Comparisons of Several Pareto Population Means

Lakhana Watthanacheewakul and Prachoom Suwattee

***Abstract*—Analysis of variance (ANOVA), the likelihood ratio test (LRT) and the Kruskal-Wallis test for Pareto populations are investigated. Since Pareto data sets are non-normal distributions, they must be transformed to normal distributions with a constant variance. The power of the ANOVA, likelihood ratio and Kruskal-Wallis tests was compared in a number of different situations and different sample sizes. It was found that the results depended on the location and shape parameters. They are set in three cases; the location parameters are the same and the shape parameters are different, both the location and shape parameters are different, and the shape parameters are the same but the location parameters are different. It seems that the likelihood ratio test was a good choice in almost every case, but it is difficult to find a test based upon the likelihood ratio.**

***Index Terms*—Pareto populations, Power, Shape parameters, Location parameters**

I. Introduction

There are many methods for comparing of more than two population means such as the ANOVA and the LRT. If the data are non-normally distributed, it is usual to apply the likelihood ratio test to scale, location or shape parameters. However, it is difficult to find the exact distribution of the generalized likelihood ratio statistic, Λ. For large samples the statistic $-2\ln\Lambda$ is approximately distributed as chi-squared with k–1 degrees of freedom, whereas for small samples the approximated chi-square may be inaccurate. Alternatively, the Kruskal-Wallis non-parametric test is used and, in case where a comparison of the means of a population with one or more parameters, it is easier to apply the ANOVA. However, in the latter case non-normal data must be transformed to normal distributions so that the required assumptions for the ANOVA fit. In the ANOVA, the usual basic assumptions are that the model is additive and the errors are randomly, independently and normally distributed with zero mean and constant variance. When analyzing data that do not match the assumptions of a conventional method of analysis, there are two choices; transform the data to fit the assumptions or develop some new robust methods of analysis [1]. If a satisfactory transformation can be found, it will almost always be easier to use it rather than to develop a new method of analysis. Transformations are used for three purposes; stabilizing response variance, making the distribution of the response variable closer to a normal distribution and improving the fit of the model to the data [2]. Choosing an appropriate transformation depends on the probability distribution of the sample data, e.g., the square root transformation is used for Poisson data. Furthermore, it is possible to transform the data using a family of transformations already extensively studied over a long period of time, e.g., [3], [4]. A well-known family of transformations often used in previous studies was proposed by Box and Cox. However, Box-Cox transformation is not always applicable. The Box-Cox transformation should be used with caution in some cases such as failure time and survival data [5]. The Box-Cox transformation was not satisfactory even when the best value of transformation parameter had been chosen. From studies of income distribution, [6], it is well established that a high proportion of a population will have low income. Pareto distributions have an important role in these studies. It is a very right long tailed distribution with the shape parameter γ, which determines the concentration of data and the location parameter θ, which is the minimum value of the variable [7]. Besides the distribution of income, the Pareto distribution is indeed common in various fields, e.g., a queuing system [8], mobile communication networks [9]. As mentioned, the ANOVA requires certain basic assumptions. Hence, the Pareto-distributed data must be transformed before the ANOVA is applied. The ANOVA test, the likelihood ratio test and the Kruskal-Wallis test are applied for comparisons of several Pareto population means. In addition, the power of the ANOVA, generalized likelihood ratio and Kruskal-Wallis tests is compared using a simulation study.

II. Tests for Comparisons of Several Means

A. *The ANOVA Test*

Usually, a Box-Cox transformation is used to transform data to normality. For the ANOVA, the Box-Cox transformation is in the form

$$Y_{ij} = \begin{cases} \dfrac{X_{ij}^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln X_{ij}, & \lambda = 0, \end{cases} \tag{1}$$

for $x_{ij} > 0$, where X_{ij} is a random variable in the j^{th} trial from the i^{th} population, Y_{ij} the transformed variable of X_{ij} and λ a transformation parameter, but sometimes this

Manuscript received August 4, 2009.

Lakhana Watthanacheewakul is with the Faculty of Science, Maejo University, Chiang Mai, Thailand (phone: 66-53-873-540; fax: 66-53-878-225; e-mail: lakhanaw@yahoo.com, lakhana@mju.ac.th).

Prachoom Suwattee is with the School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand (e-mail: prachoom@as.nida.ac.th).

transformation when applied to the Pareto-distributed data does not change the data sufficiently so that all the requirements for the ANOVA are met. Moreover, the transformed data are sometimes almost the same value when the location parameter is large; they have a variance close to zero, so that the Kolmogorov-Smirnov (K-S) test for checking normality and the Levene test for checking the homogeneity of variances cannot be performed. In order to cope with these problems, an alternative transformation proposed by Watthanacheewakul and Suwattee [10] is in the form

$$Y_{ij} = \begin{cases} \dfrac{X_{ij}^{\lambda} - (X_{ij} - 0.99\theta_i)^{\lambda}}{\lambda} & ,\lambda \neq 0 \\ \ln\left(\dfrac{X_{ij}}{(X_{ij} - 0.99\theta_i)}\right) & ,\lambda = 0 \end{cases} \tag{2}$$

where X_{ij} is a random variable in the j^{th} trial from the i^{th} Pareto distribution, Y_{ij} the transformed variable of X_{ij}, θ_i the minimum value of X_{ij} from the i^{th} Pareto distribution, and λ a transformation parameter. It is applied to transform any number of sets of Pareto data to normal with equal variance. The null hypothesis is $H_0 : \mu_1 = \mu_2 ... = \mu_k$. It performs better than the Box-Cox transformation for some specific sets of data and so allows the normality and homogeneity of variances to be checked.

B. The Likelihood Ratio Test

For Pareto data, testing the equality of means is equivalent to testing the equality of the shape parameters under the condition that the location parameters are equal.

Theorem If X_{ij} is $\mathrm{Par}(\theta_i, \gamma_i)$, i=1,…,k, then, for large sample, the statistic of the likelihood ratio test for testing the null hypothesis $H_0 : \gamma_1 = \gamma_2 = ... = \gamma_k$ and $\theta_1 = \theta_2 = ... = \theta_k$ is

$$G = -2n\ln n + 2n\ln\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x^*}\right)\right] - 2\left[\frac{n^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x^*}\right)}\right]\ln x^*$$
$$+2\left[\frac{n}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x^*}\right)}\right]\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln x_{ij} - 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[\frac{n_i}{\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x_{i(1)}}\right)}\right]\ln x_{ij}$$
$$+2\sum_{i=1}^{k}\left[n_i\ln\left(\frac{n_i}{\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x_{i(1)}}\right)}\right) + \left[\frac{n_i^2}{\sum_{j=1}^{n_i}\ln\left(\frac{x_{ij}}{x_{i(1)}}\right)}\right]\ln x_{i(1)}\right] \tag{3}$$

where x^* is the minimum value of all observations and $x_{i(1)}$ is the minimum value of each sample.

The test statistic G has an approximate chi-square distribution with k–1 degrees of freedom when H_0 is true.

C. The Kruskal-Wallis Test

This is a non-parametric analogue, based on rank, of one-way analysis of variance. The test statistic is

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{k} n_i\left(\bar{R}_i - \frac{n+1}{2}\right)^2 \tag{4}$$

where $\bar{R}_i$ is the average rank of the member of the i^{th} sample obtained after ranking all of the $n = \sum_{i=1}^{k} n_i$ observations. Kruskal [11] proves that if H_0 is true, the statistic H has a limiting chi-square distribution with k–1 degrees of freedom as $n_i \to \infty$ simultaneously. The null hypothesis is $H_0 : \xi_1 = \xi_2 = ... = \xi_k$ where ξ_i is the median of the i^{th} Pareto distribution, $\xi_i = \dfrac{\theta_i}{0.5^{1/\gamma_i}}$.

III. Example with Real-Life Data

Data on the major rice crop in the crop year 2001/2002 (April 1st, 2001 to March 31st, 2002) from three Tambols, Nonghan, Nongyang and Nongjom, of Amphoe Sansai in the Chiang Mai province, Thailand, were used as populations. Each Tambol consists of a set of 228, 281 and 294 farmer households, respectively. Samples of sizes 23, 28 and 30 were drawn from each Tambol. The sampled data are shown in Table I.

Table I The Major Rice Crop in Kilograms from the three Tambols for the Crop Year 2001/2002 (April 1st, 2001 to March 31st, 2002)

Nonghan	Nongyang	Nongjom
2,300	3,440	3,600
1,800	3,200	5,000
2,000	5,400	7,500
2,400	3,800	7,800
12,000	4,300	3,600
2,800	7,000	4,000
3,000	3,700	4,000
2,250	6,000	4,800
2,100	3,250	4,900
2,700	3,500	4,100
1,800	3,000	4,500
3,500	3,400	4,200
2,500	5,000	6,800
2,600	3,600	4,000
3,200	18,000	7,000
1,900	3,150	4,500
2,200	8,500	8,300
3,500	4,500	3,800
6,000	4,250	24,000
2,300	3,500	5,800
2,850	3,000	3,720
3,700	3,600	6,000
1,900	5,000	5,400
	3,000	3,600
	4,000	4,500
	3,600	4,220
	4,270	3,600
	5,800	4,800
		4,200
		4,113

They are Pareto data and they can be checked using a Pareto probability plot. The value of transformation parameter = 0.4372. The ANOVA assumptions of the transformed data have been checked to be valid. For significance level $\alpha = 0.05$, $F_{2,78} = 3.11$. Since F= 8.4935 >

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

3.11, it is concluded that there is a significant difference in at least one pair among the three population means. Multiple comparisons should now be applied to identify the difference amongst the population means. The value of the statistic G is 9.6764 and H_0 is rejected. There is a significant difference in at least one pair among the three population means. The value of the statistic H is 32.3261 and H_0 is rejected. It is concluded that there is a significant difference in at least one pair among the three population means.

IV. Simulation Studies

Since the power functions of the three tests are in different forms [12]–[14], they cannot be compared directly. To draw a conclusion, the power of the three tests is studied numerically for particular cases by simulation method.

A. Simulation Method

A power comparison of several tests was suggested in two steps [15].

1) To find the critical value for reject on of the null hypothesis, Pareto populations of size N_i = 4,000 are generated for $\theta_1 = \theta_2 = ... = \theta_k$ and $\gamma_1 = \gamma_2 = ... = \gamma_k$, $i = 1, 2,.., k$. From each generated population, 1,000 random samples, each of size n_i, $i = 1, 2,..., k$, are drawn. The test statistics of the three tests are calculated from each of the 1,000 samples. For each test, the 1,000 values of the test statistics are arranged in an increasing order and the 95th percentile is identified. This gives the critical values at $\alpha = 0.05$ for the three tests.

2) Since the proportion of rejections of the null hypothesis when the alternative is true is needed, Pareto populations of size $N_i = 4{,}000$ are generated for various parameter values θ_i and γ_i, $i = 1, 2,.., k$. Since means of Pareto distributions depend on both location parameter, θ_i , and shape parameter, γ_i , $i = 1, 2,.., k$ are set for three cases: the location parameters are the same and the shape parameters are different, both the location and shape parameters are different, and the shape parameters are the same but the location parameters are different. Since there are many values of location parameter and shape parameter, the difference of means of several Pareto distributions is considered. The difference of means is measured by the coefficient of variation (C.V.),

$$\text{C.V.}(\mu) = \frac{\text{S.D.}(\mu_1, \mu_2, ..., \mu_k)}{\text{Mean}(\mu_1, \mu_2, ..., \mu_k)} . \tag{5}$$

To make it clear, the examples of the calculation of the coefficient of variation are illustrated in Appendix.

From each generated $\text{Par}(\theta_i, \gamma_i)$ population, 10,000 random samples, each of size n_i , are drawn. The test statistics of the three tests are calculated. If the value of the test statistic is in the rejection area, as defined by the critical values, then the null hypothesis is rejected. The power of the test is the proportion of times that the null hypothesis is rejected.

Let $\hat{\beta}_F$ be the power of the ANOVA test, $\hat{\beta}_G$ the power of the likelihood ratio test, and $\hat{\beta}_H$ the power of the Kruskal -Wallis test .The values of parameters and the significant value are set as follows:

1) k = number of the populations = 3
2) n_i = sample sizes from the ith Pareto population between 10 and 30, for $i = 1, 2,..., k$
3) θ_i , the location parameter of the ith Pareto population, is between 1,030 and 3,253
4) γ_i , shape parameter of the *ith* Pareto population, is between 1.72 and 50 and fixed by the location parameter and the coefficient of variation of population means
5) $\alpha = 0.05$.

B. Results of the Power Comparison

For a fixed null hypothesis, the power of the three tests is obtained as the proportion of rejection when the alternative hypothesis is true with different values of the coefficients of variation of population means. When the coefficient of variation of population means is zero, the null hypothesis is true. Hence, the proportion of rejecting the null hypothesis is equal to the level of significance α . When the coefficient of variation of population means is greater than zero, the null hypothesis is false or the alternative hypothesis is true. Thus the proportion of rejection of the null hypothesis is the estimate of the power of the test. The results of the power of tests for different sample sizes and different coefficients of variation of population means are compared when α is set equal to 0.05.

The results of comparison are divided into three cases.

1) Same Location Parameters but Different Shape Parameters

From the values of location parameters and shape parameters as Table A.I in Appendix, and samples of sizes 10 to 30, the power of the three tests is shown in Table II-Table V.

Table II Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 10

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0495	0.0521	0.0504
0.10	0.0651	0.1248	0.0962
0.20	0.2191	0.4705	0.3347
0.30	0.5082	0.8139	0.6012
0.40	0.9778	0.9999	0.9806
0.50	0.9965	1.0000	0.9946

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table III Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 20

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0486	0.0548	0.0526
0.10	0.1092	0.1527	0.1199
0.20	0.5350	0.7128	0.5605
0.30	0.9121	0.9848	0.9231
0.40	1.0000	1.0000	1.0000
0.50	1.0000	1.0000	1.0000

Table IV Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 30

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0539	0.0514	0.0490
0.10	0.2285	0.2982	0.2589
0.20	0.7479	0.9196	0.7823
0.30	0.9816	0.9984	0.9733
0.40	1.0000	1.0000	1.0000
0.50	1.0000	1.0000	1.0000

Table V Power of the Tests When the C.V. of Population Means Varies with Unequal Sample Sizes of $n_1 = 10$, $n_2 = 20$, $n_3 = 30$

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0522	0.0527	0.0524
0.10	0.0694	0.2494	0.1869
0.20	0.3966	0.7908	0.5650
0.30	0.5473	0.8988	0.7124
0.40	0.9886	1.0000	0.9922
0.50	0.9996	1.0000	0.9998

In this case, if the differences among the population means are small, the likelihood ratio test has the highest power for samples of sizes 10 to 30. The power of all three tests is almost the same, if the differences among the population means are large.

2) Different Location Parameters and Different Shape Parameters

From the values of location parameters and shape parameters as Table A.II in Appendix, and samples of sizes 10 to 30, the power of the three tests is shown in Table VI-Table IX.

Table VI Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 10

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0488	0.0516	0.0476
0.10	0.1833	0.1979	0.1040
0.20	0.3580	0.4109	0.1960
0.30	0.6475	0.8362	0.5171
0.40	0.9959	0.9999	0.9361
0.50	0.9998	1.0000	0.9665

Table VII Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 20

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0519	0.0495	0.0491
0.10	0.3414	0.2971	0.1375
0.20	0.7485	0.6357	0.3956
0.30	0.9915	0.9808	0.8123
0.40	1.0000	1.0000	0.9994
0.50	1.0000	1.0000	0.9999

Table VIII Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 30

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0516	0.0508	0.0493
0.10	0.6738	0.5525	0.2236
0.20	0.9041	0.8888	0.4601
0.30	0.9994	0.9988	0.9406
0.40	1.0000	1.0000	1.0000
0.50	1.0000	1.0000	1.0000

Table IX Power of the Tests When the C.V. of Population Means Varies with Unequal Sample Sizes of $n_1 = 10$, $n_2 = 20$, $n_3 = 30$

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0543	0.0524	0.0472
0.10	0.4065	0.4091	0.1905
0.20	0.4727	0.5476	0.2209
0.30	0.6868	0.8565	0.4820
0.40	0.9953	0.9988	0.9534
0.50	0.9992	0.9997	0.9575

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

In this case, if the differences among the population means are small, the likelihood ratio test has the highest power for the simulation with equal sample sizes of n=10 and unequal sample sizes but the power of ANOVA test is higher than that of the likelihood ratio test for the simulation with equal sample sizes of n=20 and n=30. However, all the three tests have almost the same power when the differences among the population means are large.

3) Different Location Parameters and Same Shape Parameters

From the values of location parameters and shape parameters as Table A.III in Appendix, and samples of sizes 10 to 30, the power of the three tests is shown in Table X-Table XIII.

Table X Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 10

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0514	0.0506	0.0484
0.10	0.1093	0.1165	0.3725
0.20	0.2939	0.3922	0.8705
0.30	0.4328	0.6215	0.9856
0.40	0.5163	0.8921	0.9993
0.50	0.5403	0.9796	0.9997

Table XI Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 20

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0480	0.0495	0.0527
0.10	0.1770	0.2100	0.7421
0.20	0.4243	0.7299	0.9973
0.30	0.6285	0.9317	1.0000
0.40	0.6980	0.9990	1.0000
0.50	0.7667	1.0000	1.0000

Table XII Power of the Tests When the C.V. of Population Means Varies with Equal Sample Sizes of n = 30

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0498	0.0524	0.0506
0.10	0.1954	0.4780	0.9171
0.20	0.5677	0.9044	0.9999
0.30	0.7025	0.9978	1.0000
0.40	0.8368	1.0000	1.0000
0.50	0.8896	1.0000	1.0000

Table XIII Power of the Tests When the C.V. of Population Means Varies with Unequal Sample Sizes of $n_1 = 10$, $n_2 = 20$, $n_3 = 30$

C.V. of Population Means	Test		
	$\hat{\beta}_F$	$\hat{\beta}_G$	$\hat{\beta}_H$
0.00	0.0507	0.0518	0.0491
0.10	0.0741	0.1622	0.4961
0.20	0.0947	0.4707	0.9243
0.30	0.2964	0.7885	0.9982
0.40	0.3419	0.9280	0.9999
0.50	0.5741	0.9854	1.0000

In this case, if the differences among the population means are small, the Kruskal-Wallis test has the highest power. If the differences among the population means are large, the likelihood ratio test and the Kruskal-Wallis test have almost the same power and higher power than ANOVA test.

V. Conclusion

To test the equality of means of Pareto data, ANOVA test, likelihood ratio test and Kruskal-Wallis test are investigated. The use of ANOVA test for testing equality of means for several Pareto distributions, transformed data to normal with constant variances is needed. The data sets transformed by an alternative transformation meet the assumptions required for the application of ANOVA test. The power of ANOVA test is compared to those of the other two existing tests, the likelihood ratio and the Kruskal-Wallis tests. The power functions of the three tests are of different forms and cannot be compared explicitly. A numerical method is then used for comparison purposes. It is found that if the location parameters are the same and the shape parameters are different, the likelihood ratio test has the highest power for samples of sizes 10 to 30. The power of all three tests is almost the same if the differences among the population means are large.

If the populations have different values of both location and shape parameters, the likelihood ratio test has the highest power for the simulations with equal samples of sizes 10 and unequal sample sizes but ANOVA test is higher than the likelihood ratio test a little for the simulations with equal samples of sizes 20 and 30. However, all the three tests have almost the same power when the differences among the population means are large.

If the populations have the same shape parameters but different location parameters, the Kruskal-Wallis test has the highest power. In this case, as the differences among the population means are large, the likelihood ratio test and Kruskal-Wallis test have the same power and higher power than ANOVA test.

To choose the appropriate test, the location parameter and the shape parameter for each group should be checked. For the same location parameter but the different shape parameter, sample sizes do not affect on choosing the test. If the differences among the population means are small, the likelihood ratio test should be selected. If the differences among the population means are large, all three tests can be selected.

For both the different location parameter and shape parameter, samples of sizes 10 and unequal sample sizes, if the differences among the population means are small, the likelihood ratio test should be selected. If the differences among the population means are large, anyone can be selected. For samples of sizes 20 and 30, if the differences among the population means are small, the ANOVA test should be selected. If the differences among the population means are large, all three tests can be selected.

For the different location parameter and the same shape parameter, sample sizes do not affect on choosing the test. If the differences among the population means are small, the Kruskal-Wallis test should be selected. If the differences among the population means are large, the Kruskal-Wallis test or the likelihood ratio test can be selected.

It seems that the likelihood ratio test is a good choice in almost every case. However, it is difficult to find a test based on the likelihood ratio.

APPENDIX

Table A.I Calculation of the C.V. of Population Means for the Same Location and Different Shape Parameters with k=3

Location parameters			Shape parameters			Mean	S.D.	C.V
θ_1	θ_2	θ_3	γ_1	γ_2	γ_3			
1800	1800	1800	3.00	3.01	3.02	2695.54	4.45	0.00
1800	1800	1800	2.70	3.30	4.33	2593.99	259.33	0.10
1800	1800	1800	2.49	2.18	5.31	2849.02	569.93	0.20
1800	1800	1800	2.10	6.00	7.81	2553.56	766.03	0.30
1800	1800	1800	2.01	20.0	25.2	2450.43	980.17	0.40
1800	1800	1800	1.78	29.5	50.0	2600.56	1300.17	0.50

Table A.II Calculation of the C.V. of Population Means for the Different Location and Different Shape Parameters with k=3

Location parameters			Shape parameters			Mean	S.D.	C.V
θ_1	θ_2	θ_3	γ_1	γ_2	γ_3			
1800	1850	1900	2.99	3.17	3.36	2703.86	0.61	0.00
1800	1850	1900	3.00	3.00	6.00	2585.00	266.79	0.10
1800	1850	1900	2.20	3.50	6.50	2711.82	537.72	0.20
1800	1850	1900	2.10	8.00	9.00	2562.72	756.69	0.30
1800	1850	1900	1.95	25.0	30.0	2529.11	1009.64	0.40
1800	1850	1900	1.72	28.0	40.0	2722.41	1366.31	0.50

Table A.III Calculation of the C.V. of Population Means for the Different Location and Same Shape Parameters with k=3

Location parameters			Shape parameters			Mean	S.D.	C.V
θ_1	θ_2	θ_3	γ_1	γ_2	γ_3			
1800	1801	1802	3	3	3	2701.50	1.50	0.00
1642	1900	2000	3	3	3	2771.00	277.08	0.10
1350	1900	2000	3	3	3	2625.00	525.00	0.20
1300	2006	2438	3	3	3	2872.00	861.71	0.30
1044	2115	2500	3	3	3	2829.50	1131.68	0.40
1030	2500	3253	3	3	3	3391.50	1695.91	0.50

REFERENCES

[1] W. Tukey, "On the comparative anatomy of transformations," *Annals of Mathematical Statistics*, vol. 23, 1957, pp. 525-540.
[2] D. C. Montgomery, *Design and Analysis of Experiments*, 5th ed. New York: Wiley, 2001, pp. 590.
[3] G. E. P. Box and D. R. Cox, "An analysis of transformations (with discussion)," *Journal of the Royal Statistical Society, Ser.B.* vol. 26, 1964, pp.211-252.
[4] J. A. John and N. R. Draper, "An alternative family of transformations," *Applied Statistics*, vol. 29(2), 1980, pp.190-197.
[5] K. A. Doksum, and C. Wong, "Statistical tests based on transformed data," *Journal of the American Statistical Association*, vol. 78, 1983, pp. 411-417.
[6] P. R. Fisk, "The graduation of income distributions," *Econometric*, vol. 29(2), 1961, pp.171-185.
[7] N. L. Johnson, and S. Kotz, *Continuous Univariate Distributions*. New York: Wiley, 1970.
[8] C. M. Harris, "The Pareto distribution as a queue service discipline," *Operations Research*, vol. 16, 1968, pp. 307-313.
[9] M. Chatterjee, S. K. Das, "Optimal MAC state switching for CDMA 2000 networks," *Paper Research in Wireless Mobility and Networking*, University of Texas at Arlington, 2004.
[10] L. Watthanacheewakul and P. Suwattee, "Analysis of variance with Pareto data," *Proceedings of the 11th Annual Conference of Asia Pacific Decision Sciences Institute*, Hong Kong, 2006, pp. 346-349.
[11] W. H. Kruskal, "A Nonparametric Test for the Several Sample Problem," *The Annals of Mathematical Statistics*, vol. 23(4), 1952, pp. 525-540.
[12] P. B. Patnaik, "The non-central χ^2 and F-distributions and their Applications," *Biometrika*, vol. 36(2), 1949, pp. 202-232.
[13] E. S. Pearson and H. O. Hartley, "Charts of the power function for analysis of variance tests derived from the non-central F-distribution," *Biometrika*, vol. 38(2), 1951, pp.112-130.
[14] H. O. Posten and R.E. Bargemann, "Power of the likelihood ratio test of the general linear hypothesis in multivariate analysis," *Bimometrika*, vol. 51(3), 1964, pp. 467- 480.
[15] D. Karlis and E.Xekalaki, "A simulation comparison of several procedures for testing the Poisson assumption," *The Statistician*, vol. 49, 2000, pp. 365-382.

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)