

Bhoris Dhanjal

Industrial statistics

Lecture Notes
for SSTA403

Contents

1	CPM and PERT	1
1.1	Network analysis	1
1.1.1	Required definitions:	2
1.1.2	Rules for drawing the network	4
1.2	Fulkerson's Rule for Numbering the events	5
1.3	Management tools for scheduling a project	6
1.4	Time estimates for the execution of each activity	7
1.5	PERT	7
1.5.1	Calculation of the critical path	8
1.6	CPM	8
1.6.1	Example of CPM	10
1.7	Project crashing	11
1.7.1	Procedure of crashing	12
1.8	something	15
2	Quality control	17
2.1	Introduction	17
2.2	Statistical quality control	18
2.3	Control chart	18
2.4	Control charts for variable	18
2.4.1	For mean	18
2.4.2	For range	19
2.4.3	Construction of R -chart	19
2.4.4	For standard deviation	19
2.5	Control charts for attributes	20
2.5.1	p -chart	20
2.5.2	np -chart/ d -chart	21
2.5.3	c -chart	22
2.6	Objectives of statistical quality control	24

3	Acceptance sampling	25
3.1	OC curve	26
3.2	Sampling plan for attributes	27
3.2.1	Single sampling plan (SSP)	27
3.2.2	Double sampling plan (DSP)	29

Chapter 1

CPM and PERT

1.1 Network analysis

Large and complex projects of any organization involve a number of inter-related activities with limited resources such as labour, machines, material, money and time. It is not possible for the management to find and execute an optimum solution just by intuition and work experience.

The main objective of project management can hence be described in terms of successful completion of a project within time, budgeted cost and the technical specifications/ It also becomes essential to incorporate any change in the initial plan and immediately know the effect of the change from time to time.

Definition 1.1 (Network analysis). *The statistical tool developed to help in planning, managing and controlling the project is known as **network analysis**.*

Thus network analysis deals with dividing the given project into smaller well-defined tasks and studying the interdependencies of these tasks. When the interdependencies of these tasks are best represented, they form a **network**; and hence the name **network analysis**. It involves three main steps:

1. It defines the jobs to be done.
2. It integrates the jobs in a logical time sequence.
3. It controls the progress of the project plan.

When all the tasks are accomplished, the project is said to be completed.

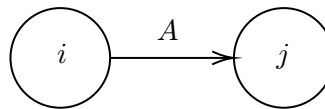
Project scheduling consists of four main steps:

1. **Planning:** The project is divided into a number of well defined tasks / activities. The time estimation for the completion of each activities and the sequential order in which the activities should be executed is indicated.
2. **Scheduling:** The objective of this phase is to construct a time chart showing the start and finish time for each activity by applying forward and backward pass techniques and identify the critical path to indicate the critical activities; which require special attention if the project is to be completed on time; along with the slack and float for the non-critical paths.
3. **Allocation of resources:** A resource is a physical variable required such as labour, equipment, space, etc. In this phase, the resources are allocated to different activities to achieve the desired objective.
4. **Controlling:** It refers to analysing and evaluating the actual progress against the plan. In this final phase of project management, a financial and technical control is obtained over the project, by having progress reports from time to time and updating the networking continuously. Reallocation of resources, crashing and review of project with periodical reports are carried out in this phase of the project.

1.1.1 Required definitions:

1. **Activity:** An individual operation which consumes resources, has a beginning and an end, is called an activity.

An arrow is used to represent an activity, with its head indicating the direction of the progress of the project.



Where A is the activity, i and j are events.

2. **Event:** An event represents a point in time satisfying the completion of some activities and the beginning of the new ones.

It is also called as a **node** or **connector** and is usually represented by a circle in the network.

3. **Tail event:** It is the event from which the activity begins, i.e. the event from which the arrow emerges.
4. **Head event:** It is the event into which the activity ends, i.e. the event into which the arrow enters.
5. **Dummy activity:** An activity which does not consume any resources but only shows the technological dependence of activities is called a dummy activity. It is represented by a dotted line in the network.
6. **Network:** When all the activities and events are connected logically and sequentially, they form a network.
7. **Start event:** An event which does not have a preceding activity is called the start event of the network.
Thus, the start event will have no arrows entering into it, but will only have arrows emerging out of it.
8. **End event:** An event which does not have a succeeding activity is called the end event of the network.
9. **Merge event:** Merge event is the event into which more than one activities end, i.e. the event into which many arrows enter.
10. **Burst event:** Burst event is the event from which more than one activities start, i.e. the event from which many arrows emerge.
11. **Predecessor activity:** Activity which must be completed before a particular activity starts is known as a predecessor activity.
12. **Successor activity:** Activity which must follow a particular activity is known as a successor activity.
13. **Critical activity:** An activity which cannot be delayed if the project is to be completed on time is known as a critical activity.
Total float for every critical activity is zero.
14. **Critical event:** An event which cannot be delayed if the project is to be completed is known as a critical event.
Slack of every critical event is zero.
15. **Path:** A path is a connected sequence of activities from start event to end event.

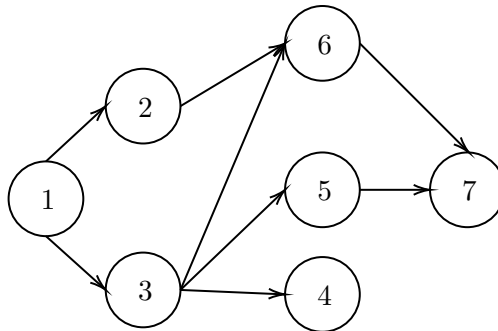
16. **Critical path:** A path connecting the start event and the end event through the critical events or critical activities is called the critical path.

Duration-wise it is the longest path on the network. It is represented using **bolder** lines on the network.

1.1.2 Rules for drawing the network

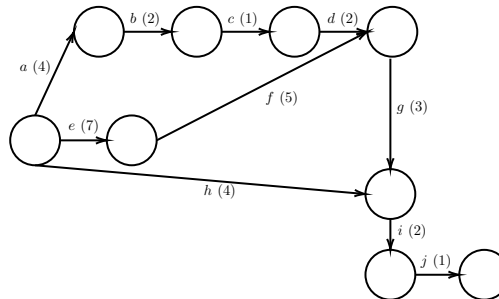
1. Each activity is represented by only one arrow in the network.
2. The arrows should be drawn as straight lines.
3. The arrows should not cross each other.
4. The length of an arrow is of no significance.
5. Time flows from left to right.
6. Arrows pointing in opposite direction are to be avoided.
7. Angles between arrows should be as large as possible.
8. Every network should have a unique start event.
9. Every network should have a unique end event.
10. Before an activity can be undertaken, all activities preceding it must be completed.
11. There should be no duplication in event numbers.
12. No two activities can have same start event and end event..
13. No loops allowed.

Example 1.2. *Event–2; 3; 4; 5; 6; 7, preceded by 1; 1; 3; 3, 4; 2, 3, 5; 5, 6*



Example 1.3. Make a diagram based on the following table

Sr. No.	Activity	Duration (hrs)
a)	Preparing a pattern for casting item 'A'	4
b)	Preparing a mould	2
c)	Casting and cleaning item 'A'	1
d)	Heat treatment for 'A'	2
e)	Obtaining and installing machine 'M'	7
f)	Preparing item 'B' by 'M'	5
g)	Assembling 'A' and 'B'	3
h)	Preparing test rig	4
i)	Testing the assembly	2
j)	Packing	1



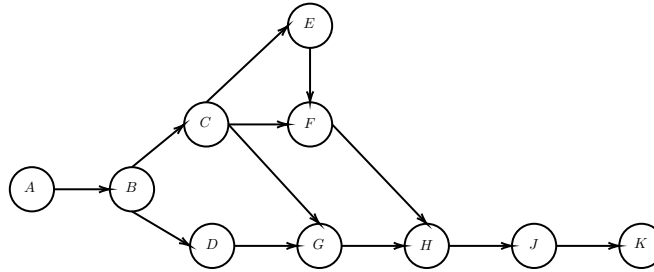
1.2 Fulkerson's Rule for Numbering the events

1. A start event is an event which has only emerging arrows, and no entering arrows. The start event is numbered as 1.
2. Strike out all the arrows that emerge out of event 1. This will create at least one event that has no entering arrows. Number these events as 2, 3, 4, ...
3. Continue this procedure till all the events are numbered.

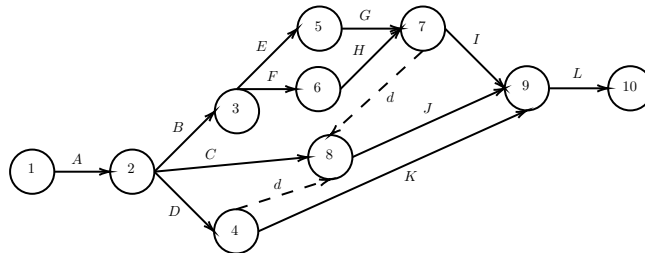
Note: Events can be numbered as 1, 2, 3, ... or 10, 20, 30, ...

Example 1.4. Draw a network for the following project and number the events according to Fulkerson's rule.

'A' is a start event and, 'K' is end event. 'A' precedes the event 'B', 'J' is the successor event to 'F', 'C' and 'D' are successor events to 'B', 'D' is preceding event to 'G', 'E' and 'F' occurs after event 'C', 'E' precedes 'F', 'C' restrains the occurrence of 'G' and 'G' precedes 'F'. 'H' precedes 'J' and 'K' succeeds 'J', 'F' restrains the occurrence of 'H'.



Example 1.5. B,C,D preceding A; E,F preceding B; G preceding E; H preceding F; I preceding G,H; J preceding G,H,C,D; K preceding D; L preceding I,J,K



Proof.

□

1.3 Management tools for scheduling a project

1. Project evaluation and review technique (PERT)
2. Critical path method (CRT)

They help in reducing project execution time.

1.4 Time estimates for the execution of each activity

Activities which can be performed but no data is available for their time estimates such activities are called variable time or probabilistic activity.

On the other hand, there are activities for which the time estimates can be accurately determined.

Such activities are called fixed type or deterministic activities.

The projects with probabilistic activities employ PERT, whereas the projects which involve deterministic activities are handled by CPM techniques.

PERT, developed during 1950's was first used in conjunction with planning and designing of the Polaris missile project.

CPM was developed by DuPont company and applied first to the construction projects in the chemical industry.

1.5 PERT

This network analysis technique gives more emphasis on events. Since the time estimates are not known, three different types of time estimates are defined for each activity.

1. Optimistic time estimate (t_0)

This is the shortest time within which an activity can be completed.

2. Pessimistic time estimate (t_p)

This is the maximum time required for the completion of an activity.

3. Most likely time estimate (t_m)

This is the estimate of time, which the activity will require under normal circumstances. The frequency curve of the activity time estimates resembles the beta distribution curve.

The mean time required for the completion of the activity (t^{ij}) is given by

$$t^{ij} = \frac{t_0 + 4t_m + t_p}{6}$$

$$\text{Variance} = \left(\frac{t_p - t_0}{6} \right)^2$$

t0	tm	tp	expected time	variance
4	5	8	5.333333333	0.4444444444
5	7	10	7.166666667	0.6944444444
2	3	7	3.5	0.6944444444
8	11	12	10.66666667	0.4444444444
4	7	10	7	1
6	8	15	8.833333333	2.25
8	12	16	12	1.777777778
5		9	2.333333333	0.4444444444
3		7	1.666666667	0.4444444444
3		11	2.333333333	1.777777778
6		13	3.166666667	1.361111111

1.5.1 Calculation of the critical path

Definition 1.6 (Earliest occurrence time). *Earliest occurrence time (T_E) is the earliest time at which an event can occur.*

Definition 1.7 (Latest Occurrence time). *Latest occurrence time (T_L) is the latest time at which an event can occur.*

Definition 1.8 (Forward pass). *The process of calculating T_E from start event to end event. Initially T_E is taken as zero. Then it is calculated by adding the previous T_E and duration of the activity. If there are more than one activity preceding then the earliest time for each event is taken as the maximum of T_E .*

Definition 1.9 (Backward pass). *content...*

Definition 1.10 (Slack). *content...*

Definition 1.11 (Critical path). *content...*

Example 1.12.

1.6 CPM

1. **Earliest start time (EST)** The EST of an activity is the earliest time of the event from which it emerges. For activity $i - j$, $(EST)_{ij} = T_E^i$
2. **Earliest finish time (EFT)** The EFT of an activity is defined as the sum of the earliest time at which the activity starts and the duration of the activity. For activity $i - j$, $(EFT)_{ij} = T_E^i + t^{ij}$

3. **Latest finish time (LFT)** The latest time by which an activity should be completed is LFT, i.e. the latest time at which the event indicating the end of the activity should occur. For activity $i - j$; $(LFT)_{ij} = T_L^j$
4. **Latest start time (LST)** The LST for an activity is defined as the difference between LFT and its duration. For activity $i - j$; $(LST)_{ij} = T_L^j - t^{ij}$

Definition 1.13 (Float). *It indicates the time by which the activity can be delayed without affecting the schedule of the project. Float for the activity has the same significance as slack for an event.*

Definition 1.14 (Total float). *It is defined as the difference between the maximum time available to perform an activity and the actual duration of an activity.*

$$\begin{aligned}(TF)_{ij} &= (T_L^j - T_E^i) - t^{ij} \\ &= T_L^j - (T_E^i + t^{ij}) \\ &= (LFT)_{ij} - (EFT)_{ij}\end{aligned}$$

For a particular activity, if the total float is zero, it indicates that it is not affordable to delay the activity.

Definition 1.15 (Critical activity). *The activity for which float is zero.*

Definition 1.16 (Critical path). *A path connecting all the critical activities.*

Definition 1.17 (Free float (FF)). *The time by which an activity can be delayed under the assumption that all events in the network occur at their earliest is known as the Free Float.*

$$\begin{aligned}(FF)_{ij} &= (T_E^j - T_E^i) - t^{ij} \\ &= (T_E^j - T_E^i) - t^{ij} + T_L^j - T_L^j \\ &= \{(T_L^j - T_E^i) - t^{ij}\} - (T_L^j - T_E^j) \\ &= (TF)_{ij} - \text{slack of the head event}\end{aligned}$$

Definition 1.18 (Independent float (IF)). *The time by which an activity can be delayed without affecting the preceding and the succeeding activities, is known as the independent float.*

$$\begin{aligned}(IF)_{ij} &= (T_E^j - T_L^i) - t^{ij} \\ &= (T_E^j - T_L^i) - t^{ij} + T_E^i - T_E^i \\ &= (FF)_{ij} - \text{slack of the tail event}\end{aligned}$$

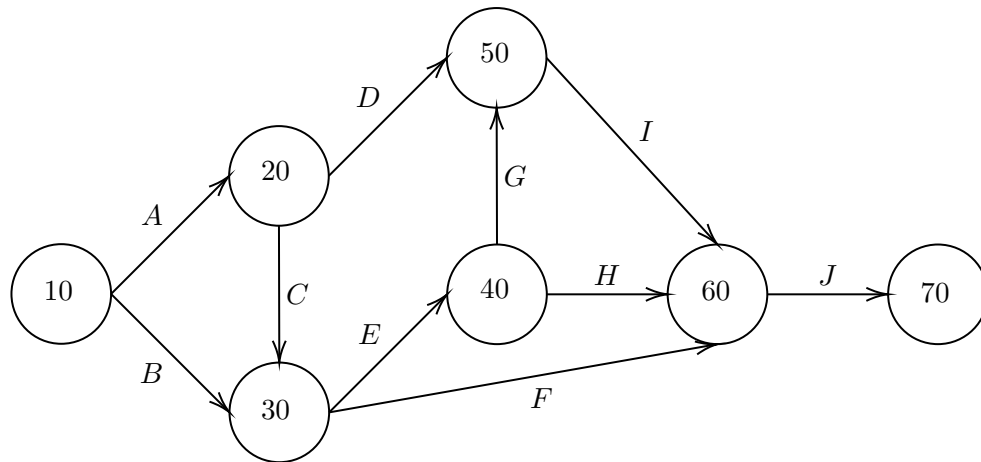
Independent float can be negative, if the duration of the activity is greater than the minimum time required to complete the activity. If negative, independent float is considered as zero.

1.6.1 Example of CPM

Activity	Duration	Preceding Activity
A	15	—
B	15	—
C	3	A
D	5	A
E	8	B, C
F	12	B, C
G	1	E
H	4	E
I	3	D, G
J	14	F, H, I

Draw the network for above project and find EST , LST , EFT , LFT and the three types of float. Hence, determine the critical path.

Proof. First draw the network diagram as follows,



$$T_E = 54$$

Events	t^{ij}	$EST = T_E^i$	$LFT = T_L^j$	$EFT = T_E^i + t_{ij}$	$LST = T_L^j - t^{ij}$	$TF = LFT - EFT$	$FF = TF - (T_L^j - T_E^i)$
10-20	15	0	15	15	0	0	0-(15-15)=0
10-30	15	0	18	15	3	3	3-(0)=3
20-30	3	15	18	18	15	0	0-(18-18)=0
20-50	5	15	37	20	32	17	7
30-40	6	18	26	24	20	0	0
30-60	12	18	40	30	28	10	10
40-50	1	26	37	27	36	10	0
40-60	14	26	40	40	26	0	0
50-60	3	27	40	30	37	10	10
60-70	14	40	54	54	40	0	0

Now we find total float and free float.

Total float = $LFT - EFT$. If total float is zero it implies its a critical activity.

Free float = $TF - \text{Head slack}$

Independent float = $FF - \text{Tail slack}$

Head slack = $T_L^j - T_E^j$

Tail slack = $T^i - T_E^i$

□

1.7 Project crashing

Two types

Definition 1.19 (Project crashing). *Process of shortening the critical path to achieve completion of the project earlier is called project crashing.*

There are two types of cost,

1. **Direct cost:** It includes cost of materials, machinery, man hours etc. When we estimate the duration of various activities in the project it is assumed that various activities in the project it is assumed that the normal amount of labour and machine required to complete these activities. But when we want to complete the project in the shorter period, than the critical path we will need to employ more resources. Hence direct cost will increase. **yeah this makes no sense dw**
2. **Indirect cost:** It includes rent, overheads, administrative costs. Indirect costs vary with time. They are expressed on per day or per week basis

Total cost = Direct cost + indirect cost

Normal time: Under normal circumstances. **Crash time:** Minimum possible time in which activity can be completed. **Crash cost:** The direct cost associated with crash time. When the activity is crashed its direct cost will increase.

Crash time ; Normal time
 Crash cost ; Normal cost

Definition 1.20 (Crash slope/Cost slope). *The increase in direct cost per day. It is Crash cost - normal cost divided by normal time - crash time*

$$\frac{\text{Crash cost} - \text{Normal cost}}{\text{Normal time} - \text{Crash time}}$$

1.7.1 Procedure of crashing

1. Find the normal critical path and identify the critical activities.
2. Calculate the cost slope for the critical activities.
3. Rank the critical activities in the ascending order of their cost slopes.
4. Crash critical activities as per the ranking (i.e. the activity with the least cost slope first) to its maximum possible extent. Calculate the revised cost by adding the cost of crashing and subtracting the indirect cost; both proportion to the number of days crashed.
5. As the critical path duration gets reduced in step 4, some other path may also become critical, i.e. there may be parallel critical paths. This implies that the project duration now has to be reduced simultaneously by crashing the activities on the parallel critical paths.
6. After crashing the activities in the above manner, a point is reached where further crashing is either not possible or it does not result in reduction of project duration or cost. For different project or durations, project cost should be calculated.

Activity	t^{ij}	Activity	t^{ij}
1-2	2	4-6	3
1-3	2	5-8	1
1-4	1	6-9	5
2-5	4	7-8	4
3-6	8	8-9	3
3-7	5		

Example 1.21. *Draw a network for this project. Find total float for each activity. Hence determine critical path for network.*

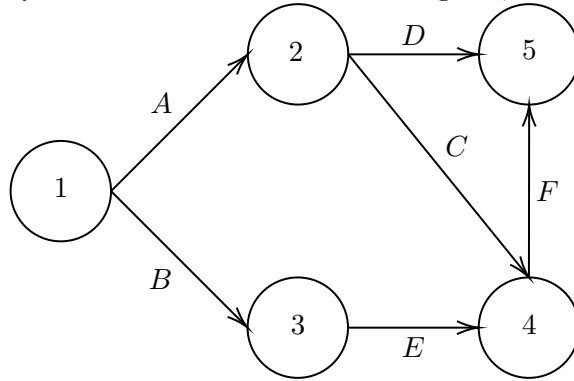
Proof. $T_E = T_L = 15$

□

Example 1.22. *Indirect cost is Rs. 70 per day.*

Activity	Normal time	Crash time	Normal cost	Crash cost
1-2	8	6	100	200
1-3	4	2	150	350
2-4	2	1	50	90
2-5	10	5	100	400
3-4	5	1	100	200
4-5	3	1	80	100

Proof. Consider first the network diagram,



Critical path is 1 – 2 – 5 : $A - D = 18$ days.

Other paths are 1 – 3 – 4 – 5 : $B - E - F = 12$ days

1 – 2 – 4 – 5 : $A - C - F = 13$ days.

Total normal cost is = $DC + IDC$

$$= DC + IDC$$

$$= 580 + 70 * 8$$

$$= 1140$$

Now see the cost slope

$$\text{Cost slope} = \frac{\Delta \text{time}}{\Delta \text{cost}}$$

Minimum total cost = optimum project cost = 1760

□

Activity	Maximum crashable days	Cost Slope / day	Rank
A(1-2)	8-6=2	50	4
B(1-3)	4-2=2	100	6
C(2-4)	2-1=1	40	3
D(2-5)	10-5=5	60	5
E(3-4)	5-1=4	25	2
F(4-5)	3-1=2	10	1

Crashing	Paths			Cost (Rs.)		
	A-D	A-C-F	B-E-F	D.C.	I.D.C	T.C.
Before Crashing	18	13	12	580	1266	1840
1st Crashing A by 1	18-1= 17	13-1= 12	12	580+(50*1) =630	17*70 =1190	1820
2nd Crashing A by 2	17-1= 16	12-1= 11	12	630+(50) =680	16*70 =1120	1800
3rd Crashing D by 1	16-1= 15	11	12	680+60 =740	15*70 =1050	1790
4th Crashing D by 1	15-1= 14	11	12	740+60 =800	14*70 =980	1780
5th Crashing D by 1	13	11	12	860	910	1770
6th Crashing D by 1	12	11	12	920	840	1760
7th Crashing D by 1 and F by 1	11	10	11	920+60+10 =990	11*70 =770	1760

1.8 something

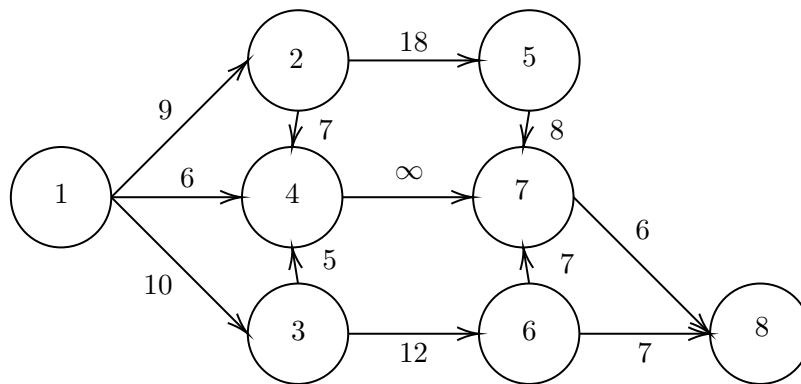
...which have occurred due to re planning or re scheduling of ...
Updating can be done in two ways,

1. Use the revised time estimates of incomplete activities and calculate from the completion of time of each event in the usual manner to know the project completion time. and represent all the activities already finished by elapsed time. Events in the revised network diagram are re numbered. And the completion of the completed activities (????) are taken as the revised time.

Example 1.23. After 15 days working the following progress is noted of an erection job.

1. Activity 1-2,1-3 and, 1-4 completed as per the original schedule.
2. Act 2-4, is in progress and will be completed in 3 more days
3. Act 3-6 is in progress and require 18 days more.
4. Act 6-7 appears to present some problem and its new estimated time is 12 days.
5. Act 6-8 can be completed in 5 days instead of 7 days.

Proof. Consider the network diagram

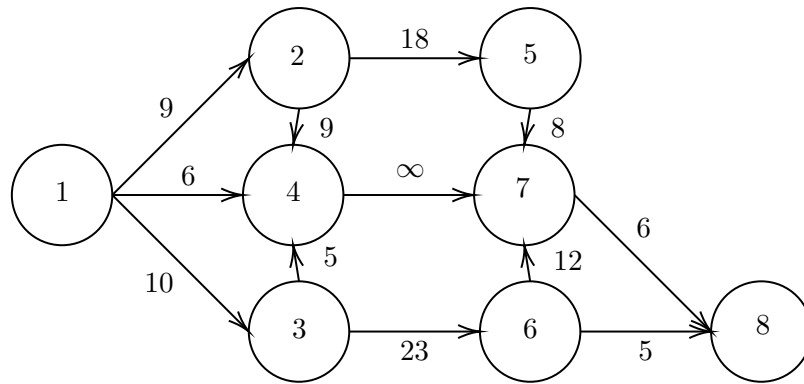


$$T_E = 42, T_L = 42 \text{ new duration of } 2 - 4 \text{ is } 15 + 3 - 9 = 9$$

1. Act 2-4 needs $15+3-9=9$ days instead of 7.
2. Act 3-6 needs $15+18-10 = 23$ days instead of 12.

3. Act 6-7=12 days instead of 7 days.
4. Act 6-8=5 days instead of 7
5. Act 2-5, 4-7, 5-7, 7-8, 3-4 remain as it is.

New network,



New $T_E, T_L = 51$ with critical path given as $1 - 3 - 6 - 7 - 8 = 51$. □

Chapter 2

Quality control

2.1 Introduction

Quality here means a level/standard which in turn depends on four M's

- Materials
- Manpower
- Machines
- Management

A customer expects a good quality product and hence the progress of industry depends on the marketing as well as quality.

Definition 2.1 (Chance cause of variation). *The variation due to these causes is beyond the control of human hand and cannot be prevented or eliminated under any circumstances.*

It is also called as allowable as we have to allow these variation. The range of such variation is known as natural tolerance of the process.

Definition 2.2 (Assignment cause of variation). *It is due to non random preventable variation.*

The assignable cause may creep in at any stage of the process, right from the arrival to the raw materials to the final delivery of goods. These causes can be identified and eliminated and are to be discovered in a production process before it goes wrong.

2.2 Statistical quality control

The aim in SQC is to control the manufacturing process so that the proportion of items beyond specified limits is not very large. This is known as process control.

The other type of problem is to ensure that lots of manufactured goods do not contain large proportion of defective items. This is known as product control or lot control.

These two are distinct problems. Process control is achieved mainly through the technique of control charts where lot control is achieved through sampling inspection.

Many quality characteristics are measurable quantitatively, e.g. length of a screw, life of bulb etc. Most of them are continuous variables.

Sometimes the quality of the products cannot be measured in such cases the items can be classified as defective or non defective.

2.3 Control chart

A typical control chart consists of three horizontal lines.

- **Control line:** It indicates the desired standard of the process.
- **Upper control limit:** It indicates the upper limit of tolerance.
- **Lower control limit:** It indicates the lower limit of tolerance.

2.4 Control charts for variable

2.4.1 For mean

The $3 - \sigma$ limits are $E[\bar{X}] \pm 3\sqrt{V(\bar{X})}$.

If μ and σ are known as μ', σ' then $UCL = \mu' + 3\frac{\sigma'}{\sqrt{n}}$ and $LCL = \mu' - \frac{3\sigma'}{\sqrt{n}}$. Here n is sample number not sample size.

If μ, σ are unknown, then $\hat{\mu} = \bar{\bar{X}} = \frac{\bar{X}_1 + \dots + \bar{X}_k}{k}$ and $\hat{\sigma} = \frac{\bar{s}}{c_2} = \frac{\bar{R}}{d_2}$. And we have $UCL = \bar{\bar{X}} + \frac{3\bar{s}}{c_2\sqrt{n}} = \bar{\bar{X}} + A_1\bar{s}$ and similar for LCL .

Construction of \bar{X} chart

Sample No. along the X -axis and \bar{X} along the Y -axis.

Sample points are then plotted against corresponding sample numbers.

The central line is drawn bold.

UCL and LCL are then plotted.

If point falls beyond the control line it is regarded that an assignable cause has thrown the process out of control.

The next step is to remove out sample results which are outside the control lines for the remaining samples.

$$UCL \text{ is } \bar{\bar{X}} + A_2\bar{R}$$

2.4.2 For range

$$UCL = \bar{\bar{x}} + \frac{3\bar{R}}{d_2\sqrt{n}} = \bar{\bar{X}} + A_2\bar{R} \text{ and similar with negative for LCL. Or } UCL = \bar{\bar{X}} + \frac{3\bar{S}}{c_2\sqrt{n}} = \bar{\bar{X}} + A_1\bar{S}$$

2.4.3 Construction of R-chart

$\bar{R} = \frac{1}{k}(R_1 + R_2 + \dots + R_k)$. So $UCL = \bar{R} + 3\sigma_R$. $\sigma_R = d_2\hat{\sigma}$ or $d_3\bar{R}/d_2$.

If sigma is known we will take $UCL = D_2\sigma$, $LCL = D_1\sigma$. If sigma is known then $UCL = D_4\bar{R}$, $LCL = D_3\bar{R}$.

We get D_1, D_2, D_3, D_4 from the table for different n values, for $2 < n < 25$.

2.4.4 For standard deviation

Combination of control chart for mean \bar{X} and s.d. S known as $\bar{X} - S$ chart. It is theoretically more appropriate than a combination of \bar{X} and R .

$$E[s] = c_2\sigma, \sigma_s = c_3\sigma$$

When standards are known $CL = E[s] = c_2\sigma$, $UCL = c_2 + 3c_3\sigma = B_2\sigma$, $LCL = c_2 - 3c_3\sigma = B_1\sigma$ when sigma is known.

When sigma is unknown $\hat{\sigma} = \bar{s}/c_2$

$$CL = \bar{s} = c_2\sigma = c_2\bar{s}/c_2 = \bar{s}$$

$$UCL = B_4\bar{S} \text{ and } LCL = B_3\bar{S} \text{ where } \bar{S} = \frac{\sum S_i}{n}$$

Example 2.3. A machine is said to deliver packets of 10 samples of size 5 each were examined and the following results were obtained.

Proof. Compute $\bar{\bar{X}} = 44.2$ and $\bar{R} = 5.8$.

First do for \bar{X} , $CL = \bar{\bar{X}} = 44.2$, $UCL = \bar{\bar{X}} + \frac{3\bar{R}}{d_2\sqrt{n}} = 44.2 + \frac{3 \times 5.8}{2.326 \times \sqrt{5}} = 47.546$,

and $LCL = \bar{\bar{X}} - \frac{3\bar{R}}{d_2\sqrt{n}} = 40.85455$

Sample No.	Mean	Range
1	43	5
2	49	6
3	37	5
4	44	7
5	45	7
6	37	4
7	51	8
8	46	6
9	43	4
10	47	6

Then do for \bar{R} , we have $CL = \bar{R} = 5.8$, and $UCL = D_4\bar{R} = 2.115 \times 5.8 = 12.267$ and $LCL = D_3\bar{R} = 0 \times 5.8 = 0$ \square

2.5 Control charts for attributes

The most commonly used control charts under this category are,

2.5.1 p -chart

Control chart for fraction defective. The objective of this chart is to evaluate the quality of the items, i.e. the average fraction defective or percent defective. And note the changes in quality over period of time. The chart can also be used in routine control and guide for correcting causes of bad quality. In many situations these charts suggest where \bar{X} and R charts could be used.

Steps to make a p -chart,

1. $P = \frac{\text{No. of defectives in the sample}}{\text{Sample size}} = \frac{np}{n} = \frac{d}{n}$, average $f.d.p = \frac{\sum d}{\sum n}$
2. $UCL_P = P + 3\sigma_p$, $LCL = P - 3\sigma_p$ where $\sigma_p = \sqrt{\frac{P(1-P)}{n}}$

Case I:

$P' = P$ then $LCL_p = P' - 3\sqrt{P'(1-P')/n} = P' - A\sqrt{P'(1-P')}$ similar for UCL and CL .

Case II:

$P = \bar{P}$ same formula as above but with \bar{P} .

If LCL_p is found to be negative according to the formulas it is taken to be 0.

If the point goes above the UCL it reflects the lack of statistical control since it has changed for the worse. Such a point known as high spot, indicates deterioration in the lot quality.

If the point goes below the LCL it again exhibits lack of control as the process has changed for the better. Such a point is known as low spot, indicates improvement in the lot quality.

2.5.2 np -chart/ d -chart

Control chart for number of defectives.

If ??????? on each occasion is the same the plotting of actual number of defectives may be more convenient and meaningful than the fraction defective. The difference is that the actual number of defectives (np) in samples of fixed size is plotted instead of fractional defectives and the central line is drawn at np instead of p .

Following values required for np ,

1. n = sample size constant
2. The number of defectives in each sample $d = np$
3. Average number of defectives per sample of constant size,

$$nP = \frac{\text{total no. of defectives}}{\text{number of samples}}$$

The $3 - \sigma$ limits are as follows,

1. $UCL_{np} = nP + 3\sigma_{np}$
2. $LCL_{np} = nP - 3\sigma_{np}$

Where, P = the fraction defective true in process (population), nP = the average number of defectives per sample based on all possible sample size n from the population and $\sigma_{np} = \sqrt{nP(1 - P)}$.

Case I: If $P = p'$ then $LCL = np' - 3\sqrt{np'(1 - p')}$ with CL UCL as expected.

Case II: Suppose we have k different samples, $LCL = np - 3\sqrt{np(1 - p)}$ with CL UCL as expected.

np -chart vs p -chart ... control line and control limits will vary, in such case p -chart is better to use. However, in case of constant sample size for all the sample, any one of p -chart or np -chart can be used. But in practice np -chart is most commonly used.

2.5.3 c -chart

Control chart for number of defects per unit.

The c -chart is used for control of number of defects per unit. Although the applications of c -chart is somewhat limited compared with p -chart there are many instances in industry where it is useful. E.g. for control of number of defects of a material, number of soiled packages in a given consignment or number of weak spots in a given length.

....based on number of defectives per unit. E.g. in case of inspection of fairly complex assembled units such as T.V. set.

The construction of control chart.

The number of defects where the sample was constant was as follows,

1. Count the number of defects individually. Obtain average number of defects $\bar{c} = \frac{\text{no of defects in all samples}}{\text{total no of samples}}$.
2. Create UCL, LCL with $3 - \sigma$ limits. $UCL_c = \hat{c} + 3\sqrt{\hat{c}}$ and $LCL_c = \hat{c} - 3\sqrt{\hat{c}}$. Where $\hat{c} = \bar{c} = \frac{c_1 + c_2 + \dots + c_k}{k}$ is true average no of defects per product.

When the estimate value is used, the two limits may be written as $UCL_c = \bar{c} + 3\sqrt{\bar{c}}$ and $LCL_c = \bar{c} - 3\sqrt{\bar{c}}$. Where c_1, c_2, \dots, c_k are the number of defects observed in $1^{st}, \dots, k^{th}$ sample respectively.

The interpretation of this chart is similar to the chart for fraction defective and number of defectives. The evaluation of the standards are also done in the same line.

Although the applications for the c -chart is somewhat limited compared with the p -chart still there are many instances in industry where it is useful. It can be used to identify or to highlight the number of surface(?????) effects observed in roll of galvanized sheet of printed/plated ????? or sheet photographic film.??????????????

Example 2.4. A sample of 500 wiring boards are taken from the wiring assembly line each day. The number and percentage rejected during 20 consecutive days of production are shown below.

Proof. We compute $\bar{P} = \frac{599}{10000} = 0.0599$

□

Date sampled	No of Samples	Number rejected	Defective percent
8	500	40	8
9	500	30	6
10	500	20	4
11	500	60	12
12	500	30	6
15	500	20	4
16	500	90	18
17	500	50	10
18	500	30	6
19	500	20	4
22	500	20	4
23	500	10	2
24	500	30	6
25	500	20	4
26	500	20	4
29	500	10	2
30	500	30	6
31	500	20	4
1	500	10	2
2	500	39	7.8

Type	Standards given	Standards are not given
Mean	$CL = \mu'$	$CL = \bar{\bar{X}}$
	$UCL = \mu' + \frac{3\sigma'}{\sqrt{n}}$	$UCL = \bar{\bar{X}} + A_2\bar{R}$
	$LCL = \mu' - \frac{3\sigma'}{\sqrt{n}}$	$LCL = \bar{\bar{X}} - A_2\bar{R}$
Range	$CL = \sigma'$	$CL = \bar{R}$
	$UCL = D_1\sigma'$	$UCL = D_4\bar{R}$
	$LCL = D_2\sigma'$	$LCL = D_3\bar{R}$

1

2.6 Objectives of statistical quality control

1. To assess the quality of raw material, semi finished goods and finished products at various stages of production process. To achieve better utilization of raw materials, more efficient utilization of equipment etc. To locate and identify the process faults. In order to control the scrap and waste.
2. To reduce the waste men and machine during the process of production.
3. To see whether the product conforms to the predetermined standards and specifications and whether it satisfies the needs of the customer.
4. To take necessary corrective steps to maintain the quality of the product.
5. To protect the manufacturers as well as the consumers against heavy losses due to a rejection of a large quantity of product.

¹ $A_2 = \frac{3}{d_2\sqrt{n}}$

Chapter 3

Acceptance sampling

If the plan is ????? qualitatively then the plan is referred as attribute sampling plan.

And if the plan is ????? then the plan is referred as variable sampling plan

An attribute sampling plan may be given as follows,

Select a sample of size $n = 50$ and count the number of defectives c if $c \leq 3$ accept the lot otherwise reject.

A variable sampling plan typically uses the normal distribution and might be stated as follows,

Select a random sample of size $n = 50$ and determine the mean tensile strength x if $x > 12000$ accept the lot otherwise reject it.

While in attribute sampling plan the number of defectives in the parent population are estimated from binomial and poisson distribution.

Actual condition	Lot accepted	Rejected
Good lot	Favourable outcome	Producer risk, type I error
Bad lot	Consumer risk, Type II Error	Favourable outcome

Producers risk - Type I error (α) , Consumers risk - Type II error p is lot tolerance (β)

AQL- Acceptable quality level - good lot.

By producer we mean any person, firm or department that prepares goods to be supplied to another person, firm or department.

Any sampling inspection plan for acceptance or rejection of a lot possesses the disadvantage of occasionally rejecting a lot of satisfactory quality

P is called producers process average,

P [rejecting a lot, under the sampling inspection plan]

If a good lot is rejected we refer it as a type I error.

A person, firm, department that receives a product from the producer is called the consumer.

P [accepting a lot of unsatisfactory quality on the basis of sampling inspection] $=\beta$.

If p is the maximum fraction defective in a lot then the consumer can tolerate. Then it is called as consumers risk.

Consumers want to keep risk low and we refer it as a type II error.

Acceptable quality level is the lot quality which the supplier believes will be accepted by the consumer. That is he believes that a lot with quality p will have a small chance of being rejected, i.e. his risk is very small. Hence it is called producers risk where p denotes the proportion of defectives in the lot.

P [the customer will reject a lot with quality p] $=\alpha$ which is very small.

LTPD - Lot tolerance perfect defective. Can also be called as rejection quality level (RQL) and it is denoted as p' . It is the maximum perfect defective that the customer is ready to accept, in an accepted lot. He believes that the probability of accepting a lot which quality more than p is very small.

The probability of quality p' is equal to β .

Process average fraction defective which is observed over a long period of time which is given as \bar{p} .

3.1 OC curve

The OC curve for a sampling inspection plan is a curve that shows probability of acceptance. $P_a(p)$ varies with product quality p is p^* against P_a or p^* against $P_a(p^*)$.

$P_a(p^*)$ can be calculated in any one of the following ways,

1. From a sample of size n we can have d defectives $d = 0, 1, \dots, n$. Each unit has a probability p^* of being a defective, $d \sim B(n, p^*)$.

Therefore, probability of obtaining d defectives out of n is given by pdf of binomial.

And accepted if,

$$P_a(p^*) = \sum_{d=0}^c \binom{n}{d} (p^*)^d (q^*)^{n-d}$$

2. Poisson distribution as an approximation to binomial distribution.

If $n \geq 10$ and $p^* \leq 0.10$ then we can use Poisson distribution as an approximation to binomial distribution.

$$P_a(p') = \sum_{d=0}^c \frac{e^{-np} np^d}{d!}$$

where np^* is mean.

3. If we take the lot size into account then we use hypergeometric distribution.

Let N =lot size, M =no of defectives in the lot= Np^* . Then probability of r defective is $= \frac{\binom{M}{r} \binom{N-M}{n-r}}{\binom{N}{n}}$

$$P_a(p^*) = \sum_{r=0}^c \frac{\binom{Np^*}{r} \binom{N-Np^*}{n-r}}{\binom{N}{n}}$$

The values of p^* are plotted on the x -axis and corresponding $P_a(p^*)$ on y -axis. The locus of points thus obtained is called as *OC*-curve. Such a curve gives the probability of acceptance of a lot for various submitted lot quality p^* and vice versa.

Ideal OC curve: If a particular lot quality say $p^* = p$ is considered to be satisfactory, then producer would wish all the lots with quality better than p to be accepted in the market with probability 1. On the other hand the lot whose lot quality is inferior to p (i.e. $p^* > p$), would be accepted with probability zero.

3.2 Sampling plan for attributes

3.2.1 Single sampling plan (SSP)

In this plan, the decision about accepting or rejecting a lot is taken on the basis of a single sample alone.

Let N =lot size, d =no. of defectives in sample, n =sample size, c =acceptance number.

Single sampling plan is described as follows,

1. Select a random sample of size n from the lot.

2. Inspect all items in the sample. Let d be the number of defectives in the sample.
3. If $d \leq c$ accept the lot replacing all the defectives pieces in the sample by non-defective pieces.
4. If $d > c$, reject the lot. In this case inspect the entire lot and replace all the defective pieces by non-defectives pieces.

Average outgoing quality

The quality of the lot before it passes through the inspection plan is called **incoming quality** of the lot. The quality of the lot after it passes through the inspection plan is called as **outgoing quality** of the lot.

The expected or average fraction defective remaining in the lot after the application of the sampling plan is called as *AOQ*. This is naturally a function of p^* . Curve is obtained by plotting p^* against *AOQ*.

1. n pieces inspected are all good pieces
2. Of the $N - n$ pieces either
 - (a) All are inspected and replaced if necessary (i.e. $d > c$)
 - (b) accepted without further inspection (i.e. if $d \leq c$)

I and II contains all non defective items

But II b is accepted on the basis of the sample. i.e. $N - n$ pieces are passed without inspection. Therefore there will be $(N - n)p^*$ defectives in the lot after inspection. That is, the proportion of defective in the lot after inspection is

$$\frac{(N - n)p^*}{N}$$

and the probability of accepting such a lot is $P_a(p^*)$

Therefore,

$$AOQ = \frac{(N - n)p^*}{N} P_a(p^*)$$

Average outgoing quality limit

The maximum value of *AOQ* when subjected to various values of p^* is called *AOQL*. That is, the maximum value of *AOQ* in the entire range p^* is called *AOQL*. If *AOQL* is at $p' = p_m$ then,

$$AOQL = \frac{(N - n)p_{*m}^*}{N} P_a(p^*)$$

Average total inspection

ATI is the average number of units that are inspected under the plan. It is also a function of incoming lot quality p . It includes inspection of N units if lot is rejected. Curve is obtained by plotting p^* against *ATI*.

For a single S.P. Either i) n units are inspected (if lot is accepted $d \leq c$) or ii) N units are inspected (if lot is rejected $d > c$) with probability of rejection $(1 - P_a(p^*))$.

$$\begin{aligned}\therefore ATI &= nP_a(p^*) + N(1 - P_a(p^*)) \\ &= N - P_a(p^*)[N - n]\end{aligned}$$

Average sample number (ASN)

Expected value of the sample size required for coming to a decision of accepting or rejecting the lot under the sampling inspection plan is called ASN. ASN is a function of lot function defectives (p^*). The ASN curve is obtained by plotting p^* against *ASN*. Hence lower the *ASN* better is the plan.

For SSP, $ASN = n$.

Note: When incoming quality is good p is small the outgoing quality is also small. As p increases *AOQ* (average outgoing quality) also increases but beyond certain limit if p increases then decision of rejection is taken. And suddenly *AOQ* drops to zero.

Hence the maximum *AOQ* is defined as *AOQL*.

3.2.2 Double sampling plan (DSP)

In double sampling plan decision using single sample is not acceptable in majority of the process, also if in the first sample the size is n_1 , the number of defectives is neither too large nor too small for acceptance.

Dodge and ??? suggested the double sampling plan in which decision is taken on the basis of number of defectives found in two samples.

Notations are as follows, N = lot size,

n_1 = size of sample 1,

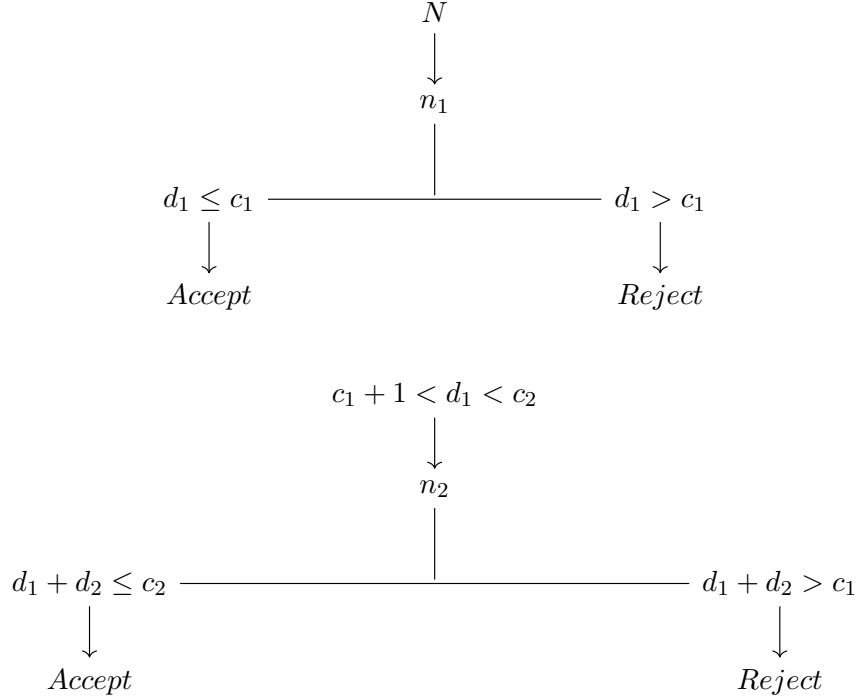
n_2 = size of sample 2,

c_1 = acceptance no for 1st sample and maximum permissible range of defective in 1st sample,

c_2 = acceptance no of sample 1, 2 combined and maximum permissible range of defectives in both samples if lot is accepted,

d_1 =no. of defectives in 1st sample.

d_2 =no. of defectives in 2nd sample.



Double sampling plan on the average requires less amount of inspection than single sampling plan for given values of α (producers risk) and β (consumers risk).

Usually double sampling plan results in 25% to 33% reduction in cost.

OC Curve

If p is incoming quality with incoming fraction defective p .

$$OC = P_a(p) = P(\text{accept on 1st} \cup \text{2nd sample}) = P_{a_1}(p) + P_{a_2}(p) \quad (3.1)$$

$$P_{a_1}(p) = P(d_1 \leq c_1).$$

Let $g(d_1, p)$ be the pmf of d_1 so

$$P_{a_1}(P) = P(d_1 \leq c_1) = \sum_{d_1=0}^{c_1} g(d_1, p) \quad (3.2)$$

$d_1 \rightarrow$ hypergeometric situation (???) represented as N goes to Np and $N - np$. And n_1 going to $d_1, n_1 - d_1$ ¹.

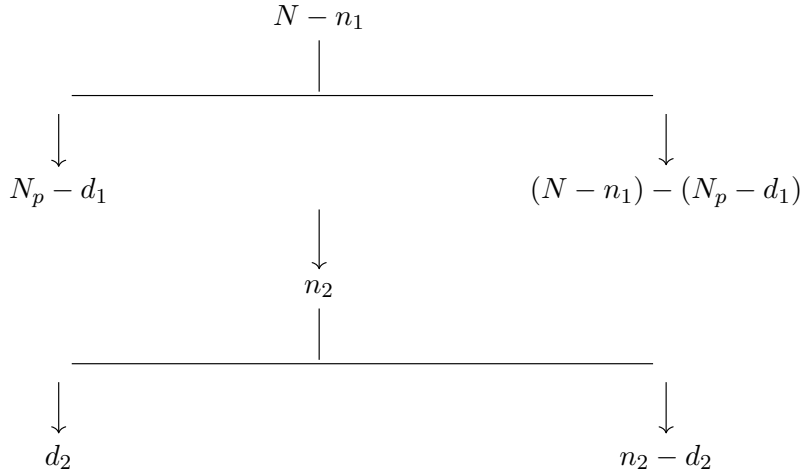
$$P_{a_1}(p) = \sum_{d_1=0}^{c_1} \frac{\binom{Np}{d_1} \binom{N-np}{n_1-d_1}}{\binom{N}{n_1}} \quad (3.3)$$

$$\begin{aligned} P_{a_2}(p) &= P[c_1 + 1 \leq d_1 \leq c_2 \text{ and } d_1 + d_2 \leq c_2] \\ &= P[c_1 + 1 \leq d_1 \leq c_2 \cap d_1 + d_2 \leq c_2] \end{aligned}$$

Let $g(d_1, d_2, p)$ be the joint distribution of d_1 and d_2 so that

$$\begin{aligned} P_{a_2}(p) &= \sum_{d_1=c_1+1}^{c_2} \sum_{d_2=0}^{c_2-d_1} g(d_1, d_2, p) \\ &= \sum_{d_1=c_1+1}^{c_2} \sum_{d_2=0}^{c_2-d_1} g_1(d_1, p) g_2(d_2|d_1) \end{aligned} \quad (3.4)$$

Consider conditional distribution of $d_2|d_1 \rightarrow g_2(d_2|d_1)$



To obtain d_2 in the second sample we have to select a sample of size n_2 from remaining elements i.e. from $N - n_1$ elements. In entire lot of size N Np items were defective. Now to get conditional distribution of $d_2|d_1$.

¹make a flowchart thing later

We know that out of N_p defectives n_1 defectives have already been eliminated. That is now the remaining lot of $N - n_1$ are defective and $(N - n_1) - (N_p - d_1)$ are non defective. Hence the situation is hypergeometric situation.

Hence conditional distribution of $d_2|d_1$ is hypergeometric distribution of d_1

We get,

$$\begin{aligned}
 P_{a_2}(p) &= \sum_{d_1=c_1+1}^{c_2} \sum_{d_2=0}^{c_2-d_1} g_1(d_1, p) g_2(d_2|d_1) \\
 &= \sum_{d_1} \sum_{d_2} \frac{\binom{N_p}{d_1} \binom{N-N_p}{n_1-d_1}}{\binom{N}{n_1}} \times \frac{\binom{N_p-d_1}{2} \binom{N-n_1-N_p+d_1}{n_2-d_2}}{\binom{N-n_1}{n_2}} \\
 P_a(p) &= P_{a_1}(p) + P_{a_2}(p)
 \end{aligned} \tag{3.5}$$

$$\begin{aligned}
 \text{Producers risk} &= P_a(\bar{p}) = P_a(\text{rejecting good quality lot}) \\
 &= 1 - P_a(p)
 \end{aligned}$$

$$\begin{aligned}
 \text{Consumers risk} &= P_o(p_t) = P_a(\text{accept bad quality lot}) \\
 &= \beta
 \end{aligned}$$

$$AOQ = \frac{E[X]}{N}$$

$$\begin{array}{cc}
 x & P(x) \\
 0 & P(\text{reject lot}) = 1 - P_a(p) \\
 (N - n_1)p & P_{a_1}(p) \\
 (N - n_1 - n_2)P & P_{a_2}(p)
 \end{array}$$

$$AOQ = \frac{\sum xP(x)}{N}$$

$ASN = E[x]$, X takes two possible values n_1 if the decision is to be taken on the basis of first sample and n_2 if decision is to be taken on second sample.

$P_{a_1}(p)$ probability of accepting a lot on the basis of first sample and $P_{R_1}(P)$ be the probability of rejecting a lot

$$\begin{aligned}
 P_1 &= P_{a_1}(p) + P_{R_1}(p) \\
 E[x] &= n_1 P_1 + (n_1 + n_2)(1 - P_1)
 \end{aligned}$$

$$\begin{array}{rcl}
 ASN & & E[x] \\
 DC & & P(x) \\
 n_1 & P_1 = P(\text{not going to 2nd sample}) & \\
 n_1 + n_2 & & 1 - P_1
 \end{array}$$

Average total inspection (ATI) which is

$$E[\text{no of units inspected to take decision to accept or reject the lot}]$$

Let X be the number of units inspected to take a decision so that ATI is $E[X]$.

Since X is the total number of units inspected to take a decision. X assumes the value n_1 - if we accept the lot on the basis of first sample with $P_{a_1}(P)$.

$n_1 + n_2$ if Probability of accepting a lot on the basis of second sample $P_{a_2}(P)$.

N if the decision of rejection is taken with the probability of $P_{R_1}(P)$ $1 - P_a(P)$.

$$ATI = n_1(P_{a_1}(p)) + (n_1 + n_2)(P_{a_2}(p)) + a_1(1 - P_a(p))$$