

# Statistics Definitions

This provides a definition of terms used in QW6 as well as information on the calculations and rules employed in developing the statistics that appear within QW6.

Only entered values (non-null) are included in calculations in QW6.

## How to use this table

1. Match the combination of **fixed** limits you have with the Fixed Limits column.
2. Go right to the column of the statistic you want to calculate.

	<b>O % OOL</b>	<b>C % OOL</b>	<b>Cr</b>	<b>Tz</b>	<b>Cpk</b> (replaced by Ppk)	<b>Ppk</b> (replaces Cpk calculations)
<b>Fixed USL</b> <b>Fixed Target</b> <b>Fixed LSL</b>	Use both Limits Out Upper Out Lower	Use both Limits Z Upper Z Lower	$6S / (USL - LSL)$	$(AVG - TGT) / S$	Lesser of $(USL - AVG) / (3*SD MR)$ and $(AVG - LSL) / (3*SD MR)$	Lesser of $(USL - AVG) / (3*SD Pop)$ and $(AVG - LSL) / (3*SD Pop)$
<b>Fixed USL</b> <b>Fixed Target</b> <b>Fixed LSL</b>	Use USL Only	Use USL Only Z Upper	$3S / (USL - TGT)$	$(AVG - TGT) / S$	$(USL - AVG) / (3*SD MR)$	$(USL - AVG) / (3*SD Pop)$
<b>Fixed Target</b> <b>Fixed LSL</b>	Use LSL Only	Use USL Only Z Lower	$3S / (TGT - LSL)$	$(AVG - TGT) / S$	$(AVG - LSL) / (3*SD MR)$	$(AVG - LSL) / (3*SD Pop)$
<b>Fixed Target</b>	N/A	N/A	N/A	$(AVG - TGT) / S$	N/A	N/A
<b>Fixed USL</b>	Use USL Only	Use USL Only Z Upper	$3S / (USL - AVG)$	N/A	$(USL - AVG) / (3*SD MR)$	$(USL - AVG) / (3*SD Pop)$
<b>Fixed LSL</b>	Use LSL Only	Use USL Only Z Lower	$3S / (AVG - LSL)$	N/A	$(AVG - LSL) / (3*SD MR)$	$(AVG - LSL) / (3*SD Pop)$
<b>Fixed USL</b> <b>Fixed LSL</b>	Use USL & LSL	Use both Limits Z Upper Z Lower	$6S / (USL - LSL)$	N/A	Lesser of $(USL - AVG) / (3*SD MR)$ and $(AVG - LSL) / (3*SD MR)$	Lesser of $(USL - AVG) / (3*SD Pop)$ and $(AVG - LSL) / (3*SD Pop)$

### Example

If you have fixed Target and LSL values and want to calculate Cr, the calculation would be  $3S / (Tgt - LSL)$ .

(Divide 3 standard deviations by the difference between the target and the lower spec limit.)

# Statistics Definitions

**Average** represented by **Avg**

The mathematical average of all data points in a field. Also known as the Arithmetic Mean

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad \text{where } n = \text{number of points in the field}$$

**Calc % above TGT** represented by **C%>TGT**

Calculated percent above target.

**Calc % above UCL** represented by **C%>UCL**

Calculated percent above upper control limit.

**Calc % above USL** represented by **C%>USL**

Calculated percent above upper specification limit.

**Calc % above UWL** represented by **C%>UWL**

Calculated percent above upper warning limit.

**Calc % below LCL** represented by **C%<LCL**

Calculated percent below lower control limit.

**Calc % below LSL** represented by **C%<LSL**

Calculated percent below upper specification limit.

**Calc % below LWL** represented by **C%<LWL**

Calculated percent below lower warning limit.

**Calc % below TGT** represented by **C%<TGT**

Calculated percent below target.

**Calc % OSL** represented by **C%OSL**

This statistic uses a table to infer the theoretical %OSL. The table assumes normality, control and the right number of samples so that the Average and SD are representative of the true population. Unless these assumptions hold true, the number will be inaccurate. When you have calculated quantities and no observed quantities, first suspect the assumptions above have been violated ... particularly

with large sample sizes. It is calculated as the value of the percentage (or parts per million) of a normal distribution curve that will fall outside the specification limits (points in the white band of the chart) if a large number of samples are taken. This value is theoretical and is accurate only if the variable is in statistical control and the data is normally distributed.

To calculate the C%OSL you will need:

- the average (**Avg**)
- the standard deviation (**Std Dev - pop**) of the population
- the Upper Specification Limit (**USL**)
- the Lower Specification Limit (**LSL**)
- a **Z-Table** to find the proportion out (**Pz**).

$$\text{Calculate } P_{z_{\text{lower}}} = \frac{|\bar{X} - \text{LSL}|}{S}$$

Look up  $P_{z_{\text{lower}}}$  in the Z - Table

If  $\bar{X} < \text{LSL}$  then  $P_{z_{\text{lower}}} = 1 - P_{z_{\text{lower}}}$

$$\text{Calculate } P_{z_{\text{upper}}} = \frac{|\text{USL} - \bar{X}|}{S}$$

Look up  $P_{z_{\text{upper}}}$  in the Z - Table

If  $\bar{X} > \text{USL}$  then  $P_{z_{\text{upper}}} = 1 - P_{z_{\text{upper}}}$

$$\text{Calculate } C\%OOL = (P_{z_{\text{lower}}} + P_{z_{\text{upper}}}) * 100$$

**Calc N above TGT** represented by **Cn>TGT**

Calculated number above target.

**Calc N above UCL** represented by **Cn>UCL**

Calculated number above upper control limit.

**Calc N above USL** represented by **Cn>USL**

Calculated number above upper specification limit.

**Calc N above UWL** represented by **Cn>UWL**

Calculated number above upper warning limit.

**Calc N below LCL** represented by **Cn<LCL**

# Statistics Definitions

Calculated number below upper control limit.

**Calc N below LSL** represented by **Cn<LSL**

Calculated number below lower specification limit.

**Calc N below LWL** represented by **Cn<LWL**

Calculated number below lower warning limit.

**Calc N below TGT** represented by **Cn<TGT**

Calculated number below target.

**Calc N OSL** represented by **CnOSL**

**Calc ppm above TGT** represented by **Cppm>TGT**

Calculated parts per million above target.

**Calc ppm above UCL** represented by **Cppm>UCL**

Calculated parts per million above upper control limit.

**Calc ppm above USL** represented by **Cppm>USL**

Calculated parts per million above upper specification limit.

**Calc ppm above UWL** represented by **Cppm>UWL**

Calculated parts per million above upper warning limit.

**Calc ppm below LCL** represented by **Cppm<LCL**

Calculated parts per million below lower control limit.

**Calc ppm below LSL** represented by **Cppm<LSL**

Calculated parts per million below lower specification limit.

**Calc ppm below LWL** represented by **Cppm<LWL**

Calculated parts per million below lower warning limit.

**Calc ppm below TGT** represented by **Cppm<TGT**

Calculated parts per million below target.

## **Calc ppm OSL** represented by **CppmOSL**

Calculated parts per million out of specification limits.

## **Calc ppm UCI** represented by **CppmUCI**

The Calculated parts per million Upper Confidence Interval is an adjustment to the defect estimate to account for sampling error using the normal distribution theory.

The UCI ppm values are defect estimates that represent the worst-case scenario of the actual defect level since it takes into account the potential error associated with low sample size. The larger the sample size, the lower the UCI ppm value. For example, if random samples were repeatedly taken from the same population and used to determine the defect estimate in ppm, the 95% UCI ppm would represent the ppm number that 95% of the samples estimates would fall below and 5% would fall above.

## **Calc Lower Warning Limit** represented by **CLWL**

Calculated lower warning limit.

## **Calc Upper Warning Limit** represented by **CUWL**

Calculated upper warning limit.

**Clearance** see [Ppk](#)

## **Co-efficient Variation** represented by **%CV**

The absolute value of (the standard deviation divided by the average) multiplied by 100

$$\%CV = \left| \frac{SD}{X} \right| \times 100$$

## **Control Limits**

Upper Control Limit represented by **UCL**

Lower Control Limit represented by **LCL**

Control limits reflect target +/- 3S based on the historical performance of each line in question. Choosing a best time (absence of special causes) for basing the historical

performance on is preferred if special causes can be eliminated or corrected. Control limits are based on the best capability of each line are preferably symmetric. +/- 3S can be used for any chart with 98-100% confidence without assumptions as to normality and control. Control Limits in QW 5.0 can be fixed, blank or continuously calculated values. These values are used in the program to set the values where color changes occur in the Charts. In QW 5.0 the control limits are relative to the target and specification limits and establish a "warning track" inside the specification limits.

## Correlation Coefficient

All Null values within variable X are filled with extrapolated values

All Null values within variable Y are filled with extrapolated values

All leading and trailing pairs of X, Y data are removed if either X or Y are null.

Determine R (Correlation Coefficient)

$$R = \frac{N * \sum xy - (\sum x * \sum y)}{\sqrt{|N * \sum x^2 - (\sum x)^2|} + \sqrt{|N * \sum y^2 - (\sum y)^2|}}$$

If  $R > 1$  then  $R = 1$

If  $R < -1$  then  $R = -1$

Determine T (Significance Test)

$$T = \frac{R}{\sqrt{\frac{1 - R^2}{N - 2}}}$$

Determine if Significant.

If  $|T| > 1.96$  highlight the value (95% significant)

Display Correlation Coefficient as  $R^2 * 100$

# Statistics Definitions

**Current Average** represented by **Avg-current**

**Current Standard Deviation** represented by **S-current**

## f-calc

Tests whether the Variation of two populations are different.

$$\text{if } SD_1 > SD_2 \text{ fcalc} = \frac{SD_1^2}{SD_2^2}$$
$$\text{if } SD_1 < SD_2 \text{ fcalc} = \frac{SD_2^2}{SD_1^2}$$

**Last Value** represented by **Last Value**

**Lower Capability Limit** represented by **Cpl**

$$(AVG - LSL) / (3 * SD MR)$$

**Lower Control Limit** represented by **LCL**

**Lower Performance Limit** represented by **Ppl**

$$(AVG - LSL) / (3 * SD Pop)$$

**Lower Process Capability (-3S)** represented by **Avg-3S**

**Lower Process Capability (-4S)** represented by **Avg-4S**

**Lower Spec. Limit** represented by **LSL**

**Lower Warning Limit** represented by **LWL**

**Maximum Value** represented by **Max**  
The highest value found in the data selected.

**Median** represented by **M**

The median is a measure of control tendency derived from taking the middlemost or most central point in a set of sorted points. Half of the points will lie above the median, and the other half will lie below the median. If the number of points is odd the middle point is the median. If there is an even number of points the median is the average of the middle points.

**Minimum Value** represented by **Min Value**  
The lowest value found in the data selected.

**Moving Range Average** represented by  $\overline{MR}$   
The Moving Range Average is based on successive differences between individual values.

$$\overline{MR} = \frac{\sum_{i=1}^{n-1} |X_{i+1} - X_i|}{n - 1}$$

**Moving Range Upper Control Limit** represented by **MRUCL**  
Is the Moving Range Upper Control Limit

$$MRUCL = \overline{MR} * 3.267$$



# Statistics Definitions

**Number of Points** represented by **N**

The number of non-null values found in the data selected.

**Obs % above TGT** represented by **O%>TGT**

Observed percent above target.

**Obs % above UCL** represented by **O%>UCL**

Observed percent above upper control limit.

**Obs % above USL** represented by **O%>USL**

Observed percent above upper specification limit.

**Obs % above UWL** represented by **O%>UWL**

Observed percent above upper warning limit.

**Obs % below LCL** represented by **O%<LCL**

**Obs % below LSL** represented by **O%<LSL**

**Obs % below LWL** represented by **O%<LWL**

**Obs % below TGT** represented by **O%<TGT**

**Obs % on TGT** represented by **O%=TGT**

**Obs % OSL** represented by **O%OSL**

**Obs N above TGT** represented by **On>TGT**

**Obs N above UCL** represented by **On>UCL**

**Obs N above USL** represented by **On>USL**

# Statistics Definitions

**Obs N above UWL** represented by **On>UWL**

**Obs N below LCL** represented by **On<LCL**

**Obs N below LSL** represented by **On<LSL**

**Obs N below LWL** represented by **On<LWL**

**Obs N below TGT** represented by **On<TGT**

**Obs N on TGT** represented by **On=TGT**

**Obs N OSL** represented by **OnOSL**

**Obs ppm above TGT** represented by **Oppm>TGT**

**Obs ppm above UCL** represented by **Oppm>UCL**

**Obs ppm above USL** represented by **Oppm>USL**

**Obs ppm above UWL** represented by **Oppm>UWL**

**Obs ppm below LCL** represented by **Oppm<LCL**

**Obs ppm below LSL** represented by **Oppm<LSL**

**Obs ppm below LWL** represented by **Oppm<LWL**

**Obs ppm below TGT** represented by **Oppm<TGT**

**Obs ppm on TGT** represented by **Oppm=TGT**

**Obs ppm OSL** represented by **OppmOSL**

**Obs ppm UCI** represented by **OppmUCI**

The Observed parts per million Upper Confidence Interval is an adjustment to the defect estimate to account for sampling error using the binomial distribution theory. The UCI ppm values are defect estimates that represent the worst-case scenario of the actual defect level since it considers the potential error associated with low sample size. The larger the sample size, the lower the UCI ppm value. For example, if random samples were repeatedly taken from the same population and used to determine the defect estimate in ppm, the 95% UCI ppm would represent the ppm number that 95% of the samples estimates would fall below and 5% would fall above.

## Point

The value of each cell in a data collection.

## Process Capability Index represented by Cpk

Requires two separate calculations and then you select the smaller of the two results.

These equations describe the clearance between the process distribution curve and the specification limits in terms of standard deviation. The concern would be on whichever end of the distribution curve is closest to the specification limit (could be equal if the process is perfectly centered within the specification limits). A value of 1.33 or greater is considered good. A value of 1.33 would mean that the process average is four deviations away from the nearest specification limit. Compare Cr with Cpk. The process variation is in the numerator of Cr, and the denominator of Cpk (also Cp). Low Cr and high Cpk numbers are good.

$$Cpk = \frac{USL - \bar{X}}{3 * SD MR} \text{ and } \frac{\bar{X} - LSL}{3 * SD MR}$$

## Process Capability Rate represented by Cp

The inverse of [Capability Ratio](#).

$$Cp = \frac{1}{6 * SD MR / (USL - LSL)}$$

## Process Performance Index represented by Ppk

Requires two separate calculations and then you select the smaller of the two results.

These equations describe the clearance between the process distribution curve and the specification limits in terms of standard deviation. The concern would be on whichever end of the distribution curve is closest to the specification limit (could be equal if the process is perfectly centered within the specification limits). A value of 1.33 or greater is considered good. A value of 1.33 would mean that the process average is four deviations away from the nearest specification limit. Compare Cr with Ppk. The process variation is in the numerator of Cr, and the denominator of Ppk (also Pp). Low Cr and high Ppk numbers are good.

$$Ppk = \frac{USL - \bar{X}}{3 * SD Pop} \text{ and } \frac{\bar{X} - LSL}{3 * SD Pop}$$

## Process Performance Rate represented by Pp

The inverse of [Capability Ratio](#).

$$Pp = \frac{1}{6 * SD Pop / (USL - LSL)}$$

## Range AVG represented by R-AVG

## Range LWL represented by R-LWL

## Range UCL represented by R-UCL

## Range UWL represented by R-UWL

## Rule Violation represented by Rule

## SD Low

SD Low is defined as the lowest standard deviation of a consecutive group of **m** points with a population of **n** points.

**m** is determined by multiplying **n** by 20%

if the result is < 10 then **m** = 10

if the result is > 100 then **m** = 100

**Sigma** represented by **Sigma**

## Specification Limits

Upper Spec. Limit represented by **USL**

Lower Spec. Limit represented by **LSL**

Fixed specification limits must be outside fixed control limits and the target. They reflect a shutdown scenario or a high probability that a real process or product problem exists that must be corrected immediately. They are defined via product and process research. All lines have the process capability to run within specifications with a margin that allows for control charts to provide their intended benefit of early warning. A Cr of < .75 and Tz of < .5 is preferred to allow early warning. avoid over adjustment and reactive operation. Specifications should not be calculated based on lead line capability. This will be a dilution of the quality system in terms of real meaning to the business.

**Std Dev - mr** represented by **S-mr**

The standard deviation of the Moving Range Average.

$$SD MR = \overline{MR} / 1.128$$

## Std Dev - pop represented by S-pop

The measurement of variability around the average

Take each individual data point and calculate the deviation of that point from the average, square the deviation (multiply it by itself), add up all the results, divide the total by one less than the number of points, then take the square root of the result.

$$SD = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}}$$

## Sum represented by Sum

### t-calc

Tests whether the Average of two populations are different.

$$\frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{SD_1^2}{N_1} + \frac{SD_2^2}{N_2}}}$$

## Target represented by TGT

The target should reflect the optimum place to average the process relative to product and process research. The process should have the capability to average a target. If not, the 3, 5 and 7 point rules cannot be used without creating over adjustment. Each lines' capability may require a unique target in order to use 3, 5 and 7 point control rules for certain variables. The target can be fixed, blank or continuously calculated.

## Target Deviation represented by T-Dev

In the statistics of QW 5 it is the separation between the process average and the Target value of the variable. The value shown is in units of measure of the variable and can be positive or negative, depending on whether the process average is above or below target.

# Statistics Definitions

$$\text{T-Dev} = \text{Avg} - \text{TGT}$$

where **TGT** = the target value for the variable

## **Targeting** represented by **Tz**

In QW6 this is determined by dividing the Target Deviation by the Standard Deviation

This produces a number that measures how well the process is centered on target in standard deviations. A good process should be less than 0.5 away from target (can be plus or minus). A good Target Z does not necessarily mean the process is always in specification.

If  $C_r \leq .5$  then  $T_z$  is calculated depending on the type of spec. limits and calculated Average.

If TGT and both spec. limits are fixed the Avg is compared against Tgt +/- 4.2% of the spec. range.

If TGT and only one spec. limits are fixed the Avg is compared against Tgt +/- 8.4% of Tgt to spec. range.

If AVG is within these limits, then  $T_z = 0$ .

If AVG is greater than these limits, then  $T_z = +1$ .

If AVG is less than these limits then  $T_z = -1$ .

$$T_z = \frac{\bar{X} - Tgt}{S}$$

## **Upper Capability Limit** represented by **Cpu**

$$(\text{USL} - \text{AVG}) / (3 * \text{SD MR})$$

## **Upper Performance Limit** represented by **Ppu**

# Statistics Definitions

$(USL - AVG) / (3 * SD Pop)$

**Upper Process Capability (+3S)** represented by **Avg +3S**

**Upper Process Capability (+4S)** represented by **Avg +4S**

**Upper Spec Limit** represented by **USL**

**Upper Warning Limit** represented by **UWL**

**Variation** represented by **Cr**

A term that describes the size of the process variability relative to the size of the specification range. It does not consider where the process is centered so a good value for capability ratio does not mean that all the values are within specification.

When both specifications are fixed Quality Window calculates CR as

$$Cr = \frac{6S}{USL - LSL}$$