional- towards cell-based layouts, or cell-based- to line layouts. Towards one that brings more systematization and control to process flow.

Throughput time often taken as measure for waste in a process.

Value stream mapping -> simple but effective approach to understand the flow. It visually maps a product/service production path, records direct and indirect information systems. Distinguishes value-adding and non-value-adding activities. Differs from process mapping:

- Uses broader range of information
- Usually at a higher level (5-10 activities)
- Often has a wider scope, frequently spanning the whole supply chain
- Can often be used to identify where to focus future improvement activities.

Value stream perspective -> working on and improving the 'big picture', rather than optimizing individual processes. (1). Identifying the value stream to map. (2). Physically mapping a process, mapping the information flow that enables the process to occur: 'current state' map. (3). Problems are diagnosed and changes suggested -> future state map which represents the improved version. Finally, changes are implemented.

Visual management is a lean technique designed to make the current and planned state transparent to everyone. Benefits; it can:

- Act as a common focus for team meetings
- Demonstrate methods for safe and effective working practices
- Communicate to everyone how performance is being judged
- Assess at a glance the current status of the operation
- Understand tasks and work priorities
- Judge your and others' performance
- Identify the flow of work, namely what has been and is being done
- Identify when something is not going to plan
- Show what agreed standards should be
- Provide real-time feedback on performance to everyone involved
- Reduce the reliance on formal meetings.

Important technique used to ensure flow visibility -> simple, highly visual signals to indicate a problem, e.g. an Andon light.

May be possibilities to smooth streamlined flow through use of small machines. Advantages: they can process different products/services simultaneously, system more robust, easily moved -> layout flexibility enhanced, risks of making errors in investment decisions are reduced. Utilization may be lower.

… matching supply and demand exactly – value of supply is always time dependent, something delivered early or late often has less value than something delivered exactly when it is needed. Exact match, often best served by using 'pull control' -> let the downstream stage in a process pull items through the system rather than have them 'pushed'.

Method of operationalizing pull control -> use of kanbans (card or signal). Kanban serves 3 purposes:

- It's an instruction for the preceding process to send more
- It's a visual control tool to show up areas of overproduction and lack of synchronization
- It's a tool for kaizen (continuous improvement), Toyota: 'number of kanbans should reduce over time.'

… flexible processes – flexible processes can significantly enhance smooth and synchronized flow. Responding exactly and directly to customer demand implies they need to be sufficiently flexible to change what and how much they do.

For many processes, increasing flexibility means reducing changeover time, methods:

- Measure and analyze changeover activities
- Separate external and internal activities – internal: cannot be carried out while process in going on, external: can -> identifying and separating, do as much as possible while process is continuing
- Convert internal to external activities – Methods: 1. Pre-prepare activities/equipment, 2. Make changeover process intrinsically flexible, 3. Speed up any required changes of equipment/info/staff.
- Practice changeover routines – practice makes perfect

… minimizing variability – One of biggest causes: variation in quality of items.

Levelling product/service schedules – keeping the mix and volume of flow between stages at an even rate over time. Disadvantages of large batches: large amounts of inventory within units and days very different in what they are expected to produce. Smaller and even batches -> reduce overall level of WIP, regularity of production (easier planning and control).

Levelling delivery schedules – E.g. dispatch smaller quantities of all products in a single truck more frequently -> inventory lower, system can respond quickly to trends

Adopting mixed modelling – principle of levelled scheduling and give a repeated mix of outputs. Ideally, sequence items as smoothly as possible, BACBACBAC -> smooth flow.

5S – Methodology: by eliminating what is unnecessary, and making it clear, work is easier and faster.

- Sort (seiri) – eliminate what is not needed
- Straighten (seiton) – position items such that they can be easily reached whenever needed
- Shine (seiso) – keep things clean and tidy
- Standardize (seiketsu) – maintain cleanliness and order – perpetual neatness
- Sustain (shitsuke) – develop a commitment and pride in keeping to standards

Adopting total productive maintenance (TPM) – TPM aims to eliminate the variability in processes caused by the effect of breakdowns.

Principles of lean are the same for a supply chain as they are for a process. One weakness of lean, especially in context of whole supply network, is when conditions are subject to unexpected disturbance. Less control, outside own operation. For entire network it is more difficult, takes longer to achieve, but just as valuable as for individual operation. How most supply chains traditionally operated -> daily snapshot of their ERP, limited visibility means operations must space out to avoid collisions, thus reducing output or they must fly blind and thus jeopardizing reliability.

<u>Lean and the theory of constraints</u> – **TOC** -> focus attention on capacity constraints/bottleneck of operation. This helps lean in obtaining a smooth flow. Always looking for constraints: optimized production technology (**OPT**). Principles OPT which demonstrate focus on bottlenecks:

- Balance flow, not capacity -> more important to reduce throughput time rather than achieving a notional capacity balance between stages or processes.
- The level of utilization of a non-bottleneck is determined by some other constraint in the system, not by its own capacity.
- Utilization and activation of a resource are not the same. According to TOC, resource is being utilized only if it contributes to the entire process creating more output. A process/stage can be activated in the sense that it is working, but it may only be creating stock or performing other non-value-added activity.
- An hour lost at a bottleneck is an hour lost forever out of the entire system.
- An hour saved at a non-bottleneck is a mirage.
- Bottlenecks govern both throughput and inventory in the system.
- You don't have to transfer batches in the same quantities as you produce them. Flow will probably be improved by dividing large production batches into smaller ones for moving through a process.
- The size of the process batch should be variable, not fixed.
- Fluctuations in connected and sequence-dependent processes add to each other rather than averaging out.
- Schedules should be established by looking at all constraints simultaneously. Because of bottlenecks and constraints within complex systems, it is difficult to work out schedules according to a simple system of rules.

OPT uses 'drum, buffer, rope' to explain planning and control approach. Steps of TOC:

1. Identify the system constraint
2. Decide how to exploit the constraint – obtain as much capability as possible from the constraint, preferably without expensive changes.
3. Subordinate everything to the constraint – other elements adjusted to the level where the constraint can operate at maximum effectiveness
4. Elevate the constraint – eliminating the constraint, only is 2 and 3 not successful.
5. Start again from step 1

<u>Lean and MRP</u> - JIT scheduling aims to connect the new network of internal and external supply processes by means of invisible conveyors so that parts only move in response to coordinated and synchronized signals derived from end-customer demand. MRP seeks to meet demand by directing that items are only produced as needed. While MRP is excellent at planning, it is weak at control. While lean synchronization may be good at control, it is week on planning. For relatively simple structures and routines, lean and pull are good. Things get more complex -> opportunities for using pull scheduling decrease. Very complex structures require network planning methods (Ch. 19).

Chapter 16

Various reasons to explain the shift towards a focus on improvement:

- There is a perception of increased competitive pressure
- The nature of world trade is changing –> new countries emerging as producers and consumers
- New technology has both introduced opportunities to improve and disrupted existing markets
- Interested resulted in development of many new ideas and approaches to improve, the more ways to improve, the more operations will be improved
- The scope of operations managements has widened to all function of the enterprise

Firms that have improved their competitive position have improved their performance more than their competitors.

Radical/breakthrough change -> the main vehicle of improvement is major and dramatic change in the way the operation works. The impact is abrupt, often high investment of capital. High value on creative solutions and free-thinking.

Continuous/incremental change -> improving by many small incremental steps. Also known as kaizen. The rate of improvement is not important, the momentum is.

Exploitation -> activity of enhancing processes that already exist within a firm, focus on creating efficiencies rather than radically changing resources/processing, benefits tend to be relatively immediate, incremental and predictable.

Exploration -> searching for and recognizing new mindsets and way of doing things, taking risks, innovation, benefits principally long term but can be difficult to predict.

Organizational ambidexterity -> ability of a firm both to exploit and explore as it seeks to improve; improving existing resources/processes while also competing in new technologies/markets where innovation is required.

Four aspects of improvement:

- <u>Elements</u> of improvement approaches – fundamental ideas, 'building blocks' of improvement
- <u>Broad approaches</u> – underlying set of beliefs that form a coherent philosophy and shape how improvement should be accomplished
- Improvement <u>techniques</u> – formal methods that help to produce improvements
- <u>Management</u> – way the process of improvement is organized, resourced and controlled, otherwise not effective

Key elements

<u>Improvement cycles</u> – repeatedly questioning detailed working of process. 2 widely used: **PDCA**: Plan -> Do -> Check -> Act. **DMAIC**: Defining -> Measurement -> Analyzed -> Improving -> Controlled/Check. For both -> cycle starts again.

<u>Process perspective</u> – Advantages: if improvement is not reflected in the process of creating products/services, then it is not really improvement and all parts of the business manage processes.

<u>End-to-end processes</u> – identified customer needs are entirely fulfilled by an 'end-to-end' business process

Evidence-based problem solving – emphasis on scientific method, responding only to hard evidence, using statistical software to facilitate analysis

Customer centricity – see customer as most important part of organization. **Voice of the customer** (VOC) = capturing a customers' requirements, expectations, perceptions and preferences in some depth.

Systems and procedures – some type of system that supports the improvement effort may be needed, improvement system or quality system.

Reduce process variation – some aspect of process performance is measured periodically -> plotted on a timescale. Advantages: check that the performance of the process is acceptable (capable), to check if process performance is changing over time, to check on the extent of variation in process performance.

Synchronized flow – items in a process, operation, supply network flow smoothly and with even velocity

Emphasize education/training – education and training important in motivating all staff towards seeing improvement as a worthwhile activity

Perfection is the goal – current situation calibrated against 'perfection target' so visible how much more improvement is possible

Waste identification – central in order to eliminate waste

Include everybody – harnessing the skills and enthusiasm of every person and all parts of the organization

Develop internal customer-supplier relationship – ensure external customers are satisfied -> every part of the organization contributes to external customer satisfaction by satisfying its own internal customer

Broad approaches

Total quality management (TQM) – approach that puts quality at the heart of everything that is done. Stress of following elements:

- Meeting the needs and expectation of customers
- Improvement covers all parts of the organization (group based)
- Improvement includes every person in the organization
- Including all costs of quality
- Getting things 'right first time', -> designing in quality rather than inspecting it in
- Developing the systems and procedures which support improvement

Lean – key elements when used as an improvement approach:

- Customer-centricity
- Internal customer-supplier relationships
- Perfection is the goal
- Synchronized flow
- Reduce variation
- Include all people

- Waste elimination

Business process re-engineering (BPR) – all work that does not add value should be eliminated. Advocated radical change rather than incremental. Main principles:

- Rethink business processes in a cross-functional manner which organizes work around the natural flow of information
- Strive for dramatic improvements in performance by radically rethinking and redesigning the process
- Have those who use the output from a process perform the process
- Put decision points where the work is performed, do not separate those how do the work from those who control and manage the work

Controversy of BPR -> only looks at work activities rather tan at people who do the work.

Six Sigma – specification range of any part of a product/service should be 6 the standard deviation of the process. Emphasize drive towards a virtually zero defect objective. General Electric (GE) early adopter. Now broad improvement concept rather than simple examination of process variation (however, that is still an important part).

Measures to assess performance used:

- A defect – failure to meet customer-required performance
- A defect unit or item
- A defect opportunity – number of different ways a unit/output can fail to meet customer requirements
- Proportion defective – percentage/fraction of units that have one or more defects
- Process yield – percentage/fraction of units defect-free
- Defect per unit (DPU) – average number of defects on a unit of output
- Defects per opportunity – proportion/percentage of defects / total number of defect opportunities
- Defects per million opportunities (DPMO)
- The Sigma measurement – number of standard deviations of the process variability that will fit within the customer specification limits

Following elements often associated with Six Sigma:

- Customer-driven objectives – Six Sigma sometimes 'process of comparing process outputs against customer requirements', in particular it expresses performance in DPMO
- Use of evidence – lots of quantitative evidence
- Structured improvement cycle – in Six Sigma de DMAIC cycle
- Process capability and control
- Process design – latterly Six Sigma proponents also include process design in the collection of elements
- Structured training and organization of improvement – improvement initiatives only successful if significant resources and training are devoted to their management

Master Black Belt -> experts in use of Six Sigma tools and techniques, how such techniques can be used and implemented, seen as teachers to guide improvement projects. Need quantitative analytical skills and organizational and interpersonal skills.

Black Belt -> can take a direct hand in organizing improvement teams. Develop their quantitative analytical skills and act as coaches for green belt.

Green Belt -> work within improvement teams, possibly as team leaders, not in full-time positions, spend about 20% on improvement projects.

Criticisms of Six Sigma:

- Does not offer anything that was not available before, only new is gimmicky martial arts analogy of Black belts etc.
- Too hierarchical in the way it structures its various levels of involvement in improvement activity
- Expensive
- DPMO too onerous

There is a significant overlap between the various approaches to improvement in terms of the improvement elements they contain.

Lean Sigma -> waste reduction, fast throughput time and impact of lean with data-driven, rigor and variation control of Six Sigma.

Techniques

Improvement techniques: 'step-by-step' methods and tools that can be used to help find improved ways of doing things;

Scatter diagrams – provide quick and simple method of identifying whether there is evidence of a connection between 2 sets of data. Only evidence of a relationship, not necessarily cause-effect relationship.

Process maps (**flow charts**) – used to give a detailed understanding prior to improvement, quickly shows poorly organized flows. Clarify improvement opportunities and shed light on workings of an operation. Most importantly, highlight problem areas where no procedure exists to cope with a particular set of circumstances

Cause-effect diagrams (**Ishikawa diagrams**) – effective of helping to search for the root causes of problems. Used to identify areas where further data is needed. They provide a way of structuring group brainstorming sessions, often involves identifying possible causes.

Pareto diagrams – distinguish between the 'vital few' issues and the 'trivial many'. Arranging items in order of importance. Based on the fact that relatively few causes explaining majority of effects.

Why-why analysis – stating the problem -> why occurred? -> reasons -> why have reasons occurred?

How to be managed?

Should be clear why improvement is happening and what is consists of -> linking improvement to overall strategy. Generally held that ability to improve its operations performance depends to a large extent on its 'culture'. Some argue that (legally) 'copying' from outsiders can be effective. 3 strategic types of imitators:

- *Pioneer importer*: imitator that is the pioneer in another place

- *Fast second*: rapid mover arriving quickly after an innovator but before other rival imitators take large share of the market
- *The come from behind*: late entrant

Benchmarking is the process of learning from others. Based on ideas that problems are almost certainly shard by processes elsewhere and there is probably another operation somewhere that has developed a better way of doing things. Different types of benchmarking:

- Internal benchmarking – comparison between (parts of) operations within same organization
- External benchmarking – comparison between operations from you and another organization
- Non-competitive benchmarking – against external organizations which don't compete directly in the same market
- Competitive benchmarking – comparison directly between competitors in the same/similar markets
- Performance benchmarking – comparison between levels of achieved performance in different operations
- Practice benchmarking – comparison between organization's operations practices and those adopted by another operation

Benchmarking is best practiced as a continuous process of comparison. It does not provide solution but ideas/information. It is not simply copying, more like learning and adapting. It needs some investment but can be advantages in organizing staff at all levels.

Critics of benchmarking -> always limiting themselves to currently accepted methods. Methods appropriate in one operation may not be in another.

There can be no intentional improvement without learning.

Single-loop learning -> when there is a repetitive and predictable link between cause and effect. Error is corrected without questioning/altering the underlying values and objectives of the process.

Double-loop learning -> questions fundamental objectives, ability to challenge existing operating assumptions, remain open to any changes in the competitive environment.

Some tangible causes for implementation failure:

- *Top-management support*

They must understand and believe the benefits, communicate principles and techniques, participate and formulate and maintain a clear 'improvement strategy'. Strategy is necessary to provide goals and guidelines which help to keep effort in line.

- *Senior managers may not fully understand the improvement approach*
- *Avoid excessive 'hype'*

Most new ideas have something to say, but jumping from one fad to another will generate backlash against any new idea, and destroy the ability to accumulate the experience that comes from experimenting with each one.

Chapter 17

Quality -> consistent conformance of customers' expectation. Quality is multi-faceted; its individual elements differ for different operations. Customer's view is what he/she perceives the service/product to be. So quality: degree of fit between customers' expectations and customer perception of the service/product. Customers' expectations and perceptions are influenced by a number of factor, some can be controlled by operation and some can't. Perceived quality is governed by the magnitude and direction of the gap between customers' expectations and their perceptions. If perceived quality gap is such that it does not match expectations, then reasons must lie in other gaps elsewhere:

1. Gap 1: The customer's specification-operation's specification gap

Perceived quality could be poor because of mismatch internal quality and expected quality. Main organizational responsibility: marketing, operations, product/service development.

2. Gap 2: The concept-specification gap

Perceived quality may be poor because of mismatch between service/product concept and way the organization has specified quality internally. Main responsibility: marketing, operations, product/service development

3. Gap 3: The quality specification-actual quality gap

Perceived quality could be poor because of mismatch between actual quality and internal quality specification. Main responsibility: operations

4. Gap 4: The actual quality-communicated image gap

Perceived quality could be poor because there is a gap between the organization's external communication/market image and the actual quality delivered to the customer. Main responsibility: marketing.

The **sandcone theory** -> there is a generic 'best' sequence of improvement. Building up improvement is a cumulative process. *Quality* should be first priority. Next -> internal *dependability* -> next *speed* of internal throughput, only while continuing to improve quality and dependability further etc. Next is *flexibility* -> and last is *cost*.

Achieving conformance to specification requires the following steps:

Step 1 – Define the quality characteristics

Generally useful: functionality, appearance, reliability, durability, recovery and contact

Step 2 – Decide how to measure each quality characteristic

Lots of characteristics are hard to measure, still try to find a way to break it down and attempt to measure the customer perceptions. Measures used: variables -> can be measured on a continuously variable scale, and attributes -> assessed by judgement and are dichotomous (e.g. right or wrong).

Step 3 – Set quality standards for each quality characteristic

Need quality standard against which it can be checked, otherwise they cannot indicate the performance. Quality standard: level which defines boundary between acceptable and unacceptable.

Step 4 – Control quality against those standards

Conform to those standards? -> 3 checks.

(1) *Where in the operation should it check?*

(2) *Should it check every service/product or take a sample?*

Advantages of samples; inspecting everything might be dangerous, checking everything might destroy the product/service, checking everything can be time consuming and costly. Also 100% checking may not guarantee that all defects will be identified. **Type I error**: when a decision was made to do something and the situation did not warrant it. **Type II error**: nothing was done, yet a decision to do something should have been taken as the situation did indeed warrant it.

(3) *How should the checks be performed?*

Most common approach as for checking quality of a sample so as to make inferences about all the output from an operation: statistical process control (SPC). SPC looks at variability in performance to check.

<u>Step 5</u> – <u>Find and correct causes of poor quality</u>

<u>Step 6</u> – <u>Continue to make improvements</u>

There is an aspect of quality management that has been particularly important -> **total quality management** (**TQM**).

TQM -> effective system for integrating the quality development, quality maintenance and quality improvement efforts of the various groups in an organization so as to enable production/service at the most economical levels which allow for full customer satisfaction. It is a philosophy of how to approach quality improvement that stresses the 'total' of TQM, and puts quality over everything. Stress the following:

- <u>Meeting the needs and expectations of customers</u>

TQM stresses the importance of starting with an insight into customers needs -> translated into quality objectives and used to drive quality improvement

- <u>Covering all parts of the organization</u>

Powerful concept -> everyone is an internal customer and supplier. Some degree of formality to that concept with service-level agreements (SLAs) -> formal definitions of the dimensions of service and the relationship between two parts of an organization. Type of issues e.g.: response times, range of services, dependability of service supply.

Criticisms of SLAs: 1. 'pseudocontractual' nature can sometimes inhibit rather than encourage improvement. 2. Tend to emphasize the 'hard' and measurable aspects of performance.

- <u>Including every person in the organization</u>

Every person has the potential to contribute to quality.

- <u>Examining all costs which are related to quality (esp. failure costs and getting thing 'right first time')</u>

Usually categorized as:

- Prevention costs -> costs incurred in trying to prevent problems, failures and errors
- Appraisal costs -> costs associated with controlling quality to check if problems have occurred during and after creation of the product/service

- Internal failure costs -> failure costs associated with errors which are dealt with inside the operation
- External failure costs -> costs associated with an error going out of the operation to a customer

Effective investment in preventing quality errors can significantly reduce appraisal and failure costs. TQM emphasizes prevention -> the more internal and external costs can be reduced -> once confidence has been established, appraisal costs can be reduced. It shifts the emphasis from reactive to proactive -> from inspect-in to design-in (first time right).

- Developing the systems and procedures which support quality and improvement

ISO 9000 provides the set of standardized requirements for a quality management system which should apply to any organization, regardless of size, or private/public sector. Purpose -> provide assurance to purchasers that it has been produced in such a way that meets their requirements. It takes a 'process' approach that focused on outputs from any operation's process. Stresses:

1. Quality management should be customer focused
2. Quality performance should be measured
3. Quality management should be improvement driven
4. Top management must demonstrate their commitment to maintaining and continually improving management systems.

Benefits: to organizations adopting it and customers, provide useful discipline to stick to 'sensible' process-orientated procedures, gaining the certificate shows quality is taken seriously -> marketing benefit.

Criticism: encourages 'management by manual' and over-systemization, expensive and time consuming, cost and time maintaining the certificate are excessive, it is too formulaic.

- Developing a continuous process of improvement

Quality awards ->

The Deming Prize – successfully applied 'company-wide quality control' based upon statistical quality control

The Malcolm Baldrige National Quality Award – like the Deming Prize but in the USA

The EFQM Excellence Model - 14 Western European companies formed the European Foundation for Quality Management (EFQM) -> launched European Quality Award (EQA) for most successful component of TQM in Europe each year. Emphasis now more on customers. The EFQM defines **self-assessment** as 'a comprehensive, systematic, and regular review of an organization's activities and results referenced against a model of business excellence'.

ISO 14000 – largely limited to Europe, it has a three-section environmental management system which covers initial planning, implementation and objective assessment.

Supplement to Chapter 17

Statistical process control (SPC) is concerned with checking a service/product during its creation. Value is to monitor quality of period of time, using control charts -> steps can be taken before there is a problem. Looking for trends.

All processes vary to some extent, if derive from common causes they can never be entirely eliminated but can be reduced. Histogram -> process variation distribution -> normal distribution with 99.7% and within ±3 standard deviations. If that is acceptable depends on the specification range of the operation.

Process capability is a measure of the acceptability of the variation of the process. The simplest measure of capability (C_p) is given by the ratio of the specification range to the natural variation (±3) → $C_p = \frac{UTL-LTL}{6s}$. UTL: upper tolerance limit. LTL: lower tolerance limit. s: standard deviation of the process variability. Usually when $C_p < 1$, the process is not 'capable'. One-sided capabilities:

Upper one-sided index $C_{pu} = \frac{UTL-X}{3s}$

Lower one-sided index $C_{pl} = \frac{X-LTL}{3s}$. Where X is the process average.

When only the lower of the two one-sided indices is used to indicate its capability (C_{pk})= min(C_{pu},C_{pl}).

Variation not result of common causes -> then it are assignable causes. To help decide if something is due to a common-cause of assignable-cause, control limits can be added to the control chart. If any points lie outside these limits -> assignable causes, it is 'out of control'. Set limits intuitively or for example, ±3 standard deviations from the mean. If process is stopped when the actual state is in control -> type I error. If process is left alone but actual state is out of control -> Type II error.

High levels of variation reduce the ability to detect changes in process performance.

Attributes have only 2 states. The population mean ($\bar{p}$) may be unknown -> mean can be estimated from the average of the proportion of 'defectives' ($\bar{p}$), from m (≥30) samples each of n (≥100) items:

$\bar{p} = \frac{p_1+p_2+p_3+\cdots+p_n}{m}$. One standard deviation can then be estimated from: $\sqrt{\frac{\bar{p}(1-p^n)}{n}}$. The upper and lower control limits can then be set as: UCL = $\bar{p}$ + 3σ and LCL = $\bar{p}$ - 3σ. LCL cannot be negative so round up to 0.

Commonly used type of control chart is the $\bar{X}$-R chart. The means chart can pick up changes in the average output, the R chart plots the range and gives an indication of whether the variability of the process is changing, even when the process average remains constant.

Calculating the control limits, first is to estimate the population mean and average range using m samples of sample size n.

$\bar{X} = \frac{\bar{X}_1+\bar{X}_2+\cdots+\bar{X}_m}{m}$. $\bar{R} = \frac{R_1+R_2+\cdots+R_m}{m}$.

Control limits of sample means: UCL = $\bar{X} + A_2\bar{R}$, LCL = $\bar{X} - A_2\bar{R}$ (may be negative)

Control limits of range charts: UCL = $D_4\bar{R}$, LCL = $D_3\bar{R}$ (may not be negative)

Other reasons for being out of control: alternating behavior, two points near control limit, apparent trend in one direction, suspiciously average behavior, five points one side of center line, sudden change in level.

Traditionally SPC was seen as very operational, now more operation's strategic capabilities. Process control leads to learning which enhances process knowledge and builds difficult-to-copy process capability.

Chapter 18

Risk management -> identifying that things could go wrong, stopping them going wrong, reducing the consequences when things go wrong and recovering after things have gone wrong. Failure will always occur in operations; recognizing this does not imply accepting or ignoring it. Managing risk and recovery generally involves 4 sets of activities:

1. <u>Understanding what failures could potentially occur and assessing their seriousness</u>

Understand potential sources of risk. A failure to understand failure can be the root cause of a lack of resilience. Failure sources are classified as:

- **Supply failure** -> any failure in the timing/quality of goods/service delivered into an operation. Important source because of increasing dependence on outsourcing in many industries, and emphasis on keeping the supply chain 'lean'
- **Human failures** -> two types:
    - Where key personnel leave, ill, die, cannot fulfil their role
    - People making mistakes
        - Errors = mistakes in judgement, person should do something different
        - Violations = acts that are clearly contrary to defined operating procedure

  Catastrophic failures are often caused by a combination of both.

- **Organizational failure** -> failures of procedures and processes, and failures that derive from a business's organizational structure and culture. Especially failure in the design of processes and failures in the resourcing of processes. This can derive from an organizational culture that minimizes consideration of risk, or lack of clarity in reporting relationships.
- **Technology/facilities failures** -> all IT systems, machines, equipment and buildings. All are liable to failure or breakdown.
- **Product/service design failures** -> in its design stage, a product/service might look fine on paper; only when it has to cope with real circumstances might inadequacies become evident
- **Customer failures** -> customers may 'fail' in that they misuse products/services
- **Environmental disruption** -> all causes outside an operation's direct influence. E.g. political upheaval, hurricanes, fire, fraud etc.
- **E-security** -> trade-off between providing wider access through the internet and the security concerns it generates. 3 developments:
    - Increased connectivity means that everyone has at least the potential to 'see' everyone else -> more available for employees etc
    - There has been a loss of perimeter security as more people work from home or through mobile communications
    - Sometimes unregulated technologies, it takes times to discover all possible sources of risk

Use previous failures to learn about sources of potential risk: post-failure analysis:

- **Accident investigation** – because of their infrequency it's hard to identify new sources of risks in advance
- **Failure traceability** – traced back to the process which produced them
- **Complaint analysis** – analyzing the number and content of complaints over time to understand better the nature of the failure, as the customer perceives it
- **Fault-tree analysis** – made up of branches connected by 2 types of nodes: AND and OR nodes. Branches below an AND node all need to occur for the event above the node to occur. Only one branch below an OR node need to occur for the event above to occur

Estimates of failure based on historical performance can be measured in 3 ways:

- **Failure rates** – how often a failure occurs
  - FR = number of failures/total number of products tested x 100%
  - FR (in time) = number of failures/operating time

  Sometimes failure is a function of time. The curve which describes failure probability of this type: bath-tub curve. It compromises 3 distinct stages:

  - **Infant-mortality/early life** – where early failures occur caused by defective parts or improper use
  - **Normal life** stage – when the failure rate is usually low and reasonably constant, caused by random factors
  - **Wear-out** stage – when the failure rate increases as the part approaches the end of its working life and failure is caused by the ageing and deterioration of parts.
- **Reliability** – the chances of a failure occurring, measures the ability to perform as expected over time. With interdependence, a failure in 1 part will cause the whole system to fail. So the reliability of the whole system is $R_s = R_1 \times R_2 \times R_3 \times \ldots \times R_n$. The more interdependent components, the lower its reliability will be.

  Alternative measure is mean time between failures (MTBF) = $\frac{Operating\ hours}{Number\ of\ failures}$ = $\frac{1}{failure\ rate\ (in\ time)}$.
- **Availability** – the amount of available useful operating time = $\frac{MTBF}{MTBF+MTTR}$. MTBF: mean time between failures. MTTR: mean time to repair (average time taken to repair, from time it fails to time it is operational again)

Subjective estimates of failure probability are better than no estimates at all.

Most well-known approach for doing failure mode and effect analysis (**FMEA**), assigning relative priorities to risk -> identify factors that are critical to various types of failure as a means of identifying failures before they happen. 3 key questions;

- What is the likelihood that failure will occur?
- What would the consequence of the failure be?
- How likely is such a failure to be detected before it affects the customer?

Based on a quantitative evaluation -> risk priority number (**RPN**) is calculated for each potential cause.

2. <u>Examine ways of preventing failures occurring </u>– 3 approaches to reducing risk by trying to prevent failure:
   a. **Redundancy** -> having back-up systems or components in case of failure. Calculated:

$R_{a+b} = R_a + (R_b \times P(failure))$, where R_{a+b} = reliability of component a with its back-up component b. R_a = reliability of a alone. R_b = reliability of back-up component b alone. P(failure) = probability that component a will fail and therefore component b will be needed.

Industry used 3 main types of redundancy:

- Hot standby – both primary and secondary systems run simultaneously
- Warm standby – secondary system runs in the background to the primary
- Cold standby – secondary system only called upon when primary system fails

b. **Fail-safeing** -> **poka-yoke** in Japan, human mistakes are to some extent inevitable -> important to prevent them becoming defects. Poka-yokes: simple devices/systems that are incorporated into a process to prevent inadvertent mistakes by those providing a service as well as customers receiving a service.

c. Maintenance -> how organizations try to avoid failure by taking care of their physical facilities. 3 approaches:
   - **Run to breakdown maintenance** (**RTB**) – allowing facilities to continue operating until they fail.
   - **Preventive maintenance** (**PM**) – eliminate or reduce the chances of failure by servicing the facilities at pre-planned intervals.
   - **Condition-based maintenance** (**CBM**) – perform maintenance only when the facilities require it

Total productive maintenance (TPM) -> productive maintenance carried out by all employees through small-group activities. It is seen as the natural extension in the evolution from run-to-breakdown to preventive maintenance. The five goal of TPM:

1. **Improve equipment effectiveness** by examining all the losses which occur
2. **Achieve autonomous maintenance** by allowing staff to take responsibility for some of the maintenance tasks and for the improvement of maintenance performance
3. **Plan maintenance with a fully worked out approach** to all maintenance activities.
4. **Train all staff** in relevant maintenance skills so that both maintenance and operating staff have all the skills to carry out their roles.
5. **Achieve early equipment management** by 'maintenance prevention' (MP), which involves considering failure causes and the maintainability of equipment during its design, manufacture, installation and commissioning

3. <u>Minimize the negative consequence of failure: failure or risk 'mitigation'</u>

Failure or risk mitigation -> isolating a failure from its negative consequences. The nature of action taken to mitigate failure will depend on the nature of the risk. Some types of mitigation actions that may be generally applicable:

- **Mitigation planning** – ensuring that all possible failure circumstances have been identified and the appropriate mitigation actions identified, may be described in the form of a decision tree, also provides mitigation action in its own right.
- **Economic mitigation** – actions such as insurance against losses from failure, spreading the financial consequences of failure, 'hedging' against failure
- **Containment (spatial)** – stopping the failure physically spreading to affect other parts of an internal or external supply network.
- **Containment (temporal)** – containing the spread of a failure over time, particularly when information about a (potential) failure needs to be transmitted without undue delay
- **Loss reduction** – any action that reduces the catastrophic consequences of failure by removing the resources that are likely to suffer those consequences

- **Substitution** – compensating for failure by providing other resources that can substitute for those rendered less effective by the failure

4. Devise plans and procedures that will help the operation to recover from failures when they do occur

Failure recovery is the set of actions taken to reduce the impact of failure once the customer has experienced its negative effects. Recovery needs to be planned and procedures put in place. Mistakes may be inevitable, dissatisfied customers are not. Failure may turn into positive experiences, good recovery can satisfy customers.

The **complaint value chain** helps to visualize the potential value of good recovery at different stages.

Organization need to design appropriate responses to failure that are suitably aligned with the cost and the inconvenience caused by the failure to their customers. Start where failure is recognized:

- **Discover** – discover exact nature of failure. 3 pieces of information needed: What happened? Who will be affected? Why did the failure occur?
- **Act** – 3 actions: (1) tell significant people involved what you are proposing to do. (2) effects of failure need to be contained in order to stop the consequences spreading and causing further failures. (3) some kind of follow-up to make sure that the containment actions really have contained the failure
- **Learn** – revisiting the failure to find out its root cause and then engineering out the causes of the failure so it will not happen again
- **Plan** – identifying all possible failures which might occur and formally defining the procedures which the organization should follow in the case of each type of identified failure.

Summary 3: Operations Research 1

Chapter 1 An Introduction to Model Building

Operations research (or Management science) is simply a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

A **prescriptive model** "prescribes" behaviour for an organization that will enable it to best meet its goal(s). The components of a prescriptive model include:

* **Objective function(s)**
* **Decision variables:** the variables whose values are under our control and influence the performance of the system.
* **Constraints:** restrictions on the values of decision variables.

An **optimization model** seeks to find values of the decision variables that optimize (maximize or minimize) an objective function among the set of all values for the decision variables that satisfy the given constraints.

Any specification of the decision variables that satisfies all of the model's constraints is said to be in the **feasible region.** An **optimal solution** to an optimization model is any point in the feasible region that optimizes the objective function.

A **static model** is one in which the decision variables do not involve sequences of decisions over multiple periods. A **dynamic model** is a model in which the decision variables do involve sequences of decisions over multiple periods.

Suppose that whenever decision variables appear in the objective function and in the constraints of an optimization model, the decision variables are always multiplies by constants and added together. Such a model is a **linear model**. If an optimization model is not linear, then it is a **nonlinear model**.

If one or more decision variables must be integers, then we say that an optimization model is an **integer model**. If all the decision variables are free to assume fractional values, then the model is a **non-integer model**.

Suppose that for any value of the decision variables, the values of the objective function and whether or not the constraints are satisfied is known with certainty. We then have a **deterministic model**. If this is not the case, then we have a **stochastic model**.

Seven step to build a model:

* **Step 1: Formulate the Problem** -> define the organization's problem.
* **Step 2: Observe the System** -> collecting data to estimate the value of parameters that affect the organization's problem.
* **Step 3: Formulate a Mathematical Model of the Problem.**
* **Step 4: Verify the Model and Use the Model for Prediction.**
* **Step 5: Select a Suitable Alternative** -> given a model and a set of alternatives, the operations researcher now choose the alternative that best meets the organization's objectives.
* **Step 6: Present the Results and Conclusion of the Study to the Organization.**
* **Step 7: Implement and Evaluate Recommendations.**

Chapter 3 Introduction to Linear Programming

Linear programming (LP) is a tool for solving optimization problems.

Characteristics of Linear Programming Problems:

* **Decision variables:** We begin by defining the relevant decision variables. In any linear programming model, the decision variables should completely describe the decisions to be made.
* **Objective function:** In any linear programming problem, the decision maker wants to maximize or minimize some function of the decision variables. The function to be maximized or minimized is called the **objective function**.
  The coefficient of a variable in the objective function is called the **objective function coefficient** of the variable.
* **Constraints:** The values of the decision variables are limited by constraints. The coefficients of the decision variables in the constraints are called **technological coefficients**. This is because the technological coefficients often reflect the technology used to produce different products.
* **Sign restrictions:** To complete the formulation of a linear programming problem, the following question must be answered for each decision variable: 'Can the decision variable only assume nonnegative values, or is the decision variable allowed to assume both positive and negative values?' If a decision variable x_i can only assume nonnegative values, then we add the **sign restriction** $\boldsymbol{X_i \geq 0}$. If a variable x_i can assume both positive and negative (or zero) values, then we say that x_i is **unrestricted in sign (urs)**.

* A function $f(x_1, x_2, \dots, x_n) of\ x_1, x_2, \dots, x_n$ is a **linear function** if and only if for some set of constants $c_1, c_2, \dots, c_n$; $f(x_1, x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$.
* For any linear function $f(x_1, x_2, \dots, x_n)$ and any number b, the inequalities $f(x_1, x_2, \dots, x_n) \leq b$ $and\ f(x_1, x_2, \dots, x_n) \geq b$ are **linear inequalities**.

A linear programming problem (LP) is an optimization problem for which we do the following:

1. We attempt to maximize (or minimize) a linear function of the decision variables. The function that is to be maximized or minimized is called the objective function.
2. The values of the decision variables must satisfy the set of constraints. Each constraint must be a linear equation of linear inequality.
3. A sign restriction is associated with each variable. For any variable x_i, the sign restriction specifies that x_i must be either nonnegative or unrestricted in sign.

The fact that the objective function for an LP must be a linear function of the decision variables has two implications:

1. The contribution of the objective function from each decision variable is proportional to the value of the decision variable. (**Proportionality Assumption of Linear Programming**)
2. The contribution to the objective function for any variable is independent of the value of the other decision variables. (**Additivity Assumption of Linear Programming**) The fact that each LP constraints must be a linear inequality of linear equation has two implications:

1. The contribution of each variable to the left-hand side of each constraint is proportional to the value of the variable. (**Proportionality Assumption of Linear Programming**)

2. The contribution of a variable to the left-hand side of each constraint is independent of the values of the variable. (**Additivity Assumption of Linear Programming**) The **Divisibility Assumption** requires that each decision variable be allowed to assume fractional values. A linear programming problem in which some or all of the variables must be nonnegative integer values is called an **integer programming problem**.

The **Certainty Assumption** is that each parameter (objective function coefficient, right-hand side, and technological coefficient) is known with certainty.

The **feasible region** for an LP is the set of all points that satisfies all the LP's constraints and sign restriction. Any other point that is not in the LP's feasible region is said to be in the **infeasible region**.

For a maximization problem, an **optimal solution** to an LP is a point in the feasible region with the largest objective function values. Similarly, for a minimization problem, an optimal solution is a point in the feasible region with the smallest objective function value.

A line in a **graphical solution for an LP problem** on which all the point have the same zvalues (profit) is in a maximization problem called the **iso-profit line**. In a minimization problem this line is called the **iso-cost line**.

A constraint is **binding** if the left-hand side and the right-hand side of the constraint are equal when the optimal values of the decision variables are substituted into the constraint. A constraint is **nonbinding** if the left-hand side and the right-hand side of the constraint are unequal when the optimal values of the decision variables are substituted into the constraint.

A set of points S is a **convex set** if the line segment joining any pair of points in S is wholly contained in S.

For any convex set S, a point P in S is an **extreme point** is each line segment that lies completely in S and contains the point P has P as an endpoint of the line segment. When the value of at least one variable can assume arbitrarily large values, the feasible region is called an **unbounded feasible region**.

* Some LPs have an infinite number of optimal solutions (alternative or multiple optimal solutions).
* Some LPs have no feasible solutions (infeasible LPs).
* Some LPs are unbounded: There are points in the feasible region with arbitrarily large z-values.

A **static scheduling problem** is a problem in which we assume that the schedule is every week the same.

You can use linear programming with the following problems:

* A diet problem
* A work-scheduling problem
* A Capital budgeting problem
* Short-term financial planning
* Blending problems

Blending problems are situations in which various inputs must be blended in some desired proportion to produce goods for sale.

Dynamic models arise when the decision maker makes decisions at more than one point in time. In a dynamic model, decisions made during the current period influence decisions made during future periods.

Chapter 9 Integer Programming

An **integer programming problem (IP)** is an LP in which some or all of the variables are required to non-negative integers.

An IP in which all variables are required to be integers is called **pure integer programming problem**.

An IP in which only some of the variables are required to be integers is called a **mixed integer programming problem**.

An integer programming problem in which all the variable must be equal 0 or 1 is called a **0-1 integer programming problem**.

The LP obtained by omitting integer or 0-1 constraint on variables is called the **LP relaxation**.

Integer programming problem	LP relaxation
max $z = 3x_1 + 2x_2$ s.t. $x_1 + x_2 \leq 6$ $x_1, x_2 \geq 0$, x_1, x_2 $integer$	max $z = 3x_1 + 2x_2$ s.t. $x_1 + x_2 \leq 6$ $x_1, x_2 \geq 0$,

Any Integer Programming problem may be viewed as the LP relaxation plus additional constraints (the constraints that state which variables must be integers or be 0 or 1). Hence, the LP relaxation is a less constrained, or more relaxed, version of the IP. This means that the feasible region for any IP must be contained in the feasible region for the corresponding LP relaxation. For any IP that is a max problem, this implies that

$$Optimal\ z - value\ for\ LP\ relaxation \geq optimal\ z - value\ for\ IP$$

In a **fixed-charge problem**, there is a cost associated with performing an activity at a nonzero level that does not depend on the level of the activity.

In a **set-covering problem**, each member of a given set 1 must be 'covered' by an acceptable member of some set 2. The objective in a set-covering problem is to minimize the number of elements in set 2 that are required to cover all the elements in set 1. Set-covering problems have many applications in areas such as airline crew scheduling, political districting, airline scheduling and truck routing.

Either-or constraints:

* At least one of the constraints $\boldsymbol{f(x_1, x_2, \dots, x_n) \leq 0\ and\ g(x_1, x_2, \dots, x_n) \leq 0}$ holds.
* Choose a **large M** s.t. $f(x_1, x_2, \dots, x_n) \leq M\ and\ g(x_1, x_2, \dots, x_n) \leq M$ * Let y be a 0-1 variable * Add the constraints:

$$\boldsymbol{f(x_1, x_2, \dots, x_n) \leq My}$$
$$\boldsymbol{g(x_1, x_2, \dots, x_n) \leq M(1 - y)}$$

If-then constraints:

- * If the constraint $\boldsymbol{f(x_1, x_2, \ldots, x_n) > 0}$ holds, then the constraint $\boldsymbol{g(x_1, x_2, \ldots, x_n) \geq 0}$ must be satisfied, while if $f(x_1, x_2, \ldots, x_n) > 0$ is not satisfied, then $g(x_1, x_2, \ldots, x_n) \geq 0$ may or may not be satisfied. So $f(x_1, x_2, \ldots, x_n) > 0$ implies $g(x_1, x_2, \ldots, x_n) \geq 0$.
- * Choose a **large M** s.t. $f(x_1, x_2, \ldots, x_n) \leq M \text{ and} - g(x_1, x_2, \ldots, x_n) \leq M$ * Let y be a 0-1 variable * Add the constraints:

$$\boldsymbol{f(x_1, x_2, \ldots, x_n) \leq M(1 - y)}$$
$$\boldsymbol{-g(x_1, x_2, \ldots, x_n) \leq My}$$

Chapter 18 Deterministic Dynamic Programming

Dynamic programming is a technique that can be used to solve many optimization problems. In most applications, dynamic programming obtains solutions by working backwards from the end of the problem toward the beginning, thus breaking up a large, unwieldy problem into a series of smaller, more tractable problems.

Many applications of dynamic programming reduce to finding the shortest (or longest) path that joins two points in a given network.

Characteristics of dynamic programming problems:

- * The problem can be divided into stages with a decision required at each stage.
- * Each stage has a number of states associated with it. By a state, we mean the information that is needed at any stage to make an optimal decision.
- * The decision chosen at any stage describes how the state at the current stage is transformed into the state at the next stage.
- * Given the current state, the optimal decision for each of the remaining stages must not depend on previously reached states or previously chosen decisions. This idea is known as the principle of optimality.
- * If the states for the problem have been classified into one of T stages, there must be a recursion that relates the cost or reward earned during stages t, t+1, …, T to the cost or reward earned from stages t+1, t+2, …, T. In essence, the recursion formalizes the working-backward procedure.

Inventory problems with the following characteristics can be solve by using dynamic programming:

1. Time is broken up into periods, the present period 1, the next period 2, and the final period T. At the beginning of period 1, the demand during each period is known.
2. At the beginning of each period, the firm must determine how many units should be produced. Production capacity during each period is limited.
3. Each period's demand must be met on time from inventory or current production. During any period in which production takes place, a fixed cost of production as well as a variable per-unit cost is incurred.
4. The firm has limited storage capacity. This is reflected by a limit on end-of-period inventory. A per-unit holding cost is incurred on each period's ending inventory.
5. The firm's goal is to minimize the total cost of meeting on time the demands for periods 1, 2, …, T.

In a **periodic review model**, the firm's inventory position is reviewed at the end of each period, and then the production decision is made.

Resource-allocation problems, in which limited resources must be allocated among several activities, are often solved by dynamic programming. To use linear programming to do resource allocation, three assumptions must be made:

1. The amount of a resource assigned to an activity may be any nonnegative number.
2. The benefit obtained from each activity is proportional to the amount of the resource assigned to the activity.
3. The benefit obtained from more than one activity is the sum of the benefits obtained from the individual activities.

For a **knapsack problem**, let

$$c_j = benefit\ obtained\ from\ each\ type\ j\ item$$

$$w_j = weight\ of\ each\ type\ j\ item$$

Many companies and customers face the problem of determining how long a machine should be utilized before it should be traded in for a new one. Problem of this types are called **equipment-replacement problems** and can often be solved by dynamic programming.

In many dynamic programming problem, a given stage simply consists of all possible states that the system can occupy at that stage. If this is the case, then the dynamic programming recursion (for a min problem) can often be written in the following form:

$$\boldsymbol{f_t(i) = \min\{(cost\ during\ stage\ t) + f_{t+1}(new\ state\ at\ stage\ t+1)\}}$$

Where the minimum is over all decisions that are allowable, or feasible, when the state at stage t is i. correct formulation of a recursion of the this form requires that we identify **three important aspects of the problem:**

1. The set of decisions that is allowable, or feasible, for the given state and stage.
2. We must specify how the cost during the current time period (stage t) depends on the value of t, the current state, and the decision chosen at stage t.
3. We must specify how the state at stage t+1 depends on the value of t, the state at stage t, and the decision chosen at stage t.

Summary 4: Operations Research 2

Chapter 15 Deterministic EOQ Inventory Model

The purpose of inventory theory is to determine rules that management can use to minimize the costs associated with maintaining inventory and meeting customer demand. Inventory models answer the following questions:

* When should an order be placed for a product?
* How large should each order be?

Many costs associated with placing an order or producing a good internally do not depend on the size of the order or on the production run. Costs of this type are referred to as the **ordering and setup cost**. Examples of ordering cost are cost for paperwork and billing, etc. examples of setup cost are the labour cost for setting up and shutting down a machine, etc.

Unit purchasing costs are the variable cost associated with purchasing a single unit. Example are variable labour cost, variable overhead cost, etc.

Holding or carrying cost are the cost of carrying one unit of inventory for one period of time. The holding cost usually includes storage cost, insurance cost, taxes on inventory, and a cost due to the possibility of spoilage, theft, or obsolescence.

When a customer demands a product and the demand is not met on time, a stockout, or shortage, is said to occur. If customers will accept delivery at a later date, we say that demand may be **back-ordered**. The case in which back-ordering is allowed is often referred to as the **backlogging demand** case. If no customer will accept late delivery, we are in the **lost sales** case. Most costs are associated with stockouts. If back-ordering is allowed, placement of back orders usually result in an extra cost. Stockouts often cause customers to go elsewhere to meet current and future demands, resulting in lost sales and lost goodwill. These costs are the **stockout of shortage costs**.

Assumptions of EOQ **(Economic Order Quantity)** models:

* **Repetitive ordering:** the ordering decision is repetitive, in the sense that it is repeated in a regular fashion.
* **Constant demand:** demand is assumed to occur at a known, constant rate.
* **Constant lead time:** the lead time for each order is an known constant, *L*. By the **lead time** we mean the length of time between the instant when an order is placed and the instant at which the order arrives.
* **Continuous ordering:** an order may be placed at any time. Inventory models that allow this are called **continuous review models**. If the amount of on-hand inventory is reviewed periodically and orders may be placed only periodically, we are dealing with a **periodic review model**.

Assumption of the basic EOQ model:

1. Demand is deterministic and occurs at a constant rate.
2. If an order of any size (q units) is placed, an ordering and setup cost *K* is incurred.
3. The lead time for each order is zero.
4. No shortage are allowed.
5. The cost per unit-year of holding inventory is *h*.

We define *D* the be the number of units demanded per year. Then assumption 1 implies that during any time interval of length *t* years, an amount of *Dt* is demanded.
The setup cost *K* of assumption 2 is in addition to a cost *pq* of purchasing or producing the *q* unit ordered.

Assumption 5 implies that if *I* units are held for *T* years, a holding cost of *ITh* is incurred.
The EOQ model determines an ordering policy that minimizes the yearly sum of ordering, purchasing and holding cost.

Derivation of a basic EOQ model:
The value of q that minimizes annual cost is called **q***. ***TC(q)*** is the total annual cost incurred if q units are ordered each time that $I = 0$.

$$TC(q) = annual\ cost\ of\ placing\ orders + annual\ purchasing\ cost + annual\ holding\ cost.$$

$$\frac{ordering\ cost}{year} = \left(\frac{ordering\ cost}{order}\right)\left(\frac{orders}{year}\right) = \frac{KD}{q}$$
$$\frac{purchasing\ cost}{year} = \left(\frac{purchasing\ cost}{unit}\right)\left(\frac{units\ purchased}{year}\right) = pD$$

Any interval of time that begins with the arrival of an order and ends the instant before the next order is received is called a cycle.

$$\frac{holding\ cost}{year} = \left(\frac{holding\ cost}{cycle}\right)\left(\frac{cycles}{year}\right) = \frac{q}{2}\left(\frac{q}{D}\right)h = \frac{q^2h}{2D} = \frac{q^2h}{2D}\left(\frac{D}{q}\right) = \frac{hq}{2}$$

- $*\quad the\ average\ inventory\ level\ during\ each\ cyle = \frac{q}{2}$
- $*\quad each\ cycle\ is\ of\ length \frac{q}{D}$

So, $TC(q) = \frac{KD}{q} + pD + \frac{hq}{2}$

To find the value of q that minimizes TC(q), we set TC'(q) equal to zero:
$TC'(q) = -\frac{KD}{q^2} + \frac{h}{2} = 0$

So the **economic order quantity (EOQ) =** $\boldsymbol{q^* = \left(\frac{2KD}{h}\right)^{\frac{1}{2}}}$

Often, the annual holding cost is expressed in term of the cost of holding one dollar's worth of inventory for one year. Suppose that $h_d =$ cost of holding one dollar in inventory for one year. Then the cost of holding one unit of inventory for one year will be ph_d and so
$q^* = \left(\frac{2KD}{ph_d}\right)^{\frac{1}{2}}$
The inventory level at which an order should be placed is the **reorder point**.

To determine the reorder point for a basic EOQ model, when the lead time *L* is greater than zero, two cases must be considered.
Case 1: demand during the lead time does not exceed the EOQ, so $LD < EOQ$. This means that the reorder point occurs when the inventory level equals *LD*.
Case 2: demand during lead time exceeds the EOQ, so $LD > EOQ$. In this case, the reorder point does not equal *LD*. In general, it can be shown that the reorder point equals the remainder when *LD* is divided by the EOQ.

Computing the optimal order quantity when quantity discounts are allowed:
If we let q be the quantity ordered each time an order is placed, the general quantity discount model may be described as follows:

If $q < b_1$, each item cost p_1 dollars.
If $b_1 \leq q < b_2$, each item costs p_2 dollars.
If $b_{k-2} \leq q < b_{k-1}$, each item costs p_{k-1} dollars.

If $b_{k-1} \le q < b_k = \infty$, each item costs p_k dollars.
We refer to $b_1, b_2, \ldots, b_{k-1}$ as **price break points**.

1. $TC_i(q)$ = total annual cost if each order is for q units at a price p_i.
2. EOQ_i = quantity that minimizes total annual cost if, for any order quantity, the purchasing cost of the item is p_i.
3. EOQ_i is admissible if $b_{i-1} \le EOQ_i < b_i$.
4. $TC(q)$ = actual annual cost if q items are ordered each time an order is placed.

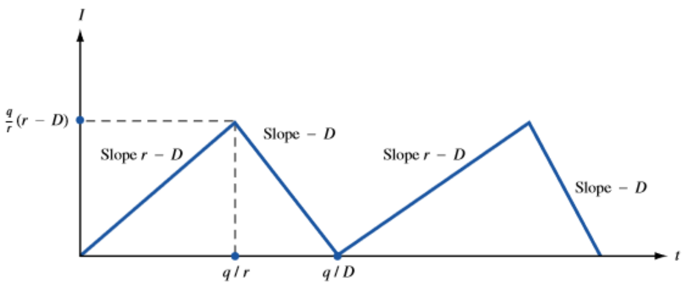
The following observations are helpful in determining the point that minimizes *TC(q).*

1. For any value of q, $TC_k(q) < TC_{k-1}(q) < \cdots < TC_2(q) < TC_1(q)$.
2. If EOQ_i is admissible, then minimum cost for $b_{i-1} \le q < b_i$ occurs for $q = EOQ_i$. If $EOQ_i < b_{i-1}$, the minimum cost for $b_{i-1} \le q < b_i$ occurs for $q = b_{i-1}$.
3. If EOQ_i is admissible, then TC(q) cannot be minimized at an order quantity for which the purchasing price per item exceeds p_i. Thus, if EOQ_i is admissible, the optimal order quantity must occur for either price $p_i, p_{i+1}, \ldots, or\ p_k$.

The continuous rate EOQ model:
Many goods are produced internally rather than purchased from an outside supplier. In this situation, the EOQ assumption that each order arrives at the same instant seems unrealistic. If a company meets demand by making its own products, the **continuous rate EOQ model** will be more realistic than the traditional EOQ model. Again, we assume that demand is deterministic and occurs at a constant rate; we also assume that shortages are not allowed. The continuous rate EOQ model assumes that a firm can produce a good at a rate of r units per time period. This means that during any time period of length *t*, the firm can produce *rt* units. We define:

* *q = number of units produced during each production run*
* *K = cost of setting up a production run*
* *h = cost of holding one unit in inventory for one year*
* *D = annual demand for the product*

FIGURE 10
Variation of Inventory for Continuous Rate EOQ Model



Assuming that per-unit production costs are independent of run size, we must determine the value of q that minimizes:

$$\frac{holding\ cost}{year} + \frac{setup\ cost}{year}$$

$$the\ average\ inventory\ level = \left(\frac{1}{2}\right)(maximum\ inventory\ level) = \left(\frac{1}{2}\right)\left(\frac{q}{r}\right)(r - D)$$

$$\frac{holding\ cost}{year} = h(average\ inventory\ level) = \frac{h(r-D)q}{2r}$$

$$\frac{ordering\ cost}{year} = \left(\frac{ordering\ cost}{cycle}\right)\left(\frac{cycles}{year}\right) = \frac{KD}{q}$$

$$\frac{holding\ cost}{year} + \frac{setup\ cost}{year} = \left(\frac{hq(r-D)}{2r}\right) + \frac{KD}{q}$$

$$\boldsymbol{optimal\ run\ size} = \left(\frac{\boldsymbol{2KD}}{\frac{\boldsymbol{h(r-D)}}{\boldsymbol{r}}}\right)^{\frac{1}{2}} = \left(\frac{\boldsymbol{2KDr}}{\boldsymbol{h(r-D)}}\right)^{\frac{1}{2}}$$

$$\boldsymbol{EOQ} = \left(\frac{\boldsymbol{2KD}}{\boldsymbol{h}}\right)^{\frac{1}{2}}$$

$$\boldsymbol{optimal\ run\ size} = \boldsymbol{EOQ}\left(\frac{\boldsymbol{r}}{\boldsymbol{r-D}}\right)^{\frac{1}{2}}$$

The EOQ model with back orders allowed:

In many real-life situations, demand is not met on time, and shortages occur. When a shortage occurs, costs are incurred. We assume that all demand is backlogged and no sales are lost. To determine the order policy that minimizes annual costs, we define

* q = order quantity
* q – M = maximum shortage that occurs under an ordering policy
* s = the cost of being short one unit for one year

Equivalently (assuming a zero lead time), the firm will be q – M units short each time an order is placed.

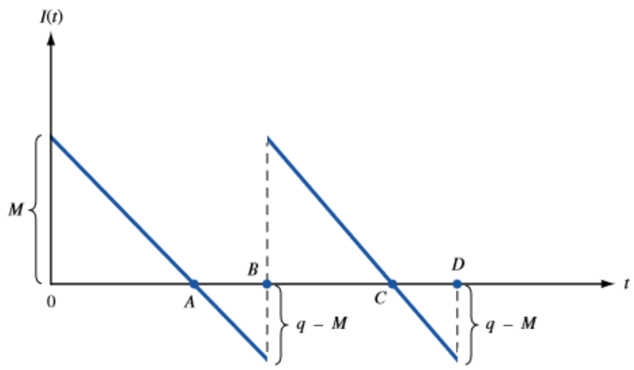


FIGURE 12 Evolution of Inventory over Time for EOQ Model with Back Orders Allowed

Since purchasing costs do not depend on q and M, we can minimize annual costs by determining the values of q and M that minimize

$$\frac{holding\ cost}{year} + \frac{shortage\ cost}{year} + \frac{order\ cost}{year}$$

* $length\ of\ AB = (length\ of\ 0B) - (length\ of\ 0A) = \frac{q-M}{D}$
* $\frac{holding\ cost}{year} = \left(\frac{holding\ cost}{cycle}\right)\left(\frac{cycles}{year}\right)$
* $\frac{holding\ cost}{cycle} = holding\ cost\ from\ time\ 0\ to\ time\ A$
* $the\ average\ inventory\ level\ between\ time\ 0\ and\ A = \frac{M}{2}\ and\ length\ 0A = \frac{M}{D}$
* $\frac{holding\ cost}{cycle} = \left(\frac{M}{2}\right)\left(\frac{M}{D}\right)h = \frac{M^2h}{2D}$
* $cycles\ per\ year = \frac{D}{q}$
* $\frac{shortage\ cost}{year} = \left(\frac{shortage\ cost}{cycle}\right)\left(\frac{cycles}{year}\right)$
* $the\ average\ level\ on\ the\ time\ interval\ AB\ is\ \frac{q-M}{2}$
* $\frac{shortage\ cost}{year} = \frac{1}{2}(q-M)\left(\frac{q-M}{D}\right)s = \frac{(q-M)^2s}{2D}$
* $\frac{shortage\ cost}{year} = \frac{(q-M)^2s}{2D}\left(\frac{D}{q}\right) = \frac{(q-M)^2s}{2D}$

$$\boldsymbol{TC(q,M)=\frac{M^2h}{2q}+\frac{(q-M)^2s}{2q}+\frac{KD}{q}}$$

$$\boldsymbol{q^* = \left(\frac{2KD(h+s)}{hs}\right)^{\frac{1}{2}} = EOQ\left(\frac{h+s}{s}\right)^{\frac{1}{2}}}$$
$$\boldsymbol{M^* = \left(\frac{2KDs}{h(h+s)}\right)^{\frac{1}{2}} = EOQ\left(\frac{s}{h+s}\right)^{\frac{1}{2}}}$$
$$\boldsymbol{maximum\ shortage = q^* - M^*}$$

Suppose that production is not instantaneous and we can produce at a rate of r units per year. If shortages are allowed, it can be shown that

$$q^* = \left(\frac{2KDr(h+s)}{h(r-D)s}\right)^{\frac{1}{2}}$$
$$M^* = \frac{q^*(r-D)}{r} - \left(\frac{2KD(r-D)h}{sr(h+s)}\right)^{\frac{1}{2}}$$

The maximum shortage occurring in this case (S^*) is given by $S^* = \left(\frac{2KD(r-D)h}{sr(h+s)}\right)^{\frac{1}{2}}$

To determine whether the assumption of constant demand is reasonable, suppose that during n periods of time, demands $d_1, d_2, \dots, d_n$ have been observed. Also, enough is known about future demands to make the assumption of deterministic demand a realistic one. To decide whether demand is sufficiently regular to justify use of EOQ models. The following computations can be done:

1. Determine the estimate $\bar{d}$ of the average demand per period given by
   $\bar{d} = \frac{1}{n}\sum_{i=1}^{i=n} d_i$
2. Determine an estimate of the variance of the per-period demand **D** from
   $Est.var\ \boldsymbol{D} = \frac{1}{n}\sum_{i=1}^{i=n} {d_i}^2 - \bar{d}^2$
3. Determine an estimate of the relative variability of demand (called the **variability co-efficient**). This quantity is labelled VC, where
   $VC = \frac{Est.var\ \boldsymbol{D}}{\bar{d}^2}$

Research indicates that the EOQ should be used if VC < 0.20; otherwise, demand is too irregular to justify the use of an EOQ model.
If all the d_i are equal, the estimate of the variance of **D** will equal zero. This will also make VC = 0.

Chapter 16 Probabilistic Inventory Models

In many situations, a decision maker is faced with the problem of determining the value *q* for a variable. After q has been determined, the value *d* assumed by a random variable **D** is observed. Depending on the values of *d* and *q*, the decision maker incurs a cost *c(d, q)*. we assume that the person is risk-neutral and want to choose *q* to minimize his/her expected cost. Since the decision is made only once, we call a model of this type a **single-period decision model.**

For the single-period model, we assume that **D** is an integer-valued discrete random variable with $P(\boldsymbol{D} = d) = p(d)$. Let *E(q)* be the decision maker's expected cost if *q* is chosen. Then
$$\boldsymbol{E(q) = \sum_d p(d)c(d,q)}$$
Let q^* be the value of q that minimizes *E(q).* If *E(q)* is a convex function, the graph of *E(q)* must look something like the following picture.
From the figure, we see that q^* is the smallest value of q for which
$$\boldsymbol{E(q^*+1) - E(q^*) \geq 0}$$

To determine q^*, we begin with
$q = 0$. If $E(1) - E(0) \leq 0$, we can benefit by increasing *q* from 0 to 1. Now we check to see whether $E(2) - E(1) \leq 0$. If this is true, then increasing *q* from 1 to 2 will reduce expected cost. Continuing in this fashion, we see that increasing *q* by 1 will reduce expected costs up to the point where we try to increase q from $q^*\ to\ q^* + 1$. In this case, increasing q by 1 will increase expected cost. From the graph, we see if $E(q^* + 1) - E(q^*) \geq 0$, then for $q \geq q^*, E(q+1) - E(q) \geq 0$. Thus, q^* must be the value of q that minimizes *E(q).*

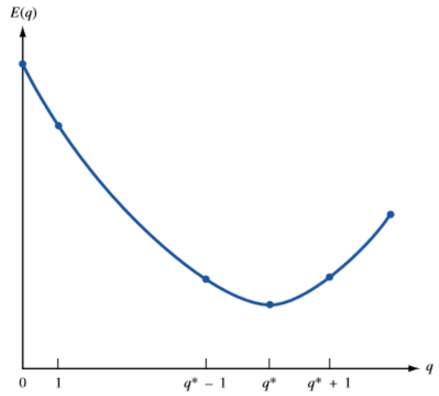


FIGURE 1
Determination of q^* by Marginal Analysis

Our approach determines q^* by repeatedly computing the effect of adding a marginal unit to the value of *q*. For this reason, it is often called **marginal analysis**. Marginal analysis is very useful if it is easy to determine a simple expression for $E(q + 1) - E(q)$.

Organizations often face inventory problems where the following sequence of events occurs:

1. The organization decides how many units to order. We let *q* be the number of units ordered.
2. With probability *p(d)*, a demand of d units occurs. We assume that *d* must be a nonnegative integer. We let **D** be the random variable representing demand.
3. Depending on *d* and *q*, a cost *c(d, q)* is incurred.

Problems that following this sequence are often called **news vendor problems**. A marginal analyses can be used to solve news vendor problems when demand is a discrete random variable and *c(d, q)* had the following form:

1. $c(d, q) = c_o q + (terms\ not\ involving\ q) \qquad (d \leq q)$
2. $c(d, q) = -c_u q + (terms\ not\ involving\ q) \qquad (d \geq q + 1)$

c_o is the per-unit cost of being overstocked. If $d \leq q$, we have ordered more than was demanded – that is, overstocked. If the size of the order is increased from *q* to *q + 1*, then the cost increased by c_o. Hence, c_o is the cost due to being overstocked by one extra unit. We refer to c_o as the **overstocking cost**. Similarly, if $(d \geq q + 1)$, we have understocked (ordered an amount less than demand). If $(d \geq q + 1)$ and we increase the size of the order by one unit, we are understocked by one less unit. Then formula 2 implies that the cost is reduced by c_u, so c_u is the per-unit cost of being understocked. We call c_u the **understocking cost**. To derive the optimal order quantity via marginal analysis, let *E(q)* be the expected cost if an order is placed for *q* units. The goal is to find the value q^* that minimizes *E(q)*. If $c(d, q)$ can be described by formulas 1 and 2 (from above), and *E(q)* is a convex function of q, then marginal analysis can be used to determine q^*. We must determine the smallest value of q^* for which $E(q + 1) - E(q) \geq 0$. To calculate $E(q + 1) - E(q)$, we must consider two possibilities:

1. $d \leq q$. In this case, ordering *q + 1* units instead of q units causes us to be overstocked by one more unit. This increases cost by c_o. The probability that case 1 will occur is simply $P(\boldsymbol{D} \leq q)$, where **D** is the random variable representing demand.
2. $(d \geq q + 1)$. In this case, ordering *q + 1* units instead of q units enables us to be short one less unit. This will decrease cost by c_u. The probability that case 2 will occur is $P(\boldsymbol{D} \geq q + 1) = 1 - P(\boldsymbol{D} \leq q)$.

In summary, a fraction $P(\boldsymbol{D} \leq q)$ of the time, ordering *q + 1* units will cost c_o more than ordering q units; and a fraction $1 - P(\boldsymbol{D} \leq q)$ of the time, ordering *q + 1* units will cost c_u less than ordering q units. Thus, on the average, ordering *q + 1* units will cost

$c_oP(\boldsymbol{D} \le q) - c_u(1 - P(\boldsymbol{D} \le q))$ more than ordering q units.

$E(q+1) - E(q) = c_oP(\boldsymbol{D} \le q) - c_u\big(1 - P(\boldsymbol{D} \le q)\big) = (c_o + c_u)P(\boldsymbol{D} \le q) - c_u$
Then $E(q+1) - E(q) \ge 0$ will hold if $(c_o + c_u)P(\boldsymbol{D} \le q) - c_u \ge 0 \; or \; P(\boldsymbol{D} \le q) \ge \frac{c_u}{c_o+c_u}$

Let $F(q) = P(\boldsymbol{D} \le q)$ be the demand distribution function. Since marginal analysis is applicable, we have just shown that *E(q)* will be minimized by the smallest value of *q* $(= q^*)$ satisfying $F(q^*) \ge \frac{c_u}{c_o+c_u}$.

The EOQ model with uncertain demand: the (r, q) and (s, S) models:
K = ordering cost.
h = holding cost/unit/year.
L = lead time for each order.
q = quantity ordered each time an order takes place.
D = random variable (assumed continuous) representing annual demand, with mean *E(**D**),* variance *var(**D**)*, and standard deviation $\sigma_{\boldsymbol{D}}$.
$c_b =$ cost incurred for each unit short, which does not depend on how long it takes to make up stockout.
OHI(t) = on-hand inventory (amount of stock on hand) at time t.
B(t) = the number of outstanding back orders at time t.
I(t) = net inventory level at time t $= OHI(t) - B(t)$.
r = inventory level at which order is placed (reorder point).

X = random variable representing demand during lead time
We assume that X is a continuous random variable having density function f(x) and mean, variance, and standard deviation of E(X), var(X) and $\sigma_{\boldsymbol{X}}$. If we assume that the demands at different points in time are independent, then it can be shown that the random lead time demand X satisfies (lead time fixed):

* $E(X) = LE(D)$
* $var(X) = L\big(var(D)\big)$
* $\sigma_X = \sigma_D\sqrt{L}$

We assume that if **D** is normally distributed, then X will also be normally distributed.
Suppose we allow the leas time *L* to be a random variable (**L**), with mean E(**L**), variance var(**L**), and standard deviation $\sigma_{\boldsymbol{L}}$. If the length of the lead time is independent of the demand per unit time during the lead time, then (lead time variable):

* $E(\boldsymbol{X}) = E(\boldsymbol{L})E(\boldsymbol{D})$
* $var(\boldsymbol{X}) = E(\boldsymbol{L})\big(var(\boldsymbol{D})\big) + E(\boldsymbol{D})^2(var(\boldsymbol{L}))$

A **cycle** is defined to be the time interval between any two instants at which an order is received.

Determination of reorder point: the back-ordering case:
The situation is which all demand must eventually be met and no sales are lost is called the **back-ordered case**, for which we show how to determine the reorder point and order quantity that minimize annual expected cost.
We assume each unit is purchased for the same price, so purchasing costs are fixed. Define *TC(q, r)* = expected annual cost (excluding purchasing cost) incurred if each order is for *q* units and is placed when the reorder point is *r*. Then ***TC(q, r) = (expected annual holding cost) + (expected annual ordering cost) + (expected annual cost due to shortages)***. To determine the optimal reorder point and order quantity, we assume that the average number of back orders is small relative to the

average on-hand inventory level. Then $I(t) = OHI(t) - B(t)$ yields $\boldsymbol{Expected\ value\ of\ I(t)} \cong$ $\boldsymbol{expected\ value\ of\ OHI(t).}$
The $expected\ annual\ holding\ cost = h(expected\ value\ of\ OHI)$,
then $expected\ annual\ holding\ cost = h(expected\ value\ of\ I(t)$

$expected\ value\ of\ I(t) during\ a\ cycle = \frac{1}{2}((expected\ value\ of\ I(t)\ at\ beginning\ of\ cycle) +$ $(expected\ value\ of\ I(t)\ at\ end\ of\ a\ cycle))$

At the end of a cycle, the inventory level will equal the inventory level at the reorder point (*r*) less the demand **X** during lead time. So, $Expected\ value\ of\ I(t)\ at\ end\ of\ a\ cycle = r - E(\boldsymbol{X})$.
At the beginning of a cycle, the inventory level at the end of the cycle is augmented by the arrival of an order of size *q*. So, $expected\ value\ of\ I(t)\ at\ beginning\ of\ cycle = r - E(\boldsymbol{X}) + q$

$$expected\ value\ of\ I(t)\ during\ a\ cycle = \frac{1}{2}(r - E(\boldsymbol{X}) + r - E(\boldsymbol{X}) + q) = \frac{q}{2} + r - E(\boldsymbol{X})$$

$expected\ annual\ holding\ cost \cong h(\frac{q}{2} + r - E(\boldsymbol{X})$

To determine the expected annual cost due to stockouts or back orders, we must define
$\boldsymbol{B_r} =$ random variable representing the number of stockouts or back orders during a cycle if the reorder point is *r*.

- $expected\ annual\ shortage\ cost = \left(\frac{expected\ shortage\ cost}{cycle}\right)\left(\frac{expected\ cycles}{year}\right)$
- $\frac{expected\ shortage\ cost}{cycle} = c_B E(\boldsymbol{B_r})$
- $an\ average\ of\ \frac{E(\boldsymbol{D})}{q}\ orders\ will\ be\ placed\ each\ year$
- $\frac{expected\ shortage\ cost}{year} = \frac{c_B E(\boldsymbol{B_r}) E(\boldsymbol{D})}{q}$
- $expected\ annual\ order\ cost = K\left(\frac{expected\ orders}{year}\right) = \frac{KE(\boldsymbol{D})}{q}$
- $\boldsymbol{TC(q,r) = h\left(\frac{q}{2} + r - E(X)\right) + \frac{c_B E(B_r) E(D)}{q} + \frac{KE(D)}{q}}$

We assume that the optimal order quantity q^* may be adequately approximated by the EOQ. Given a value *q* for the order quantity, we now show how marginal analysis can be used to determine a reorder point r^* that minimizes *TC(q, r)*.
If we assume a given value of *q*, the expected annual ordering cost is independent of *r*. thus, in determining a value of *r* that minimizes *TC(q, r)*, we may concentrate on minimizing the sum of the expected annual holding an shortage cost.

If we increase the reorder point from $r\ to\ r + \Delta$, the expected annual holding cost will increase by
$\boldsymbol{h\left(\frac{q}{2} + r + \Delta - E(X)\right) - h\left(\frac{q}{2} + r - E(X)\right) = h\Delta.}$
If we increase the reorder point from $r\ to\ r + \Delta$, expected annual stockout costs will be reduced, because of the fact that during any cycle in which lead time demand is at least r, the number of stockouts during the cycle will be reduced by Δ units. In order words, increasing the reorder point from $r\ to\ r + \Delta$ will reduce stockout cost by $c_b\Delta$ during a fraction $P(X \geq r)$ of all cycles. Since there are an average of $\frac{E(\boldsymbol{D})}{q}$ cycles per year, increasing the reorder point from $r\ to\ r + \Delta$ will reduce expected annual stockout cost by
$\frac{\Delta E(\boldsymbol{D}) c_{\boldsymbol{B}} P(\boldsymbol{X} \geq r)}{q}$. Observe that as r increases, $P(\boldsymbol{X} \geq r)$ decreasing, the expected reduction in expected annual shortage cost resulting from increasing the reorder point by Δ will decrease. Let r^* be the value of *r* for which marginal benefit equals marginal cost, of

$\frac{\Delta \boldsymbol{E}(\boldsymbol{D})\boldsymbol{c_B}\boldsymbol{P}(\boldsymbol{X}\geq \boldsymbol{r}^*)}{\boldsymbol{q}} = \boldsymbol{h}\Delta$

$\boldsymbol{P}(\boldsymbol{X} \geq \boldsymbol{r}^*) = \frac{\boldsymbol{hq}}{\boldsymbol{c_B}\boldsymbol{E}(\boldsymbol{D})}$

In summary, if we assume that the order quantify can be approximated by $\boldsymbol{EOQ} = \left(\frac{\boldsymbol{2KE}(\boldsymbol{D})}{\boldsymbol{h}}\right)^{\frac{1}{2}}$, then we have the reorder point r^* and the order quantity q^* for the back-ordered case: $\boldsymbol{q}^* = \left(\frac{\boldsymbol{2KE}(\boldsymbol{D})}{\boldsymbol{h}}\right)^{\frac{1}{2}}$

$\boldsymbol{P}(\boldsymbol{X} \geq \boldsymbol{r}^*) = \frac{\boldsymbol{hq}^*}{\boldsymbol{c_B}\boldsymbol{E}(\boldsymbol{D})}$

If $\frac{hq^*}{c_B E(\boldsymbol{D})} > 1$ then $P(\boldsymbol{X} \geq r^*) = \frac{hq^*}{c_B E(\boldsymbol{D})}$ will have no solution, and holding cost is prohibitively high relative to the stockout cost.

Determination of reorder point: the lost sales case:
We now assume that all stockouts result in lost sales and that a cost of c_{LS} dollar is incurred for each lost sale. As in the back-ordered case, we assume that the optimal order quantity can be adequately approximated by the EOQ and attempt to use marginal analysis to determine the optimal reorder point r^*. The optimal order quantity q^* and the reorder point r^* for the lost sales case are

$$\boldsymbol{q}^* = \left(\frac{\boldsymbol{2KE}(\boldsymbol{D})}{\boldsymbol{h}}\right)^{\frac{1}{2}}$$

$$\boldsymbol{P}(\boldsymbol{X} \geq \boldsymbol{r}^*) = \frac{\boldsymbol{hq}^*}{\boldsymbol{hq}^* + \boldsymbol{c_{LS}}\boldsymbol{E}(\boldsymbol{D})}$$

The key to the derivation of these formula's is to realize that expected inventory in lost sales case = (expected inventory in back-ordered case) + (expected number of shortages per cycle). This equation follows because in the lost sales case, we find that during each cycle, an average of (expected shortages per cycle) fewer orders will be filled from inventory, thereby raising the average inventory level by an amount equal to expected shortages per cycle. The lost sales assumption will yield a lower stockout probability than the back-ordered assumption.

A continuous review inventory policy, in which we order a quantity *q* whenever our inventory level reaches a reorder level *r*, is often called an **(r, q) policy**. An *(r, q)* policy is also called a **two-bin policy**, because it can easily be implemented by using two bins to store an item.

In our derivation of the best *(r, q)* policy, we assumed that an order could be placed exactly at the point when the inventory level reached the reorder point *r*. We used this assumption to compute the expected inventory level at the beginning and end of a cycle. Suppose that a demand for more than one unit can arrive at a particular time. Then an order may be triggered when the inventory is less than *r*, and our computation of expected inventory level at the end and beginning of a cycle is then incorrect. We see that if demands of size greater than one unit can occur at a point in time, then the *(r, q)* model may not yield a policy that minimizes expected annual cost.
In such situation, it has been shown that an **(s, S) policy** is optimal. To implement an *(s, S)* policy, we place an order whenever the inventory level is less than or equal to *s*. The size of the order is sufficient to raise the inventory level to *S* (assuming zero lead time).
Set $\boldsymbol{S} - \boldsymbol{s} = \boldsymbol{equal\ to\ the\ economic\ order\ quantiy\ q}$. Then set ***s* equal to the reorder point *r***.
Finally, we obtain $\boldsymbol{S} = \boldsymbol{r} + \boldsymbol{q}$.

Chapter 18 Deterministic Dynamic Programming

Description of Dynamic Lot-Size Model:

1. Demand d_t during period *t* ($t = 1, 2, \ldots, T$) is known at the beginning of period 1.
2. Demand for period *t* must be met on time from inventory or from period *t* production. The cost *c(x)* of production *x* units during any period is given by $c(0) = 0$, and for $x > 0, c(x) = K + cx$, where *K* is a fixed cost for setting up production during a period, and *c* is the variable per-unit cast of production.
3. At the end of period *t*, the inventory level i_t is observed, and a holding cost hi_t is incurred. We let i_0 denote the inventory level before period 1 production occurs.
4. The goal is to determine a production level x_i for each period *t* that minimizes the total cost of meeting (on time) the demands for periods 1, 2, …, T.
5. There is a limit c_t placed on period *t*'s ending inventory.
6. There is a limit r_t placed on period *t*'s production.

Wagner and Whitin developed a recursion that can be used to determine an optimal production policy. We assume that the initial inventory level is zero. Define f_t as the minimum cost incurred during periods *t, t + 1, …, T*, given that at the beginning of period *t,* the inventory level is zero. Then $f_1, f_2, \ldots, f_T$ must satisfy

$$\boldsymbol{f_t = \min_{j=0,1,2,\ldots,T-t} (c_{tj} + f_{t+j+1})}$$

Where $f_{T+1} = 0 \; and \; c_{ij}$ is the total cost incurred during periods $t, t+1, \ldots, t+j$ if production during period *t* is exactly sufficient to meet demands for periods $t, t+1, \ldots, t+j$. Thus, $\boldsymbol{c_{tj} = K + c(d_t + d_{t+1} + \cdots + d_{t+j}) + h(jd_{t+j} + (j-1)d_{t+j-1} + \cdots + d_{t+1})}$ where *K* is the setup cost incurred during period *t*, $c(d_t + d_{t+1} + \cdots + d_{t+j})$ is the variable production cost incurred during period *t*, and $h(jd_{t+j} + (j-1)d_{t+j-1} + \cdots + d_{t+1})$ is the holding cost incurred during periods $t, t+1, \ldots, t+j$.

The **Silver-Meal (S-M)** heuristic involves less work than the Wagner-Whitin algorithm and can be used to find a near-optimal production schedule. The S-M heuristic is based an the fact that our goal is to minimize average cost per period. Suppose we are at the beginning of period 1 and are trying to determine how many periods of demand should be satisfies by period 1's production. During period 1, if we produce an amount sufficient to meet demand for the next *t* periods, then a cost of $\boldsymbol{TC(t) = K + HC(t)}$ will be incurred. Here *HC(t)* is the holding cost incurred during the next *t* periods if production during the current period is sufficient to meet demand for the next *t* periods.

Let $\boldsymbol{AC(t) = \frac{TC(t)}{t}}$ be the average per-period cost incurred during the next t periods. In most situations, an integer t^* can be found such that for $\boldsymbol{t < t^*, AC(t+1) \leq AC(t)}$ **and** $\boldsymbol{AC(t^* + 1 \geq AC(t^*)}$. The S-M heuristic recommends that period 1's production should satisfy the demand for periods $1, 2, \ldots, t^*$. Since t^* is a local minimum for AC(t), it seems reasonable that producing $d_1 + d_2 + \cdots + d_{t^*}$ units during period 1 will come close to minimizing the average per-period cost incurred during periods $1, 2, \ldots, t^*$. Next we apply the S-M heuristic while considering period $t^* + 1$ as the initial period. We find that during period $t^* + 1$, the demand for period T has been produced.

Chapter 20 Queuing Theory

The **input process** is usually called the **arrival process**. Arrivals are called **customers**. In all models that will discuss, we assume that no more than one arrival can occur at a given instant. For a case like a restaurant, this is a very unrealistic assumption. If more than one arrival can occur at a given instant, we say that **bulk arrivals** are allowed. Usually, we assume that the arrival process is unaffected by the number of customers present in the system.
There are two common situations in which the arrival process may depend on the number of customers present. The first occurs when arrivals are drawn from a small population. Models in which arrivals are drawn from a small population are called **finite source models**. Another situation in which the arrival process depends on the number of customers present occurs when the rate at which customers arrive at the facility decreases when the facility becomes too crowded. If a customers arrives but fails to enter the system, we say that the customer has **balked**.
If the arrival process is unaffected by the number of customers present, we usually describe it by specifying a probability distribution that governs the time between successive arrivals.

To describe the **output process (service process)** of a queuing system, we usually specify a probability distribution – the **service time distribution** – which governs a customer's service time. In most cases, we assume that the service time distribution is independent of the number of customers present. We study to arrangements of servers: **servers in parallel** and **serves in series**. Servers are in parallel if all servers provide the same type of service and a customer need only pass trough one server to complete service. Servers are in series if a customer must pass through several servers before completing service.

The **queuing discipline describes** the method used to determine the order in which customers are served. The most common queue discipline is the **FCFS discipline** (first come, first served), in which customers are served in the order of their arrival. Under the **LCFS discipline** (last come, first served), the most recent arrivals are the first to enter service. Sometimes the order in which customers arrive has no effect on the order in which they are served. This would be the case if the next customer to enter service is randomly chosen from those customers waiting for service. Such a situation is referred to as the **SIRO discipline** (service in random order). Finally, we consider **priority queuing disciplines**. A priority discipline classifies each arrival into one of several categories. Each category is then given a priority level, and within each priority level, customers enter service on an FCFS basis.

Another factor that has an important effect on the behaviour of a queuing system is the method that customers use to determine which line to join.

Modelling the arrival process:
We assume that at most one arrival can occur at a given instant of time. We define t_i to be the time at which the *i*th customer arrives. For $i \geq 1$, we define $T_i = t_{i+1} - t_i$ to be the *i*th interarrival time. In modelling the arrival process, we assume that the $T_i's$ are independent, continuous random variables described by the random variable **A**. The assumption that each interarrival time is governed by the same random variable implies that the distribution of arrivals is independent of the time of day or the day of the week. This is the assumption of stationary interarrival times.
We assume that **A** has a density function *a(t)*. For small $\Delta t, P(t \leq A \leq t + \Delta \text{t})$ is approximately $\Delta ta(t)$. Of course, a negative interarrival time is impossible. This allows us to write $\boldsymbol{P(A \leq c) = \int_0^c a(t)dt} \ \boldsymbol{and\ P(A > c) = \int_c^\infty a(t)dt}$.
We define $\frac{1}{\lambda}$ to be the mean or average interarrival time. Without loss of generality, we assume that time is measured in units of hours. Then $\frac{1}{\lambda}$ will have units of hours per arrival. We may compute $\frac{1}{\lambda}$ from a(t) by using the following equation: $\boldsymbol{\frac{1}{\lambda} = \int_0^\infty ta(t)dt}$.

We define λ to be the arrival rate, which will have units of arrivals per hour. In most applications of queuing, an important question is how to choose **A** to reflect reality and still be computationally tractable. The most common choice for **A** is the **exponential distribution**. An exponential distribution with parameter λ has a density $a(t) = \lambda e^{-\lambda t}$. Using $\frac{1}{\lambda} = \int_0^\infty ta(t)dt$ and integration by part, we can show that the average of mean interarrival time (E(**A**)) is given by $\boldsymbol{E(A) = \frac{1}{\lambda}}$.
$var(A) = E(A^2) - \big(E(A)\big)^2 \rightarrow \boldsymbol{var(A) = \frac{1}{\lambda^2}}$

If **A** has an exponential distribution, then for all nonnegative values of *t* and *h*,
$\boldsymbol{P(A > t + h|A \geq t) = P(A > h).}$ A density that satisfies this formula is said to have the no-memory property. Suppose we are told that there has been no arrival for the last *t* hours ($\boldsymbol{A} > t$) and are asked what the probability is that there will be no arrival during the next *h* hours ($\boldsymbol{A} > t + h$). Then $P(A > t + h|A \geq t) = P(A > h)$ implies that this probability does not depend on the value of *t*, and for all values of *t*, this probability equals $P(\boldsymbol{A} > h)$. So, if we know that at least t time units have elapsed since the last arrival occurred, then the distribution of the remaining time until the next arrival (*h*) does not depend on *t*.
The no-memory property of the exponential distribution is important, because it implies that if we want to know the probability distribution of the time until the next arrival, then it does not matter how long it has been since the last arrival.

If interarrival time are exponential, the probability distribution of the number of arrivals occurring in any time interval of length *t* is given by the following theorem:
Interarrival times are exponential with parameter λ if and only if the number of arrivals to occur in an interval of length t follows a Poisson distribution with parameter λt.

A discrete random variable **N** has a Poisson distribution with parameter λ if, for $n = 0, 1, 2, \ldots$,
$P(\boldsymbol{N} = n) = \frac{e^{(-\lambda)}\lambda^n}{n!}$.
$E(\boldsymbol{N}) = var(\boldsymbol{N}) = \lambda$

If we define $\boldsymbol{N}_t$ to be the number of arrivals to occur during any time interval of length t, the theorem states that $P(\boldsymbol{N}_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$ and $E(\boldsymbol{N}_t) = var(\boldsymbol{N}_t) = \lambda t$. An average of λt arrivals occur during a time interval of length t, so λ may be though of as the average number of arrivals per unit time, or the arrival rate.

There are two assumptions required for interarrival time to be exponential:

1. Arrivals defined on nonoverlapping time intervals are independent.
2. For small Δt, the probability of one arrival occurring between time t and $t + \Delta t$ is $\lambda\Delta t + o(\Delta t)$, where $o(\Delta t)$ refers to any quantity satisfying $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$.

Also, the probability of no arrival during the interval between t and $t + \Delta t$ is $1 - \lambda\Delta t + o(\Delta t)$, and the probability of more than one arrival occurring between t and $t + \Delta t$ is $o(\Delta t)$.

If the assumptions hold, then $\boldsymbol{N}_t$ follows a Poisson distribution with parameter λt, and interarrival times are exponential with parameter λ; that is, $a(t) = \lambda e^{-\lambda t}$.

If interarrival times do not appear to be exponential, they are often modelled by an **Erlang distribution**. An Erlang distribution is a continuous random variable (**T**) whose density function *f(t)* is specified by two parameters: a rate parameter *R* and a shape parameter *k* (positive integer). Given values of *R* and *k*, the Erlang density has the following probability density function: $\boldsymbol{f(t)} = \frac{\boldsymbol{R(Rt)^{k-1}e^{-Rt}}}{\boldsymbol{(k-1)!}} \quad (\boldsymbol{t \geq 0})$

* $\boldsymbol{E(T) = \frac{k}{R}}$
* $\boldsymbol{var(T) = \frac{k}{R^2}}$

Modelling the service process:
We assume that the service times of different customers are independent random variables and that each customer's service time is governed by a random variable **S** having a density function *s(t)*. We let $\frac{1}{\mu}$ be the mean service time for a customer. Of course, $\frac{1}{\mu} = \int_0^\infty ts(t)dt$.
The variable $\frac{1}{\mu}$ will have units of hours per customer, so μ has units of customers per hour. We call μ the service rate.
If we can model a customer's service time as an exponential random variable, we van determine the distribution of a customer's remaining service time without having to keep track of how long the customer has been in service. If service times follow an exponential density $s(t) = \mu e^{-\mu t}$, then a customer's mean service time will be $\frac{1}{\mu}$.
Actual service times may not be consistent with the no-memory property. For this reason, we often assume that *s(t)* is an Erlang distribution with shape parameter *k* and rate parameter $k\mu$.
In certain situations, interarrival of service times may be modelled as having zero variance; in this case, interarrival of service times are considered to be **deterministic**.

The Kendall-Lee notation for queuing systems:
The **first characteristic specifies the nature of arrival process**. The following standard abbreviations are used:

* *M* = interarrival timer are independent, identically distributed (iid) random variables having an exponential distribution.
* *D* = interarrival times are iid and deterministic.
* $E_k =$ interarrival times are iid Erlangs with shape parameter *k*.
* GI = interarrival times are idd and governed by some general distribution.

The **second characteristic specifies the nature of the service times**:

* *M* = service times are idd and exponentially distributed.
* *D* = service times are idd and deterministic.
* $E_k =$ service times are idd Erlangs with shape parameter *k*.
* *G* = service times are idd and follow some general distribution.

The **third characteristic is the number of parallel servers**. The **fourth characteristic describes the queue discipline**:

* **FCFS** = First come, first served.
* **LCFS** = Last come, first served.
* **SIRO** = Service in random order.
* **GD** = General queue discipline.

The **fifth characteristic specifies the maximum allowable number of customers in the system** (including customers who are waiting and customers who are in service).
The sixth characteristic gives the size of the population form which customers are drawn. Unless the number of potential customers is of the same order of magnitude as the number of serves, the population size is considered to be infinite.

The waiting time paradox:

Suppose the time between the arrival of buses at the student centre is exponentially distributes, with a mean of 60 minutes. If we arrive at a random chosen instant, what is the average amount of time that we will have to wait for a bus?
The **no-memory property of the exponential distribution** implies that no matter how long it had been since the last bus arrived, we would still expect to wait an average of 60 minutes until the next bus arrived. This answer is correct, but is appears to be contradicted by the following argument. On the average, somebody who arrives at a random time should arrive in the middle of a typical interval between arrivals of successive buses. If we arrive at the midpoint of a typical interval and the average time between buses is 60 minutes, then we should have to wait 30 minutes for the next bus. This is correct because the typical interval between buses is longer than 60 minutes. The reason for this is anomaly is that we are more likely to arrive during a longer interval than a shorter interval.
In general, it can be shown that if **A** is the random variable for the time between buses, then the average time until the next bus is given by $\frac{1}{2}\left(E(A)+\frac{var(A)}{E(A)}\right)$.

The M/M/1/GD/∞/∞ queuing system and the queuing formula $L = \lambda W$:
The **M/M/1/GD/∞/∞ queuing system** has exponential interarrival times (we assume that the arrival time is λ) and a single server with exponential serves times (we assume that each customer's service time is exponential with rate μ).
We define $\rho = \frac{\lambda}{\mu}$ and **ρ** is called the **traffic intensity** of the queuing system.

* **π_j = the steady-state probability that *j* customers will be present.**
* $\pi_0 = 1 - \rho \qquad for\ 0 \le \rho \le 1$
* $\pi_j = \rho^j(1-\rho) \qquad for\ 0 \le \rho \le 1$

We assume that $\rho < 1$, ensuring that a steady-state probability distribution does exist.

* $L = (1-\rho)\sum_{j=0}^{j=\infty} j\rho^j = \frac{\rho}{1-\rho}\left(=\frac{\lambda}{\mu-\lambda}\right)$
* $L_q = L - \rho = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho}\ \left(=\frac{\lambda^2}{\mu(\mu-\lambda)}\right)$
* $L_s = \rho$
* $L = L_s + L_q$

We define *W* as the expected time a customer spends in the queuing system, including the time in line plus time in service, and W_q as the expected time a customer spend waiting in line. Both *W* and W_q are computed under the assumption that the steady state has been reached. By using a powerful result known as **Little's queuing formula**, *W* and W_q may easily be computed from *L* and L_q. We first define the following quantities:

* λ = average number of arrivals entering the system per unit time.
* **L** = average number of customers present in the queuing system.
* L_q = average number of customers waiting in line.
* L_s = average number of customers in service.
* W= average time a customer spends in the system.
* W_q = average time a customer spends in line.
* W_s = average time a customer spends in service.

For any queuing system in which a steady-state distribution exists, the following relations hold:

* $L = \lambda W$
* $L_q = \lambda W_q$
* $L_s = \lambda W_s$

L is expressed is terms of number of customers, λ is expressed is terms of customers per hour and *W* is expressed in hours. Thus, λW has the same units as *L*.

* $L = \frac{\rho}{1-\rho}$
* $W = \frac{L}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu-\lambda}$
* $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$
* $W_q = \frac{L_q}{\lambda} = \frac{\lambda}{\mu(\mu-\lambda)}$

Problems in which a decision maker must choose between alternative queuing systems are called **queuing optimization problems**.

The queuing formula $L = \lambda W$ is very general and can be applied to many situations that do not seem to be queuing problems.

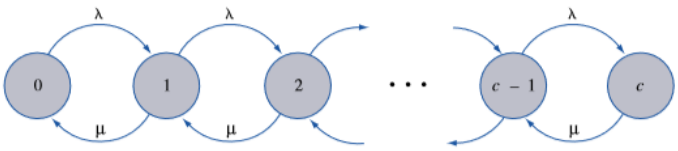
The **M/M/1/GD/c/∞ queuing system** is a M/M/1/GD/∞/∞ system except from the fact that when *c* customers are present, all arrivals are turned away and are forever lost to the system.

* $\lambda_j = \lambda \quad (j = 0, 1, \dots, c-1)$
* $\lambda_c = 0$
* $\mu_0 = 0$
* $\mu_j = \mu \quad (j = 1, 2, \dots, c)$

The steady state probabilities for the M/M/1/GD/c/∞ model ae given by:

* $\pi_0 = \frac{1-\rho}{1-\rho^{c+1}}$
* $\pi_j = \rho^j \pi_0 \quad (j = 1, 2, \dots, c)$
* $\pi_j = 0 \quad (j = c+1, c+2, \dots)$

FIGURE 14
Rate Diagram for
M/M/1/GD/c/∞
Queuing System



$L = \frac{\rho(1-(c+1)\rho^c + c\rho^{c+1})}{(1-\rho^{c+1})(1-\rho)}$ **when** $\lambda \neq \mu$

$L = \frac{c}{2}$ *and* $\pi_j = \frac{1}{c+1}$ **when** $\lambda = \mu$

$L_q = L - L_s$

λ represents the average number of customers per unit time who actually enter the system. In our finite capacity model, an average of λ arrivals per unit time arrive, but $\lambda\pi_c$ of these arrivals find the system filled to capacity and leave. Thus, an average of $\lambda - \lambda\pi_c = \lambda(1-\pi_c)$ arrivals per unit time will actually enter the system. This gives use:

$W = \frac{L}{\lambda(1-\pi_c)}$ **and** $W_q = \frac{L_q}{\lambda(1-\pi_c)}$.

For an M/M/1/GD/c/∞ system, a steady state exist even if $\lambda \geq \mu$. This is because, even if $\lambda \geq \mu$, the finite capacity of the system prevents the number of people in the system from 'blowing up'.

The M/M/s/GD/∞/∞ queuing system:

We now consider the **M/M/s/GD/∞/∞ system**. We assume that interarrival times are exponential (with rate λ), service times are exponential (with rate μ), and there is a single line of customers waiting to be served at one of ***s* parallel servers**. If $j \leq s$ customers are present, then all *j* customers are in service; if $j > s$ customers are present, then all s servers are occupied, and $j - s$ customers are waiting in line. Any arrival who finds an idle server enter service immediately, but an arrival who does not find an idle server joins the queue of customers awaiting service.

* $\lambda_j = \lambda \quad (j = 0, 1, 2, \dots)$

If *j* servers are occupied, then service completions occur at a rate $j\mu$.

Whenever *j* costumers are present, $\min(j, s)$ servers will be occupied. Thus, $\mu_j = \min(j, s)\mu$.

* $\boldsymbol{\lambda_j = \lambda} \qquad (\boldsymbol{j = 0, 1, \dots})$
* $\boldsymbol{\mu_j = j\mu} \qquad (\boldsymbol{j = 0, 1, \dots, s})$
* $\boldsymbol{\mu_j = s\mu} \qquad (\boldsymbol{j = s+1, s+2, \dots})$

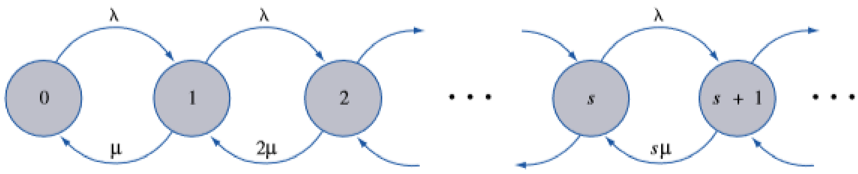
We define $\boldsymbol{\rho = \frac{\lambda}{s\mu}}$. For $\rho < 1$ are the following steady-state probabilities:

* $\boldsymbol{\pi_0 = \frac{1}{\sum_{i=0}^{i=(s-1)}\left(\frac{(s\rho)^i}{i!} + \frac{(s\rho)^s}{s!(1-\rho)}\right)}}$
* $\boldsymbol{\pi_j = \frac{(s\rho)^j \pi_0}{j!}} \qquad (\boldsymbol{j = 1, 2, \dots, s})$
* $\boldsymbol{\pi_j = \frac{(s\rho)^j \pi_0}{s! s^{j-s}}} \qquad (\boldsymbol{j = s, s+1, s+2, \dots})$

If $\rho \geq 1$, no steady-state exists. In other words, if the arrival rate is at least as large as the maximum possible service rate ($\lambda \geq s\mu$), the system 'blows up'.

The steady-state probability that all servers are busy is given by $\boldsymbol{P(j \geq s) = \frac{(s\rho)^s \pi_0}{s!(1-\rho)}}$.

FIGURE 16
Rate Diagram for *M/M/s/GD/∞/∞* Queuing System



* $\boldsymbol{L_q = \frac{P(j \geq s)\rho}{1-\rho}}$
* $\boldsymbol{W_q = \frac{L_q}{\lambda} = \frac{P(j \geq s)}{s\mu - \lambda}}$
* $\boldsymbol{L = L_q + L_s}$
* $\boldsymbol{W_s = \frac{1}{\mu}}$
* $\boldsymbol{L_s = \frac{\lambda}{\mu}}$
* $\boldsymbol{L = L_q + \frac{\lambda}{\mu}}$
* $\boldsymbol{W = \frac{L}{\lambda} = \frac{L_q}{\lambda} + \frac{1}{\mu} = W_q + \frac{1}{\mu} = \frac{P(j \geq s)}{s\mu - \lambda} + \frac{1}{\mu}}$

In addition to a customer's expected time in the system, the distribution of a customer's waiting time is of interest. To determine the probability that the waiting time is too long, we need to know the distribution of a customer's waiting time. For an M/M/s/FSFS/∞/∞ queuing system, it can be shown that

$$\boldsymbol{P(W > t) = e^{-\mu t}\left(1 + P(j \geq s)\frac{1 - \exp(-\mu t(s - 1 - s\rho))}{s - 1 - s\rho}\right)}$$

$$\boldsymbol{P(W_q > t) = P(j \geq s)\exp(-s\mu(1-\rho)t)}$$

Finite source models: The machine repair model:

With the exception of the M/M/1/GD/c/∞ model, all the models we have studied have displayed arrival rate that were independent of the state of the system. As discussed previously, there are two situations where the assumption of the state-independent arrival rates may be invalid:

1. If customers do not want to buck long lines, the arrival rate may be a decreasing function of the number of people present in the queuing system.
2. If arrivals to a system are drawn from a small population, the arrival rate may greatly depend on the state of the system.

Models in which arrivals are drawn from a small population are called **finite source models**. An important finite source model is **the machine repair model**.

In the machine repair problem, the system consists of *K* machines and *R* repair people. At any instant in time, a particular machine is in either good or bad condition. The length of time that a machine remains in good condition follows an exponential distribution with rate λ. Whenever a machine breaks down, the machine is sent to a repair centre consisting of *R* repair people. The repair centre services the broken machines as if they were arriving at an M/M/R/GD/∞/∞ system.

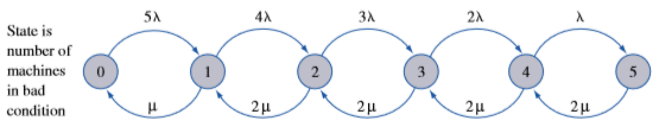
Thus, if $j \leq R$ machines are in bad condition, a machine that has just broken down will immediately be assigned for repair; if $j > R$ machines are broken, $j - R$ machines will be waiting in a single line for a repair worker to become idle. The time it takes to complete repairs on a broken machine is assumed exponential with rate μ (or mean repair time $\frac{1}{\mu}$). Once a machine is repaired, it returns to good condition and is again susceptible to breakdown.

The model just described may be expressed as an M/M/R/GD/K/K model. The first *K* indicates that at any time, no more than *K* customers may be present, and the second *K* indicates that arrivals are drawn from a finite source of size *K*.

TABLE 8
Possible States in a Machine Repair Problem When $K = 5$ and $R = 2$

State	No. of Good Machines	Repair Queue	No. of Repair Workers Busy
0	*G G G G G*		0
1	*G G G G*		1
2	*G G G*		2
3	*G G*	*B*	2
4	*G*	*B B*	2
5		*B B B*	2

FIGURE 22
Rate Diagram for *M/M/R/GD/K/K* Queuing System When $R = 2$, $K = 5$



When the state is *j*, there are $K - j$ machines in good condition. Since each machine breaks down at rate λ, the total rate at which breakdowns occur when the state is *j* is $(K - j)\lambda$.

When the state is j, $\min(j, R)$ repair people will be busy. Since each occupied repair worker completes repairs at rate μ, μ_j is given by:

* $\mu_j = j\mu \qquad (j = 0, 1, \ldots, R)$
* $\mu_j = R\mu \qquad (j = R + 1, R + 2, \ldots, K)$

We define $\boldsymbol{\rho} = \frac{\lambda}{\mu}$, yields the following steady-state probability distribution:

* $\boldsymbol{\pi_0 + \pi_1 + \cdots + \pi_k = 1}$
* $\boldsymbol{\pi_j = \binom{K}{j} \rho^j \pi_0 \qquad (j = 0, 1, \ldots, R)}$
* $\boldsymbol{= \frac{\binom{K}{j}\rho^j j! \pi_0}{R! R^{j-R}} \qquad (j = R + 1, R + 2, \ldots, K)}$
* $\boldsymbol{\binom{K}{j} = \frac{K!}{j!(K-j)!}}$

Using the steady-state probabilities, we can determine the following quantities of interest:

* $\boldsymbol{L}$ = expected number of broken machines.

* L_q = expected number of machines waiting for service.
* W = average time a machine spends broken (down time).
* W_q = average time a machine spends waiting for service.

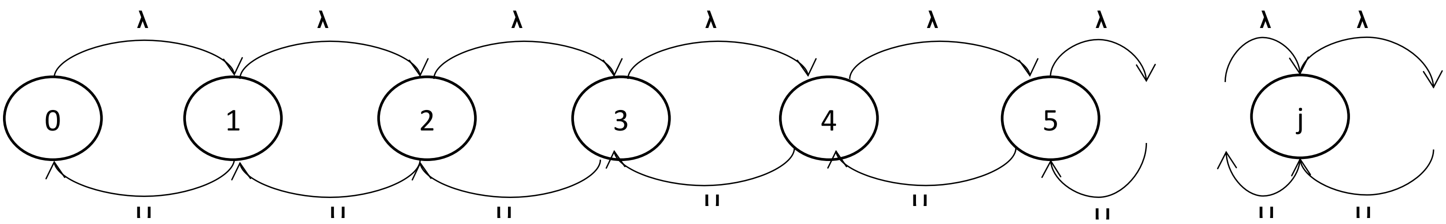
* $L = \sum_{j=0}^{j=K} j\pi_j$
* $L_q = \sum_{j=R}^{j=K} (j - R)\pi_j$

Since the arrival rate is state-dependent, the average number of arrivals per unit time is given by $\bar{\lambda}$, where

* $\bar{\lambda} = \sum_{j=0}^{j=K} \pi_j \lambda_j = \sum_{j=0}^{j=K} \lambda(K - j)\pi_j = \lambda(K - j)$

* $W = \frac{L}{\bar{\lambda}}$
* $W_q = \frac{L_q}{\bar{\lambda}}$

Example M/M/1 model:
$\lambda = 6\ and\ \mu = 8, so\ \rho = \frac{3}{4}$



* **Define P_n (n = 0,1,2,…) = probability of having n customers in the system (n is called the state) = fraction of time that n customers are present in the system**
* **λP_n is the average number of transitions per hour from state n to state n+1**
* **μP_n is the average number of transitions per hour from state n to state n-1**

$P_2 = \pi_2 = \frac{1}{2}$
How many times per hour from $j = 2\ to\ j = 3$?
$\lambda P_2 = 6 * \frac{1}{2} = 3$

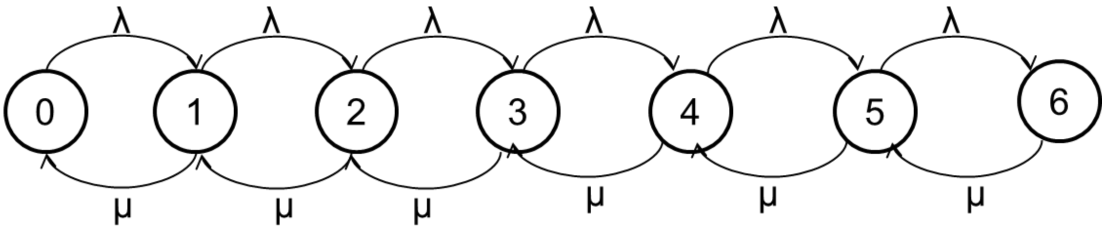
$P_3 = \pi_3 = \frac{1}{3}$
How many times per hour from $j = 3\ to\ j = 4$?
$\lambda P_3 = 6 * \frac{1}{3} = 2$

How many times per hour from $j = 2\ to\ j = 1$?
$\mu P_2 = 8 * \frac{1}{2} = 4$

Number of transitions out of stage *j* is the number of transitions into stage *j*.

* $(\lambda + \mu)P_n = \lambda P_{n-1} + \mu P_{n+1}; n = 1, 2, 3, \ldots$
* $\lambda P_0 = \mu P_1$
* $(\lambda - \mu) * \frac{\lambda}{\mu} P_0 = \lambda P_0 + \mu P_2 \to P_2 = \frac{\lambda^2}{\mu^2} P_0 = \rho^2 P_0$
* $P_n = \rho^n P_0; n = 1, 2, 3, \ldots$
* $utilization = 1 - P_0 = \rho$

Example M/M/1/6 model:



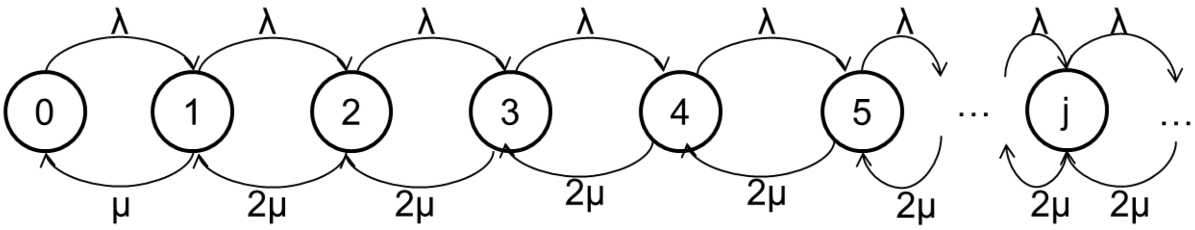
* $\lambda P_5 = \mu P_6$
* $1 = \sum_{i=0}^{6} P_1 = P_0 \sum_{i=0}^{6} \rho^i = P_0 \frac{1-\rho^7}{1-\rho}$
* $L = 0P_0 + 1P_1 + 2P_2 + 3P_3 + 4P_4 + 5P_5 + 6P_6$
* $\bar{\lambda} = \lambda(1 - P_6)$
* $W = \frac{L}{\bar{\lambda}}$

Important formula's:

* $\boldsymbol{\sum_{j=0}^{\infty} \rho^j = 1 + \rho + \rho^2 + \rho^3 + \cdots = \frac{1}{(1-\rho)} \; if \; \rho < 1}$
* $\boldsymbol{\sum_{j=0}^{\infty} j\rho^j = \rho + 2\rho^2 + 3\rho^3 + \cdots = \frac{\rho}{(1-\rho)^2} \; if \; \rho < 1}$
* $\boldsymbol{\sum_{j=0}^{n} \rho^j = \frac{1-\rho^{n+1}}{1-\rho}}$

Example M/M/s model:

* 9 requests for help on average per hour (λ = 9)
* Average service time of an employee is 6 minutes (μ = 10)
* 2 employees available (s = 2)



* $\lambda P_0 = \mu P_1$
* $(\lambda + \mu)P_1 = \lambda P_0 + 2\mu P_2$
* $(\lambda + 2\mu)P_i = \lambda P_{i-1} + 2\mu P_{i+1}$,i = 2, 3,…
* $P_1 = \rho P_0$; $P_2 = \frac{1}{2}\rho^2 P_0$; $P_3 = \frac{1}{2^2}\rho^3 P_0$; $P_j = \frac{1}{2^{j-1}}\rho^j P_0, \; j = 1, \, 2, \, \ldots$
* $1 = \sum_{j=0}^{\infty} P_j = P0 + P_0 \sum_{j=1}^{\infty} \frac{1}{2^{j-1}} \rho^j = P_0 + P_0 . 2 \sum_{j=1}^{\infty} \left(\frac{\rho}{2}\right)^j = P_0 + P_0 \left[2 \frac{1}{1-\rho/2} - 2\right] = -P_0 + P_0 \frac{4}{2-\rho} = \frac{2+\rho}{2-\rho} P_0 \; \rightarrow \; P_0 = \frac{2-\rho}{2+\rho}$
* $P_j = \frac{1}{2^{j-1}} \rho^j P_0$, $j = 1, \, 2, \, \ldots$
* $L = \sum_{j=0}^{\infty} jP_j = P_1 + 2P_2 + 3P_3 + \cdots = \frac{2-\rho}{2+\rho}\left\{\rho + \sum_{j=2}^{\infty} j \left(\frac{1}{2}\right)^{j-1} \rho^j\right\} == \frac{2-\rho}{2+\rho}\left\{\rho + 2\sum_{j=2}^{\infty} j \left(\frac{\rho}{2}\right)^j\right\} = \frac{2-\rho}{2+\rho}\left\{\rho + 2\frac{\rho/2}{\left(1-\rho/2\right)^2} - 2\,\rho/2\right\} = \frac{4\rho}{(2+\rho)(2-\rho)} = \frac{4*0.9}{2.9*1.1} = 1.129$
* $W = \frac{L}{\lambda} = \frac{1.129}{9} = 0.125 \; hour$
* $1 - P_0 - P_1 = 1 - (1+\rho)\frac{2-\rho}{2+\rho} = \frac{\rho^2}{2+\rho} = \frac{0.81}{2.9} = 0.28$
* **Probability all servers are busy = probability that a customer has to wait**
* **(PASTA)!**

- **What is the average number of busy servers = average number of customers in service?**
- $L_s = 0P_0 + 1P_1 + 2\sum_{j=2}^{\infty} P_j = 0 + \rho\frac{2-\rho}{2+\rho} + 2\frac{\rho^2}{2+\rho} = \frac{(\rho^2+2\rho)}{2+\rho} = \rho = 0.9$

Example machine repair model:

- $\lambda_{av} = \sum_{j=0}^{K} P_j\lambda_j = \lambda(K-L)$
- $L = \sum_{j=0}^{j=K} jP_j$
- $W = \frac{L}{\lambda_{av}}$
- $L_q = \sum_{j=R}^{j=K}(j-R)\,P_j$

$$3\lambda P_0 = \mu P_1$$
$$(2\lambda+\mu)P_1 = 3\lambda P_0 + 2\mu P_2$$
$$(\lambda+2\mu)P_2 = 2\lambda P_1 + 2\mu P_3$$
$$\lambda P_2 = 2\mu P_3$$

$\longrightarrow$

$$P_1 = 3\rho P_0 = \binom{3}{1}\rho^1 P_0$$
$$P_2 = 3\rho^2 P_0 = \binom{3}{2}\rho^2 P_0$$
$$P_3 = \frac{3}{2}\rho^3 P_0 = \frac{3!\binom{3}{3}\rho^3 P_0}{2!\,2}$$

$$P_0 + P_1 + P_2 + P_3 = 1 \rightarrow P_0 = \frac{1}{1 + 3\rho + 3\rho^2 + \frac{3}{2}\rho^3}$$

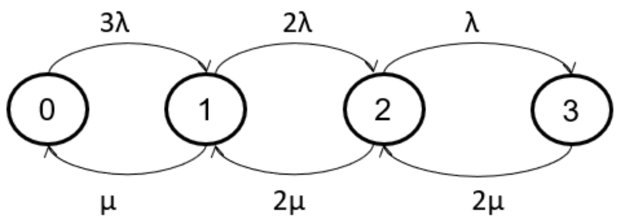
So: $P_0 = \frac{128}{251}\,;P_1 = \frac{96}{251}\,;P_2 = \frac{24}{251}\,;P_3 = \frac{3}{251}$

- K = 3 machines, break down once in 8 days on average
- R = 2 repair workers, repair takes 2 days on average
- State n = number of broken machines

λ = 1/8 per day

μ =1/2 per day

ρ = λ/μ = 2/8 = 1/4



- Average number of broken machines

$$L = \sum_{j=0}^{3} jP_j = P_1 + 2P_2 + 3P_3 = \frac{153}{251}$$

- Average time a machine is down

$$W = \frac{L}{\lambda_{av}} = \frac{153}{75} > 2\ days!$$

where λ_{av} = average number of "arriving customers"

$$= 3\lambda P_0 + 2\lambda P_1 + \lambda P_2 + 0P_3 = \frac{75}{251}$$

- In general:

$$\sum_{j=0}^{3} P_j\lambda_j = \sum_{j=0}^{3} P_j\lambda(K-j) = \lambda(K-L) = \frac{1}{8}\left(3 - \frac{153}{251}\right) = \frac{75}{251}$$

Go rock your studies

GOOD LUCK!

Stress