

Journal of Applied Statistical Science

Volume 19, Number 3, 2011

Table of Contents

Error Analysis on The Quasi-Negative Binomial Distribution <i>Felix Famoye and Adeyinka Ogunsanya</i>	245
A New Family of Skew Slash Logistic Distributions <i>Bindu Punathumparambath and Sebastian George</i>	253
On the Moments of Sample Covariance from a Bivariate Normal Population <i>Anwar H. Joarder and M. Hafidz Omar</i>	261
Bayesian Identification of Two Dimensional Moving Average Processes <i>Samir M. Shaarawy, Mohammad S. Albassam and Sherif S. Ali</i>	273
On the Use of Moment Estimators to Test for a Zero Inflated Poisson (ZIP) Distribution <i>A.Nathakumar and Z. Govindarajulu</i>	289
Optimal Designing of a Variables Quick Switching Sampling System by Minimzing the Averde Sample Number <i>S. Balamurali and M. Usha</i>	295
Characteristics of BIBD and Complementary BIBD in Terms of Triplets of Treatments <i>S. Mandal, C. Zhu, D. K. Ghosh and S. C. Bagui</i>	311

Bayesian Inference for an Impatient M M 1 Queue with Balking	317
<i>Paul R. Savariappan, P. Chandrasekhar and Ambily Jose</i>	
Interval Estimation for Weibull Distribution Based on a Sample of Size One	325
<i>S.V. Sabnis and A.D. Dharmadhikari</i>	

Journal of Applied Statistical Science

The *Journal of Applied Statistical Science* is a peer-reviewed journal that publishes original articles that make significant contributions to the advancement of statistical science. Articles include theory oriented to real life problems, comparison of existing methods, reviews of special topics, etc.

EDITOR-IN-CHIEF

M. Ahsanullah

Department of Management Sciences
Rider University
Lawrenceville, New Jersey, 08648-3099, U.S.A.
Telephone: 1 (609) 895-5530
Fax: 1 (609) 896-5304
E-Mail: ahsan@rider.edu

Journal of Applied Statistical Science
is published quarterly by

Nova Science Publishers, Inc.

400 Oser Avenue, Suite 1600
Hauppauge, New York, 11788-3619, U.S.
Telephone: (631) 231-7269
Fax: (631) 231-8175
E-mail: nova.main@novapublishers.com
Web: www.novapublishers.com

ISSN: 1067-5817

Subscription Price per Volume (2012)

Print: \$695 Electronic: \$695 Combined Print + Electronic: \$1042

Additional color graphics might be available in the e-version of this journal.

Copyright © 2012 by Nova Science Publishers, Inc. All rights reserved. Printed in the United States of America. No part of this Journal may be reproduced, stored in a retrieval system, or transmitted in any form or by any means: electronic, electrostatic, magnetic, tape, mechanical, photocopying, recording, or otherwise without permission from the Publisher. The Publisher assumes no responsibility for any statements of fact or opinion expressed in the published papers.

EDITOR-IN-CHIEF

M. Ahsanullah

Department of Management Sciences
Rider University
Lawrenceville, New Jersey, 08648-3099, U.S.A.
Telephone: 1 (609) 895-5530 Fax: 1 (609) 896-5304
E-Mail: ahsan@rider.edu

ASSOCIATE EDITORS

Rosihan M. Ali, School of Mathematical Sciences, Univ. of Sains, Malaysia, Penang, Malaysia
E. A. A. Aly, Department of Statistics and Operations Research, Kuwait University, Safat, Kuwait
A. A. Alzaid, Dept. of Statistics and Operations Research, King Saud Univ., Riyadh, Saudi Arabia
B. C. Arnold, Statistics Department, University of California, Riverside, USA
S. C. Bagui, Dept. of Mathematics & Statistics, University of West Florida, Pensacola, USA
I. G. Bairamov, Department of Mathematics, Izmir University of Economics, Izmir, Turkey
N. Balakrishnan, Dept. of Mathematics and Statistics, McMaster Univ., Hamilton, ON, Canada
N. K. Bansal, Dept. of Mathematics, Stat. & Comp. Sci., Marquette Univ., Milwaukee, WI, USA
M. I. Bhatti, Department of Op. Mg & B-Stat, SQ University, Muscat, Oman
W. K. T. Fung, Department of Statistics, The University of Hong Kong, Hong Kong
J. Galambos, Department of Mathematics, Temple University, Philadelphia, PA, USA
J. K. Ghosh, Stat. Math. Div., Indian Statistical Institute, Barrackpore, INDIA
A. K. Gupta, Dept. of Math and Stat., Bowling Green State Univ., Bowling Green, OH, USA
R. C. Gupta, Department of Mathematics, University of Maine, Orono, USA
G. G. Hamedani, Dept. of Mathematics, Stat. and Comp. Sci., Marquette Univ., Milwaukee, USA
M. Hanif, National College of Business Administration and Economics, Gulberg Lahore, Pakistan
S. Huda, Department of Statistics and Operations Research, Kuwait University, Safat, Kuwait
S. N. U. A. Kirmani, Dept. of Math and Comp. Sci., Univ. of Northern Iowa, Cedar Falls, USA
J. Laüter, Inst. of Biometrics and Informatics, Magdeburg University, Germany
J. S. Mehta, Department of Mathematics, Temple University, Philadelphia, PA, USA
V. B. Nevzorov, Dept. of Mathematics and Mechanics, St. Petersburg State University, Russia
A. G. Pakes, Mathematics Department, University of Western Australia, Nedlands, Australia
P. R. Parthasarathy, Department of Mathematics, Indian Institute of Technology, Madras, India
I. Rahimov, Department of Mathematics, KFUPM, Dhahran, Saudi Arabia
M. Z. Raqab, Department of Mathematics, University of Jordan, Amann, Jordan
Dilip Roy, Department of Business Administration, Burdwan University, India
P. K. Sen, Biostatistics and Statistics Department, University of North Carolina, Chapel Hill, USA
B. K. Sinha, Mathematics and Statistics Department, University of Maryland, Baltimore, USA
J. Stoyanov, Institute of Mathematics, University of New Castle upon Tyne, U. K.
C. P. Tsokos, Department of Mathematics & Statistics, University of South Florida, USA
W. W. S. Wei, Department of Statistics, Temple University, Philadelphia, PA, USA
J. Wesolowski, Mathematical Institute, Warsaw University of Technology, Poland

ERROR ANALYSIS ON THE QUASI-NEGATIVE BINOMIAL DISTRIBUTION

Felix Famoye¹ and Adeyinka Ogunsanya²

¹Department of Mathematics, Central Michigan University,
Mt. Pleasant, Michigan 48859

²Department of Mathematics, University of Lagos, Akoka-Yaba, Lagos, Nigeria

Abstract

The quasi-negative binomial distribution, characterized by three parameters, has been used to fit data arising from many fields of study. The distribution gets truncated under certain conditions. We investigate the truncation error by carrying out a detailed error analysis. Through this analysis, we determine the parameter space when the model can be used instead of using a truncated quasi-negative binomial distribution. We define a truncated quasi-negative binomial distribution that can be used when the quasi-negative binomial distribution gets truncated on the right hand side. A numerical example is used for illustration.

Keywords and phrases: Truncation error; parameter space; maximum likelihood estimates.

1. Introduction

A discrete random variable X is said to have a quasi-negative binomial distribution (QNBD) if its probability function, (Li et al. 2008), is given by

$$f(x) = f_x = \begin{cases} \frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)} \frac{1}{1+cx} \left(\frac{1+cx}{1+b+cx} \right)^x \left(\frac{b}{1+b+cx} \right)^\alpha, & x = 0, 1, 2, \dots \\ 0, & \text{for } x > k \text{ if } c < 0, \end{cases} \quad (1.1)$$

where $\alpha > 0$, $b > 0$ and k is the largest positive integer for which $1 + kc \geq 0$ when $c < 0$. The QNBD in (1.1) reduces to the negative binomial distribution (NBD) when the parameter $c = 0$. For the non-truncated QNBD, the mean and the variance as provided by Li et al. (2008) are not in closed form. The probability model in (1.1) gets truncated on the right hand side when $c < 0$. If the truncation error is negligible, one may still be able to use the model in (1.1). However, if the truncation error is substantial, it is important to consider a truncated quasi-negative binomial distribution as an alternative. This paper will study the parameter space of QNBD to identify where the truncation error is negligible. The paper will also define

a truncated quasi-negative binomial distribution that can be used when the truncation error is not negligible.

In section 2, we conduct a detailed error analysis of the QNBD. In section 3, we define a truncated quasi-negative binomial distribution. In section 4, a numerical data set on the number of absenteeism among shift-workers is used for illustration. Finally in section 5, we provide some concluding remarks.

2. Error Analysis of the QNBD

Due to the truncation in QNBD, the sum of the probabilities in (1.1) may differ from unity. The difference between 1 and the sum of the probabilities ($\sum f_x$) is the truncation error. The percentage truncation error will be computed as $100(1 - \sum f_x)$. Some illustrative examples for $k \leq 3$ are presented in Table 1. For two classes only, the truncation leads to only two probabilities f_0 and f_1 . The sum of the two probabilities could be very small or very large as shown in Table 1. In Table 1, the sum of the non-negative probabilities could be larger than 3.0 and could be smaller than 0.2. It is hard to see a pattern in the truncation errors.

One of the motivations for this paper is to determine various conditions for which the truncation error will be negligible. For example in row 5 for $k = 2$, the sum of the probabilities is more than 3 because $f(2)$ leads to a value that is close to 3.0 on the account that $1 + 2c$ is 0.05.

When $c < 0$, we require that $1 + cx \geq 0$ in the QNBD in (1.1). If this is not satisfied, we set f_x to 0 as shown in (1.1). Thus, the largest x value can be obtained from $0 \leq 1 + cx \Rightarrow -cx \leq 1 \Rightarrow x \leq -1/c$ since $-c > 0$. Thus, the largest x value, k , is given by the integer part of $-1/c$.

Through computation, we make a detailed error analysis on the QNBD model when $c < 0$. For this analysis, we consider the values of α and b in the parameter space and the values of $c < 0$. Thus, we consider the values of $\alpha > 0$, $b > 0$ and $c < 0$. In the analysis, we compute the values of $f(x)$ for $x = 0, 1, 2, \dots, k$ where k is such that $k \leq -1/c$ when $c < 0$. After computing the probabilities, we compute the percentage truncation error by using the formula $100(1 - \sum f_x)$.

In fitting the QNBD to an observed data set, the three parameters α , b and c have to be estimated. In order to have at least 1 degree of freedom for the chi-square goodness-of-fit test, we need at least five non-zero probability classes. Thus, we need the smallest value of x to be 4. Therefore, in the error analysis, the smallest x value is required to be 4. The percentage error of truncation will be said to be negligible if it is below 0.5%. Thus, the difference between 1 and the sum of all non-negative probabilities is below 0.005. This was the value used by Consul and Shoukri (1985) in their error analysis for the generalized Poisson distribution. In view of this, we will conduct the error analysis for $k \geq 4$.

The maximum truncation error for the different values of α , b and c are provided in Table 2. Since we need to have at least five non-zero probability classes, we examine the different errors for cases where x is at least 4. In the error analysis we consider the values of $\alpha = (.01)(.01)(5.0)$, $b = (0.1)(0.1)(10.0)$ and $c = (-0.25)(0.01)(-0.01)$.

Table 1. Sum of probabilities for some QNBD parameter sets

k	Parameters			Probabilities				Σf_x
	α	b	c	f_0	f_1	f_2	f_3	
1	1.0	0.1	-0.90	0.0909	2.5000			2.5909
	1.0	0.1	-0.75	0.0909	0.8163			0.9072
	1.0	0.1	-0.60	0.0909	0.4000			0.4909
	1.0	0.5	-0.90	0.3333	1.3889			1.7222
	1.0	0.5	-0.75	0.3333	0.8889			1.2222
	1.0	0.5	-0.60	0.3333	0.6173			0.9506
	1.5	0.1	-0.90	0.0274	2.6517			2.6791
	1.5	0.1	-0.75	0.0274	0.6545			0.6819
	1.5	0.1	-0.60	0.0274	0.2683			0.2957
	1.5	0.5	-0.90	0.1925	1.9018			2.0943
	1.5	0.5	-0.75	0.1925	1.0887			1.2812
	1.5	0.5	-0.60	0.1925	0.6901			0.8826
	1.5	1.0	-0.90	0.3536	1.1820			1.5356
	1.5	1.0	-0.75	0.3536	0.8587			1.2123
	1.5	1.0	-0.60	0.3536	0.6468			1.0004
	2.0	0.1	-0.90	0.0083	2.5000			2.5083
	2.0	0.1	-0.75	0.0083	0.4665			0.4748
	2.0	0.1	-0.60	0.0083	0.1600			0.1683
	2.0	0.5	-0.90	0.1111	2.3148			2.4259
	2.0	0.5	-0.75	0.1111	1.1852			1.2963
	2.0	0.5	-0.60	0.1111	0.6859			0.7970
	2.0	1.0	-0.90	0.2500	1.5026			1.7526
	2.0	1.0	-0.75	0.2500	1.0240			1.2740
	2.0	1.0	-0.60	0.2500	0.7289			0.9789
2	1.0	0.1	-.475	0.0909	0.2560	1.4815		1.8284
	1.0	0.1	-.45	0.0909	0.2367	1.2500		1.5776
	1.5	0.1	-.475	0.0274	0.1536	2.2680		2.4490
	1.5	0.1	-.45	0.0274	0.1392	1.6573		1.8239
	2.0	0.1	-.475	0.0083	0.0819	2.9630		3.0532
	2.0	0.1	-.45	0.0083	0.0728	1.8750		1.9561
3	1.0	0.1	-.33	0.0909	0.1687	0.3991	0.0683	0.7270
	1.0	0.1	-.30	0.0909	0.1563	0.3200	0.6250	1.1922
	1.5	0.1	-.33	0.0274	0.0912	0.3568	0.1425	0.6179
	1.5	0.1	-.30	0.0274	0.0829	0.2683	0.9667	1.3453
	2.0	0.1	-.33	0.0083	0.0438	0.2721	0.2484	0.5726
	2.0	0.1	-.30	0.0083	0.0391	0.1920	1.2500	1.4894

In Table 2, we report the ranges for the parameters that produce the maximum percentage error in the sum of the non-zero probabilities. In the error analysis the value of k is taken to be the integer part of $-1/c$. We also report the specific parameter values at which the maximum truncation error occurs. For example, when $0.01 < \alpha \leq 0.50$, $0.1 \leq b \leq 10.0$ and $-0.18 \leq c \leq -0.01$, the maximum truncation error with at least 5 non-zero probability classes is -0.4834 . When $0.1 < b < 1.0$, the maximum truncation error in the probabilities occurs at the maximum value of α in the interval and the minimum value of $b = 0.1$. As α values increase, the range of c values decreases in order to have a maximum truncation error of less than 0.5%. For fixed α , as the number of non-zero probability classes increases, the truncation error decreases.

Table 2. Maximum percentage error and the corresponding parameter values

Range of parameters values			% Truncation error (α, b, c)	k
α	b	c		
[0.01,0.50]	[0.10,1.0]	[-0.18,0)	-0.4834 (0.50, 0.1, -0.17)	5
[0.51,1.00]	[0.10,1.0]	[-0.15,0)	-0.3314 (1.00, 0.1, -0.15)	6
[1.01,1.50]	[0.10,1.0]	[-0.14,0)	-0.2209 (1.50, 0.1, -0.11)	9
[1.51,3.00]	[0.10,1.0]	[-0.10,0)	-0.4571 (3.00, 0.1, -0.10)	10
[3.01,6.50]	[0.10,1.0]	[-0.08,0)	-0.3695 (5.06, 0.1, -0.08)	12
[0.01,1.00]	[1.01,10]	[-0.25,0)	-0.0268 (1.00, 1.01, -.25)	4
[1.01,2.00]	[1.01,10]	[-0.25,0)	-0.1347 (2.00, 1.01, -.25)	4
[2.01,3.00]	[1.01,10]	[-0.25,0)	-0.3663 (0.39, 1.01, -.25)	4
[3.01,3.41]	[1.01,10]	[-0.25,0)	-0.4977 (0.36, 1.01, -.25)	4

When $c < 0$ and $k \geq 4$, the QNBD can be used in general when $0.01 < \alpha \leq 3.0$ for any value of $b > 1$. If $b < 1$ and $.01 < \alpha < 3$, the range of c values decreases to $-.18 \leq c < 0$. When $b < 1$ and $k \geq 4$, the QNBD can be used in general when $.01 < \alpha \leq 6.5$ and $-0.08 \leq c < 0$.

3. Truncated Quasi-Negative Binomial Distribution

A truncated quasi-negative binomial distribution (TQNBD) is defined as

$$P(X = x) = P_x = f(x) / Q_k, x = 0, 1, 2, \dots, k, \quad (2.1)$$

and zero otherwise, where $f(x)$ is given by (1.1), $Q_k = \sum_{x=0}^k f(x)$, $\alpha > 0$, $b > 0$ and $-\infty < c < \infty$. The model in (2.1) is truncated on the right hand side. One may define a model that is truncated on the left hand side or truncated on both sides. However, these other truncation types will not be considered in this paper. The quantity k in (2.1) is any positive integer less than or equal to the largest observed value of x such that $1 + ck \geq 0$ when $c < 0$. We now consider the different cases for the values of c in the parameter space.

(i) When $c < -1$:

The condition $1 + ck \geq 0$ is satisfied only when $k = 0$. This case leads to a degenerate probability function as the whole probability mass is concentrated at the point $x = 0$.

(ii) When $-1 \leq c < -0.5$:

The maximum value that k can have is 1 and so the TQNBD reduces to the Bernoulli distribution with $P_0 = \left[1 + \frac{\alpha}{1+b+c} \left(\frac{1+b}{1+b+c} \right)^\alpha \right]^{-1}$ and $P_1 = \frac{\alpha}{1+b+c} \left[\frac{\alpha}{1+b+c} + \left(\frac{1+b+c}{1+b} \right)^\alpha \right]^{-1}$.

(iii) When $-0.5 \leq c < 0$:

The value of k can be any positive integer $\leq -1/c$. The largest observed value of x in a random sample is usually taken as an estimate of k .

(iv) When $c \geq 0$:

The quantity $f(x)$ is positive for all values of x and so the largest value of k is $+\infty$. By using Theorem 3 of Li et al. (2008), $Q_k = 1$ for $k = \infty$. Hence, $f(x)$, for $x = 0, 1, 2, \dots$ in (2.1) will provide a true probability distribution.

For any positive integer k , the probability distribution in (2.1) represents a TQNBD and it is a true probability model for any value of c . The first two probabilities of the TQNBD are $P_0 = f_0 / Q_k = (b / (1+b))^\alpha / Q_k$ and $P_1 = f_1 / Q_k = \alpha b^\alpha / [(1+b+c)^{\alpha+1} Q_k]$. The ratio of the first two probabilities is $P_1 / P_0 = \alpha(1+b)^\alpha (1+b+c)^{-\alpha-1}$. The population mean μ_k and the population variance σ_k^2 of the TQNBD can be expressed as

$$\mu_k = E(X) = Q_k^{-1} \sum_{x=0}^k x f(x) \text{ and } \sigma_k^2 = E(X - \mu_k)^2 = Q_k^{-1} \sum_{x=0}^k x^2 f(x) - \mu_k^2. \quad (2.2)$$

The mean and variance in (2.2) are not in closed forms and hence it is difficult to determine the parameter region where the variance is less than (or larger than) the mean. Based on (2.2) and the ratio P_1 / P_0 , it is not easy to estimate the TQNBD parameters by using the first two moments and the ratio of the first two frequencies. Suppose a random sample of size n is taken from the TQNBD in (2.1) and let the observed frequencies be denoted by n_x , $x = 0, 1, 2, \dots, k$ for the different classes so that $\sum_{x=0}^k n_x = n$, where k is the largest observed value of x .

The log-likelihood function of the TQNBD is given by

$$\begin{aligned} \log L = \sum_{x=0}^k n_x \{ & \log(\Gamma(x+\alpha) / (x! \Gamma(\alpha))) - \log(1+cx) + x \log((1+cx) / (1+b+cx)) \\ & + \alpha \log(b / (1+b+cx)) - \log Q_k \}. \end{aligned} \quad (2.3)$$

The likelihood equations can be obtained by taking the partial derivatives of (2.3) with respect to the parameters α , b and c . The likelihood equations have to be solved numerically either by the Newton-Raphson method or some other iterative technique. We have not been able to show if the solutions to the likelihood equations are unique. The initial estimates for the parameters of TQNBD can be taken as the final estimates when a non-truncated QNBD is fitted to the data. We now describe the initial estimates for finding the maximum likelihood estimates (MLE) of the non-truncated QNBD.

Since the QNBD in (1.1) reduces to the negative binomial distribution (NBD) when $c = 0$, we find the moment estimates of α and b when $c = 0$. These estimates with $c = 0$ can be used as the initial estimates for finding the MLE of QNBD parameters. Alternatively, one can take the moment estimates of α and b plus the estimate of c from equating the QNBD

probability of 1 to the sample proportion of 1. This initial estimate of c is given by $\tilde{c} = (n\tilde{\alpha}\tilde{b}^{\tilde{\alpha}} / n_1)^{1/(1+\tilde{\alpha})} - 1 - \tilde{b}$, where $\tilde{\alpha}$ and $\tilde{b}$ are the moment estimates from the NBD.

Table 3. Number of absenteeism among shift-workers

Count	Observed frequency	Expected frequency			
		NBD	GNBD (GO ^a)	QNBD	TQNBD
0	7	11.13	9.23	12.40	10.10
1	16	15.74	16.18	15.92	16.76
2	23	17.77	19.86	17.25	19.40
3	20	18.36	21.06	17.52	19.60
4	23	18.10	20.50	17.19	18.60
5	24	17.32	18.78	16.50	17.12
6	12	16.24	16.46	15.59	15.51
7	13	15.01	14.02	14.56	13.96
8	9	13.72	11.79	13.46	12.52
9	9	12.43	9.95	12.35	11.23
10	8	11.19	8.55	11.26	10.09
11	10	10.01	7.54	10.20	9.09
12	8	8.91	6.84	9.19	8.21
13	7	7.90	6.33	8.24	7.43
14	2	6.98	5.94	7.36	6.75
15	12	6.14	5.61	6.54	6.16
16	3	5.40	5.29	5.79	5.63
17	5	4.73	4.97	5.11	5.16
18	4	4.13	4.64	4.48	4.75
19	2	3.61	4.28	3.92	4.38
20	2	3.14	3.92	3.42	4.05
21	5	2.73	3.55	2.97	3.75
22	5	2.37	3.19	2.57	3.49
23	2	2.06	2.84	2.22	3.25
24	1	1.78	2.50	1.91	3.03
25 – 48	16	11.10	14.13	10.08	7.98
Total	248	248.00		248.00	248.00
$\hat{\alpha}$		1.679 (0.177)		1.447 (0.290)	5.188 (4.674)
$\hat{b}$		0.187 (0.022)		0.144 (0.047)	0.943 (1.219)
$\hat{c}$				-0.007 (0.008)	0.155 (0.075)
χ^2		15.97	8.27	19.91	18.66
df		17	15	16	16
p-value		0.5260	0.9125	0.2243	0.2867

^aGupta and Ong (2004)

^cWe combined adjacent classes for chi-square values as it was done in Gupta and Ong (2004).

4. Application to the Number of Absenteeism among Shift-Workers

Gupta and Ong (2004) defined a new generalization of the negative binomial distribution (GNBD) by mixing the mean of Poisson distribution with that of a generalized gamma distribution. The probability mass function of the new generalized negative binomial distribution, characterized by four parameters, is in terms of the confluent hyper geometric function of the second kind. The new distribution is fitted to a data set (from Arbous and Sichel (1954)) on absenteeism among shift-workers in a steel industry.

By using the SAS PROC NLMIXED, the ML estimates for the parameters of the QNBD are obtained and these are reported in Table 3. Since NBD is a special case of the QNBD, the estimates from the NBD are also reported in Table 3. The log-likelihood statistics for the NBD and the QNBD are respectively -794.00 and -793.68 . In comparing the fit from NBD to the fit from QNBD, it appears there is no difference between the two fits. Furthermore, the estimate of parameter c with its standard error in the QNBD model shows that the parameter is not significantly different from zero. Thus, the QNBD model does not provide any improvement over the NBD. The results from the chi-square goodness of fit seem to suggest that the fit of the NBD is slightly better since the chi-square statistic is 15.97 compared to the 19.91 for QNBD. These chi-square values depend on how the adjacent classes are combined.

On fitting the QNBD, we compute the truncation error to see if this is negligible. The percentage truncation error is 2.366%, which is substantially larger than the 0.5% used in our error analysis. Based on this result, we fit the TQNBD to the data on number of absenteeism and the expected frequencies under the TQNBD are provided in Table 3. The final ML estimates from QNBD are used as the initial for the ML estimates of TQNBD. The log-likelihood for the TQNBD is -777.41 , which is substantially smaller than the result from the QNBD. Based on the results in Table 3, the TQNBD model seems to provide better fit than the NBD and the QNBD models.

5. Conclusion

In this paper, we provide the region of the parameter space for which the percentage truncation error is below 0.5%. One needs to ensure that the number of non-zero probability classes is at least five (that is, $k \geq 4$). By using the parameter region specified in Table 2, one can determine if the estimated parameter values are in the region where the truncation error is negligible. For an observed data set, the observed value of k needs to be close to the integer part of $-1/\hat{c}$ in order for the percentage truncation error to be negligible. When $c < 0$, one can use the truncated QNBD. However, one should consider the non-truncated QNBD if the truncation error is negligible.

For the parameter estimates of QNBD in Table 3, the maximum percentage error of truncation is -0.2209% . This value is under the assumption that QNBD probabilities are computed for all integer values of $k \leq -1/\hat{c} \approx 142$. However, the data get truncated well below 142. In view of this situation, the truncation error is no longer negligible and a TQNBD is considered.

By using the results in Li et al. (2008), the mean and variance of the QNBD in (1.1) are given by

$$\mu_* = E(X) = \sum_{i=0}^{\infty} (c/b)^i \Gamma(\alpha + i + 1) / [b\Gamma(\alpha)], \quad (4.1)$$

$$\sigma_*^2 = \sum_{i=0}^{\infty} \left(\frac{c}{b}\right)^i \frac{(i+1)(i+2)\Gamma(\alpha + i + 1)}{b\Gamma(\alpha)} + \sum_{i=0}^{\infty} \left(\frac{c}{b}\right)^i \frac{(i+1)\Gamma(\alpha + i + 2)}{b^2\Gamma(\alpha)} - \mu_*^2. \quad (4.2)$$

From (4.1) and (4.2), the mean and variance do not exist when the parameter $c > 0$. The QNBD is truncated at a point k such that $1 + kc \geq 0$ when $c < 0$. When $c = 0$, QNBD reduces to the NBD and it is over-dispersed (i.e. the variance is greater than the mean). It is not easy to show if the QNBD has the property of both over-dispersion and under-dispersion. In the error analysis, we noticed that for some parameter values the mean is smaller than the variance and in others the mean is larger than the variance. Based on this, we conjecture that the QNBD has both under-dispersion and over-dispersion properties. Since the TQNBD is more difficult to estimate, one should apply the QNBD whenever the truncation error is negligible. A recommendation is to first fit the QNBD to an observed data set and check if the percentage truncation error is less than 0.5%. If not, consider fitting a TQNBD to the data set.

Acknowledgments

This work was done while Felix Famoye, Central Michigan University, was on sabbatical leave at the Department of Mathematics, University of Lagos, Nigeria. The author gratefully acknowledges the support received from the U.S. Department of State, Bureau of Education and Cultural Affairs under the grant #09-78737.

References

- Arbous, A.G. & Sichel, H.S. (1954). New techniques for the analysis of absenteeism data. *Biometrika*, 41, 77-90.
- Consul, P.C. & Shoukri, M.M. (1985). The generalized Poisson distribution when the sample mean is larger than the sample variance. *Communications in Statistics - Simulation and Computation*, 14(3), 667-681.
- Gupta, R.C. & Ong, S.H. (2004). A new generalization of the negative binomial distribution. *Computational Statistics & Data Analysis*, 45, 287-300.
- Li, S., Lee, C. & Famoye, F. (2008). On certain mixture distributions based on Lagrangian probability models. *Journal of Probability and Statistical Science*, 6(1), 91-100.

A NEW FAMILY OF SKEW SLASH LOGISTIC DISTRIBUTIONS

Bindu Punathumparambath and Sebastian George*

Department of Statistics, St. Thomas College, Pala, Kerala, India

Abstract

In this paper we present a new family of univariate slash distributions, which arises as the ratio of logistic distribution and independent uniform power function distribution. Slash logistic and skew Slash logistic distributions were derived and the properties were studied. Logistic and skew logistic distributions are special cases of this new family. These distributions provide us an alternative choices in simulation study and in particular, in fitting skewed data sets with heavy tails.

Keywords: Logistic distribution; slash distribution; skew logistic distribution; slash logistic distribution; slash skew logistic distribution.

1. Introduction

Univariate skew-symmetric models have been considered by several authors. In the last two decades there has been substantial work in the area of skew-normal (SN) and related distributions (see, Azzalini (1985)). For a more detailed discussion refer to Azzalini (2005). The main feature of these models is that a new parameter λ is introduced to control skewness and kurtosis. Gupta et al. (2002) provided detailed discussion of skew-symmetric models based on the normal, Student t, Cauchy, Laplace, logistic and uniform distributions. Also similar constructions were developed by Balakrishnan and Ambagaspitiya (1994), Arnold and Beaver (2000), Branco & Dey (2001) and Jones & Faddy (2003). Symmetric distributions generalizing normality have got more attention in the statistical literature. One such extension is the class of standard slash distribution proposed by Kafadar (see, 1982 and 1988), which have heavier tails than normal distribution. Genton and Wang (2006) generalized the univariate slash normal of Kafadar to multivariate skew-slash and investigated its properties. They applied their skew-slash distributions to fit AIS and glass-fiber data. Tan and Feng (2005) introduced the multivariate slash student distributions and the skew-slash student distributions. Arslan and Genc (2009) introduced a generalization of the multivariate slash distribution. Recently Bindu (2011(a) and (b)) introduced a new family of skewed slash distributions generated by normal and Cauchy kernels. Also Bindu (2011(c)) investigated several properties of multivariate asymmetric slash Laplace distribution and Bindu (2011 (d)) introduced the multivariate skew-slash t and skew-slash Cauchy distributions.

*E-mail address: ppbindukannan@gmail.com

The present paper introduces a new skewed family of distributions; skewed slash logistic distribution. This article is organized as follows. Section 2 gives the overview of the slash distribution. Section 3 defines the slash Logistic distribution and discuss various of its properties. In section 4, introduces the slash skew Logistic distribution and explore its properties. Finally, Section 5 discusses the possible applications of the model.

2. Overview of Slash Distribution

Kafadar (1988) proposed the univariate standard slash distribution as a heavy tailed alternative to the normal distribution. The standard slash distribution is obtained as the distribution of the ratio $X = \frac{Y}{U^{1/q}}$, where Y is a standard normal random variable, U is an independent uniform random variable over the interval $(0,1)$ and $q > 0$. The probability density function (pdf) of the univariate standard slash distribution is symmetric about the origin and has heavier tails than those of the normal density. When $q = 1$ we obtain the canonical slash, which has the same tail heaviness as the Cauchy and for $q \rightarrow \infty$ yields the normal distribution. The probability density function (pdf) of the univariate slash distribution is given by

$$g(x; q) = \int_0^1 u^{1/q} f(xu^{1/q}) du, \quad -\infty < x < \infty, \quad (2.1)$$

where $f(\cdot)$ is the standard normal pdf.

Then its cdf is

$$G(x; q) = \int_{-\infty}^x g(x; q) dx = \int_0^1 F(xu^{1/q}) du, \quad (2.2)$$

where $F(\cdot)$ denote the standard normal cdf. In terms of $v = u^{1/q}$ the pdf of standard slash distribution can be obtained as

$$h(x; q) = q \int_0^1 v^q f(xv) dv; \quad -\infty < x < \infty. \quad (2.3)$$

In particular for $x = 0$ we have

$$h(0; q) = \frac{q}{(q+1)\sqrt{2\pi}}$$

The closed form expression for the pdf can computed for various values of q ,

$$h(x, 1) = \begin{cases} \frac{\phi(0) - \phi(x)}{x^2}; & x \neq 0 \\ \frac{\phi(0)}{2}; & x = 0 \end{cases}$$

$$h(x, 2) = \begin{cases} \frac{2((\Phi(x) - \Phi(x))/x - \phi(x))}{x^2}; & x \neq 0 \\ \frac{2\phi(0)}{3}; & x = 0 \end{cases}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denotes the standard normal pdf and cdf respectively.

The expectation and the variance of the standard slash distribution are given by

$$E(X) = 0, \text{ for } q > 1,$$

$$\text{Var}(X) = q/(q-2), \text{ for } q > 2.$$

A general slash distribution is obtained by scale multiplication and location shift of a standard slash random variable. That is, X has a general slash distribution with parameters μ , σ and q if,

$$X = \sigma \frac{Y}{U^{1/q}} + \mu,$$

Where Y is a standard normal random variable, U is an independent uniform random variable over the interval $(0, 1)$, $\mu \in \mathfrak{R}$, $\sigma > 0$ and $q > 0$.

General properties of this distribution are studied in Rogers and Tukey (1972) and in Mostelles and Tukey (1977). Maximum likelihood estimates (MLEs) of the related location-scale family are discussed in Kafadar (1982). Recently, Wang and Genton (2006) described multivariate and skew-multivariate extensions of the slash distribution.

3. Slash Logistic Distribution

The slash logistic distribution can be defined as the distribution of the ratio $X = \frac{Y}{U^{1/q}}$ where Y is a logistic random variable and U is an independent uniform random variable over $(0,1)$ and $q > 0$. It is denoted by $X \sim SLO(\beta; q)$.

Definition 3.1. A random variable X denoted by $X \sim SLO(\beta; q)$ is said to have a slash logistic distribution if its probability density function is

$$h(x; \beta) = \frac{q}{\beta} \int_0^1 v^q \exp\left(\frac{-xv}{\beta}\right) \left\{ 1 + \exp\left(\frac{-xv}{\beta}\right) \right\}^{-2}, -\infty < x < \infty, \quad (3.1)$$

where $q > 0$, $\beta > 0$, $q > 0$ and $\lambda \in \mathbb{R}$.

The cdf is given by

$$H(x; q) = q \int_0^1 v^{q-1} \left\{ 1 + \exp\left(\frac{-xv}{\beta}\right) \right\}^{-1} dv. \quad (3.2)$$

Figure (1) gives the probability density curves of the slash logistic distribution along with Logistic distribution.

We can see that the slash logistic density is symmetric, kurtosis and tail. Thus, (3.1) allows for a greater degree of flexibility and is useful in many more practical situations, such as in the modeling of detector relative efficiencies, extreme wind speeds, measurement errors, position errors in navigation, stock return, the Earth's magnetic field and wind shear data.

Remark 3.1. Note that the slash Logistic random variable in (3.1) is a scale mixture of the logistic random variable and so it can be represented as,

$$\mathbf{X}|(U = u) \sim LO(0, u^{-1/q}\beta).$$

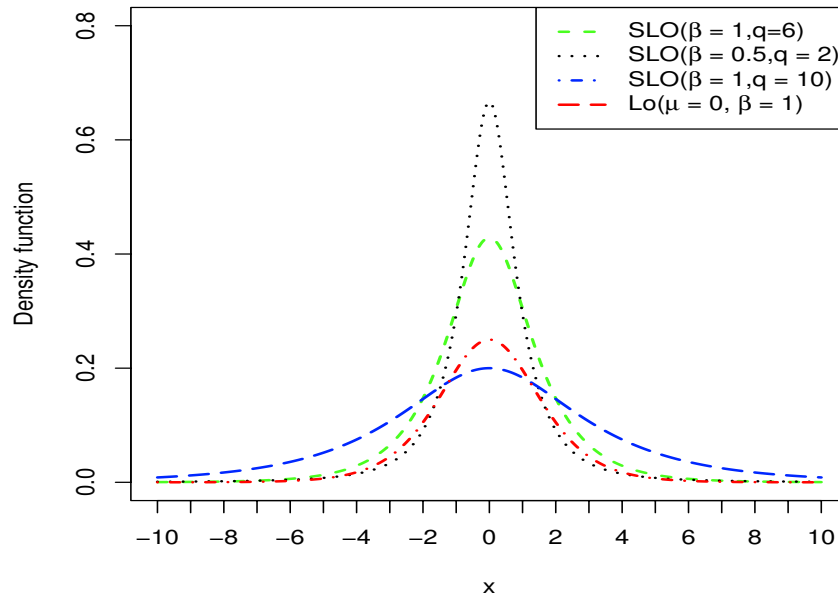


Figure 1. Slash logistic density functions for various values of parameters, along with logistic density.

Remark 3.2. If $X \sim SLO(0, 1; q)$ then its pdf $h(x; q)$ and cdf $H(x; q)$ has point mass at the origin. That is, when $x = 0$,

$$h(0, q) = q \int_0^1 v^q f(0) dv = \frac{q}{4(q+1)},$$

and

$$H(0, q) = \frac{1}{4}.$$

Remark 3.3. The limiting distribution of the standard slash logistic distribution $SLO(0, 1; q)$, $q \rightarrow \infty$, is the standard logistic distribution $LO(0, 1)$.

Proof. If $X \sim CLO(0, 1; q)$ then consider the distribution function $H(x; q)$ given by (3.2).

$$\lim_{q \rightarrow \infty} H(x; q) = \lim_{q \rightarrow \infty} \int_0^1 F(xu^{1/q}) du = \int_0^1 F(x) du = F(x),$$

where $F(x)$ is the cdf of the standard logistic distribution. Hence as $q \rightarrow \infty$, the standard slash logistic yields the logistic density itself.

Remark 3.4. If $X \sim SLO(0, 1; q)$ with pdf $h(x; q)$ given by (3.1). Then $Y = |X|$, has the pdf $2 h(y; q) I_{(y>0)}$.

Proof. If $Y = |X|$, the pdf of Y can be obtained as,

$$f(y) = h(-y; q) + h(y; q) = 2 h(y; q) I_{(y>0)},$$

which is the half slash logistic density.

4. Skew Slash Logistic Distribution

In this section we define the skewed generalization of the slash logistic distribution using the method of Azzalini (1985). A result which is central to the development of the skew-normal and related distribution appeared in Azzalini (1985), and we refer to Azzalini (2005) for a more general form, is given below.

Lemma 4.1. *Let X be a random variable with density function $f(y)$ symmetric about 0, and Z a random variable with absolutely continuous distribution function $G(y)$ such that $\check{G}(y)$ is symmetric about 0, then for any $\lambda > 0$,*

$$h(y/\lambda) = 2f(y)G(\lambda y), \quad -\infty < y < \infty, \quad (4.1)$$

is a density function for any real λ .

For example, the skew normal distribution is obtained by taking $f \equiv \phi$ and $G \equiv \Phi$ in (4.1).

The random variable $X = \frac{Y}{U^{\frac{1}{q}}}$ is said to follow a skew-slash logistic distribution, if Y follow a skew-logistic distribution and U is independent uniform distribution over the interval $(0, 1)$. Then the probability density function of X is given by,

$$f(x; \lambda, \beta, q) = 2q \int_0^1 v^q \frac{\exp\left(\frac{-xv}{\beta}\right)}{\beta \left\{1 + \exp\left(\frac{-xv}{\beta}\right)\right\}^2 \left\{1 + \exp\left(\frac{-\lambda xv}{\beta}\right)\right\}} dv, \quad x \in \mathbb{R}, \quad (4.2)$$

where $\lambda \in \mathbb{R}$, $\beta > 0$ and $q > 0$.

Figure (2) gives the probability density curves of the skew-slash logistic distribution along with Laplace distribution.

We can see that the skew-slash logistic density is symmetric, negatively skewed and positively skewed for $\lambda = 0$, $\lambda < 0$ and $\lambda > 0$ respectively. The main feature of the skew-slash logistic distribution in (4.2) is that the parameters λ and q control skewness, kurtosis and tail. Thus (4.2) allows for a greater degree of flexibility and is useful in many more practical situations, such as in the modeling of detector relative efficiencies, extreme wind speeds, measurement errors, position errors in navigation, stock return, the Earth's magnetic field and wind shear data.

4.1. Properties

The following properties of the probability density function (4.2) hold. For most of them the proof is immediate.

- (i) If X has the skew-slash logistic distribution with skew parameter λ . Then for $\lambda = 0$, X has the slash logistic distribution.

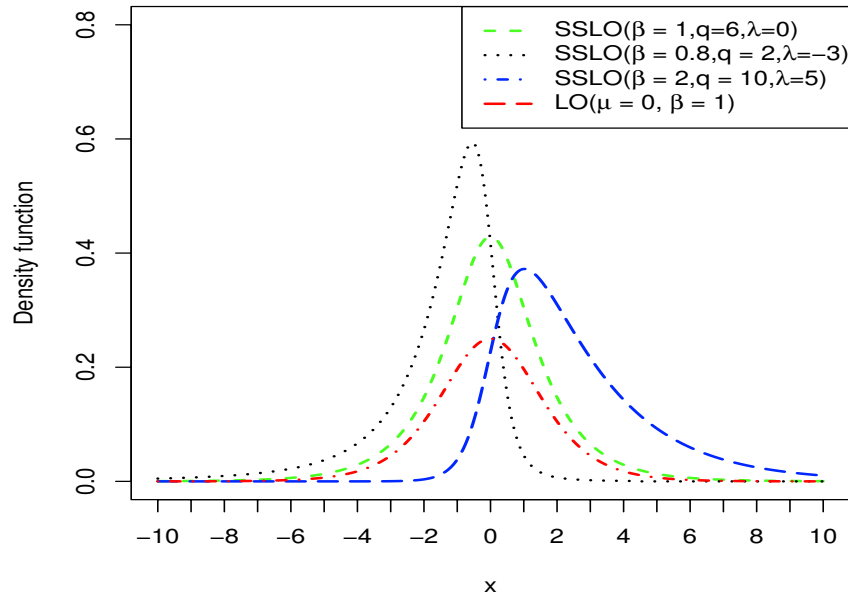


Figure 2. Skew slash logistic density functions for various values of parameters, along with logistic density.

- (ii) If X has a skew-slash logistic distribution with skew parameter λ , then $-X$ also has a skew slash logistic distribution with skewness parameter $-\lambda$.
- (iii) If $Y_\lambda(y)$ and $Y_0(y)$ are respectively the skew-slash logistic distribution with skewness parameter λ and the slash logistic random variables. Then

$$|Y_\lambda(y)| \rightarrow |Y_0(y)| \text{ as } \lambda \rightarrow \infty.$$

That is, when $\lambda \rightarrow \infty$ the half skew-slash logistic distribution tends to half slash logistic distribution. This result implies that all even moments of $Y_\lambda(x)$ and $Y_0(x)$ are the same. This property is referred to as Perturbation invariance.

- (iv) If Y_λ follows a skew-slash logistic distribution and Y is slash logistic distribution. Then for any even function h , $h(Y_\lambda) \stackrel{d}{=} h(Y)$, since from [p3.] we have $|Y_\lambda| \stackrel{d}{=} |Y|$. Also

$$Y_\lambda^2 \stackrel{d}{=} Y^2.$$

5. Discussion

The skew slash logistic distributions is a family of slash distributions. The logistic distribution belongs to this family, since when the skewness parameter is zero and tail parameter tends to infinity the skew slash logistic distributions reduces to the Logistic distribution. The slash Logistic family is also belongs to this family when the skewness parameter is

zero. This new family of distributions has wide flexibility in tail behavior and peakedness and can be applied to data from different contexts especially for heavy-tailed data sets. Heavy-tailed distributions are commonly found in complex multi-component systems like ecological systems, biometry, economics, sociology, internet traffic, finance, business etc. Finally, skew-slash logistic distribution could be employed to deal with skewed data, heavy-tailed data and data which are both skewed and heavy-tailed.

Acknowledgment

The first author is grateful to the Department of Science & Technology, Government of India, New Delhi, for financial support under the Women Scientist Scheme (WOS-A (2008)), Project No: SR/WOS-A/MS-09/2008. The authors are very thankful to the referees and the editor of the journal for their valuable suggestions and comments.

References

- [1] Arnold, B. C. and Beaver, R. J. (2000). The skew-Cauchy distribution. *Statist. Probab. Lett.* 49, 285-290.
- [2] Arslan, O. and Genc, A. I. (2009). A generalization of the multivariate slash distribution. *Journal of Statistical Planning and Inference* 78, 1164-1170.
- [3] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist.* 12, 171-178.
- [4] Azzalini, A. (2005). The Skew-normal Distribution and Related Multivariate Families. *Scand. J. Statist.* 32, 159-188.
- [5] Balakrishnan, N. and Ambagaspitiya, R. S. (1994). *On skew-Laplace distributions*. Technical Report, Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada.
- [6] Bindu, P. P. (2011(a)). A new family of skewed slash distributions generated by the normal kernel. *STATISTICA* (to appear).
- [7] Bindu, P. P. (2011(b)). A new family of skewed slash distributions generated by the cauchy kernel. *Communications in Statistics- Theory and Methods* (in press).
- [8] Bindu, P. P. (2011(c)). The multivariate asymmetric slash Laplace distribution and its applications. *STATISTICA* (accepted for publication).
- [9] Bindu, P. P. (2011 (d)). The multivariate skew-slash t and skew-slash Cauchy distributions. *Model Assisted Statistics and Applications* 6, 18.
- [10] Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *J. Multivariate Anal.* 79, 99-113.
- [11] Gupta, A. K., Chang, F. C., Huang, W. J. (2002). Some skew-symmetric models. *Random Operators and Stochastic Equations*. 10, 133140.

- [12] Jones, M.C., Faddy, M.J. (2003). A skew extension of the t distribution, with applications. *J. Roy. Statist. Soc. Series B*, 65, 159-174.
- [13] Kafadar, K. (1982). A biweight approach to the one-sample problem. *J. Amer. Statist. Assoc.* 77, 416-424.
- [14] Kafadar, K. (1998). Slash Distribution. *Encyc. Statist. Sciences* 8.
- [15] Rogers, W.H., Tukey, J.W. (1972). Understanding some long-tailed symmetrical distributions. *Statist. Neerlandica* 26, 211-226.
- [16] Tan, F., Peng, H. (2005). The slash and skew-slash student t distribution. Pre-print.
- [17] Wang, J., Genton, M. G. (2006). The multivariate skew-slash distribution. *Journal of Statistical Planning and Inference* 136, 209-220.

ON THE MOMENTS OF SAMPLE COVARIANCE FROM A BIVARIATE NORMAL POPULATION

Anwar H. Joarder^{*} and M. Hafidz Omar[†]

Department of Mathematics and Statistics
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia

Abstract

The distribution of the sample covariance based on a sample from a bivariate normal population has been derived in the literature by many authors. Some used conditionality principle, some used inverse Mellin transform and some used simple transformations. The main contribution of the paper is the direct derivation of the general moment structure of sample covariance via a lemma dealing with a infinite sum of product of two gamma functions. Integer valued moments of any order are also derived in terms of hypergeometric function. Some corollaries have been deduced which have potential for further use in mathematics and statistics. The probability density function, skewness and kurtosis have been graphed.

AMS Mathematics Subject Classification: 60E05, 60E10, 62E15.

Keywords and Phrases: Sample covariance, Wishart distribution, Macdonald function, hypergeometric function.

1. Introduction

Let a_{11} and a_{22} be the mean centered sum of squares, and a_{12} , the mean centered sum of products for samples of size $N (> 2)$ drawn from a bivariate normal distribution with means θ_1 and θ_2 , variances σ_1^2 and σ_2^2 , and covariance $\sigma_{12} = \rho\sigma_1\sigma_2$. Then $a_{ii} = ms_i^2$, $a_{12} = mrs_1s_2$, $m = N - 1$, ($i = 1, 2$), s_1^2 and s_2^2 are sample variances and r is the product moment correlation coefficient. The joint density function of mean centered sum of squares and sum of products at a_{11} and a_{22} a_{12} , originally due to Fisher (1915), is given by

^{*} E-mail address: anwarj@kfupm.edu.sa

[†] E-mail address: omarmh@kfupm.edu.sa

$$f(a_{11}, a_{22}, a_{12}) = \frac{(1-\rho^2)^{-m/2} (\sigma_1 \sigma_2)^{-m}}{2^m \sqrt{\pi} \Gamma(m/2) \Gamma((m-1)/2)} (a_{11} a_{22} - a_{12}^2)^{(m-3)/2} \times \exp\left(-\frac{a_{11}}{2(1-\rho^2)\sigma_1^2} - \frac{a_{22}}{2(1-\rho^2)\sigma_2^2} + \frac{\rho a_{12}}{(1-\rho^2)\sigma_1 \sigma_2}\right), \quad (1.1)$$

$a_{11} > 0, a_{22} > 0, -\infty < a_{12} < \infty, m > 2$ and $-1 < \rho < 1$ (Anderson, 2003, 123).

Because of the importance of Wishart distribution in multivariate statistical analysis, various authors have given different derivations. See Gupta and Nagar (2000, 87-88) and the references therein for a good update on the Wishart distribution.

The distribution of the sample covariance based on a sample from a bivariate normal population has been known. Pearson, Jeffery, and Elderton (1929) used simple transformation on the distribution of sample variances and covariance. Wishart and Bartlett (1932) also studied the distribution of covariance by inverting the characteristic function of sample variances and covariance. Mahalanobis, Bose and Roy (1937) reviewed the distribution of the generalized variance (determinant of variance covariance matrix) due to Wilks (1932), distribution of the ratio of correlated standard deviations due to Bose (1935), the distribution of sample covariance due to Pearson, Jeffery, and Elderton (1929) and many other aspects of the distribution of sample variances and correlation coefficient. Springer, (1979, 343) derived the distribution of sample covariance by inverse Mellin transform.

We sketch the derivation of the density function of sample covariance by making simple transformations on the distribution of sample variances and covariance as was done by Pearson, Jeffery, and Elderton (1929) just to make the paper self-contained for today's readers. Standard notations have been used throughout the paper. The main contribution of the paper is the derivation of general moments of the distribution with the help of a lemma that involves infinite sum of a product of gamma functions. Integer valued moments of sample covariance of any order have been derived in terms of hypergeometric function. A number of corollaries have been deduced which has potential for applications in mathematics and statistics. The probability density function, the coefficient of skewness and coefficient of kurtosis have been graphed.

The following results will be used in sequel:

$$(2z)! \sqrt{\pi} = 2^{2z} z! \Gamma(z + \frac{1}{2}). \quad (1.2)$$

The modified Bessel function of the second kind or the Macdonald function admits the following integral representation (Watson, 1993, 183):

$$K_\alpha(y) = \frac{1}{2} \left(\frac{1}{2} y\right)^\alpha \int_0^\infty t^{-(\alpha+1)} \exp\left(-t - \frac{y^2}{4t}\right) dt. \quad (1.3)$$

The hypergeometric function ${}_pF_q(a_1, a_2, \dots, a_p; b_1, b_2, \dots, b_q; z)$ is defined by

$${}_pF_q(a_1, a_2, \dots, a_p; b_1, b_2, \dots, b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_{\{k\}} (a_2)_{\{k\}} \dots (a_p)_{\{k\}}}{(b_1)_{\{k\}} (b_2)_{\{k\}} \dots (b_q)_{\{k\}}} \frac{z^k}{k!}, \quad (1.4)$$

where $a_{\{k\}} = a(a+1)\dots(a+k-1)$.

2. Distribution of the Sample Covariance

In this section, we derive the distribution of sample covariance $W = (a_{12} / m)$ by integrating out a_{11} and a_{22} from the density function (1.1) as was done by Pearson, Jeffery and Elderton (1929).

Theorem 2.1. Let $w = a_{12} / m$, be the sample covariance based on a sample from a bivariate normal distribution. Then the density function of W is given by

$$f_1(w) = \frac{(m / \sigma_1 \sigma_2)^{(m+1)/2}}{2^{(m-1)/2} \sqrt{\pi(1-\rho^2)} \Gamma(m/2)} |w|^{(m-1)/2} \exp\left(\frac{\rho m w}{(1-\rho^2)\sigma_1 \sigma_2}\right) K_{(m-1)/2}\left(\frac{m |w|}{(1-\rho^2)\sigma_1 \sigma_2}\right), \quad (2.1)$$

where $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1, m > 2$ and $K_\alpha(x)$ is the modified Bessel function of the second kind represented by (1.3).

Proof. It follows from (1.1) that the density function of $a_{12} = Z$ is given by

$$f_2(z) = \frac{(1-\rho^2)^{-m/2} (\sigma_1 \sigma_2)^{-m}}{2^m \sqrt{\pi} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{m-1}{2}\right)} \exp\left(\frac{\rho z}{(1-\rho^2)\sigma_1 \sigma_2}\right) I(m, \rho), \quad (2.2)$$

$$\text{where } I(m, \rho) = \iint_{z^2 \leq xy} (xy - z^2)^{(m-3)/2} \exp\left(-\frac{x}{2(1-\rho^2)\sigma_1^2} - \frac{y}{2(1-\rho^2)\sigma_2^2}\right) dy dx.$$

Letting $xy - z^2 = u$, i.e., $y = (z^2 + u) / x$ with Jacobian $J(y \rightarrow u) = 1/x$, we have

$$I(m, \rho) = \int_{x=0}^{\infty} \int_{u=0}^{\infty} u^{(m-3)/2} \exp\left(-\frac{x}{2(1-\rho^2)\sigma_1^2} - \frac{z^2 + u}{2(1-\rho^2)\sigma_2^2 x}\right) \frac{1}{x} du dx.$$

Completing the integral in u and substituting $[2(1-\rho^2)\sigma_1^2]^{-1} x = t$ in the resulting integral, we have

$$I(m, \rho) = \Gamma\left(\frac{m-1}{2}\right) [2(1-\rho^2)\sigma_1 \sigma_2]^{m-1} \int_{t=0}^{\infty} t^{(m-3)/2} \exp\left(-t - \frac{z^2}{4(1-\rho^2)^2 \sigma_1^2 \sigma_2^2 t}\right) dt.$$

Putting $-(\alpha+1) = (m-3)/2$ and $y^2 = [(1-\rho^2)\sigma_1 \sigma_2]^{-2} z^2$, in the Bessel function of the second kind given by (1.3), and using the property $K_{-\alpha}(z) = K_\alpha(z)$, we have

$$I(m, \rho) = 2^{(m+1)/2} \Gamma\left(\frac{m-1}{2}\right) [(1-\rho^2)\sigma_1\sigma_2]^{(m-1)/2} \times |z|^{(m-1)/2} K_{(m-1)/2}\left(\frac{|z|}{(1-\rho^2)\sigma_1\sigma_2}\right). \quad (2.3).$$

By using (2.3) in (2.2) and performing a bit of simplification, we have (2.1).

The probability density function in (2.1) has been graphed in Figure 1 below:

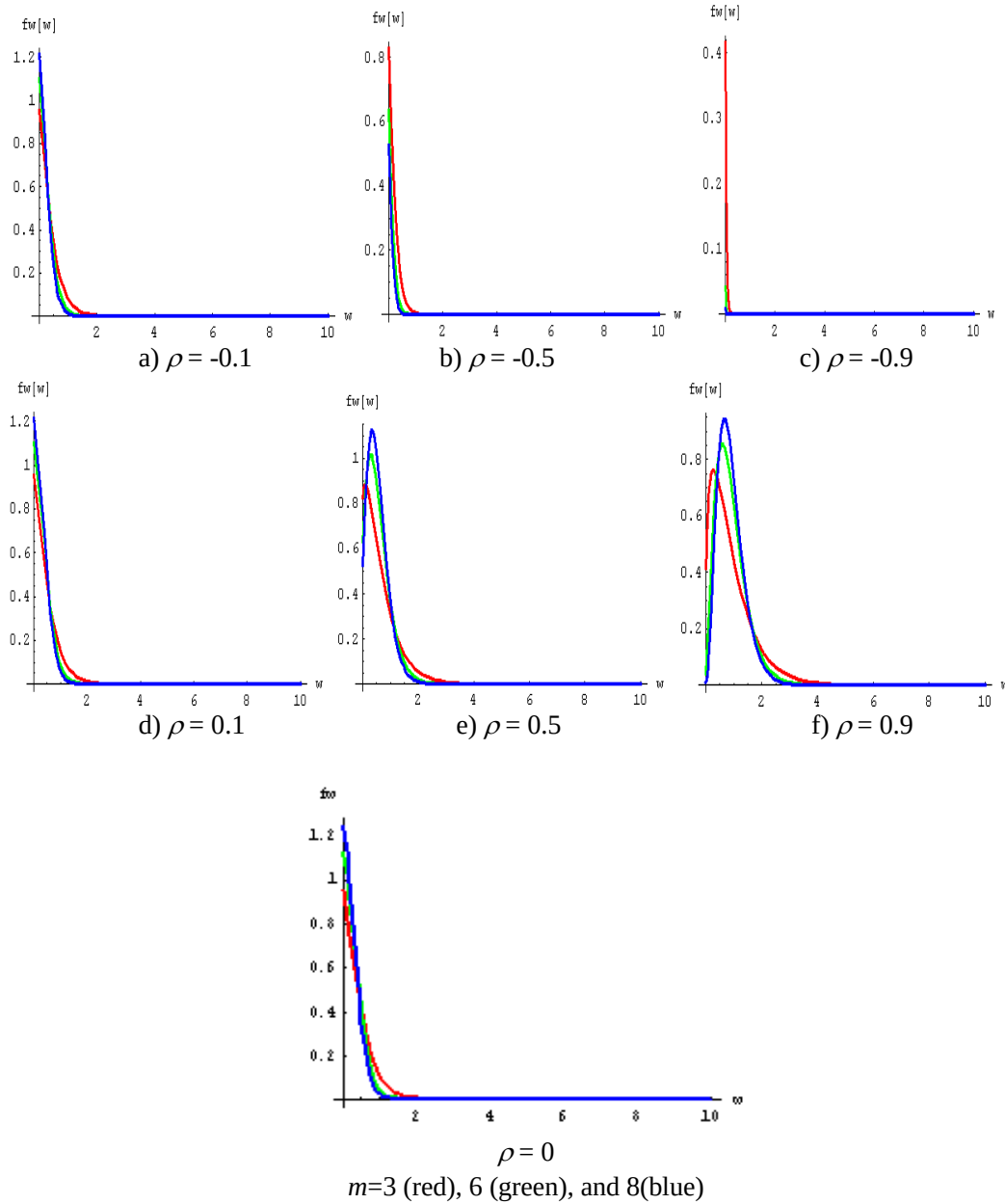


Figure 1. Density function for the sample covariance at various values of ρ and m .

The density function of $Y = m[(1-\rho^2)\sigma_1\sigma_2]^{-1}W$ at ν is given by

$$f_2(\nu) = \frac{(1-\rho^2)^{m/2}}{2^{(m-1)/2}\sqrt{\pi}\Gamma(m/2)} |\nu|^{(m-1)/2} e^{\rho} K_{(m-1)/2}(|\nu|),$$

which was obtained by Springer (1979, 343) by inverse Mellin transform. Some interesting forms of the density function of sample covariance are pointed by Press (1967).

3. Moments of the Sample Covariance

The higher order moments of the sample covariance $W = S_1 S_2 R$ can be calculated directly from the density function derived in Theorem 2.1. In this section we will derive them from (1.1). Since the moments of sample covariance involves infinite sum of product of gamma functions, we have the following lemma to have closed form expressions.

Lemma 3.1. Let $-1 < \rho < 1$, $m > 2$, and c and d are real numbers. Also let

$$g(k; c, d) = \frac{(2\rho)^k}{k!} \Gamma\left(\frac{k+m+2h+c}{2}\right) \Gamma\left(\frac{k+2h+d}{2}\right).$$

Then

$$\text{a. } \sum_{j=0}^{\infty} g(2j; c, d) = \Gamma\left(h + \frac{m+c}{2}\right) \Gamma\left(h + \frac{d}{2}\right) {}_2F_1\left(h + \frac{m+c}{2}, h + \frac{d}{2}; \frac{1}{2}; \rho^2\right),$$

$$2h + m + c \geq 0,$$

$$\text{b. } \sum_{j=0}^{\infty} g(2j+1; c, d) = 2\rho \Gamma\left(h + \frac{m+c+1}{2}\right) \Gamma\left(h + \frac{d+1}{2}\right) {}_2F_1\left(h + \frac{m+c+1}{2}, h + \frac{d+1}{2}; \frac{3}{2}; \rho^2\right),$$

$$2h + m + c + 1 \geq 0, \quad 2h + d + 1 \geq 0.$$

Proof. By using (1.2) in $g(k; c, d)$, we have

$$\sum_{j=0}^{\infty} g(2j; c, d) = \sum_{j=0}^{\infty} \frac{\sqrt{\pi} \rho^{2j}}{j! \Gamma(j + (1/2))} \Gamma\left(j + h + \frac{m+c}{2}\right) \Gamma\left(j + h + \frac{d}{2}\right)$$

which can be expressed as what we have in part (a) by virtue of (1.4). By using (1.2) in $g(k; c, d)$, we have

$$\sum_{j=0}^{\infty} g(2j+1; c, d) = \sum_{j=0}^{\infty} \frac{\sqrt{\pi} \rho^{2j+1}}{j! \Gamma(j + (3/2))} \Gamma\left(j + h + \frac{m+c+1}{2}\right) \Gamma\left(j + h + \frac{d+1}{2}\right),$$

which can be written as part (b) by virtue of (1.4).

Theorem 3.1. Let W be the sample covariance based on a bivariate normal distribution. Then, for any real number h , the h -th moment of W is given by

$$E(W^h) = \frac{2^{h-1}(1-\rho^2)^{h+(m/2)}(\sigma_1\sigma_2)^h}{m^h\Gamma(m/2)\sqrt{\pi}} \sum_{k=0}^{\infty} \{1+(-1)^{k+h}\} \frac{(2\rho)^k}{k!} \Gamma\left(\frac{k+m+h}{2}\right) \Gamma\left(\frac{k+1+h}{2}\right),$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \rho < 1$ and $m > 2$.

Proof. Under the transformation $a_{11} = mS_1^2$, $a_{22} = mS_2^2$, $a_{12} = mrs_1s_2$ in (1.1) with Jacobian $4m^3s_1^2s_2^2$, the joint density function of the sample standard deviations (S_1, S_2) and correlation coefficient R is given by

$$\begin{aligned} f_2(s_1, s_2, r) &= \frac{m^m (1-\rho^2)^{-m/2} (\sigma_1\sigma_2)^{-m}}{\pi \Gamma(m-1)} \exp\left(-\frac{m}{2-2\rho^2} \left(\frac{s_1^2}{\sigma_1^2} + \frac{s_2^2}{\sigma_2^2}\right)\right) \\ &\times \sum_{k=0}^{\infty} \frac{(m\rho)^k (s_1s_2)^{m-1+k}}{(1-\rho^2)^k (\sigma_1\sigma_2)^k k!} r^k (1-r^2)^{(m-3)/2}. \end{aligned} \quad (3.1)$$

Since $W = S_1S_2R$, the h -th moment of W is given by $E(W^h) = E(S_1^h S_2^h R^h)$ which is (h, h, h) -th moment of the joint density function of S_1, S_2 and R . Hence, we have

$$\begin{aligned} E(W^h) &= \frac{m^m (1-\rho^2)^{-m/2} (\sigma_1\sigma_2)^{-m}}{\pi \Gamma(m-1)} \sum_{k=0}^{\infty} \frac{(m\rho)^k}{(1-\rho^2)^k (\sigma_1\sigma_2)^k k!} \\ &\times \int_0^{\infty} s_1^{(m+k+h)/2} \exp\left(\frac{-ms_1^2}{(2-2\rho^2)\sigma_1^2}\right) ds_1 \int_0^{\infty} s_2^{(m+k+h)/2} \exp\left(-\frac{1}{2} \frac{ms_2^2}{(1-\rho^2)\sigma_2^2}\right) ds_2 \\ &\times \{1+(-1)^{k+h}\} \int_0^1 u^{(k+h-1)/2} (1-u)^{(m-3)/2} \frac{1}{2} du. \end{aligned}$$

Having evaluated the gamma and beta integrals, we have the theorem.

Theorem 3.2. Let W , be the sample covariance based on a sample from a bivariate normal distribution. Then for any real number h , the $(2h)$ -th moment of W is given by

$$\begin{aligned}
E(W^{2h}) &= \frac{2^{2h-1}(1-\rho^2)^{2h+(m/2)}}{m^{2h}\Gamma(m/2)\sqrt{\pi}}(\sigma_1\sigma_2)^{2h} \\
&\times \left[\{1+(-1)^{2h}\}\Gamma\left(h+\frac{m}{2}\right)\Gamma\left(h+\frac{1}{2}\right) {}_2F_1\left(h+\frac{m}{2}, h+\frac{1}{2}; \frac{1}{2}; \rho^2\right) \right. \\
&\quad \left. + 2\{1-(-1)^{2h}\}\rho\Gamma\left(h+\frac{m+1}{2}\right)\Gamma(h+1) {}_2F_1\left(h+\frac{m+1}{2}, h+1; \frac{3}{2}; \rho^2\right) \right], \quad (3.2)
\end{aligned}$$

where $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1, m > 2$ and ${}_pF_q(a_1, a_2; b_1; z)$ is the generalized hypergeometric function defined by (1.4).

Proof. From Theorem 3.1, we have

$$E(W^{2h}) = \frac{2^{2h-1}(1-\rho^2)^{2h+(m/2)}}{m^{2h}\Gamma(m/2)\sqrt{\pi}}(\sigma_1\sigma_2)^{2h} \sum_{k=0}^{\infty} \{1+(-1)^{2h}(-1)^k\} g(k; 0, 0)$$

where $g(k; 0, 0)$ is defined in Lemma 3.1. Note that

$$\sum_{k=0}^{\infty} \{1+(-1)^{2h}(-1)^k\} g(k; 0, 1) = \{1+(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j; 0, 1) + \{1-(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j+1; 0, 1).$$

Since

$$E(W^{2h}) = \frac{2^{2h-1}(1-\rho^2)^{2h+(m/2)}}{m^{2h}\Gamma(m/2)\sqrt{\pi}}(\sigma_1\sigma_2)^{2h} \left[\{1+(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j; 0, 1) + \{1-(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j+1; 0, 1) \right],$$

the proof is complete by using Lemma 3.1.

Corollary 3.1. Let W be the sample covariance based on a sample from a bivariate normal distribution, and h be an integer. Then the $2h$ -th moment of sample covariance is given by

$$E(W^{2h}) = \frac{(1-\rho^2)^{(m+4h)/2}}{\Gamma(m/2)\sqrt{\pi}} \left(\frac{2\sigma_1\sigma_2}{m} \right)^{2h} \Gamma\left(\frac{m}{2}+h\right) \Gamma\left(\frac{1}{2}+h\right) {}_2F_1\left(\frac{1}{2}+h, \frac{m+2h}{2}; \frac{1}{2}; \rho^2\right), \quad (3.3)$$

where $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1, m > 2$.

Proof. If h is an integer in (3.2), then $1+(-1)^{2h} = 2$, and $1-(-1)^{2h} = 0$. Then the corollary follows from Theorem 3.2.

Theorem 3.3. Let W , be the sample covariance based on a sample from a bivariate normal distribution. Then for any real number h , the $(2h+1)$ -th moment of W is given by

$$\begin{aligned} E(W^{2h+1}) &= \frac{2^{2h}(1-\rho^2)^{(2h+1)+(m/2)}}{\Gamma(m/2)\sqrt{\pi}} \left(\frac{\sigma_1\sigma_2}{m} \right)^{2h+1} \\ &\times \left[\{1-(-1)^{2h}\}h\Gamma\left(h+\frac{m+1}{2}\right)\Gamma(h) {}_2F_1\left(h+\frac{m+1}{2}, h+1; \frac{1}{2}; \rho^2\right) \right. \\ &\left. + 2\{1+(-1)^{2h}\}\rho\Gamma\left(h+\frac{m+2}{2}\right)\Gamma\left(h+\frac{3}{2}\right) {}_2F_1\left(h+\frac{m+2}{2}, h+\frac{3}{2}; \frac{1}{2}; \rho^2\right) \right], \end{aligned} \quad (3.4)$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \rho < 1$ and $m > 2$.

Proof. From Theorem 3.1, we have

$$E(W^{2h+1}) = \frac{2^{2h}(1-\rho^2)^{(2h+1)+(m/2)}}{m^{2h+1}\Gamma(m/2)\sqrt{\pi}} (\sigma_1\sigma_2)^{2h+1} \sum_{k=0}^{\infty} \{1+(-1)^k(-1)^{2h+1}\} \frac{(2\rho)^k}{k!} g(k; 1, 2)$$

where $g(k; 1, 2)$ is given by Lemma 3.1. Since

$$\sum_{k=0}^{\infty} \{1+(-1)^k(-1)^{2h+1}\} g(k; 1, 2) = \{1-(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j; 1, 2) + \{1+(-1)^{2h}\} \sum_{j=0}^{\infty} g(2j+1; 1, 2),$$

we have

$$\begin{aligned} E(W^{2h+1}) &= \frac{2^{2h}(1-\rho^2)^{(2h+1)+(m/2)}}{\Gamma(m/2)\sqrt{\pi}} \left(\frac{\sigma_1\sigma_2}{m} \right)^{2h+1} \\ &\times \left[[1+(-1)^{2h}] \sum_{j=0}^{\infty} g(2j; 1, 2) + [1-(-1)^{2h}] \sum_{j=0}^{\infty} g(2j+1; 1, 2) \right], \end{aligned}$$

the proof is then complete by Lemma 3.1.

Corollary 3.2. Let W be the sample covariance based on a sample from a bivariate normal distribution. Then for any integer h , the $(2h+1)$ -th moment of sample covariance is given by

$$\begin{aligned} E(W^{2h+1}) &= \frac{\rho(1-\rho^2)^{(2h+1)+(m/2)}}{\Gamma(m/2)\sqrt{\pi}} \left(\frac{2\sigma_1\sigma_2}{m} \right)^{2h+1} \\ &\times \Gamma\left(h+\frac{m+2}{2}\right)(2h+1)\Gamma\left(h+\frac{1}{2}\right) {}_2F_1\left(h+\frac{m+2}{2}, h+\frac{3}{2}; \frac{1}{2}; \rho^2\right), \end{aligned} \quad (3.5)$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \rho < 1$ and $m > 2$.

Proof. If h is an integer in (3.4), then $1+(-1)^{2h}=2$, and $1-(-1)^{2h}=0$. Then the corollary follows from Theorem 3.3.

Corollary 3.3. Let h be an integer, $m > 2$ and $-1 < \rho < 1$. Then we have the following integrals:

$$\begin{aligned} a. \int_{-\infty}^{\infty} w^{2h} |w|^{(m-1)/2} \exp\left(\frac{\rho m w}{(1-\rho^2)\sigma_1\sigma_2}\right) K_{(m-1)/2}\left(\frac{m|w|}{(1-\rho^2)\sigma_1\sigma_2}\right) dw \\ = \frac{2^{2h+(m-1)/2}(1-\rho^2)^{(m+4h+1)/2}}{m^{2h+(m+1)/2}(\sigma_1\sigma_2)^{2h-(m+1)/2}} \Gamma\left(\frac{m}{2}+h\right) \Gamma\left(\frac{1}{2}+h\right) {}_2F_1\left(\frac{1}{2}+h, \frac{m+2h}{2}; \frac{1}{2}; \rho^2\right), \end{aligned} \quad (3.6)$$

$$\begin{aligned} b. \int_{-\infty}^{\infty} w^{2h+1} |w|^{(m-1)/2} \exp\left(\frac{\rho m w}{(1-\rho^2)\sigma_1\sigma_2}\right) K_{(m-1)/2}\left(\frac{m|w|}{(1-\rho^2)\sigma_1\sigma_2}\right) dw \\ = \frac{2^{2h+(m+1)/2} \rho \{(1-\rho^2)\sigma_1\sigma_2\}^{2h+(m+3)/2}}{m^{2h+(1-m)/2}} \\ \times \Gamma\left(h+\frac{m+2}{2}\right) (2h+1) \Gamma\left(h+\frac{1}{2}\right) {}_2F_1\left(h+\frac{m+2}{2}, h+\frac{3}{2}; \frac{1}{2}; \rho^2\right). \end{aligned} \quad (3.7)$$

Proof. Part (a) follows from Corollary 3.1. Part (b) follows from Corollary 3.2.

A more general integral can be derived from Theorem 3.1 for any real h .

Corollary 3.4. Let W be the sample covariance based on a sample from a bivariate normal distribution. Then first four moments of sample covariance are given below:

- a. $E(W) = \rho\sigma_1\sigma_2$,
- b. $E(W^2) = \frac{1}{m}[(m+1)\rho^2 + 1]\sigma_1^2\sigma_2^2$,
- c. $E(W^3) = \frac{(m+2)\rho}{m^2}[(m+1)\rho^2 + 3]\sigma_1^3\sigma_2^3$,
- d. $E(W^4) = \frac{(m+2)}{m^3}[(m+1)(m+3)\rho^4 + 6(m+3)\rho^2 + 3]\sigma_1^4\sigma_2^4$,

where $\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$, and $m > 2$.

The centered moments of sample covariance of order a is given by $\mu_a = E(W - \mu)^a$, $a=1,2,\dots$, that is the second, third and fourth order mean corrected moments are given by $\mu_2 = E(W^2) - \mu^2$, $\mu_3 = E(W^3) - 3E(W^2)\mu + 2\mu^3$, and $\mu_4 = E(W^4) - 4E(W^3)\mu + 6E(W^2)\mu^2 - 3\mu^4$.

Corollary 3.5. The first three non-zero centered moments of sample covariance based on bivariate normal distribution are given by

a. $\mu_2 = \frac{1}{m}(1+\rho^2)\sigma_1^2\sigma_2^2,$

b. $\mu_3 = \frac{2}{m^2}\rho(3+\rho^2)\sigma_1^3\sigma_2^3,$

c. $\mu_4 = \frac{3}{m^3}[(m+2)\rho^4 + (2m+12)\rho^2 + (m+2)]\sigma_1^4\sigma_2^4,$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \rho < 1$ and $m > 2$ which matches with (32.126b) of Johnson, Balakrishnan and Johnson (1995, 601).

The skewness and kurtosis are given by the moment ratios $\alpha_i(W) = \mu_i \mu_2^{-i/2}$, $i = 3, 4$.

That is, they are given by

$$\alpha_3(W) = \frac{2\rho(3+\rho^2)}{\sqrt{m}(1+\rho^2)^{3/2}}, \text{ and } \alpha_4(W) = 3 + \frac{6(1+6\rho^2+\rho^4)}{m(1+\rho^2)^2} \text{ respectively.}$$

Note that the moment ratios match with Johnson, Kotz and Balakrishnan (1995, 601).

If $\rho = 0$, then $\alpha_3(W) = 0$ and $\alpha_4(W) = 3$. Moreover if $m \rightarrow \infty$, then $\alpha_3(W) = 0$ and $\alpha_4(W) = 3$.

The graph of skewness as a function of ρ at different values of m is produced in Figure 2 below.

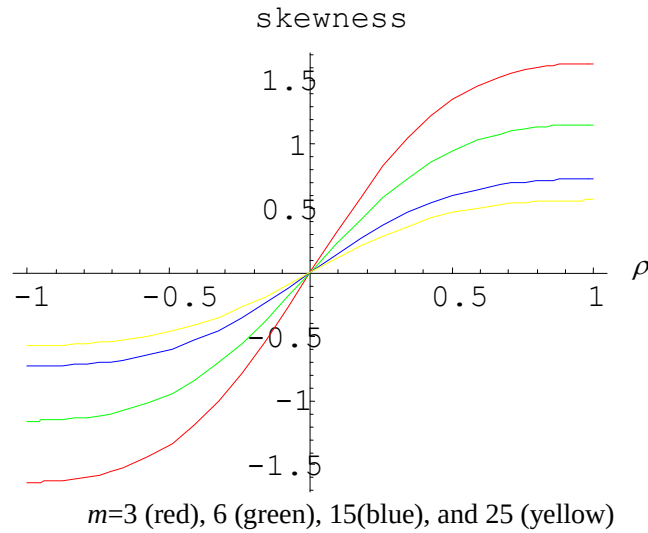


Figure 2. Coefficient of Skewness for different ρ and m .

The graph of kurtosis as a function of ρ at different values of m is produced in Figure 3 below.

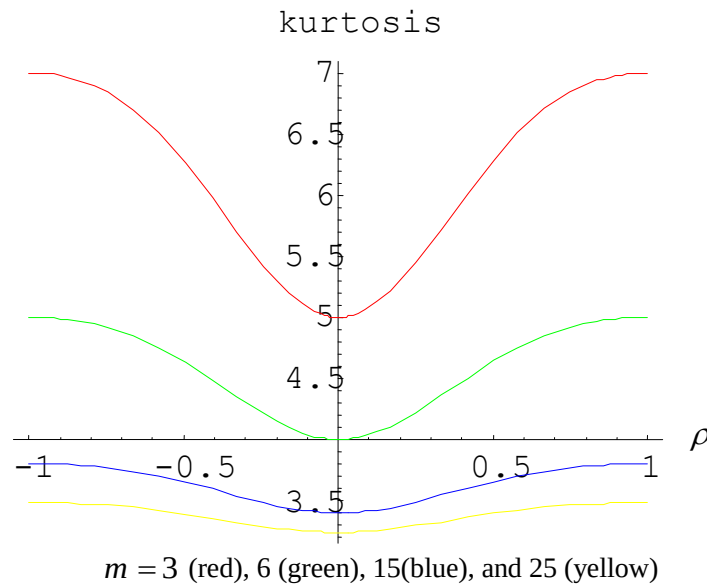


Figure 3. Coefficient of Kurtosis for different ρ and m .

4. Conclusion

In this paper we have derived closed form expressions of integer valued moments of higher order. This may lead to further investigation in statistical inference.

Acknowledgments

The authors gratefully acknowledge the research support provided by King Fahd University of Petroleum & Minerals, Saudi Arabia.

References

- Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons. New York.
- Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10, 507-521.
- Gupta, A.K. and Nagar, D.K. (2000). *Matrix Variate Distributions*. Chapman and hall, London, UK.
- Bose, S.S. (1935). On the distribution of the ratio of variances of two samples drawn from a given normal bivariate correlated population. *Sankhya*, 2, 65-72.

- Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, v2. John Wiley.
- Mahalanobis, P.C., Bose, R.C. and Roy, S.N. (1937). Normalization of statistical variates and the use of rectangular coordinates in the use of sampling distributions. *Sankhya*, 3, 1-40.
- Pearson, K., Jeffery, G.B. and Elderton, E.M. (1929). On the distribution of the first product moment-coefficient in samples drawn from an indefinitely large normal population. *Biometrika*, 21, 164-193. (tables computed by E.M. Elderton, pp 194-201).
- Press, S.J. (1967). On the sample covariance from a bivariate normal distribution. *Annals of the Institute of Statistical Mathematics*, 19, 355-361.
- Springer, M.D. (1979). *The Algebra of Random Variables*. John Wiley and Sons.
- Watson, G.N. (1993). *A Treatise on the Theory of Bessel Functions*. Cambridge University Press.
- Wilks, S.S. (1932). Certain generalizations in the Analysis of variance. *Biometrika*, 24, 476-477.
- Wishart, J. and Bartlett, M.S. (1932). The distribution of second order moment statistics in a normal system. *Proceedings of the Cambridge Philosophical Society*, 28, 455-459.

BAYESIAN IDENTIFICATION OF TWO DIMENSIONAL MOVING AVERAGE PROCESSES

Samir M. Shaarawy^{1,*}, Mohammad S. Albassam² and Sherif S. Ali^{2,†}

¹Department of Mathematical Sciences, Umm Al-Qura University,
Makkah, Saudi Arabia

²Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

The main objective of this article is to develop a direct Bayesian technique to determine or estimate the order of two dimensional moving average processes. The proposed technique is based on approximating the likelihood function by a matrix normal-Wishart distribution on the parameter space. Combining the approximate likelihood with a matrix normal-Wishart (or vague prior) and an indirect technique to estimate an initial value for the order, the marginal posterior probability mass function of the model order is developed in a convenient form. Then one investigates the approximate posterior probability mass function of the model order over its grid and selects the order with the maximum posterior probability to solve the identification problem. A simulation study, with three different priors, is conducted in order to inspect the numerical efficiency of the proposed Bayesian technique and compare it with the well known AIC automatic technique. The numerical results show that the proposed Bayesian technique is simple, efficient and easy to program.

Keywords: Identification, Two dimensional moving average processes, likelihood function, matrix normal-Wishart distribution, Probability mass function, AIC technique.

1. Introduction

Moving average processes have been widely used successfully to model time series data arise in many areas of application such as economics, business, agriculture, engineering, biometrics, geophysics, chemistry and environmental studies. The statistical analysis of such data often requires modeling two time series jointly. In a study of consumer behavior, as an example, one may record spending on education $y_1(t)$ and spending on health services $y_2(t)$. In a study of traffic accidents, as another example, one may record the number of accidents $y_1(t)$ and the number of fatalities because of the accidents $y_2(t)$. As a third example, one may record the temperature $y_1(t)$ and humidity $y_2(t)$. These variables are

*Permanent Address: Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt

†Permanent Address: Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt

modeled and examined jointly in order to understand the nature of dynamic relationship among the variables and to increase the precision of the estimates of the parameters and future observations.

Of great importance in analyzing the two dimensional time series is identifying the order of the process. In identification process, a time series analyst constructs the tools that will be used to absorb information from the data to determine or estimate the model order. However, it is well known that the solution of identification problem depends on subjective opinions as well as mathematical and statistical arguments. It is well known also that there is no optimum method is completely agreed upon to identify the model order.

With respect to univariate moving average processes, one may trace three different non-Bayesian approaches to identify the model order. In 1970, Box and Jenkins presented the first and most popular approach by matching the autocorrelation function computed from the data with its theoretical counterpart. Their technique has been widely used by many others such as Chatfield (1980), Harvey (1981), Priestely (1981) and Liu (2009). The second non-Bayesian approach to identify the order of univariate moving average processes is what so called the automatic approach. Assuming the maximum order of the process is known, the automatic approach is to fit all possible moving average models and compute a certain criterion for each model. The identified model is then the model for which the proposed criterion attains its optimal value. However, one may say that there is no complete agreement on the form of the criterion to be optimized. Examples of the automatic identification technique may be found in Akaike (1974), Schwarz (1978), Rissanen (1978), Hannan and Quinn (1979), Terasvirta and Mellin (1986), Mills and Passad (1992), and Beveridge and Oickle (1994). A third non-Bayesian approach is called a goodness of fit approach. Assuming the correct value of the order q , the asymptotic distribution of the residual correlations have a zero mean. Box and Pierce (1970) proposed a chi-square test to identify the order of the process. The Chi-square test has been modified and improved by Ljung and Box (1978).

On the other hand, Bayesian identification of univariate moving average processes is being developed and the Bayesian literature devoted to the identification of moving average processes is sparse. One of the most important Bayesian contribution has been done by Monahan (1983) who found a way to work with univariate autoregressive moving average processes with low order. Assuming stationarity and invertibility, he has developed a numerical technique to implement the identification phase of low order autoregressive moving average processes which include moving average processes as a special case. Broemeling and Shaarawy (1988) have developed an approximate technique to identify the order of ARMA processes. Shaarawy and Ali (2003) have developed a direct Bayesian technique to identify seasonal autoregressive models. Shaarawy et. al. (2007) found a direct way to identify the order of moving average processes. They developed an approximate posterior probability mass function of the model order in a convenient form. Then, one may choose the order at which the posterior probability mass function attains its maximum to be the identified order of the time series being analyzed.

With respect to the two dimensional and multivariate versions, the identification problem, from a non-Bayesian view point, has been studied by Granger and Newbold (1973), Tiao and Box (1981) and Tiao and Tsay (1983) by matching the cross correlation functions of the data with their theoretical counterparts. On the other hand, Shaarawy et. al. (2006)

have introduced a direct Bayesian technique to identify the order of bivariate autoregressive processes. Their technique has been extended by Shaarawy and Ali (2008) to the case of general vector autoregressive processes. However, one may say that the direct Bayesian approach to identify the order of two dimensional moving average processes has not been explored yet.

The main objective of this article is to develop an approximate direct Bayesian technique to determine or estimate the order of two dimensional moving average processes. Using n vectors of 2×1 observations, the marginal posterior probability mass function of the model order is derived in an approximate convenient form using an approximate likelihood function and a matrix normal-Wishart prior class of densities or a vague prior. Then the time series analyst may inspect the approximate marginal posterior mass function of the model order over its grid and selects the order with the maximum posterior probability to be the identified or estimated order. A simulation study is conducted, using the modern specialized package SCA, in order to inspect the numerical efficiency of the proposed Bayesian technique and compare it with some popular non-Bayesian technique.

The rest of this paper is structured as follows: Section 2 introduces the definition of the two dimensional moving average model of order q in matrix notations. Section 3 presents an indirect Bayesian technique to identify an initial order for the two dimensional moving average processes. Section 4 develops an approximate marginal posterior probability mass function of the order q and explains the idea of the proposed identification technique. Section 5 is devoted to investigate the numerical efficiency of the proposed Bayesian technique in solving the identification problems. The performance of the proposed Bayesian technique is compared with the AIC in the same section.

2. Two Dimensional Moving Average Processes

As Box and Tiao (1981) have shown, a very useful class of models is the two dimensional moving average processes, denoted by $MA_2(q)$ for short, defined by

$$\underline{y}(t) = \theta_q(B) \varepsilon(t) \quad (2.1)$$

where $\underline{y}(t)$ is the vector of (2×1) observations at time t , $t = \dots, -1, 0, 1, 2, \dots$, $\varepsilon(t)$ is a sequence of independent and normally distributed (2×1) unobservable random vector with zero mean and a (2×2) unknown precision matrix T ,

$$\theta_q(B) = I - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

is a matrix polynomial in B , the θ_j 's are (2×2) matrices of real unknown constants, q is a positive integer, and the backshift operator B is such that

$$B^j y(t) = y(t - j) \quad j = 1, 2, \dots$$

The $MA_2(q)$ processes are always stationary no matter the parameters matrices θ_j 's are, and are invertible if the roots of the determinant $|\theta_q(B)|$, which is a polynomial in B , lie outside the unit circle. The parameter space is constrained by invertibility conditions for identifiability reasons, see Box and Jenkins (1970) and Box and Tiao (1981).

In general, one can write the $MA_2(q)$ process as

$$\underline{y}(t) = \varepsilon(t) - \theta_1 \varepsilon(t-1) - \theta_2 \varepsilon(t-2) - \dots - \theta_q \varepsilon(t-q)$$

Thus, if we have n observations, we can express the general class of $MA_2(q)$ model as

$$Y = X_1 \theta'_1 + X_2 \theta'_2 + \dots + X_q \theta'_q + U \quad (2.2)$$

where

$$Y = \begin{bmatrix} y(1) & y(2) & \dots & y(n) \end{bmatrix}' = \begin{bmatrix} y(1,1) & y(1,2) \\ y(2,1) & y(2,2) \\ \vdots & \vdots \\ y(n,1) & y(n,2) \end{bmatrix}_{n \times 2}$$

$$X_1 = \begin{bmatrix} -\varepsilon'(0) \\ -\varepsilon'(1) \\ \vdots \\ -\varepsilon'(n-1) \end{bmatrix}_{n \times 2}, X_2 = \begin{bmatrix} -\varepsilon'(-1) \\ -\varepsilon'(0) \\ \vdots \\ -\varepsilon'(n-2) \end{bmatrix}_{n \times 2}, \dots, X_q = \begin{bmatrix} -\varepsilon'(1-q) \\ -\varepsilon'(2-q) \\ \vdots \\ -\varepsilon'(n-q) \end{bmatrix}_{n \times 2}$$

In more compact form, one may write the general class of $MA_2(q)$ models as

$$Y_{n \times 2} = X_{n \times 2q} \Theta_{2q \times 2} + U_{n \times 2} \quad (2.3)$$

where

$$X = \begin{bmatrix} X_1 & X_2 & \dots & X_q \end{bmatrix} = \begin{bmatrix} -\varepsilon'(0) & -\varepsilon'(-1) & \dots & -\varepsilon'(1-q) \\ -\varepsilon'(1) & -\varepsilon'(0) & \dots & -\varepsilon'(2-q) \\ \vdots & \vdots & \vdots & \vdots \\ -\varepsilon'(n-1) & -\varepsilon'(n-2) & \dots & -\varepsilon'(n-q) \end{bmatrix},$$

$$\Theta = \begin{bmatrix} \theta'_1 \\ \theta'_2 \\ \vdots \\ \theta'_q \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(1)} & \theta_{21}^{(1)} \\ \theta_{12}^{(1)} & \theta_{22}^{(1)} \\ \dots & \dots \\ \theta_{11}^{(2)} & \theta_{21}^{(2)} \\ \theta_{12}^{(2)} & \theta_{22}^{(2)} \\ \dots & \dots \\ \vdots & \vdots \\ \theta_{11}^{(q)} & \theta_{21}^{(q)} \\ \theta_{12}^{(q)} & \theta_{22}^{(q)} \end{bmatrix}_{2q \times 2}$$

$$U = \begin{bmatrix} \varepsilon(1) & \varepsilon(2) & \dots & \varepsilon(n) \end{bmatrix}' \quad \text{and} \quad \varepsilon(t) = \begin{bmatrix} \varepsilon(t,1) & \varepsilon(t,2) \end{bmatrix}'$$

The expression (2.3) represents the general class of two dimensional moving average models of order q . This class of model is quite useful in modelling and forecasting a lot of two dimensional time series arise in many fields of application. In real applications, the order q is usually unknown and should be identified or determined using the observed data

about the two time series being analyzed. For Bayesians the order q is considered, in effect, as additional unknown parameter for which the posterior probability mass function should be developed in a convenient form. The main objective of this research is to develop such mass function.

In order to achieve our main objective, let $S_n = [y(1) \ y(2) \ \dots \ y(n)]'$ be a matrix of $n \times 2$ observations generated from the two dimensional moving average process of order q on the form (2.3) where the order q is an unknown positive integer. Assuming that $\varepsilon(1-q) = \varepsilon(2-q) = \dots = \varepsilon(0) = \underline{0}$, (see Hillmer and Tiao (1979)), the likelihood function of the parameters $\Theta(q)$, q , and T can be written as

$$L(\Theta(q), q, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr} \sum_{t=1}^n \varepsilon(t) \varepsilon'(t) T\right) \quad (2.4)$$

where $\Theta(q) \in \mathbb{R}^{2q \times 2}$, $T > 0$, $q = 1, 2, \dots, Q$, and Q is the maximum order of the process. In general the likelihood function (2.4) is very complicated because the disturbances $\varepsilon(t)$ are non-linear function of the coefficients Θ . To see this, one may write the disturbances vectors, form (2.2), as

$$\varepsilon'(t) = y'(t) - x_q'(t-1) \Theta(q), \quad t = 1, 2, \dots, n \quad (2.5)$$

where

$$x_q'(t-1) = [-\varepsilon'(t-1) \ -\varepsilon'(t-2) \ \dots \ -\varepsilon'(t-q)]$$

Thus, one may write the m^{th} component of the residual vector $\varepsilon(t)$ as

$$\varepsilon(t, m) = y(t, m) + \sum_{k=1}^q \sum_{j=1}^2 \theta_{mj}^{(k)} \varepsilon(t-k, j), \quad m = 1, 2. \quad (2.6)$$

The expression (2.6) is a recurrence relation for the residuals. This recurrence causes the main problem in developing the exact Bayesian solution for the identification problem of the two dimensional moving average processes. However (2.6) can be used to evaluate the residuals recursively if one knows Θ and initial value of the residuals.

Using (2.6), the likelihood function (2.4) can be rewritten in the form

$$L(\Theta(q), q, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr} \sum_{t=1}^n H(\Theta(q), q, t) T\right) \quad (2.7)$$

where $H(\Theta(q), q, t) = (h_{rs})$ is 2×2 matrix and

$$h_{rs} = \left[y(t, r) + \sum_{k=1}^q \sum_{j=1}^2 \varepsilon(t-k, j) \theta_{rj}^{(k)} \right] \left[y(t, s) + \sum_{k=1}^q \sum_{j=1}^2 \varepsilon(t-k, j) \theta_{sj}^{(k)} \right] \quad (2.8)$$

The form (2.8) is not quadratic in the parameters θ 's because $\varepsilon(t-k, j)$ is a function of θ 's through the recurrence (2.6). If $\varepsilon(t-k, j)$ are known, $H(\Theta(q), q, t)$ would be a quadratic form in the parameters. The proposed approximation is based on replacing the exact residuals $\varepsilon(t)$ by their least squares estimates. The least square estimates, say $\hat{\varepsilon}(t)$, are obtained by searching over the parameter space for the value of Θ , say Θ_0 , which minimizes the

residual sum of squares $\sum_{t=1}^n \hat{\varepsilon}^2(t, i)$, $i = 1, 2$. Before doing this process, one should have an initial adequate value for the order q . It is proposed to obtain such value, say q_0 , by the indirect technique presented in the next section. The least squares estimate $\Theta_0(q_0)$ and the assumed initial values, namely zero, are then substituted in (2.6) to obtain the least squares estimates of the residuals recursively. Substituting these estimates in $\hat{x}'_q(t-1)$, one can write (2.5) as

$$\hat{\varepsilon}'(t) = y'(t) - \hat{x}'_{q_0}(t-1) \Theta(q) \quad (2.9)$$

where $t = 1, 2, \dots, n$ and $\hat{x}'_q(t-1)$ is the same as $\hat{x}'_q(t-1)$ but using the residuals estimates instead of the exact ones,

Using the estimates of the residuals, one may rewrite the likelihood function (2.7) approximately as

$$L^*(\Theta(q), q, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr} \left\{ \sum_{t=1}^n [\underline{y}(t) - \Theta'(q) \hat{x}_q(t-1)][\underline{y}(t) - \Theta'(q) \hat{x}_q(t-1)]' T \right\}\right) \quad (2.10)$$

A convenient choice of the conditional prior of $\Theta(q)$ given q and T is

$$f_1(\Theta(q) | q, T) = (2\pi)^{\frac{-4q}{2}} |R(q)|^{\frac{2}{2}} |T|^{\frac{2q}{2}} \exp\left(-\frac{1}{2} \text{tr} [\Theta(q) - D(q)]' R(q) [\Theta(q) - D(q)] T\right) \quad (2.11)$$

where the hyperparameters $D(q) \in \mathbb{R}^{2q \times 2}$ and $R(q)$ is a $(2q \times 2q)$ positive definite matrix. The precision matrix T is assigned, a priori, the following Wishart distribution

$$f_2(T) \propto |T|^{(a-3)/2} \exp\left(-\frac{1}{2} \text{tr} \mathbf{V} T\right)$$

where $\mathbf{V}$ is a (2×2) positive definite matrix. Thus the joint prior distribution of $\Theta(q)$ and T given q is assumed to be

$$f_3(\Theta(q), T | q) \propto f_1(\Theta(q) | q, T) f_2(T) \quad (2.12)$$

The class of prior distribution (2.12) is called matrix normal-Wishart class of distributions. Let β_i be the prior probability that the time series realization Y is generated from a two dimensional moving average process of order i . Thus the prior probability mass function of q is

$$\beta_i = P_r[q = i], \quad i = 1, 2, \dots, Q \quad (2.13)$$

From (2.12) and (2.13), the joint prior distribution of the parameters $\Theta(q)$, T , and q is

$$f(\Theta(q), q, T) \propto \beta_i (2\pi)^{\frac{-4q}{2}} |R(q)|^{\frac{2}{2}} |T|^{\frac{1}{2}[2q+a-3]} \exp\left(-\frac{1}{2} \text{tr} \{[\Theta(q) - D(q)]' R(q) [\Theta(q) - D(q)] + \mathbf{V}\} T\right) \quad (2.14)$$

If one can not or unwilling to specify the hyperparameters $D(q)$, $R(q)$, a , $\mathbf{V}$ and β_i , one might use Jeffrey's vague prior

$$f(\Theta(q), q, T) \propto |T|^{-3/2} \quad (2.15)$$

3. Indirect Bayesian Identification Technique

The $MA_2(q)$ class of models is quite useful in modeling two dimensional time series data and frequently q is not excess of 4. In practice the value of the model order q is unknown and one has to identify it using the observed n vectors of 2×1 observations. The direct Bayesian approach to identify the order q is to find the marginal posterior probability mass function of the unknown q and use its mode to be the identified value. The approach taken here is somewhat different from the direct approach. Instead of working directly with the posterior distribution of q , it is proposed to focus on the marginal posterior distribution of the coefficients

$$\Theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_Q \end{bmatrix}_{2Q \times 2} \quad (3.1)$$

where θ_i is a square matrix of order 2, $i = 1, 2, \dots, Q$. The maximum order Q is assumed to be known.

Assuming that $\varepsilon(0) = \varepsilon(-1) = \dots = \varepsilon(1 - Q) = 0$, the likelihood function of the parameters Θ and T is

$$L(\Theta, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left\{\sum_{t=1}^n \varepsilon(t) \varepsilon'(t) T\right\}\right) \quad (3.2)$$

where $\Theta \in \mathbb{R}^{2Q \times 2}$, T is positive definite matrix, and

$$\varepsilon'(t) = y'(t) - x'(t-1) \Theta, \quad t = 1, 2, \dots, n \quad (3.3)$$

where

$$x'(t-1) = \begin{bmatrix} -\varepsilon'(t-1) & -\varepsilon'(t-2) & \dots & -\varepsilon'(t-Q) \end{bmatrix}$$

The expression (3.3) is a recurrence relation for the residuals and the m -th component of the residual $\varepsilon(t)$ can be written as

$$\varepsilon(t, m) = y(t, m) + \sum_{k=1}^Q \sum_{j=1}^2 \theta_{mj}^{(k)} \varepsilon(t-k, j), \quad m = 1, 2 \quad (3.4)$$

The recurrence relation (3.4) causes the main problem in developing the exact Bayesian analysis of two dimensional moving average processes. However, this recurrence may be used to evaluate the residuals recursively if one knows Θ and the initial values of the residuals. The proposed approximation is based on replacing the exact residuals by their least squares estimated and assuming that $\varepsilon(0) = \varepsilon(-1) = \dots = \varepsilon(1 - Q) = 0$. Thus, we estimate the residuals recursively by

$$\hat{\varepsilon}(t, m) = y(t, m) + \sum_{k=1}^Q \sum_{j=1}^2 \hat{\varepsilon}(t-k, j) \hat{\theta}_{mj}^{(k)} \quad (3.5)$$

where $t = 1, 2, \dots, n$; $m = 1, 2$, and $\hat{\theta}_{mj}^{(k)}$ are the nonlinear least squares estimates of $\theta_{mj}^{(k)}$. Using the estimates of the residuals, one may rewrite the likelihood function (3.2) approximately as

$$L^*(\Theta, T | S_n) \propto |T|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr} \left\{ \sum_{t=1}^n [\underline{y}(t) - \Theta' \hat{x}(t-1)][\underline{y}(t) - \Theta' \hat{x}(t-1)]' T \right\}\right) \quad (3.6)$$

where $\hat{x}(t-1)$ is the same as $x(t-1)$ but using the estimate residuals instead of the exact ones.

A convenient choice of the prior density of the parameters Θ and T is the following matrix normal-Wishart distribution:

$$\xi(\Theta, T) \propto \xi_1(\Theta | T) \xi_2(T) \quad (3.7)$$

where

$$\xi_1(\Theta | T) \propto |T|^{\frac{2Q}{2}} \exp\left(-\frac{1}{2} \text{tr} [\Theta - D]' W [\Theta - D] T\right)$$

and

$$\xi_2(T) \propto |T|^{(a-3)/2} \exp\left(-\frac{1}{2} \text{tr} \mathbf{V} T\right)$$

where the hyperparameters $D \in \mathbb{R}^{2Q \times 2}$, W is a $(2Q \times 2Q)$ positive definite matrix, $a > 0$ and Σ is a (2×2) positive definite matrix. If one has "little" information about the parameters, a priori, he may use the Jeffreys' prior

$$\xi(\Theta, T) \propto |T|^{-3/2}, \quad \Theta \in \mathbb{R}^{2Q \times 2}, T > 0 \quad (3.8)$$

Theorem 3.1. *Using the approximate likelihood function (3.6) and the matrix normal-Wishart prior density (3.7), the marginal posterior distribution of Θ is a matrix t with parameters $(A^{-1} B, A^{-1}, C - B' A^{-1} B, \mathbf{v})$ where*

$$\begin{aligned} A &= W + \sum_{t=1}^n \hat{x}(t-1) \hat{x}'(t-1), \\ B &= WD + \sum_{t=1}^n \hat{x}(t-1) y'(t), \\ C &= D' W D + \mathbf{V} + \sum_{t=1}^n y(t) y'(t), \end{aligned}$$

and $\mathbf{v} = n + a - 1$

Corollary 3.1. *Using the approximate likelihood function (3.6) and the Jeffreys' prior density (3.8), the marginal posterior distribution of Θ is a matrix t with parameters $(A^{-1} B, A^{-1}, C - B' A^{-1} B, \mathbf{v})$. However, A , B , C and $\mathbf{v}$ will be modified by letting $W \rightarrow 0_{2Q \times 2Q}$, $a \rightarrow -2Q$, and $\mathbf{V} \rightarrow 0_{2 \times 2}$.*

The reader is referred to Box and Tiao (1973) for the form and the properties of the matrix t density function. Since Θ has a matrix t distribution, any subset of 2 rows has a matrix t distribution. Also, the conditional distribution of a subset of rows given any other subset of rows is a matrix t distribution, see Box and Tiao (1973). Furthermore, one can test any subset of rows to be zero (marginally or conditionally) using exact F statistics. The forms of the F statistics can be found in Box and Tiao (1973, pp. 451-453).

The proposed indirect Bayesian procedure to identify the order q is to do the backward elimination procedure. One first tests $H_0 : \theta_Q = 0$ versus $\theta_Q \neq 0$ using the marginal posterior distribution of θ_Q which is a matrix t distribution. If H_0 is not rejected, one tests $H_0 : \theta_{Q-1} = 0$ given $\theta_Q = 0$ using the conditional posterior distribution of θ_{Q-1} given $\theta_Q = 0$ which is also a matrix t distribution. If the last H_0 is not rejected, one tests $H_0 : \theta_{Q-2} = 0$ given $\theta_Q = \theta_{Q-1} = 0$ using the conditional posterior distribution of θ_{Q-2} given $\theta_Q = \theta_{Q-1} = 0$ which is also a matrix t distribution. The procedure is continued in this fashion until the hypothesis $H_0 : \theta_{q_0} = 0$ is rejected for some q_0 where $0 < q_0 \leq Q$. The value q_0 is the proposed indirect Bayesian solution of the two dimensional identification problem and the proposed initial value of our proposed direct Bayesian approach.

4. The Posterior Probability Mass Function of the Model Order

The main objective of this section is to develop an approximate posterior probability mass function of the order q under the conditions outlined above. In order to do that, one must combine the approximate likelihood function (2.10), via Bayes theorem, with the prior distribution of the parameters (2.14). Thus the joint posterior distribution of the parameters $\Theta(q)$, T , and q is

$$\begin{aligned} g(\Theta(q), T, q | S_n) \propto & \beta_i (2\pi)^{-\frac{4q}{2}} |R(q)|^{\frac{2}{2}} |T|^{\frac{\alpha(q)}{2}} \\ & \exp\left(-\frac{1}{2} \text{tr} \{[\Theta(q) - D(q)]' R(q) [\Theta(q) - D(q)] + \mathbf{V}\right. \\ & \left. + \sum_{t=1}^n [y(t) - \Theta'(q) \hat{x}_q(t-1)][y(t) - \Theta'(q) \hat{x}_q(t-1)]'\} T\right) \end{aligned} \quad (4.1)$$

where

$$\alpha(q) = n + 2q + a - 3 \quad (4.2)$$

Theorem 4.1. *Using the approximate likelihood function (2.10) and prior distribution (2.14), the marginal posterior probability mass function of the order q is*

$$h(q | S_n) \propto \beta_i |R(q)| |A(q)|^{-1} |C(q)|^{-\frac{1}{2}(n+a)} \prod_{j=1}^k \Gamma\left(\frac{n+a+j-2}{2}\right), \quad n > 1-a$$

where

$$A(q) = R(q) + \sum_{t=1}^n \hat{x}_q(t-1) \hat{x}_q'(t-1) \quad (4.3)$$

$$B(q) = R(q) D(q) + \sum_{t=1}^n \hat{x}_q(t-1) y'(t) \quad (4.4)$$

and

$$C(q) = D'(q) R(q) D(q) + \mathbf{V} + \sum_{t=1}^n y(t) y'(t) - B'(q) A^{-1}(q) B(q) \quad (4.5)$$

It might be noted that the proof of Theorem (4.1.) can be done by integrating (4.2) with respect to $\Theta(q)$ and T respectively. However, the integral with respect to $\Theta(q)$ is done using the matrix normal integral and the integral with respect to T is done using the generalized gamma function. The reader is referred to Box and Tiao (1973) for the forms of the matrix normal density and the generalized gamma function.

Theorem 4.2. *Using the approximate likelihood function (2.10) and the non-informative prior distribution (2.15), the marginal posterior probability mass function of the model order q is*

$$h_1(q|S_n) \propto (\pi)^{2q} |A^*(q)|^{-1} |C^*(q)|^{-\frac{1}{2}[n-2q]} \prod_{j=1}^2 \Gamma\left(\frac{n+j-2q-2}{2}\right), \quad n > 1+2q$$

where

$$A^*(q) = \sum_{t=1}^n \hat{x}_q(t-1) \hat{x}_q'(t-1),$$

$$B^*(q) = \sum_{t=1}^n \hat{x}_q(t-1) y'(t),$$

and

$$C^*(q) = \sum_{t=1}^n y(t) y'(t) - B^{*'}(q) A^{*-1}(q) B^*(q)$$

The proof of Theorem (4.2.) can be done using the same procedures used to prove Theorem(4.1.).

The forms of the posterior probability mass function given by Theorem (4.1.) and (4.2.) are convenient and could be easily programmed. Then one may inspect the posterior probability mass function over the grid of the order q and select the value of q at which the posterior mass function attains its maximum to be the identified order of the two dimensional time series data being analyzed.

5. A Numerical Study

One of the main objectives of this research is to study the effectiveness of the proposed technique in identifying the order of two dimensional moving average processes. In order to achieve this goal, six simulations have been conducted. The proposed technique is employed, with three different prior distributions, to identify the order of MA₂(1) model with various parameter values. All computations were performed by using the modern package SCA.

Our main concern is to study the effectiveness of the proposed Bayesian technique by calculating the percentage of correct identification. Such effectiveness will be examined with respect to the time series length (sample size) as well as the parameters of the selected

model. For all sample sizes and parameters, the covariance matrix of the noise term is fixed at $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Simulation 1, as an illustration, begins with the generation of pairs of 500 data sets of bivariate normal variates, each of size 300 to represent $\varepsilon_1(t)$ and $\varepsilon_2(t)$ respectively. These data sets are then used to generate a pair of 500 realizations, each of size 300, from $MA_2(1)$ process with $\Theta = \begin{pmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{pmatrix}$ assuming the starting values are zero. For a specific prior, the second step of simulation 1 is to carry out all computations, assuming the maximum order $Q = 4$, required to identify each of the 500 realizations and finding the percentage of correct identification. Note that the computations include the application of the indirect Bayesian technique in order to get an adequate initial value q_0 , and use it on applying our proposed direct Bayesian technique. Such computations are done for a specific time series length n using the first n observations of each generated set. This second step is repeated for each chosen sample size, maximum order, and prior combinations. The sample size is taken to be 50, 100, 150, 200 and 300, while the maximum order Q is taken to be 4. With respect to the prior probability mass of the order q , which is combined with the vague prior $\Theta(q)$ and T , the following three priors are used

Prior 1 : $\xi(q) = \frac{1}{Q}$, $q = 1, 2, \dots, Q$

Prior 2 : $\xi(q) \propto (0.5)^q$, $q = 1, 2, \dots, Q$

Prior 3 : $\xi(q=1) = 0.4$; $\xi(q=2) = 0.3$; $\xi(q=3) = 0.2$; $\xi(q=4) = 0.1$,

The first prior assigns equal probabilities to the all possible values of the order q . The second prior is chosen in such a way to give probabilities that decline exponentially with the order, while the third prior is chosen in such a way to give probabilities that decrease with an absolute amount 0.1 as the order increases.

Simulation 2 is done in a similar way but using $\Theta = \begin{pmatrix} -0.5 & 0.4 \\ 0.3 & -0.2 \end{pmatrix}$. The results of simulations 1 and 2 are presented in Table 1. The other simulations are done in similar way but using different values of Θ and their results are presented in Tables 2 and 3. The parameters of the simulations have been chosen in such a way to satisfy the invertibility conditions, see Harvey (1981). In some simulations, the parameters are chosen to be close from non-invertibility region, see Table 3.

An inspection of the numerical results shows an increasing trend for the efficiency of the proposed Bayesian technique as the sample size n increases. The percentages of correct identification are reasonably high, being greater than 68%, for all models and priors, for samples of size 150 or more no matter what the parameters are. In addition, the percentages of correct identification achieved using the third prior is higher than the corresponding percentages achieved by the first prior, while the corresponding percentages achieved by the second prior are the highest. Finally, for sufficiently large n , it is to be noticed that the posterior results are not very sensitive to the minor changes between the second and third prior distributions.

For the matter of comparison, the well known Akaike's information criterion, denoted by AIC, has been used to identify the data sets, which have been used with simulations 1-6, and the percentages of choosing the correct orders have been calculated. For all considered models, the results of the proposed technique are less than the results of AIC if a small

Table 1. Percentages of Correct Identification of the proposed Bayesian Technique and AIC for MA₂(1) Processes for Simulations 1 and 2

PARAMETERS	n	PRIOR1	PRIOR2	PRIOR3	AIC
$\Theta = \begin{bmatrix} 0.5 & -0.4 \\ -0.3 & 0.2 \end{bmatrix}$	50	53.8	60.4	57.2	66.2
	100	76.4	82.6	79.8	83.2
	150	88.4	92.2	90.4	84.4
	200	92.8	95.2	94.0	89.2
	300	93.6	94.6	94.2	90.0
$\Theta = \begin{bmatrix} -0.5 & 0.4 \\ 0.3 & -0.2 \end{bmatrix}$	50	56.6	61.6	59.2	73.6
	100	75.4	81.8	78.8	82.8
	150	87.4	90.8	88.8	85.2
	200	93.4	95.6	94.6	88.6
	300	96.2	97.0	96.8	88.2

Table 2. Percentages of Correct Identification of the proposed Bayesian Technique and AIC for MA₂(1) Processes for Simulations 3 and 4

PARAMETERS	n	PRIOR1	PRIOR2	PRIOR3	AIC
$\Theta = \begin{bmatrix} 0.3 & 0.5 \\ 0.4 & 0.3 \end{bmatrix}$	50	63.4	68.0	65.4	70.0
	100	75.2	81.4	79.2	81.6
	150	88.8	92.6	91.0	83.6
	200	95.2	96.4	95.6	86.4
	300	95.8	97.0	96.8	88.6
$\Theta = \begin{bmatrix} 0.7 & 0.5 \\ 0.4 & 0.3 \end{bmatrix}$	50	64.6	69.8	67.4	76.4
	100	76.8	83.0	79.8	84.6
	150	85.6	89.6	88.2	85.8
	200	88.8	92.0	90.8	87.0
	300	88.6	90.6	89.6	88.4

sample size is used. As the sample size increases, the results of the proposed technique become superior to the results of AIC. It might be important to mention that the correct percentages achieved by the proposed technique, with simulations 1-4, were higher than

Table 3. Percentages of Correct Identification of the proposed Bayesian Technique and AIC for MA₂(1) Processes for Simulations 5 and 6

PARAMETERS	n	PRIOR1	PRIOR2	PRIOR3	AIC
$\Theta = \begin{bmatrix} 1 & -0.2 \\ 1.1 & -1 \end{bmatrix}$	50	52.6	58.6	54.6	76.0
	100	58.0	64.6	62.4	80.6
	150	68.4	74.2	71.6	84.4
	200	80.0	83.6	82.0	85.0
	300	84.8	88.2	86.4	88.2
$\Theta = \begin{bmatrix} 0.9 & -0.2 \\ 1.1 & -0.9 \end{bmatrix}$	50	46.8	52.2	48.6	71.8
	100	58.0	64.0	60.2	82.4
	150	76.6	81.8	79.8	84.2
	200	84.8	88.2	86.8	88.6
	300	89.8	92.8	91.4	88.4

the correct percentages achieved by AIC for $n \geq 200$, regardless of the prior distribution.

In general the numerical results support the adequacy of using the proposed Bayesian technique in solving the identification problems of two dimensional moving average processes. The percentages of correct identification depends on the sample size and the choice of the prior distribution. A moderate sample size is needed in order to have a better chance of getting correct identification no matter what the parameters are.

Summary and Conclusions

The main objective of this article is to develop a Bayesian methodology to choose the order q of two dimensional moving average processes. The marginal posterior probability mass function of the model order is developed in a convenient form using an approximate conditional likelihood function and a matrix- normal Wishart or Jeffreys' vague prior. Then one may easily calculate the posterior probabilities of all possible values of the order q and choose the value of q at which the posterior probability mass function attains its maximum to be the identified order.

In order to demonstrate the performance of the proposed Bayesian procedure and test its adequacy in identifying the order of two dimensional moving average processes, a simulation study with three different prior mass functions has been carried out. The numerical results show that the proposed procedure can efficiently identify the orders of two dimensional moving average processes with high precision for moderate and large sample size.

Acknowledgment

The project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant no. 3-013/429. The authors, therefore, acknowledge with thanks DSR support for Scientific Research.

References

- [1] Akaike, H. (1974). *A New Look at Statistical Model Identification*. IEEE Transaction on Automatic Control, Vol. 19, pp. 716-723.
- [2] Beveridge, S. and Oickle, C. (1994). *A Comparison of Box-Jenkins and Objectives Methods for Determining the Order of a Non-seasonal ARMA Model*. Journal of Forecasting, Vol. 13, pp. 419-434, (1994).
- [3] Box, G. and Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco.
- [4] Box, G. and Pierce, D.A. (1970). *Distribution of residual autocorrelation in autoregressive integrated moving average time series models*. J. of the American Stats. Associ., Vol. 65, pp. 1509-1526.
- [5] Box, G. and Tiao, G. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, MA.
- [6] Broemling, L.D. and Shaarawy, S. (1988). *Time Series Analysis: Bayesian Analysis in the Time Domain*. Bayesian Analysis of Time Series and Dynamic Models, edited by Spall, J.
- [7] Chatfield, C. (1980). *The Analysis of Time Series: Theory and Practice*. Chapman and Hall Ltd., London.
- [8] Granger, C.W.J and Newbold, P. (1973). *Forecasting Economic Time Series*. New York: Academic Press.
- [9] Hannan, E. J. and Quinn, B. G. (1979). *The Determination of an Autoregression*. Journal of the Royal Statistical Society, Series B, Vol. 42, pp. 190-195.
- [10] Harvey, A.C. (1981). *Time Series Models*. Wiley, New York.
- [11] Hillmer, S.C. and Tiao, G.C. (1979). *Likelihood function of stationary multiple autoregressive moving average models*. J. Am. Stat. Assoc. 74, 652-660.
- [12] Liu, L.M. (2009). *Time Series Analysis and forecasting*. 2nd edition. Scientific Computing Association Corp.
- [13] Ljung, G. and Box, G.E. (1978). *On a measure of lack of fit in time series models*. Biometrika, Vol. 65, pp. 297-303.

-
- [14] Mills, J. and Prasad, K.(1992). *A Comparison of Model Selection Criteria*. Econometric Reviews, Vol. 11, p. 201-233, (1992).
 - [15] Monahan, J.E. (1983). *Fully Bayesian Analysis of ARIMA Time Series Models*. Journal of Econometrics, Vol. 21, pp. 307-331.
 - [16] Rissanen, J. (1978). *Modelling by Shortest Data Description*. Automatica, Vol. 14, pp. 465-471.
 - [17] Schwarz, G. (1978). *Estimating the Dimension of a Model*. Annals of the Institute of Statistical Mathematics, Vol. 6, pp. 461-464.
 - [18] Shaarawy, S. and Ali, S. (2003). *Bayesian Identification of Seasonal Autoregressive Models*. Communications in Statistics- Theory and Methods, Vol. 32, Issue 5, pp. 1067-1084.
 - [19] Shaarawy, S., Albassam, M. and Ali S. (2006). *A Direct Bayesian Identification to Select the Order of Bivariate Autoregressive Processes*. The Egyptian Statistical Journal, Institute of Statistical Studies and Research, Cairo University, Vol. 50, Issue 1.
 - [20] Shaarawy, S., Soliman, E. and Ali S. (2007). *Bayesian Identification of Moving Average Models*. Communications in Statistics- Theory and Methods, Vol. 36, Issue 12.
 - [21] Shaarawy, S. and Ali S.(2008). *Bayesian Identification of Multivariate Autoregressive Processes*. Communications in Statistics- Theory and Methods, Vol. 37, Issue 5, (2008).
 - [22] Terasvirta, T. and Mellin, I. (1986). *Model Selection Criteria and Model Selection Tests in Regression Models*. Scandinavian Journal of Statistics: Theory and Applications, Vol. 13, pp. 159-171.
 - [23] Tiao, G. and Box, G. (1981). *Modeling Multiple Time Series with Multiplications*. Journal of the American Statistical Association, Vol. 76, No. 376, pp.802-816.
 - [24] Tiao, G. C. and Tsay, R.(1983). *Multiple Time Series Modeling and Extended Sample Cross Correlation*. Journal of Business and Economic Statistics, Vol.1, No.1, (1983).

ON THE USE OF MOMENT ESTIMATORS TO TEST FOR A ZERO INFLATED POISSON (ZIP) DISTRIBUTION

A. Nanthakumar¹ and Z. Govindarajulu²

¹Department of Mathematics, SUNY-Oswego, Oswego, NY 13126

²Department of Statistics, University of Kentucky, Lexington, KY 40506

Abstract

The paper studies two simple approaches that are very powerful and computationally less intensive to test poisson versus zero inflated poisson distributions.

AMS Subject Classification: 62F03.

Keywords: Moment Estimator, Hypothesis Test, Distance, Zero Inflated Poisson.

1. Introduction

For a long time, the occurrence of large number of zeroes was observed in the context of counting processes which otherwise could be easily found to follow a regular Poisson distribution. This phenomenon occurs regularly in physical and biological systems where the Poisson counts appear when a system functions properly and additional zeroes are observed when the system malfunctions. Cochran (1954) was the first to develop a hypothesis-test for testing Poissonness against ZIP and later Rao and Chakravarty (1956), Singh (1963) developed their own tests for testing Poissonness against ZIP. The test developed by Cochran (1954) was later shown to be equivalent to the score test developed by Van den Broeck (1995). El-Shaarawi (1985) conducted a simulation study for comparing the power of all the tests that existed until that time to test for Poissonness, and based on his study, he concluded that the score test is more powerful while being computationally less intensive. Most of the hypothesis tests, for example by Lambert (1992), Van den Broeck (1995) were developed in the context of regression models where the parameters depend on the covariates. Recently, Thas and Rayner (2005) developed a partial sum based score tests for testing Poissonness against a class of alternatives (including ZIP as a possible alternative) and their test seems to be reasonably powerful.

In this paper, we present two simple tests which are very simple and at the same time very powerful. One of these tests use the idea of "distance" to perform the test. The other one is a locally most powerful test (LMP). The LMP test (also called likelihood derivative test) is due to Rao (1948). According to Moran (1970), the LMP test is very powerful for

testing mixing proportions. One can describe a ZIP distribution (indeed a mixture distribution) as follows.

$$P(Y=0) = \theta + (1-\theta)e^{-\lambda} \quad (1)$$

$$P(Y=y) = \theta + (1-\theta)\frac{e^{-\lambda}\lambda^y}{y!}, \quad y > 0 \quad (2)$$

where θ is the mixing proportion.

2. Methodology

We use the method of moments for deriving the estimator. For the above described ZIP distribution, one can easily show that

$$E(Y) = (1-\theta)\lambda \quad (3)$$

and

$$E(Y^2) = (1-\theta)(\lambda + \lambda^2) \quad (4)$$

This clearly implies that the moment estimator of λ and θ are as follows.

$$\hat{\lambda}_n = \left(\frac{\sum_{i=1}^n y_i^2}{\frac{1}{n}} \right) - 1 \quad (5)$$

and

$$\hat{\theta}_n = 1 - \left(\frac{\left(\sum_{i=1}^n y_i \right)^2}{n \left(\sum_{i=1}^n y_i^2 - \sum_{i=1}^n y_i \right)} \right) \quad (6)$$

We are led to the following results as $n \rightarrow \infty$.

Theorem 1.

$\left(\frac{\sum_{i=1}^n y_i^2}{n}, \frac{\sum_{i=1}^n y_i}{n} \right)$ follows a bivariate normal distribution with mean $= (\mu_1, \mu_2)$ and

variance-covariance matrix $\begin{bmatrix} \frac{\sigma_{11}}{n} & \frac{\sigma_{12}}{n} \\ \frac{\sigma_{12}}{n} & \frac{\sigma_{22}}{n} \end{bmatrix}$ as n tends to ∞ , where

$$\mu_1 = (1 - \theta)\lambda(1 + \lambda), \quad \mu_2 = (1 - \theta)\lambda \quad (7)$$

$$\sigma_{11} = (1 - \theta)\lambda(1 + 6\lambda + 4\lambda^2 + \theta\lambda(1 + \lambda)^2) \quad (8)$$

$$\sigma_{22} = (1 - \theta)\lambda(1 + \theta\lambda) \quad (9)$$

and

$$\sigma_{12} = (1 - \theta)\lambda(1 + 2\lambda + \theta\lambda(1 + \lambda)) \quad (10)$$

Proof. Since each component in the vector is a linear combination of iid random variables, the result follows due to the multivariate central limit theorem.

Theorem: Suppose that $X_n = (X_{n1}, X_{n2}, \dots, X_{nk})$ is $AN\left(\mu, \frac{1}{n}\Sigma\right)$ as n tends to ∞ then for any continuous function $g(x)$,

$$g(X_n) = AN\left(g(\mu), \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^k \sigma_{ij} \frac{\partial g}{\partial x_i} \bigg|_{x=\mu} \frac{\partial g}{\partial x_j} \bigg|_{x=\mu}\right) \text{ as } n \text{ tends to } \infty.$$

where AN means asymptotically normally distributed.

Corollary 1.

$$\hat{\lambda} \sim AN\left(\lambda, \frac{(\lambda + 2)}{n(1 - \theta)}\right) \text{ as } n \text{ tends to } \infty$$

Proof: In order to apply the above quoted theorem, let us define

$$X_{n1} = \sum_i^n y_i^2 \text{ and } X_{n2} = \sum_i^n y_i. \text{ Note that } \hat{\lambda} = \frac{X_{n1}}{X_{n2}} - 1 = g(X_{n1}, X_{n2}).$$

Note that

$$\frac{\partial g}{\partial X_{n1}} \bigg|_{x_{n1}=\mu, x_{n2}=\mu_2} = \frac{1}{n(1 - \theta)\lambda} \quad (11)$$

$$\frac{\partial g}{\partial X_{n2}} \bigg|_{x_{n1}=\mu, x_{n2}=\mu_2} = \frac{-(1 + \lambda)}{n(1 - \theta)\lambda} \quad (12)$$

After simplification, the asymptotic variance of $\hat{\lambda}$ is seen to be $\frac{\lambda + 2}{n(1 - \theta)}$.

Proof. The proof can be mimicked along the same lines as in Corollary 1.

Corollary 2.

$$\text{Let } D = \frac{\overline{Y^2} - \bar{Y}^2}{\bar{Y}}$$

Then, $D \sim AN \left(1 + \theta\lambda, \frac{2 + \theta^2\lambda + \theta(1-\theta)^2\lambda^2}{n(1-\theta)} \right)$ as n tends to ∞ .

Proof. The proof can be mimicked along the same lines as in Corollary 1.

3. Hypothesis Testing Procedures

3.1. Distance Based Testing

Suppose that we want to test $H_0 : \theta = 0$ (Poisson) versus $H_1 : \theta > 0$ (ZIP)

$$\text{Under } H_0, \hat{\lambda} = \bar{Y} \text{ and Under } H_1, \hat{\lambda} = \frac{\bar{Y}^2}{\bar{Y}} - 1 \quad (13)$$

Now, let us define the distance between these estimates as follows.

$$\text{Distance, } D = \bar{Y}^2 - \bar{Y}^2 - \bar{Y} \quad (14)$$

We aim to derive a decision rule based on this distance. Note that when is close to 0 then this supports the null hypothesis. However, this distance measure can be negative and moreover the null distribution for this distance measure depends on the nuisance parameter which is unknown. Therefore, we re-define the distance measure as follows.

$$D = \frac{(\bar{Y}^2 - \bar{Y}^2)}{\bar{Y}} \quad (15)$$

Corollary 3.

For the distance measure D given by (15), the null distribution is asymptotically normal with mean = 1 and variance = $\frac{2}{n}$ as n tends to ∞ .

4. Locally Most Powerful Test (LMP Test)

The likelihood function L is given by

$$L = \left(\theta + (1-\theta)e^{-\lambda} \right)^{n_1} \cdot (1-\theta)^{n-n_1} \cdot \frac{e^{(-n-n_1)\lambda} \lambda^{\sum y_i}}{\prod y_i!} \quad (16)$$

Note that n_1 represents the number of '0's in a sample of size $= n$. Next, evaluate the partial derivative of the natural logarithm of L with respect to θ at $\theta = 0$.

$$\frac{\partial \ln L}{\partial \theta} = \frac{n_1(1 - e^{-\lambda})}{(\theta + (1-\theta)e^{-\lambda})} - \frac{(n - n_1)}{(1-\theta)} \quad (17)$$

Reject the null hypothesis H_0 when $n_1 > C_1$, where n_1 represents the number of '0's in the sample and C_1 is the critical value for the test.

Example. (Leroux and Puterman (1992))

Count data for fetal lamb movements.

Outcome:	0	1	2	3	4	5	6	7
Count:	182	41	12	2	2	0	0	1

Based on this data, $\bar{Y} = 0.358$, $\overline{Y^2} = 0.783$.

For the LMP test, the number of '0's = 182. The critical value at 5% significance level is C_1 179. Hence, we reject the null hypothesis. For the distance based test, the distance measure $D = 1.829$ and the critical value at 5% level is 1.126 for the exact test. So, we reject H_0 and accept H_1 . These results agree with the findings of Thas and Rayner (2005) and Leroux and Puterman (1992).

5. Power Computation

We computed the power for both the locally most powerful (LMP) test and this distance based test for different choices of λ and θ where $0 \leq \theta \leq 1$. The power is nearly the same for both tests. These tests are far less computationally intensive while being very powerful. As λ and θ move away from 0, the power of both these tests is almost 1.

6. Conclusion

In this paper, we have presented two simple approaches that are very powerful and less computationally intensive for testing Poisson versus Zero Inflated Poisson (ZIP). The locally most powerful test (LMP test) seems to be very powerful for testing Poisson versus Zero Inflated Poisson (ZIP). Also, the power of the distance based test seems to be very high except for values of θ in the neighborhood of 1. This could be due to the difficulty in separating an observation '0' as coming from either regular Poisson or ZIP, but the power of separation seems to get better for large values of the Poisson mean λ . Note that when θ is equal to 1 this results in a degenerate distribution. Also note that the null distribution (when $\theta = 0$) for the distance based test does not require any knowledge about the nuisance parameter λ and so the critical value for testing Poisson versus ZIP does not involve λ . Although, the power of the partial sum score test developed by Thas and Rayner (2005) and the score test developed by Van den Broeck (1995) seems to have a similar power pattern as our distance based test, the distance based test and the Locally most powerful test (LMP test) seem to outperform the partial sum score test and the score test on the basis of power. In fact, the LMP is a better test for testing the mixing proportions in the context of mixture distributions.

Acknowledgment and Dedication

I wish to dedicate this paper to the memory of my PhD advisor and my mentor, Professor Zakkula Govindarajulu for his contribution to the development of Statistics as a subject. Professor Govindarajulu joined me in writing this paper at his old age of 77 before his sudden death on December 5, 2010. Also, I wish to thank the Editor and the referees for their comments to improve the presentation of this paper on behalf late Professor Z. Govindarajulu and myself. Thank you.

References

- Cochran, W. (1954). Some methods of strengthening tests, *Biometrics* 10, 417 - 451.
- El-Shaarawi, A. (1985). Some goodness-of-fit methods for the Poisson plus added zero distribution, *Applied and Environmental Microbiology* 49, 1304 - 1306.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics* 34, 1 - 13.
- Leroux, B. and Puterman, M. (1992). Maximum-penalized likelihood estimation for independent and Markov dependent mixture models, *Biometrics*, 48, 545 - 588.
- Moran, P.A.P. (1970). On asymptotically optimal tests for composite hypothesis, *Biometrika*, 57, 47 - 55.
- Rao, C.R. (1948). Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation, *Proceedings of the Cambridge Philosophical Society*, 44, 50 - 57.
- Rao, C.R. and Chakravarti, I.M. (1956). Some small sample tests of significance for a Poisson distribution, *Biometrics*, 12, 264 - 282.
- Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley.
- Singh, S. (1963). A note on inflated Poisson distribution, *Journal of the Indian Statistical Association*, 1, 140 - 144.
- Thas, O. and Rayner, J.C.W. (2005). Smooth Tests for the Zero-Inflated Poisson distribution, *Biometrics*, 61, 808 - 815.
- Van den Broeck, J. (1995). A score test for zero inflation in a Poisson distribution, *Biometrics*, 51, 738 - 743.

OPTIMAL DESIGNING OF A VARIABLES QUICK SWITCHING SAMPLING SYSTEM BY MINIMIZING THE AVERAGE SAMPLE NUMBER

S. Balamurali and M. Usha

Department of Mathematics, Kalasalingam University,
Krishnan Koil 626 190, TN, India

Abstract

This article proposes variables quick switching system (VQSS) where the quality characteristic follows normal distribution or lognormal distribution and has upper or lower specification limit. The variables quick switching system will be effective when testing is costly and destructive. The advantages of the variables QSS over the variables single and double sampling plans and attributes QSS are discussed. Tables are also constructed for the selection of parameters of known and unknown standard deviation VQSS for given acceptable quality level and limiting quality level. The problem is formulated as a nonlinear programming where the objective function to be minimized is the average sample number and the constraints are related to lot acceptance probabilities at acceptable quality level and limiting quality level under the operating characteristic curve .

1. Introduction

A sampling system consists of two or more sampling plans and the rules for switching between the sampling plans to achieve a blending of the advantageous features of each of the sampling plans. In general, any sampling system of sampling inspection involving only normal and tightened inspection will be referred to as a two-plan system. Quick switching system (QSS) developed by Dodge (1967) is one of the two-plan systems for the application of attributes quality characteristics. In any two plan system, the tightened inspection can be used when the quality of a product deteriorated and normal inspection are used when the quality is found to be good. Dodge (1965), Hald and Thyregod (1965) and Stephen and Larson (1967) have investigated the two-plan systems using different switching criteria to achieve the desired discrimination on the operating characteristic (OC) curve. Romboski (1969) has investigated the quick switching system of type QSS-1 by taking single sampling plan as the reference plan.

Whenever the normal distribution for the quality characteristic is justifiable, variables sampling plans are used. The main advantage of the variables sampling plan is that the same operating characteristic (OC) curve can be obtained with a smaller sample size than would be

required by an attributes sampling plan. Thus, a variables acceptance sampling plan would require less sampling. The measurements data required by a variables sampling plan would probably cost more per observation than the collection of attributes data. However, the reduction in sample size obtained may more than offset this increased cost. When destructive testing is employed, variables sampling is particularly useful in reducing the costs of inspection. Another advantage is that measurements data usually provide more information about the manufacturing process or lot than do attributes data. Generally, numerical measurements of quality characteristics are more useful than simple classification of the item as conforming or non-conforming. Another advantage of the variables sampling plan is that when acceptable quality levels (AQLs) are very small, the sample size required by it is very less than the attributes sampling plans. Under these circumstances, the variables sampling plans have significant advantages.

For compliance testing of a measurable characteristic, a variable sampling plan may be preferred. Based on this idea, Govindaraju and Balamurali (1998) extended the concept of chain sampling of Dodge (1955) to the measurable quality characteristics. Later, Balamurali and Jun (2006) developed a sampling plan called variables repetitive group sampling for the application variables inspection. Soundararajan and Palanivel (1997) extended the concept of QSS to the normally distributed quality characteristics. They have also developed tables for selecting parameters of the variables QSS, but they didn't follow any optimization techniques. So, this paper attempts to design a variables QSS by minimizing average sample number (ASN). Obviously, a sampling plan having smaller ASN would be more desirable. For the selection of the parameters of the VQSS, the problem is formulated by a nonlinear programming where the objective function to be minimized is the ASN and the constraints are related to lot acceptance probabilities at acceptable quality level (AQL) and limiting quality level (LQL). Comparisons are also made to show the advantages of the proposed system with other conventional variables plans such as single and double sampling plans.

2. Conditions of Application and Operating Procedure

The following assumptions should be valid for the application of the variables quick switching system.

- (i) Production is in a steady state, so that results of past, present and future lots are broadly indicative of a continuing process.
- (ii) Lots are submitted for inspection serially either in the order of production or in the order of being submitted for inspection.
- (iii) Inspection is by measurements, with quality is defined as the fraction nonconforming, p .

In addition, the usual conditions for the application of single sampling variables plans with known or unknown standard deviation should also be valid. The operating procedure of the VQSS is as follows.

Suppose that the quality characteristic of interest has the upper specification limit U and follows a normal distribution with known standard deviation σ . Then the following two procedures of the variables QSS are proposed.

2.1. Procedure 1: k Method

Step 1. Start with normal inspection. During normal inspection, take a random sample of size n_σ , say $(X_1, X_2 \dots X_{n_\sigma})$ and compute $v = \frac{(U - \bar{X})}{\sigma}$, where $\bar{X} = \frac{1}{n_\sigma} \sum_{i=1}^{n_\sigma} X_i$.

Step 2. Accept the lot if $v \geq k_{N\sigma}$ and reject the lot if $v < k_{N\sigma}$. If a lot is rejected on normal inspection, then switch to tightened inspection as in Step 3.

Step 3. During tightened inspection, take a random sample of size n_σ , say $(X_1, X_2 \dots X_{n_\sigma})$ and compute $v = \frac{(U - \bar{X})}{\sigma}$, where $\bar{X} = \frac{1}{n_\sigma} \sum_{i=1}^{n_\sigma} X_i$.

Step 4. Accept the lot if $v \geq k_{T\sigma}$ and reject the lot if $v < k_{T\sigma}$ ($k_{T\sigma} > k_{N\sigma}$). If a lot is accepted on tightened inspection, then immediately switch to normal inspection as in Step 1.

Thus, the variables QSS system is characterized by three parameters, namely n_σ , $k_{N\sigma}$ and $k_{T\sigma}$. If $k_{N\sigma} = k_{T\sigma}$, then the system will reduce to the variables single sampling plan.

2.2. Procedure 2: M Method

Step 1. Start with normal inspection. During normal inspection, take a random sample of size n_σ , say $(X_1, X_2 \dots X_{n_\sigma})$ and compute $Q_U = \frac{(U - \bar{X})}{\sigma} \sqrt{\frac{n}{n-1}}$, where $\bar{X} = \frac{1}{n_\sigma} \sum_{i=1}^{n_\sigma} X_i$.

Also determine the minimum variance unbiased estimate of the fraction non-conforming p , as

$$\hat{p}_U = \int_{Q_U}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Step 2. Accept the lot if $\hat{p}_U \leq M_{N\sigma}$ and reject the lot if $\hat{p}_U > M_{N\sigma}$. Where $M_{N\sigma}$ is the proportion of the area under the normal curve beyond $k_{N\sigma} \sqrt{\frac{n}{n-1}}$. If the lot is rejected on normal inspection, then switch to tightened inspection as in Step 3.

Step 3. During tightened inspection, take a random sample of size n_σ , say $(X_1, X_2 \dots X_{n_\sigma})$ and compute $Q_U = \frac{(U - \bar{X})}{\sigma} \sqrt{\frac{n}{n-1}}$, where $\bar{X} = \frac{1}{n_\sigma} \sum_{i=1}^{n_\sigma} X_i$. Also

determine the minimum variance unbiased estimate of the fraction non-conforming p , as

$$\hat{p}_U = \int_{Q_U}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Step 4. Accept the lot if $\hat{p}_U \leq M_{T\sigma}$ and reject the lot if $\hat{p}_U > M_{T\sigma}$. Here $M_{T\sigma}$ is the proportion of the area under the normal curve beyond $k_{T\sigma} \sqrt{\frac{n}{n-1}}$. (Note: $M_{T\sigma} < M_{N\sigma}$). If a

lot is accepted on tightened inspection then immediately switch to normal inspection as in Step 1.

Thus, the parameters of the variables QSS under the M method are the sample size n_σ , acceptable criterion $M_{T\sigma}$ and $M_{N\sigma}$.

It is also worthwhile to note that the same VQSS can be applied to the case where the quality characteristic follows a log-normal distribution, since the logarithmic transformation leads to a normal distribution. Let $\hat{\mu}$ be an estimate of mean which is calculated from the logarithmic transformation of the data. Then the VQSS is employed as follows.

Step 1. Start with the normal inspection level. Select a random sample of size n_σ and compute $v = \frac{\log(U) - \hat{\mu}}{\sigma}$. Accept the lot if $v \geq k_{N\sigma}$ and reject the lot if $v < k_{N\sigma}$. If a lot is rejected under normal inspection, then switch to tightened inspection.

Step 2. During the tightened inspection, inspect the lots using the single sampling variables plan with a sample size n_σ and the acceptance criterion $k_{T\sigma}(>k_{N\sigma})$. Accept the lot if $v \geq k_{T\sigma}$ and reject the lot if $v < k_{T\sigma}$, where $v = \frac{\log(U) - \hat{\mu}}{\sigma}$. If a lot is accepted on tightened inspection, then switch to normal inspection as in Step 1.

Thus, the parameters of the VQSS under the log-normal distribution are the sample size n_σ , acceptable criterion $k_{T\sigma}$ and $k_{N\sigma}$. The rest of the paper deals with k method under normal distribution only.

3. Operating Characteristics of a Variables Quick Switching System

The OC function of the variables QS system, which gives the proportion of lots that are expected to be accepted for given product quality, p under known sigma case is given by

$$P_a(p) = \frac{P_T}{1 - P_N + P_T} = \frac{\Pr(v \geq k_{T\sigma})}{1 - \Pr(v \geq k_{N\sigma}) + \Pr(v \geq k_{T\sigma})} \quad (1)$$

where $P_N = \Pr(v \geq k_{N\sigma})$ is the probability of accepting a lot based on a single sample with parameters $(n_\sigma, k_{N\sigma})$ and $P_T = \Pr(v \geq k_{T\sigma})$ is the probability of accepting a lot based on a single sampling plan with parameters $(n_\sigma, k_{T\sigma})$. Under type B situation (i.e. a series of lots of the same quality), forming lots of N items from a process and then drawing random sample of size n_σ from these lots is equivalent to drawing random samples of size n_σ directly from the process. Hence the derivation of the OC function is straightforward. The fraction non-conforming in a lot will be determined as

$$p = 1 - \Phi\left(\frac{U - \mu}{\sigma}\right) = 1 - \Phi(v) = \Phi(-v) \quad (2)$$

where $\Phi(y)$ is given by

$$\Phi(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz, \quad (3)$$

provided that the quality characteristic of interest is normally distributed with mean μ and standard deviation σ , and the unit is classified as non-conforming if it exceeds the upper specification limit U . Then its probability of acceptance is written as

$$P_a(p) = \frac{\Phi(w_T)}{1 - \Phi(w_N) + \Phi(w_T)} \quad (4)$$

where $w_T = (v - k_{T\sigma})\sqrt{n_\sigma}$ and $w_N = (v - k_{N\sigma})\sqrt{n_\sigma}$

4. Designing of a Known Sigma Variables Quick Switching System

The OC function of a known sigma VQSS is given in (4). If two points on the OC curve namely, AQL(=p₁), LQL(=p₂), the producer's risk α and the consumer's risk β are prescribed then the OC function can be expressed as

$$\frac{\Phi(w_{T1})}{1 - \Phi(w_{N1}) + \Phi(w_{T1})} = 1 - \alpha \quad (5)$$

and

$$\frac{\Phi(w_{T2})}{1 - \Phi(w_{N2}) + \Phi(w_{T2})} = \beta \quad (6)$$

Here w_{T1} is the value of w_T at $p=p_1$, w_{N1} is the value of w_N at $p=p_1$, w_{T2} is the value of w_T at $p=p_2$ and w_{N2} is the value of w_N at $p=p_2$. That is,

$$\begin{aligned} w_{T1} &= (v_1 - k_{T\sigma})\sqrt{n_\sigma}, w_{N1} = (v_1 - k_{N\sigma})\sqrt{n_\sigma} \\ w_{T2} &= (v_2 - k_{T\sigma})\sqrt{n_\sigma} \text{ and } w_{N2} = (v_2 - k_{N\sigma})\sqrt{n_\sigma} \end{aligned} \quad (7)$$

where v_1 is the value of v at AQL and v_2 is the value of v at LQL. For given AQL or LQL, the values of $k_{N\sigma}$, $k_{T\sigma}$ and the sample size n_σ are determined by using a search procedure.

The average sample number (ASN), by definition, means the expected number of sampled units required for making decisions about the lot. The concept of ASN is meaningful under Type B sampling situations. It is also known that the ASN of the known sigma VQSS is

$$ASN(p) = \frac{n\xi + n\delta}{\xi + \delta} = n \quad (8)$$

The ASN given above can be used as an objective function to solve for the parameters $(n_\sigma, k_{N\sigma}, k_{T\sigma})$. Since there are several choices to obtain the objective function, it is considered here to minimize ASN at AQL. If the objective is to minimize the ASN at AQL, then the problem will be reduced to the following nonlinear optimization problem.

$$\begin{aligned} &\text{Minimize } ASN(p_1)=n_\sigma \\ &\text{Subject to} \\ &P_a(p_1) \geq 1 - \alpha \\ &P_a(p_2) \leq \beta \\ &n_\sigma > 1, k_{1\sigma} > k_{2\sigma} > 0 \end{aligned} \quad (9)$$

where $P_a(p_1)$ and $P_a(p_2)$ are the lot acceptance probabilities at AQL and LQL respectively and are given in (5) and (6) respectively.

5. Designing of Unknown Sigma Variables QS System

Whenever the standard deviation is unknown, we should use the sample standard deviation S instead of σ . In this case, the operation of the scheme is as follows.

Step 1. Start with the normal inspection level using the single sampling variables plan with a sample size n_s and the acceptance criterion k_{NS} . Accept the lot if $v \geq k_{NS}$ and reject

the lot if $v < k_{NS}$, where $v = \frac{(U - \bar{X})}{S}$, $\bar{X} = \frac{1}{n_s} \sum_{i=1}^{n_s} X_i$ and $S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n_s - 1}}$. If a lot is

rejected under normal inspection, then switch to tightened inspection.

Step 2. During the tightened inspection, inspect the lots using the single sampling variables plan with a sample size n_s and the acceptance criterion $k_{TS}(>k_{NS})$. Accept the lot if

$v \geq k_{TS}$ and reject the lot if $v < k_{TS}$, where $v = \frac{(U - \bar{X})}{S}$, $\bar{X} = \frac{1}{n_s} \sum_{i=1}^{n_s} X_i$ and

$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n_s - 1}}$. If a lot is accepted on tightened inspection, then immediately switch to

normal inspection as in Step 1.

Thus, the unknown sigma VQSS has the parameters namely the sample size n_s , and the acceptable criterion k_{NS} and k_{TS} . If $k_{NS}=k_{TS}$, then the VQSS reduced to the variables single

sampling plan with unknown standard deviation. The relationship between known and unknown sigma plan parameters is true only for single sampling plan. Hamaker (1979) has given an approximation for finding the parameters of the unknown sigma single sampling plan from the parameters of the known sigma single sampling plan. Soundarajan and Palanivel (1997) have followed the same approximation for selecting the parameters of unknown sigma variables QSS. However it is to be pointed out that Hamaker's (1979) results must be extended to variables QS system rather than wrongly assuming that the same approximation is valid for VQSS. So the entire design of unknown sigma schemes provided in Soundarajan and Palanivel (1997) seems faulty. So we will follow a different procedure for the unknown sigma case. The determination of parameters for the unknown sigma case namely (n_S , k_{NS} , k_{TS}) is slightly different from the known sigma case. It is known that $\bar{X} \pm k_{NS}S$ is approximately normally distributed with mean $\mu \pm k_{NS}E(S)$ and variance $\frac{\sigma^2}{n_S} + k_{NS}^2 Var(S)$ (see Duncan (1986)).

That is.,

$$\bar{X} + k_{NS}S \sim N\left(\mu + k_{NS}\sigma, \frac{\sigma^2}{n_S} + k_{NS}^2 \frac{\sigma^2}{2n_S}\right)$$

Therefore, the probability of accepting a lot under normal inspection is given by

$$\begin{aligned} P\{\bar{X} \leq U - k_{NS}S | p\} &= P\{\bar{X} + k_{NS}S \leq U | p\} \\ &= \Phi\left(\frac{U - k_{NS}\sigma - \mu}{(\sigma/\sqrt{n_S})\sqrt{1 + \frac{k_{NS}^2}{2}}}\right) = \Phi\left((v - k_{NS})\sqrt{\frac{n_S}{1 + \frac{k_{NS}^2}{2}}}\right) \end{aligned}$$

If we let, $w_{NS} = (v - k_{NS})\sqrt{\frac{n_S}{1 + \frac{k_{NS}^2}{2}}}$ then the probability of acceptance under tightened

inspection is considered $\Phi(w_{NS})$.

Similarly if we let, $w_{TS} = (v - k_{TS})\sqrt{\frac{n_S}{1 + \frac{k_{TS}^2}{2}}}$ then the probability of acceptance under

tightened inspection is taken as $\Phi(w_{TS})$. Hence the lot acceptance probability for sigma unknown case under two-point situation is given by

$$P_a(p_1) = \frac{\Phi(w_{T1S})}{1 - \Phi(w_{N1S}) + \Phi(w_{T1S})} \quad (10)$$

and

$$P_a(p_2) = \frac{\Phi(w_{T2S})}{1 - \Phi(w_{N2S}) + \Phi(w_{T2S})} \quad (11)$$

We obtain $w_{N1S}, w_{T1S}, w_{N2S}, w_{T2S}$ corresponding to $w_{N1}, w_{T1}, w_{N2}, w_{T2}$ respectively by

$$w_{N1S} = \left((v_1 - k_{NS}) \sqrt{\frac{n_s}{1 + \frac{k_{NS}^2}{2}}} \right), \quad w_{T1S} = \left((v_1 - k_{TS}) \sqrt{\frac{n_s}{1 + \frac{k_{TS}^2}{2}}} \right)$$

$$w_{N2S} = \left((v_2 - k_{NS}) \sqrt{\frac{n_s}{1 + \frac{k_{NS}^2}{2}}} \right) \text{ and } w_{T2S} = \left((v_2 - k_{TS}) \sqrt{\frac{n_s}{1 + \frac{k_{TS}^2}{2}}} \right)$$

In this case, the optimization problem becomes,

$$\begin{aligned} &\text{Minimize } \text{ASN}(p_1) = n_s \\ &\text{Subject to} \\ &P_a(p_1) \geq 1 - \alpha \\ &P_a(p_2) \leq \beta \\ &n_\sigma > 1, k_{NS} > k_{TS} > 0 \end{aligned} \quad (12)$$

where $P_a(p_1)$ and $P_a(p_2)$ are the lot acceptance probabilities at AQL and LQL respectively and are described in (10) and (11).

We may determine the parameters of the known sigma and unknown sigma VQSS by solving the nonlinear equation given in (9) and (12) respectively. There may exist multiple solutions since there are three unknowns with only two equations. Generally a sampling would be desirable if the required number of sampled is small. So, in this paper, we consider the ASN as the objective function to be minimized with the probability of acceptance along with the corresponding producer's and consumer's risks as constraints. To solve the above nonlinear optimization problems given in (9) and (12), the sequential quadratic programming (SQP) proposed by Nocedal and wright (1999) can be used. The SQP is implemented in Matlab software using the routine "fmincon". By solving the nonlinear problem mentioned above, the parameters ($n_\sigma, k_{N\sigma}$ and $k_{T\sigma}$) for known sigma plan and the parameters (n_s, k_{NS} and k_{TS}) for unknown sigma plan are determined and these values are tabulated in Table 2.

6. Examples

6.1. Selection of Known Sigma VQSS Indexed by AQL and LQL

Table 1 is used to determine the parameters of the known σ variables quick switching system for specified values of AQL and LQL when $\alpha = 5\%$ and $\beta = 10\%$. For example, if $p_1 = 2\%$, $p_2 = 8\%$, $\alpha = 5\%$ and $\beta = 10\%$, Table 2 gives the parameters as $n_\sigma = 12$, $k_{N\sigma} = 1.552$ and $k_{T\sigma} = 1.817$.

For the above example, the operation of the VQSS is as follows.

Step 1. Take a random sample of size 12 and compute $v = \frac{(U - \bar{X})}{\sigma}$, where $\bar{X} = \frac{1}{12} \sum_{i=1}^{12} X_i$.

Accept the lot if $v \geq 1.552$ and reject the lot if $v < 1.552$. If a lot is rejected, then switch to tightened inspection as in step 2.

Step 2. Select a random sample of size 12 and compute $v = \frac{(U - \bar{X})}{\sigma}$, where $\bar{X} = \frac{1}{12} \sum_{i=1}^{12} X_i$.

Accept the lot if $v \geq 1.817$ and reject the lot if $v < 1.817$. Switch to normal inspection as in step 1, if a lot accepted on this inspection phase.

6.2. Selection of Unknown Sigma VQSS Indexed by AQL and LQL

Table 1 can also be used for the selection of the parameters of the unknown σ VQSS for given values of AQL and LQL. Suppose that AQL=1%, LQL=5%, $\alpha=5\%$ and $\beta=10\%$. From Table 2, the parameters of the VQSS can be determined as $n_S = 24$, $k_{NS} = 1.729$ and $k_{TS} = 2.214$.

7. Advantages of the Variables Quick Switching System

In this section, we will discuss the advantages of the variables QSS over attributes QSS and variables single sampling plans. For the purpose of comparison, we will consider the plans which have the same AQL and LQL. Suppose that for given values of AQL=0.01, $\alpha=5\%$, LQL=0.02 and $\beta = 10\%$, one can find the parameters of the attributes quick switching system under the application of Poisson model as

(i) $n = 251$, $c_T = 5$ and $c_N = 9$

For the same AQL and LQL, we can determine the parameters of the variables single sampling plan (from Sommers (1981)) and variables QSS (from Table 2) respectively as follows.

(ii) $n_\sigma = 116$ and $k = 2.17$

(iii) $n_\sigma = 28$, $k_{N\sigma} = 1.923$ and $k_{T\sigma} = 2.418$

By comparing the above, it is clear that the variables QS System achieves a reduction of over 89% in sample size than the attributes QSS and about 76% than the variables single sampling plan with same AQL and LQL conditions. In order to show the better efficiency of the variables QSS, three OC curves are considered. Fig.1. shows the OC curves of the variables single sampling plans with parameters (10, 1.754) and (10, 2.179) and the variables QSS with parameters (10; 1.754, 2.179). The variables quick switching system (10; 1.754, 2.179) is selected in such a way that it satisfies the two-points on the OC curve condition ($p_1 = 0.01$, $1-\alpha = 0.95$) and ($p_2 = 0.045$, $\beta = 0.10$).

Table 1. Variables Quick Switching Sampling Systems Indexed by AQL and LQL for $\alpha=5\%$ and $\beta=10\%$ Involving Minimum ASN

p_1	p_2	MinASN(p_1) Known Sigma			MinASN(p_1) Unknown Sigma		
		n_{σ}	$k_{T\sigma}$	$k_{N\sigma}$	n_s	k_{TS}	k_{NS}
0.001	0.002	134	2.998	2.943	234	3.193	2.773
	0.003	45	2.958	2.833	108	3.166	2.666
	0.004	31	2.899	2.784	104	2.999	2.704
	0.005	14	2.994	2.599	71	2.998	2.628
	0.006	11	2.985	2.535	60	2.955	2.600
	0.007	10	2.939	2.519	45	2.985	2.525
	0.008	14	2.769	2.639	38	2.982	2.482
	0.009	8	2.897	2.457	35	2.947	2.467
	0.010	7	2.899	2.409	32	2.926	2.446
	0.012	9	2.710	2.525	28	2.880	2.415
	0.015	6	2.751	2.381	24	2.819	2.379
	0.025	4	2.672	2.222	16	2.715	2.260
0.0025	0.004	96	2.841	2.596	326	2.917	2.542
	0.005	35	2.917	2.432	156	2.945	2.445
	0.006	26	2.878	2.398	113	2.910	2.410
	0.0075	23	2.773	2.413	82	2.861	2.371
	0.010	14	2.762	2.302	56	2.809	2.309
	0.015	9	2.693	2.193	37	2.710	2.235
	0.020	8	2.572	2.182	33	2.563	2.233
	0.025	6	2.569	2.079	28	2.490	2.200
	0.030	8	2.353	2.213	19	2.564	2.074
	0.040	5	2.374	2.039	15	2.485	2.010
0.005	0.0075	100	2.626	2.361	330	2.679	2.329
	0.010	31	2.683	2.183	123	2.705	2.205
	0.012	23	2.641	2.146	89	2.664	2.169
	0.015	19	2.549	2.139	63	2.617	2.122
	0.020	9	2.657	1.897	45	2.533	2.073
	0.030	9	2.377	1.982	28	2.443	1.978
	0.035	7	2.385	1.895	23	2.424	1.929
	0.040	7	2.298	1.913	20	2.394	1.894
	0.060	6	2.109	1.884	15	2.224	1.834

Table 1. Continued

p_1	p_2	MinASN(p_1) Known Sigma			MinASN(p_1) Unknown Sigma		
		n_σ	$k_{T\sigma}$	$k_{N\sigma}$	n_S	k_{TS}	k_{NS}
0.0075	0.010	149	2.499	2.244	575	2.502	2.252
	0.012	58	2.527	2.137	209	2.549	2.144
	0.015	34	2.485	2.080	107	2.549	2.059
	0.020	23	2.391	2.041	64	2.486	1.996
	0.025	14	2.402	1.922	47	2.426	1.951
	0.030	11	2.367	1.867	36	2.395	1.900
	0.035	11	2.263	1.893	31	2.337	1.877
	0.040	11	2.179	1.909	26	2.313	1.838
	0.050	8	2.151	1.816	20	2.266	1.776
	0.060	7	2.086	1.781	17	2.196	1.741
	0.070	6	2.036	1.686	14	2.175	1.685
0.010	0.015	68	2.427	2.047	243	2.434	2.064
	0.020	28	2.418	1.923	94	2.441	1.946
	0.025	21	2.334	1.899	62	2.393	1.893
	0.030	15	2.314	1.829	47	2.344	1.854
	0.035	13	2.255	1.810	38	2.304	1.819
	0.040	14	2.141	1.856	32	2.268	1.788
	0.045	10	2.179	1.754	27	2.247	1.752
	0.050	14	2.005	1.875	24	2.214	1.729
	0.060	7	2.123	1.648	20	2.148	1.693
	0.070	6	2.084	1.599	17	2.100	1.655
	0.080	6	1.987	1.617	14	2.089	1.599
	0.090	6	1.904	1.629	11	2.138	1.513
	0.100	3	1.906	1.557	11	2.027	1.532
0.015	0.020	129	2.236	1.971	402	2.248	1.973
	0.025	43	2.265	1.835	121	2.317	1.827
	0.030	26	2.252	1.757	78	2.279	1.779
	0.035	23	2.163	1.768	59	2.233	1.748
	0.040	22	2.082	1.795	48	2.188	1.723
	0.045	21	2.012	1.787	39	2.167	1.687
	0.050	13	2.083	1.658	34	2.131	1.666
	0.060	11	2.009	1.629	26	2.088	1.613
	0.070	8	2.021	1.526	22	2.025	1.585
	0.080	7	1.952	1.526	18	2.005	1.535
	0.090	7	1.888	1.508	16	1.956	1.511
	0.100	5	1.856	1.376	14	1.927	1.477
0.020	0.030	87	2.070	1.840	170	2.167	1.772
	0.035	33	2.158	1.673	96	2.172	1.702
	0.040	26	2.108	1.653	68	2.153	1.658
	0.045	21	2.076	1.621	54	2.120	1.630
	0.050	17	2.060	1.580	50	2.046	1.636

Table 1. Continued

p_1	p_2	MinASN(p_1) Known Sigma			MinASN(p_1) Unknown Sigma		
		n_σ	$k_{T\sigma}$	$k_{N\sigma}$	n_s	k_{TS}	k_{NS}
0.020	0.060	17	1.926	1.618	33	2.044	1.554
	0.070	17	1.813	1.638	26	2.004	1.510
	0.080	12	1.817	1.552	24	1.910	1.510
	0.090	8	1.877	1.417	19	1.911	1.451
	0.100	7	1.851	1.376	16	1.897	1.407
	0.120	5	1.700	1.345	18	1.664	1.474
0.030	0.040	116	1.934	1.679	269	1.965	1.665
	0.045	57	1.957	1.592	133	1.999	1.584
	0.050	33	1.998	1.493	89	1.999	1.535
	0.060	22	1.947	1.442	54	1.974	1.469
	0.070	17	1.894	1.404	40	1.925	1.430
	0.080	14	1.844	1.374	31	1.893	1.388
	0.090	12	1.800	1.345	26	1.846	1.361
	0.100	10	1.781	1.296	29	1.695	1.415
	0.120	11	1.597	1.362	24	1.598	1.388
	0.150	6	1.647	1.152	13	1.644	1.224
	0.200	5	1.470	1.110	10	1.458	1.173
0.040	0.060	58	1.801	1.476	104	1.892	1.427
	0.070	29	1.826	1.361	63	1.869	1.369
	0.080	21	1.796	1.311	47	1.819	1.339
	0.090	17	1.754	1.279	37	1.778	1.308
	0.100	14	1.725	1.240	30	1.749	1.274
	0.110	14	1.639	1.264	26	1.704	1.254
	0.120	14	1.565	1.280	24	1.644	1.249
	0.140	10	1.546	1.191	17	1.634	1.169
	0.160	8	1.512	1.127	17	1.195	1.485
	0.18	7	1.460	1.090	12	1.521	1.096
	0.200	4	1.317	0.882	10	1.496	1.046
0.050	0.060	280	1.666	1.521	462	1.708	1.488
	0.070	66	1.723	1.373	132	1.761	1.366
	0.080	35	1.735	1.280	73	1.768	1.293
	0.090	25	1.711	1.231	53	1.730	1.260
	0.100	23	1.630	1.245	40	1.709	1.219
	0.120	14	1.619	1.134	28	1.640	1.170
	0.140	11	1.560	1.085	21	1.585	1.120
	0.160	9	1.511	1.036	16	1.560	1.060
	0.200	6	1.432	1.006	15	1.328	1.083
	0.250	5	1.324	0.860	8	1.375	0.895

From this figure, it can be easily observed that, for good quality, i.e. for smaller values of fraction nonconforming, the composite OC curve (OC curve of the variables QSS) coincides with the OC curve of the variables single sampling plan (10, 1.754). As quality deteriorates

the OC curve of the composite OC curve moves toward that for the single sampling plan (10, 2.179) and comes close to it beyond the indifference quality level.

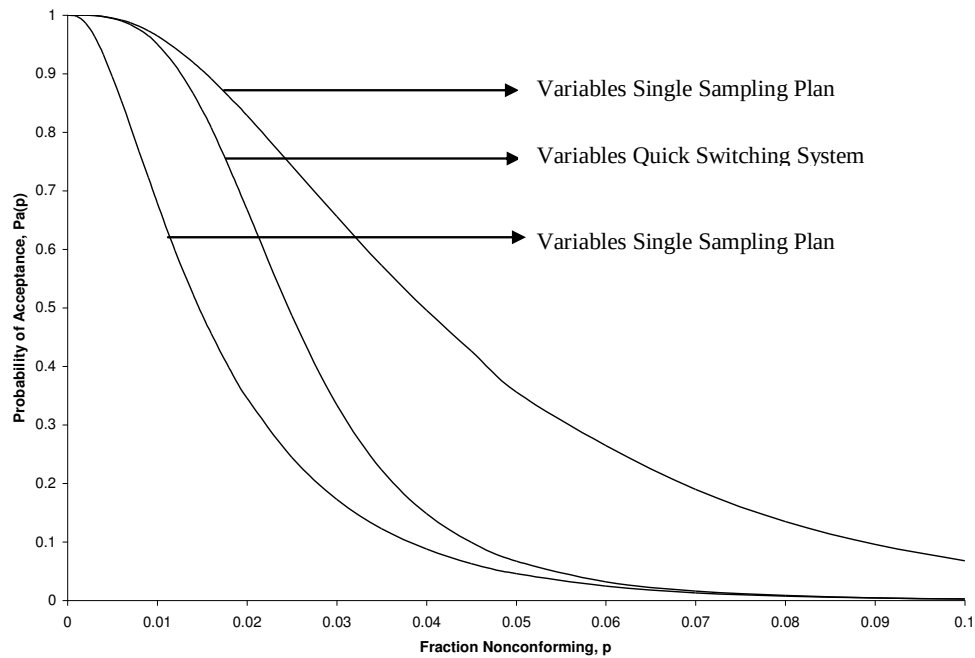


Figure 1. Operating Characteristic (OC) Curves of Single Sampling Normal Plan (10, 1.754), Quick Switching System (10; 1.754, 2.179) and Single Sampling Tightened Plan (10, 2.179).

Table 2. ASN Values of the Known Sigma Variables Single Sampling Plan, Variables Double Sampling Plan and Variables Quick Switching System

P ₁	P ₂	Average Sample Number			
		Variables SSP	Variables DSP	Variables* QSS	Variables ** QSS
0.001	0.002	191	154.9	191	134
0.001	0.003	74	59.4	74	45
0.005	0.010	138	112.0	138	31
0.005	0.012	85	69.5	85	23
0.03	0.04	506	434.6	506	116
0.03	0.06	81	127.7	81	22
0.04	0.07	114	180.6	114	29
0.04	0.08	72	58.4	72	21
0.05	0.07	300	246.7	300	66
0.05	0.08	149	122.3	149	35

* Average sample number given in Soundararajan and Palanivel (1997).

** Average sample number given in Table 1.

Further, it is also to be pointed that the variables QSS is economically superior to the variables double sampling plan in terms of average sample number (ASN). Obviously, a sampling plan having smaller ASN would be more desirable. The variables double or multiple

sampling plans are not practically very useful. Variables sampling Standards avoid presenting such plans due to increased complexity involved in operating them.

Table 2 shows the ASN values of the variables single sampling plan and the variables double sampling plan along with the variables QSS for some arbitrarily selected combinations of AQL and LQL under known sigma case. Table 3 gives the ASN values of the above said plans when sigma is unknown. These ASN values are calculated at the producer's quality level for both known and unknown sigma plans. The sample size of the variables single sampling plan and the ASN of the variables double sampling plan can be found in Sommers (1981).

Table 3. ASN Values of the Unknown Sigma Variables Single Sampling Plan, Variables Double Sampling Plan and Variables Quick Switching System

P ₁	P ₂	Average Sample Number			
		Variables SSP	Variables DSP	Variables* QSS	Variables ** QSS
0.001	0.002	1032	829.1	1032	234
0.001	0.003	381	302.4	412	108
0.005	0.010	547	437.1	941	123
0.005	0.012	327	263.0	823	89
0.03	0.04	1333	1138.7	945	269
0.03	0.06	197	316.5	357	89
0.04	0.07	258	417.6	263	104
0.04	0.08	159	125.8	201	47
0.05	0.07	660	535.4	768	132
0.05	0.08	319	258.0	572	73

* Average sample number given in Soundararajan and Palanivel (1997).

** Average sample number given in Table 1.

These tables apparently show that the variables QSS will have minimum ASN when compared to the variables single and double sampling plans for both known and unknown sigma cases. Similar reduction in ASN can be achieved for any combination of AQL and LQL values. This implies that variables QSS scheme will give desired protection with minimum inspection so that the cost of inspection will greatly be reduced. Thus the variables QSS provides better protection than the variables single sampling plans and variables double sampling plans.

8. Comparison of Parameters of Variables QSS

It is to be pointed out that the ASNs of variables QSS provided in Soundararajan and Palanivel (1997) are equal or greater than the ASN of single sampling plans for some combinations of p_1 and p_2 . For example, for given $p_1=0.01$ and $p_2=0.04$, ASN of single sampling variables plan is 506, but the ASN of variables QSS given in Soundararajan and Palanivel (1997) is 1856. This is a contradiction, since in the attributes case, quick switching system will always have minimum ASN than the attributes single sampling plan (see Romboski (1969) and Soundararajan and Arumainayagam (1992)). This should be valid for

variables sampling also. Hence the entire design of variables QSS provided in Soundararajan and Palanivel (1997) for both known and unknown sigma seems faulty or doubtful. Hence the parameters given in this paper are more reliable and optimum.

Table 4. Parameters of Known Sigma Variables QSS for Some Selected Combination of AQL and LQL Values

P ₁	P ₂	*Parameters of VQSS			**Parameters of VQSS		
		n _σ	k _{Nσ}	k _{Tσ}	n _σ	k _{Nσ}	k _{Tσ}
0.001	0.002	370	2.93	3.10	134	2.943	2.998
0.001	0.003	95	2.87	3.00	45	2.833	2.958
0.005	0.010	399	2.40	2.60	31	2.183	2.683
0.005	0.012	126	2.39	2.50	23	2.146	2.641
0.03	0.04	1856	1.73	2.00	116	1.679	1.934
0.03	0.05	515	1.73	1.90	33	1.493	1.998
0.04	0.06	965	1.62	1.78	58	1.476	1.801
0.04	0.08	121	1.55	1.68	21	1.311	1.796
0.05	0.06	1311	1.55	1.66	280	1.521	1.666
0.05	0.07	853	1.53	1.66	66	1.373	1.723

*Parameters from Soundararajan and Palanivel (1997).

**Parameters from Table 1.

Table 5. Parameters of Unknown Sigma Variables QSS for Some Selected Combination of AQL and LQL Values

P ₁	P ₂	*Parameters of VQSS			**Parameters of VQSS		
		n _s	k _{NS}	k _{TS}	n _s	k _{NS}	k _{TS}
0.001	0.002	2053	2.93	3.10	234	2.773	3.193
0.001	0.003	504	2.87	3.00	108	2.666	3.166
0.005	0.010	1647	2.40	2.60	123	2.205	2.705
0.005	0.012	502	2.39	2.50	89	2.169	2.664
0.03	0.04	5100	1.73	2.00	269	1.665	1.965
0.03	0.05	1365	1.73	1.90	89	1.535	1.999
0.04	0.06	2362	1.62	1.78	104	1.427	1.892
0.04	0.08	279	1.55	1.68	47	1.339	1.819
0.05	0.06	3001	1.55	1.66	462	1.488	1.708
0.05	0.07	1939	1.53	1.66	132	1.366	1.761

*Parameters from Soundararajan and Palanivel (1997).

**Parameters from Table 1.

9. Conclusions

In this paper, the problem of minimizing the average sample number is considered in variables quick switching system. In general, variables sampling plans require a smaller sample size than do attributes plans. It has also been shown that the variables QSS provided in this paper has smaller ASN than the ASN of the existing single and double sampling plans. The

variables QSS provided in this paper also ensure the protection for the consumers in their point of view. This kind of variables QSS scheme will be effective and useful for compliance testing.

Acknowledgments

The authors would like to thank the reviewers and the Editor for their valuable comments which led to improve the manuscript.

References

- Balamurali, S and Jun, C-H.(2006), "Repetitive Group Sampling Procedure for Variables Inspection", *Journal of Applied Statistics*, Vol.33, pp.327-338.
- Dodge, H. F. (1955), "Chain Sampling Inspection Plan", *Industrial Quality Control*, Vol.11, No.4, pp. 10-13.
- Dodge, H. F. (1965), "Evaluation of a Sampling System Having Rules for Switching Between Normal and Tightened Inspection", *Technical Report No.14*, Statistics Center, Rutgers University, Piscataway, NJ.
- Dodge, H. F. (1967), "A Dual System of Acceptance Sampling", *Technical Report No.16*, The Statistics Center, Rutgers- The State University, New Brunswick, New Jersey.
- Duncan, A. J. (1986): *Quality Control and Industrial Statistics*, 5th Ed., Richard D. Irwin, Homewood, Illinois.
- Govindaraju, K and Balamurali, S. (1998), "Chain Sampling Plan for Variables Inspection", *Journal of Applied Statistics*, Vol.25, pp. 103-109.
- Hald, A and Thyregod, P. (1965), "The Composite Operating Characteristic under Normal and Tightened Sampling Inspection by Attributes", *Bulletin of the International Statistical Institute*, Vol. 41, pp. 517-529.
- Hamaker, H. C. (1979), "Acceptance Sampling for Percent Defective by Variables and by Attributes", *Journal of Quality Technology*, Vol.11, pp. 139-148.
- Nocedal, J. and Wright, S.J. (1999): *Numerical Optimization*, Springer.
- Romboski, L. D. (1969): *An Investigation of Quick Switching Acceptance Sampling System*, Doctoral Dissertation, Rutgers- The State University, New Brunswick, New Jersey.
- Sommers, D. J. (1981), "Two-Point Double Variables Sampling Plans", *Journal of Quality Technology*, Vol.13, pp. 25-30.
- Soundararajan, V and Arumainayagam, S. D. (1992), "Some Sampling Plans with Identical Operating Characteristics Curves", *Journal of Applied Statistics*, Vol.19, No.1, pp.141-153.
- Soundararajan, V and Palanivel, M (1997), "Quick Switching Variables Single Sampling System Indexed by AQL and LQL: Acceptance Criterion Tightening", *Journal of Applied Statistical Science*, Vol.6, No.1, pp.45-57.
- Stephens, K. S. and Larson, K. E. (1967), "An Evaluation of the MIL-STD 105D System of Sampling Plans", *Industrial Quality Control*, Vol.23, pp. 310-319.

CHARACTERISTICS OF BIBD AND COMPLEMENTARY BIBD IN TERMS OF TRIPLETS OF TREATMENTS

S. Mandal¹, C. Zhu¹, D.K. Ghosh² and S.C. Bagui³

¹Department of Statistics, University of Manitoba,
Winnipeg, MB, R3T 2N2, Canada

²Department of Statistics, Saurashtra University,
Rajkot - 360005, India

³Department of Mathematics and Statistics, University of West Florida,
Pensacola, FL-32514, USA

Abstract

It is well known that in a balanced incomplete block (BIB) design two treatments occur together in the same block equal number of times. Thus the restriction on triplets of treatments is the second condition for balance in the design. In this article, we establish some conditions on the occurrence of triplets of treatments in a BIB design, its complementary design and the design formed by the union of the BIB design and the complementary design. We establish the conditions in terms of the parameters of the original BIB design.

Keywords: BIB design; Characteristic matrix; Complementary design; Resolvable design; Triplets of treatments.

1. Introduction

A balanced incomplete block (BIB) design, introduced by Yates (1936), is a binary and connected block design with parameters v, b, r, k and λ , where v is the number of treatments, b is the number of blocks, r is the number of times each treatment is repeated, $k(< v)$ is the block size and λ is the number of times a pair of treatments occur together in the same block. BIB design is a simple and important class of incomplete block design from the construction and analysis point of view. The characteristic matrix (C -matrix) of a BIB design is easily expressed as $C = \theta(I_v - 1/v E_{vv})$, where I_v is the identity matrix of order $v \times v$, E_{vv} is a $v \times v$ matrix with all elements 1, and θ is non-zero eigenvalue of C -matrix with multiplicity $(v-1)$ as the design is connected. The motivation of this paper is to obtain some characteristics of a BIB design, its complementary design and the union of the BIB design and its complementary design in terms of occurrence of triplets of treatments.

BIB designs have also very nice properties in terms of connectedness, non-orthogonality and balancedness. It is well known that in a BIB design two treatments occur together in

the same block equal number of times. Thus the restriction on triplets of treatments is the second condition for balance in the design. Designs that are balanced for the pairs as well as for the triplets of treatments are called doubly balanced incomplete block designs (Calvin, 1954). Motivated by this fact, in the present work we establish some criteria by which we can say that, for a given BIB design, number of triplets of treatments occur together an equal number of times. However, for all BIB designs, all the triplets do not occur an equal number of times. Let d and d_1 denote a BIB design and its complementary design respectively. Given the parameters (v, b, r, k, λ) of the BIB design d , it is well known that the parameters of the complementary design d_1 are $v_1 = v, b_1 = b, r_1 = b - r, k_1 = v - k, \lambda_1 = b - 2r + \lambda$. We have shown that the incomplete block design obtained by taking the union of designs d and d_1 all triplets occur an equal number of times. We have also shown that the incomplete block design obtained by taking the union of designs d and d_1 are either another BIBD (with or without repeated blocks) or a variance balanced design with equireplicated and with unequal block sizes.

2. Triplets of Balanced Incomplete Block Design

For any BIB design triplets of treatments occur if the block size is greater than or equal to 3. In this section we establish a condition in terms of the parameters of BIBD, by which we can express that in the existing BIBD all the triplets occur a constant number of times.

Theorem 2.1. For a BIB design d with parameters v, b, r, k, λ , and with $k \geq 3$, all the triplets occur together a constant number of times provided $\lambda(k-2)$ is divisible by $(v-2)$.

Proof. For a BIB design with parameters v, b, r, k, λ , and with $k \geq 3$, number of triplets for any block is ${}^kC_3$. So the total number of triplets for the b blocks is $b {}^kC_3$. These triplets will be distinct provided two blocks are distinct. However, in any two blocks, if at least three treatments are identical, then the triplets obtained from these two blocks are not distinct. Again for any BIB design with v treatments, if all triplets occur in the design, the number of distinct triplets are ${}^vC_3$. Hence, the triplets obtained from b blocks of the BIB design are repeated $b {}^kC_3 / {}^vC_3 = \lambda(k-2)/(v-2)$ times.

Since all the triplets being repeated is always an integer, $\lambda(k-2)$ must be divisible by $(v-2)$. Hence the theorem.

Now we will consider a series of cases.

Corollary 2.1. Any BIB design constructed using the series $v, b = {}^vC_k, k, r = {}^{v-1}C_{k-1}$ and $\lambda = {}^{v-2}C_{k-2}$, all triplets occur a constant number of times.

Proof. For a BIB design, all triplets occur a constant number of times if

$$\lambda(k-2) = P(v-2) \quad (1)$$

holds true, where P is an integer. For the given series, (1) becomes ${}^{v-2}C_{k-2}(k-2) = P(v-2)$. After simplification, we have, $P = {}^{v-3}C_{k-3}$, which is always an integer. Hence, all triplets occur a constant number of times.

Corollary 2.2. If a BIB design is resolvable satisfying $v = 2k$, all triplets occur together a constant number of times.

Proof. For a BIB design with $v = 2k$, we have from (1), $\lambda(k-2) = P(2k-2)$, that is, $P = \lambda(k-2)/2(k-1)$, which is always an integer for any possible values of λ and k of a BIB design with $v = 2k$. Hence, all triplets occur a constant number of times.

We have verified this result by taking all the BIB designs listed in Cochran and Cox (1957), Raghavarao (1971) and Das and Giri (1986). As an example, consider a BIB design with parameters $v = 8, b = 14, r = 7, k = 4, \lambda = 3$. Here P comes out as 1 so all triplets occur only once.

Remark 1.1. If a BIB design is resolvable and does not satisfy $v = 2k$, all triplets do not occur together a constant number of times.

As an example, consider a BIB design with parameters $v = 9, b = 12, r = 4, k = 3, \lambda = 1$. This is a resolvable BIB design and it does not satisfy $v = 2k$. Here, $P = \lambda(k-2)/2(k-1)$ comes out as a fraction (1/7), and hence all triplets do not occur together a constant number of times.

Remark 1.2. If a BIB design is neither resolvable nor constructed using the series discussed in Corollary 2.1, but satisfies $v = 2k$, all triplets do not occur together a constant number of times. As an example consider a BIB design with parameters $v = 6, b = 10, r = 5, k = 3, \lambda = 2$.

Remark 1.3. For a BIB design, occurrence of each triplet a constant number of times implies the occurrence of each triplet another constant number of times in the corresponding complementary design provided that in both BIB design and its complementary design the block sizes are at least three.

3. Triplets of Union of BIB Design and Its Complementary BIB Design

Consider a BIB design, say d , with parameters v, b, r, k, λ . Then the parameters of the complementary design, say d_1 , are $v_1 = v, b_1 = b, r_1 = b - r, k_1 = v - k, \lambda_1 = b - 2r + \lambda$. In the complementary design triplets will occur if $k_1 \geq 3$, that is, $v \geq k + 3$. Let us construct a design by taking the union of d and d_1 . Call this design as d_2 . This combined design is an incomplete block design with equal/unequal block sizes and equal/unequal replication numbers. In the following theorem, we establish a condition by which we can say that all triplets of treatments occur together a constant number of times.

Theorem 3.1. For design d_2 , all triplets occur together a constant number of times if $b(v-1) - 3r(v-k)$ is divisible by $(v-1)$, that is, $b(v-1) - 3r(v-k) / (v-1) = P$, where P is an integer.

Proof. In each block of design d , there will be ${}^kC_3$ triplets. There are b blocks in design d . So all together we have $b {}^kC_3$ triplets in design d . In the complementary design d_1 , we also have b blocks each of size $v - k$. So all together we have $b {}^{v-k}C_3$ triplets in the complementary design d_1 . Thus, in design d_2 , there will be $b({}^kC_3 + {}^{v-k}C_3)$ triplets. Further for design d_2 with v treatments, number of distinct triplets is ${}^vC_3$. So, number of times each triplet being repeated is obtained by

$$P = \frac{b({}^kC_3 + {}^{v-k}C_3)}{{}^vC_3} = \frac{b(v-1) - 3r(v-k)}{v-1}. \quad (2)$$

Here P must be an integer, which is only possible if $b(v-1) - 3r(v-k)$ is divisible by $(v-1)$, in which case all triplets occur together P times. Hence the theorem.

We have verified this result by taking all the BIB designs listed in Cochran and Cox (1957), Raghavarao (1971) and Das and Giri (1986).

Now we establish some nice characteristics of a BIB design along with its complementary BIB design.

Theorem 3.2. Consider a BIB design d with parameters $v, b = {}^vC_k, r = {}^{v-1}C_{k-1}, k, \lambda = {}^{v-2}C_{k-2}$ with $v = 2k$. Then the design d_2 obtained by combining the BIB design d and its complementary design d_1 is another BIB design with repeated blocks. The parameters of this resulting BIB design is given by $v_2 = v, b_2 = 2b, r_2 = b, k_2 = k, \lambda_2 = 2\lambda$.

Proof. Let N, N_1 and N_2 be the incidence matrices of design d, d_1 and d_2 respectively. As the design d_2 is the union of d and d_1 , the incidence matrix N_2 can be written as $N_2 = [N : N_1]$. The concurrence matrix of d_2 can then be written as

$$N_2 N_2' = N N' + N_1 N_1' = 2[(r - \lambda)I_v + \lambda E_{vv}] \quad (3)$$

which is the concurrence matrix of a BIB design with parameters $v_2 = v, b_2 = 2b, r_2 = b, k_2 = k, \lambda_2 = 2\lambda$. Hence the theorem.

Example. Consider a BIB design d with parameters $v = 6, b = 20, r = 10, k = 3, \lambda = 4$. The parameters of the corresponding complementary design d_1 are $v_1 = 6, b_1 = 20, r_1 = 10, k_1 = 3, \lambda_1 = 4$. Using Theorem 3.2, d_2 gives a BIB design with parameters $v_2 = 6, b_2 = 40, r_2 = 20, k_2 = 3, \lambda_2 = 8$. The resulting design is also a resolvable BIB design.

Remark 3.1. BIB designs obtained by using Theorem 3.2 with $v = 2k$ are either resolvable or affine resolvable. For odd k , the design is resolvable (as $v = 2k$ implies $b = 2r$). For even k , the design is affine resolvable (as k^2/v is integer).

Corollary 3.1. Consider a BIB design that does not satisfy $v = 2k$. Then the design d_2 gives an incomplete block design that is a variance balanced design with unequal block sizes and equal replications.

Proof. The C -matrix of design d_2 is obtained by

$$C = \theta(I_v - \frac{1}{v}E_{vv})$$

where θ is the non-zero eigenvalue of the C -matrix with multiplicity $(v - 1)$. So the design d_2 is variance balanced. The parameters of d_2 are easily obtained by $v_2 = v, b_2 = 2b, r_2 = b, k_2 = (k, v - k), \lambda_2 = \lambda + \lambda_1$.

As a concluding remark, in this article, our intention was not to construct a balanced incomplete block design (with or without repeated blocks) and a variance balanced design with equireplicated and unequal block sizes. Our main goal was to explain and establish conditions on the occurrences of the triplets of treatments in a BIB design, its complementary BIB design and the design formed by taking the union of a BIB design and its complementary design.

Acknowledgments

The research of S. Mandal is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

- [1] Calvin, L.D. (1954). Doubly Balanced Incomplete Block Designs for Experiments in which the Treatment Effects are Correlated. *Biometrics*, 10, 61-88.
- [2] Cochran, W.G. and Cox, G.M. (1957). *Experimental Designs*. 2nd edition. John Wiley & Sons, New York.
- [3] Das, M.N. and Giri, N.C. (1986). *Design and analysis of experiments*. 2nd edition. John Wiley & Sons, New York.
- [4] Raghavarao, D. (1971). *Constructions and combinatorial problems in design of experiments*. John Wiley & Sons, New York.
- [5] Yates, F. (1936). Incomplete randomized blocks. *Annals of Eugenics*, 7, 121-140.

BAYESIAN INFERENCE FOR AN IMPATIENT M|M|1 QUEUE WITH BALKING

Paul R. Savariappan¹, P. Chandrasekhar² and Ambily Jose²

¹Department of Mathematics, Luther College, Decorah, Iowa-52101, USA

²Department of Statistics, Loyola College, Chennai-600 034, India

Abstract

Assuming that the stationary distribution in an M|M|1 balking situation is Negative Binomial, maximum likelihood estimator (MLE) and Bayes estimator of the parameter p based on the number of observations present at several sampled time points are obtained. Further, the minimum posterior risk associated with Bayes estimator and minimum Bayes risk of the estimator are obtained. In addition, the maximum likelihood and consistent asymptotically normal (CAN) estimators and $100(1-\alpha)\%$ asymptotic confidence interval for the expected number of customers in the system are obtained.

Keywords: Bayes estimator – maximum likelihood estimator – minimum posterior risk-M|M|1 queue with balking – multivariate central limit theorem – Slutsky theorem.

AMS 2000 subject classification: Primary 60 K 25, Secondary 90 B 2.

1. Introduction

Most of the studies on queuing models are confined to only obtaining expressions for transient or stationary (steady state) solutions and do not consider the associated statistical inference problems. Parametric estimation, interval estimation and Bayesian estimation are some of the essential tools to understand any random phenomena using stochastic models. Analysis of queuing systems in all these directions has not received much attention in the past. Whenever the systems are fully observable in terms of their basic random components such as interarrival times and service times, standard parametric techniques of statistical theory are quite appropriate.

An important aspect of queuing theory is to estimate queuing parameters for which both classical and Bayesian approaches are useful. It is often the case that some information is available on the parameters of the interarrival time or service time distributions from prior experiments or prior analysis of the interarrival time or service time data. The Bayesian approach provides the methodology for formal incorporation of prior information with the

current data. Table 1 indicates the present state of work of queuing systems, wherein both classical and Bayesian approaches are used for the estimation of queuing parameters.

Table 1. Confidence Limits for L_S , W_S and W_Q of queueing systems

S.No	System description	Authors	Estimators
1	M M 1	Clarke (1957)	MLEs of λ and μ
2	M M 1	Muddapur (1972)	Bayes estimators of λ and μ
3	M M 1 ∞ and M M 1 N	Yadavalli et al (2004)	MLE of W_Q
4	M M c ∞ and M M c N	Yadavalli et al (2006)	MLE of W_Q
5	Tandem queue with blocking and dependent structure for service times	Chandrasekhar et al (2006)	Moment estimators of L_S and W_S
6	Tandem queue with zero queue capacity and with blocking	Chandrasekhar et al (2008)	MLE of W_S

Thiruvaiyaru and Basawa (1992) adopted an empirical Bayesian approach to estimate the parameters of various queuing systems, where they used arrival and service times as the observed data. In all these models considered so far, it may be noted that MLE and Bayes estimator of queuing parameters are obtained by observing mainly number of arrivals and number of service completions in a discrete set up and interarrival time, service time completion, waiting time or sojourn time in the continuous setup. But in a real life situation the number of customers at different time points is easy to observe. Recently, Mukherjee and Chowdhry (2005) have obtained MLE and Bayes estimator of traffic intensity in an M|M|1 queueing model based on the number of customers present at several sampled time points. An attempt is made in this paper to obtain MLE of the parameter p based on the number of observations present at several sampled time points assuming that the stationary distribution in an M|M|1 balking situation is Negative Binomial. Further, Bayes estimator of p under the same set up, minimum Bayes risk and minimum posterior risk associated with Bayes estimator of p are obtained. In the next section, the MLE of p and its exact distribution are obtained.

2. MLE OF p and Its Exact Distribution

Assume that for an M|M|1 balking situation, the stationary distribution is given by

$$p_r = \binom{r+N-1}{N-1} p^N q^r, \quad r = 0, 1, 2, 3 \dots \quad (2.1)$$

where $p+q = 1$.

Let $(x_1, x_2, \dots, x_n)$ denote the number of customers present at different sampled time points $(t_1, t_2, \dots, t_n)$. The likelihood function of the number of customers $(x_1, x_2, \dots, x_n)$ present at $(t_1, t_2, \dots, t_n)$ is given by

$$L(p|x_1, x_2, \dots, x_n) = \left\{ \prod_{i=1}^n \binom{x_i + N - 1}{N - 1} \right\} p^{nN} (1 - p)^{\sum_{i=1}^n x_i} \quad (2.2)$$

Clearly,

$$\left(\frac{\partial \log L}{\partial p} \right) = 0 \Rightarrow \hat{p} = \frac{N}{(N + \bar{x})} = \frac{nN}{(nN + Y)}, \text{ where } Y = \sum_{i=1}^n X_i \sim \text{NB}(nN, p)$$

with the probability mass function given by

$$\begin{aligned} P[Y = y] &= \binom{y + nN - 1}{nN - 1} p^{nN} q^y, \quad y = 0, 1, 2, \dots \\ &= f(y; p) \text{ (say)} \end{aligned}$$

It readily follows that, $E(Y) = \frac{nNq}{p}$ and $\text{Var}(Y) = \frac{nNq}{p^2}$.

Since $\hat{p}$ is a one-to-one function of y , it is clear that $y = 0, 1, 2, 3, \dots$ with probability mass function given by

$$\begin{aligned} P[\hat{p} = u] &= P\left[\frac{nN}{nN + Y} = u\right] = P\left[Y = \frac{nN(1-u)}{u}\right] \\ &= \binom{\frac{nN(1-u)}{u} + nN - 1}{nN - 1} p^{nN} q^{\frac{nN(1-u)}{u}} \end{aligned} \quad (2.3)$$

Now, for large n , $E(\hat{p}) \cong \frac{nN}{nN + E[Y]} = p$

and

$$\begin{aligned} \text{Var}(\hat{p}) &\cong \left[\left(\frac{d\hat{p}}{dy} \right)^2 \text{Var}(Y) \right]_{E(Y) = \frac{nNq}{p}} \\ &= \frac{q}{nNp^2} \frac{1}{\left(1 + \frac{E(Y)}{nN} \right)^4} = \frac{qp^2}{nN} \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

i.e., $\hat{p}$ is a consistent estimator of p .

3. Bayes Estimator of p

In this section, Bayes estimator of p and its Bayes risk are obtained by using the same data of section 2 namely the number of customers present at several sampled time points. Beta distribution of first kind is taken as natural conjugate prior density for the parameter p. Assume that, p has a prior distribution Beta of first kind with the parameters α_1 and α_2 .

$$i.e., \tau(p|\alpha_1, \alpha_2) = \frac{1}{\beta(\alpha_1, \alpha_2)} p^{\alpha_1-1} (1-p)^{\alpha_2-1}, \quad 0 < p < 1, \alpha_1, \alpha_2 > 0 \quad (3.1)$$

The marginal pdf of Y, which is called the predictive pdf is given by

$$\begin{aligned} f^*(y) &= \int_0^1 f(y; p) \tau(p|\alpha_1, \alpha_2) dp \\ &= \frac{\beta(nN + \alpha_1, y + \alpha_2)}{\beta(\alpha_1, \alpha_2)} \binom{y + nN - 1}{nN - 1} \end{aligned}$$

Hence, the posterior distribution of p is given by

$$\begin{aligned} q(p|x_1, x_2, \dots, x_n) &= \frac{f(y; p) \tau(p|\alpha_1, \alpha_2)}{\int_0^1 f(y; p) \tau(p|\alpha_1, \alpha_2) dp} \\ &= \frac{1}{\beta(nN + \alpha_1, y + \alpha_2)} p^{(nN + \alpha_1)-1} (1-p)^{(y + \alpha_2)-1}, \\ &0 < p < 1 \end{aligned} \quad (3.2)$$

It may be noted that the posterior distribution of p is also that of Beta distribution of I kind with the parameters $(nN + \alpha_1)$ and $(y + \alpha_2)$.

Thus, the Bayes estimator of p under squared error loss is given by

$$\begin{aligned} E(p|x_1, x_2, \dots, x_n) &= \int_0^1 p \cdot q(p|x_1, x_2, \dots, x_n) dp \\ &= \frac{(nN + \alpha_1)}{(nN + \alpha_1 + \alpha_2 + y)} \end{aligned} \quad (3.3)$$

Further, the minimum posterior risk associated with this Bayes estimator is given by

$$\begin{aligned} V_p(\hat{p}^B|x_1, x_2, \dots, x_n) &= E[\hat{p} - p]^2 \\ &= \int_0^1 (\hat{p} - p)^2 \cdot q(p|x_1, x_2, \dots, x_n) dp \\ &= \frac{n^2 N^2}{(nN + y)^2} + \frac{(nN + \alpha_1)}{(nN + \alpha_1 + \alpha_2 + y)} \left\{ \frac{y(\alpha_1 + 1) - nN(nN + \alpha_1 + 2\alpha_2 + y + 1)}{(nN + \alpha_1 + \alpha_2 + y + 1)(nN + y)} \right\} \end{aligned} \quad (3.4)$$

Now, the marginal distribution $h(x_1, x_2, \dots, x_n)$ of $f(x_1, x_2, \dots, x_n)$ is given by

$$\begin{aligned} h(x_1, x_2, \dots, x_n) &= \int_0^1 L(p|x_1, x_2, \dots, x_n) \cdot \tau(p|\alpha_1, \alpha_2) dp \\ &= \frac{\beta(nN + \alpha_1, y + \alpha_2)}{\beta(\alpha_1, \alpha_2)} \prod_{i=1}^n \binom{x_i + N - 1}{N - 1} \end{aligned}$$

Now, the minimum Bayes risk $r_{\tau, \hat{p}^B}$ of $\hat{p}^B$ is given by

$$\begin{aligned} r_{\tau, \hat{p}^B} &= E[V_p(\hat{p}^B|x_1, x_2, \dots, x_n)] \\ &= \sum_{y=0}^{\infty} V_p(\hat{p}^B|x_1, x_2, \dots, x_n) \cdot h(x_1, x_2, \dots, x_n) \\ &= \frac{\prod_{i=1}^n \binom{x_i + N - 1}{N - 1}}{\beta(\alpha_1, \alpha_2)} \sum_{y=0}^{\infty} \left\{ \frac{\alpha_1(\alpha_1 + 1)y^2 + n^2 N^2 \alpha_2(\alpha_2 + 1) + nNy(nN + y - 2\alpha_1\alpha_2)}{(nN + y)^2(nN + \alpha_1 + \alpha_2 + y)(nN + \alpha_1 + \alpha_2 + y + 1)} \right\} \times \\ &\quad \beta(nN + \alpha_1, y + \alpha_2) \end{aligned}$$

4. Expected Number of Customers in the System

Suppose in M|M|1 balking model, the discouragement function b_n is given by

$$b_n = e^{-\frac{\alpha n}{\mu}}, \quad n = 1, 2, \dots; \quad b_0 \equiv 1.$$

Hence, the stationary distribution $\{p_n, n \geq 0\}$ is given by

$$p_n = p_0 \left[\frac{\lambda}{\mu} \right]^n \prod_{i=1}^n b_{i-1}, \quad n = 0, 1, 2, \dots$$

$$\text{Now, } \sum_{n=0}^{\infty} p_n = 1 \Rightarrow p_0 = \frac{1}{\left[1 + \sum_{j=1}^{\infty} \left(\frac{\lambda}{\mu} \right)^j \cdot e^{-\frac{j(j-1)\alpha}{2\mu}} \right]}$$

$$\text{In other words, } p_n = \frac{\left[\frac{\lambda}{\mu} \right]^n \cdot e^{-\frac{n(n-1)\alpha}{2\mu}}}{\left[1 + \sum_{j=1}^{\infty} \left[\frac{\lambda}{\mu} \right]^j \cdot e^{-\frac{j(j-1)\alpha}{2\mu}} \right]}, \quad n = 0, 1, 2, \dots$$

Now, the expected number of customers in the system is given by

$$L_s = \sum_{n=0}^{\infty} np_n = \frac{1}{\left[1 + \sum_{j=1}^{\infty} \left[\frac{\lambda}{\mu} \right]^j \cdot e^{-\frac{j(j-1)\alpha}{2\mu}} \right]} \sum_{n=0}^{\infty} n \cdot \left[\frac{\lambda}{\mu} \right]^n \cdot e^{-\frac{n(n-1)\alpha}{2\mu}} \quad (4.1)$$

5. MLE and CAN Estimator for the Expected Number of Customers in the System

5.1. ML Estimator

Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_n$ be two independent random samples each of size n drawn from exponential interarrival time and exponential service time populations with the parameters λ and μ respectively. It is clear that $E[\bar{X}] = \frac{1}{\lambda}$ and $E[\bar{Y}] = \frac{1}{\mu}$, where $\bar{X}$ and $\bar{Y}$ are the sample means of interarrival times and service times respectively. It can be shown that $\bar{X}$ and $\bar{Y}$ are the MLE's of $\frac{1}{\lambda}$ and $\frac{1}{\mu}$ respectively.

Let $\theta_1 = \frac{1}{\lambda}$ and $\theta_2 = \frac{1}{\mu}$. Clearly, the expected number of customers in the system given in (4.1) reduces to

$$L_s = \frac{1}{\left[1 + \sum_{j=1}^{\infty} \left[\frac{\theta_2}{\theta_1}\right]^j \cdot e^{-\frac{j(j-1)\alpha\theta_2}{2}}\right]} \cdot \sum_{n=0}^{\infty} n \cdot \left[\frac{\theta_2}{\theta_1}\right]^n \cdot e^{-\frac{n(n-1)\alpha\theta_2}{2}} \quad (5.1)$$

and hence MLE of L_s is given by

$$\hat{L}_s = \frac{1}{\left[1 + \sum_{j=1}^{\infty} \left[\frac{\bar{Y}}{\bar{X}}\right]^j \cdot e^{-\frac{j(j-1)\alpha\bar{Y}}{2}}\right]} \cdot \sum_{n=0}^{\infty} n \cdot \left[\frac{\bar{Y}}{\bar{X}}\right]^n \cdot e^{-\frac{n(n-1)\alpha\bar{Y}}{2}}$$

It may be noted that $\hat{L}_s$ is a real valued function in $\bar{X}$ and $\bar{Y}$, which is also differentiable. Consider the following application of multivariate central limit theorem. see Rao(1974).

5.2. Application of Multivariate Central Limit Theorem

Suppose $T_1', T_2', T_3', \dots$ are independent and identically distributed k - dimensional random variables such that

$$T_n' = (T_{1n}, T_{2n}, T_{3n}, \dots, T_{kn}), n = 1, 2, 3, \dots$$

having the first and second order moments $E(T_n) = \mu$ and $\text{var}(T_n) = \Sigma$. Define the sequence of random variables

$$\bar{T}_n' = (\bar{T}_{1n}, \bar{T}_{2n}, \dots, \bar{T}_{kn}), n = 1, 2, 3, \dots$$

where $\bar{T}_{in} = \frac{\sum_{j=1}^n T_{ij}}{n}$, $i = 1, 2, 3, \dots, k$; $j = 1, 2, 3, \dots, n$.

Then, $\sqrt{n}(\bar{T}_n - \mu) \xrightarrow{d} N_k(0, \Sigma)$ as $n \rightarrow \infty$.

5.3. CAN Estimator

By applying the multivariate central limit theorem given in section 5.2, it is readily seen that $\sqrt{n}[(\bar{X}, \bar{Y}) - (\theta_1, \theta_2)] \xrightarrow{d} N_2(0, \Sigma)$ as $n \rightarrow \infty$, where the dispersion matrix $\Sigma = ((\sigma_{ij}))$ is given by $\Sigma = \text{diag}(\theta_1^2, \theta_2^2)$. Again from Rao (1974), we have

$$\sqrt{n}(\hat{L}_s - L_s) \xrightarrow{d} N(0, \sigma^2(\theta)) \text{ as } n \rightarrow \infty.$$

where $\theta = (\theta_1, \theta_2)$ and

$$\begin{aligned} \sigma^2(\theta) &= \sum_{i=1}^2 \left(\frac{\partial L_s}{\partial \theta_i} \right)^2 \sigma_{ii} \\ &= \sum_{i=1}^2 \left(\frac{\partial L_s}{\partial \theta_i} \right)^2 \theta_i^2 \end{aligned} \quad (5.2)$$

By substituting the partial derivatives $\left(\frac{\partial L_s}{\partial \theta_i} \right)$, $i = 1, 2$ in (5.2), we obtain an expression for $\sigma^2(\theta)$. Thus, $\hat{L}_s$ is a CAN estimator of L_s . There are several methods for generating CAN estimators and the Method of moments and the Method of Maximum likelihood are commonly used to generate such estimators. see Sinha(1986).

5.4. Confidence Limits for the Expected Number of Customers in the System

Let $\sigma^2(\hat{\theta})$ be the estimator of $\sigma^2(\theta)$ obtained by replacing θ by a consistent estimator $\hat{\theta}$ namely $\hat{\theta} = (\bar{X}, \bar{Y})$. Let $\widehat{\sigma^2} = \sigma^2(\hat{\theta})$. Since $\sigma^2(\theta)$ is a continuous function of θ , $\widehat{\sigma^2}$ is a consistent estimator of $\sigma^2(\theta)$.

$$\text{i.e., } \widehat{\sigma^2} \xrightarrow{P} \sigma^2(\theta) \text{ as } n \rightarrow \infty.$$

By Slutsky theorem

$$\left(X_n \xrightarrow{d} X, Y_n \xrightarrow{P} b \Rightarrow \frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{b}, b \neq 0 \right),$$

we have

$$\begin{aligned} \sqrt{n} \left(\frac{\hat{L}_s - L_s}{\widehat{\sigma}} \right) &\xrightarrow{d} N(0, 1) \\ \text{i.e., } \Pr \left[-k_{\frac{\alpha}{2}} < \frac{\sqrt{n}(\hat{L}_s - L_s)}{\widehat{\sigma}} < k_{\frac{\alpha}{2}} \right] &= (1 - \alpha), \end{aligned}$$

where $k_{\frac{\alpha}{2}}$ is obtained from normal tables. Hence, a $100(1 - \alpha)\%$ asymptotic confidence interval for L_s is given by

$$\hat{L}_s \pm k_{\frac{\alpha}{2}} \frac{\widehat{\sigma}}{\sqrt{n}}. \quad (5.3)$$

where $\hat{\theta}$ is obtained from (5.2).

Acknowledgment

The research of P.Chandrasekhar has been sponsored by the University Grants Commission, New Delhi.

References

- [1] Chandrasekhar, P., Chandrasekar, B. and Yadavalli, V.S.S. (2006)-Statistical Inference for a tandem queue with dependent structure for service times, *Proceedings of the Sixth IASTED International Conference on Modelling, Simulation and Optimization*, September 11-13, 2006, Gaborone, Botswana, pp 233-238.
- [2] Chandrasekhar, P., Natarajan, R. and Yadavalli, V.S.S. (2008)-Statistical analysis for a tandem queue with blocking, *Proceedings of the second National Conference on Management Science and Practice* (N. Ravichandran, Ed), Allied Publishers Pvt. Ltd, March 9-11, 2007, Chennai, India, pp 65-72.
- [3] Clarke, A. B. (1957)- Maximum Likelihood estimates in a simple queue, *Ann. Math Stat.*, 28, pp 1036-1040.
- [4] Muddapur, M. V. (1972)-Bayesian estimates of parameters in some queuing models, *Ann. Inst. Stat. Math.*, 24, pp 327-331.
- [5] Mukherjee, S. P. and Shovan Chowdhury (2005)-Maximum Likelihood and Bayes estimators in M|M|1 queue, *Stochastic Modelling and Applications*, Vol 8, No 2, pp 47-55.
- [6] Radhakrishna Rao, C. (1974)-*Linear Statistical Inference and its applications*, Wiley Eastern Pvt. Ltd., New Delhi.
- [7] Sinha, S. K. (1986)-*Reliability and Life Testing*, Wiley Eastern Pvt. Ltd., New Delhi.
- [8] Thiruvaiyaru, D. and Basawa, I. V. (1992)-Empirical Bayes estimation for queuing systems and networks, *Queuing Sys.*, 11, pp 179-202.
- [9] Yadavalli, V. S. S., Adendorff, K., Erasmus, G., Chandrasekhar, P. and Deepa, S. P. (2004)-Confidence limits for expected waiting time of two queuing models, *OriON* (South Africa), Vol. 20, No 1, pp 1-6.
- [10] Yadavalli, V.S.S., Natarajan, R., and Chandrasekhar, P. (2006)-Confidence limits for expected waiting time of M|M|c| ∞ and M|M|c|N queuing models, *Pak. J. Statist*, Vol. 22(2), pp 171-178.

INTERVAL ESTIMATION FOR WEIBULL DISTRIBUTION BASED ON A SAMPLE OF SIZE ONE

S. V. Sabnis¹ and A. D. Dharmadhikari²

¹Dept. of Mathematics, IIT Bombay, Powai, Mumbai, India

²Quality & Reliability Division, Tata Motors, Pune, India

Abstract

In this paper, we demonstrate the method of construction of $100(1 - \delta)\%$ confidence interval for parameter λ (scale parameter) of a Weibull distribution based on a sample of size one. We take a cue from this method to construct a similar $100(1 - \delta)\%$ confidence intervals for these parameters based on $X_{1:n} = \min\{X_1, X_2, \dots, X_n\}$ where $X_1, X_2, \dots, X_n$ are $n(\geq 2)$ independent random variables from a Weibull distribution having parameters λ and β . For $\beta = 1$, i.e. for exponential distribution, although the proposed confidence interval for the mean λ overlaps with the confidence interval based on chi-square distribution, the former one is significantly narrower than the latter one. For $n \geq 2$ additional numerical results that compare proposed $100(1 - \delta)\%$ confidence intervals for λ with the ones that are given in Bain and Engelhardt (1991) are provided. It is interesting to note that the former ones turn out to be significantly narrower than the latter ones for samples of sizes 1 through 5 for $\delta = 0.02, 0.05$ and 0.10 .

1. Introduction

In many industrial situations, development of a product involves use of extremely expensive prototypes. Even the accelerated life testing requires large amount of time to yield complete life. In such situations the user is required to make inferences about the product as and when the first observation becomes available. Such inferential work based on sample size one is available for random variables having normal distribution. Machol and Rosenblatt (1966) provided $100(1 - \delta)\%$ confidence interval (CI) of finite length for the mean (μ) for the given standard deviation, σ , when $X \sim N(\mu, \sigma^2)$. For this case, Balchman and Machol (1987) obtained σ -free bounds for μ . Edelman (1990) showed that for every value of δ , $100(1 - \delta)\%$ confidence interval can be constructed for the mean of any unimodal distribution. Recently, Wall, Boen and Tweedie (2001) produced σ -free bounds for μ based on samples of size $n=2$ that are better than traditional bounds based on t-distribution.

In the context of reliability theory parametric distributions such as exponential, Gamma and Weibull distributions play very vital role. In this paper some inferential work based on sample of size one as well as sample of size $n \geq 2$ for Weibull distribution is presented. To be specific we consider the problem of providing $100(1 - \delta)\%$ confidence interval for the scale parameter λ of a Weibull distribution on the basis of sample sizes $n=1$ and $n \geq 2$ when the shape parameter β is unknown. The only information required about β is whether $\beta < 1$

or $\beta > 1$, i.e., essentially the information in terms of the aging properties of the component. It should be emphasized that when $n=1$ and $\beta = 1$ (exponential distribution), although the proposed confidence interval overlaps with the traditional confidence interval based on chi-square distribution, it is significantly narrower than the latter one. It is interesting to note that for $n \geq 2$, the pivot for the confidence interval is taken to be the first order statistics $X_{1:n}$ which represents the lifetime of a series system of n independent components. For $n \geq 2$ additional numerical results that compare proposed $100(1 - \delta)\%$ confidence intervals for λ with the ones that are given in Bain and Engelhardt (1991) are provided. It should be highlighted that the former ones turn out to be significantly narrower than the latter ones for samples of sizes 1 through 5 for $\delta = 0.02, 0.05$ and 0.10 .

The organization of this paper is as follows. Section 1 contains introduction. Various subsections of Section 2 contain discussion on $100(1 - \delta)\%$ confidence intervals for various ranges of β . Specifically, Subsections 2.1, 2.2 and 2.3 deal with construction of $100(1 - \delta)\%$ confidence intervals for the scale parameter of Weibull distribution based on a single observation and when $\beta = 1$, $\beta > 1$ and $\beta < 1$ respectively. Subsection 2.4 discusses numerical results corresponding to theoretical results presented in Subsections 2.2 and 2.3. Section 3 gives similar results when the sample size n is at least 2.

2. Confidence Interval for λ Based on Sample of Size One

It is known that the cumulative distribution function (CDF) of a Weibull random variable X with parameters λ and β is given by

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^\beta\right), \quad x > 0; \lambda > 0, \beta > 0. \quad (2.1)$$

To decide the nature of the confidence interval, we use the fact that $F(X)$ is $U[0, 1]$ random variable. Now for given $\delta \in (0, 1)$ there exist a and b ($0 < a < b < 1$) such that

$$P[a < F(X) < b] = 1 - \delta. \quad (2.2)$$

Using $F(x)$ as given in Equation 2.1, one can simplify Equation 2.2 as

$$P[k_1 X \leq \lambda \leq k_2 X] = 1 - \delta \quad (2.3)$$

where k_1 and k_2 depend upon shape parameter β . Further k_2 may be infinite and k_1 may be zero for a given value of δ . The question is - can one get β -free bounds for λ ? We answer this question in affirmative. We in fact show that if limited information about β , say $\beta < 1$ or $\beta > 1$, is available then for δ in disjoint subintervals of $(0, 1)$ it is possible to provide β -free bounds for λ . It is worthwhile to note that $\beta > 1$ ($\beta < 1$) is a well known case of IFR (DFR) distribution. If the problem is related to the life of a mechanical component whose lifetime has adverse effect on aging then assuming $\beta > 1$ is natural, where as, if the problem is related to repair (replacement) times then assuming $\beta < 1$ may be more adequate. Below, we attempt to provide CI for both these situations.

Note that when $k_1 > 0$ and $k_2 > 0$, in view of Equation 2.1, Equation 2.3 simplifies to

Table 2.1.

	CI based on chi-square distribution	New proposed CI
$\delta = .01$	(0.1887X, 199.4992X)	(0.1086X, 100.5092X)
$\delta = .05$	(0.2711X, 39.5257X)	(0.1669X, 20.5488X)
$\delta = 0.1$	(0.3338X, 19.4956X)	(0.2171X, 10.6032X)

$$e^{-(\frac{1}{k_2})^\beta} - e^{-(\frac{1}{k_1})^\beta} = 1 - \delta. \quad (2.4)$$

This, in turn, yields

$$k_2(\beta) = [-\ln\{(1 - \delta) + e^{-(\frac{1}{k_1})^\beta}\}]^{-1/\beta}. \quad (2.5)$$

Note that $k_2(\beta)$ in the above equation is defined if and only if

$$k_1 < q_1^{1/\beta} \quad (2.6)$$

where $q_1 = (-\ln \delta)^{-1}$. Hence for a known β , one may select $k_1 = q_1^{1/\beta}$ and the corresponding $100(1 - \delta)\%$ confidence interval will become (k_1X, k_2X) . Observe that $q_1 < 1 \Leftrightarrow \delta < e^{-1}$ and $q_1 \geq 1 \Leftrightarrow \delta \geq e^{-1}$. However for this choice of k_1 , $k_2(\beta) = \infty$ for all values of β . This suggests that we need to consider values of $k_1 \in (0, q_1^{1/\beta})$ for which k_2 is finite. As mentioned earlier, our aim is to construct β -free confidence interval when the value of β is unknown. We, first begin with the case $\beta = 1$ which corresponds to exponential distribution having mean λ .

2.1. Confidence Interval for λ when $\beta = 1$

As noted above, $k_2(1)$ is defined if and only if $k_1 \leq q_1^1$ where $q_1 = (-\ln \delta)^{-1}$. Therefore $100(1 - \delta)\%$ CI exists and is given by $(k_1X, k_2(1)X)$ where $k_1 = a(-\ln \delta)^{-1}$ for $a \in (0, 1)$ and $k_2(1)$ is as defined in Equation 2.5. The comparison of this CI with the CI based on chi-square distribution is made in the following table by taking $a = 1/2$ and it reveals that although the former CI is not completely contained in the latter CI, it is significantly narrower than the CI's based on chi-square distribution. Moreover, it turns out to be narrower irrespective of the value of a . As the value of a increases, the confidence interval becomes more and more narrower, and eventually it gets completely contained in the confidence interval based on chi-square distribution. X appearing in the following table represents a single observation from exponential distribution with mean λ .

Remark 2.1. *It is true that the CI based on chi-square distribution can be made shorter by moving the cut-off points to the left. However, note that in the methodology given above k_1 depends on $a \in (0, 1)$ and thus the choice of a can be made so as to ensure that the proposed CI is shorter than the former one.*

Remark 2.2. *It may be noted that, unlike the traditional CI, the proposed CI is not constructed using maximum likelihood estimator.*

Next, we consider the case $\beta > 1$.

2.2. Confidence Interval for λ when $\beta > 1$

It follows from Inequality (2.6) that $100(1 - \delta)\%$ CI (k_1X, k_2X) has an infinite length when $k_1 = q_1^{1/\beta}$. Thus it is of interest to determine values of k_1 in $(0, q_1^{1/\beta})$ for which $0 < k_1 < k_2 < \infty$. The natural β -free choice for k_1 is $\inf_{\beta > 1} q_1^{1/\beta}$. Thus

$$\inf_{\beta > 1} q_1^{1/\beta} = \begin{cases} q_1 & \text{if } \delta < e^{-1} \quad (\text{i.e. } q_1 < 1) \\ 1 & \text{if } \delta \geq e^{-1} \quad (\text{i.e. } q_1 \geq 1). \end{cases}$$

We need to consider two cases.

Case I. $\delta < e^{-1}$ (i.e. $q_1 < 1$)

In this case $\inf_{\beta > 1} q_1^{1/\beta} = q_1$, and if we take $k_1 = q_1$ then

$$\begin{aligned} k_2(\beta) &= [-\ln(1 - \delta + e^{-(\frac{1}{q_1})^\beta})]^{-1/\beta} \\ &< q_3^{1/\beta} \end{aligned} \quad (2.7)$$

where $q_3 = [-\ln(1 - \delta + e^{-(\frac{1}{k_1})})]^{-1}$ for $k_1 \in (0, q_1)$. Hence any value in the interval $(k_2(\beta), \infty)$ is a candidate for the coefficient of an upper bound. In order to have a β -free choice of the coefficient of an upper bound, one needs to consider $\bigcap_{\beta > 1} (k_2(\beta), \infty)$ which is

equal to $(\sup_{\beta > 1} k_2(\beta), \infty)$. However, $\sup_{\beta > 1} k_2(\beta) = \infty$ for $q_1 < 1$. This shows that there is no finite

upper bound when $q_1 < 1$. Further, it is worthwhile to add that for any value of $k_1 \in (0, q_1)$, the corresponding β -free upperbound is ∞ .

Thus, in this case, CI is of the type (aq_1X, ∞) , $a \in (0, 1)$, having at least $(1 - \delta)$ confidence coefficient. Note that (q_1X, ∞) is the optimal CI having $(1 - \delta)$ confidence coefficient.

Next we consider the case $\delta > e^{-1}$.

Case II. $\delta \geq e^{-1}$ (i.e. $q_1 \geq 1$)

In this case $\inf_{\beta > 1} q_1^{1/\beta} = 1$ and, therefore, the coefficient, k_1 , of lower bound is a number between $(0, 1]$. When $k_1 = 1$, $k_2(\beta) = q_2^{1/\beta}$ where

$$q_2 = [-\ln(1 - \delta + e^{-1})]^{-1}.$$

The following observations, namely,

$$q_2 > 1 \Leftrightarrow \delta < 1,$$

and

$$\delta \geq e^{-1} \Leftrightarrow q_2 \geq 0$$

ensure that an upper bound $q_2^{1/\beta}X$ is finite provided $\delta \in [e^{-1}, 1)$. Now, using arguments similar to Case I, it follows that for β -free CI the coefficient k_2 must belong to $\bigcap_{\beta > 1} (q_2^{1/\beta}, \infty)$ and

this intersection is equal to $[\sup_{q_2 > 1, \beta > 1} q_2^{1/\beta}, \infty)$. In otherwords, the coefficient $k_2 \in [q_2, \infty)$.

When $k_2 = q_2$, the CI with at least $(1 - \delta)$ confidence coefficient is given by (X, q_2X) .

Results obtained in Case I and Case II are summarized below in the form of a theorem.

Theorem 2.1. *If $X \sim W(\lambda, \beta)$, $\beta > 1$ with $F_X(x) = 1 - \exp(-(\frac{x}{\lambda})^\beta)$, then β -free confidence interval with at least $(1 - \delta)$ confidence coefficient for the scale parameter λ is (i) (q_1X, ∞) for $\delta \in (0, e^{-1})$ with $q_1 = (-\ln \delta)^{-1}$ and (ii) (X, q_2X) for $\delta \in (e^{-1}, 1)$ with $q_2 = [-\ln(1 - \delta + e^{-1})]^{-1}$.*

We note that the above theorem provides a finite length CI for $\delta \in (e^{-1}, 1)$, while for $\delta \in (0, e^{-1})$ it does not have a finite upper bound. We address the latter issue in Theorems 2.3 and 2.4. Prior to that, below, we examine the case for $\beta < 1$.

2.3. Confidence Interval for λ when $\beta < 1$

The rationale employed in Section 2.2 to obtain β -free upper and lower bounds, wherever possible, is used in this section also. To avoid repetition, it is not stated explicitly below. In this case

$$\inf_{\beta < 1} k_1 = \begin{cases} 0 & \text{if } \delta < e^{-1} \\ q_1 & \text{if } \delta \geq e^{-1}. \end{cases}$$

When $\delta < e^{-1}$, using $k_1 = 0$, Equation 2.5 yields,

$$k_2(\beta) = [-\ln(1 - \delta)]^{-1/\beta}.$$

For $0 < \delta < e^{-1}$, $[-\ln(1 - \delta)]^{-1} > 1$ we have $\sup k_2(\beta) = \infty$. Hence $100(1 - \delta)\%$ confidence interval for $\delta < e^{-1}$ and $\beta < 1$ is $(0, \infty)$.

For $\delta > e^{-1}$, we have $\inf q_1^{1/\beta} = q_1$ which is greater than or equal to one. Hence any number aq_1 with $0 < a < 1$ is a candidate for k_1 . One may verify that β -free bound for $k_2(aq_1)$ turns out to be ∞ when supremum is taken over entire range $(0, 1)$ of β . Hence in this case, the result can be summarized as given below.

Theorem 2.2. *If $X \sim W(\lambda, \beta)$, $\beta < 1$ with $F_X(x) = 1 - e^{-(\frac{x}{\lambda})^\beta}$, then β -free confidence interval with at least $(1 - \delta)$ confidence coefficient for the scale parameter λ is (i) $(0, \infty)$ for $\delta \in (0, e^{-1})$ and (ii) (q_1X, ∞) for $\delta \in (e^{-1}, 1)$ with $q_1 = (-\ln \delta)^{-1}$.*

Note that Theorem 2.2 states that with minimal information, one can not provide a finite length β -free confidence interval for the scale parameter λ of a Weibull distribution having β as its shape parameter.

These several $100(1 - \delta)\%$ CIs obtained in Theorems 2.1 and 2.2 can be summarized in the following table.

Table 2.2.

	$\delta \in (0, e^{-1})$	$\delta \in (e^{-1}, 1)$
$\beta \in (0, 1)$	$(0, \infty)$	$(q_1 X, \infty)$
$\beta \in (1, \infty)$	$(q_1 X, \infty)$	$(X, q_2 X)$

where q_1 and q_2 are as defined in Theorems 2.1 and 2.2. Next we give below a theorem which shows that if an additional information about β such as $1 < \beta_0 \leq \beta$ or $0 < \beta_0 < \beta < 1$ for known β_0 is available then it is possible to provide finite length CI for λ .

Theorem 2.3. For $\beta \in [\beta_0, \infty)$, $\beta_0 > 1$, $100(1 - \delta)\%$ CI is given by (i) $(aq_1 X, q_4^{1/\beta_0} X)$ if $\delta \in (0, e^{-1})$, (ii) $(X, q_2 X)$ if $\delta \in (e^{-1}, 1 - e^{-1})$, and, (iii) $(\max\{aq_1 X, X\}, \min\{q_2 X, X\})$ if $\delta \in (1 - e^{-1}, 1)$ where $q_1 = (-\ln \delta)$, $q_2 = [-\ln(1 - \delta + e^{-1})]^{-1}$, and $q_4 = [-\ln(1 - \delta + e^{-n})]^{-1}$. Here $a \in (0, 1)$, and $n > 1$ are such that $(\frac{1}{aq_1})^{\beta_0} > n$.

Proof. Recall that

$$k_2 = \left[-\ln \left(1 - \delta + e^{-\left(\frac{1}{q_1}\right)^\beta} \right)^{-1} \right]^{1/\beta}$$

is defined only if $k_1 \leq q_1^{1/\beta}$ where $q_1 = (-\ln \delta)^{-1}$. Now,

$$\inf_{\beta \in [\beta_0, \infty)} k_1 = \begin{cases} q_1 & \text{if } q_1 < 1 \\ 1 & \text{if } q_1 \geq 1. \end{cases}$$

Here again two cases arise.

Case 1. If $q_1 \geq 1$ let $k_1 = 1$. Then $k_2 = q_2^{1/\beta}$ where $q_2 = [-\ln(1 - \delta + e^{-1})]^{-1}$, and,

$$\sup_{\beta \in [\beta_0, \infty)} q_2^{1/\beta} = \begin{cases} 1 & \text{if } q_2 \leq 1 \\ q_2 & \text{if } q_2 \geq 1 \end{cases}$$

Hence for $\delta \in (e^{-1}, 1)$, $100(1 - \delta)\%$ CI for λ is $(X, q_2 X)$.

Case 2. If $q_1 < 1$ then $k_1 = q_1$. This yields the following inequality, namely,

$$\begin{aligned} \left[-\ln(1 - \delta + e^{-\left(\frac{1}{q_1}\right)^\beta})^{-1} \right]^{1/\beta} &= k_2(q_1) \\ &< q_3^{1/\beta} \end{aligned} \quad (2.8)$$

where $q_3 = [-\ln(1 - \delta + e^{-\left(\frac{1}{q_1}\right)})]^{-1}$. But $k_1 = q_1$ implies that $q_3 = \infty$. Hence $k_1 = q_1$ does not provide finite upper bound. To get a finite upper bound, we find a, n such that

$$\frac{1}{q_1} \left(\frac{1}{n+1} \right)^{1/\beta} < a < \frac{1}{q_1} \left(\frac{1}{n} \right)^{1/\beta}$$

and this, in turn, leads to

$$[-\ln(1 - \delta + e^{-(n+1)})^{-1}]^{1/\beta} < [-\ln(1 - \delta + e^{-\frac{1}{aq_1}})^{-1}]^{1/\beta} < [-\ln(1 - \delta + e^{-(n)})^{-1}]^{1/\beta}.$$

Therefore, $k_2(aq_1) \leq q_4^{1/\beta}$ where $q_4 = [-\ln(1 - \delta + e^{-n})]^{-1}$. Therefore,

$$\sup_{\beta \in [\beta_0, \infty)} q_4^{1/\beta} = \begin{cases} 1 & \text{if } q_4 \leq 1 \\ q_4^{1/\beta_0} & \text{if } q_4 \geq 1. \end{cases}$$

Therefore, for $\delta \in (0, e^{-1})$ $100(1 - \delta)\%$ CI for λ is given by $(aq_1X, q_4^{1/\beta_0}X)$. If $q_4 < 1$ then $100(1 - \delta)\%$ CI for λ is given by (aq_1X, X) .

Theorem 2.4. For $\beta \in (\beta_0, 1)$, $\beta_0 > 0$, $100(1 - \delta)\%$ CI for λ is (i) $(aq_1X, q^{1/\beta_0}X)$ if $\delta \in (e^{-1}, 1 - e^{-1})$, (ii) (aq_1X, qX) if $\delta > 1 - e^{-1}$ where $q = [-\ln(1 - \delta + e^{-m})]^{-1}$, $0 < a < \frac{1}{q_1}$ and $m > 1$ are numbers such that $a < \frac{1}{m^{\frac{1}{\beta_0}} q_1}$.

Proof. Let $\beta \in (\beta_0, 1)$. Note that for $a \in (0, \frac{1}{q_1})$, there exists $m(\beta_0, \delta)$ such that $a < \frac{1}{m^{\frac{1}{\beta_0}} q_1}$.

This leads to

$$[-\ln(1 - \delta + e^{-\left(\frac{1}{aq_1}\right)^\beta})^{-1}]^{1/\beta} < [-\ln(1 - \delta + e^{-m})]^{-1/\beta} = q^{1/\beta}$$

where $q = [-\ln(1 - \delta + e^{-m})]^{-1}$. Therefore,

$$\sup_{\beta \in (\beta_0, 1)} q^{\frac{1}{\beta}} = \begin{cases} q & \text{if } q \leq 1 \\ q^{1/\beta_0} & \text{if } q \geq 1. \end{cases}$$

Therefore, for $\delta > 1 - e^{-1}$, $100(1 - \delta)\%$ CI for λ is given by (aq_1X, qX) whereas the same for $\delta \in (e^{-1}, 1 - e^{-1})$ is equal to $(aq_1X, q^{1/\beta_0}X)$. In view of Theorems 2.3 and 2.4, Table 2.2 is modified accordingly to contain more precise information on bounds on λ and is given below in the form of Table 2.3.

Here q , q_1 , q_2 and q_4 are as defined in Theorems 2.3 and 2.4.

Remark 2.3. For completeness sake, confidence intervals are provided for various values of $\delta \in (0, 1)$. For $\beta < 1$ and for smaller values of δ , finite intervals do not exist and hence are not available. However for $\beta > 1$, finite confidence intervals having confidence coefficient to be at least 0.63 can be obtained. It may, further, be noted that in practice there are instances in which actual coverage probabilities are as low as 30% for 95% or 99% confidence intervals for parameters of interest. Thus with confidence intervals having at least 0.63 confidence coefficient, it should be considered quite reasonable even if the coverage probability is about 50%.

Remark 2.4. As mentioned earlier, it may be noted that the issue of optimality of CIs has not been considered in this paper.

Table 2.3.

	$\delta \in (0, e^{-1})$	$\delta \in (e^{-1}, 1 - e^{-1})$	$\delta \in (1 - e^{-1}, 1)$
$0 < \beta_0 < \beta < 1$	$(0, \infty)$	$(aq_1X, q^{\frac{1}{\beta_0}}X)$	(aq_1X, qX)
$\beta > \beta_0 > 1$	$(aq_1X, q^{\frac{1}{\beta_0}}X)$	(X, q_2X)	(X, q_2X)

Remark 2.5. *Edelman (1990) showed that if X is a continuous random variable with uni-modal density that has mode θ , then for any fixed a and $t > 1$,*

$$P[X - t|X - a| \leq \theta \leq X + t|X - a|] \geq 1 - \frac{2}{t+1}.$$

Note that for Weibull distribution with parameters λ and β ,

$$\theta = \begin{cases} 0 & \text{if } \beta \leq 1 \\ \lambda(\frac{\beta-1}{\beta})^{1/\beta} & \text{if } \beta > 1. \end{cases}$$

Thus if $\beta > 1$, then Edelman's bounds with $a = 0$ acquire the following form, namely,

$$X - t|X| \leq \lambda(\frac{\beta-1}{\beta})^{1/\beta} \leq X + t|X|.$$

This implies that

$$(X - t|X|)(\frac{\beta}{\beta-1})^{1/\beta} \leq \lambda \leq (X + t|X|)(\frac{\beta}{\beta-1})^{1/\beta}.$$

Therefore, β -free bounds on λ can be obtained by taking the infimum and supremum w.r.t. $\beta(> 1)$ of the lower and upper bounds, respectively, of the above inequality. On comparing, it is found that β -free CI of λ obtained using Edelman's result is of infinite length while CIs obtained using Theorems 2.3, 2.4 all have finite lengths.

2.4. Numerical Results

From Table 2.3 it is clear that the meaningful numerical results are possible only for $\beta > \beta_0 > 1$ and $(1 - \delta) \in (0.63212, 1)$ and the same are reported below for $\beta_0 = 2$ and $(1 - \delta) = 0.95$. It may, further, be noted that the meaningful confidence intervals are possible only when a is smaller than the value shown against each confidence interval.

Table 2.4.

m	confidence interval
3	$(a(0.3338)X, 68.4981X), a < 1.7296$
4	$(a(0.3338)X, 5.5730X), a < 1.4979$
5	$(a(0.3338)X, 4.7551X), a < 1.3398$
6	$(a(0.3338)X, 4.5300X), a < 1.2230$

3. Bounds for λ Based on $n(\geq 2)$

Here, unlike in Section 2, we assume that $n(\geq 2)$ independent observations $X_1, X_2, \dots, X_n$ are available. Further assume that X_i observation follows Weibull distribution with λ_i as its scale parameter, and β as its shape parameter for $i = 1, 2, \dots, n$. We see in this section that a closure property of Weibull distribution together with results in Subsections 2.1, 2.2 and 2.3 yields a number of very useful confidence intervals for various functions of the original scale parameter λ_i . These, in turn, provide several important results in the context of reliability studies. We first state below a closure property of a Weibull distribution in the form of a theorem.

Theorem 3.1. *If $X_1, X_2, \dots, X_n$ are independent random variables such that X_i has Weibull distribution with parameters λ_i & β , $i = 1, 2, \dots, n$ then $X_{1:n} = \min(X_1, \dots, X_n)$ follows $W(\lambda_0, \beta)$ where $\lambda_0 = \left(\sum_{i=1}^n \lambda_i^\beta\right)^{1/\beta}$.*

In view of this result (and the results in Subsections 2.2-2.3) it is quite clear that $100(1 - \delta)\%$ CIs for λ_0 can easily be constructed along the lines of the ones in these two sections. We emphasize that these are $100(1 - \delta)\%$ CIs based on n independent observations from Weibull distribution and are in terms of the minimum of n observations.

Remark 3.1. *Note that $X_{1:n} = \min(X_1, X_2, \dots, X_n)$ can be viewed as the lifetime of a series system comprising of components having independent Weibull lifetimes. As the confidence bounds for λ_0 are based on $X_{1:n}$, they, in turn, can be used to yield confidence bounds for reliability of the corresponding system. This observation has significant practical value as these bounds on reliability of the system are available for any life distribution which is a member of the Weibull family of distribution functions and not necessarily that of only exponential family of distribution functions.*

Remark 3.2. *Let $\lambda_{\min} = \min\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and $\lambda_{\max} = \max\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. As $n^{1/\beta}(\lambda_{\min}) \leq \lambda_0 \leq n^{1/\beta}(\lambda_{\max})$, suitable confidence intervals for $\lambda_{\min}, \lambda_{\max}$ can be determined. However, it should be noted that these confidence intervals are one-sided and have confidence coefficient greater than or equal to $(1 - \delta)$. Further, the above remark regarding confidence intervals on the reliability of the system is applicable when component lifetimes are independent, Weibully distributed random variables having $\lambda_{\min}$ or $\lambda_{\max}$ and β as parameters.*

Remark 3.3. If in addition $X_1, X_2, \dots, X_n$ are also identically distributed then $\lambda_0 = n^{1/\beta}\lambda$. Therefore using results of Section 2, CIs for λ would become available for different ranges of β and δ .

Remark 3.4. The dependence of λ_0 on β , in general, does not pose any problem because metallurgical properties suggest that β can be determined using Hooke's law (Barlow (1998)) and thus, in principle, it can be assumed to be known.

3.1. Numerical Results

Let $\beta = 1$. A two-sided $100(1 - \delta)\%$ confidence interval for λ based on the i^{th} order statistic is given by

$$\left[\frac{X_{i:n}}{-\ln[B_{\delta/2}(n-i+1, i)]}, \frac{X_{i:n}}{-\ln[B_{1-\delta/2}(n-i+1, i)]} \right]$$

where $B_r(p, q) = \frac{pF_r(2p, 2q)}{q + pF_r(2p, 2q)}$. $B_r(p, q)$ and $F_r(2p, 2q)$ are the percentage points of the beta and F distributions respectively. This result is available in Bain and Engelhardt (1991, page 185). For $i = 1$, this two-sided $100(1 - \delta)\%$ confidence interval for λ becomes

Table 3.1.

n	δ	Value of $X_{1:n}$ in a sample of size n drawn from Exp(1) distribution	100(1 - δ)% confidence	
			interval given in Bain & Engelhardt	interval proposed in this paper
1	0.02	1.9229	[0.4174, 191.3241]	(0.2457, 96.4481)
	0.05		[0.5213, 75.9495]	(0.3209, 39.5128)
	0.10		[0.6419, 37.4960]	(0.4175, 20.3886)
2	0.02	1.3196	[0.5731, 262.5657]	(0.0843, 33.3308)
	0.05		[0.7154, 104.0652]	(0.1101, 13.5576)
	0.10		[0.8812, 51.4936]	(0.1432, 6.9957)
3	0.02	1.0184	[0.6636, 308.1093]	(0.0434, 17.1500)
	0.05		[0.8282, 120.6740]	(0.0567, 6.9759)
	0.10		[1.0204, 59.3955]	(0.0737, 3.5996)
4	0.02	0.4055	[0.3521, 166.9972]	(0.0130, 5.1203)
	0.05		[0.4397, 64.1621]	(0.0169, 2.0831)
	0.01		[0.5413, 31.7260]	(0.0220, 1.0750)
5	0.02	1.4198	[1.5415, 706.0310]	(0.0363, 14.3456)
	0.05		[1.9235, 277.6901]	(0.04739, 5.8352)
	0.10		[2.3710, 138.4346]	(0.0617, 3.0110)
6	0.02	0.3865	[0.5033, 227.1625]	(0.0082, 3.2542)
7	0.02	0.2339	[0.3557, 166.9687]	(0.00427, 1.6882)

$$\left[\frac{X_{1:n}}{-\ln[B_{\delta/2}(n, 1)]}, \frac{X_{1:n}}{-\ln[B_{1-\delta/2}(n, 1)]} \right]. \quad (3.1)$$

In view of Theorem 3.1 and Remarks 3.3 and 3.4, confidence intervals indicated in Table 2.3 would become applicable for $n > 1$ with X being replaced with $X_{1:n}$. The comparison of these resulting confidence intervals with those given in (3.1) is carried out here for various values of δ by drawing random samples of sizes ranging from 1 to 10 from the exponential distribution with mean 1. These results are presented below in Table 3.1.

It is quite evident from this table that for samples having sizes 1 through 5 and for $\delta = 0.02, 0.05, 0.1$, the newly proposed confidence intervals are significantly narrower than those given in Bain and Engelhardt (1991, page 185) inspite of the fact that the latter ones are obtained using maximum likelihood approach. However, for samples of sizes 6 through 8, similar observation is available only for delta $\delta = 0.02$. This is because for other values of $\delta = 0.05$ and 0.1 , and for $n = 6, 7, 8$, the confidence intervals proposed in this paper do not include the value of the parameter $\lambda = 1$.

Acknowledgment

The 1st and 2nd authors would like to thank Department of Science and Technology, and CSIR, Government of India, New Delhi, respectively, for providing grant to carry out this research work.

References

- Abbot, J.H., and Rosenblatt, J. (1963). Two Stage Estimation With One Observation on the First Stage. *Annals of the Institute of Statistical Mathematics*. 14, 229-235.
- Bain, L.J., and Engelhardt M. (1991). *Statistical Analysis of Reliability And Life-Testing Models*, 2nd Ed, Marel Dekker, New York.
- Barlow, R.E. (1998). Engineering Reliability. *ASA-SIAM Series on Statistics and Applied Probability*.
- Blachman, N.M., and Machol, R.E. (1987). Confidence Intervals Based on One or More Observations. *IEEE Transactions on Information Theory*. IT-33, No.3, 373-382.
- Edelman, D. (1990). A Confidence Interval for the Center of an Unknown Unimodal Distribution Based on a Sample Size 1. *The American Statistician*. 44, 285-287.
- Machol, R.E., and Rosenblatt, J. (1966). Confidence Interval Based on Single Observation. *Proceedings of IEEE*. 54, 1087-1088.
- Wall, M.M., Boen J. and Tweedie R. (2001). An Effective Confidence Interval for the Mean With Samples of Size One and Two. *The American Statistician*. 55, 102-105.