Estimation of Industrial Process Capability Indices for Non-Normal Distributions

DUMITRAȘCU ADELA-ELIZA¹, NEDELCU ANIȘOR¹, DUMITRAȘCU DORIN²,
FOTA ADRIANA¹
Technological Engineering Faculty¹, Automotive Engineering Department²
“Transilvania” University of Brasov
29 Eroilor Street, 500036, Brasov
ROMANIA
dumitrascu_a@unitbv.ro¹; a.nedelcu@unitbv.ro¹; d.dumitrascu@unitbv.ro²; fota.a@unitbv.ro¹

Abstract: - Not all measurements are normally distributed. When they are not, use of the normal distribution for the estimation of capability indices and the construction of control charts can be very misleading. This paper contains the estimation of process capability indices for non-normal distributions using Monte-Carlo simulation. The non-normal distributions considered are: uniform, exponential, Weibull, \( \chi^2 \) and Gamma. Principal advantage of this specific formulation is that it facilitates evaluation of the degradation of the capability of the process due to bias between process mean and the process target.

Key-Words: - Monte-Carlo simulation, histogram, process capability indices.

1 Introduction
Manufacturing process would commonly be described in terms of three parameters: the finite target value (T), an upper specification limit (USL), and a lower specification limit (LSL). All the parts for which the measured value “x” for a certain specification exceeds the USL or falls below the LSL are rejected. If the process target value for the product characteristic is centered between the USL and LSL, then the tolerances are said to be symmetric. The capability index is of interest to the manufacturing community because it consolidates the details in a complicated multifaceted manufacturing process down to one quantity which can be used to predict the fraction of parts rejected [1].

2 Process Capability Indices for Non-Normal Distribution
Paper [2] proposed generalizations of Cp for any underlying distribution as follows:

\[
Cnp = \frac{USL - LSL}{(F_{99,865} - F_{0.135})}, \quad (1)
\]

where \( F_a \) is the (100a)th percentile of the distribution. The following indices used for estimation of process capability are:

\[
Cnpk = \frac{\min\{USL - M, M - LSL\}}{2}, \quad (2)
\]

\[
Cnpm = \frac{USL - LSL}{6\left(\frac{F_{99.865} - F_{0.135}}{6}\right)^2 + (M - T)^2}, \quad (3)
\]

\[
Cnmpm = \frac{\min\{USL - M, M - LSL\}}{3\left(\frac{F_{99.865} - F_{0.135}}{6}\right)^2 + (M - T)^2}, \quad (4)
\]

\[
Spnk = \Phi^{-1}\left(1 + \frac{F(USL) - F(LSL)}{2}\right), \quad (5)
\]

There are two methods of determination of the values \( x_{0.00135}, x_{0.5}, x_{0.99865} \) [3]:
- Clements’ method; it is an approximate method basing on values of shape parameters i.e. kurtosis and skewness; in this method there is no need to know the form of distribution of the analyzed parameter.
- Exact method requiring knowledge of a density function \( f(x) \) determining a distribution of the analyzed parameter; in this case the values \( x_{0.00135}, x_{0.5}, x_{0.99865} \) we determine from the relationships:

\[
\int_{-\infty}^{x_{0.00135}} f(x)dx = 0.00135 \quad (6)
\]

\[
\int_{-\infty}^{x_{0.5}} f(x)dx = 0.5 \quad (7)
\]
3 Simulated Researches

Estimation of capability indices was based on Monte-Carlo simulation. It were simulated a sample size of 100 values for uniform, exponential, Weibull and Gamma distribution. Also, it were determined the process capability indices: Cnp, Cnpk, Cnpmk and Snpmk.

Uniform distribution:
Minimum and maximum value:

\[ n = 100 \]
\[ x_{\text{min}} = 16.99 \]
\[ x_{\text{max}} = 25.72 \]

Range: \( r = 8.735 \)
\[ nc = 9 \]
\[ lc = 1.092 \]

<table>
<thead>
<tr>
<th>Cell Boundary</th>
<th>fa</th>
<th>fr</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.4441</td>
<td>17.5359</td>
<td>7</td>
</tr>
<tr>
<td>17.5359</td>
<td>18.6278</td>
<td>12</td>
</tr>
<tr>
<td>18.6278</td>
<td>19.7196</td>
<td>12</td>
</tr>
<tr>
<td>19.7196</td>
<td>20.8115</td>
<td>18</td>
</tr>
<tr>
<td>20.8115</td>
<td>21.9034</td>
<td>5</td>
</tr>
<tr>
<td>21.9034</td>
<td>22.9952</td>
<td>12</td>
</tr>
<tr>
<td>22.9952</td>
<td>24.0871</td>
<td>13</td>
</tr>
<tr>
<td>24.0871</td>
<td>25.1789</td>
<td>16</td>
</tr>
<tr>
<td>25.1789</td>
<td>26.2708</td>
<td>5</td>
</tr>
</tbody>
</table>

where \( fa \) – absolute frequency; \( fr \) – relative frequency.

Exponential distribution:
Minimum and maximum value:

\[ n = 100 \]
\[ x_{\text{min}} = 1.77 \]

Range: \( r = 198.422 \)
\[ nc = 9 \]
\[ lc = 24.803 \]

<table>
<thead>
<tr>
<th>Cell Boundary</th>
<th>fa</th>
<th>fr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7676</td>
<td>26.5703</td>
<td>49</td>
</tr>
<tr>
<td>26.5703</td>
<td>51.3731</td>
<td>19</td>
</tr>
<tr>
<td>51.3731</td>
<td>76.1758</td>
<td>15</td>
</tr>
<tr>
<td>76.1758</td>
<td>100.9785</td>
<td>9</td>
</tr>
<tr>
<td>100.9785</td>
<td>125.7813</td>
<td>3</td>
</tr>
<tr>
<td>125.7813</td>
<td>150.5840</td>
<td>2</td>
</tr>
<tr>
<td>150.5840</td>
<td>175.3868</td>
<td>1</td>
</tr>
<tr>
<td>175.3868</td>
<td>200.1895</td>
<td>1</td>
</tr>
<tr>
<td>200.1895</td>
<td>224.9922</td>
<td>0</td>
</tr>
</tbody>
</table>

Weibull distribution:
Minimum and maximum value:

\[ n = 100 \]
\[ x_{\text{min}} = 0.58 \]
\[ x_{\text{max}} = 1.27 \]

Range: \( r = 0.687 \)
\[ nc = 9 \]
\[ lc = 0.086 \]

<table>
<thead>
<tr>
<th>Cell Boundary</th>
<th>fa</th>
<th>fr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5356</td>
<td>0.6214</td>
<td>1</td>
</tr>
<tr>
<td>0.6214</td>
<td>0.7073</td>
<td>5</td>
</tr>
<tr>
<td>0.7073</td>
<td>0.7932</td>
<td>9</td>
</tr>
<tr>
<td>0.7932</td>
<td>0.8791</td>
<td>13</td>
</tr>
<tr>
<td>0.8791</td>
<td>0.9649</td>
<td>18</td>
</tr>
<tr>
<td>0.9649</td>
<td>1.0508</td>
<td>21</td>
</tr>
<tr>
<td>1.0508</td>
<td>1.1367</td>
<td>22</td>
</tr>
<tr>
<td>1.1367</td>
<td>1.2226</td>
<td>8</td>
</tr>
<tr>
<td>1.2226</td>
<td>1.3085</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
\int_{-\infty}^{x_{\text{max}}} f(x)dx = 0.99865
\]
$\chi^2$ distribution:
Minimum and maximum value:
\[ n = 100 \]
\[ x_{\text{min}} = 0.04 \]
\[ x_{\text{max}} = 14.60 \]
Range: \( r = 14.556 \)
\[ nc = 9 \]
\[ lc = 1.819 \]

4. Capability Indices Estimation for Non-Normal Distributions

4.1 Uniform distribution:
\[ C_{np} = 1.003 \]
\[ Z(0.99865) = 25.764183 \]
\[ Z(0.00135) = 16.992597 \]
\[ M = \text{Median} = 21.59601152 \]
\[ C_{npk} = 0.998391926 \]
\[ C_{npkm} = 0.580646257 \]
\[ S_{npkm} = 0.545790202 \]

4.2 Exponential distribution:
\[ C_{np} = 1.025 \]
\[ Z(0.99865) = 195.44536 \]
\[ Z(0.00135) = 1.7894444 \]
\[ M = \text{Median} = 26.52521349 \]
\[ C_{npk} = 1.000786082 \]
\[ C_{npkm} = 0.478120126 \]
\[ S_{npkm} = 0.505790202 \]
4.3 Weibull distribution:
\[ Cnp = 1.008 \]
\[ Z(0.99865) = 1.2649852 \]
\[ Z(0.00135) = 0.5902886 \]
\[ M = \text{Median} = 0.988992785 \]
\[ Cnpk = 0.981937198 \]
\[ Cnpmk = 0.956576864 \]
\[ Snpmk = 0.538950491 \]

4.4 \( \chi^2 \) distribution:
\[ Cnp = 0.964 \]
\[ Z(0.99865) = 14.595556 \]
\[ Z(0.00135) = 0.07206885 \]
\[ M = \text{Median} = 2.157075183 \]
\[ Cnpk = 0.794757866 \]
\[ Cnpmk = 0.47940711 \]
\[ Snpmk = 0.375861697 \]

4.5 Gamma distribution:
\[ Cnp = 0.976 \]
\[ Z(0.99865) = 12.763703 \]
\[ Z(0.00135) = 1.4899535 \]
\[ M = \text{Median} = 5.362288161 \]
\[ Cnpk = 0.964371291 \]
\[ Cnpmk = 0.513038538 \]
\[ Snpmk = 0.444715725 \]

5 Conclusions
In the figure 6 it can be observed the differences between industrial process capability indices. Compared with normal distribution, the non-normal distribution gives an accurate image of the analyzed process. The process should be off-center to the high side because the pieces are likely to come off the backstop only a small amount.

We can still calculate the process capability if we can first find a curve that fits our data distribution and then use this model to predict yields.

The techniques of process capability analysis, including process capability indices, however:

- First and foremost, get the process into statistical control.
- Be sure that the curve fits the data, especially in the distribution tails.
- Remember that yields are more direct measurements of process performance than process capability indices. Be sure that actual counts confirm the yield predictions based on models.
- Verify that the predicted performance actually occurs with future runs. This is much more likely to be the case if the process is kept in statistical control.

References: