Percentile Control Charts for the Pareto Distribution

Ruth Burkhalter

University of South Dakota
Introduction
The generalized Pareto Distribution

The generalized Pareto distribution is:
- A skewed distribution
- Can be used to model and predict lifetime data
- Useful in industry for quality standards
- Control charts will assist in monitoring a product’s lifetime
The probability distribution function (PDF), the cumulative distribution function (CDF), and the quantile function for the generalized Pareto distribution can be written as follows:

\[ f(x; \alpha, \lambda) = \alpha \lambda (1 + x \lambda)^{-(\alpha + 1)}; x > 0, \]  

\[ F(x; \alpha, \lambda) = 1 - (1 + x \lambda)^{-\alpha}; x > 0, \]  

\[ Q(p; \alpha, \lambda) = \frac{1}{\lambda} ((1 - p)^{-1/\alpha} - 1); 0 < p < 1. \]
Maximum Likelihood Estimate

Likelihood and Log Likelihood Function

Let \( X = \{x_1, x_2, \ldots, x_n\} \) be a random sample of size \( n \) from the Pareto distribution. Then the likelihood function is defined as follows:

\[
L(X; \alpha, \lambda) = \prod_{i=1}^{n} \alpha \lambda (1 + x_i \lambda)^{-\alpha-1}.
\] (4)

Taking the natural logarithm of the likelihood function, one produces the log likelihood function which is defined as follows:

\[
l(X; \alpha, \lambda) = n \ln(\alpha) + n \ln(\lambda) - (\alpha + 1) \sum_{i=1}^{n} \ln(1 + x_i \lambda).
\] (5)
In order to find the maximum likelihood estimator (MLE), we take the partial derivative of $l(X; \alpha, \lambda)$ with respect to $\alpha$ and $\lambda$ and set those partial derivatives equal to 0.

$$\frac{\partial}{\partial \alpha} l(X; \alpha, \lambda) = \frac{n}{\alpha} - \sum_{i=1}^{n} \ln(1 + x_i \lambda) = 0 \quad (6)$$

$$\frac{\partial}{\partial \lambda} l(X; \alpha, \lambda) = \frac{n}{\lambda} - (\alpha + 1) \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} = 0 \quad (7)$$
Maximum Likelihood Estimate

Solve for $\alpha$ and $\lambda$

Solving equation (6) for $\alpha$ in terms of $\lambda$ and plugging our solution into equation (7), we have the following:

\[
\frac{n}{\lambda} - \left( \frac{n}{\sum_{i=1}^{n} \ln(1 + x_i \lambda)} + 1 \right) \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} = 0.
\]  

(8)
Maximum Likelihood Estimate

Estimations for $\alpha$ and $\lambda$

We use this equation to estimate $\lambda$ and then use the estimation of $\lambda$ to estimate $\alpha$. These estimates, denoted $\hat{\alpha}_n$ and $\hat{\lambda}_n$, are then used to find the MLE of the $100p$th percentile by using the quantile function as shown:

$$\hat{Q}_n(p; \hat{\alpha}_n, \hat{\lambda}_n) = \frac{1}{\hat{\lambda}_n}((1 - p)^{-1/\hat{\alpha}_n} - 1); 0 < p < 1.$$  \hspace{1cm} (9)
It can be shown that $\sqrt{n}((\hat{\alpha}, \hat{\lambda}) - (\alpha, \lambda)) \rightarrow N(0, I^{-1}(\alpha, \lambda))$ where $0$ is the zero two-dimensional column vector and $I(\alpha, \lambda)$ is the Fischer matrix. The Fischer matrix is defined as follows:

$$I(\alpha, \lambda) = -\frac{1}{n} \begin{bmatrix}
E\left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha^2} \right) & E\left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha \partial \lambda} \right) \\
E\left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \lambda \partial \alpha} \right) & E\left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \lambda^2} \right)
\end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} \\
I_{21} & I_{22}\end{bmatrix}. \quad (10)$$
Fischer Matrix

Second Partial Derivatives

\[ \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha^2} = - \frac{n}{\alpha^2} \]

\[ \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha \partial \lambda} = - \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} \]

\[ \frac{\partial^2 l(\alpha, \lambda)}{\partial \lambda \partial \alpha} = - \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} \]

\[ \frac{\partial^2 l(\alpha, \lambda)}{\partial \lambda^2} = - \frac{n}{\lambda^2} + (\alpha + 1) \sum_{i=1}^{n} \frac{x_i^2}{(1 + x_i \lambda)^2} \]
Fischer Matrix

Solutions for $l_{11}, l_{12}, l_{21},$ and $l_{22}$

\[
\begin{align*}
l_{11} &= -\frac{1}{n}E \left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha^2} \right) = \frac{1}{\alpha^2} \\
l_{12} &= l_{21} = -\frac{1}{n}E \left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \alpha \partial \lambda} \right) = \frac{1}{\lambda(\alpha + 1)} \\
l_{22} &= -\frac{1}{n}E \left( \frac{\partial^2 l(\alpha, \lambda)}{\partial \lambda^2} \right) = \frac{1}{\lambda^2} - \frac{2}{\lambda(\alpha + 2)}
\end{align*}
\]

\[
I(\alpha, \lambda) = \begin{bmatrix}
\frac{1}{\alpha^2} & \frac{1}{\lambda(\alpha + 1)} \\
\frac{1}{\lambda(\alpha + 1)} & \frac{1}{\lambda^2} - \frac{2}{\lambda(\alpha + 2)}
\end{bmatrix} \tag{11}
\]
Finding $l_{11}$ and $l_{21}$

$$l_{12} = l_{21} = -rac{1}{n} E \left( - \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \int_{0}^{\infty} \frac{x}{1 + x \lambda} \cdot f(x) \, dx$$

$$= \int_{0}^{\infty} \frac{x}{1 + x \lambda} \cdot \frac{\alpha \lambda (1 + x \lambda)^{-\alpha-1}}{\alpha \lambda (1 + x \lambda)^{-\alpha-1}} \, dx$$

$$= \alpha \lambda \int_{0}^{\infty} x (1 + x \lambda)^{-\alpha-2} \, dx.$$
Fischer Matrix

U substitution for \( l_{11} \) and \( l_{21} \)

Let \( u = (1 + x\lambda)^{-1} \). So \( u^{\alpha+2} = (1 + x\lambda)^{-\alpha-2} \) and \( u^{-1} = 1 + x\lambda \). This means \( (-1)u^{-2}du = \lambda dx \) and \( x = \frac{u^{-1} - 1}{\lambda} = \frac{1 - u}{u\lambda} \).

\[
l_{12} = l_{21} = \alpha \int_1^0 \frac{1 - u}{u\lambda} u^{\alpha+2}(-1)u^{-2}du
\]

\[
= \frac{\alpha}{\lambda} \int_0^1 (1 - u)u^{\alpha-1}du
\]

\[
= \frac{\alpha}{\lambda} \beta(\alpha, 2) = \frac{\alpha}{\lambda} \left( \frac{\Gamma(\alpha)\Gamma(2)}{\Gamma(\alpha + 2)} \right)
\]

\[
= \frac{\alpha}{\lambda} \left( \frac{\Gamma(\alpha)}{\alpha(\alpha + 1)\Gamma(\alpha)} \right)
\]

\[
= \frac{1}{\lambda(\alpha + 1)}
\]
Confidence Interval Based for the Shewhart-Type Chart

**Standard Error**

It can be shown that

\[
\frac{\hat{Q}_n(p; \hat{\alpha}_n, \hat{\lambda}_n) - Q(p; \alpha, \lambda)}{\sigma_{p,n}^2} \rightarrow N(0, 1) \quad (12)
\]

where

\[
\sigma_{p,n}^2 = \frac{1}{n} \nabla Q(p; \alpha, \lambda)^T \mathbf{I}^{-1}(\alpha, \lambda) \nabla Q(p; \alpha, \lambda) \quad (13)
\]

and \( \nabla Q(p; \alpha, \lambda) \) is the gradient of \( Q(p, \alpha, \lambda) \) with respect to \( \alpha \) and \( \lambda \).

The asymptotic standard error can be estimated by the following:

\[
SE_{\hat{Q}_m} = \sqrt{\frac{1}{m} \nabla Q^T(p; \hat{\alpha}_n, \hat{\lambda}_n) \hat{I}_n(\hat{\alpha}_n, \hat{\lambda}_n) \nabla Q(p; \hat{\alpha}_n, \hat{\lambda}_n)} \quad (14)
\]
Confidence Interval for the Shewhart-Type Chart

Phase I

We assume $k$ in-control pre-samples, each of size $m$, are taken from the generalized Pareto distribution for creating the control chart. Let $n = m \times k$ be the total sample size of Phase I.

We obtain the MLE for $j$th pre-sample of size $m$ for $j = 1, 2, ..., k$ and take the sample mean as follows:

$$
\bar{\hat{Q}}_{p,m}(\hat{\alpha}_m, \hat{\lambda}_m) = \frac{1}{k} \sum_{j=1}^{k} \hat{Q}_{p,m}^j(\hat{\alpha}_m^j, \hat{\lambda}_m^j) \quad (15)
$$

This sample mean becomes the center line of our control chart.
Confidence Interval for the Shewhart-Type Chart

Control Limits

The upper and lower control limits are as follows:

\[
UCL_{SH} = \bar{Q}_{p,m}(\hat{\alpha}_m, \hat{\lambda}_m) + z_{1-\gamma/2} \cdot SE_{Q_m},
\]

\[
LCL_{SH} = \bar{Q}_{p,m}(\hat{\alpha}_m, \hat{\lambda}_m) - z_{1-\gamma/2} \cdot SE_{Q_m},
\]

where \( z_{1-\gamma/2} \) satisfies \( \Phi(z_{1-\gamma/2}) = 1 - \gamma/2 \) with \( 0 < \gamma < 1 \) and \( \Phi \) is the standard normal CDF and \( \gamma \) is called the false alarm rate (FAR).
Parametric Bootstrap Charts
For the MLE

1. Use the $n$ sample observations from Phase I to find the MLE.
2. Generate $m$ parametric bootstrap observations from the generalized Pareto distribution of (2), replacing $\alpha$ and $\lambda$ by the corresponding MLEs, found in Step 1. Denote the parametric bootstrap observation by $x^*_1, x^*_2, \ldots, x^*_m$.
3. Find the MLEs of $\alpha$ and $\lambda$ using $x^*_1, x^*_2, \ldots, x^*_m$ from Step 2 and denote them $\hat{\alpha}_m^*$ and $\hat{\lambda}_m^*$.
4. Compute the bootstrap estimate of the 100$p$th percentile using $\hat{\alpha}_m^*$ and $\hat{\lambda}_m^*$ as following:

$$\hat{Q}_{p,m}^*(\hat{\alpha}, \hat{\lambda}) = ((1.0 - p)^{-1/\hat{\alpha}_m^*} - 1)^{1/\hat{\lambda}_m^*}$$ (16)
We repeat Steps 2 through 4 $M$ times to obtain a size $M$ bootstrap sample $\hat{q}_{p,1}^*, \hat{q}_{p,2}^*, \ldots, \hat{q}_{p,M}^*$, where $M$ is any given large, positive integer.

Then given a FAR, $\gamma$, find the $(\gamma/2)$th and the $(1-\gamma/2)$th empirical quantiles of the bootstrap sample in Step 5 as the LCL and UCL.
Average Run Length for the MLE Parametric Bootstrap Chart and the Shewhart-Type Chart

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$n = 5$</th>
<th>$n = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARL</td>
<td>SERL</td>
</tr>
<tr>
<td>$\gamma_0 = 0.1$ (FAR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>9.0851</td>
<td>0.1250</td>
</tr>
<tr>
<td>$p = 0.05$</td>
<td>9.2691</td>
<td>0.1294</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>9.1935</td>
<td>0.1302</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>9.3039</td>
<td>0.1285</td>
</tr>
<tr>
<td>$\gamma_0 = 0.01$ (FAR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>89.7949</td>
<td>1.5565</td>
</tr>
<tr>
<td>$p = 0.05$</td>
<td>90.1718</td>
<td>1.5683</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>91.3064</td>
<td>1.5865</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>90.6854</td>
<td>1.5839</td>
</tr>
<tr>
<td>$\gamma_0 = 0.0027$ (FAR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>336.3652</td>
<td>6.4181</td>
</tr>
<tr>
<td>$p = 0.05$</td>
<td>333.6685</td>
<td>6.5899</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>339.3910</td>
<td>6.4750</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>340.8538</td>
<td>6.7502</td>
</tr>
</tbody>
</table>
Upper and Lower Control Limits for the MLE Parametric Bootstrap Chart and the Shewhart-Type Chart

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$n = 5$</th>
<th>$n = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LCL</td>
<td>UCL</td>
</tr>
<tr>
<td>$\gamma_0 = 0.1$ (FAR)</td>
<td>0.0012</td>
<td>0.0106</td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>0.0064</td>
<td>0.0546</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>0.1034</td>
<td>0.1127</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>0.0388</td>
<td>0.3166</td>
</tr>
<tr>
<td>$\gamma_0 = 0.01$ (FAR)</td>
<td>0.0007</td>
<td>0.0181</td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>0.0036</td>
<td>0.926</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>0.0075</td>
<td>0.1913</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>0.0214</td>
<td>0.5284</td>
</tr>
<tr>
<td>$\gamma_0 = 0.0027$ (FAR)</td>
<td>0.0005</td>
<td>0.0232</td>
</tr>
<tr>
<td>$p = 0.01$</td>
<td>0.0028</td>
<td>0.1185</td>
</tr>
<tr>
<td>$p = 0.10$</td>
<td>0.0058</td>
<td>0.2452</td>
</tr>
<tr>
<td>$p = 0.25$</td>
<td>0.0163</td>
<td>0.6761</td>
</tr>
</tbody>
</table>
Conclusion

- The Parametric Bootstrap Chart is better than the Shewhart-Type Chart when the sampling distribution is unknown.
- The ARL for the MLE is better than the ARL of the Shewhart-Type Chart.