Journal Of Harmonized Research (JOHR)

Journal Of Harmonized Research in Management
2(1), 2016, 79-88

ISSN 2454-5384

Original Research Article

LINEAR DISCRIMINATORY FUNCTION APPROACH TO EMPLOYMENT SEX RATIO VARIATION: EVIDENCE FROM ASIA

Jayadevan CM

Charles Stuart University, Australia

Abstract: - Employment plays an important role in everyone's daily life as it is an instrument for achieving any accomplishments. The main purpose of this paper is to analyze the variation in average employment sex ratio among the countries of Asia on the basis of available statistical data for the period 1991-2013 and shed some light for women's employment management. Using a linear discriminant function analysis technique, the study has shown that significant discriminating factors responsible for the variation in women's employment are gross parity index for secondary enrollment, employer's sex ratio and the percentage of women employed in service sector.

Key Words: Employment sex ratio, primary, secondary, tertiary, GPI

Introduction: Employment patterns of men and women have changed over the last 25 years. Significantly, the proportion of women who were employed has increased over the period. Changing social attitudes and smaller families have contributed to these changes in women's employment. Greater proportions of women now have higher education qualifications. Education appears to draw women into the workforce by instilling in them more career oriented attitudes and by enhancing their potential wages in the labour market. Paid work may also provide women with opportunities for social interaction and job satisfaction (ABS, 2006).

The main purpose of this paper is to analyze the variation in average overall employment sex ratio among the countries of Asia on the basis of available statistical data for the period 1991-2013. The average overall employment sex ratio have been classified into three categories. The study makes an attempt to find the factors responsible for the variation in average overall employment sex ratio. Higher the value of factors like gross parity index for enrollment, employers sex ratio, %share of females in agriculture, industry and services sectors, gender wage gap and adult literacy, higher the overall employment sex ratio. In other words, factors like gross parity index for enrollment, employers

For Correspondence:
devcmj@gmail.com
Received on: February 2016
Accepted after revision: February 2016
Downloaded from: www.johronline.com

sex ratio, %share of females in agriculture, industry and services sectors, gender wage gap and females adult literacy play a positive role in enhancing the overall employment sex ratio.

Materials and Methods: The main source of data for this study is taken from online statistical database published by United Nations ESCAP. 1)EMP_SEX_RATIO: Employment sex ratio, overall employment [Employed females per 100 employed males], 2)WAGE_GAP: Gender wage gap [Percentage]: The gender wage gap is the difference between gross average nominal monthly wages of male and those of female employees expressed as a percentage of gross average nominal monthly wages of male employees. Indicator calculations: Gender pay gap (%) = 100*(Em – Ew)/Em where Em is the gross average nominal monthly wages of men in any given population group and Ew is the gross average nominal monthly wages of women, 3)EMPRS_SEX_RATIO: Employers sex ratio [Female employers per 100 male employers], 4)Agriculture employment, female [% of employed females], 5)IND_EMP_F: Industry employment, female [% of employed females], 6)SER_EMP_F: Services employment, female [% of employed females], 7)GPI_PRIM:School enrollment, primary (gross), gender parity index (GPI)- Gender parity index for gross enrollment ratio in primary education is the ratio of girls to boys enrolled at primary level in public and private schools, 8)GPI_SEC: School enrollment, secondary (gross), gender parity index (GPI)- Gender parity index for gross enrollment ratio in secondary education is the ratio of girls to boys enrolled at secondary level in public and private schools, 9)GPI_TER:School enrollment, tertiary (gross), gender parity index (GPI)- Gender parity index for gross enrollment ratio in tertiary education is the ratio of women to men enrolled at tertiary level in public and private schools.

In this paper we will use the linear discriminant analysis (LDA) as a technique for analyzing overall employment sex ratio variation. LDA is a statistical technique designed to investigate the differences between two or more groups of people with respect to several underlying variables. Because the variable being predicted is categorical, LDA technique is more appropriate than commonly used measures. LDA performs a multivariate test of differences between groups. In addition, LDA is used to determine the minimum number of dimensions needed to describe these differences.

LDA is used to analyze relationships between a dependent variable and independent variables. Overall employment sex ratio has been considered as the dependent variable. Since this is a discrete variable, this has been classified into three categories, that is 1) 0-50, 2) above 50 to 80 and 3) above 80 to 100.There are 9 predictor variables:1) gender wage gap, 2)employers sex ratio, 3) percentage of employed women in agriculture, 4) percentage of employed women in industry, 5) percentage of employed women in services, 6) primary GPI, 7) secondary GPI, 8)tertiary GPI and 9) female adult literacy rate. LDA analysis attempts to use the predictor variables to distinguish among the groups of the response variable. If LDA is able to distinguish among groups, it must have a strong relationship to at least one of the predictor variables. Using LDA, a series of statistical tests are conducted to test the overall relationship among the predictor variables and groups defined by the response variable.

This paper is mainly concerned with an analysis to determine if there is a significant effect of factors like gender wage gap, employers sex ratio, percentage of employed women in agriculture, percentage of employed women in industry, percentage of employed women in services, primary GPI, secondary GPI, tertiary GPI and adult literacy rate on the overall employment sex ratio.

The hypothesis of interest is:

$$H_0: \beta_1 = \beta_2 = \beta_3 \ldots = \beta_8 = 0$$

$$H_a: \text{Not all } \beta_i \text{ equal zero}$$

This hypothesis has been tested using LDA. The test statistic used for LDA is $\textit{Wilk's Lambda } \Lambda = \prod_i \frac{1}{1+\lambda_i}$. where λ_i are the

eigen values of the corresponding design matrices. There are three main assumptions for LDA: they are 1) Multivariate Normality (MVN): To test for MVN, we begin by examining the marginal distributions of each univariate variable using box plots. If any of these plots show non-normality, then MVN is suspect and we use a procedure based on Mahalanobis distance, in which we construct a $\chi 2$ probabilities to determine conformity with multivariate normality. 2) Equality of covariances: the test for equality of covariances is based on Box's M-test and 3) Independence of observations: This test is a function of the experimental design, or data collection method and hence is not tested. For the purposes of this paper we assume that it is true.

Empirical Results: The average overall employment sex ratio was 69.65 females per 100 males during the period 1991-2013. However, the overall employment sex ratio varied across countries of Asia. On the basis of average overall employment sex ratio, countries of Asia have been divided into three categories, i.e. 1) 50 or below 2) Above 50, but upto 80 and 3) above 80. The average overall employment sex ratio was 92.36 females per 100 males for the third group, 67.60 females per 100 males for the second group and 31.60 females per 100 males for the first group (Table 1).

Table 1: Summary Statistics of Overall Employment Sex Ratio

Rank	Mean	N	Std. Deviation	Minimum	Maximum	Median	Kurtosis	Skewness
1	31.60	161	14.28	10.60	63.70	35.20	-1.03	.21
2	67.60	344	8.78	48.80	92.20	67.30	.02	.44
3	92.36	301	6.59	78.60	104.90	93.50	-.88	-.14
Total	69.65	806	23.95	10.60	104.90	70.70	-.18	-.74

Countries like Afghanistan(15.73), Pakistan(18.17), Iran (18.80), India(37.95), Turkey(40.29), Sri Lanka(44.82) and Maldives(45.47) had, generally, an average overall employment sex ratio of below 50 females per 100 males. Countries like Malaysia, Indonesia, Philippines, Bangladesh, Bhutan, Brunei Dar., Turkmenistan, Uzbekistan, Singapore, Rep. of Korea, Japan, Hong Kong, Kyrgyzstan, Tajikistan, and Armenia had an average overall employment sex ratio between 50 to 80 females per 100 males during the same period. On the other hand, countries like China, Thailand, Mongolia, Azerbaijan, Georgia, Kazakhstan, Nepal, Viet Nam, DPR Korea, Russian Fed., Myanmar, Cambodia and Lao PDR had an average overall employment sex ratio between 80 or above females per 100 males. Country-wise average overall employment sex ratio is shown in Fig.1.

Fig.1: Average Employment Sex Ratio in Asia during 1991-2013

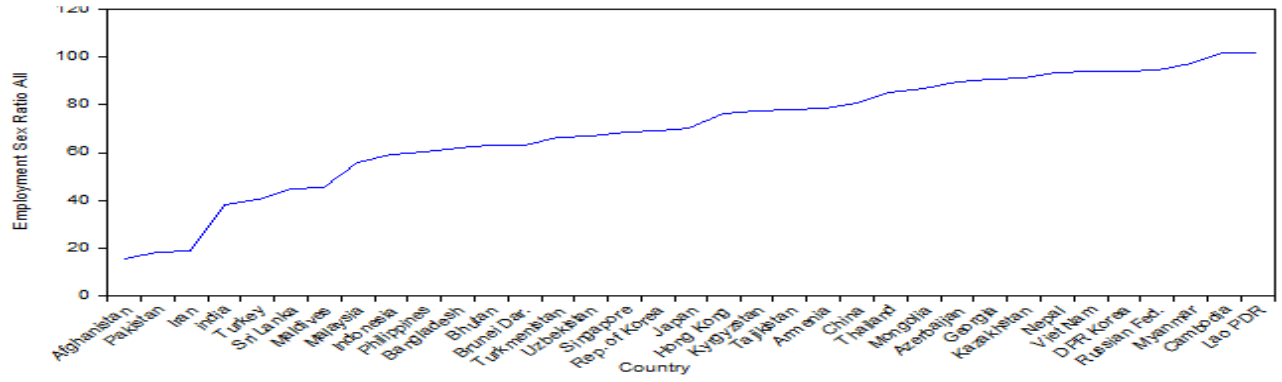


Table 2:Groups Statistics

	Group 1		Group 2		Group 3		Total	
	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
WAGE_GAP	16.47	15.77	22.25	25.71	31.94	18.94	25.37	20.71
EMPRS_SEX_RATIO	6.35	4.08	21.42	7.10	47.46	26.16	29.96	26.35
GPI_PRIM	.92	.12	1.04	.03	.96	.07	.96	.09
GPI_SEC	.88	.13	1.16	.03	1.05	.11	1.02	.15
GPI_TER	1.06	.47	1.34	.12	1.09	.47	1.13	.44
LIT	70.32	23.28	96.64	2.77	86.59	23.54	83.66	23.06
AGRI_EMP_F	56.11	14.86	35.24	9.46	53.15	15.85	50.52	16.36
IND_EMP_F	17.00	6.33	10.74	2.13	10.45	5.53	12.48	6.08
SER_EMP_F	25.01	8.80	54.19	8.52	36.73	14.25	36.63	15.50

The average gender wage gap for group 3 (31.94) is higher than group 2(22.25) and group 1(16.47), however, the standard deviation is higher for group 2. Employers sex ratio is higher for group 3 (47.46 females for 100 males) than the group 2 (21.42 females for 100 males) and group 1(6.35 females for 100 males). The average secondary GPI is 1.05 for group 3, 1.16 for group 2 and 0.88 for group 1. The average percentage of females employed in agriculture is higher for group 1 (56.11) and group 3(53.15) and it is very low for group 2(35.24).The average percentage of females employed in industry is higher for group 1 (17.00) than for group 2 and 3. The average percentage of females employed in service sector is higher for group 2 (54.19) than for group 3(36.73) and group 1 (25.01). The average adult literacy rate was 96.64% for group 2 and 86.59% for group 3 and 70.32% for group 1(Table 2).

Group-wise box plots for different variables are shown below. For overall employment sex ratio and wage gap, extreme values are not observed, but higher median and variance is observed for groups 3 and 2. For employers sex ratio, extreme values are observed for group 2, but higher median and variance is observed for groups 3 and 2 (Fig.4).

Fig2: Box Plot for Overall Employment Sex Ratio

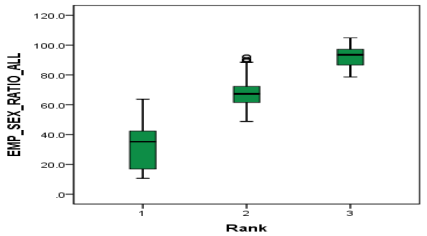


Fig 3: Box Plot for Wage gap

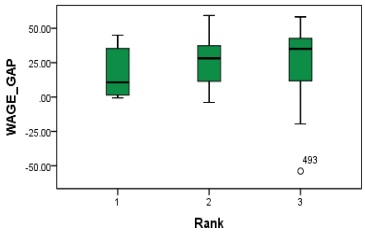
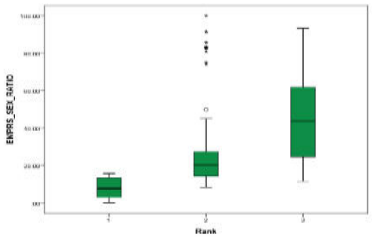


Fig 4: Box Plot for Employers Sex ratio



For % share of female's agricultural employment, no extreme values are observed, but higher median values is observed for group 1, however, higher variance is observed for groups 2 and 3 (Fig 5). For % share of industrial employment, extremes values are observed for group 2, but higher median is observed for group 1, however, higher variance is observed for groups 1 and 2(Fig 6). For the % share of female's service employment, no extreme values are observed, but higher median and variance is observed for group 2 followed by group 3(Fig 7). For primary GPI, extreme values are observed for all groups (Fig 8). For secondary GPI, extreme values are observed for group 3 (Fig 9). For tertiary GPI, extreme values are observed for groups one and three (Fig 10).

Fig 5: Box Plot for % share of Female's Agricultural. Employment.

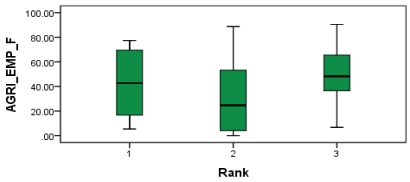


Fig 6: Box Plot for % share of Women's Industrial Employment

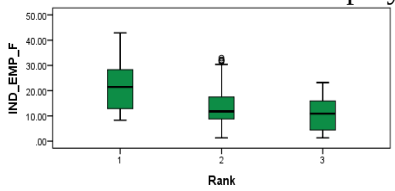


Fig 7: Box Plot for % share of Women's Service Employment

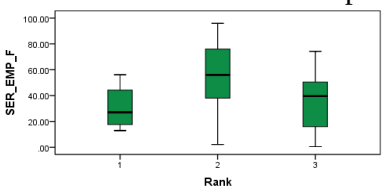


Fig 8:Box Plot for Primary GPI

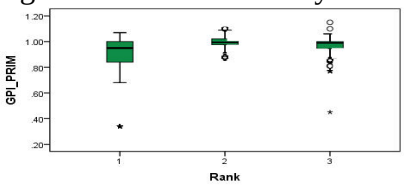


Fig 9:Box Plot for Secondary GPI

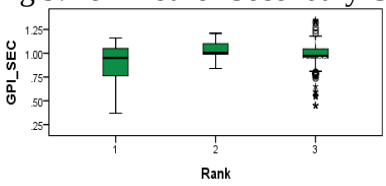


Fig 10:Box Plot for Tertiary GPI

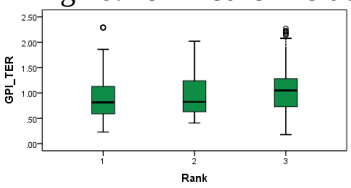


Table 3:Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		229	28.4
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	578	71.6
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	578	71.6
Total		807	100.0

The minimum ratio of valid cases to independent variables for LDA is 5 to 1. In this case, it is 229/9≈ 25 to 1, which satisfies the minimum requirement and it does satisfy the preferred ratio of 20 to 1(Table 3).

Table 4:Prior Probabilities for Groups

Rank	Prior	Cases Used in Analysis	
		Unweighted	Weighted
1	.333	69	69
2	.333	45	45
3	.333	115	115
Total	1.000	229	229

The number of cases in the smallest group in this problem is 45, which is larger than the number of predictor variables (9), satisfying the minimum requirement. In addition, the number of cases in the smallest group satisfies the preferred minimum of 20 cases (Table 4).In this analysis there were 3 groups defined by category of overall employment sex ratio, 9 independent variables, so the maximum possible number of discriminant functions was 2. The canonical correlations for the dimensions one and two are 0.95 and 0.79, respectively (Table 5).

Table 5:Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	8.54	83.66	83.66	.95
2	1.67	16.34	100.00	.79

Table 6:Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.039	720.21	16	.000
2	.375	218.36	7	.000

In the table of Wilk's lambda which tested functions for statistical significance, the stepwise analysis identified 2 discriminant functions that were statistically significant. The Wilk's lambda statistic for the test of function 1 through 2 functions (chi-square=720.21) had a probability of 0.000 which was less than the level of significance of 0.05. The Wilk's lambda statistic for the test of function 2 (chi-square=218.36) had a probability of 0.000 which was less than the level of significance of 0.05. The significance of the maximum possible number of discriminant functions supports the interpretation of a solution using 2 discriminant functions (Table 6).

Table 7:Functions at Group Centroids (Unstandardized canonical discriminant functions evaluated at group means

	Function	
Rank	1	2
1	-4.329	-.396
2	2.992	-2.232
3	1.427	1.111

Table 7 shows unstandardized canonical discriminant functions evaluated at group means. Function 1 separates the overall employment sex ratio category 1(the negative value of 4.329) from overall employment sex ratio category 2(positive value of 2.992) and overall employment sex ratio category 3(positive value of 1.427). Function 2 separates the overall employment sex ratio category 3(the positive value of 1.111) from overall employment sex ratio category 1(negative value of -0.396) and overall employment sex ratio category 3 (negative value of -2.232).

Table 8:Variables Entered/Removed[a,b,c,d]

		Min. D Squared					
			Between	Exact F			
Step	Entered	Statistic	Groups	Statistic	df1	df2	Sig.
1	GPI_SEC	1.095	2 and 3	35.42	1	226.00	.000
2	EMPRS_SEX_RATIO	2.905	2 and 3	46.77	2	225.00	.000
3	IND_EMP_F	4.025	2 and 3	43.01	3	224.00	.000
4	SER_EMP_F	4.721	2 and 3	37.67	4	223.00	.000
5	LIT	7.855	2 and 3	49.91	5	222.00	.000
6	GPI_PRIM	12.385	2 and 3	65.28	6	221.00	.000
7	GPI_TER	12.923	2 and 3	58.13	7	220.00	.000
8	WAGE_GAP	13.626	2 and 3	53.38	8	219.00	.000

At each step, the variable that maximizes the Mahalanobis distance between the two closest groups is entered. a. Maximum number of steps is 16. b. Maximum significance of F to enter is .05, c. Minimum significance of F to remove is .10, d. F level, tolerance, or VIN insufficient for further computation.

When we use the stepwise method of variable inclusion,we limit our interpretation of predictor variables to those listed as statistically significant in the table of variables Entered/Removed. We will interpret the impact on membership in groups defined by the response variable by the predictor variables:1)Secondory GPI 2)employers sex ratio, 3)%share of female employment in industrial sector, 4))%share of female employment in service sector, 5)Adult literacy rate, 6)Primary GPI, 7)Tertiary GPI and 8)Wage gap(Table 8).Differences between overall employment sex ratio observed for groups 2 and 3 is mainly caused by the higher GPI at seconday, primary and tertiary levels, employer's sex ratio, %share of female employment in industrial sector, %share of female employment in service sector, adult literacy rates and wage gap.

Table 9:Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
WAGE_GAP	.889	14.134	2	226	.000
EMPRS_SEX_RATIO	.514	106.881	2	226	.000
IND_EMP_F	.761	35.516	2	226	.000
SER_EMP_F	.577	82.968	2	226	.000
GPI_PRIM	.802	27.934	2	226	.000
GPI_SEC	.529	100.608	2	226	.000
GPI_TER	.945	6.557	2	226	.002
LIT	.828	23.483	2	226	.000

Using Wilk's lambda and step-wise LDA, the variables that minimizes the overall Wilk's lambda is entered. In our case, employers sex ratio, secondary GPI and %share of female employment in service sector are significant (Table 9).

Table 10:Structure Matrix

	Function	
	1	2
GPI_SEC	.314*	-.166
IND_EMP_F	-.186*	-.103
LIT	.153*	-.071
EMPRS_SEX_RATIO	.240	.523*
SER_EMP_F	.256	-.322*
GPI_PRIM	.147	-.196*
WAGE_GAP	.088	.189*
GPI_TER	.051	-.147*

Based on the structure matrix, the predictor variables strongly associated positively with discriminant function 1 which distinguished between overall employment sex ratio categories are Secondary GPI(r=0.314).Based on the structure matrix, the predictor variable strongly associated positively with discriminant function 2 which distinguished between overall employment sex ratio categories is employers sex ratio (r=0.523).Other predictor variable

strongly associated with discriminant function 2 which is strongly associated negatively with overall employment sex ratio is %share of female employment in service sector(Table 10).

The number of discriminant dimensions is the number of groups minus 1. However, some discriminant dimensions may not be statistically significant. In this example, there are two discriminant dimensions, both of which are statistically significant. The Coefficients of linear discriminants are reported in Table 11.

The equations of the linear discriminante function are:

1)discriminant_score_1=0.593*WAGE_GAP+0.226*EMPRS_SEX_RATIO-0.830*IND_EMP_F+0.490*SER_EMP_F-0.140*GPI_PRIM+2.145*GPI_SEC-1.225*GPI_TER+0.141*LIT

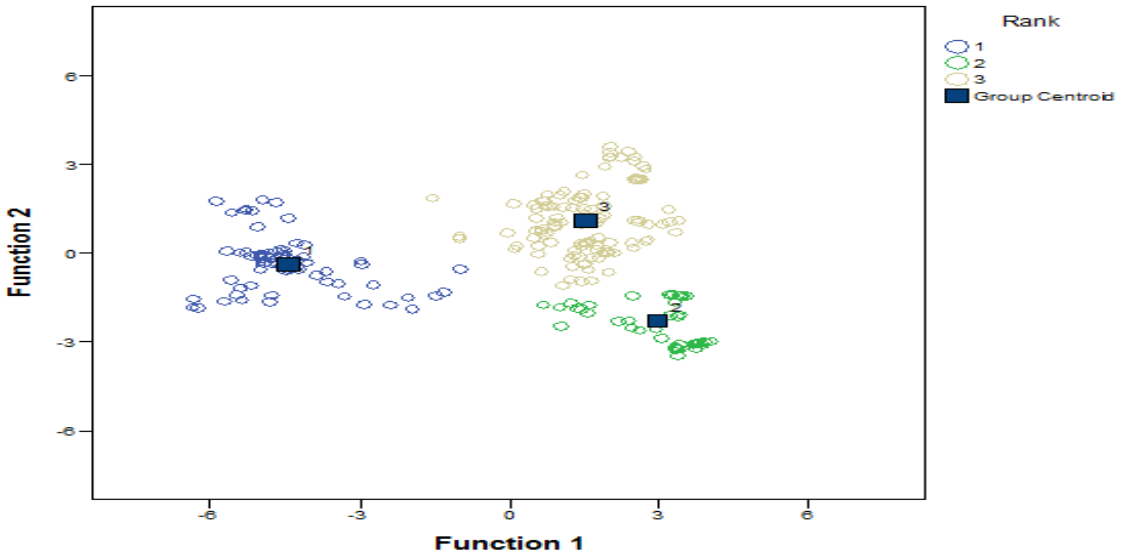
2)2)discriminant_score_2= -.070*WAGE_GAP+0.908*EMPRS_SEX_RATIO+0.003*IND_EMP_F-1.203*SER_EMP_F-1.574*GPI_PRIM-0.214*GPI_SEC+0.061*GPI_TER+2.57*LIT

Table 11:Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
WAGE_GAP	.593	-.070
EMPRS_SEX_RATIO	.226	.908
IND_EMP_F	-.830	.003
SER_EMP_F	.490	-1.203
GPI_PRIM	-.140	-1.574
GPI_SEC	2.145	-.214
GPI_TER	-1.225	.061
LIT	.141	2.570

As you can see, the overall employment sex ratio category 1 tend to be more at the GPI, tertiary (negative) end of dimension 1. The overall employment sex ratio categories 2 and 3 tend to be at the opposite end in the dimension one. On dimension 2, the overall employment sex ratio category 3 tend to be higher on literacy and categories 1 and 2 lower on GPI, primary and SER_EMP_F (Fig 11)

Fig.11: Canonical Discrimination Functions



The cross validated accuracy rate computed by SPSS was 98.3% which was greater than the proportional by chance accuracy criteria of 41.25% (1.25*33.0=41.25). The criteria for classification accuracy is satisfied (Table 12). The proportional by chance accuracy rate was computed by squaring and summing the proportion of cases in each group from the table of prior probabilies for groups (0.333^2 + 0.333^2 + 0.333^2 =33.0).

Table 12:Classification Results[a,c]

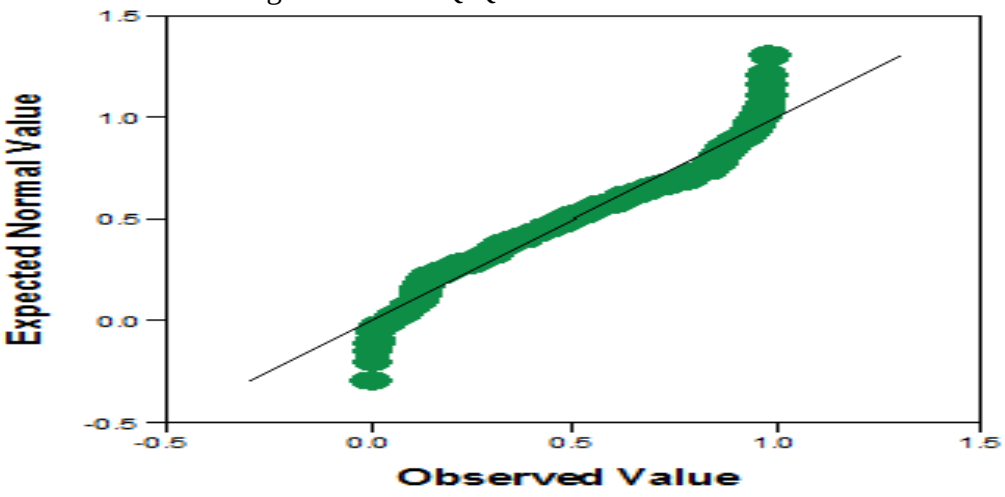
		Rank	Predicted Group Membership			Total
			1	2	3	
Original	Count	1	68	0	1	69
		2	0	45	0	45
		3	0	1	114	115
	%	1	98.6	.0	1.4	100.0
		2	.0	100.0	.0	100.0
		3	.0	.9	99.1	100.0
Cross-validated[b]	Count	1	68	0	1	69
		2	0	45	0	45
		3	0	3	112	115
	%	1	98.6	.0	1.4	100.0
		2	.0	100.0	.0	100.0
		3	.0	2.6	97.4	100.0

Note:a. 99.1% of original grouped cases correctly classified. b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case. c. 98.3% of cross-validated grouped cases correctly classified.

Apart from linearity the main assumptions in LDA are:

1) MVN errors: The first assumption can be checked using Mahalanobis plot although symmetry is probably more important. If normality can not be induced by transformation or if the data are seriously non normal ie categorical, then the alternative of logistic regression should be used. It is worth pointing out that if all the assumptions are satisfied, lda is the optimal procedure and so should be used.

Fig.15: Normal Q-Q Plot for Multivariate Data



The plot of ordered Mahalanobis distances against their expected values under the assumption of Multivariate Normality clearly shows slight deviation from the straight line. So we conclude that the assumption of multivariate normality is approximately upheld (Fig.15).

2) **Box's Test of Equality of Covariance Matrices**.

For the second assumption there is a test of equality of covariances matrices, Box's M test. Violation of this assumption can affect significance tests of classification results. The significance level can be inflated (false positives) when the number of variables is large and the sample sizes of the groups differ. Quadratic methods can be used if the covariance matrices are unequal but a large number of parameters are involved and LDA is thus superior for small sample sizes. Overall LDA is robust to both the assumption of MVN and equality of covariance matrices, especially if the sample sizes are equal. The formal hypothesis forBox's M test for Equality of covariance would be:

$H_0: \Sigma 1 = \Sigma 2 = \Sigma 3$, $H_0: \Sigma 1 \neq \Sigma 2 \neq \Sigma 3$

α = 0.05, $Fobs = \frac{MS_{Regression}}{MS_{Residual}}$

Reject H_0 if p-value <0.05

Do reject H_0 as p-value = 0.000<0.05

Table 13:Test Results

Box's M		1795.691
F	Approx.	23.454
	df1	72
	df2	61823.925
	Sig.	.000

Test Statistic

$$M = \sum n_i \ln|s| - \sum_{i=1}^{k} n_i \ln|s_i|$$

$$C^{-1} = 1 - \frac{2p^2+3p-1}{6(p+1)(k-1)}\left(\sum_{n-1}^{k}\frac{1}{n_i} - \frac{1}{\sum n_i}\right)$$

Sampling Distribution

$$MC^{-1} \sim \frac{\chi^2 (k-1)(p)(p+1)}{2} \text{ if } k, p < 5 \text{ and } n_i \approx 20 \text{ else } F \text{ distribution}$$

To test the assumption of equality of Co-variances, we use Box's M-test. If the Box's M Test shows p <.05, the covariances are significantly different and the null hypothesis is NOT rejected. If the Box's M Test shows p >.05, the covariances are not significantly different and the null hypothesis is not rejected. The value of Box's M is 1795.69, with a p-value of 0.00, indicating that the assumption of equal co-variances is not satisfied and null hypothesis is rejected. So the assumption of homoscedasticity is violated. That is we do not reject the null hypothesis of $H_0: \Sigma 1 = \Sigma 2 = \Sigma 3$.Thus, the assumption of equality of covariance matrices is not satisfied.

Conclusion: Using the LDA technique for countries in Asia for the period 1991-2013, the study has shown that significant factors responsible for the variation in overall employment sex ratio are GPI for secondary enrollment, employers sex-ratio and percentage of women employed in service sector. In other words, most discriminating factors of variation in overall employment sex ratio are GPI for secondary enrollment, employers sex-ratio and the percentage of women employed in service sector. The overall employment sex ratio is high where the gross parity index at secondary level of education is high across countries. Similarly, overall employment sex ratio is high where the employers sex ratio is high. Similarly, the overall employment sex ratio is high where the percentage share of females in service sector is low. So in order to achieve higher employment sex ratio for countries in first and second rank categories, gender parity index for gross enrollment ratio at secondary education level need to be increased. Also, employers sex ratio need to be increased, i.e. more women need to be encouraged to become employers or entrepreneurs. There is a further scope for increasing the employment of women in service sector.

References:

[1] Australian Bureau of Statistics (2006), Trends in Women's Employment

[2]UNESCAP, online statistics