CMM MCA Report: Mathematical Calculations

May 19, 2025

1 Introduction

Coordinate Measuring Machines (CMMs) measure the geometrical characteristics of manufactured parts to ensure compliance with design specifications. The Measurement Capability Analysis (MCA) report quantifies a process's ability to meet tolerance requirements using statistical methods. This document details the mathematical calculations involved in MCA, focusing on process capability indices, distribution fitting, statistical metrics, and visualization, with an emphasis on their theoretical foundations in statistical process control (SPC).

2 Process Capability Indices

Process capability indices evaluate whether a manufacturing process can consistently produce parts within specified tolerance limits, defined by the Upper Specification Limit (USL) and Lower Specification Limit (LSL).

2.1 Process Capability Index (C_p)

The C_p index measures the potential capability of a process by comparing the specification width to the process variability:

$$C_p = \frac{\text{USL} - \text{LSL}}{6\sigma} \tag{1}$$

where:

- USL: Upper Specification Limit
- LSL: Lower Specification Limit
- σ: Process standard deviation, estimated from sample data

The denominator 6σ represents the spread covering 99.73% of a normally distributed process ($\pm 3\sigma$ from the mean). A $C_p \geq$ cp_target indicates a process capable of meeting specifications, assuming the process is centered at the target value.

2.2 Process Capability Index (C_{pk})

The C_{pk} index adjusts for the process mean's deviation from the target, providing a realistic measure of capability:

$$C_{pk} = \min\left(\frac{\bar{x} - \text{LSL}}{3\sigma}, \frac{\text{USL} - \bar{x}}{3\sigma}\right) \tag{2}$$

where $\bar{x}$ is the sample mean. The C_{pk} reflects the process's ability to produce within specifications, accounting for both variability and centering. A $C_{pk} \geq$ cpk_target is typically required for a process to be considered capable.

2.3 Confidence Intervals for C_p and C_{pk}

Since C_p and C_{pk} are estimated from sample data, confidence intervals quantify the uncertainty in these estimates.

2.3.1 C_p Confidence Interval

The confidence interval for C_p is based on the chi-squared distribution, as the sample variance s^2 follows a scaled chi-squared distribution:

$$C_p\sqrt{\frac{\chi^2_{1-\alpha/2,n-1}}{n-1}} \leq C_p \leq C_p\sqrt{\frac{\chi^2_{\alpha/2,n-1}}{n-1}} \tag{3}$$

where:

- $\chi^2_{1-\alpha/2,n-1}$, $\chi^2_{\alpha/2,n-1}$: Chi-squared critical values for a confidence level of $1-\alpha$ (typically $\alpha = 0.05$)
- n: Sample size

This interval accounts for the variability in estimating σ from the sample standard deviation.

2.3.2 C_{pk} Confidence Interval

The confidence interval for C_{pk} is approximated using the standard error of the C_{pk} estimator:

$$\text{SE}_{C_{pk}} = \sqrt{\frac{1}{n} + \frac{C_{pk}^2}{2(n-1)}} \tag{4}$$

The confidence interval is:

$$C_{pk} \pm z_{1-\alpha/2} \cdot \text{SE}_{C_{pk}} \tag{5}$$

where $z_{1-\alpha/2}$ is the critical value from the standard normal distribution (e.g., 1.96 for $\alpha = 0.05$). This approximation assumes normality and a sufficiently large sample size.

3 Distribution Fitting

The calculations for C_p and C_{pk} assume the process data follows a normal distribution. To validate this or identify a better-fitting distribution, statistical tests and model selection criteria are applied.

3.1 Anderson-Darling Test for Normality

The Anderson-Darling test assesses whether the data is normally distributed by comparing the empirical distribution to the normal distribution:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln \Phi(y_i) + \ln(1-\Phi(y_{n+1-i}))\right] \tag{6}$$

where:

- y_i: Ordered observations transformed to the standard normal scale
- Φ: Standard normal cumulative distribution function
- n: Sample size

The test produces a p-value, where $p \geq 0.05$ suggests the data is consistent with normality. If normality is rejected, alternative distributions are evaluated.

3.2 AIC-Based Distribution Selection

When normality is not supported, multiple distributions (e.g., Normal, Log-Normal, Exponential, Gamma, Weibull, Rayleigh, Beta) are fitted to the data. The best distribution is selected using the Akaike Information Criterion (AIC):

$$\text{AIC} = 2k - 2\ln(L) \tag{7}$$

where:

- k: Number of parameters in the distribution model
- L: Maximum likelihood of the fitted distribution

The distribution with the lowest AIC is chosen, balancing goodness of fit and model complexity. For distributions requiring positive data (e.g., Log-Normal, Gamma), a small constant is added to shift non-positive measurements, ensuring valid parameter estimation.

4 Statistical Metrics

A comprehensive set of statistical metrics is calculated to describe the measurement data and assess process performance:

- **Sample Size (n)**: The number of valid measurements.

- **Mean ($\bar{x}$)**: The arithmetic average, calculated as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{8}$$

- **Median**: The 50th percentile, robust to outliers.
- **Standard Deviation (σ)**: The sample standard deviation, adjusted for $n - 1$ degrees of freedom:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{9}$$

- **Minimum, Maximum, Range**: The smallest and largest measurements and their difference.
- **Percentiles**: The 0.135th, 50th, and 99.865th percentiles, corresponding to $\pm 3\sigma$ bounds in a normal distribution.
- **Out-of-Specification Counts and Proportions**: The number and percentage of measurements below LSL or above USL:
  - $n_{\text{below}} = \sum_{i=1}^{n} \nVdash(x_i < \text{LSL})$
  - $n_{\text{above}} = \sum_{i=1}^{n} \nVdash(x_i > \text{USL})$
  - $p_{\text{below}} = \frac{n_{\text{below}}}{n} \times 100$
  - $p_{\text{above}} = \frac{n_{\text{above}}}{n} \times 100$
- **Within-Specification Proportion**: The percentage of measurements within LSL and USL:

$$p_{\text{within}} = \frac{\sum_{i=1}^{n} \nVdash(\text{LSL} \leq x_i \leq \text{USL})}{n} \times 100 \tag{10}$$

- **Process Capability Metrics**: C_p, C_{pk}, and their confidence intervals.
- **Distribution Metrics**: The identified distribution and its p-value from the Anderson-Darling test.
- **Capability Requirements**: Boolean checks for $C_p \geq$ cp_target and $C_{pk} \geq$ cpk_target.

These metrics provide a holistic view of process variability, centering, and compliance with specifications.

5 Visualization of Measurements

A scatter plot visualizes the measurement data to identify trends and specification compliance.

5.1 Scatter Plot Structure

The plot displays measurements against their sequence number (1 to n), with the following elements:

- **Points**: Each measurement x_i is plotted at position (i, x_i), colored based on its relation to statistical and specification limits:
  - Red: Measurements below $\bar{x} - 3\sigma$ or above $\bar{x} + 3\sigma$, but within LSL and USL
  - Green: Measurements within $\bar{x} \pm 3\sigma$
  - Blue: Measurements between $\bar{x} - 3\sigma$ and LSL or between $\bar{x} + 3\sigma$ and USL
  - Purple: Measurements outside specification limits (below LSL or above USL)
- **Horizontal Lines**: Reference lines at:
  - LSL and USL: Specification boundaries
  - $\bar{x}$: Process mean
  - $\bar{x} \pm 3\sigma$: Expected process bounds under normality
- **Axes**: The x-axis represents the measurement sequence, and the y-axis represents measurement values, scaled to include all measurements and specification limits with padding.

5.2 Mathematical Basis

The coloring of points is based on statistical process control principles, where $\bar{x} \pm 3\sigma$ defines the expected range for 99.73% of measurements under normality. Measurements outside LSL or USL indicate non-conformance, while those outside $\bar{x} \pm 3\sigma$ but within specifications suggest potential process instability.

6 Metadata and Parameter Calculations

Metadata and parameter-specific calculations provide context and detailed analysis for each measured characteristic.

6.1 Metadata

Metadata includes descriptive information such as inspection agency, component details, and order quantities. These are extracted from predefined locations in the data source and used for traceability and reporting.

6.2 Parameter-Specific Calculations

For each measured parameter, the following are calculated:

- **Characteristic Description**: A unique identifier for the parameter.
- **Specification Limits**: USL, LSL, and nominal value (midpoint of the tolerance range).

- **Tolerance**: The specification width, USL − LSL.
- **Measurement Count**: The number of valid measurements.
- **Temporal Bounds**: The earliest and latest measurement times, if available.
- **Statistical Metrics**: As described in Section 4.
- **Process Capability**: C_p, C_{pk}, and their confidence intervals.
- **Distribution Fit**: The best-fitting distribution and its p-value.
- **Visualization**: A scatter plot specific to the parameters measurements.

The tolerance calculation ensures the process capability indices are contextualized relative to the allowable variation.

7 Conclusion

The mathematical calculations for CMM MCA reports provide a robust framework for assessing manufacturing process capability. Process capability indices (C_p, C_{pk}) quantify the processs ability to meet specifications, while distribution fitting validates distributional assumptions. Statistical metrics and visualizations offer insights into variability, centering, and compliance. Grounded in statistical process control and probability theory, these calculations enable quality engineers to ensure product reliability and process stability.