Part I

Statistical Process and Quality Control

Introduction

Quality and Process Control

Definition of Quality
- Quality means fitness for use. and
- Quality is inversely proportional to variability.
- My definition: *Is* and *should* are the same.

Statistical Process Control (SPC)
- Statistical process control is, first and foremost, a way of thinking which happens to have some tools attached.

The Magnificent Seven
1. histogram 2. check sheet 3. Pareto chart 4. defect concentration diagram 5. cause-and-effect diagram 6. control chart 7. scatter diagram

Control Charts

The basis of a control chart is a statistical hypothesis test.

Hypothesis test

Question:
is $|\bar{x} - \mu_0|$ significant?
- μ_0 target value
- $\bar{x}$ Arithmetic mean of the measurements

Hypothesis:
Two-sided statistical test to check the two alternative hypotheses:

$$\begin{aligned} H_0 &: mu_0 = \bar{x} \text{ i.e. process is not disturbed} \\ H_1 &: mu_0 \neq \bar{x} \text{ i.e. process is disturbed} \end{aligned} \tag{1}$$

Statistical test:

Test statistic: (z-test, since σ is known)

$$z = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} \tag{2}$$

Critical value: i.e. q-quantile of the normal distribution

$$P(|z| \leq z_q) = q \tag{3}$$

with $q = 1 - \frac{\alpha}{2}$, where $\alpha = 0.0027 \rightarrow z_q \approx 3$

Statistical conclusion:
If $\leq |z|$ $z_q \rightarrow$ accept null hypothesis, i.e. process is not disturbed.
If $< |z|$ $z_q \rightarrow$ reject null hypothesis, i.e. process is disturbed.

Reverse it and determine acceptance and rejection limits of the test!

Control limits:

$$UCL = \mu_0 + z_q \frac{\sigma}{\sqrt{n}} \qquad LCL = \mu_0 - z_q \frac{\sigma}{\sqrt{n}} \tag{4}$$

Statistical conclusion:
If $LSL \leq \bar{x} \leq UCL \rightarrow$ process is not disturbed.
If $\bar{x} < LCLorUCL < \bar{x} \rightarrow$ process is disturbed.

Problem: In general, process standard deviation is unknown.
- Monitoring the mean and the variation of a process.
- First, monitoring the variation, then (if variation under control) monitoring the mean.

Solution: Control charts! :)

The Control Chart

No.	sample values						mean	sd	range
1	x_{11}	x_{12}	$\cdots$	x_{1j}	$\cdots$	x_{1n}	$\bar{x}_1$	s_1	R_1
2	x_{21}	x_{22}	$\cdots$	x_{2j}	$\cdots$	x_{2n}	$\bar{x}_2$	s_2	R_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	x_{i1}	x_{i2}	$\cdots$	x_{ij}	$\cdots$	x_{in}	$\bar{x}_i$	s_i	R_i
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	x_{k1}	x_{k2}	$\cdots$	x_{kj}	$\cdots$	x_{kn}	$\bar{x}_k$	s_k	R_k

Figure 1: Data Set with Mean, Standard Deviation and Range

Mean values

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \tag{5}$$

Standard deviations

$$s_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2} \tag{6}$$

Ranges

$$R_i = max\{x_{ij} | j \in \{1, ..., n\}\} - min\{x_{ij} | j \in \{1, ..., n\}\} \tag{7}$$

for all $i \in \{1, ..., k\}$

Control Chart for $\bar{x}$ and R

R Chart Centerline

$$\bar{R} = \frac{1}{k}\sum_{i=1}^{k} R_i \tag{8}$$

Control limits

$$UCL = D_4\bar{R}; \qquad LCL = D_3\bar{R} \tag{9}$$

$\bar{x}$ Chart based on R chart

Control limits

$$UCL = \mu + 3\frac{\sigma}{\sqrt{n}}; \qquad LCL = \mu - 3\frac{\sigma}{\sqrt{n}} \tag{10}$$

Problem: μ and σ are in general unknown and must be estimated from the process data.
Two-stage process:
- Make sure that the process standard deviation (R chart) is under statistical control. That is, if some samples are out of bounds, it is recommended to omit these measurements and recalculate the limits.
- Use $\bar{R}$ to estimate the process standard deviation.

Centerline
We know that for an independent sample $x_1, ..., x_n$ from a normal distribution with parameters μ and σ the mean

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{11}$$

satisfies

$$E(\bar{x}) = \mu \quad and \quad Var(\bar{x}) = \frac{\sigma^2}{n} \tag{12}$$

The mean is an unbiased estimator with the standard error

$$SE(\bar{x}) = \frac{\sigma}{\sqrt{n}} \tag{13}$$

Assumption: R chart is under statistical control.
- The value $\bar{R}$ is a reliable estimate for the mean range.
- The value $\bar{R}$ is a reliable estimate for the process standard deviation

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \tag{14}$$

Any samples excluded for construction of the R chart should also be disregarded for construction of the $\bar{x}$ chart. This results in a sample of $k^\star$ valid samples, (where $k^\star$ denotes the reduced number of samples). Mean values of $\bar{x}_1, ..., \bar{x}_{k^\star}$ provide an estimate of μ, i.e

$$\bar{\bar{x}} = \frac{1}{k^\star}\sum_{i=1}^{k^\star} \bar{x}_i \tag{15}$$

Control limits

$$\begin{aligned} UCL &= \bar{\bar{x}} + 3\frac{\bar{R}}{d_2}\frac{1}{\sqrt{n}} \approx \bar{\bar{x}} + A_2\bar{R} \\ UCL &= \bar{\bar{x}} - 3\frac{\bar{R}}{d_2}\frac{1}{\sqrt{n}} \approx \bar{\bar{x}} - A_2\bar{R} \end{aligned} \tag{16}$$

Control Chart with $\bar{x}$ and s

s Chart

Centerline The centreline of the s chart is denoted by $\bar{s}$ and is calculated from the arithmetic mean of the standard deviations

$$\bar{s} = \frac{1}{k}\sum_{i=1}^{k} s_i \tag{17}$$

Control limits

$$UCL = B_4\bar{s}; \quad LCL = B_3\bar{s} \tag{18}$$

$\bar{x}$ Chart based on s chart

Using an s chart of a process that is under control, the process standard deviation can be estimated by

$$\hat{\sigma} = \frac{\bar{s}}{c_4} \tag{19}$$

Any samples excluded for construction of the s chart should also be disregarded for construction of the $\bar{x}$ chart. This results in a sample of $k^\star$ valid samples, (where $k^\star$ denotes the reduced number of samples). Mean values of $\bar{x}_1, ..., \bar{x}_{k^\star}$ provide an estimate of μ, i.e

$$\hat{\mu} = \bar{\bar{x}} = \frac{1}{k^\star}\sum_{i=1}^{k^\star} \bar{x}_i \tag{20}$$

Control limits

$$\begin{aligned} UCL =& \bar{\bar{x}} + 3\frac{\bar{s}}{c_4}\frac{1}{\sqrt{n}} \approx \bar{\bar{x}} + A_3\bar{s} \\ UCL =& \bar{\bar{x}} - 3\frac{\bar{s}}{c_4}\frac{1}{\sqrt{n}} \approx \bar{\bar{x}} - A_3\bar{s} \end{aligned} \tag{21}$$

Individual Control Charts

Individual control charts have exactly one measurement per sample.
Problem: You cannot estimate variability from a single measurement.
Idea: Use variation of two adjacent measurements.

Moving ranges

$$MR_i = |x_{i+1} - x_i| \tag{22}$$

for all $i \in \{1, ..., n-1\}$.

Arithmetic mean of the moving ranges

$$\overline{MR} = \frac{1}{n-1}\sum_{i=1}^{n-1} MR_i \tag{23}$$

Estimated process standard deviation

$$\hat{\sigma} = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128} \tag{24}$$

Since two neighboring measurements were used to calculate the moving ranges we have $d_2 = 1.128$.

Centerline
The centerline for the individuals control chart is the arithmetic mean of the measured values.

$$\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i \tag{25}$$

Control limits

$$UCL = \bar{x} + 3\frac{\overline{MR}}{1.128}; \quad LCL = \bar{x} - 3\frac{\overline{MR}}{1.128} \tag{26}$$

Control Charts for Attributes Data – p Chart

Number of defectives under number tested is a discrete random variable.

Given: Random sample of size n, of which D parts are defective We know: The number of defective D under n examined parts follows a binomial distribution with the unknown probability p of success.

Estimated probability

$$\hat{p} = \frac{D}{n} \tag{27}$$

Variance

$$Var(\hat{p}) = \frac{p(1-p)}{n} \tag{28}$$

Given:
- k random samples with $n_1, ..., n_k$ values.
- Each of these samples contains $d_1, ..., d_k$ defective products.

k relative frequencies

$$p_1 = \frac{d_1}{n_1}, ..., p_k = \frac{d_k}{n_k} \tag{29}$$

Centerline
The centreline and the control limits of a p chart are again determined from a stable trial run with $k^\star$ valid samples.
Again $k^\star \leq k$ is the reduced number of samples.
Distinguish 2 cases:
1. The sample sizes $n_1, ..., n_k$ are all equal to n.
2. The sample sizes are not all equal.

Case 1

Centerline

$$\bar{p} = \frac{1}{k^\star}\sum_{i=1}^{k^\star} p_i \tag{30}$$

Control limits

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}; \quad LSL = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{31}$$

Case 2

Centerline

$$\bar{p} = \frac{d_1 + \cdots + d_{k^\star}}{n_1 + \cdots + n_{k^\star}} \tag{32}$$

Control limits

$$UCL_i = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}; \quad LSL_i = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}} \tag{33}$$

The control limits now depend on the index i.

Part II
Multiple Regression

Part III
Design of Experiment