Control Chart Concepts

- Introduction
- Types of Control Charts
- Basic Principles
- General Form of Control Limits
- Choice of Control Limits
- Rational Subgroups
- Subgroup Size Considerations
- Variable Subgroup Sizes
- Average Run Length
- Detecting a Mean Shift
- The $p$ chart
- Process Capability
- Analysis of Patterns on Control Charts

This covers sections 4.3

**Introduction**

- Principal purpose: early detection of an ‘out-of-control’ process
- A process is out of control if it is producing items which are
  - off target or
  - too variable
- An out-of-control process is likely to produce many nonconforming items
- If an assignable cause can be found, the process can be corrected and brought back into control.
- A capable, in-control process will produce fewer nonconforming items.

**Types of Control Charts**

1. Variables - $\bar{x}$, $R$, $s$ - applicable to continuous measurement data.
2. Attributes - $p$, $c$, $u$ - applicable to count-type data. e.g. nos. of defective items in a sample, etc.
   - e.g. $p$-chart: $p \pm 3\sqrt{p(1-p)}/n$ where $p$ is the probability of nonconformance, and $n$ is the number of items sampled.
3. Cusum - cumulative sum charts - used to detect small shifts in the process.
4. EWMA - exponentially weighted moving averages charts - small shift detection.

**Basic Principles**

- Samples of measurements are periodically taken at one or more stages of a production process to provide data for the monitoring of the process.
- Based on each sample, a statistic is computed and plotted against time.
- The result is a time series of the observed statistic values.
• Because of chance variation, these observed statistic values will naturally fluctuate.

• When the process is in control, the fluctuations will be around a common mean $\mu$ (the in-control mean).

• A horizontal centerline is plotted on the chart representing the in-control mean.

• This centerline is used as a reference line to compare the plotted points. Sometimes, a target value is used.

• Occasionally, a problem develops causing a change in the process.

• When the process changes, the mean of the process usually changes. This is considered to be an out of control situation.

• If the process is out of control at one or more of the plotted time points, the corresponding fluctuations are likely to be larger. Often, a trend or a shift may become evident.

• Points far from the centerline signal that the process may be out of control.

• Control limits are horizontal lines drawn on either side of the centerline.

• When in control, most of the points lie between the control limits. Variation among such points is attributed to chance.

• The process is judged out of control when a point plots outside the control limits. The extra variation is attributed to an assignable cause.

• The process is also likely out of control if there is a systematic pattern (e.g. cyclic, increasing, decreasing) in the plotted points.

• Control Charts and Hypothesis Testing:
  
  - Null hypothesis $H_0$: process is in control
  - Alternative hypothesis $H_1$: process is out of control
  - When a point plots within the control limits, the null hypothesis is not rejected
  - When a point plots outside the control limits, the null hypothesis is rejected
  - Type I error:
    1. Rejecting the null hypothesis when it is true
    2. Concluding the process is out of control when it isn’t
    3. False Alarm: an in-control point plots outside the control limits
  - Type II error:
    1. Not rejecting the null hypothesis when it is false
    2. Failing to detect an out of control condition: an out-of-control point plots inside the control limits

• Upper control limit:
  
  \[
  \text{UCL} = \text{E[statistic]} + k(\text{sd of stat.})
  \]

• Centerline:
  
  \[
  \text{Centerline} = \text{E[statistic]}
  \]

• Lower control limit:
  
  \[
  \text{LCL} = \text{E[statistic]} - k(\text{sd of stat.})
  \]
Recall the Central Limit Theorem: If $X_1, X_2, \ldots, X_n$ are independent random variables with mean $\mu$ and variance $\sigma^2$, then $\bar{X}$ is approximately normally distributed with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$.

$\Rightarrow \bar{x}$-chart:

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

Centerline = $\mu$

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

The sample means $\bar{x}_1, \bar{x}_2, \ldots$ are plotted on this chart.

Example.

- Samples of 4 plastic cords are taken every hour for tensile testing.
- The process mean (when in control) is 480 kg. The standard deviation $\sigma$ is 8. We use $k = 3$.
- $LCL = 480 - 3(8)/2 = 468$ and $UCL = 480 + 3(8)/2 = 492$

<table>
<thead>
<tr>
<th>time</th>
<th>$\bar{X}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>476</td>
</tr>
<tr>
<td>2</td>
<td>466</td>
</tr>
<tr>
<td>3</td>
<td>484</td>
</tr>
<tr>
<td>4</td>
<td>466</td>
</tr>
</tbody>
</table>

Observed data: 5 470
              6 494
              7 486
              8 496
              9 488
             10 482

- We judge the process to be out of control at times 6 and 8 i.e. We reject the null hypothesis (that the mean is 480) at the .0027 % significance level at these two time points.
- An assignable cause should be found for the two points

E.g. $p$-chart:

$$UCL = p + k\sqrt{p(1-p)/n}.$$  

Centerline = $p$

$$LCL = p - k\sqrt{p(1-p)/n}.$$  

Here, $p$ denotes the proportion of the theoretical population of items which are nonconforming.

The sample proportions $\hat{p}_1, \hat{p}_2, \ldots$ are plotted on this chart.

Choice of Control Limits

$k$ is most often chosen to be 3.

Equivalent to testing whether the process mean is $\mu$ using a type I error probability ($\alpha$) of .0027.

- Moving the control limits closer to the centerline increases the risk of type I error
- Moving the control limits away from the centerline increases the risk of type II error.

Trade-off: compare costs of shutting down the process more often (due to false alarms) with costs of producing more nonconforming parts (because of more undetected out of control conditions)
Sometimes warning limits are used:

\[ UWL = \text{mean} + 2 \text{ standard deviation} \]

\[ LWL = \text{mean} - 2 \text{ standard deviation} \]

One possible response to points plotting outside these limits is to increase the subgroup size or sampling frequency.

**Expected pattern of variation when in control:**

1. about 68% of the points fall within 1 sigma of the centerline.
2. about 27% of the points fall between 1 and 2 sigma of the centerline.
3. about 5% of the points fall between 2 and 3 sigma of the centerline
   i.e. 5% of the in-control points are expected to plot outside the warning limits

**Rational Subgroups**

- Samples are referred to as subgroups.
- Choose subgroups to either
  1. minimize variation within a sample due to a process shift by sampling units which were produced at the same time
  2. represent all process output over the sampling interval

**Subgroup Size Considerations**

- Larger samples make it easier to detect small shifts in the process (since the standard error is reduced as \( n \) increases).

- Factors influencing Sampling Frequency
  - Cost of Sampling
  - Losses associated with letting the process run out of control
  - Rate of production
  - Probabilities of process mean shifts occurring.

**Variable Subgroup Sizes**

Sometimes one is forced by circumstances to use different subgroup sizes: \( n_1, n_2, \ldots, n_m \).

There are 3 ways of setting up control charts in this case:

1. Use the average sample size: \( \bar{n} = \frac{1}{m} \sum_{j=1}^{m} n_j \) to set up the control limits:

   \[ \text{UCL} = \mu + k\sigma/\sqrt{\bar{n}} \]
   \[ \text{Centerline} = \bar{x} \]
   \[ \text{LCL} = \mu - k\sigma/\sqrt{\bar{n}} \]

2. Use variable control limits:

   \[ \text{UCL}_i = \mu + k\sigma/\sqrt{n_i} \]
   \[ \text{Centerline} = \bar{x} \]
   \[ \text{LCL}_i = \mu - k\sigma/\sqrt{n_i} \]
3. Use standardized control limits:

\[ \text{UCL}_i = k \]

Centerline = 0

\[ \text{LCL}_i = -k \]

and the points

\[ z_i = \frac{\bar{x}_i - \mu}{\sigma / \sqrt{n_i}} \]

**ARL - Average Run Length**

- Consider the general control chart where subgroups of size \( n \) are taken from a process.
- For each subgroup, a statistic \( T \) is computed and plotted against the corresponding time.
- The control limits are UCL and LCL.
- For any subgroup,

\[ p_d = P(T > \text{UCL} \text{ or } T < \text{LCL}) \]

\( p_d \) is the probability that the statistic plots outside the control limits.
- Let \( Y \) be the number of subgroups until such an event occurs.
- Let’s find the distribution of \( Y \) under the assumption that the subgroups are independent of each other and process is stationary.

\[
P(Y = 1) = p_d \\
P(Y = 2) = (1 - p_d)p_d \\
P(Y = 3) = (1 - p_d)^2p_d \\
P(Y = j) = (1 - p_d)^{j-1}p_d
\]

- Thus, \( Y \) has a geometric distribution with parameter \( p_d \).
- The average run length is defined as \( E[Y] \).

\[
E[Y] = \sum_{j=1}^{\infty} jP(Y = j) \\
= p_d \sum_{j=1}^{\infty} j(1 - p_d)^{j-1} \\
= p_d \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} (1 - p_d)^{k-1} \\
= p_d \sum_{j=1}^{\infty} \left( 1 - (1 - p_d)^{j-1} \right) \\
= \frac{1}{p_d}
\]

**Plastic cord Example**

- Recall: \( \mu = 480kg, \sigma = 8 \) \( n = 4 \). (in-control)
- 3 sigma control limits for \( \bar{x} \)-chart: UCL = 492, LCL = 468
- Q: What is the ARL, when this process is in control?
• **A:** ARL = $1/p_d$

  \[
p_d = P(\bar{X} \notin (468, 492)|\text{in control})
  = P(\bar{X} < 468) + P(\bar{X} > 492)
\]

where $\bar{X}$ is N(480, 16).

\[
p_d = P(Z < -3) + P(Z > 3) = 2P(Z > 3) = .0027
\]

since $Z$ is N(0,1).

\[
ARL = 1/.0027 = 370
\]

**Detecting a Mean shift**

• **Q:** Suppose that we are interested in detecting shifts in the mean level of 5, say from 480 to 475. How long would it take (on average) for this control chart to detect such a shift?

• **A:** Compute ARL assuming that the process mean has shifted to 475 kg, but the control limits remain as before:

  \[
p_d = P(\bar{X} \notin (468, 492)|\mu = 475)
  = P(\bar{X} \leq 468|\mu = 475).
\]

\[
s = 8 \text{ and } n = 4 \text{ so } \frac{s}{\sqrt{n}} = 4.
\]

\[
p_d = P(\frac{\bar{X} - \mu}{s/\sqrt{n}} < \frac{468 - 475}{4}).
\]

\[
p_d = P(Z < -1.75) = P(Z > 1.75) = .04
\]

\[
ARL = 1/.04 = 25.
\]

• On average, 25 subgroups will be inspected until such a shift is detected. ⇒ 25 hours.

• If this is too much time, either sampling frequency must be increased or sample size must be increased.

• **Q:** What sample size is required to reduce the ARL to 3, while maintaining the in-control ARL at 370?

• **A:** Use 3 sigma limits to retain the in-control ARL at 370.

• Require

  \[
p_d = 1/3
\]

where

\[
p_d = P(\bar{X} \leq 480 - 3(8)/\sqrt{n})
\]

• $\bar{X}$ is distributed as N(475, 64/n).

\[
.33 = P \left( Z \leq \frac{5 - 24/\sqrt{n}}{8/\sqrt{n}} \right)
\]

\[
.33 = P \left( Z \geq 3 - 5\sqrt{n}/8 \right)
3 - 5\sqrt{n}/8 = .44
2.56 = 5\sqrt{n}/8
n = (2.56(8)/5)^2 = 17
\]

• A subgroup size of 17 is sufficient to reduce ARL for this size of shift to 3.

**Results for the \(\bar{x}\)-chart**
1. ARL for given mean shift and subgroup size $n$:
   - size of mean shift $= \Delta = \mu_{\text{in control}} - \mu_{\text{shifted}}$.
   - process variance $\sigma^2$.
   - in control mean $= \mu$.
   - $k\sigma$ control limits. $\mu + k\sigma/\sqrt{n}$.

   $ARL = \frac{1}{P(Z > k - \frac{\Delta\sqrt{n}}{\sigma}) + P(Z > k + \frac{\Delta\sqrt{n}}{\sigma})}$

   where $Z$ is N(0,1).

2. sample size for given ARL, and shift:

   $n = \left(\frac{(z_{pd} - k)\sigma}{|\Delta|}\right)^2$

Results for the $p$-chart

- Let $p$ be the proportion nonconforming (for the in-control process).
- Subgroups of size $n$ are inspected.
- $X$ is the number of nonconforming in a particular subgroup.
- $\hat{p} = X/n$ is the observed proportion nonconforming. This is plotted against time for the $p$-chart.
- $UCL = p + k\sqrt{p(1-p)/n}$.
- $LCL = p - k\sqrt{p(1-p)/n}$.
- $ARL = 1/p_d$.
- $p_d = P(\hat{p} > UCL) + P(\hat{p} < LCL)$.
- If process shifts to $p_s$, find ARL. Note: control limits remain at UCL and LCL.

   $p_d = P(X > nUCL) + P(X < nLCL)$

   where $X$ is binomial with parameters $n$ and $p_s$.

- For large $n$, we can use the normal approximation:

   $p_d = P\left(\frac{Z}{\sqrt{p_s(1-p_s)}} > \frac{(p - p_s)\sqrt{n} + k\sqrt{p(1-p)}}{\sqrt{p_s(1-p_s)}}\right)$

   $+ P\left(\frac{Z}{\sqrt{p_s(1-p_s)}} > \frac{(p_s - p)\sqrt{n} + k\sqrt{p(1-p)}}{\sqrt{p_s(1-p_s)}}\right)$

   where $Z$ is N(0,1).

- Required sample size can be found by solving for $n$, for given ARL and shift size.

Introduction to Process Capability

- Let $X$ denote a single measurement, and let $\bar{X}$ be the average for a sample of measurements.

- Different kinds of limits:
- **Control Limits** – these are used to indicate whether the process is in control or not. As long as LCL ≤ \( \bar{X} \) ≤ UCL, we can believe the process is in control.
- **Specification Limits** – these are requirements of the process itself: i.e. we require LSL ≤ X ≤ USL with as high probability as possible.
- **Process Limits** – these indicate where the process usually is. If \( X \) is approximately normally distributed with mean \( \mu \) and variance \( \sigma^2 \), then

\[
P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx .99
\]

- We cannot assess process capability until we are sure that the process is in control: stable and non-systematic. The control chart limits are used to determine whether the process is in control.
- **Process capability** is then determined using the specification limits and the process limits.
- The process is capable if the process limits lie inside the specification limits. Then most measurements meet specifications: LSL ≤ LPL ≤ X ≤ UPL ≤ USL. with high probability.
- Summary: A capable in-control process will produce items which usually meet specifications.

**Analysis of Patterns on Control Charts**

- Even when all points plot between the control limits, the process may still be out of control.
- **Symptoms:**
  1. Asymmetric distribution of points - i.e. a much larger number lie on one side of the centerline than the other.
  2. Runs - sequences of observations of the same type.
  e.g. run of increasing (decreasing) observations
- **Cyclic or periodic behavior**
  Removing the cause of this kind of systematic variation can reduce process variability.

**Sensitizing Rules**
Using some or all of the following rules increases the sensitivity of the control chart to process behavior which may be out of control.

1. One or more points outside the (3-sigma) control limits.
2. A run of at least 8 points on one side of the centerline.
3. Two of three consecutive points outside the 2-sigma warning limits.
4. Four of five consecutive points beyond the 1-sigma limits.
5. An unusual or nonrandom pattern in the data (e.g. periodicity).
6. Six successive points that increase (or decrease).

**Complications Arising from Sensitizing Rules**

1. Adding too many sensitizing rules can obscure the simplicity of a control chart.
2. The significance level \( \alpha \) increases as the number of rules increases. i.e. the risk of false alarms increases.

   **With enough rules, almost any sequence of observations may be judged out of control.**