International Journal of Advances in Applied Sciences (IJAAS)
Vol. 2, No. 3, September 2013, pp. 141~144
ISSN: 2252-8814

Percentiles of Range: Pareto Type Model

R. Subba Rao[1], G. Prasad[2], R. R. L. Kantam[3]
[1] Shri Vishnu Engineering College for Women, Bhimavaram - 534202, A.P., India
[2] Aditya Degree College, Rajahmundry – 533103, A.P., India
[3] Acharya Nagarjuna University, Guntur - 522510, A.P., India

Article Info

Article history:

Received Nov 11, 2012
Revised Mar 17, 2013
Accepted Mar 28, 2013

Keyword:

Order Statistics
Distribution of Range
Interval estimation
Percentiles
Control limits

ABSTRACT

Statistical quality control relates heavily on the goodness of control chart limits. The more accurate those limits are, the more likely are to detect whether a process is in control. Various procedures have been developed to compute good control limits. This paper proposes construction of Range chart by considering a Pareto distribution of IV kind. The cumulative distribution function of sample range from their distribution is derived. The percentiles of the distribution of range are worked out and are used to construct the control limits. The performance of the control chart is compared with that of gamma based control chart. Interval estimation for the scale parameter is also worked out.

Corresponding Author:

R. Subba Rao,
Shri Vishnu Engineering College for Women,
Bhimavaram - 534202,
A.P., India.
Email: rsr_vishnu@rediffmail.com

1. INTRODUCTION

Industrial statisticians are accustomed to monitoring the stability of the process output through the maintenance of control charts. Use of these control charts assumes that the process output is normally distributed or that appeal to the central limit theorem is plausible. The probability density function of a quality characteristic such as product life, a positive valued random variable is often modeled by distributions such as Gamma, Wei-bull, Pareto, Half Logistic, Inverse Gaussian or Log Normal than by a Normal probability model. In this context we focused a Pareto model of IV kind to develop the control limits of Range chart. Range of a random sample in a known population is a measure of sample dispersion. It is a statistic whose sampling distribution would be useful to study the variations in different sample ranges of any data following the statistical distribution under consideration. Its percentiles with predetermined inclusion probability can be used to develop control charts for the population range as suggested by Shewart for the well known range chart in samples from Normal distribution.

The general formula for the cumulative distribution function of sample range of a random sample of size n from continuous probability model is considered [1]. It is given by

$$G(w) = n\int f(x)[F(x+w) - F(x)]^{n-1}\,dx \qquad ...(1.1)$$

where P(.), p(.) are respectively the cumulative distribution function and probability density function of standard Pareto distribution.

The recent studies of quality control methods for non normal quality variates are: [2]-[10] and the references there in.

The cumulative distribution function of Pareto IV model is

$$P(x,\mu,\sigma,\alpha)=1-\left[1+\left(\frac{x-\mu}{\sigma}\right)\right]^{-\alpha},x\geq\mu \qquad ...(1.2)$$

The probability density function of Pareto distribution type IV with shape parameter α, threshold parameter μ and scale parameter σ is

$$p(x,\mu,\sigma,\alpha)=\frac{\alpha}{\sigma}\left[1+\left(\frac{x-\mu}{\sigma}\right)\right]^{-(\alpha+1)},x\geq\mu,\sigma>0,\alpha>1$$

$$=\alpha\left[1+x\right]^{-(\alpha+1)} \qquad if\ \sigma=1,\mu=0 \qquad ...(1.3)$$

Characteristics of this Distribution: When the location parameter μ is 0, the characteristics of Pareto model are given by

$$\text{Mean} = \frac{\sigma}{\alpha-1}$$

$$\text{Variance} = \frac{\alpha\sigma^2}{(\alpha-1)^2(\alpha-2)}$$

$$\text{Skewness} = 2\frac{(\alpha+1)}{(\alpha-3)}\sqrt{\frac{\alpha-2}{\alpha}}$$

$$\text{Kurtosis} = \frac{3(\alpha-2)(3\alpha^2+\alpha+2)}{\alpha(\alpha-3)(\alpha-4)}$$

2. PERCENTILES OF RANGE CHART

The well known Pareto distribution IV kind considered for deriving the distribution of sample range. On substituting the probability density function and cumulative distribution function p(x), P(x) (i.e.) equation (1.2) and (1.3) in the general cumulative distribution of sample range, (i.e.) equation (1.1), we get

$$F(w)=n\int_{-\infty}^{\infty}p(x)\left[P(x+w)-P(x)\right]dx,$$

$$=n\int_{-\infty}^{\infty}\alpha(1+x)^{-(\alpha+1)}\left[1-(1+x+w)^{-\infty}-1+(1+x)^{-\infty}\right]^{(n-1)}dx,$$

$$=n\int_{-\infty}^{\infty}\alpha(1+x)^{-(\alpha+1)}\left[(1+x)^{-\infty}-(1+x+w)^{-\infty}\right]^{(n-1)}dx.$$

Since the integrand is not analytically tractable, we use Gauss-Lagrange's formula for n points.

$$\int_0^{\infty}e^{-x}f(x)dx=\sum_{i=0}^{n}l_i f(x_i)\ .$$

The numerical integration coefficients are borrowed from Rao and Mitra tables. The percentiles of p(w) are calculated using numerical interpolation technique for n=2(1)10 corresponding to percentiles 0.00135, 0.01, 0.05, 0.1, 0.5, 0.9, 0.90, 0.95, 0.99865. It is known that for a Pareto model α > 1. For the sake of simplicity of calculations, α is considered as 1.2. The selected percentiles are tabulated in Table (1).

Table 1. Percentile of the Pareto distribution with α = 1.2

F(w)	n = 2	n = 3	n = 4	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10
0.00135	0.0006	0.0315	0.1088	0.2093	0.3211	0.4394	0.5623	0.6890	0.8192
0.01	0.0044	0.0899	0.2344	0.3970	0.5667	0.7410	0.9197	1.1029	1.2911
0.05	0.0223	0.2218	0.4734	0.7318	0.9924	1.2565	1.5266	1.8050	2.0941
0.1	0.0453	0.3400	0.6726	1.0046	1.3364	1.6725	2.0178	2.3766	2.7533
0.5	0.2521	1.1623	2.0502	2.9300	3.8296	4.7800	5.8184	6.9902	8.3548
0.9	0.5083	2.2749	4.2181	6.3993	8.9322	12.0364	16.1588	22.2787	33.0267
0.95	0.5445	2.4557	4.6233	7.1339	10.1515	14.0119	19.4651	28.4170	47.5715
0.99	0.5741	2.6097	4.9832	7.8146	11.3345	16.0405	23.1480	36.3010	74.1039
0.99865	0.5806	2.5959	5.0657	7.9747	11.6207	16.5493	24.1223	38.6129	84.5832

3. CONTROL LIMITS FOR RANGE CHART

The control limits for Range chart of this Pareto model are constructed on par with the inclusion probability for range chart suggested by Schewart as follows. We know that, $R = \sigma\, w$ Where w is range of standard Pareto Variant.

$$(i.e.)\, W = Z_{(n)} - Z_{(1)},$$

$$W = \frac{X_{(n)}}{\sigma} - \frac{X_{(1)}}{\sigma},$$

$$\Rightarrow \sigma W = X_{(n)} - X_{(1)},$$

$$= R,$$

$$\Rightarrow \quad R = \sigma W.$$

where $X_{(i)}$ is the i^{th} order statistics in a sample of size n from a Pareto type model. We have to find two constants C_1 and C_2 such that

$$P(C_1 \le \sigma W \le C_2) = 0.9973,$$

$$\Rightarrow P\left[\frac{C_1}{\sigma} \le W \le \frac{C_2}{\sigma}\right] = 0.9973.$$

where $\frac{C_1}{\sigma}, \frac{C_2}{\sigma}$ would respectively be the 0.00135, 0.99865 percentiles of the distribution of W. Hence the constants C_1, C_2 are given by

$$C_1 = W_{0.00135}\sigma \quad , \qquad C_2 = W_{0.99865}\sigma$$

$$C_1 = \frac{W_{0.00135}R}{\alpha_{(n)} - \alpha_{(1)}}, \qquad C_2 = \frac{W_{0.99865}R}{\alpha_{(n)} - \alpha_{(1)}}$$

$$C_1 = D_3^* R \quad , \qquad C_2 = D_4^* R$$

Where $\alpha(i)$ is the i^{th} quantile Pareto Distribution type IV in a sample of size 'n'. We obtain the value of $\alpha(i)$ by equating the c.d.f. F(x) to $\frac{i}{i+1}, i = 1, 2, ...n.$

It can be considered that $\frac{R}{\alpha_{(n)} - \alpha_{(1)}}$ is an estimator of σ, (i.e.) $P\left[D_3^* \overline{R} < W < D_4^* \overline{R}\right] = 0.9975$, where $\overline{R}$ is mean of sample ranges in the place of R. The constants D_3^* and D_4^* depend only on n and the mathematical model of standard Pareto distribution and hence can be calculated a priori outside the data set. Over repeated sub grouping, R can be replaced by the mean of the range $\overline{R}$. Thus the control limits of range chart would be $D_3^* \overline{R}$, $D_4^* \overline{R}$. The calculated values of the D_3^* and D_4^* for n = 2(1)10 are given in table (2).

Table 2. D_3^*, D_4^* Values of Pareto distribution with α = 1.2

n	D_3^*	D_4^*
2	0.0005	0.5297
3	0.0165	1.3635
4	0.0415	1.9340
5	0.0637	2.4262
6	0.0818	2.9614
7	0.0968	3.6459
8	0.1095	4.6957
9	0.1204	6.7491
10	0.1302	13.4399

4. PERFORMANCE COMPARISON

To assess the power of control limits, we have generated 10,000 random samples each of size 2(1)10 form standard Pareto model with shape parameter α = 1.2. The evaluated control limits using the constants of Pareto distribution derived in section 3 and the constants of gamma distribution taken from Kantam and Sriram (2001) are compared [8]. The coverage probability of ranges that have fallen within the control limits of Gamma distribution and those of Pareto distribution are evaluated, which are given in Table (3). It is observed that the coverage probability of Pareto distribution is more than that of Gamma modal. Hence, Pareto modal with respect to range chart is distinct from Gamma model.

Table 3. Coverage probabilities of Pareto type IV and Gamma Models.

Sample size	Coverage Probabilities with in Pareto	Coverage Probabilities with in Gamma
2	0.7495	0.7228
3	0.8743	0.7343
4	0.8833	0.7075
5	0.8633	0.6850
6	0.8483	0.6699
7	0.8397	0.6784
8	0.8031	0.7078
9	0.7762	0.7670
10	0.7703	0.8731

5. CONFIDENCE INTERVAL FOR THE SCALE PARAMETER

The percentiles of range in Pareto population can also be used for computing the confidence interval for the scale parameter of the Pareto distribution. The two sided (or equitailed) interval for the Pareto scale parameter σ is

$$P\left[T_1 < \sigma < T_2\right] = 1 - \alpha$$

The single sided upper (or right tailed) confidence interval for Pareto scale parameter σ is

$$P\left[0 < \sigma < T\right] = 1 - \alpha$$

The single sided lower (or left tailed) confidence interval for Pareto scale parameter σ is

$$P\left[L < \sigma < \infty\right] = 1 - \alpha$$

where T_1, T_2, L, T shall be the suitable percentile of the table (1), multiplied by estimate of σ.

REFERENCES

[1] David, H.A. *Order Statistics, Wiley*. New York. 1981.
[2] Clements, J.A. "Process capability calculations for non-normal distribution", *Quality Process*, Vol/Issue: 24(9). Pp. 95-100, 1989.
[3] Edgeman, R.L."Inverse Gaussian control charts",*Australian Journal of Statistics,*Vol/Issue:31(1). Pp.435-446, 1989.
[4] Sundaraiyer, V.H. "Estimation of process capability index for Inverse Gaussian distribution", *Communications Statistic Theory Meth*., Vol/Issue: 25(10). Pp. 32381-23963, 1996.
[5] Kantam and Rosaiah (a): "Estimation of process capability index for half-logistic distribution", *National Conference on Business Competitiveness and Growth through Quantative Methods*, ISI, Madras, 1998.
[6] Kantam and Rosaiah (b): "Acceptance sampling based on life tests-Inverse Rayleigh distribution", *National Conference on Applied Statistics and Operations Research*. Nagpur University, Nagpur, November. 1998.
[7] R. Subba Rao. "A quality control approach to some inferential procedures in half-logistic distribution", *M.Phil dissertation, Nagarjuna University*. 1998.
[8] Kantam,R.R,L and Sriram,B. "Variable control charts based on gamma distribution", *IAPQR Transaction,* Vol. 26. Pp. 63-78, 2001.
[9] L.L.K.Chan and H.J.Cu. "Skewness correction $\overline{X}$ and R charts for skewed distributions", *Nav Res Logist*. Pp. 555-573, 2003.
[10] Tadikamalla, Pandu R., Popescu, Dana G. "Kurtosis Correction Method for $\overline{X}$ and R control charts for Long-Tailed Symmetrical Distributions", *Published online 17 January 2007 in Wiley Inter Science,* 2007.