Manufacturing Data Science

Statistical Process Control (SPC) & Run-to-Run Control (RtR)

Chia-Yen Lee, Ph.D. (李家岩 博士)

Department of Information Management (資訊管理學系)
National Taiwan University (國立台灣大學)

Course Contents

- Data Science & Manufacturing Systems
  - Data, Information, Knowledge, and ML/DS Functions
  - Analytics Framework and Data Preprocessing
  - Manufacturing Systems and Factory Dynamics
- Diagnostic and Predictive Analytics
  - Feature Selection and Feature Engineering
  - Regression, Classification, MARS, and Symbolic Regression
  - Tree-based Methods, Random Forest and Boosting
  - Neural Network and Deep Learning
  - **Statistical Process Control & RtR Control**
  - Signal Processing, and PHM
  - Manufacturing Practice
- Prescriptive Analytics
  - Linear Programming and Capacity Planning
  - Metaheuristic Algorithm and Genetic Algorithm
  - Scheduling Optimization
- Advanced Techniques (if time permits)
  - Concept Drift and Domain Adaptation
  - Transfer Learning, Meta-Learning, Few-shot Learning, Small Samples
- Term-project Presentation

Content

- Statistical Process Control (SPC)

y 軸可能是 sensor 或參數，會隨時間跳動
可以針對產品或製程
決定生產或製成參數
也會有一些機台的參數，就是可能需要校正歸零

- Process Capability
- Multivariate Control Chart
- Tool/Chamber Matching
- Advanced Topics (Self-Study)
  - Nonparametric Control Chart
  - Concept Drift
  - Run-to-Run Control

Process Adjustment (製程控管與調整)

☐ Tool Matching- from feature selection to process adjustment

- Theoretical Framework

田口方法、品質工程

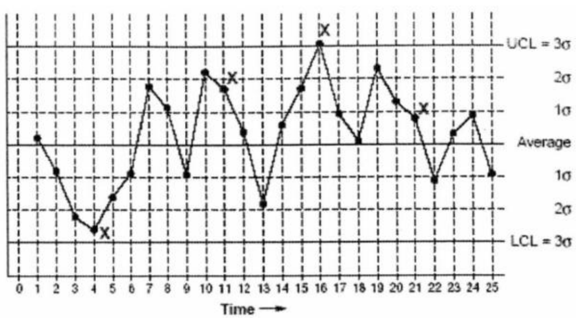
校正機差問題

產品
SPC

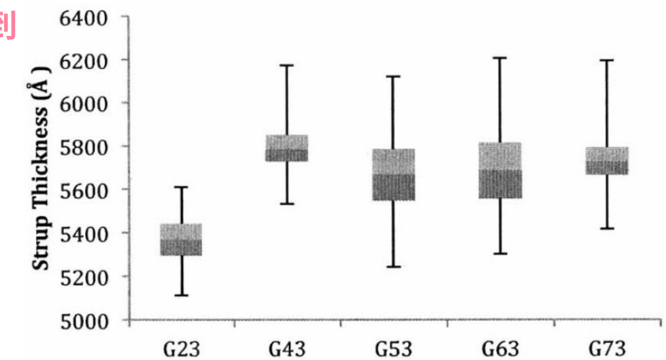
Statistical Process Control → Design of Experiments (DOE) → Tool Matching Virtual Metrology → Run-to-Run Control

在有限的實驗次數下找到最佳的參數組合

每一批進來我應該如何修正逐批補償

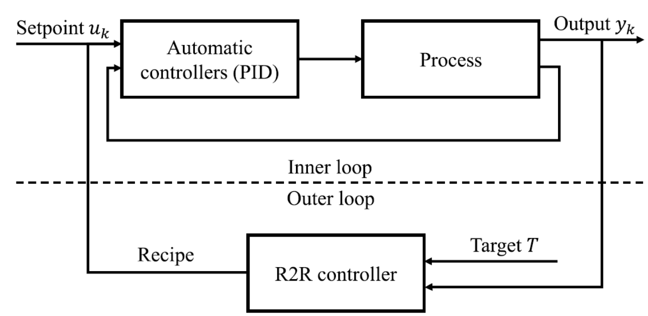


如果是平行機台，但可能有機差



機台
PHM

	A	B	AB	C	AC	BC	ABC
(1)	-1	-1	+1	-1	+1	+1	-1
a	+1	-1	-1	-1	-1	+1	+1
b	-1	+1	-1	-1	+1	-1	+1
ab	+1	+1	+1	-1	-1	-1	-1
c	-1	-1	+1	+1	-1	-1	+1
ac	+1	-1	-1	+1	+1	-1	-1
bc	-1	+1	-1	+1	-1	+1	-1
abc	+1	+1	+1	+1	+1	+1	+1



Haskaraman, F. 2016. Chamber Matching in Semiconductor Manufacturing Using Statistical Analysis and Run-to-Run Control. Master Thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology.

Liu, K., Chen, Y., Zhang, T., Tian, S., & Zhang, X. 2018. A survey of run-to-run control for batch processes. ISA Transactions, 83, 107-125.

Quality Improvement and Statistics

POLab

Definitions of Quality

- Quality means fitness for use
  - quality of design
  - quality of conformance
- Quality is inversely proportional to variability.

著重變異高於平均

Quality Improvement

重要是變異縮減，可以提高穩定度，在應用上可以節省資源

- Quality improvement is the reduction of variability in processes and products.
- Alternatively, quality improvement is also seen as “waste reduction”.

TPS Toyota 主要是在乎品質，是不用不良品，但是近期開始使用不良品，因為物料成本上漲

Statistical process control (SPC)

- is a collection of tools that when used together can result in process stability and variance reduction

Statistical Process Control (SPC)

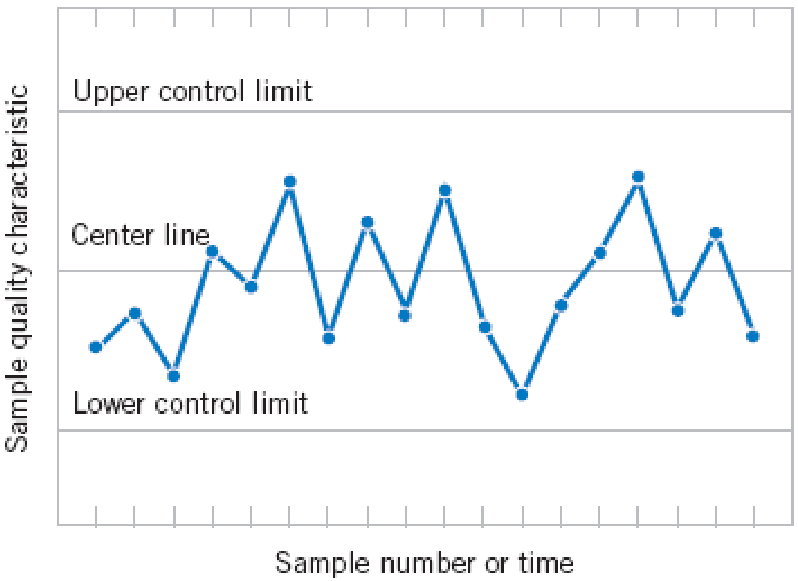
- The seven major tools are
  - 1) Histogram
  - 2) Pareto Chart
  - 3) Cause and Effect Diagram
  - 4) Defect Concentration Diagram
  - 5) Control Chart
  - 6) Scatter Diagram
  - 7) Check Sheet

Introduction to Control Charts

建構上下管制界線來監控製成的變異是不是變大

Basic Principles

- A process that is operating with only chance causes of variation present is said to be in statistical control.
- A process that is operating in the presence of assignable causes is said to be out of control.
- The eventual goal of SPC is the elimination of variability in the process.
- A typical control chart has control limits set at values such that if the process is in control, nearly all points will lie within the upper control limit (UCL) and the lower control limit (LCL).



k 要多大就看這個品管嚴格程度

$$
\begin{aligned}
UCL &= \mu_W + k\sigma_W \\
CL &= \mu_W \\
LCL &= \mu_W - k\sigma_W
\end{aligned}
$$

where

k = distance of the control limit from the center line

μ_w = mean of some sample statistic, W.

σ_w = standard deviation of some statistic, W.

Introduction to Control Charts

Types of control charts

- Variables Control Charts
  - These charts are applied to data that follow a continuous distribution.
- Attributes Control Charts 計數、計次（可數）
  - These charts are applied to data that follow a discrete distribution.

Popularity of control charts

- 1) Control charts are a proven technique for improving productivity.
- 2) Control charts are effective in defect prevention.
- 3) Control charts prevent unnecessary process adjustment.
- 4) Control charts provide diagnostic information.
- 5) Control charts provide information about process capability.

Design of a Control Chart

活塞環

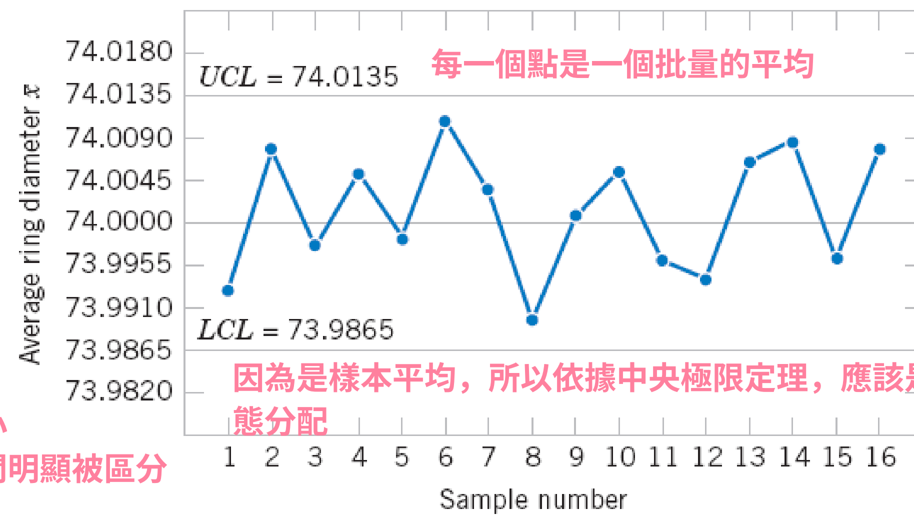
- In manufacturing automobile engine piston rings, the inside diameter of the rings is a critical quality characteristic. 內徑的直徑是重要品管
- Suppose we have a process that we assume the true process mean is $\mu = 74$ and the process standard deviation is $\sigma = 0.01$. Samples of size 5 are taken giving a standard deviation of the sample average, is

標準誤
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{5}} = 0.0045$$
一批生產出來就會產出這批的平均
run to run 才是一個一個樣本

- Control limits can be set at 3 standard deviations from the mean in both directions, i.e. "3-Sigma Control Limits"
  - UCL = 74 + 3(0.0045) = 74.0135
  - CL= 74
  - LCL = 74 - 3(0.0045) = 73.9865
- Choosing the control limits is equivalent to setting up the critical region for hypothesis testing
- $\begin{cases} H_0: \mu = 74 \\ H_1: \mu \neq 74 \end{cases}$

我們會希望一個批量內的變異小
而當有異常，可以從每個點之間明顯被區分

X-bar control chart for piston ring diameter.



每一個點是一個批量的平均
因為是樣本平均，所以依據中央極限定理，應該是常態分配
我們會假設每個點之間是獨立的，所以有點不能說是時間序列

Rational Subgrouping

- Subgroups or samples should be selected so that if assignable causes are present, the chance for differences between subgroups will be maximized, while the chance for differences due to these assignable causes within a subgroup will be minimized.

  組內差距小，異常時組間差距大

- Constructing Rational Subgroups
  - Select consecutive units of production.
    - Provides a “snapshot” of the process.
    - Good at detecting process shifts.
  - Select a random sample over the entire sampling interval.
    - Good at detecting if a mean has shifted
    - out-of-control and then back in-control

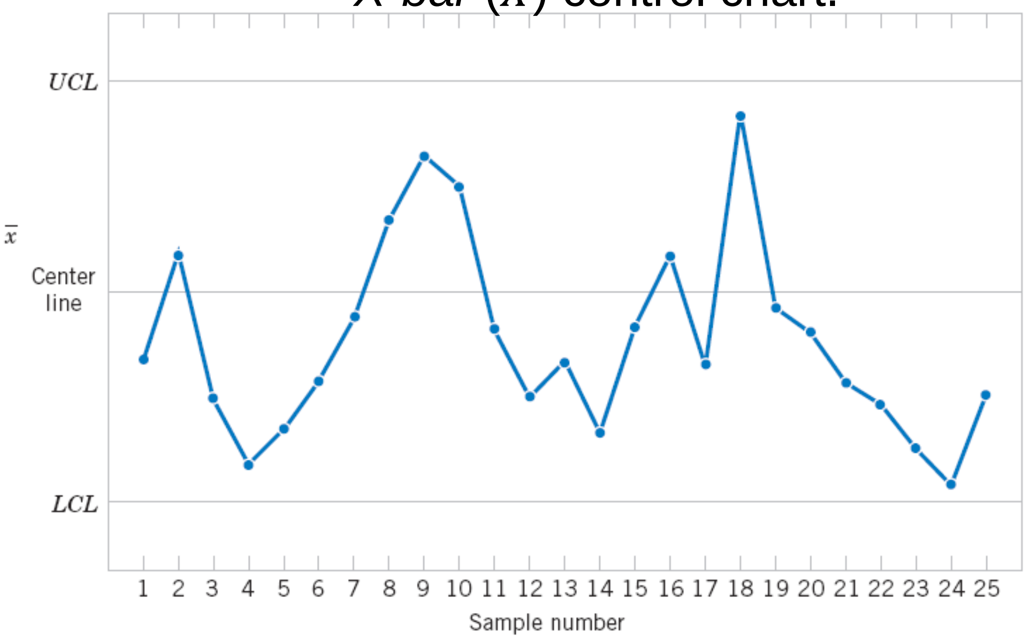
Analysis of Patterns on Control Charts

也可以應用在網路流量的管控

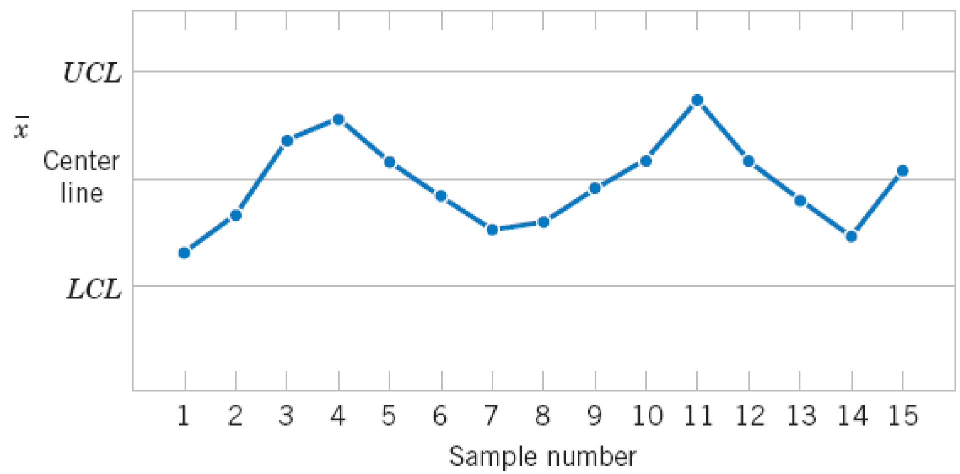
異常不一定是不良品，只是不尋常，在界線內的也不一定就不是異常的

- Look for “runs” - this is a sequence of observations of the same type (all above the center line, or all below the center line)
- Runs of say 8 observations or more could indicate an out-of-control situation.
  - Run up: a series of observations are increasing
  - Run down: a series of observations are decreasing

X-bar ($\bar{X}$) control chart.



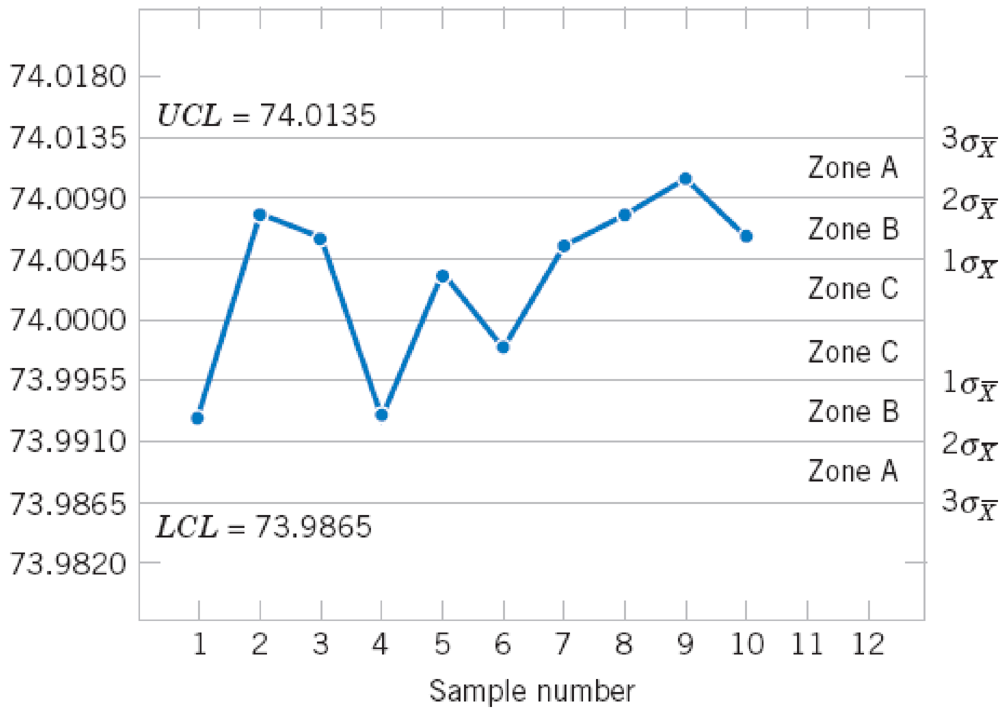
X-bar ($\bar{X}$) chart with a cyclic pattern



Western Electric Handbook Rules (Run Rules)

- A process is considered out of control if any of the following occur:
  - 1) One point plots outside the 3-sigma control limits.
  - 2) Two out of three consecutive points plot beyond the 2-sigma warning limits.
  - 3) Four out of five consecutive points plot at a distance of 1-sigma or beyond from the center line.
  - 4) Eight consecutive points plot on one side of the center line.

通常也就會在 H_0 平均數是否改變有一樣的效果



- $\bar{X}$ chart with 3-sigma control limits:

$$UCL = \mu + 3\sigma/\sqrt{n}$$

x bar 的標準差

$$LCL = \mu - 3\sigma/\sqrt{n}$$

$$CL = \mu$$

樣本估計，用樣本平均的平均估計

- The grand mean: $\hat{\mu} = \bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_i$

大致上就是小於 30

- If the sample size is relatively small…How to build $\bar{X}$ chart?
  - σ can be estimated from either the standard deviation or the range of the observations within each sample.
  - There is little loss in efficiency in estimating σ from the sample ranges.
  - S is a biased estimator of σ.

    樣本的變異數是母體的不偏估計量
    但是標準差是偏估計量，當 sample size 越小偏誤越大

    - sample variance S^2 is an unbiased estimator of the population variance σ^2.
    - S is a biased estimator of σ. For large samples, the bias is very small.
- An unbiased estimator of σ : $\hat{\sigma} = \bar{s}/c$

如果沒有校正，那標準差就會被低估，那品管界線就會比較窄，false alarm 就會很高

c 就是校正因子，一個 constant

n	c
2	0.7979
3	0.8862
4	0.9213
5	**0.9400**
6	0.9515
7	0.9594
8	0.9650
9	0.9693
10	0.9727
15	0.9823
20	0.9869
25	0.9896

應用上會希望 sample size = 1，不要等一個批量都出錯了才被發現
也就是 run to run control

n 越大，c 越接近 1，超過 30 不需要校正
樣本就可以估計母體標準差

$\bar{X}$ and R or S Control Charts

m 個點、m 個批量，每個批量的最大減最小就是 R_i

- The average range: $\bar{R} = \frac{1}{m}\sum_{i=1}^{m} R_i$
  - Since R is a random variable, the quantity $W = R/\sigma$, called the relative range, is also a random variable. The parameters of the distribution of W have been determined for any sample size n. The mean of the distribution of W is called d_2. The standard deviation of W is called d_3. Because $R = \sigma W$.

$$\mu_R = d_2\sigma \qquad \sigma_R = d_3\sigma$$

  - $\bar{R}$ is an estimator of μ_R.
- An unbiased estimator of σ: $\hat{\sigma} = \frac{\bar{R}}{d_2}$

where the constant d_2 is tabulated for various sample sizes in Appendix

- $\bar{X}$ Control Chart (from $\bar{R}$): $UCL = \bar{\bar{X}} + \frac{3}{d_2\sqrt{n}}\bar{R}$ $\quad LCL = \bar{\bar{X}} - \frac{3}{d_2\sqrt{n}}\bar{R}$

Factors for Constructing Variables Control Charts

- Appendix

$$A_2 = \frac{3}{d_2\sqrt{n}}$$

$$D_3 = 1 - 3d_3/d_2$$

$$D_4 = 1 + 3d_3/d_2$$

	Factor for Control Limits						
	$\overline{X}$ Chart			R Chart		S Chart	
n^*	A_1	A_2	d_2	D_3	D_4	c_4	n
2	3.760	1.880	1.128	0	3.267	0.7979	2
3	2.394	1.023	1.693	0	2.575	0.8862	3
4	1.880	.729	2.059	0	2.282	0.9213	4
5	1.596	.577	2.326	0	2.115	0.9400	5
6	1.410	.483	2.534	0	2.004	0.9515	6
7	1.277	.419	2.704	.076	1.924	0.9594	7
8	1.175	.373	2.847	.136	1.864	0.9650	8
9	1.094	.337	2.970	.184	1.816	0.9693	9
10	1.028	.308	3.078	.223	1.777	0.9727	10
11	.973	.285	3.173	.256	1.744	0.9754	11
12	.925	.266	3.258	.284	1.716	0.9776	12
13	.884	.249	3.336	.308	1.692	0.9794	13
14	.848	.235	3.407	.329	1.671	0.9810	14
15	.816	.223	3.472	.348	1.652	0.9823	15
16	.788	.212	3.532	.364	1.636	0.9835	16
17	.762	.203	3.588	.379	1.621	0.9845	17
18	.738	.194	3.640	.392	1.608	0.9854	18
19	.717	.187	3.689	.404	1.596	0.9862	19
20	.697	.180	3.735	.414	1.586	0.9869	20
21	.679	.173	3.778	.425	1.575	0.9876	21
22	.662	.167	3.819	.434	1.566	0.9882	22
23	.647	.162	3.858	.443	1.557	0.9887	23
24	.632	.157	3.895	.452	1.548	0.9892	24
25	.619	.153	3.931	.459	1.541	0.9896	25

$^*n > 25$: $A_1 = 3/\sqrt{n}$ where n = number of observations in sample.

R Chart: 全距是否有變異

- The standard deviation of R should be $\hat{\sigma}_R = d_3\hat{\sigma} = d_3 \dfrac{\bar{R}}{d_2}$
- the upper and lower control limits on the R chart

$$UCL = \bar{R} + \frac{3d_3}{d_2}\bar{R} = \left(1 + \frac{3d_3}{d_2}\right)\bar{R}$$

$$LCL = \bar{R} - \frac{3d_3}{d_2}\bar{R} = \left(1 - \frac{3d_3}{d_2}\right)\bar{R}$$

$$UCL = D_4\bar{r} \qquad CL = \bar{r} \qquad LCL = D_3\bar{r}$$

- The LCL for an R chart can be a negative number.
  - It is customary to set LCL to zero. Because the points plotted on an R chart are nonnegative, no points can fall below an LCL of zero.

 x bar chart 是管平均是不是有偏移

標準差的管制，管變異，製程變異是不是忽大忽小

S Chart:

- Calculate the standard deviation of each subgroup and plot these standard deviations to monitor the process standard deviation σ. Called an S chart.
- S is a biased estimator of σ.
  - the sample variance S^2 is an unbiased estimator of the population variance σ^2.
  - S is a biased estimator of σ. For large samples, the bias is very small. However, there are good reasons for using S as an estimator of σ in samples from normal distributions.

3-sigma control limits for S:

c_4 校正因子

$$LCL = c_4\sigma - 3\sigma\sqrt{1 - c_4^2} \qquad CL = c_4\sigma. \qquad UCL = c_4\sigma + 3\boxed{\sigma\sqrt{1 - c_4^2}}$$

標準差的標準差

$$\bar{S} = \frac{1}{m}\sum_{i=1}^{m} S_i$$

An unbiased estimator of σ: $\hat{\sigma} = \bar{S}/c_4$

Because $E(\bar{S}) = c_4\sigma$, an unbiased estimator of σ is $\bar{S}/c_4$

where the constant c_4 is tabulated for various sample sizes in Appendix

$\bar{X}$ and R or S Control Charts

- S Chart:

$$UCL = \bar{s} + 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2} \qquad CL = \bar{s} \qquad LCL = \bar{s} - 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$$

- $\bar{X}$ Control Chart (from $\bar{S}$):

$$UCL = \bar{\bar{x}} + 3\frac{\bar{s}}{c_4\sqrt{n}} \qquad CL = \bar{\bar{x}} \qquad LCL = \bar{s} - 3\frac{\bar{s}}{c_4\sqrt{n}}$$

Example

A component part for a jet aircraft engine is manufactured by an investment casting process. The vane opening on this casting is an important functional parameter of the part. We will illustrate the use of $\overline{X}$ and R control charts to assess the statistical stability of this process. Table presents 20 samples of five parts each. The values given in the table have been coded by using the last three digits of the dimension; that is, 31.6 should be 0.50316 inch.

The quantities $\overline{\overline{x}} = 33.3$ and $\overline{r} = 5.8$ are shown at the foot of Table . The value of A_2 for samples of size 5 is $A_2 = 0.577$. Then the trial control limits for the $\overline{X}$ chart are

$$\overline{\overline{x}} \pm A_2\overline{r} = 33.32 \pm (0.577)(5.8) = 33.32 \pm 3.35$$

這邊是用全距作範例

or

$$UCL = 36.67 \qquad LCL = 29.97$$

$\bar{X}$ and R or S Control Charts

Example- Vane-Opening Measurements

每五個點構成一個批量

Sample Number	x_1	x_2	x_3	x_4	x_5	$\bar{x}$	r	s
1	33	29	31	32	33	31.6	4	1.67332
2	33	31	35	37	31	33.4	6	2.60768
3	35	37	33	34	36	35.0	4	1.58114
4	30	31	33	34	33	32.2	4	1.64317
5	33	34	35	33	34	33.8	2	0.83666
6	38	37	39	40	38	38.4	3	1.14018
7	30	31	32	34	31	31.6	4	1.51658
8	29	39	38	39	39	36.8	10	4.38178
9	28	33	35	36	43	35.0	15	5.43139
10	38	33	32	35	32	34.0	6	2.54951
11	28	30	28	32	31	29.8	4	1.78885
12	31	35	35	35	34	34.0	4	1.73205
13	27	32	34	35	37	33.0	10	3.80789
14	33	33	35	37	36	34.8	4	1.78885
15	35	37	32	35	39	35.6	7	2.60768
16	33	33	27	31	30	30.8	6	2.48998
17	35	34	34	30	32	33.0	5	2.00000
18	32	33	30	30	33	31.6	3	1.51658
19	25	27	34	27	28	28.2	9	3.42053
20	35	35	36	33	30	33.8	6	2.38747
						$\bar{\bar{x}} = 33.32$	$\bar{r} = 5.8$	$\bar{s} = 2.345$

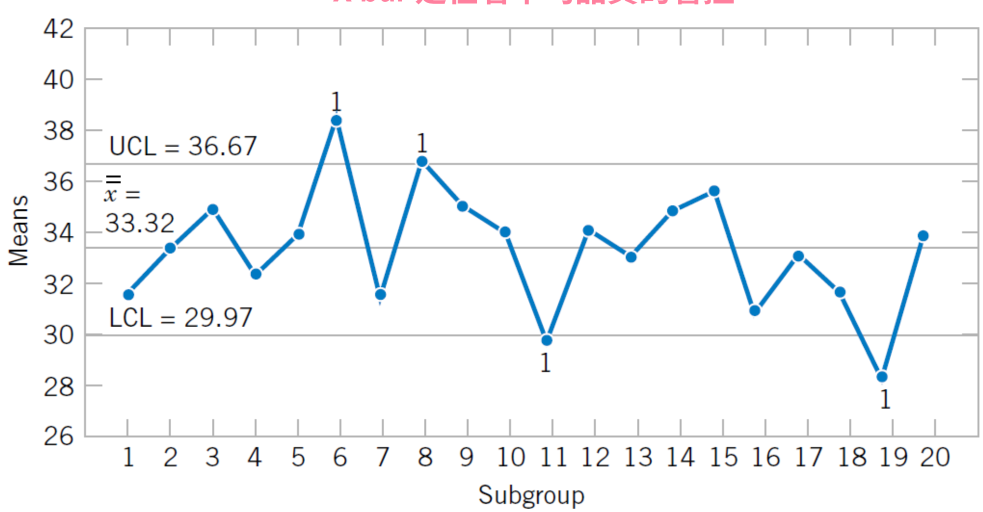
Example (cont.)

For the R chart, the trial control limits are

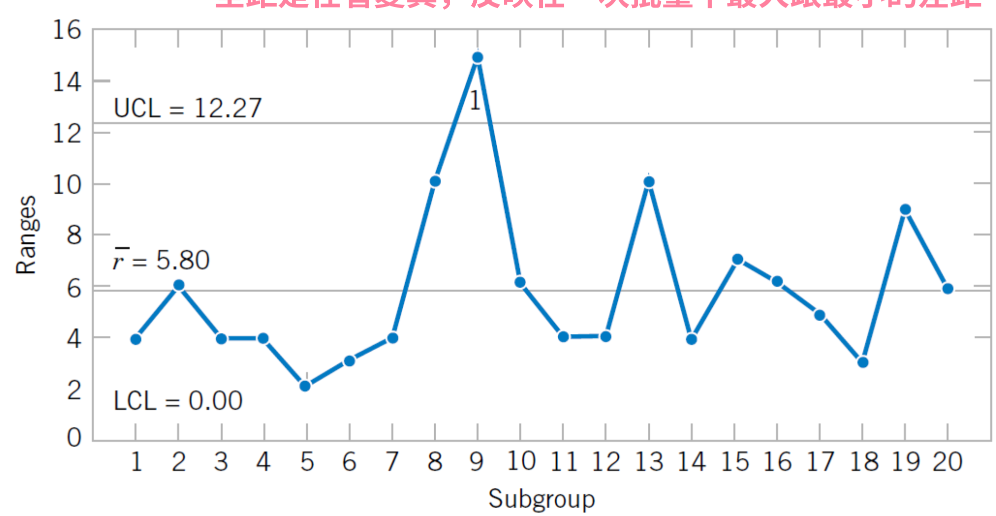
$$UCL = D_4\bar{r} = (2.115)(5.8) = 12.27$$
$$LCL = D_3\bar{r} = (0)(5.8) = 0$$

The $\overline{X}$ and R control charts with these trial control limits are shown in Fig. Notice that samples 6, 8, 11, and 19 are out of control on the $\overline{X}$ chart and that sample 9 is out of control on the R chart. (These points are labeled with a "1" because they violate the first Western Electric rule.)

X bar 是在看平均品質的管控



全距是在管變異，反映在一次批量中最大跟最小的差距



The $\bar{X}$ and R control charts for vane opening.

Example (cont.)

For the S chart, the value of $c_4 = 0.94$. Therefore,

$$\frac{3\bar{s}}{c_4}\sqrt{1 - c_4^2} = \frac{3(2.345)}{0.94}\sqrt{1 - 0.94^2} = 2.553$$

and the trial control limits are

$$UCL = 2.345 + 2.553 = 4.898$$
$$LCL = 2.345 - 2.553 = -0.208$$

The LCL is set to zero. If $\bar{s}$ is used to determine the control limits for the $\overline{X}$ chart,

$$\bar{\bar{x}} \pm \frac{3\bar{s}}{c_4\sqrt{n}} = 33.32 \pm \frac{3(2.345)}{0.94} = 33.32 \pm 3.35$$

有注意到這個校正因子所帶來的影響嗎？
若沒有除上0.94會如何？

$\bar{X}$ and R or S Control Charts

and this result is nearly the same as from $\bar{r}$. The S chart is shown in Fig. Because the control limits for the $\bar{X}$ chart calculated from $\bar{s}$ are nearly the same as from $\bar{r}$, the chart is not shown.

Suppose that all of these assignable causes can be traced to a defective tool in the wax-molding area. We should discard these five samples and recompute the limits for the $\bar{X}$ and R charts. These new revised limits are, for the $\bar{X}$ chart,

eg. use only 1 sample

把異常拿掉再做一次，因為品管圖是針對良品的，如果原本有 outlier 會導致偏移，就要先刪掉再做一次

$$UCL = \bar{\bar{x}} + A_2\bar{r} = 33.21 + (0.577)(5.0) = 36.10$$

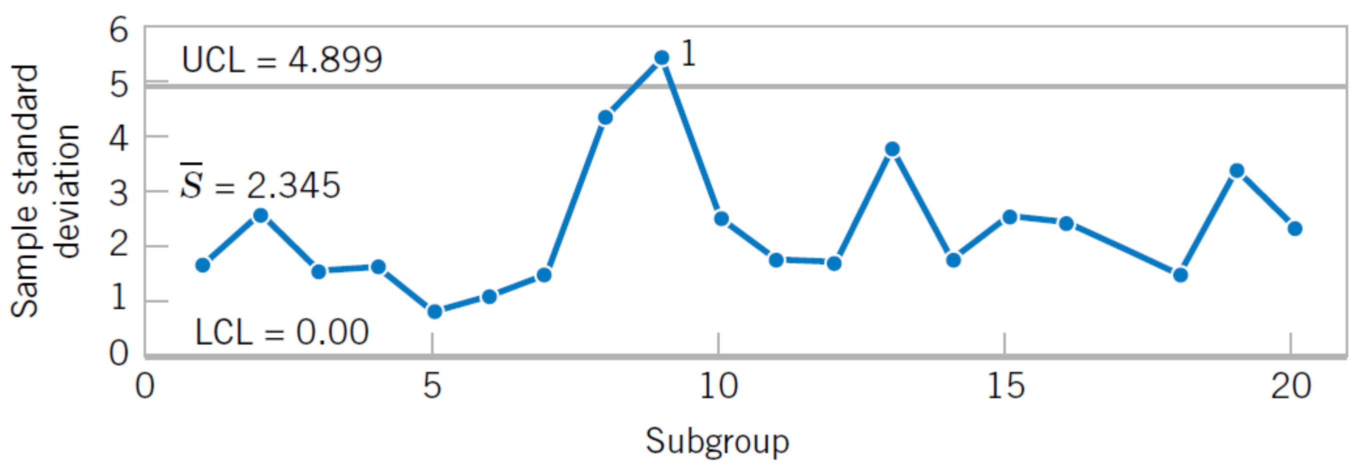
$$LCL = \bar{\bar{x}} - A_2\bar{r} = 33.21 - (0.577)(5.0) = 30.33$$

and for the R chart,

$$UCL = D_4\bar{r} = (2.115)(5.0) = 10.57$$

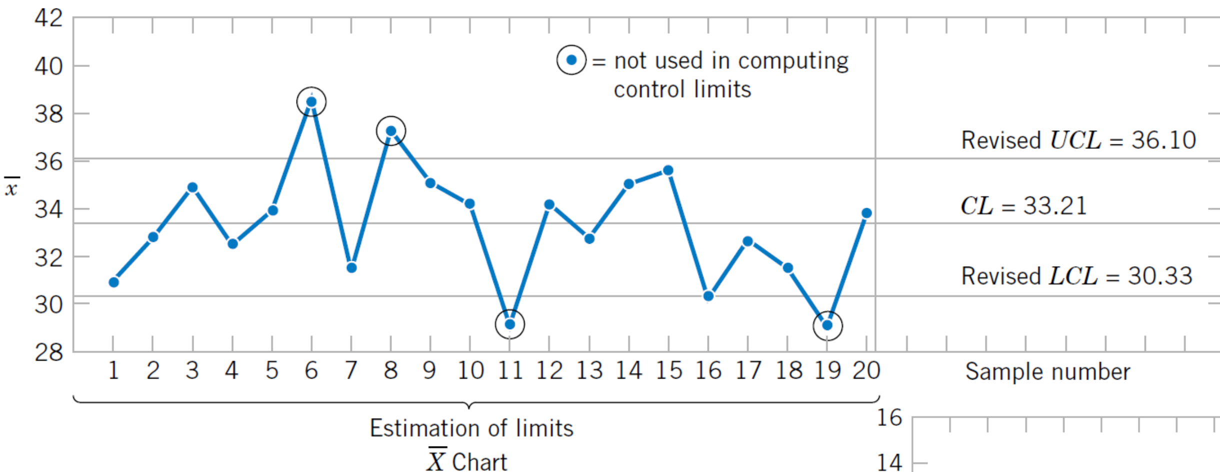
$$LCL = D_3\bar{r} = (0)(5.0) = 0$$

The S control chart for vane opening.



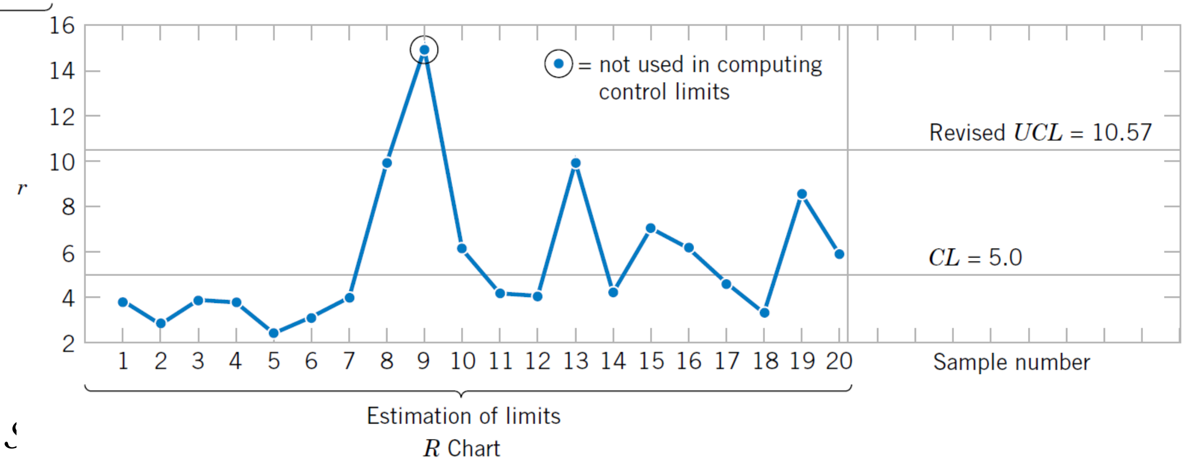
Example (cont.)

The revised control charts are shown in Fig.   Notice that we have treated the first 20 preliminary samples as **estimation data** with which to establish control limits. These limits can now be used to judge the statistical control of future production. As each new sample becomes available, the values of $\bar{x}$ and r should be computed and plotted on the control charts. It may be desirable to revise the limits periodically, even if the process remains stable. The limits should always be revised when process improvements are made.



periodically retrain model
要週期性的重訓練，即使是穩定
但只要有變異越來越大，不再適用，就一定要更新重新訓練

The $\bar{X}$ and R control charts for vane opening, revised limits.



Control Charts for Individual Measurements

- What if you could not get a sample size greater than 1 (n =1)?
- Examples

例如要求客製化、個人化
但可能就要換線

  - Automated inspection and measurement technology is used, and every unit manufactured is analyzed.
  - The production rate is very slow, and it is inconvenient to allow samples sizes of N > 1 to accumulate before analysis.
  - Repeat measurements on the process differ only because of laboratory or analysis error, as in many chemical processes.
  - In process plants, such as papermaking, measurements on some parameters such as coating thickness across the roll will differ very little and produce a standard deviation that is much too small if the objective is to control coating thickness along the roll.

- Potential for future: Industry 4.0, run-to-run (R2R) control

批量與單一樣本，差別就像機器學習更新權重方向可以是批次更新，也有一筆一筆更新，但一筆一筆更新如果遇到 outlier 那方向可能會有急遽的變化，或是有偏掉，且花比較多時間；如果是批次，那比較穩定

- The individual control charts are useful for samples of sizes n = 1.

- The moving range (MR) is defined as the absolute difference between two successive observations:

$$MR_i = |x_i - x_{i-1}|$$

which will indicate possible shifts or changes in the process from one observation to the next.

- The center line and upper and lower control limits for a control chart for individuals are

n=1 就要用 MR，只有批量才可以用 x_bar
但是 x 本身就不一定符合常態分配
x_bar 是因為中央極限定理

$$UCL = \bar{x} + 3\frac{\overline{mr}}{d_2} = \bar{x} + 3\frac{\overline{mr}}{1.128}$$
$$CL = \bar{x}$$
$$LCL = \bar{x} - 3\frac{\overline{mr}}{d_2} = \bar{x} - 3\frac{\overline{mr}}{1.128}$$

針對 mu
管平均

- and for a control chart for moving ranges

$$UCL = D_4\overline{mr} = 3.267\overline{mr}$$
$$CL = \overline{mr}$$
$$LCL = D_3\overline{mr} = 0$$

針對 range
管變異

- Example

Table shows 20 observations on concentration for the output of a chemical process. The observations are taken at one-hour intervals. If several observations are taken at the same time, the observed concentration reading will differ only because of measurement error. Since the measurement error is small, only one observation is taken each hour.

To set up the control chart for individuals, note that the sample average of the 20 concentration readings is $\bar{x} = 99.1$ and that the average of the moving ranges of two observations shown in the last column of Table is $\overline{mr} = 2.59$. To set up the moving-range chart, we note that $D_3 = 0$ and $D_4 = 3.267$ for $n = 2$. Therefore, the moving-range chart has center line $\overline{mr} = 2.59$, $LCL = 0$, and $UCL = D_4\overline{mr} = (3.267)(2.59) = 8.46$. The control chart is shown as the lower control chart in Fig.

Because no points exceed the upper control limit, we may now set up the control chart for individual concentration measurements. If a moving range of $n = 2$ observations is used, $d_2 = 1.128$. For the data in Table we have

$$UCL = \bar{x} + 3\frac{\overline{mr}}{d_2} = 99.1 + 3\frac{2.59}{1.128} = 105.99$$

$$CL = \bar{x} = 99.1$$

$$LCL = \bar{x} - 3\frac{\overline{mr}}{d_2} = 99.1 - 3\frac{2.59}{1.128} = 92.21$$

The control chart for individual concentration measurements is shown as the upper control chart in Fig. There is no indication of an out-of-control condition.

Control Charts for Individual Measurements

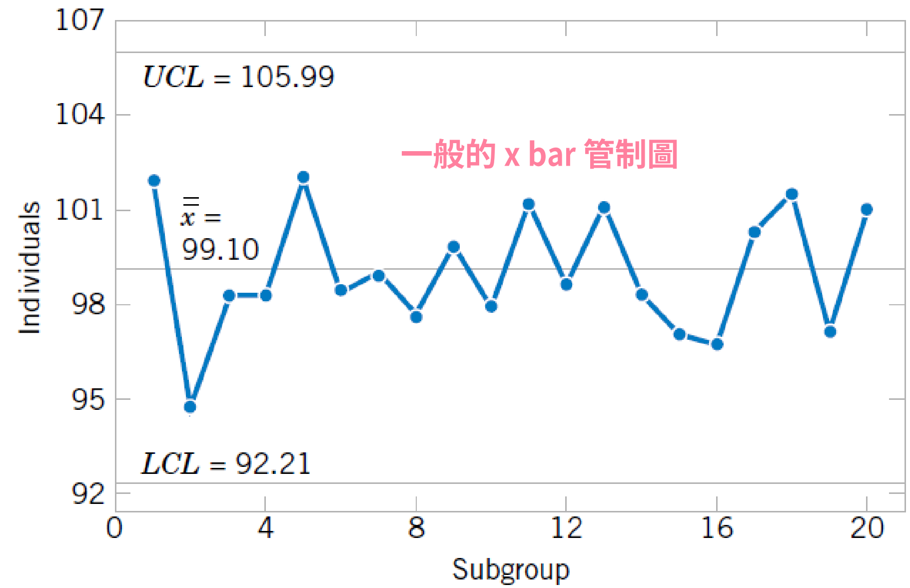
Example (cont.)

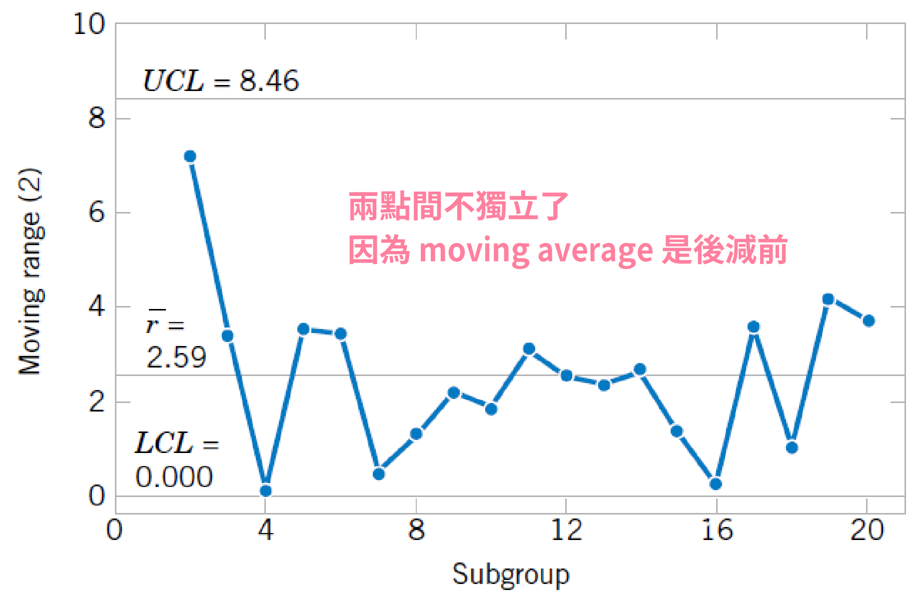
N=1 moving average 就是後減前取絕對值

Chemical Process Concentration Measurement

Observation	Concentration x	Moving Range mr
1	102.0	
2	94.8	7.2
3	98.3	3.5
4	98.4	0.1
5	102.0	3.6
6	98.5	3.5
7	99.0	0.5
8	97.7	1.3
9	100.0	2.3
10	98.1	1.9
11	101.3	3.2
12	98.7	2.6
13	101.1	2.4
14	98.4	2.7
15	97.0	1.4
16	96.7	0.3
17	100.3	3.6
18	101.4	1.1
19	97.2	4.2
20	101.0	3.8
	$\bar{x} = 99.1$	$\overline{mr} = 2.59$

Control charts for individuals and the moving range.





Control Charts for Individual Measurements

Interpretation of the Charts

資訊月少能做的處理越少

- X Charts can be interpreted similar to $\bar{X}$ charts. MR charts cannot be interpreted the same as $\bar{X}$ or R charts.

  MR 是後減前

- Since the MR chart plots data that are "correlated" with one another, then looking for patterns on the chart does not make sense. (WHY? too noisy!)
- MR chart cannot really supply useful information about process variability.
- More emphasis should be placed on interpretation of the X chart.

每個點是 x，畫的是 x chart 而不是 x bar chart，那上下三倍標準差是有意義的嗎？所以 variation 解釋很有限

管制圖顯示的是 x bar 的分布

Note

- MR control chart for individuals is very insensitive to small shifts in the process mean.

  個別的點偏一對整個製程是不敏感的

  individual 可能存在白噪音的干擾，不一定可以跟異常變異區隔

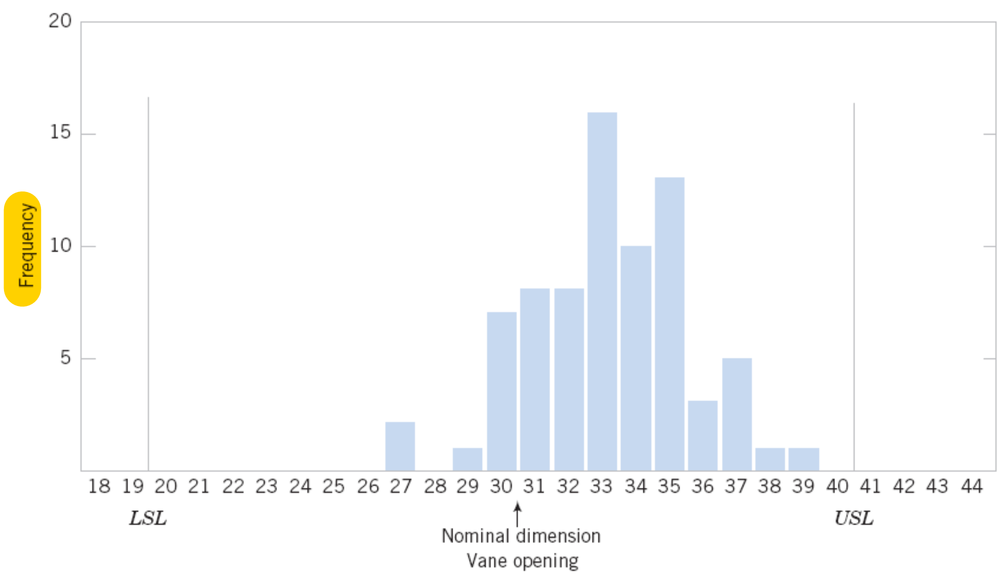
  - For example, if the size of the shift in the mean is one standard deviation, the average number of points to detect this shift is 43.9 (ARL) in MR control chart.
  - While the performance of the MR control chart for individuals is much better for large shifts.

  還勢必較常用 CUSUN EWMA

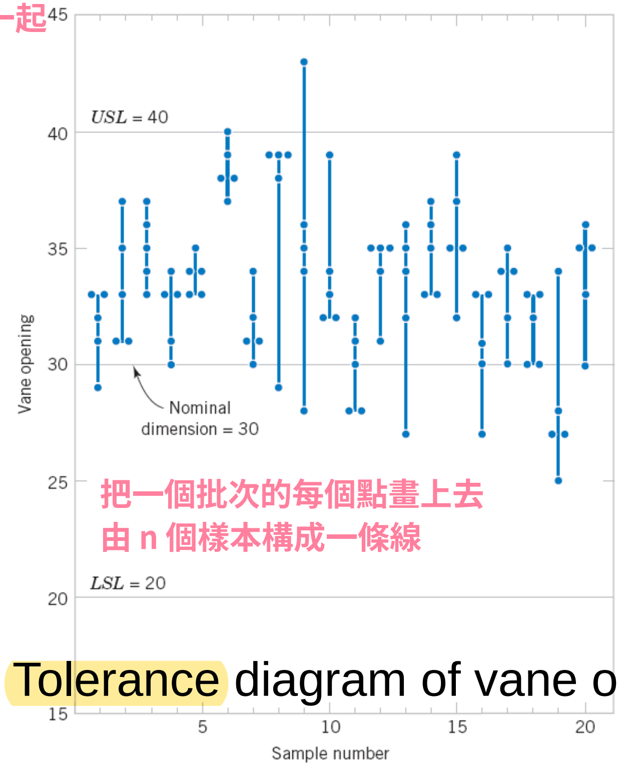
- The shift is not large but more rapid shift detection is desirable.
  - recommend the cumulative sum (CUSUM) control chart or an exponentially weighted moving-average (EWMA) chart (Montgomery, 2001).

Process Capability (Montgomery & Runger, 2018)

- Process capability refers to the performance of the process when it is operating in control. USL、LSL， S 就是指規格，例如商品的長寬
  spec 就是規格，上下規格界線
- Two graphical tools are helpful in assessing process capability:
  - Tolerance chart (or tier chart) and Histogram
  - right ← (V) Plot the specification limits on the tolerance chart, since it is a chart of individual measurements 規格
    規格不等於 control limit，control 是指製程的 recipe 溫度壓力等，是控制參數、製程能力
  - wrong ← (X) Plot specification limits on a control chart or to use the specifications in determining the control limits. Specification limits and control limits are unrelated.
    規格的資料不可以畫在 control chart 上，USL 跟 UCL 不可以劃在一起



Histogram for vane openings.



把一個批次的每個點畫上去
由 n 個樣本構成一條線

Tolerance diagram of vane openings.

Process Capability 製程能力跟規格有關

如果 USL - LSL 很小，代表對規格結果很嚴格，可以接受或要求的範圍很窄

Process Capability Ratio (PCR)

- Capability of Precision (C_p) (信度reliability)
- If the process is centered… 是對稱的

$$PCR = \frac{USL-LSL}{6\sigma}$$

規格的區間／容忍度

製程能力的變異

越大越好

PCR_k (C_{pk})

中心點沒有對齊，
我的規格對於理想值往上或往下的罰值嚴重程度不一樣
所以可以接受的上下不對稱，所以是單邊的檢查

- If the process is running off-center…

$$PCR_k = \min\left[\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right]$$

where $\hat{\mu}$ is sample average.

規格為單邊時

要嘛跟上界或下界有關（檢查）

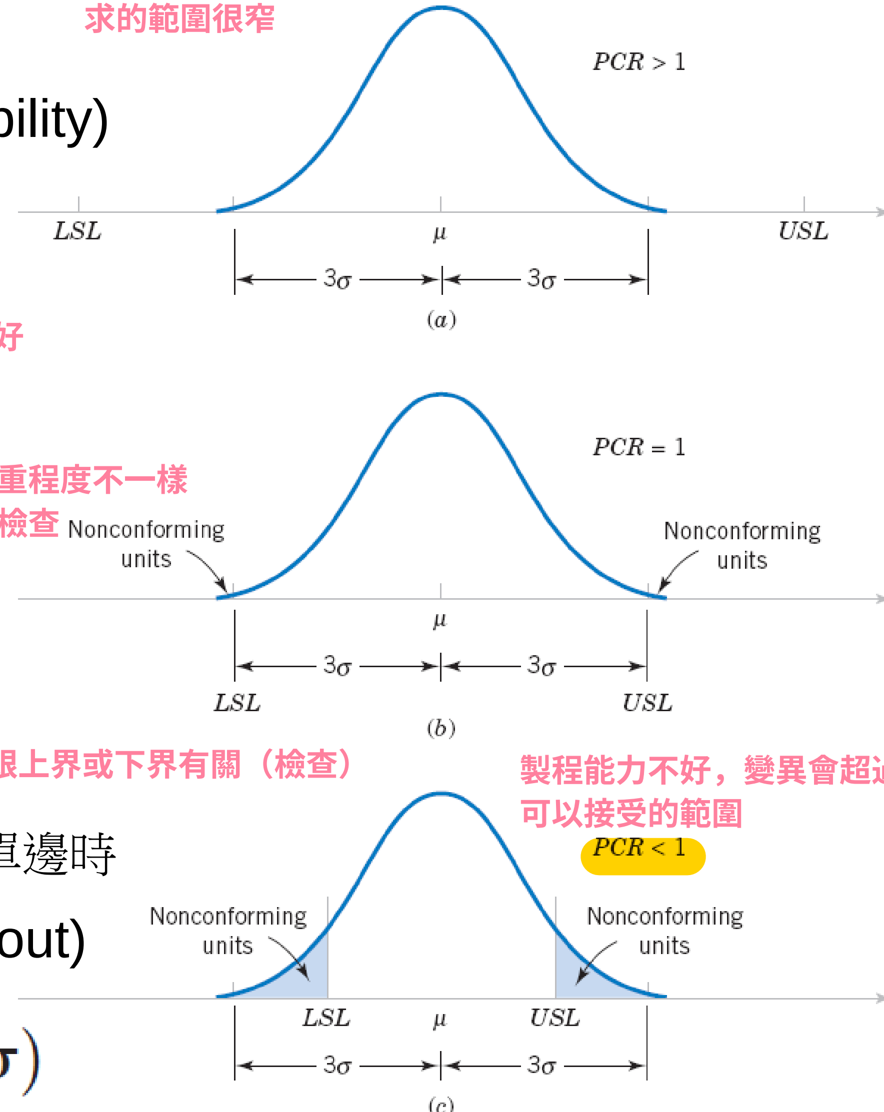
製程能力不好，變異會超過我可以接受的範圍

Fractions of nonconforming output (or fallout)

可以算 CDF

轉換 z score

$$P(X < LSL) = P(Z < (LSL - \mu)/\sigma)$$
$$P(X > USL) = P(Z > (USL - \mu)/\sigma)$$



Process Fallout and the process capability ratio (PCR).

Process Capability

Example

- For an electronic manufacturing process a current has specifications of 100±10 milliamperes. The process mean μ and standard deviation σ are 107.0 and 1.5, respectively. The process mean is nearer to the USL. Consequently,

規格

USL　LSL

信度 $$PCR = (110 - 90)/(6 \cdot 1.5) = 2.22$$

製程能力好

(1/2.22)x100%=45% the % of the specifications' width used by the process.

效度 $$PCR_k = (110 - 107)/(3 \cdot 1.5) = 0.67$$

從製程的平均其實有漂掉，既可以從 PCRk 看出來

製程能力只用到規格寬度的 45%
六倍 sigma 只佔 LSL 到 USL 這段寬度的 45%

- The small PCR_k indicates that the process is likely to produce currents outside of the specification limits. From the normal distribution…

$$P(X < LSL) = P(Z < (90 - 107)/1.5) = P(Z < -11.33) = 0$$

幾乎不會小於 LSL

$$P(X > USL) = P(Z > (110 - 107)/1.5) = P(Z > 2) = 0.023$$

– *Practical Interpretation*: The relatively large probability of exceeding the USL is a warning of potential problems with this criterion even if none of the measured observations in a preliminary sample exceeds this limit. The PCR_k would improve if the process mean were centered in the specifications at 100 milliamperes.

Process Capability

製程能力是憶一起看平均值跟變異，也就是信度效度一起看

parts per million (PPM)

PCR Related to PPM for a Normally Distributed Process

PCR	不良品的機率 P(defective product)	PPM Mean Centered	PPM Mean Shifted 1.5σ
0.5	0.1336	133,614.4	501,349.9
0.67		44,431.2	305,249.8
0.75		24,448.9	226,715.8
1	0.0027=2700ppm	2,699.8	66,810.6
1.25		176.8	12,224.5
1.33	0.000065=65ppm	66.1	6,387.2
1.5		6.8	1,349.9
1.67	0.62ppm	0.5	224.1
1.75		0.2	88.4
2 六標準差	0.0018ppm	0.0	3.4

一百萬當中有這些都是不良

Poor

Good

Excellent

製程能力差不良品會超多

有飄移，不良品也會較多

但是只要製程能力夠好，其實問題會比較不嚴重

在檢定中，平均數已經飄調

當 USL 跟 LSL 很寬，而我的 CP 峰態很高，那就算稍微有點偏移，但還是在很寬的 USL LSL 的範圍內

所以製程能力是指就算已經偏移，但只要能力好，不良品還是很少，因此也呈現出變異小的大優勢

Process Capability

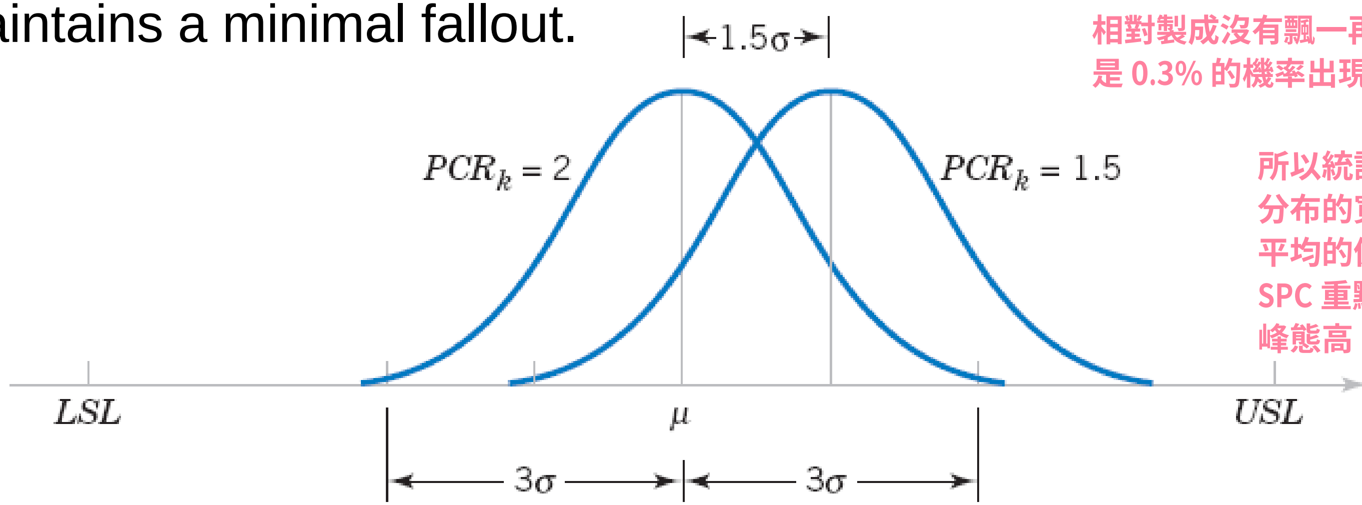
6-sigma process

- A process with PCR_k = 2.0 is referred to as a **6-sigma process** because the distance from the process mean to the nearest specification is 6 standard deviations.
- If the process mean shifts off-center by 1.5 standard deviations, the PCR_k decreases to

$$PCR_k = \frac{USL - \mu}{3\sigma} = \frac{6\sigma - 1.5\sigma}{3\sigma} = \frac{4.5\sigma}{3\sigma} = 1.5$$

- Assuming a normally distributed process, the fallout of the shifted process is **3.4 parts per million**. Consequently, the result still maintains a minimal fallout.

相對製成沒有飄一再常態分布下
是 0.3% 的機率出現不良



所以統計製程管制在意的是常態
分布的寬度也就是變異，而不是
平均的偏移
SPC 重點是 variance reduction
峰態高

Mean of a 6-sigma process shifts by 1.5 standard deviations.

POLab

Capability of Accuracy (k or C_a) (效度validity)

- Measure the degree of center shifted…as smaller as better…
- $k = \frac{\mu - m}{(USL - LSL)/2}$, where m is the specification midpoint, $m = (USL + LSL)/2$ (規格中心).

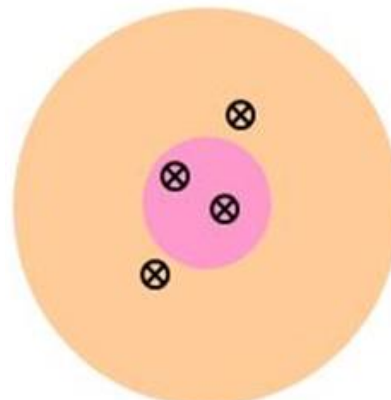
只要 mu 在 USL LSL 之間，可以用 K 把 cp 跟 cpk 的關係串起來

- If $LSL < \mu < USL$, then $|k| \leq 1$. We have a relation **C_{pk}=(1- $|k|$) C_p** and C_p as an upper limit of C_{pk}, which as larger as better…

靶心沒有校正

效度低、信度高
k差、Cp好

打得都很近



效度高、信度低
k好、Cp差

Taguchi Capability Index

- Estimates process capability around a target m. Assumes process output is approximately normally distributed. C_{pm} is also as the Taguchi capability index.

信度

- $C_{pm} = C_p / \sqrt{1 + \left(\frac{\mu - m}{\sigma}\right)^2}$

效度

可以同時看信度跟效度的指標

數據科學模型預測模型也會有信度效度的問題

P Chart (Control Chart for Proportions)

比例，是從二項分布來的

$$UCL = p + 3\sqrt{\frac{p(1-p)}{n}} \qquad LCL = p - 3\sqrt{\frac{p(1-p)}{n}}$$

- Fraction-defective control chart (for fraction nonconforming)
  - n is sample size. D is the number of defective units in the sample.
  - D is a binomial random variable with unknown parameter p.

    總共有多少不良
  - The fraction defective $p = D/n$ of each sample is plotted on the chart.
  - Variance: $\sigma_p^2 = \frac{p(1-p)}{n}$
- Suppose m preliminary samples each of size n are available, and let D_i be the number of defectives in the ith sample. That is
  - $\bar{p} = \frac{1}{m}\sum_{i=1}^{m} p_i = \frac{1}{mn}\sum_{i=1}^{m} D_i$, where $\bar{p}$ may be used as an estimator of p.

P Chart

The center line and upper and lower control limits for the P chart are

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \qquad CL = \bar{p} \qquad LCL = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

where $\bar{p}$ is the observed value of the average fraction defective.

比例

□ *P* Chart Example: Ceramic Substrate

We wish to construct a fraction-defective control chart for a ceramic substrate production line. We have 20 preliminary samples, each of size 100; the number of defectives in each sample is shown in Table. Assume that the samples are numbered in the sequence of production. Note that $\bar{p} = (800/2000) = 0.40$; therefore, the trial parameters for the control chart are

800 是每個 sample 的不良品數量加總

$$UCL = 0.40 + 3\sqrt{\frac{(0.40)(0.60)}{100}} = 0.55 \qquad CL = 0.40$$

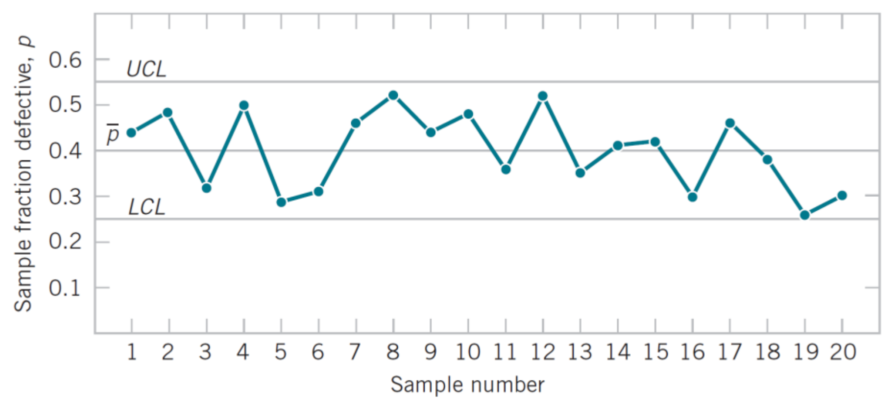
$$LCL = 0.40 - 3\sqrt{\frac{(0.40)(0.60)}{100}} = 0.25$$

The control chart is shown in Figure. All samples are in control. If they were not, we would search for assignable causes of variation and revise the limits accordingly. This chart can be used for controlling future production.

Number of Defectives in Samples of 100 Ceramic Substrates

Sample	No. of Defectives	Sample	No. of Defectives
1	44	11	36
2	48	12	52
3	32	13	35
4	50	14	41
5	29	15	42
6	31	16	30
7	46	17	46
8	52	18	38
9	44	19	26
10	48	20	30

監控比例是不是有異常



P chart for a ceramic substrate.

PO Lab

U Chart (Control Chart for Defects per Unit)

$$UCL = \lambda + 3\sqrt{\frac{\lambda}{n}} \qquad LCL = \lambda - 3\sqrt{\frac{\lambda}{n}}$$

- defects-per-unit chart
  - a hospital might record the number of cases of infection per month 會以一個晶圓為單位
  - a semiconductor manufacturer record the # of large contamination particles per wafer.
- If each subgroup consists of *n* units and there are *C* total defects in the subgroup, then, $U = C/n$ is the average number of defects per unit.
  - Poisson distribution with mean λ, and the variance equals λ.
  - Each point on the chart is an observed value of *U* (the avg. # of defects per unit) from a sample of *n* units. The mean of *U* is λ and variance of *U* is λ/n.
- Suppose *m* preliminary samples each of size *n* are available, and let C_i be the number of defectives in the ith sample. That is
  - $\overline{U} = \frac{1}{m}\sum_{i=1}^{m} U_i = \frac{1}{mn}\sum_{i=1}^{m} C_i$, where $\overline{U}$ may be used as an estimator of λ.

稀少事件的發生
應該符合 poisson
只有 lambda 這個參數
lambda 就是 rate
一段時間有幾個，
類似頻率

U Chart

The center line and upper and lower control limits on the *U* chart are

$$UCL = \overline{u} + 3\sqrt{\frac{\overline{u}}{n}} \qquad CL = \overline{u} \qquad LCL = \overline{u} - 3\sqrt{\frac{\overline{u}}{n}}$$

where $\overline{u}$ is the average number of defects per unit.

U Chart Example: Printed Circuit Boards

因為 sample size 是 5，所以 #defect 除以五

Printed circuit boards are assembled by a combination of manual assembly and automation. Surface mount technology (SMT) is used to make the mechanical and electrical connections of the components to the board. Every hour, five boards are selected and inspected for process-control purposes. The number of defects in each sample of five boards is noted. Results for 20 samples are shown in Table.

The center line for the U chart is

$$\bar{u} = \frac{1}{20}\sum_{i=1}^{20} u_i = \frac{32.0}{20} = 1.6$$

and the upper and lower control limits are

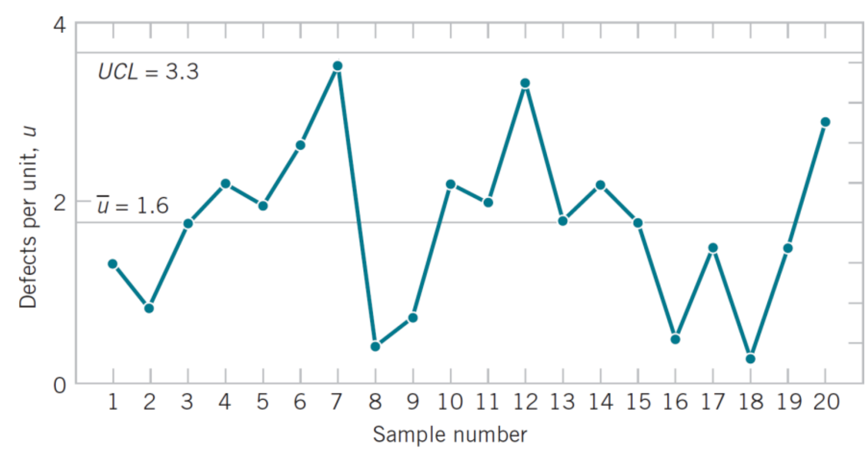
$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 1.6 + 3\sqrt{\frac{1.6}{5}} = 3.3$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 1.6 - 3\sqrt{\frac{1.6}{5}} < 0$$

The control chart is plotted in Figure. Because LCL is negative, it is set to 0. From the control chart in Figure, we see that the process is in control.

Number of Defects in Samples of Five Printed Circuit Boards

Sample	Number of Defects	Defects per Unit u_i	Sample	Number of Defects	Defects per Unit u_i
1	6	1.2	11	9	1.8
2	4	0.8	12	15	3.0
3	8	1.6	13	8	1.6
4	10	2.0	14	10	2.0
5	9	1.8	15	8	1.6
6	12	2.4	16	2	0.4
7	16	3.2	17	7	1.4
8	2	0.4	18	1	0.2
9	3	0.6	19	7	1.4
10	10	2.0	20	13	2.6



U chart of defects per unit on printed circuit boards.

Note

樣本很大 posssion 可以 normal 逼近關係
可以用 possion 來做稀少事件

- *U*-chart control limits are based on the normal approximation to the Poisson distribution.
- When λ is small, the normal approximation may not always be adequate.
  - In such cases, we may use control limits obtained directly from a table of Poisson probabilities.
- If $\bar{u}$ is small, the lower control limit obtained from the normal approximation may be a negative number. If this should occur, it is customary to use zero as the lower control limit.

Poisson Approximation to Binomial

- Binomial distribution with parameters (n, p)

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}$$

跟 n 很大 p 很小，binomial 就會逼近 poisson

- As $n\to\infty$ and $p\to 0$, with $np=\lambda$ moderate, binomial distribution converges to Poisson with parameter λ.

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}$$

$$=\frac{(n-k+1)\ldots(n-1)n}{k!}\cdot\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$\frac{(n-k+1)\ldots(n-1)n}{n^k}\xrightarrow[n\to\infty]{}1$$

$$\left(1-\frac{\lambda}{n}\right)^n\xrightarrow[n\to\infty]{}e^{-\lambda}$$

$$\left(1-\frac{\lambda}{n}\right)^k\xrightarrow[n\to\infty]{}1$$

$$P\{X=k\}\xrightarrow[n\to\infty]{}e^{-\lambda}\frac{\lambda^k}{k!}$$

Control Chart Performance

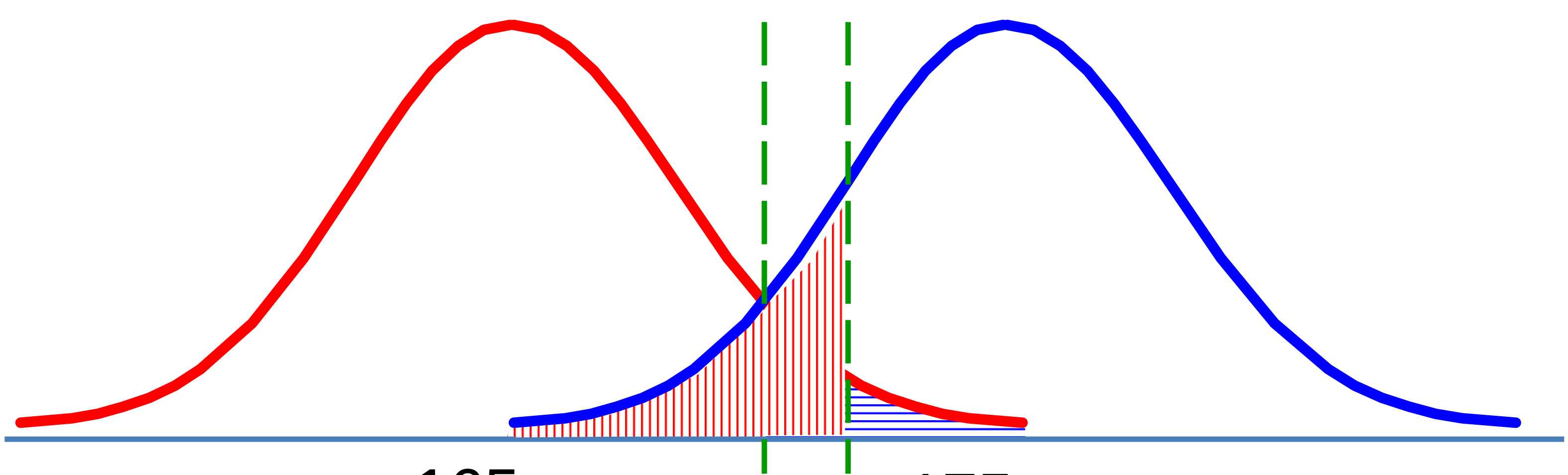
P chart 是比例 binomial；U chart 是稀少事件 poisson；但都是計次 countable

Background

品管圖的好壞指標是看 type 1 type 2 tradeoff

- Moving the control limits farther from center line → decrease type I error
  - the risk of a point falling beyond the control limits, indicating an out-of-control condition when no assignable cause is present. 真實是良，被判斷為不良
- However, widening the control limits also increases the type II error
  - the risk of a point falling between the control limits when the process is really out of control. 實際是不良，被判斷為良
- If we move the control limits closer to the center line, the opposite effect is obtained: type I error increases, and type II error decreases.

擔心誤判女生就把門檻拉高，但誤判男生的機率就變大



Control Chart Performance

Average Run Length (ARL)

- The average run length (ARL) is a very important way of determining the appropriate sample size and sampling frequency.
- ARL is the avg. # of points plotted to signal an out-of-control condition.
  - The ARL can be calculated from the mean of a geometric random variable.
- Let p = probability that any point exceeds the control limits. Then,

$$\text{ARL} = \frac{1}{p}$$

Example

應該涵蓋 99.7%，那落在外面就是 0.27%

- Thus, for an $\bar{X}$ chart with 3-sigma limits, p = 0.0027 is the probability that a normally distributed point falls outside the limits when the process is in control, so

$$\text{ARL} = \frac{1}{p} = \frac{1}{0.0027} \cong 370$$

產線上每生產 370 批，平均來說就會 alarm 一次
應該寄售 false alarm，就是其實是來自正常的分布但每 370 就會超出範圍

- is the average run length of the $\bar{X}$ chart when the process is in control. That is, even if the process remains in control, an out-of-control signal is generated every 370 points on average.

就可以思考 370 合理嗎，如果真實情況覺得 1000 個 lot 才會有一次不良品才對，那就代表品管圖用 +- 三個標準差太窄了~~可以用 3.5 之類的

Control Chart Performance

previous example

- Consider the piston ring process with sampling every hour
  - **false alarm** about every 370 hours on average.
  - the true process mean is $\mu = 74$ and the process standard deviation is $\sigma = 0.01$, with a sample size of $n = 5$
    - UCL = 74 + 3(0.0045) = 74.0135
    - CL= 74
    - LCL = 74 - 3(0.0045) = 73.9865

mean 飄到上管制界線，那超出 UCL 的機會 cdf 就是 0.5

UCL，如果變窄超出界線的 cdf 也會上升

LCL

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{5}} = 0.0045$$

  - when the process goes out of control, the mean shifts to 74.0135 millimeters.
  - Then, the probability that $\bar{X}$ falls between the control limits is equal to

$$P\,[73.9865 \le \bar{X} \le 74.0135 \text{ when } \mu = 74.0135]$$
$$= P\left[\frac{73.9865 - 74.0135}{0.0045} \le Z \le \frac{74.0135 - 74.0135}{0.0045}\right]$$
$$= P[-6 \le Z \le 0] = 0.5$$

  - Thus, p is 0.5, and the out-of-control ARL is $1/p = 1/0.5 = 2$. That is, the control chart requires 2 samples to detect the process shift, *on the average*, so 2 hours is expected to elapse between the shift and its detection.

Control Chart Performance

- Consider the piston ring process with sampling every hour
  - Suppose that this approach is unacceptable because production of piston rings with a mean diameter of 74.0135 millimeters results in excessive scrap costs and delays final engine assembly.
  - How can we reduce the time needed to detect the out-of-control condition?
    - Sample more frequently (Design 1) 方案一
      - For example, if we sample every half hour, only one hour elapses between the shift and its detection.
    - Increase the sample size (Design 2) 方案二 n 上升管制圖變窄
      - For example, if we use n = 10, the control limits narrow to 73.9905 and 74.0095. The probability of $\bar{X}$ falling between the control limits when the process mean is 74.0135 millimeters is approximately 0.1, so $p = 0.9$, and the out-of-control ARL is $1/p = 1/0.9 = 1.11$.
      - Thus, the larger sample size allow the shift to be detected about twice as quickly as the smaller one. If it became important to detect the shift in approximately the first hour after it occurred.

Design 1	**Design 2**
Sample size: $n = 5$	Sample size: $n = 10$
Sampling frequency: every half hour	Sampling frequency: every hour

Control Chart Performance

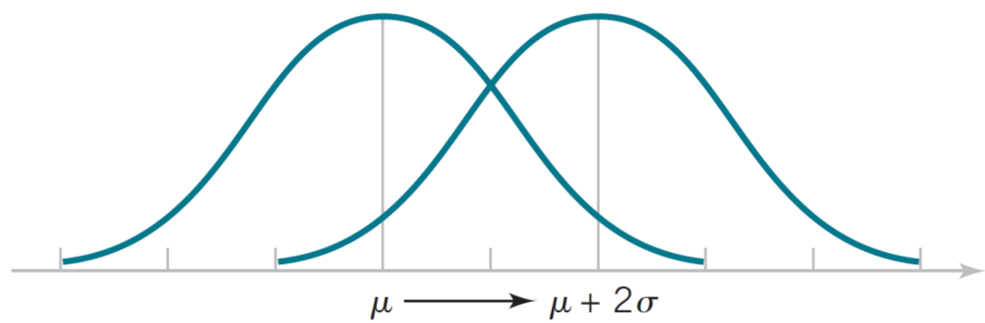
可以針對抽樣的頻率 跟抽樣的大小去調整

- Table provides ARL for an $\bar{X}$ chart with 3-sigma control limits.
  - ARL are calculated for shifts in the process mean from 0 to 3.0σ and for sample sizes of n = 1 and n = 4 by using 1/p.
- Figure illustrates a shift in the process mean of 2σ.

Average Run Length (ARL) for an $\bar{X}$ Chart with 3-Sigma Control Limits

樣本數量上升時，針對偏移的上升的敏感度變化更快

Magnitude of Process Shift	ARL $n = 1$	ARL $n = 4$
0	370.4	370.4
0.5σ	155.2	43.9
1.0σ	43.9	6.3
1.5σ	15.0	2.0
2.0σ	6.3	1.2
3.0σ	2.0	1.0



Process mean shift of 2σ.

Time-Weighted Charts

X bar chart 每個樣本被視為獨立；individual 的用 MR 才不是獨立

Motivation

- A major disadvantage of any Shewhart control chart
  因為是平均數，所以相對的平滑而且看不到偏態等資訊，而且樣本間不該是獨立的
  - relatively insensitive to small shifts in the process on the order of about 1.5σ or less, because it ignores the information in the sequence of points.
  - Solution: time-weighted chart, which integrates data over several time periods.

Exponential Weighted Moving Average (EWMA) Control Chart

Data collected in time order are often averaged over several time periods. For example, economic data are often presented as an average over the last four quarters. That is, at time *t*, the average of the last four measurements can be written as

$$\overline{x}_t = \frac{1}{4}x_t + \frac{1}{4}x_{t-1} + \frac{1}{4}x_{t-2} + \frac{1}{4}x_{t-3}$$

假設時窗為 4
時窗越大越平滑，noise smooth out，不敏感

This average places weight of 1/4 on each of the most recent observations and zero weight on older observations. It is called a **moving average** and in this case, a *window* of size 4 is used. An average of the recent data is used to smooth the noise in the data to generate a better estimate of the process mean than only the most recent observation.

Exponential Weighted Moving Average (EWMA) Control Chart

For statistical process control, rather than use a fixed window size, it is useful to place the most weight on the most recent observation or subgroup average and then gradually decrease the weights on older observations. An average of this type can be constructed by a multiplicative decrease in the weights. Let $0 < \lambda \leq 1$ denote a constant and μ_0 denote the process target or historical mean. Suppose that samples of size $n \geq 1$ are collected and $\bar{x}_t$ is the average of the sample at time t. The **exponentially weighted moving-average (EWMA)** is

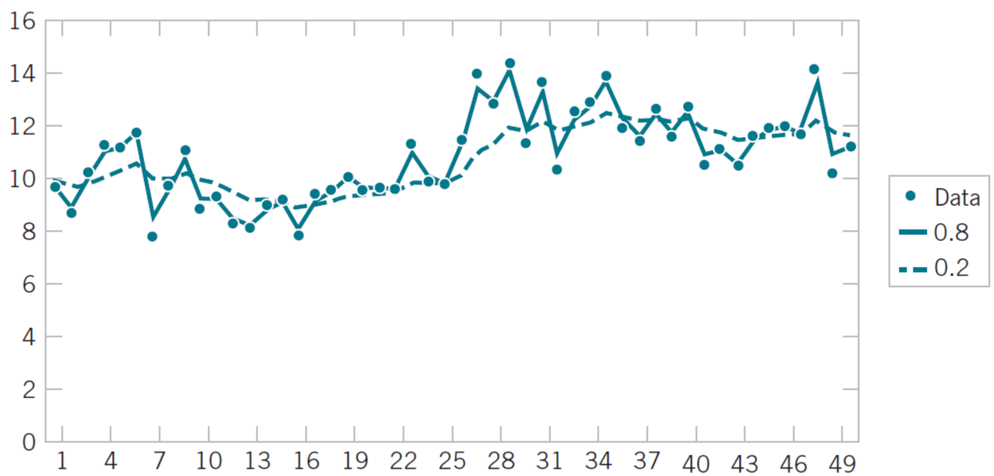
lambda 就是權重，樣本點越古老權重就越小，也就離的越遠影響力應該越小

$$z_t = \lambda\bar{x}_t + \lambda(1-\lambda)\bar{x}_{t-1} + \lambda(1-\lambda)^2\,\bar{x}_{t-2} + \cdots + \lambda(1-\lambda)^{t-1}\,\bar{x}_1 + (1-\lambda)^t\mu_0$$

$$= \sum_{k=0}^{t-1}\lambda(1-\lambda)^k\,\bar{x}_{t-k} + (1-\lambda)^t\mu_0$$

lambda 是我放給我當下現在的資料的權重有多大，所以 Lambda 越大，參考比較多現在，就會比較震盪；lambda 小就會平滑

Each older observation has its weight decreased by the factor $(1-\lambda)$. The weight on the starting value μ_0 is selected so that the weights sum to 1. An EWMA is also sometimes called a **geometric average**.



Let $z_0 = \mu_0$.
EWMA update equation.

$$z_t = \lambda\bar{x}_t + (1-\lambda)z_{t-1}$$

EWMAs with λ=0.8 and λ=0.2 show a compromise between a smooth curve and a response to a shift

E[cX] = cE[X]
Var[cX] = c^2 Var[X]

Mean of a Linear Function

If $Y = c_0 + c_1X_1 + c_2X_2 + \cdots + c_pX_p$,

$$E(Y) = c_0 + c_1E(X_1) + c_2E(X_2) + \cdots + c_pE(X_p)$$

Variance of a Linear Function

If $X_1, X_2, \ldots, X_p$ are random variables, and $Y = c_0 + c_1X_1 + c_2X_2 + \cdots + c_pX_p$,

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p) + 2\sum_{i<j}\sum c_ic_j \operatorname{cov}(X_i, X_j)$$

If $X_1, X_2, \ldots, X_p$ are *independent*,

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p)$$

EWMA Control Chart

$$LCL = \mu_0 - 3\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2t}\right]}$$

$$CL = \mu_0$$

$$UCL = \mu_0 + 3\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2t}\right]}$$

Exponential Weighted Moving Average (EWMA) Control Chart

Note that the control limits are not of equal width about the center line. The control limits are calculated from the variance of Z_t and that changes with time. However, for large t, the variance of Z_t converges to

$$\lim_{t\to\infty} V(Z_t) = \frac{\sigma^2}{n}\left(\frac{\lambda}{2-\lambda}\right)$$

so that the control limits tend to be parallel lines about the center line as t increases.

The parameters μ_0 and σ are estimated by the same statistics used in $\overline{X}$ or X charts. That is, for subgroups

$$\hat{\mu}_0 = \overline{\overline{X}} \quad \text{and} \quad \hat{\sigma} = \overline{R}/d_2 \quad \text{or} \quad \hat{\sigma} = \overline{S}/c_4$$

and for $n = 1$

$$\hat{\mu}_0 = \overline{X} \quad \text{and} \quad \hat{\sigma} = \overline{MR}/1.128$$

Example: Chemical Process Concentration EWMA

Consider the concentration data shown in Table. Construct an EWMA control chart with $\lambda = 0.2$ with $n = 1$. It was determined that $\bar{x} = 99.1$ and $\overline{mr} = 2.59$. Therefore, $\hat{\mu}_0 = 99.1$ and $\hat{\sigma} = 2.59/1.128 = 2.30$. The control limits for z_1 are

$$LCL = 99.1 - 3(2.30)\sqrt{\frac{0.2}{2-0.2}\left[1-(1-0.2)^2\right]} = 97.72$$

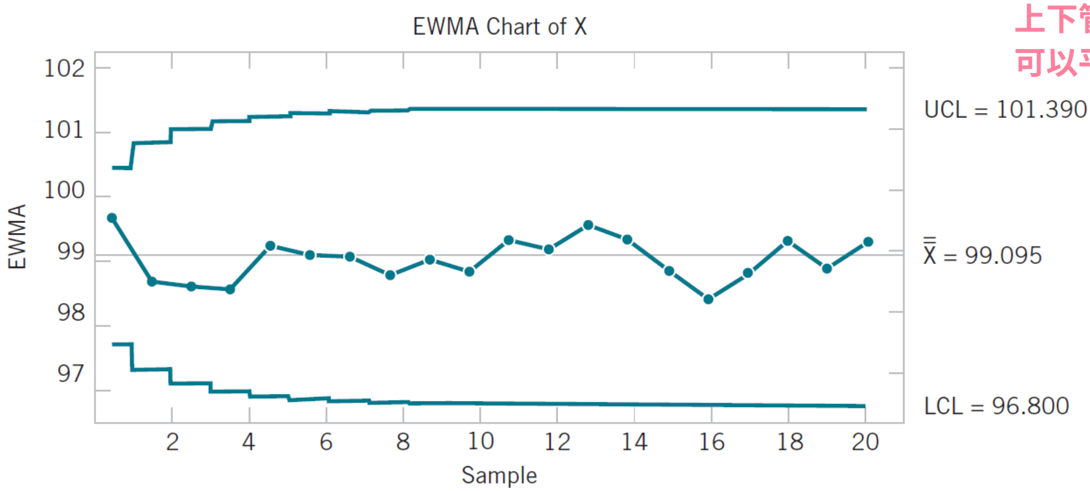
$$UCL = 99.1 + 3(2.30)\sqrt{\frac{0.2}{2-0.2}\left[1-(1-0.2)^2\right]} = 100.48$$

The first few values of z_t along with the corresponding control limits are

t	1	2	3	4	5
x_t	102	94.8	98.3	98.4	102
z_t	99.68	98.7	98.62	98.58	99.26
LCL	97.72	97.33	97.12	97	96.93
UCL	100.48	100.87	101.08	101.2	101.27

The chart generated by computer software is shown in Figure. Notice that the control limits widen as time increases but quickly stabilize. Each point is within its set of corresponding control limits so there are no signals from the chart.

上下管制界線一開始不是水平的，是因為指數平滑在一開始沒有太多過去可以平滑



EWMA control chart for the chemical process concentration data from computer software.

The points plotted on an EWMA control chart are not independent. Therefore, run rules should not be applied to an EWMA control chart. Information in the history of the data that is considered by run rules is to a large extent incorporated into the EWMA that is calculated at each time *t*.

Cumulative Sum Control Chart

- The cusum chart incorporates all information in the sequence of sample values by plotting the cumulative sums of the deviations of the sample values from a target value.
- Suppose that samples of size *n* ≥ 1 are collected. If μ_0 is the target for the process mean, $\bar{X}_j$ is the average of the jth sample, then the cumulative sum control chart is formed by plotting the quantity

$$S_i = \sum_{j=1}^{i} (\bar{X}_j - \mu_0)$$

- Now S_i is called the cumulative sum up to and including the ith sample.

Example: Chemical process concentration data

- Because the concentration readings are individual measurements, we would take $\bar{X}_j = X_j$ in computing the CUSUM. Suppose that the target value for the concentration is $\mu_0 = 99$. Then the CUSUM is

$$S_i = \sum_{j=1}^{i} (X_j - 99) = (X_i - 99) + \sum_{j=1}^{i-1} (X_j - 99)$$

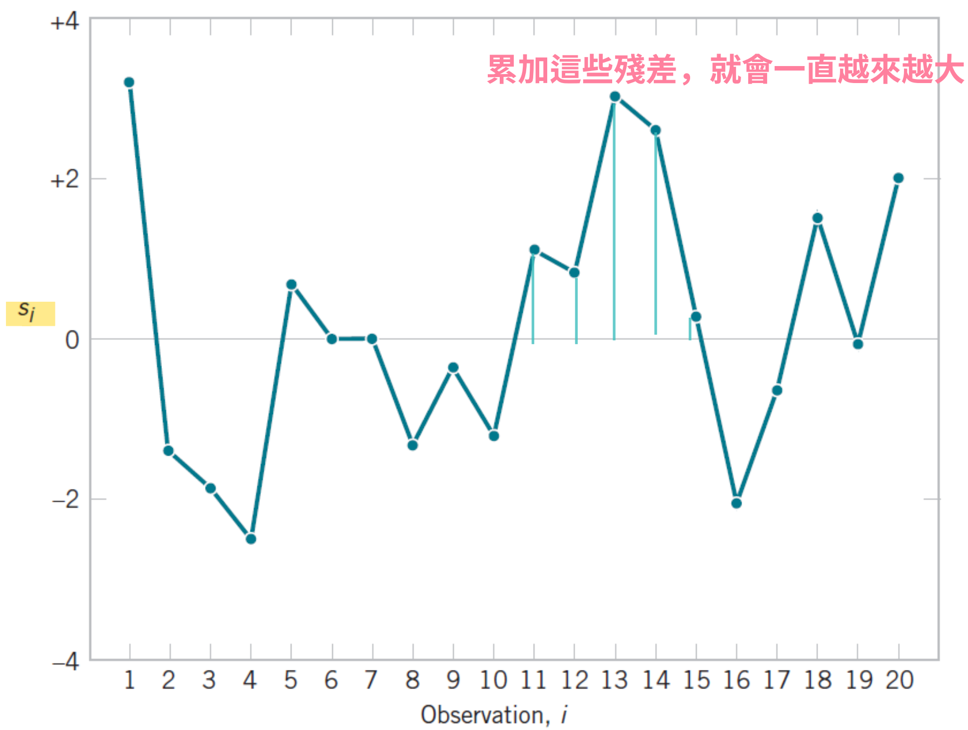
$$= (X_i - 99) + S_{i-1}$$

POLab

CUSUM for the Chemical Process Concentration Data

- Plot of the cumulative sums for the concentration data.

正負合在一起算



The graph is not a control chart because it lacks control limits.

Observation, i	x_i	$x_i - 99$	$s_i = (x_i - 99) + s_{i-1}$
1	102.0	3.0	3.0
2	94.8	–4.2	–1.2
3	98.3	–0.7	–1.9
4	98.4	–0.6	–2.5
5	102.0	3.0	0.5
6	98.5	–0.5	0.0
7	99.0	0.0	0.0
8	97.7	–1.3	–1.3
9	100.0	1.0	–0.3
10	98.1	–0.9	–1.2
11	101.3	2.3	1.1
12	98.7	–0.3	0.8
13	101.1	2.1	2.9
14	98.4	–0.6	2.3
15	97.0	–2.0	0.3
16	96.7	–2.3	–2.0
17	100.3	1.3	–0.7
18	101.4	2.4	1.7
19	97.2	–1.8	–0.1
20	101.0	2.0	1.9

K: treated as standard deviation
H: treated as $\pm$3sigma

CUSUM Control Chart

- Let $S_H(i)$ be an upper one-sided CUSUM for period i and $S_L(i)$ be a lower one-sided CUSUM for period i.

CUSUM Control Chart 個別一邊自己算，正的自己加總，負的自己加總

$$s_H(i) = \max[0, \bar{x}_i - (\mu_0 + K) + s_H(i-1)]$$

and

$$s_L(i) = \max[0, (\mu_0 - K) - \bar{x}_i + s_L(i-1)]$$

where the starting values $s_H(0) = s_L(0) = 0$.

- K is called the **reference value**, which is usually chosen about halfway between the target μ_0 and the value of the mean corresponding to the out-of-control state, $\mu_1 = \mu_0 + \Delta$. That is, K is about one-half the magnitude of the shift we are interested in, or $K = \Delta/2$.

類似標準差的另一種距離單位，是管制界線的範圍除以二，就是一個方向的界線範圍

- Notice that $S_H(i)$ and $S_L(i)$ accumulate deviations from the target value that are greater than K, with both quantities reset to zero upon becoming negative. If either $S_H(i)$ and $S_L(i)$ exceeds a constant H, the process is out of control. This constant H is usually called the **decision interval**.

Example: Chemical Process Concentration Tabular CUSUM

We illustrate the tabular CUSUM with the chemical process concentration data in Table. The process target is $\mu_0 = 99$, and we use $K = 1$ as the reference value and $H = 10$ as the decision interval. The reasons for these choices are explained later.

Table shows the tabular CUSUM scheme for the chemical process concentration data. To illustrate the calculations, note that

$$\begin{aligned} S_H(i) &= \max[0, x_i - (\mu_0 + K) + S_H(i-1)] \\ &= \max[0, x_i - (99 + 1) + S_H(i-1)] \\ &= \max[0, x_i - (100) + S_H(i-1)] \end{aligned}$$

$$\begin{aligned} S_L(i) &= \max[0, (\mu_0 - K) - x_i + S_L(i-1)] \\ &= \max[0, (99 - 1) - x_i + S_L(i-1)] \\ &= \max[0, 98 - x_i + S_L(i-1)] \end{aligned}$$

We assumed that the process standard deviation σ = 2. Then with k = 1/2 and h = 5, we would use K = $k\sigma$ = 1 and H = $h\sigma$ = 10 in the tabular CUSUM procedure.

Time-Weighted Charts

- Tabular CUSUM for the Chemical Process Concentration Data

連續出現正的次數

Observation i	x_i	Upper CUSUM $x_i - 100$	$s_H(i)$	n_H	Lower CUSUM $98 - x_i$	$s_L(i)$	n_L
1	102.0	2.0	2.0	1	−4.0	0.0	0
2	94.8	−5.2	0.0	0	3.2	3.2	1
3	98.3	−1.7	0.0	0	−0.3	2.9	2
4	98.4	−1.6	0.0	0	−0.4	2.5	3
5	102.0	2.0	2.0	1	−4.0	0.0	0
6	98.5	−1.5	0.5	2	−0.5	0.0	0
7	99.0	−1.0	0.0	0	−1.0	0.0	0
8	97.7	−2.3	0.0	0	0.3	0.3	1
9	100.0	0.0	0.0	0	−2.0	0.0	0
10	98.1	−1.9	0.0	0	−0.1	0.0	0
11	101.3	1.3	1.3	1	−3.3	0.0	0
12	98.7	−1.3	0.0	0	−0.7	0.0	0
13	101.1	1.1	1.1	1	−3.1	0.0	0
14	98.4	−1.6	0.0	0	−0.4	0.0	0
15	97.0	−3.0	0.0	0	1.0	1.0	1
16	96.7	−3.3	0.0	0	1.3	2.3	2
17	100.3	0.3	0.3	1	−2.3	0.0	0
18	101.4	1.4	1.7	2	−3.4	0.0	0
19	97.2	−2.8	0.0	0	0.8	0.8	1
20	101.0	1.0	1.0	1	−3.0	0.0	0

累計後是負的就算 0

從 0 重新

還維持正的就記

累計同意方向的偏移越多就越異常

- Example: Chemical Process Concentration Tabular CUSUM

Therefore, for observation 1, the CUSUMs are

$$s_H(1) = \max[0, x_1 - 100 + s_H(0)]$$
$$= \max[0, 102.0 - 100 + 0] = 2.0$$

and

$$s_L(1) = \max[0,98 - x_1 + s_L(0)]$$
$$= \max[0,98 - 102.0 + 0] = 0$$

as shown in Table. The quantities n_H and n_L in Table indicate the number of periods in which the CUSUM $s_H(i)$ or $s_L(i)$ have been nonzero. Notice that the CUSUMs in this example never exceed the decision interval $H = 10$. We would therefore conclude that the process is in control.

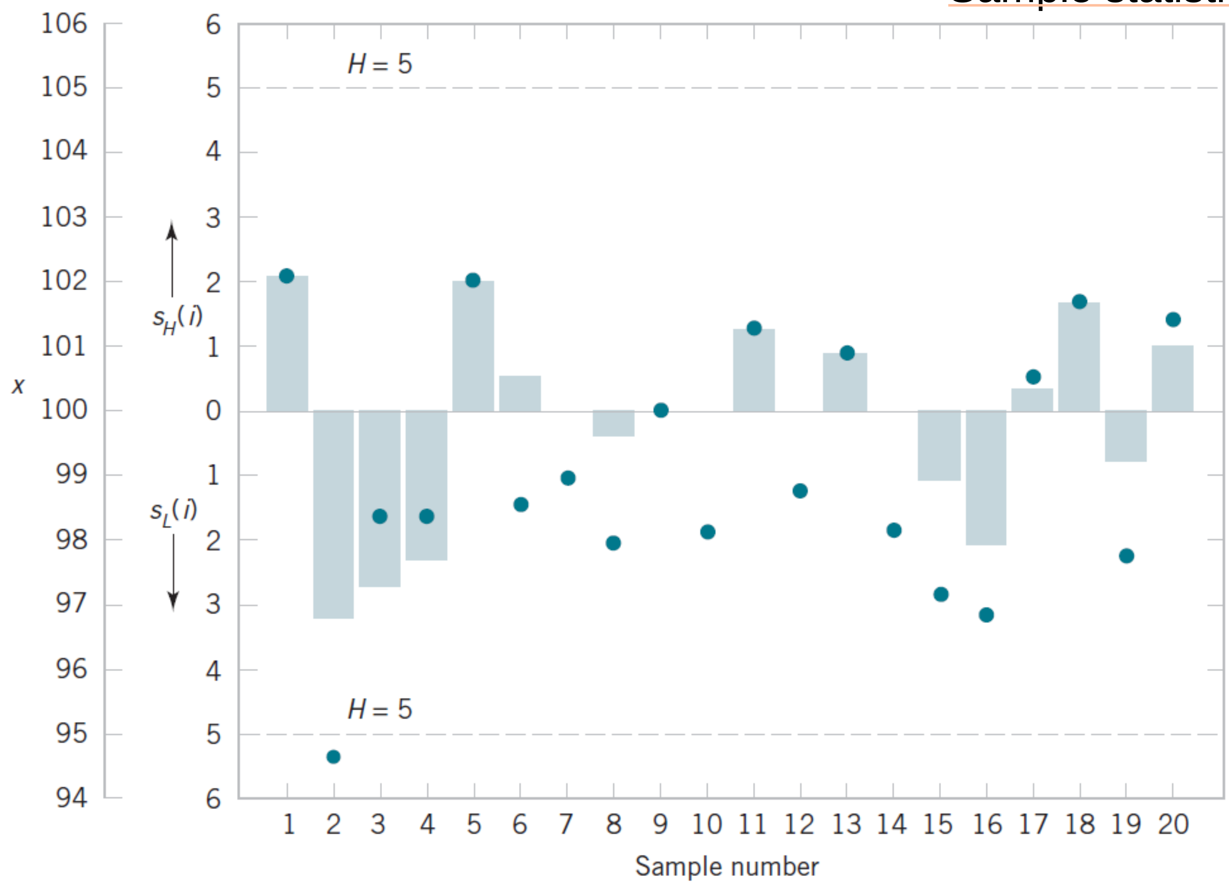
The CUSUM status chart

- When the tabular CUSUM indicates that the process is out of control, we should search for the assignable cause, take any corrective actions indicated, and restart the CUSUMs at zero.
- It may be helpful to have an estimate of the new process mean following the shift. This can be computed from

$$\hat{\mu} = \begin{cases} \mu_0 + K + \dfrac{s_H(i)}{n_H}, & \text{if } s_H(i) > H \\ \mu_0 - K - \dfrac{s_L(i)}{n_L}, & \text{if } s_L(i) > H \end{cases}$$

The CUSUM status chart

$S_H(i)$ and $S_L(i)$: vertical bar
Sample statistics x: solid dot



- Note: the Western Electric rules cannot be safely applied to the CUSUM because successive values of $S_H(i)$ and $S_L(i)$ are not independent. In fact, the CUSUM can be thought of as a weighted average, where the weights are stochastic or random. In effect, all CUSUM values are highly correlated, thereby causing the Western Electric rules to produce too many false alarms.

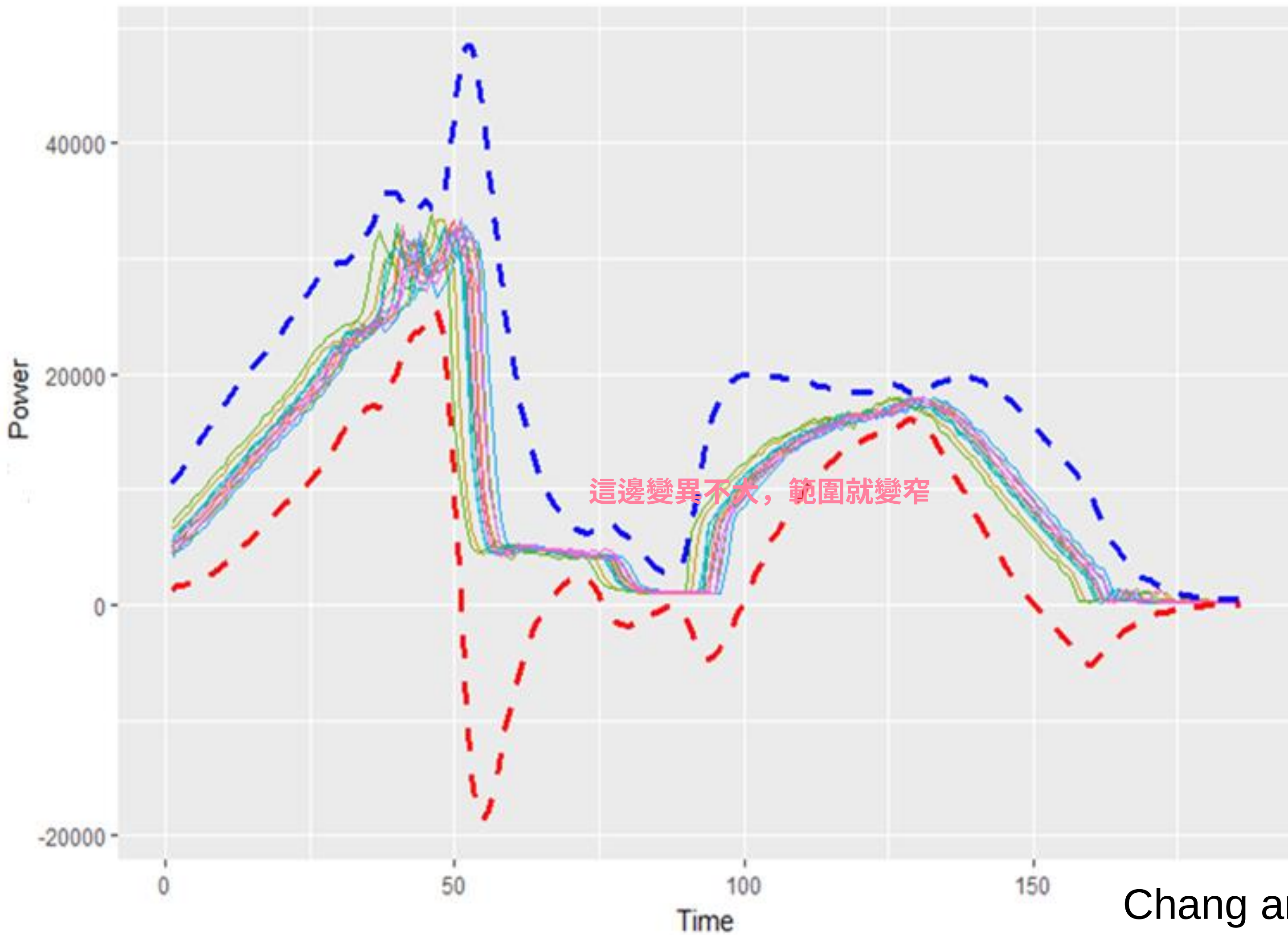
Average Run Lengths for a CUSUM Control Chart

- The tabular CUSUM is designed by choosing values for the reference value K and the decision interval H.
- Define $H = h\sigma_{\bar{X}}$ and $K = k\sigma_{\bar{X}}$ where $\sigma_{\bar{X}}$ is the standard deviation of the sample variable used in forming the CUSUM (if n = 1, $\sigma_{\bar{X}} = \sigma_X$). Using h = 4 or h = 5 and k = 1/2 generally provide a CUSUM that has good ARL properties against a shift of about $1\sigma_{\bar{X}}$ (or $1\sigma_X$) in the process mean. If much larger or smaller shifts are of interest, set k = δ/2 where δ is the size of the shift in standard deviation units.
- To illustrate how well the recommendations of h = 4 or h = 5 with k = 1/2 work, consider these ARLs for a CUSUM in Table.
  - Notice that a shift of $1\sigma_{\bar{X}}$ would be detected in either 8.38 samples (with k = 1/2 and h = 4) or 10.4 samples (with k = 1/2 and h = 5).
  - By comparison, an $\bar{X}$ chart would require approximately 43.9 samples, on the average, to detect this shift.

Shift in Mean (multiple of $\sigma_{\bar{X}}$)	Average Run Lengths $h = 4$	$h = 5$
0	168	465
0.25	74.2	139
0.50	26.6	38
0.75	13.3	17
1.00	8.38	10.4
1.50	4.75	5.75
2.00	3.34	4.01
2.50	2.62	3.11
3.00	2.19	2.57
4.00	1.71	2.01

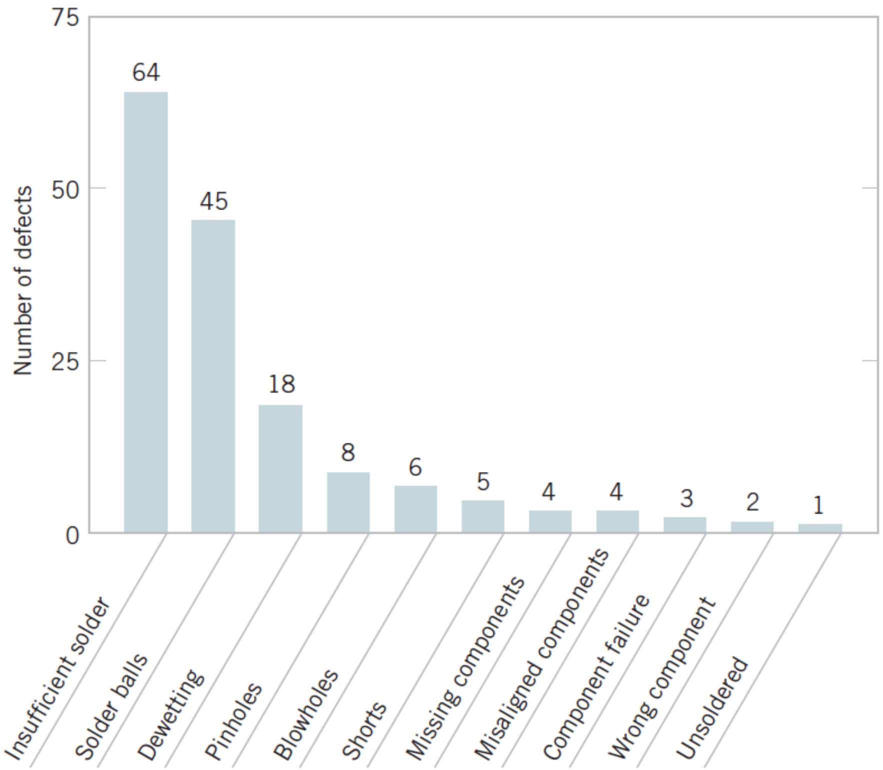
Time-Weighted Charts

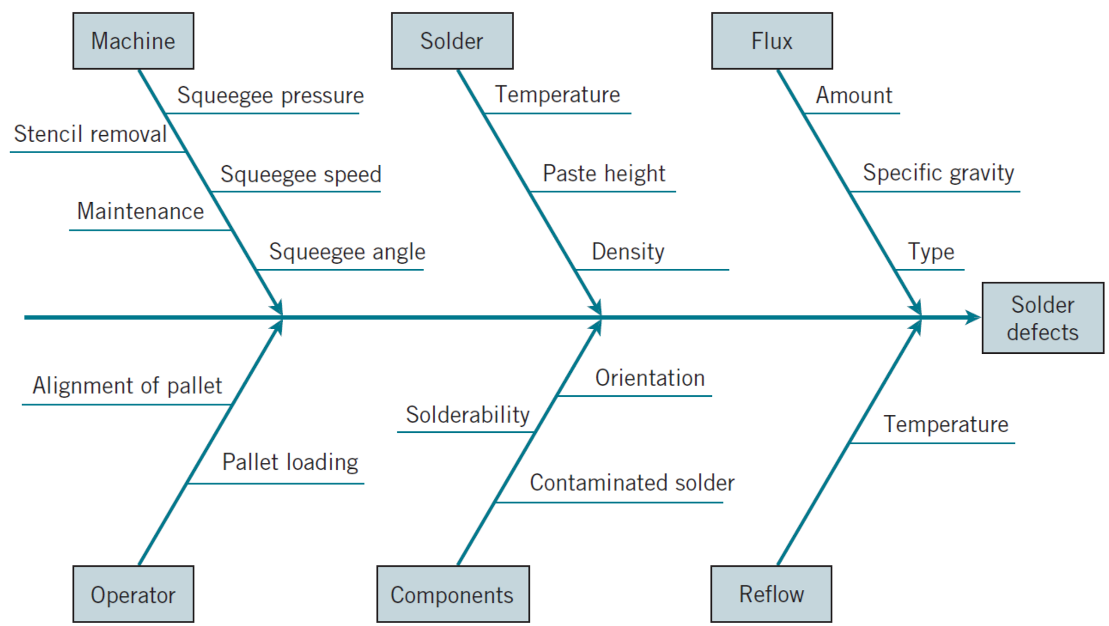
- MCEWMA capture 正常的 pattern 例如一開始加溫後來降溫，這樣的製程在參數的曲線在哪樣的範圍是異常
  - moving centerline exponentially weighted moving average



Chang and Lee (2018)

- Pareto diagram for printed circuit board defects.
- Cause-and-effect diagram for the printed circuit board flow solder process.





Deming's 14 points PDCA 循環

- 1. Create a constancy of purpose focused on the improvement of products and services
- 2. Adopt a new philosophy of rejecting poor workmanship, defective products, or bad service.
- 3. Do not rely on mass inspection to "control quality."
- 4. Do not award business to suppliers on the basis of price alone, but also consider quality.
- 5. Focus on continuous improvement.
- 6. Practice modern training methods and invest in training for all employees.
- 7. Practice modern supervision methods.

Deming's 14 points

- 8. Drive out fear.
- 9. Break down the barriers between functional areas of the business.
- 10. Eliminate targets, slogans, and numerical goals for the workforce.
- 11. Eliminate numerical quotas and work standards.
- 12. Remove the barriers that discourage employees from doing their jobs.
- 13. Institute an ongoing program of training and education for all employees.
- 14. Create a structure in top management that vigorously advocates the first 13 points.

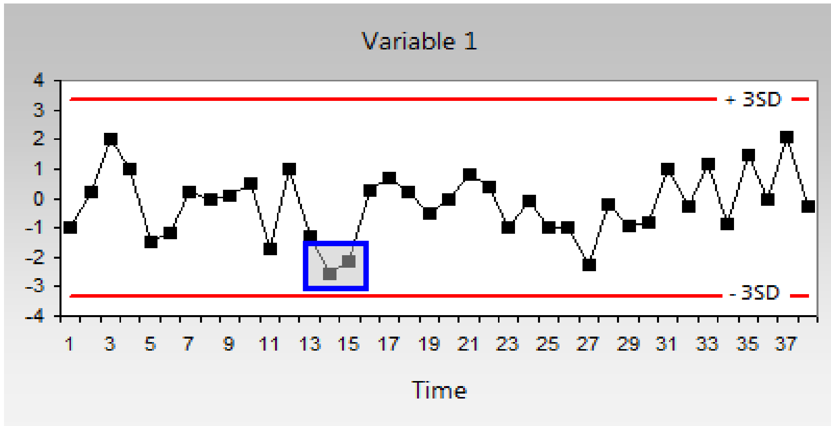
Multivariate Control Chart

Health Index Construction

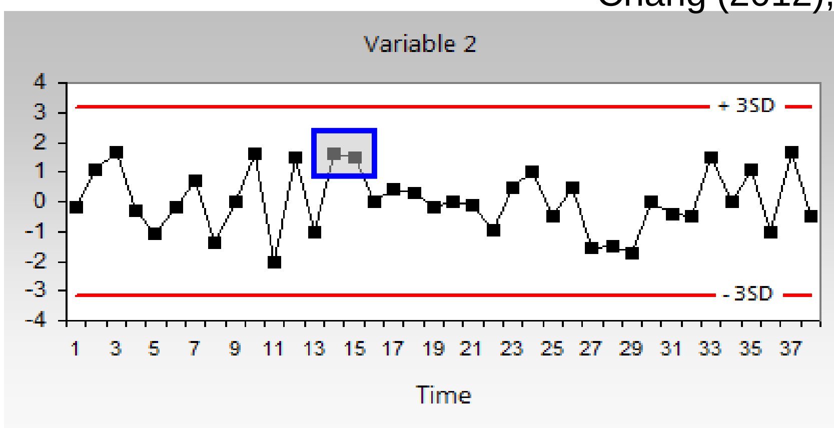
個別看每個參數的時候都是在範圍內的，但可能需要多變因的交互關係
多變因之後就有可能反而超出範圍

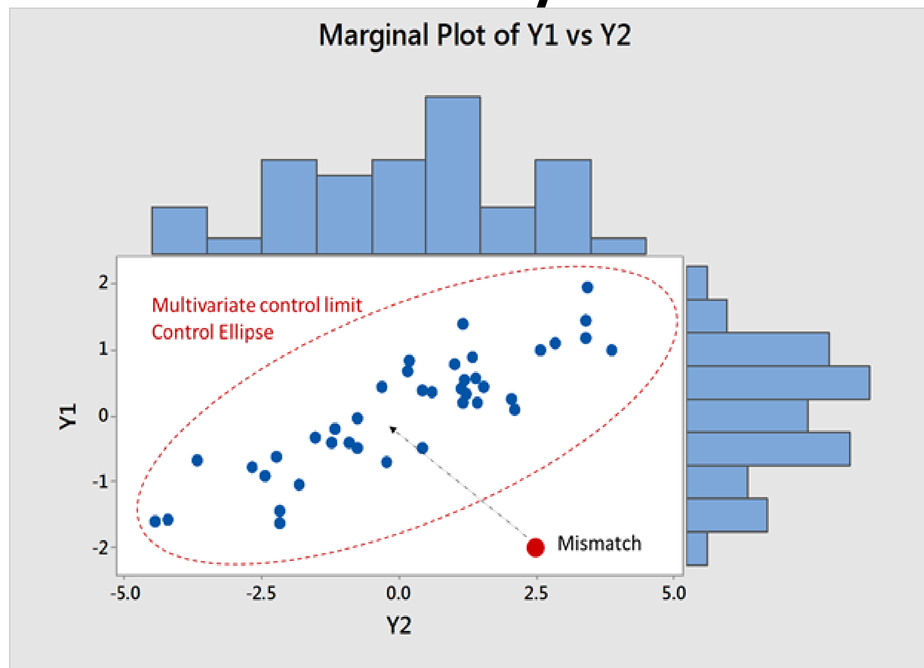
- Multivariate Control Chart (statistical-based)

Multivariate Control Limits



Chang (2012), IBM





The information is found in the correlation pattern - not in the individual variables!

Scibilia, B., 2016. A simple guide to multivariate control charts. https://blog.minitab.com/blog/applying-statistics-in-quality-projects/a-simple-guide-to-multivariate-control-charts

Multivariate Control Chart

把所有的指標融合為一個健康指標

馬氏距離 Mahalanobis Distance 就是每個維度的權重可以是不一樣的，考慮不同因子之間的相關性
權重，sqrt((x - mu)^-1 S^-1 (x - mu))
就是用共變異矩陣當權重，給一個客觀的權重

Health Index Construction

- Multivariate Control Chart (statistical-based)
- Hotelling's T-squared distribution (T^2)

$$t^2 = (\overline{\mathbf{x}} - \boldsymbol{\mu})' \hat{\boldsymbol{\Sigma}}_{\overline{\mathbf{x}}}^{-1} (\overline{\mathbf{x}} - \boldsymbol{\mu})$$

就是用馬氏距離
這樣的數值會逼近 T-square 的分布
所以就可以查表檢定

Sample covariance

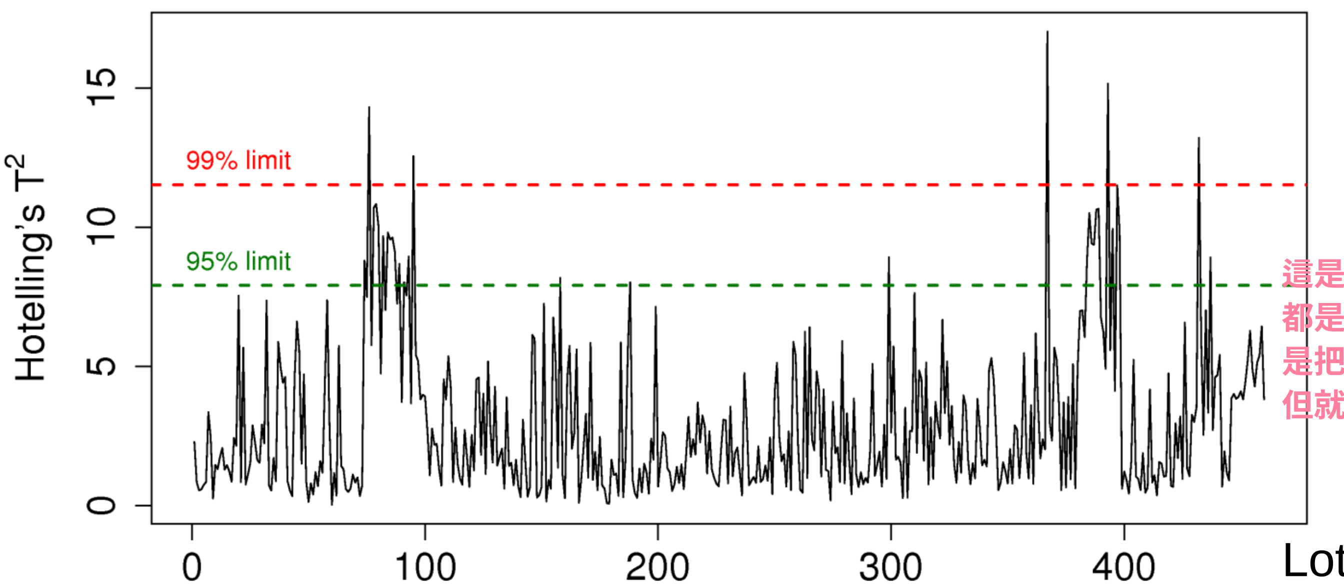
$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n-1} \sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})'$$

Also, from the distribution,

$$t^2 \sim T^2_{p,n-1} = \frac{p(n-1)}{n-p} F_{p,n-p},$$

The sample covariance matrix of the mean $\hat{\boldsymbol{\Sigma}}_{\overline{\mathbf{x}}} = \hat{\boldsymbol{\Sigma}}/n.$

where $F_{p,n-p}$ is the F-distribution with parameters p and $n - p$.



多變量的管制圖就要納入變數間的相關性，所以就用共變異矩陣計算

這是一張多變輛管制圖，因為是計算距離，所以都是正的
是把所有指標都合在一起為一個指標
但就會不知道出問題的時候是哪一個指標出問題

https://learnche.org/pid/latent-variable-modelling/principal-component-analysis/hotellings-t2-statistic

Video Time

Coca-Cola

DATALYZER International

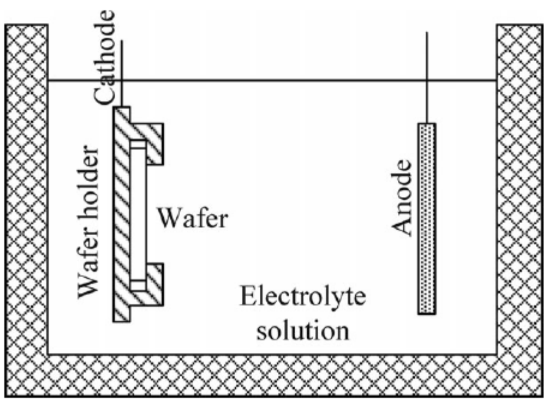
Partners in Continuous Improvement

SPC Software in Coca Cola. https://www.youtube.com/watch?v=hFoS2a7p5AE

Chamber-to-chamber matching

- Step 1: collect 1st dataset including product quality and process variable
- Step 2: calculate standard deviation (STD) of the quality, draw box plot of each chamber, and identify the smallest STD as golden chamber

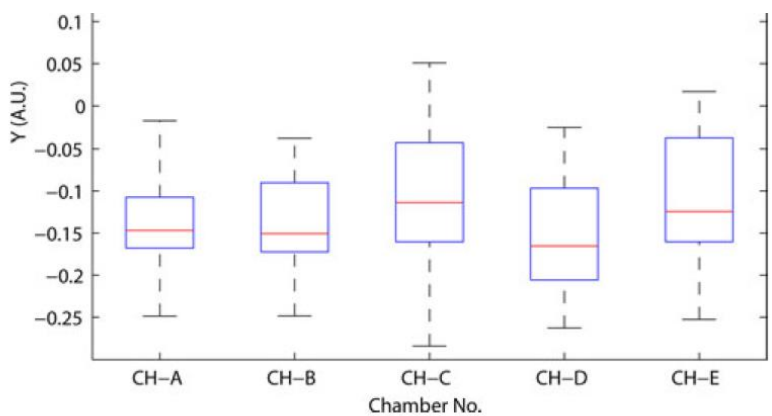
Electroplating process



Chamber name	#wafers	s_n (Y)	p-value
CH-A	18	0.0544	–
CH-B	11	0.0623	0.2997
CH-C	**18**	**0.0908**	**0.0207**
CH-D	16	0.0728	0.1244
CH-E	19	0.0807	0.0555

golden

CH-A is a golden chamber

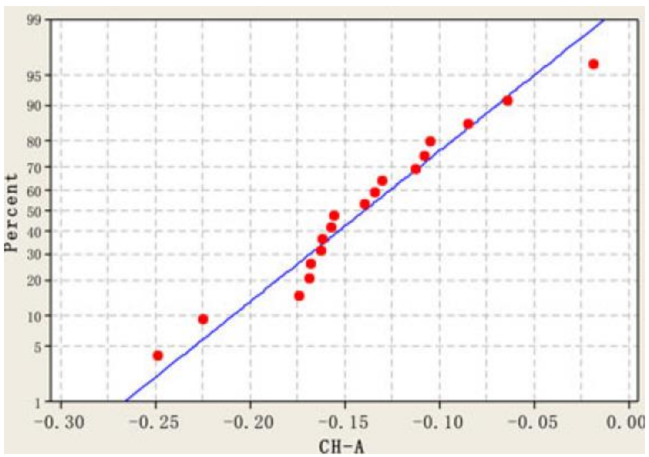


(a.u.: arbitrary unit)

Pan, T.-H., Wong, D. S., and Jang, S.-S. 2012. Chamber matching of semiconductor manufacturing process using statistical analysis. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 42, NO. 4, 571-576.

Chamber-to-chamber matching

- Step 3: normality test for all chambers and perform the F-test for product quality variations and identify the inferior chamber.
- Step 4: use F-test to determine the standard deviations of which process variables in the inferior chamber are larger than those in the golden chamber.



$$H_0 : s_n^2 = s_A^2$$

$$H_\alpha : s_n^2 > s_A^2$$

$$F = \frac{s_n^2}{s_A^2}$$

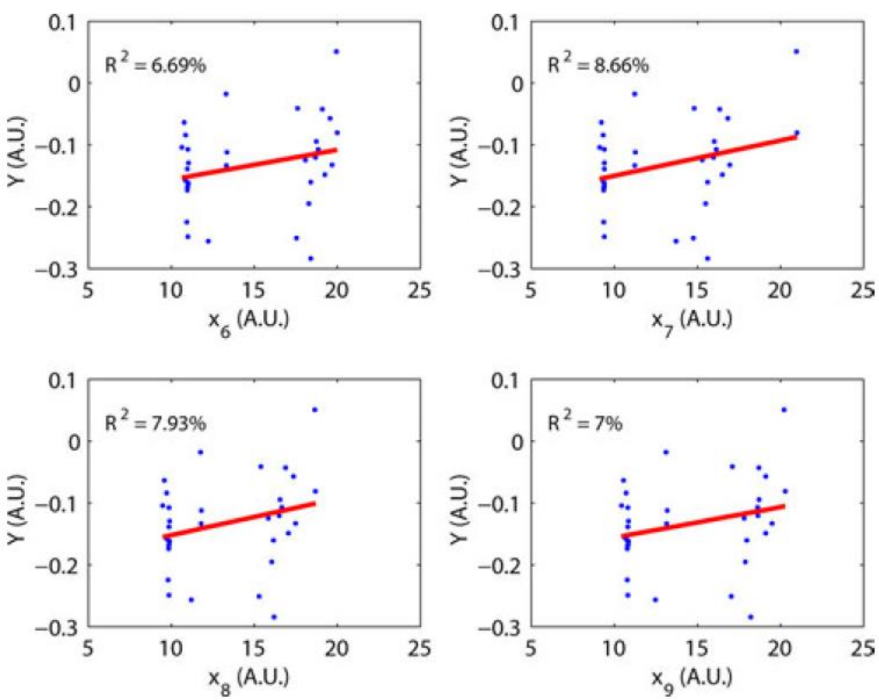
variable name	s_{x_i} (CH-A)	s_{x_i} (CH-C)	p-value
x_1	4.6E-16	2.3E-16	1
x_2	1.2E-4	4.1E-5	1
x_3	**5.1E-5**	**1.1E-4**	**0.0023**
x_4	**6.7E-6**	**1.5E-5**	**0.0037**
x_5	6.7E-5	9.4E-5	0.0827
x_6	**0.109**	**0.777**	**1.55E-9**
x_7	**0.031**	**0.881**	**1.44E-15**
x_8	**0.087**	**1.018**	**9.96E-12**
x_9	**0.085**	**0.970**	**1.32E-11**
x_{10}	3.053	0.6436	1

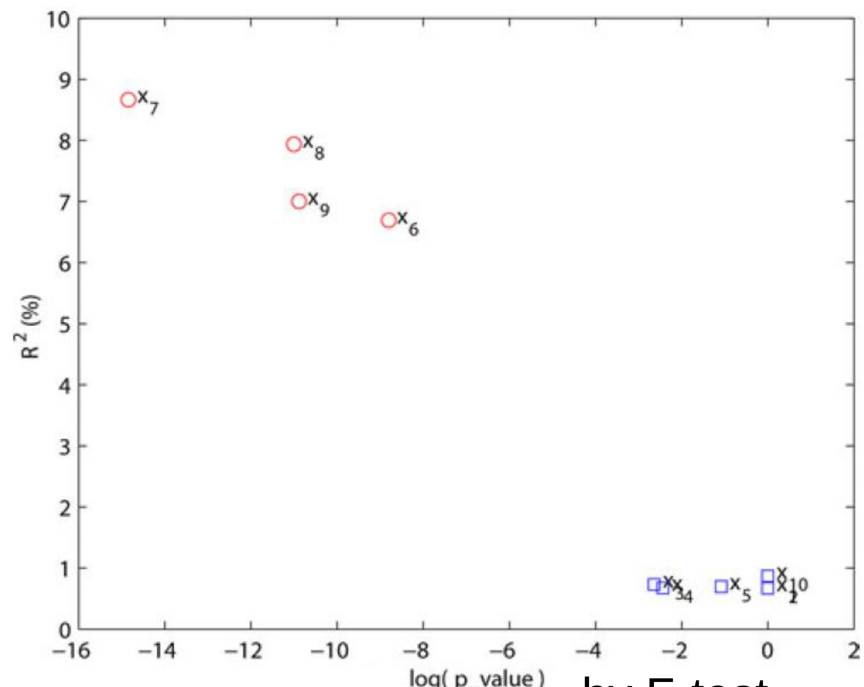
CH-C is an inferior chamber with $p < 0.05$

Pan, T.-H., Wong, D. S., and Jang, S.-S. 2012. Chamber matching of semiconductor manufacturing process using statistical analysis. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 42, NO. 4, 571-576.

Chamber-to-chamber matching

- Step 5: Perform a linear analysis to determine correlation between means of the process variables and product quality.
- Step 6: Select key variables that have larger standard deviations and sufficient correlation with the product quality.





by F-test

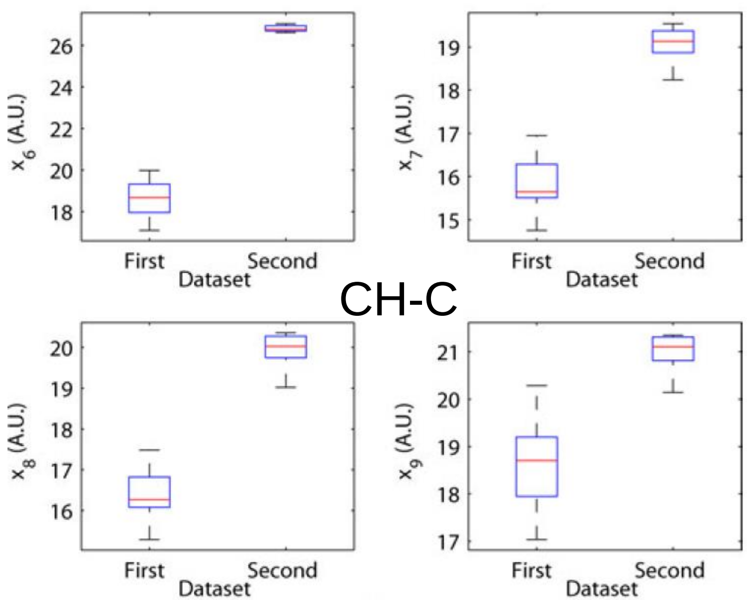
Pan, T.-H., Wong, D. S., and Jang, S.-S. 2012. Chamber matching of semiconductor manufacturing process using statistical analysis. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 42, NO. 4, 571-576.

Chamber-to-chamber matching

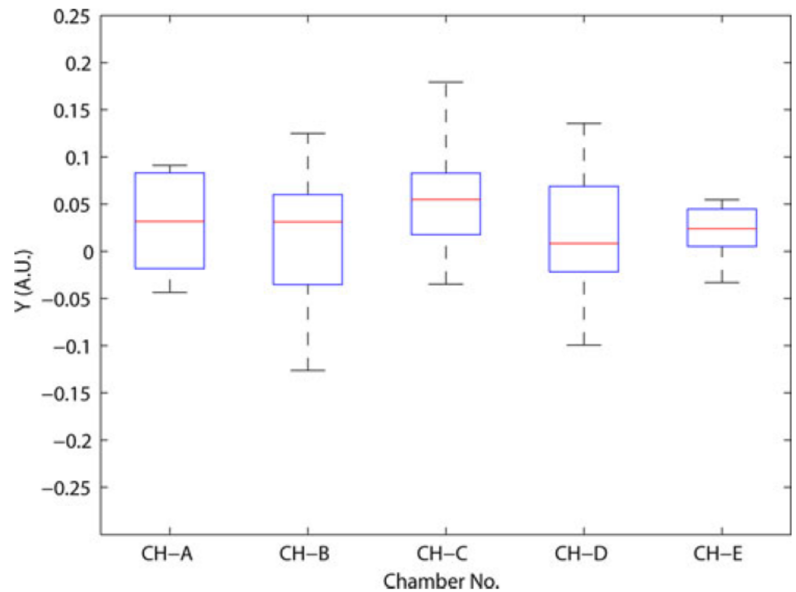
- Step 7: collect the 2nd dataset for validation after variable adjustment.

Chamber name	#wafers	s_n (Y)	p-value
CH-A	18	0.0544	–
CH-B	11	0.0623	0.2997
CH-C (As-Is)	**18**	**0.0908**	**0.0207**
CH-D	16	0.0728	0.1244
CH-E	19	0.0807	0.0555

Chamber name	#wafers	s_n (Y)	p-value
CH-A	8	0.0553	–
CH-B	9	0.0619	0.3895
CH-C (To-Be)	7	0.0679	0.3005
CH-D	10	0.0742	0.2250
CH-E	11	0.0755	0.2109



CH-C



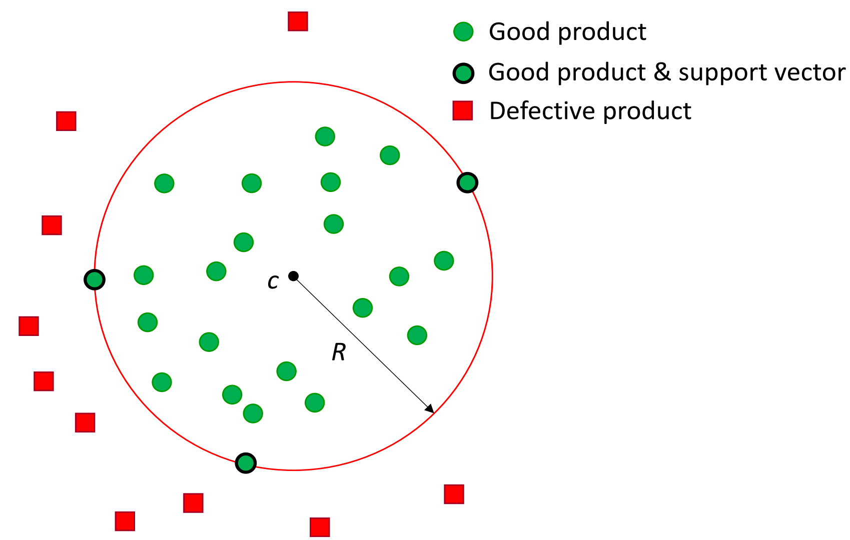
Pan, T.-H., Wong, D. S., and Jang, S.-S. 2012. Chamber matching of semiconductor manufacturing process using statistical analysis. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 42, NO. 4, 571-576.

Advanced Topics

(Self-Study)

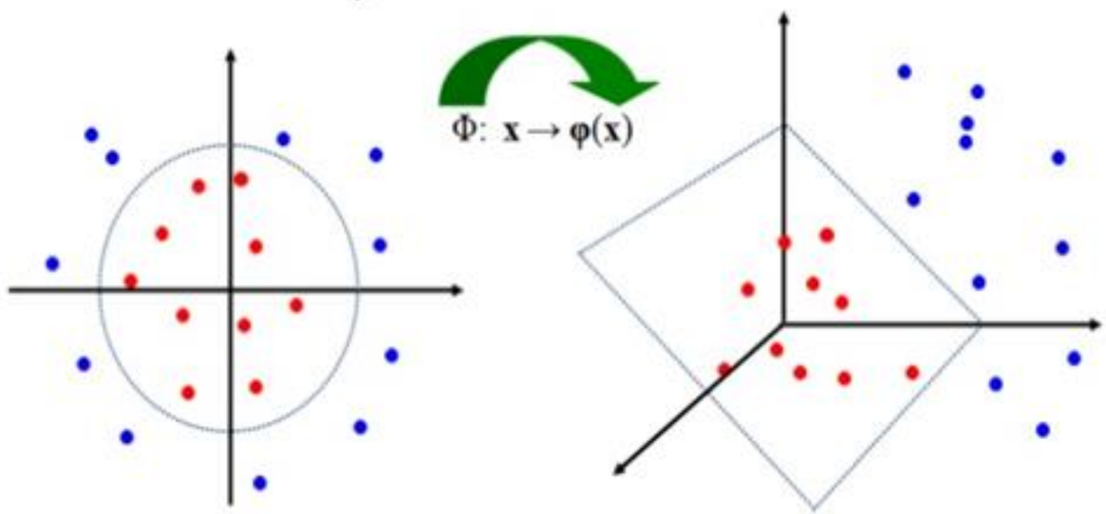
Nonparametric Control Chart

- Support Vector Data Description (SVDD) (Tax and Duin, 2004)
  - inspired by support vector machine (SVM) and used for one-class classification
  - identifies the support vectors to calculate the kernel distance
  - constructs a hypersphere with a minimal volume to envelop the samples in the same class, and thus, used for anomaly detection:
    - the good product is inside the hypersphere while the defective product is outside the hypersphere.
    - The primal problem is equation of circle, which can be transformed into dual problem by Lagrange method.



Support vector machine (SVM) can perform as a classifier for diagnostics to determine the decision boundaries among different classes

- map the input data to a higher dimension feature space, where the transformed data become linearly separable
- does not suffer from multiple local minima and its solution is global and unique; it also does not have the problem of the curse of dimensionality

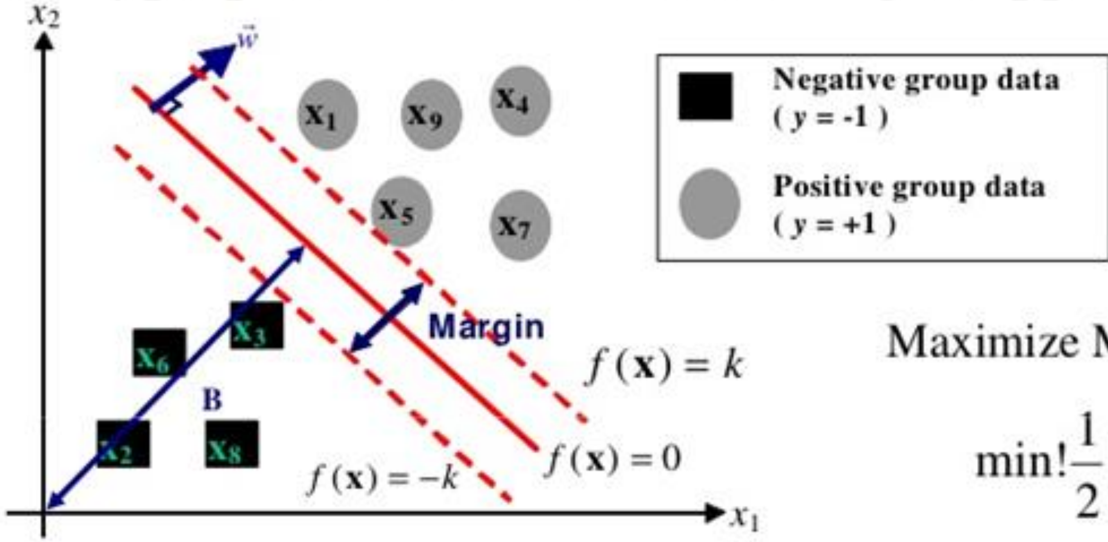


Chen, C. Prognostics and Health Management Fundamentals. Lecture Note, Center for Advanced Life Cycle Engineering, University of Maryland. https://www.slideshare.net/ASQwebinars/prognostics-and-health-management

Support Vector Machine (SVM)

- The optimal hyperplane is determined through support hyperplanes.

;



Maximize Margin $\dfrac{2}{\|\mathbf{w}\|}$

$$\min!\frac{1}{2}\|w\|^2 = \frac{1}{2}w^T w$$

- The supporting hyperplanes can be expressed as:

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b = k \qquad \mathbf{w}^T\mathbf{x}_i + b \geq 1 \quad \text{for} \quad y_i = +1$$

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b = -k \qquad \mathbf{w}^T\mathbf{x}_i + b \leq -1 \quad \text{for} \quad y_i = -1, \quad i = 1,\ldots,n$$

or $\boxed{y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 \geq 0}$ for $i = 1,\ldots,n$

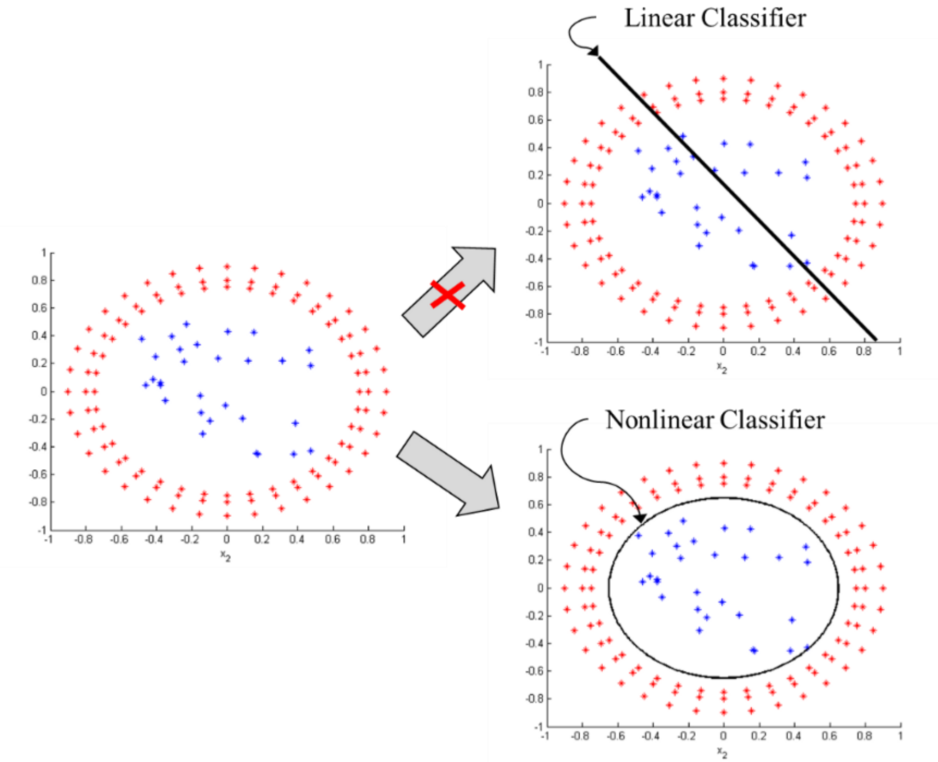
constraints

Chen, C. Prognostics and Health Management Fundamentals. Lecture Note, Center for Advanced Life Cycle Engineering, University of Maryland. https://www.slideshare.net/ASQwebinars/prognostics-and-health-management

Support Vector Machine (SVM)

Kernel Trick

- Used for non-linearly separable problem
- polynomial kernel function: $k(x, y) = (< x, y > +c)^d$
  - $d = 2$ and $c = 0$



$$k(x,y) = \langle \varphi(x), \varphi(y) \rangle = (x^T y)^2$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

此時可以經由簡單的推導得到投影函數(φ)，如下

$$(x^T y)^2 = \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right)^2 = (x_1 y_1 + x_2 y_2)^2 = x_1^2 y_1^2 + x_2^2 y_2^2 + 2 x_1 x_2 y_1 y_2$$

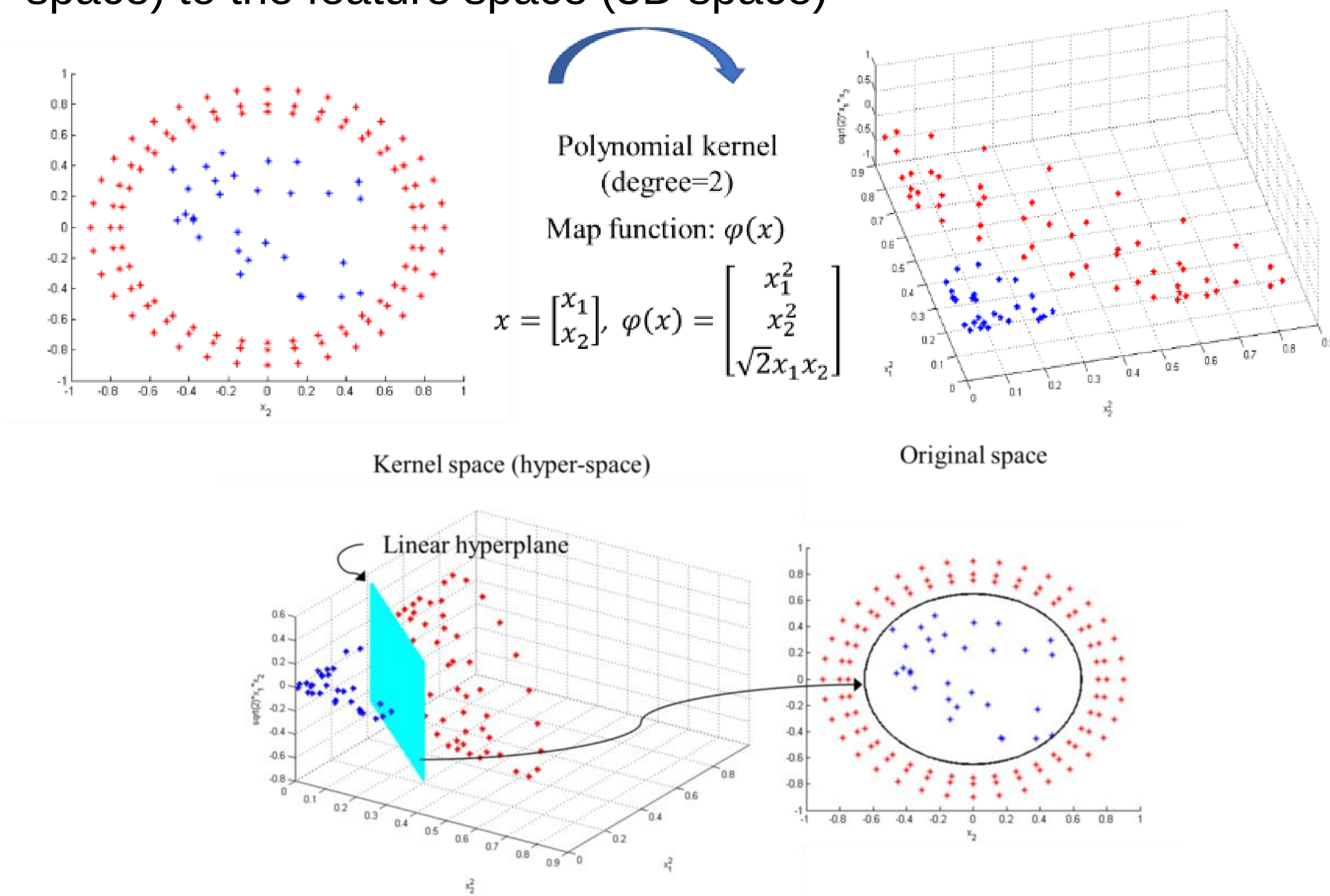
$$= \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2} x_1 x_2 \end{bmatrix}^T \begin{bmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2} y_1 y_2 \end{bmatrix} = \langle \varphi(x), \varphi(y) \rangle$$

$$\varphi(x) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2} x_1 x_2 \end{bmatrix}, \varphi(y) = \begin{bmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2} y_1 y_2 \end{bmatrix}$$

Huang, T. (2018). 機器學習: Kernel 函數. https://chih-sheng-huang821.medium.com/%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-kernel-%E5%87%BD%E6%95%B8-47c94095171

Kernel Trick

- polynomial kernel function projects the data from original space (2D space) to the feature space (3D space)



SVDD Primal Problem

- Let observation $i \in I$, and sample size is $|I|$. The objective function and constraints are as follows.

$$\underset{R}{\text{Min}}\ \mathrm{F}(R, c) = R^2$$
$$\text{subject to } (x_i - c)^T (x_i - c) \leq R^2, \quad \forall i$$

- where x_i are training samples, c is the center of hypersphere, and R is the radius of the hypersphere.

- To address the outlier or noise
  - the penalty cost λ is added into the objective function by introducing the slack variable ξ. The formulation is revised as follows.

$$\underset{R}{\text{Min}}\ \mathrm{F}(R, c, \xi) = R^2 + \lambda \sum_i \xi_i$$
$$\text{subject to } \|x_i - c\|^2 \leq R^2 + \xi_i, \quad \forall i$$
$$\xi_i \geq 0, \quad \forall i.$$

Primal Problem

- Lagrange form
- $L(R, c, \alpha, \gamma, \xi) = R^2 + \lambda \sum_i \xi_i - \sum_i \alpha_i \{R^2 + \xi_i - (\|x_i\|^2 - 2cx_i + \|c\|^2)\} - \sum_i \gamma_i \xi_i$,
  - where $\alpha \geq 0$ and $\gamma \geq 0$ are the Lagrange multipliers. Then, the first-order conditions, i.e. Karush-Kuhn-Tucker (KKT) conditions, can be used to find the optimal solutions.
  - $\frac{\partial L}{\partial R} = 0$, we derive $\sum_i \alpha_i = 1$
  - $\frac{\partial L}{\partial c} = 0$, we drive $c = \frac{\sum_i \alpha_i x_i}{\sum_i \alpha_i} = \sum_i \alpha_i x_i$
  - $\frac{\partial L}{\partial \xi} = 0$, we derive $\lambda - \alpha_i - \gamma_i = 0$

Dual Problem

Max $L = \sum_i \alpha_i (x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j)$

subject to $0 \leq \alpha_i \leq \lambda, \forall i$

$\sum_i \alpha_i = 1, \forall i.$

Nonparametric Control Chart

SVDD (discuss the relation between α_i and λ)

- Sample is inside the hypersphere
  - $\|x_i - c\|^2 < R^2 + \xi_i$, then $\alpha_i = 0$, $\gamma_i = 0$, and $\xi_i > 0$.
  - the Lagrange multipliers equal zero.
- Sample is support vector
  - $\|x_i - c\|^2 = R^2 + \xi_i$, then $0 < \alpha_i < \lambda$, $\gamma_i = 0$, and $\xi_i > 0$.
  - the sample is outside the hypersphere with the slack ξ. which allows the outlier or noise. In fact, it is still the class inside the hypersphere. The Lagrange multiplier α is not equal to zero.
- Sample is outside the hypersphere
  - $\|x_i - c\|^2 > R^2 + \xi_i$, then $\alpha_i = \lambda$, $\gamma_i > 0$, and $\xi_i = 0$.

- Distance measure
  - the distance for the new sample z and identify its class type by comparing with the radius of SVDD.
  - $\text{Distance} = \|z - c\|^2 = (z \cdot z) - 2\sum_i \alpha_i (z \cdot x_i) + \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j) \leq R^2$
    - where $R^2 = (x_i \cdot x_k) - 2\sum_i \alpha_i (x_i \cdot x_k) + \sum_{i,j} \alpha_i \alpha_j (x_i \cdot x_j)$, $\forall x_k \in SV_{<\lambda}$

SVDD

- Kernel function $K(\cdot)$
  - like SVM to project the dataset to a high-dimensional space for classification. That is, a *Kernel Trick*.

$$\text{Max } L = \sum_i \alpha_i K(x_i \cdot x_i) - \sum_{i,j} \alpha_i \alpha_j K(x_i \cdot x_j)$$

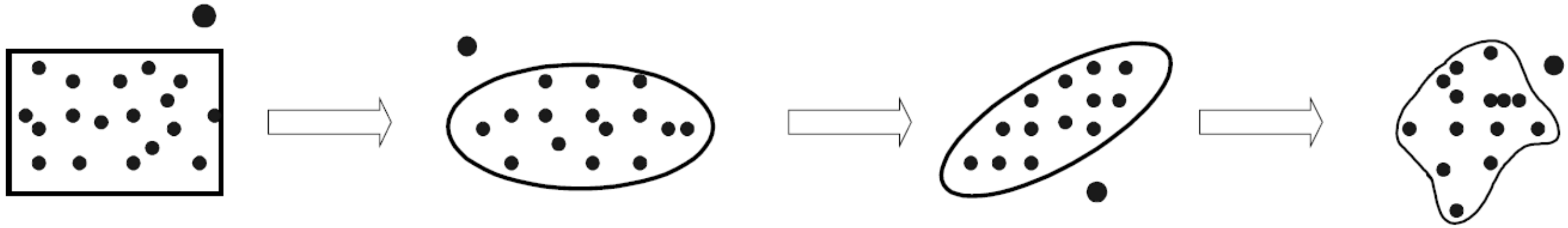
subject to $0 \leq \alpha_i \leq \lambda, \forall i$

$\sum_i \alpha_i = 1, \forall i.$

- Distance Measure
  - $\text{Distance} = \|z - c\|^2 = K(z \cdot z) - 2\sum_i \alpha_i K(z \cdot x_i) + \sum_{i,j} \alpha_i \alpha_j K(x_i \cdot x_j) \leq R^2$
    - where $R^2 = K(x_i \cdot x_k) - 2\sum_i \alpha_i K(x_i \cdot x_k) + \sum_{i,j} \alpha_i \alpha_j K(x_i \cdot x_j), \forall x_k \in SV_{<\lambda}$

Sun, R., & Tsung, F. (2003). A kernel-distance-based multivariate control chart using support vector methods. International Journal of Production Research, 41(13), 2975–2989.

Nonparametric Control Chart



a. Individually

b. Independently ; Bivariate Normal Distribution

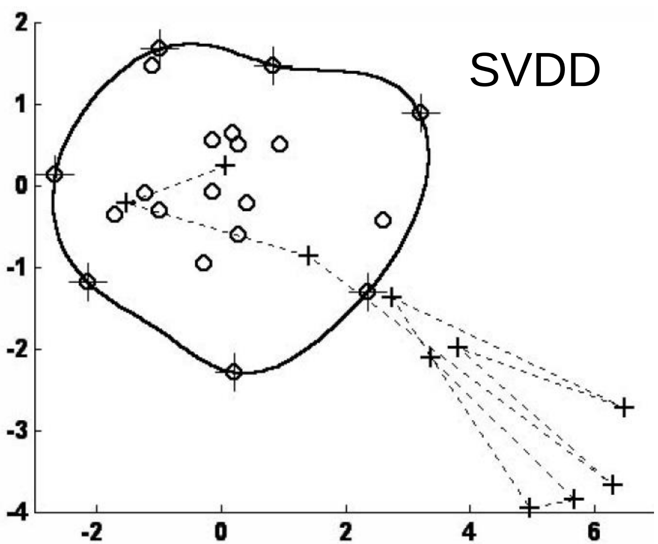
c. Corelated ; Bivariate Normal Distribution

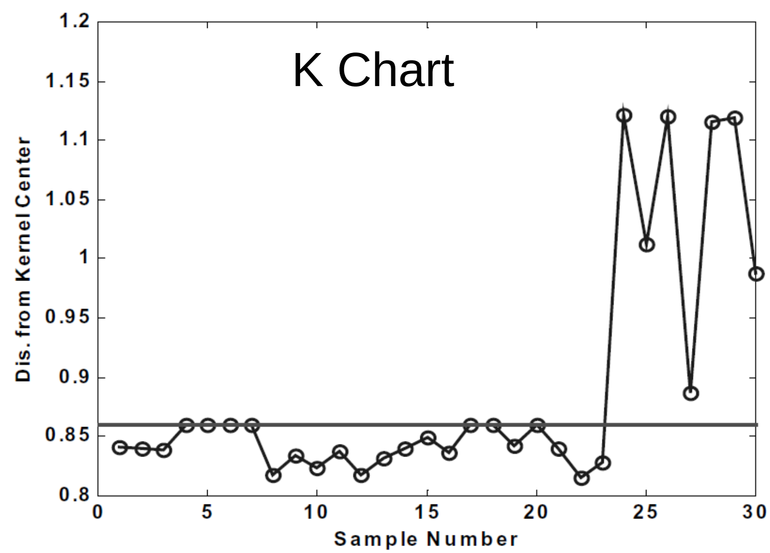
d. Kernel-distance based MSPC

Shewhart Control Chart

Chi-Square Control Chart ; Hotelling T- Square Control Chart

K-Chart





Sun, R., & Tsung, F. (2003). A kernel-distance-based multivariate control chart using support vector methods. International Journal of Production Research, 41(13), 2975–2989.

Concept Drift

Concept Drift & Domain Adaptation

- Concept drift detection (何時模型不準了需要retraining)
  - Hypothesis test and control chart
- Concept drift understanding (釐清模型不準的根本原因)
  - Time of concept drift occurs (When)
  - The severity of concept drift (How)
  - The drift regions of concept drift (Where)
- Drift adaptation (如何retrain模型)
  - Training new models for global drift
  - Model ensemble for recurring drift
  - Adjusting existing models for regional drift

Transfer Learning

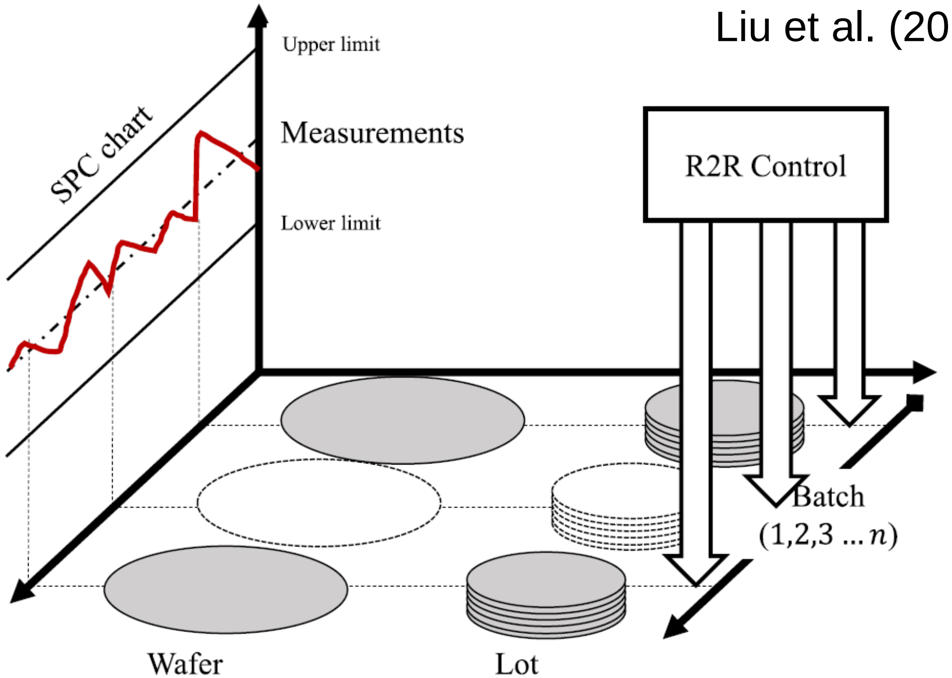
- Instance based transfer learning
- Feature based transfer learning
- Parameter based transfer learning
- Relational knowledge-based transfer learning

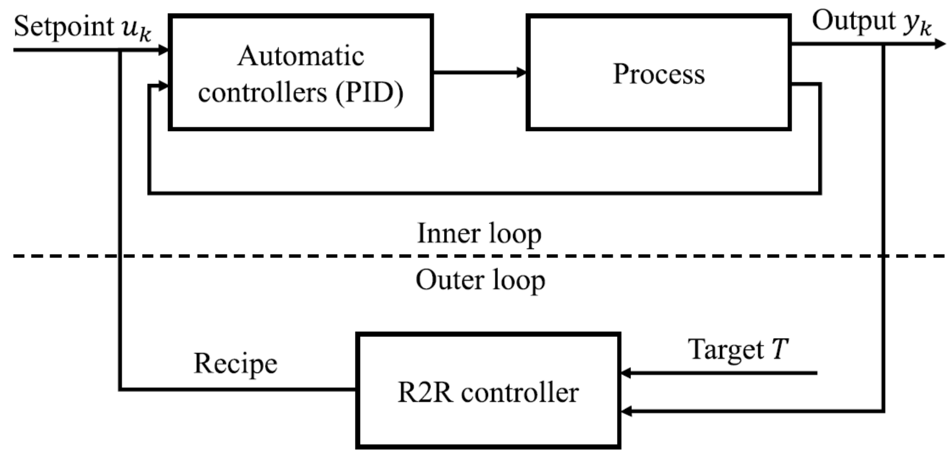
Lu et al. (2020). Learning under concept drift: a review. https://arxiv.org/abs/2004.05785v1

Ran, et al. (2019). A survey of predictive maintenance: systems, pruposes and approaches. IEEE Communications Surveys & Tutorials.

R2R Control

Liu et al. (2018)





- exponentially weighted moving average (EWMA), double EWMA
- model predictive control (MPC)
- optimizing adaptive quality controller (OAQC)
- artificial neural network (ANN)

EWMA Control

- with linear equation: $y_k = a_k + bu_k + \varepsilon_k$
- It regulates the process with time-varying intercept by adjusting a_k; i.e.,

$$u_k = \frac{T - a_k}{b} \qquad a_k = \lambda(y_{k-1} - bu_{k-1}) + (1 - \lambda)a_{k-1}$$

Liu, K., Chen, Y., Zhang, T., Tian, S., & Zhang, X. 2018. A survey of run-to-run control for batch processes. ISA Transactions, 83, 107-125.

Summary & Conclusion

Data Type	Defect Definition	Subgroup Size	Chart Type
Attribute Data Counted as Discrete Events	Defect Data -Number of defects, not number of defective units	Constant Subgroup Size (n constant)	c Chart Number of Defects
		Variable Subgroup Size (n constant or varies)	u Chart Defects per Unit
	Defective Unit Data	Constant Subgroup Size, Usually >= 50 (n constant)	np Chart Number of Defective Units
		Variable Subgroup Size, Usually >= 50 (n constant or varies)	p Chart Fraction of Defective Units
Variable Data Measured on a Continuous Scale		Subgroup Size = 1 (time order)	X and R_m Moving Range
		Subgroup Size < 10 (time order)	$\bar{X}$ and R
	(time order	or rational subgroup) Subgroup Size >= 10	$\bar{X}$ and s

For each subgroup, count is the number of occurrences. Numerator and denominator measure different things.

Each item has the attribute or it doesn’t. The numerator is a subset of the denominator.

Hart, M. and Hart, R. Statistical Process Control for Health Care. Pacific Grove, CA: Duxbury, 2002.
https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm

Summary & Conclusion

PDLab

Variable Data Chart Formulas		
Chart Type	**Subgroup Size**	**Control Limits**
$\bar{X}$ and R Average and Range Chart	< 10 (usually 3-5)	$\bar{X}$ Central Line: $\bar{\bar{X}} = \frac{(\bar{X}_1+\bar{X}_2+\dots\bar{X}_k)}{k}$ $\bar{X}$ UCL = $\bar{\bar{X}} + A_2\bar{R}$ $\bar{X}$ LCL = $\bar{\bar{X}} - A_2\bar{R}$ R Central Line: $\bar{R} = \frac{(R_1+R_2+\dots R_k)}{k}$ R UCL = $D_4\bar{R}$ R LCL = $D_3\bar{R}$
X and mR Individuals and Moving Range Chart Note: $\widetilde{mR}$ = Median Moving Range	1	X Central Line: $\bar{X} = \frac{(X_1+X_2+\dots X_k)}{k}$ X UCL = $\bar{X} + (3.14 \times \widetilde{mR})$ X LCL = $\bar{X} - (3.14 \times \widetilde{mR})$ mR Central Line: Median Moving Range mR UCL = $(3.87 \times \widetilde{mR})$

Attribute Data Chart Formulas		
Chart Type	**Subgroup Size**	**Control Limits**
p Chart Fraction Defective	Variable or Constant	Central Line: $\bar{p} = \Sigma np / \Sigma n$ UCL = $\bar{p} + 3\sqrt{(\bar{p}(1-\bar{p}))/n}$ LCL = $\bar{p} - 3\sqrt{(\bar{p}(1-\bar{p}))/n}$
np Chart Number Defective	Constant	Central Line: $\overline{np} = \Sigma np / k$ UCL = $\overline{np} + 3\sqrt{\overline{np}(1-\bar{p})}$ LCL = $\overline{np} - 3\sqrt{\overline{np}(1-\bar{p})}$
c Chart Number of Defects	Constant	Central Line: $\bar{c} = \Sigma c / k$ UCL = $\bar{c} + 3\sqrt{\bar{c}}$ LCL = $\bar{c} - 3\sqrt{\bar{c}}$
u Chart Number of Defects per Unit	Variable or Constant	Central Line: $\bar{u} = \Sigma c / \Sigma n$ UCL = $\bar{u} + 3\sqrt{\bar{u}/n}$ LCL = $\bar{u} - 3\sqrt{\bar{u}/n}$

https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm