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

Equations

The probability distribution function (PDF), the cummulative distribution function (CDF), and the quantile function for the generalized Pareto distribution can be written as follows:

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

$$F(x;\alpha,\lambda) = 1-(1+x\lambda)^{-\alpha}; x > 0, \tag{2}$$

$$Q(p;\alpha,\lambda) = \frac{1}{\lambda}((1-p)^{-1/\alpha}-1); 0 < p < 1. \tag{3}$$

Maximum Likelihood Estimate

Likelihood and Log Likelihood Function

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

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

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

$$l(X; \alpha, \lambda) = n\ln(\alpha) + n\ln(\lambda) - (\alpha + 1)\sum_{i=1}^{n} \ln(1 + x_i\lambda). \tag{5}$$

Maximum Likelihood Estimate

Partial Derivatives

In order to find the maximum likelihood estimator (MLE), we take the partial derivative of $l(X;\alpha,\lambda)$ with respect to α and λ and set those partial derivatives equal to 0.

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

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

Maximum Likelihood Estimate

Solve for α and λ

Solving equation (6) for α in terms of λ and plugging our solution into equation (7), we have the following:

$$\frac{n}{\lambda} - \left(\frac{n}{\sum_{i=1}^{n} \ln(1 + x_i \lambda)} + 1 \right) \sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} = 0. \tag{8}$$

Maximum Likelihood Estimate

Estimations for α and λ

We use this equation to estimate λ and then use the estimation of λ to estimate α. These estimates, denoted $\hat{\alpha_n}$ and $\hat{\lambda_n}$, are then used to find the MLE of the 100pth percentile by using the quantile function as shown:

$$\hat{Q}_n(p; \hat{\alpha}_n, \hat{\lambda}_n) = \frac{1}{\hat{\lambda}_n}((1-p)^{-1/\hat{\alpha}_n} - 1); 0 < p < 1. \quad (9)$$

Fischer Matrix

Fischer Matrix

It can be shown that $\sqrt{n}((\hat{\alpha}, \hat{\lambda}) - (\alpha, \lambda)) \to N(\mathbf{0}, \mathbf{I}^{-1}(\alpha, \lambda))$ where $\mathbf{0}$ if the zero two-dimensional column vector and $\mathbf{I}(\alpha, \lambda)$ is the Fischer matrix. The Fischer matrix is defined as follows:

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

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 I_{11},I_{12},I_{21}, and I_{22}

$$I_{11} = -\frac{1}{n}E\left(\frac{\partial^2 l(\alpha,\lambda)}{\partial\alpha^2}\right) = \frac{1}{\alpha^2}$$
$$I_{12} = I_{21} = -\frac{1}{n}E\left(\frac{\partial^2 l(\alpha,\lambda)}{\partial\alpha\partial\lambda}\right) = \frac{1}{\lambda(\alpha+1)}$$
$$I_{22} = -\frac{1}{n}E\left(\frac{\partial^2 l(\alpha,\lambda)}{\partial\lambda^2}\right) = \frac{1}{\lambda^2} - \frac{2}{\lambda(\alpha+2)}$$

$$\mathbf{I}(\alpha,\lambda) = \begin{bmatrix} \frac{1}{\alpha^2} & \frac{1}{\lambda(\alpha+1)} \\ \frac{1}{\lambda(\alpha+1)} & \frac{1}{\lambda^2} - \frac{2}{\lambda(\alpha+2)} \end{bmatrix} \tag{11}$$

Fischer Matrix

Finding I_{11} and I_{21}

$$
\begin{aligned}
I_{12} = I_{21} &= -\frac{1}{n} E\left(-\sum_{i=1}^{n} \frac{x_i}{1 + x_i \lambda} \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \int_0^\infty \frac{x}{1 + x\lambda} \cdot f(x) dx \\
&= \int_0^\infty \frac{x}{1 + x\lambda} (\alpha\lambda (1 + x\lambda)^{-\alpha - 1}) dx \\
&= \alpha\lambda \int_0^\infty x (1 + x\lambda)^{-\alpha - 2} dx.
\end{aligned}
$$

Fischer Matrix

U substitution for I_{11} and I_{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 $x = \dfrac{u^{-1} - 1}{\lambda} = \dfrac{1 - u}{u\lambda}$.

$$
\begin{aligned}
I_{12} = I_{21} &= \alpha \int_1^0 \frac{1-u}{u\lambda} u^{\alpha+2}(-1)u^{-2}du \\
&= \frac{\alpha}{\lambda} \int_0^1 (1-u)u^{\alpha-1}du \\
&= \frac{\alpha}{\lambda}\beta(\alpha, 2) = \frac{\alpha}{\lambda}\left(\frac{\Gamma(\alpha)\Gamma(2)}{\Gamma(\alpha+2)}\right) \\
&= \frac{\alpha}{\lambda}\left(\frac{\Gamma(\alpha)}{\alpha(\alpha+1)\Gamma(\alpha)}\right) \\
&= \frac{1}{\lambda(\alpha+1)}
\end{aligned}
$$

Confidence Interval Based for the Shewhart-Type Chart

Standard Error

It can be shown that

$$\frac{\hat{Q}_n(p; \hat{\alpha}_n, \hat{\lambda}_n) - Q(p; \alpha, \lambda)}{\sigma^2_{p,n}} \to N(0,1) \tag{12}$$

where

$$\sigma^2_{p,n} = \frac{1}{n} \nabla Q(p; \alpha, \lambda)^T \mathbf{I}^{-1}(\alpha, \lambda) \nabla Q(p; \alpha, \lambda) \tag{13}$$

and $\nabla Q(p; \alpha, \lambda)$ is the gradient of $Q(p, \alpha, \lambda)$ with respect to α and λ. The asymptotic standard error can be estimated by the following:

$$SE_{\hat{Q}_m} = \sqrt{\frac{1}{m} \nabla Q^T(p; \hat{\alpha}_n, \hat{\lambda}_n) \hat{I}_n(\hat{\alpha}_n, \hat{\lambda}_n) \nabla Q(p; \hat{\alpha}_n, \hat{\lambda}_n)} \tag{14}$$

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

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

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:

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

where $z_{1-\gamma/2}$ satisfies $\Phi(z_{1-\gamma/2}) = 1 - \gamma/2$ with $0 < \gamma < 1$ and Φ is the standard normal CDF and γ 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 α and λ by the corresponding MLEs, found in Step 1. Denote the parametric bootstrap observation by $x_1^*, x_2^*, ..., x_m^*$.
3. Find the MLEs of α and λ using $x_1^*, x_2^*, ..., x_m^*$ from Step 2 and denote them $\hat{\alpha}_m^*$ and $\hat{\lambda}_m^*$.
4. Compute the bootstrap estimate of the 100pth percentile using $\hat{\alpha}_m^*$ and $\hat{\lambda}_m^*$ as following:

$$\hat{Q}_{p,m}^*(\hat{\alpha}, \hat{\lambda}) = ((1.0 - p)^{-1/\hat{\alpha}_m^*} - 1)^{1/\hat{\lambda}_m^*} \tag{16}$$

Parametric Bootstrap Charts

For the MLE Continued

We repeat Steps 2 through 4 M times to obtain a size M bootstrap sample $\hat{q}^*_{p,1}, \hat{q}^*_{p,2}, ..., \hat{q}^*_{p,M}$, where M is any given large, positive integer.

Then given a FAR,γ, 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

Parameters	$n = 5$		$n = 5$	
	ARL	SERL	ARL	SERL
	$\gamma_0 = 0.1$ (FAR)		$1/\gamma_0 = 10$	
$p = 0.01$	9.0851	0.1250	16.1360	0.2922
$p = 0.05$	9.2691	0.1294	15.9680	0.2886
$p = 0.10$	9.1935	0.1302	15.6528	0.2785
$p = 0.25$	9.3039	0.1285	15.1762	0.2680
	$\gamma_0 = 0.01$ (FAR)		$1/\gamma_0 = 100$	
$p = 0.01$	89.7949	1.5565	43.0204	0.8511
$p = 0.05$	90.1718	1.5683	42.5070	0.8302
$p = 0.10$	91.3064	1.5865	41.5448	0.7988
$p = 0.25$	90.6854	1.5839	39.4974	0.7754
	$\gamma_0 = 0.0027$ (FAR)		$1/\gamma_0 = 370.37$	
$p = 0.01$	336.3652	6.4181	65.0647	1.3018
$p = 0.05$	333.6685	6.5899	64.7746	1.3200
$p = 0.10$	339.3910	6.4750	63.4313	1.2958
$p = 0.25$	340.8538	6.7502	59.6773	1.2202

Upper and Lower Control Limits for the MLE Parametric Bootstrap Chart and the Shewhart-Type Chart

Parameters	$n = 5$		$n = 5$	
	LCL	UCL	LCL	UCL
	$\gamma_0 = 0.1$ (FAR)		$1/\gamma_0 = 10$	
$p = 0.01$	0.0012	0.0106	-0.0003	0.0097
$p = 0.05$	0.0064	0.0546	-0.0013	0.0495
$p = 0.10$	0.0134	0.1127	-0.0022	0.1019
$p = 0.25$	0.0388	0.3166	-0.0018	0.2800
	$\gamma_0 = 0.01$ (FAR)		$1/\gamma_0 = 100$	
$p = 0.01$	0.0007	0.0181	-0.0031	0.0125
$p = 0.05$	0.0036	0.926	-0.0157	0.0639
$p = 0.10$	0.0075	0.1913	-0.0316	0.1314
$p = 0.25$	0.0214	0.5284	-0.0814	0.3597
	$\gamma_0 = 0.0027$ (FAR)		$1/\gamma_0 = 370.37$	
$p = 0.01$	0.0005	0.0232	-0.0044	0.0138
$p = 0.05$	0.0028	0.1185	-0.0222	0.0705
$p = 0.10$	0.0058	0.2452	-0.0451	0.1448
$p = 0.25$	0.0163	0.6761	-0.1177	0.3961

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