Lecture 11
Statistical process control

Quality control
- Goal is to ensure that processes function as planned
  - basic questions are: "is process controlled at which phase?", "how often?" and "in which way?"

How is control part of managing quality?
- Managing quality
  - Designing quality
    - Product design
    - Process design
    - Sourcing and purchasing
  - Quality control
    - Production
    - Inspection/packaging
    - Delivery
    - Field organization

Quality built into the process
“Modern”
Repairing during process
“Traditional”
Repairing before and after production
Many methods in quality control

Pareto-analysis  Process description
Check lists         Control charts
Graphs               Correlation diagrams
Benchmark  Cause-effect diagram
Acceptance sampling

1. Identifying of problem
   - customer complaints or control charts as starting point
2. Data gathering
   - check lists, graphs, histograms as tools
3. Data analysis
   - e.g. pareto-analysis as a good place to begin
4. Finding the causes of problems
   - e.g. cause-effect analysis
5. Developing and implementing solution
6. Continuous control and improvement of processes

How to use different tools?

Quality control tools
- process control charts -

Statistical process control (SPC)
- valuation is based on samples
Process is under control when there’s only random variation (there’s always some variation)
- for non-random variation there is usually an assignable cause that should be eliminated
- SPC doesn’t reveal the reason, that management and employees job!
- random variation can be reduced only by redesigning the process
Variation is random when sample values are inside the so called control limits
- limits usually +/- three standard deviation which means 99.74% trust
- if values are outside the limits, the process is probably not under control
One of the most difficult decisions for management is to decide if the process needs to be changed
- task is to reduce both random and non-random variation

3 standard deviations as a base for analysis

- Most of the values around the mean
- 99.74% of the values within ±3 standard deviation from the mean
Random and assignable variation

Every process has some amount of random variation

![Distribution shifted](image1)
![Distribution increased](image2)
![Distribution skewed](image3)

Three main types of assignable variation

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}
\]
\[
\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}
\]

Using process control charts

Sign that there might be something wrong with the process (highly unlikely that the sample mean would be outside control limits if the process is in control)

\[
x \pm 3\sigma
\]

Upper control limit

Hypothetical process mean \( x \)

99.74%

One sample mean

Random variation = process under control

Using three standard deviation is recommended but it is possible to use others if needed

- one sample result is outside the control limits
- two following sample results are near the same control limit
- 5 following sample results are at the same side of average
- 5 following sample results are forming a trend either up or down
- level of sample result is changing a lot very rapidly
- other non-random behavior of sample results

Using process control charts

Process is not under control when...

- one sample result is outside the control limits
- two following sample results are near the same control limit
- 5 following sample results are at the same side of average
- 5 following sample results are forming a trend either up or down
- level of sample result is changing a lot very rapidly
- other non-random behavior of sample results

Different characteristics are analyzed with different charts

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Variables</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Durability</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Surface appearance</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Packaging</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Teapot quality characteristics
Measuring continuous variables

- **X-chart**
  - is used to analyze the development of averages of continuous variables
  - the average ("middle line") of X-chart is calculated from the sample averages because the real average of process is usually not known
  - processes real standard deviation is usually not known, so x-charts control limits are calculated with samples’ average range
    - range (R); difference on largest and smallest value in the sample
      - the size of the sample is considered in control charts, relatively small samples are preferred to limit time delays
    - if deviation is known, may be calculated by using the basic formula;

- **R-chart**
  - is used to analyze the internal variability of samples
  - the real standard deviation is usually not known and therefore range R (min-max) is used instead
    - the results are almost the same

\[ \frac{\sigma_X}{\sigma} = \frac{1}{\sqrt{n}} \]

**X- and R-chart example**

Finnish Washer Ltd. produces metal sheets, that are used as components for different kinds of machines. The critical measure is metal sheet hole’s diameter, so that the component can be used for its designed purpose. Quality control has taken 10 samples during 10 days, each consisting of 5 sheets (measurements below). Find out with the help of SPC if the process is under control. Create control charts of the machine operation. Give short argument for your answer.

<table>
<thead>
<tr>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5.02</td>
<td>5.01</td>
<td>4.94</td>
<td>4.99</td>
<td>4.96</td>
</tr>
<tr>
<td>B</td>
<td>5.01</td>
<td>5.03</td>
<td>5.07</td>
<td>4.95</td>
<td>4.96</td>
</tr>
<tr>
<td>C</td>
<td>4.99</td>
<td>5.00</td>
<td>4.93</td>
<td>4.92</td>
<td>4.99</td>
</tr>
<tr>
<td>D</td>
<td>5.03</td>
<td>4.91</td>
<td>5.01</td>
<td>4.98</td>
<td>4.89</td>
</tr>
<tr>
<td>E</td>
<td>4.95</td>
<td>4.92</td>
<td>5.93</td>
<td>5.05</td>
<td>5.01</td>
</tr>
<tr>
<td>F</td>
<td>4.97</td>
<td>5.06</td>
<td>5.06</td>
<td>4.96</td>
<td>5.03</td>
</tr>
<tr>
<td>G</td>
<td>5.05</td>
<td>5.06</td>
<td>5.10</td>
<td>4.96</td>
<td>4.99</td>
</tr>
<tr>
<td>H</td>
<td>5.09</td>
<td>5.01</td>
<td>5.00</td>
<td>4.99</td>
<td>5.08</td>
</tr>
<tr>
<td>I</td>
<td>5.14</td>
<td>5.10</td>
<td>4.99</td>
<td>5.08</td>
<td>5.09</td>
</tr>
<tr>
<td>J</td>
<td>5.01</td>
<td>4.98</td>
<td>5.08</td>
<td>5.07</td>
<td>4.99</td>
</tr>
</tbody>
</table>

**Average** 5.009 0.115

**X- and R-chart example**

1. Calculate sample mean and range as well as average mean and range

<table>
<thead>
<tr>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>X</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5.02</td>
<td>5.01</td>
<td>4.94</td>
<td>4.99</td>
<td>4.96</td>
<td>4.98</td>
<td>0.08</td>
</tr>
<tr>
<td>B</td>
<td>5.01</td>
<td>5.03</td>
<td>5.07</td>
<td>4.95</td>
<td>4.96</td>
<td>5.00</td>
<td>0.12</td>
</tr>
<tr>
<td>C</td>
<td>4.99</td>
<td>5.00</td>
<td>4.93</td>
<td>4.92</td>
<td>4.99</td>
<td>5.01</td>
<td>0.13</td>
</tr>
<tr>
<td>D</td>
<td>5.03</td>
<td>4.91</td>
<td>5.01</td>
<td>4.98</td>
<td>4.89</td>
<td>5.05</td>
<td>0.11</td>
</tr>
<tr>
<td>E</td>
<td>4.95</td>
<td>4.92</td>
<td>5.03</td>
<td>5.05</td>
<td>4.99</td>
<td>5.07</td>
<td>0.10</td>
</tr>
<tr>
<td>F</td>
<td>4.97</td>
<td>5.06</td>
<td>5.06</td>
<td>4.96</td>
<td>5.03</td>
<td>5.01</td>
<td>0.14</td>
</tr>
<tr>
<td>G</td>
<td>5.05</td>
<td>5.06</td>
<td>5.10</td>
<td>4.96</td>
<td>4.99</td>
<td>5.00</td>
<td>0.15</td>
</tr>
<tr>
<td>H</td>
<td>5.09</td>
<td>5.01</td>
<td>5.00</td>
<td>4.99</td>
<td>5.08</td>
<td>5.07</td>
<td>0.10</td>
</tr>
<tr>
<td>I</td>
<td>5.14</td>
<td>5.10</td>
<td>4.99</td>
<td>5.08</td>
<td>5.09</td>
<td>5.03</td>
<td>0.11</td>
</tr>
<tr>
<td>J</td>
<td>5.01</td>
<td>4.98</td>
<td>5.08</td>
<td>5.07</td>
<td>4.99</td>
<td>5.03</td>
<td>0.115</td>
</tr>
</tbody>
</table>

**Average** 5.009 0.115
X- and R-chart example

2. Calculate control limits for X- and R-charts

Values for X- and R-charts

<table>
<thead>
<tr>
<th>n</th>
<th>A2</th>
<th>D3</th>
<th>D4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.880</td>
<td>0</td>
<td>3.268</td>
</tr>
<tr>
<td>3</td>
<td>1.023</td>
<td>0</td>
<td>2.574</td>
</tr>
<tr>
<td>4</td>
<td>0.729</td>
<td>0</td>
<td>2.282</td>
</tr>
<tr>
<td>5</td>
<td>0.577</td>
<td>0</td>
<td>2.114</td>
</tr>
<tr>
<td>6</td>
<td>0.483</td>
<td>0</td>
<td>2.004</td>
</tr>
<tr>
<td>7</td>
<td>0.419</td>
<td>0.076</td>
<td>1.924</td>
</tr>
<tr>
<td>8</td>
<td>0.373</td>
<td>0.136</td>
<td>1.864</td>
</tr>
<tr>
<td>9</td>
<td>0.337</td>
<td>0.184</td>
<td>1.816</td>
</tr>
<tr>
<td>10</td>
<td>0.308</td>
<td>0.223</td>
<td>1.777</td>
</tr>
</tbody>
</table>

When one has to use samples’ average range \( R \) as standard deviation every control limit has its own formula that has to be memorized.

When one has to use samples’ average range \( R \) as standard deviation one needs to use tabular values when calculating control limits.

3. Plot sample values and control limits

4. Interpret results and make conclusions/recommendations

- Mean not in control; one over control limit, rising trend etc.
- Variability in control

Analyzing process charts

<table>
<thead>
<tr>
<th>Graph</th>
<th>Description</th>
<th>Possible causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Random variation</td>
<td></td>
</tr>
<tr>
<td>Unstable</td>
<td>Assignable causes e.g. tools, materials, people, overreacting etc.</td>
<td>E.g. wearing of machine, employee getting tired, better working methods</td>
</tr>
<tr>
<td>Trend</td>
<td>E.g. wearing of machine, employee getting tired, better working methods</td>
<td></td>
</tr>
<tr>
<td>Cycle</td>
<td>Different working shifts, variation of electricity, periodical changes</td>
<td></td>
</tr>
</tbody>
</table>

Measuring attributes

- **p-chart** (quantity of faulty ones within the sample)
  - all the variables are not continuous. P-chart is used when observations can be divided into two groups
    - functioning vs. not functioning, good vs. bad, acceptable vs. not acceptable
    - quantity is usually given as a percentage
    \[ \sigma_p = \sqrt{\frac{p(1-p)}{n}} \]

- **c-chart** (quantity of failures per unit)
  - used when only observations can be counted ("not observations" can’t be counted)
    - phone calls, complaints, breakdowns etc. per time period
    - scratches or dents per product etc.
    - cannot be given as a percentage
    \[ \sigma_c = \sqrt{c} \]
Some activists have complained to the city council that all the city residents should have an equal right to feel safe. They think that number of policemen and investments preventing crime (e.g. lights at streets) should be made on relative basis so that so called problem areas should have more attention than safe areas. To investigate the complaint, the city council acquired data of crimes committed during the last 30 days (the city had been divided into twenty 5000 resident areas). In each area the sample size was 1000 persons. What kind of advice would you give to city council for using the information to allocate the recursos? Use 95% level ($z=1.96$) so that your recommendation would be correct with high probability. (Base your analysis on quality control methods)

### Sample Data

<table>
<thead>
<tr>
<th>Area</th>
<th>Number of crimes</th>
<th>Area</th>
<th>Number of crimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>16</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

### Calculations

1. Calculate sample probability $\hat{p}$
2. Calculate mean probability $\bar{p}$
3. Calculate sample’s standard deviation

$$p = \frac{\text{Total defectives}}{\text{Number of observations}} = \frac{300}{20 \times 1000} = 0.015 \text{ that is (1.5%)}$$

$$\sigma_p = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.015 \times (1 - 0.015)}{1000}} = 0.0038438$$

4. Calculate control limits

$$\text{UCL} = \bar{p} + z \sigma_p = 0.015 + 1.96 \times 0.0038438 = 0.0225339 \text{ about 2.25%}$$

$$\text{LCL} = \bar{p} - z \sigma_p = 0.015 - 1.96 \times 0.0038438 = 0.0074661 \text{ about 0.74%}$$

Typically $z=3$ but in this case $z=1.96$ because the question asked for only “95% confidence level”.

5. Plot sample values and control limits

6. Interpret results and make conclusions/recommendations
   - More investments to areas 6, 16 and 20
   - Reduce investments from areas 2, 10, 17 and 19

How to use different tools?

1. Identifying of problem
   - customer complaints or control charts as starting point
2. Data gathering
   - check lists, graphs, histograms as tools
3. Data analysis
   - e.g. pareto-analysis as a good place to begin
4. Finding the causes of problems
   - e.g. cause-effect analysis
5. Developing and implementing solution
6. Continuous control and improvement of processes
Quality control tools
- correlation diagrams -

- Effective graphical presentation style when there is a clear cause-effect relationship

![Graph showing number of problems vs. training hours]

How to use different tools?

1. Identifying of problem
   - customer complaints or control charts as starting point

2. Data gathering
   - check lists, graphs, histograms as tools

3. Data analysis
   - e.g. pareto-analysis as a good place to begin

4. Finding the causes of problems
   - e.g. cause-effect analysis

5. Developing and implementing solution

6. Continuous control and improvement of processes

Continuous improvement as key to success

- Activities must be improved all the time
  - observing and questioning performance essential

- Employees have central role in developing operational quality
  - management has many ways to get employees committed
    - culture change, customer focus, team work, empowerment, training, awards, incentives

Process improvement also visible in charts

- process
- people
- machines
- material
Process capability

Process capability means ability to maintain certain product / process specifications (often defined by customer)
- Focus is on random variation (common causes), assignable causes presumed to be fixed
- Measured by capability index
  \[
  C_{pk} = \min \left( \frac{\bar{X} - LTL}{3\sigma}, \frac{UTL - \bar{X}}{3\sigma} \right)
  \]
  - Index takes into account if the distribution average shifts
  - 1.33 is considered as a limit of good process (99.9936% good)
- Capability is often improved by improving the process itself

Everything has its own tolerance limits

Tolerance limits are set design specifications that define the acceptable range for a part, product or service

How well should process function?

In general 99% sounds good, but in many cases it’s not
- 3.5 days a year without electricity
- 15 minutes of undrinkable water every day
- 15 minutes without phone and television every day
- 20,000,000 wrong prescriptions a year (USA)
- Over 15,000 new born babies dropped by doctors a year (USA)
- 5,000 failures in medical operations per week (USA)
- 3 failed landings at Heathrow every day
- 20,000 lost letters (by post) every hour (USA)

According to research, basic processes of human beings are on 99.86% (3.2 sigmas) level
Process capability and number of good products

- **6 sigma level**: Tolerance limit is six standard deviation from the mean. 99.999999% good.
- **2 sigma level**: 95.45% good.
- **4 sigma level**: 99.937% good.

Sigma-levels are not linear

<table>
<thead>
<tr>
<th># sigma’s</th>
<th>Area</th>
<th>Spelling</th>
<th>Time</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Floor space of Soldier Field</td>
<td>170 misspelled words per page in a book</td>
<td>31.75 years per century</td>
<td>Here to the moon</td>
</tr>
<tr>
<td>2</td>
<td>Floor space of large supermarket</td>
<td>25 misspelled words per page in a book</td>
<td>0.4 years per century</td>
<td>1.5 times around the world</td>
</tr>
<tr>
<td>3</td>
<td>Floor space of small hardware store</td>
<td>1.5 misspelled words per page in a book</td>
<td>1.5 months per century</td>
<td>London to New York</td>
</tr>
<tr>
<td>4</td>
<td>Your living room</td>
<td>1 misspelled word per 30 pages</td>
<td>2.5 days per century</td>
<td>Basel to Zurich</td>
</tr>
<tr>
<td>5</td>
<td>The button of your telephone</td>
<td>1 misspelled word in a set of encyclopedias</td>
<td>30 minutes per century</td>
<td>Leverson to Norris</td>
</tr>
<tr>
<td>6</td>
<td>Diamond</td>
<td>1 misspelled word in a library</td>
<td>6 seconds per century</td>
<td>Four steps from your chair</td>
</tr>
</tbody>
</table>

Why is 6 sigma so desirable?

- Process capability is increased by improving the process.

Process capability is increased by improving the process

- **$C_{pk}=0.57$**
  - $\sigma=100$
  - 30.54 %
  - $3\sigma$
  - LTL 4.3
  - UTL 1.7

- **$C_{pk}=1.00$**
  - $\sigma=100$
  - 99.74 %
  - $3\sigma$
  - LTL 3.0
  - UTL 3.0

- **$C_{pk}=1.13$**
  - $\sigma=50$
  - 99.37 %
  - $3\sigma$
  - LTL 7.9
  - UTL 3.4

- **$C_{pk}=2.00$**
  - $\sigma=50$
  - 99.999 %
  - $3\sigma$
  - LTL 9.0
  - UTL 6.0
Sigma levels and capability have a relationship

\[ C_{pk} = \min \left( \frac{\bar{X} - LTL}{3\sigma}, \frac{UTL - \bar{X}}{3\sigma} \right) \]

Six sigma as a mgmt. philosophy

- Statistic oriented improvement concept that looks more and more like TQM
  - employees improve processes, focus on customer, decisions based on facts, quality controlled with statistical methods etc.
- Emphasis originally more on eliminating errors
  - "zero defects", "fix in the first time", higher yield, lower costs...
- Required process capability is 1,5
  - based on "typical" 1,5 std. movement of mean in the process
  - number of defective 3,4 for every million produced
- Limited number of wide success stories

<table>
<thead>
<tr>
<th>Processes sigma level (σ)</th>
<th>Share of defective</th>
<th>Amount of defective</th>
<th>Capability for centered process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.7310519573 %</td>
<td>317 per thousand</td>
<td>0.33</td>
</tr>
<tr>
<td>2</td>
<td>4.5500124000 %</td>
<td>45 per thousand</td>
<td>0.67</td>
</tr>
<tr>
<td>3</td>
<td>0.2699994440 %</td>
<td>2 per thousand</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>0.0083372000 %</td>
<td>63 per million</td>
<td>1.33</td>
</tr>
<tr>
<td>5</td>
<td>0.0000574000 %</td>
<td>574 per billion</td>
<td>1.67</td>
</tr>
<tr>
<td>6</td>
<td>0.0000001980 %</td>
<td>2 per billion</td>
<td>2.00</td>
</tr>
<tr>
<td>7</td>
<td>0.0000000003 %</td>
<td>3 per trillion</td>
<td>2.33</td>
</tr>
</tbody>
</table>

Lecture 11
Statistical process control

Acceptance sampling

Quality control tools - acceptance sampling -

- "Traditional quality control method"
  - batch's quality is ensured by inspecting a sample of the batch
- Batch is accepted if the sample has a low enough amount of defectives
  - rejected batches are repaired or obliterated
- Method has many disadvantages
  - expensive from total life-time cost perspective
  - assumes that a certain amount of defectives is acceptable
  - sampling method may lead to wrong decisions (limited information)
- Used in some companies
  - easy process, low direct costs (cheaper than investigating all items), only method for products that "perish/break" in inspection
  - need to decide sample size and acceptable quality level
  - motivates manufacturer to produce good quality