Process Capability Statistics for Non-Normal Distributions in R

Thomas Roth

The Department of Quality Science - Technical University of Berlin

August 18, 2011

Outline

A (Very) Short Introduction To Quality Sciences (ISO 9001:2008)

quality

degree to which a set of inherent characteristics fulfils requirements

A (Very) Short Introduction To Quality Sciences (ISO 9001:2008)

quality

degree to which a set of inherent characteristics fulfils requirements

management

coordinated activities to direct and control

A (Very) Short Introduction To Quality Sciences (ISO 9001:2008)

quality

degree to which a set of inherent characteristics fulfils requirements

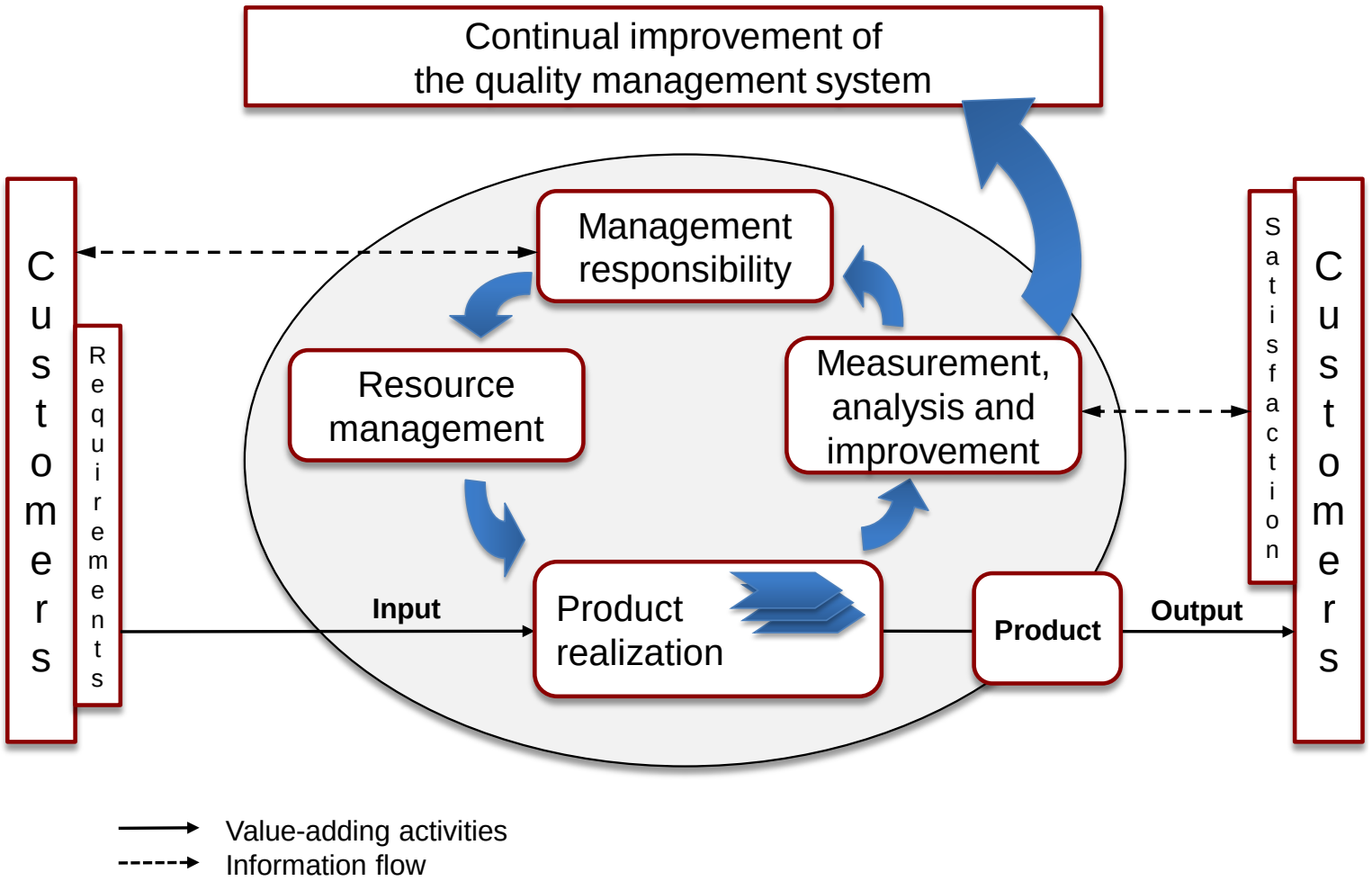
management

coordinated activities to direct and control

quality management system

to direct and control an organization with regard to quality

Process-based Quality Management System For Continual Improvement (ISO 9001:2008)



Implications

Certifications in 2009

Up to the end of December 2009, at least 1 064 785 ISO 9001 (2000 and 2008) certificates had been issued in 178 countries and economies. (ISO Survey 2009)

Implications

Certifications in 2009

Up to the end of December 2009, at least 1 064 785 ISO 9001 (2000 and 2008) certificates had been issued in 178 countries and economies. (ISO Survey 2009)

8.2.4

The organization shall [...] measure the characteristics of the product to verify that product requirements have been met. [...] carried out at appropriate stages of the product realization process (ISO 9001:2008)

Capability

ISO 25514-7

Statistical methods in process management
Capability and performance
Part 4:
Process capability estimates and performance measures

Process Capability

ability of the process to realize a characteristic that will fullfil the requirements for that characteristic (ISO 25517-4) (i.e. statistical measure of inherent process variability)

Process Capability Index For Normal Distribution

histogram

plot a histogram of the data

distribution model

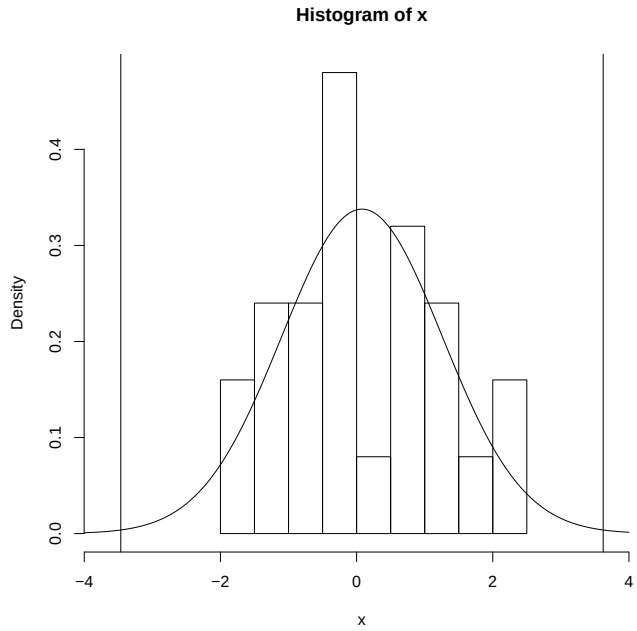
select an appropriate distribution

specification limits

identify the specification limits

process capability index

index describing process capability in relation to specified tolerance



$$c_p = \frac{USL - LSL}{Q_{0.99865} - Q_{0.00135}} = \frac{USL - LSL}{6s_p}$$

Non-Normal Distribution

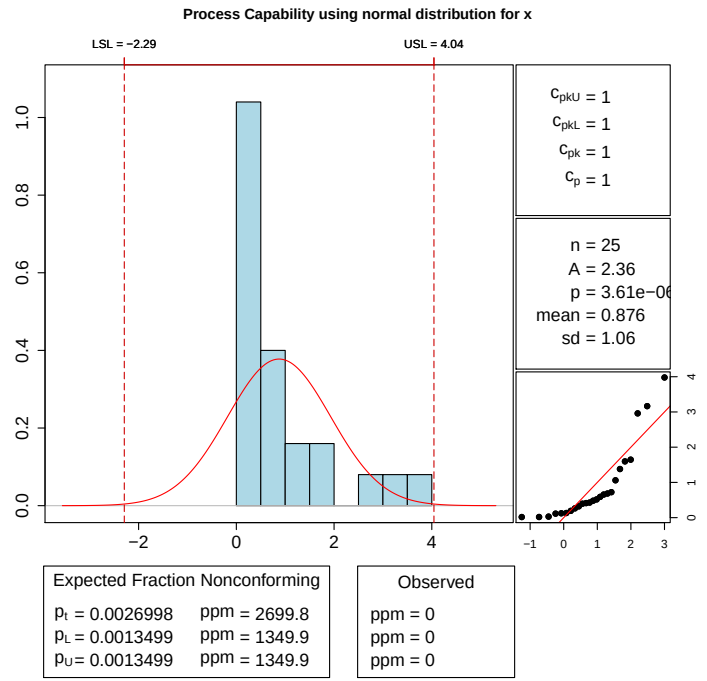
distribution identification

- Anderson Darling Test (ISO 25517-4, p.14)
- Probability Plots (ISO 25517-4, p.16)

$$c_{pkL} = \frac{Q_{0.5} - LSL}{Q_{0.5} - Q_{0.00135}}$$

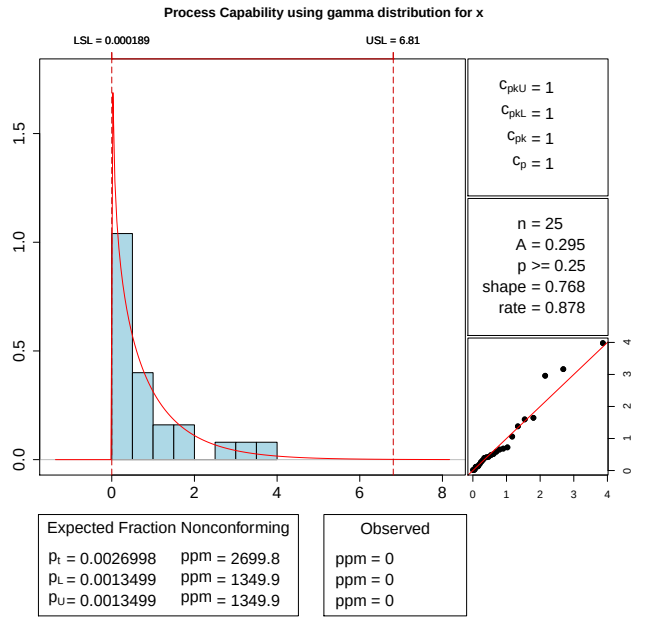
$$c_{pkU} = \frac{USL - Q_{0.5}}{Q_{0.99865} - Q_{0.5}}$$

```
> x = rexp(25)
> cp(x)

        Anderson Darling Test for normal distribution

data:  x
A = 1.8384, mean = 1.016, sd = 1.154, p-value = 7.58e-05
alternative hypothesis: true distribution is not equal to normal
```

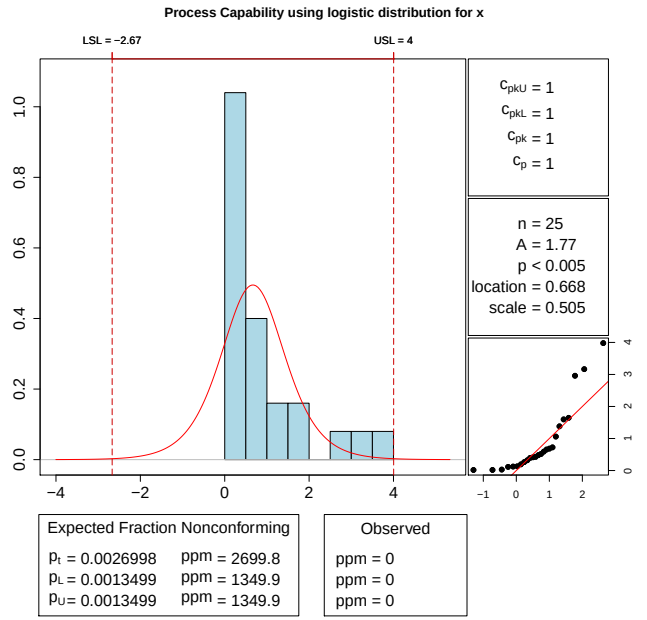


Non-Normal Distribution | Two-Sided

```
> cp(x, "gamma")

        Anderson Darling Test for gamma distribution

data:  x
A = 0.2947, shape = 0.768, rate = 0.878, p-value > 0.25
alternative hypothesis: true distribution is not equal to gamma
```

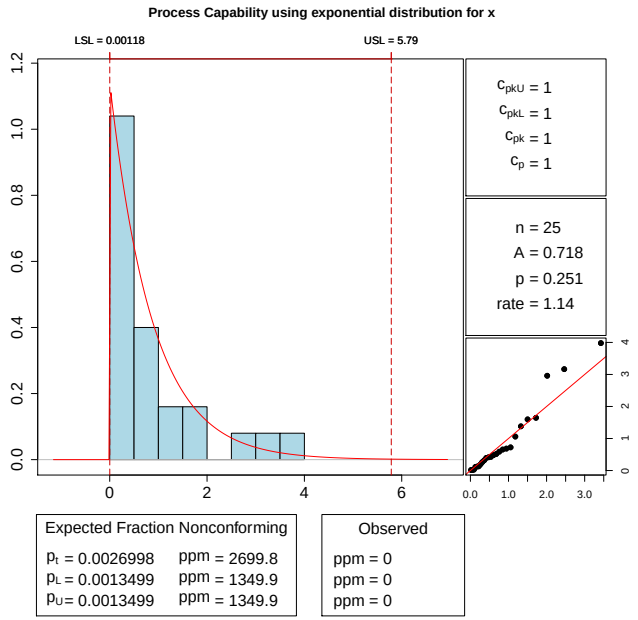



```
> cp(x, "logistic")

        Anderson Darling Test for logistic distribution

data:  x
A = 1.7745, location = 0.668, scale = 0.505, p-value <= 0.005
alternative hypothesis: true distribution is not equal to logisti
```

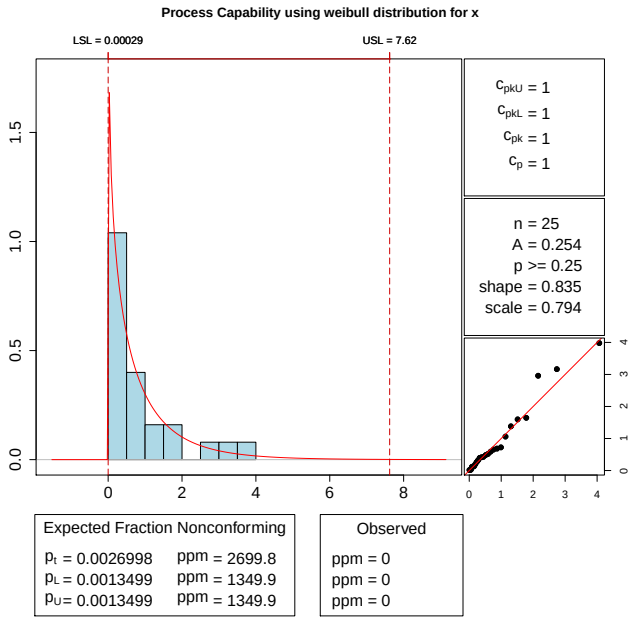


Non-Normal Distribution | Two-Sided

```
> cp(x, "exponential")

        Anderson Darling Test for exponential distribution

data:  x
A = 0.7179, rate = 1.142, p-value = 0.2511
alternative hypothesis: true distribution is not equal to exponential
```

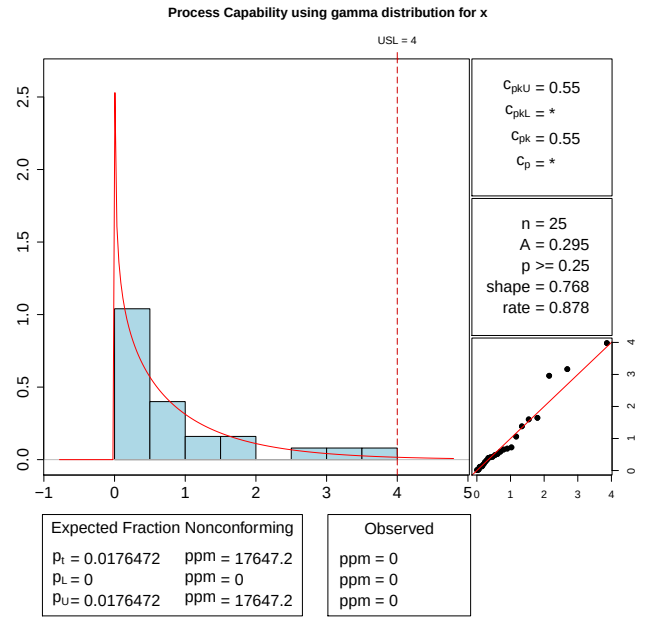



```
> cp(x, "weibull")

        Anderson Darling Test for weibull distribution

data:  x
A = 0.2539, shape = 0.835, scale = 0.794, p-value > 0.25
alternative hypothesis: true distribution is not equal to weibull
```

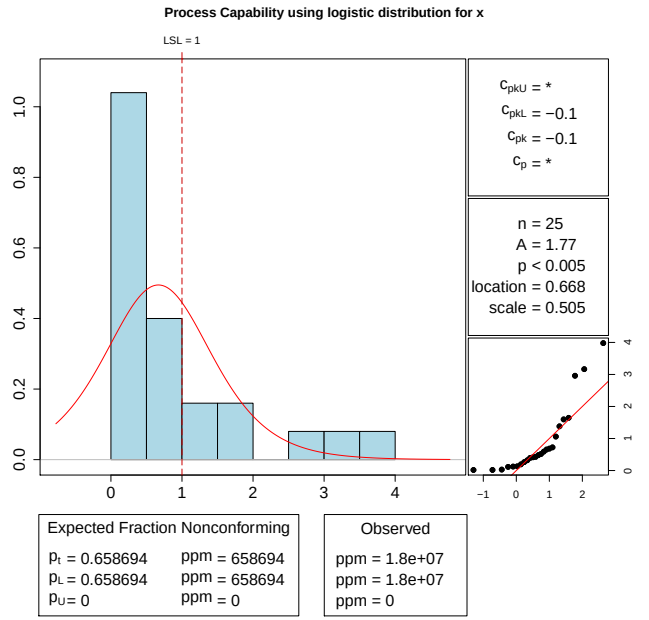


Non-Normal Distribution | One-Sided

```
> cp(x, "gamma", usl = 4)

        Anderson Darling Test for gamma distribution

data:  x
A = 0.2947, shape = 0.768, rate = 0.878, p-value > 0.25
alternative hypothesis: true distribution is not equal to gamma
```

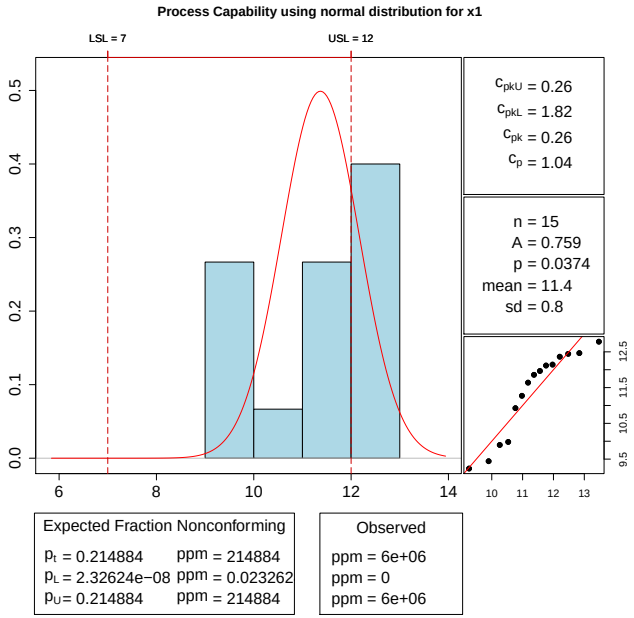



```
> cp(x, "logistic", lsl = 1)

        Anderson Darling Test for logistic distribution

data:  x
A = 1.7745, location = 0.668, scale = 0.505, p-value <= 0.005
alternative hypothesis: true distribution is not equal to logisti
```

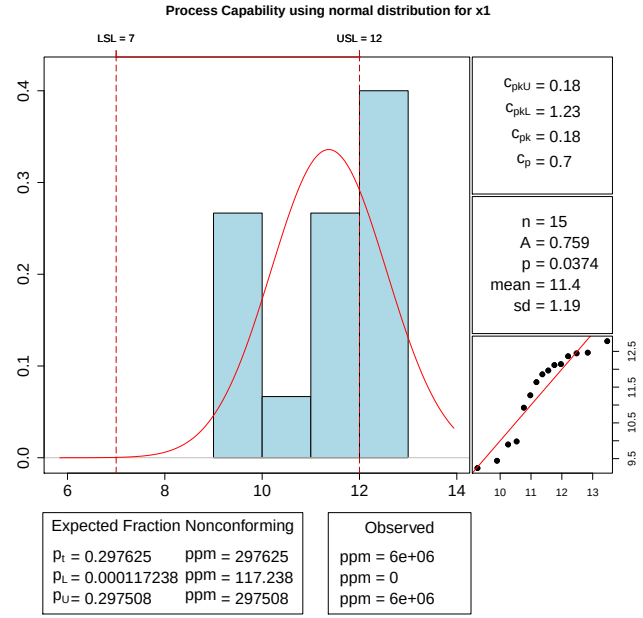


Data in subgroups

```
> x1 = c(rnorm(5, 11), rnorm(5, 12), rnorm(5, 10))
> group = c(rep(1, 5), rep(2, 5), rep(3, 5))
> cp(x1, grouping = group, lsl = 7, usl = 12)

        Anderson Darling Test for normal distribution

data:  x1
A = 0.5594, mean = 11.598, sd = 0.827, p-value = 0.1219
alternative hypothesis: true distribution is not equal to normal
```




```
> cp(x1, lsl = 7, usl = 12)

        Anderson Darling Test for normal distribution

data:  x1
A = 0.7587, mean = 11.369, sd = 1.188, p-value = 0.03743
alternative hypothesis: true distribution is not equal to normal
```



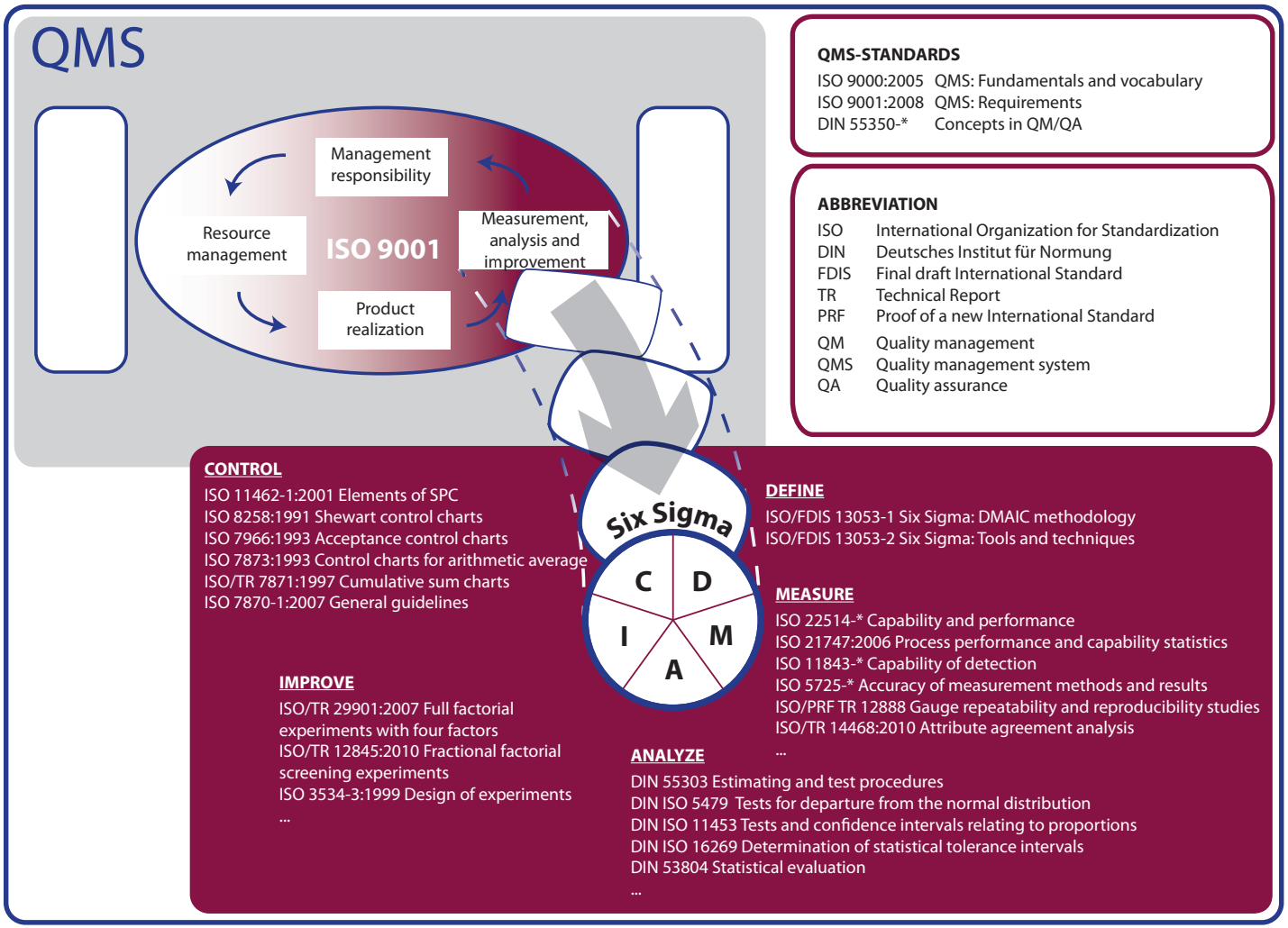
Process Capability Indices

process capability index

index describing process capability in relation to specified tolerance

- represents the capability of a process in a single number
- is used to assess and communicate the capability of internal/external suppliers
- is used to calculate the upper, lower and total fraction non-conforming (2.12, 2.13, 2.14)

Process Capability and Six Sigma



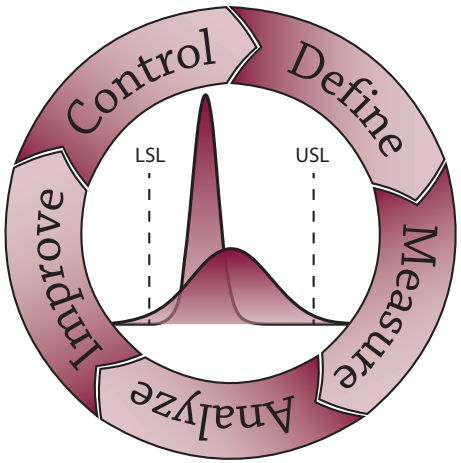
Summary

What is still missing

- direct calculation of p-Values
- calculation of confidence intervals
- three parameter weibull, loglogistic and gamma distribution

The process capability function is part of the comprehensive Six Sigma qualityTools package which contains easy to use methods associated with the different phases of the DMAIC Cycle.

`http://www.r-sixsigma.org`



References

- ISO/TR 22514-4 (2007) Statistical methods in process management - capability and performance - part 4: process capability estimates and performance, Norm.
- ISO 9001 (2008) Quality management systems - Requirements, Norm.
- The ISO-Survey of certifications 2009, 18th ed
- R. B. D'Agostino and M. A. Stephens (1986) Goodness-Of-fit techniques, M. Dekker, New York.
- Roth, T.; Herrmann Joachim: Teaching Statistics in Quality Science using the R-Package qualityTools. `http://user2010.org/slides/Roth+Herrmann.pdf` (useR! 2010 Conference)
- Roth, T. (2010). qualityTools: Statistical Methods for Quality Science. `http://www.r-qualitytools.org`