

## 최근 5년간 국내 연근해에서 발생한 해양사고에 대한 주성분분석

김 영 식†  
(부경대학교)

### Principal Component Analysis on Marine Casualties Occurred at Korean Littoral Sea in Recent 5 Years

Yeong-Sik KIM†  
(Pukyong National University)

#### Abstract

Principal Component Analysis (PCA) is useful statistical technique for finding patterns in data, and expressing the data in such a way as to highlight their similarities and differences. In this paper, 1417 marine casualties occurred in Korean littoral sea in recent 5 years, were examined by the PCA.

The main results obtained were as follows :

1. Most of marine casualties resulted from the human factors such as careless operation and insufficient engine maintenance.
2. Collision and standing mainly resulted from steering room-related human factors such as careless guard, inadequate ship-handling, however engine damage and fire explosion mainly resulted from engine room-related human factor such as bad handling of engine system.
3. No. 1 principal component represents accident frequency, No. 2 principal component represents the cause and No. 3 principal component represents the pattern of marine casualties, respectively.

**Key words :** Principal component analysis (PCA), Marine casualties, Correlation matrix, Proportion, Factor loading

#### I. 서론

우리 나라는 삼면이 바다로 둘러싸여 있어 해상교통에 의한 인명수송과 물류이동은 교통의 중요한 부분을 차지하며, 수출입 물량의 증대와 함께 그 규모도 해마다 증가하고 있다. 그러나 우리 나라 주변해역에서의 해상교통량의 증가는 필연적으로 해양사고의 요인을 증대시키고 있는데, 2014년 4월 진도 인근에서 발생한 세월호 전복사

고는 무려 300명이 넘는 인명 손실을 초래한 금세기 최악의 사고 중 하나로 기록되고 있다. 이러한 해양사고는 그 자체로서의 인명 및 재산피해도 크지만, 유류나 폐기물의 유출 등 해양환경의 보존이라는 측면에서도 많은 문제를 야기시키고 있다. 최근 수산업의 퇴조와 함께 등록선박의 척수가 2000년 이전에 비해 많이 줄어들었으나 해상수송량의 증가, 특히 유조선이나 가스운반선과 같은 특수선박의 운항 증대로 인해 앞으로도

† Corresponding author : 051-629-6155, yeongsik@pknu.ac.kr

※ 이 논문은 부경대학교 자율창의학술연구비(2015년)에 의해 연구되었음.

해양사고 발생의 개연성은 더욱 커질 수밖에 없는 실정이다.

한편 이러한 해양사고에 대하여 해양수산부 및 중앙해양안전심판원에서 매년 기본적인 통계수치가 발표되고 있으나, 해양사고에 영향을 미치는 여러 요인별 정보를 집약하여 이를 체계화시키려는 시도는 오래 전에 이루어진 Yoon(1987), Kim (1990, 1994), 및 Koh(2001)의 연구이외에는 찾아보기 힘들다.

본 연구에서는 최근 5년 동안(2010~2014년)에 우리 나라 주변해역에서 발생하여 중앙해양안전심판원의 재결을 마친 1417건((Marine Accidents Inquiry Agency, Statistics of National Approval, 2015)의 해양사고에 대해 이를 25개의 요인별로 분류하고, SPSS 통계프로그램에 의한 주성분분석(Principal Component Analysis; PCA)을 행하여 이들 각 요인들의 상관성 및 주요 해양사고의 원인을 분석·고찰하였다.

## II. 주성분 분석

주성분분석이란 서로 상관성이 있는 많은 요인들로부터 특징이 되는 본질을 합성하여 그것을 기초로 데이터를 분석하는 통계적 수법이다. 지금 서로 상관성이 있는  $p$ 개의 요인  $x_1, x_2, \dots, x_p$ 에 대해  $N$ 조의 데이터  $(x_{1\lambda}, x_{2\lambda}, \dots, x_{p\lambda}, \lambda = 1, 2, \dots, N)$ 가 얻어진다고 하자. 이들  $N$ 조의 데이터는 각각  $p$ 개의 요인과 상호관련이 있으며, 이것을 설명하는 함수로서  $p$ 개 요인의 1차 결합으로 다음의 식을 가정한다.

$$Z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \quad (1)$$

이 식에서  $l_1, l_2, \dots, l_p$ 를 여러 가지로 바꾸어  $Z$ 의 분산(variance)이 최대가 될 때의  $Z$ 가 제1주성분이다.

단,

$$\sum_{i=1}^p l_i^2 = 1 \quad (2)$$

이때의 계수를  $l_{1i} (i = 1, 2, \dots, p)$ 로 표시하면 제1주성분은 다음 식과 같이 나타낼 수 있다.

$$Z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \quad (3)$$

다음  $Z_1$ 과는 무상관인  $Z$ 중에서 두 번째로 큰 분산을 갖는  $Z$ 가 제2주성분  $Z_2$ 이다. 이와 같이 하여 제 $m$ 주성분  $Z_m$ 은 다음 식과 같이 나타낼 수 있다.

$$Z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \quad (4)$$

제1주성분에서의 계수와 마찬가지로 식 (4)에서의 계수도

$$l_{\alpha 1}^2 + l_{\alpha 2}^2 + \dots + l_{\alpha p}^2 = 1 \quad (\alpha = 1, 2, \dots, m) \quad (5)$$

를 만족해야 한다.

주성분분석의 중요한 목적은  $p$ 개의 요인이 갖는 데이터를 가능한 한 적은 수의 주성분  $Z_1, Z_2, \dots, Z_m (m < p)$ 을 취하여 데이터가 갖는 의미를 요약하여 분석하고자 하는 것이므로 주성분의 수  $m$ 을 결정하기 위해서는 다음 식으로 주어지는 기여율(propotion)을 사용한다.

$$C_\alpha = \frac{V\{Z_\alpha\}}{t_r(\Sigma)} \times 100(\%) \quad (6)$$

여기서  $V\{Z_\alpha\}$ 는 주성분  $Z_\alpha$ 의 분산이고  $t_r(\Sigma)$ 는  $p$ 개 요인  $x_1, x_2, \dots, x_p$ 의 분산의 합이므로 결국 기여율  $C_\alpha$ 란  $p$ 개 요인의 분산의 합에 대한 주성분  $Z$ 의 분산의 비율로 정의할 수 있다.

$m$ 주성분까지 각 기여율의 합을  $m$ 주성분까지의 누적기여율(accumulated propotion)이라 하고 다음의 식으로 주어진다.

$$P_r = \sum_{\alpha=1}^m C_\alpha \quad (7)$$

또 각 주성분  $Z_\alpha$ 와 변량  $x_j$ 의 상관계수를 주성분  $Z$ 의 인자부하량(factor loading)이라 하며 다음의 식으로 주어진다.

$$r_{\alpha j} = \sqrt{\lambda_\alpha l_{\alpha j}} \quad (8)$$

이 인자부하량의 값은 -1과 1 사이에 분포하는데 그 절대치가 클수록 주성분  $Z_\alpha$ 와 변량  $x_j$ 사이에 상관이 크다는 것을 의미한다.

### III. 결과 및 고찰

<Table 1>은 최근 5년 동안(2010~2014년)에 우리 나라 주변해역에서 발생하여 중앙해양안전심판원의 재결을 마친 1417건의 해양사고를 25개 요인으로 분류하여 작성한 상관빈도표(correlation frequency)이고, Table 2는 식(8)에 의해 이를 변환시킨 상관행렬(correlation matrix)이다. 이 통계는 여객선, 화물선, 유조선, 예선 등의 비어선에 대한 것으로서 어선은 제외되어 있다. <Table 1, 2>에서 1~25까지 굵게 표시된 숫자는 해양사고의 요인  $x_i$ 를 표기상의 편의를 위해 사용한 것이며, <Table 1>의 하단에 그 숫자에 대응하는 요인들이 표기되어 있다. <Table 1, 2>에서 공란은 0이다. 해양사고 중 가장 많이 발생하는 사고의 유형은 충돌로서 전체의 65%를 점하고 있는데, 그 원인의 대부분은 경계소홀에서 비롯된 것으로 이들의 상관계수는 0.85로서 매우 크다. 충돌이외에 접촉이나 좌초 등의 사고도 그 원인의 대부분은 경계소홀, 선위확인소홀, 조선부적절 등 기본적인 것을 소홀히 하는 데에서 비롯된 것이고, 전복은 황천에 대한 대응불량과 여객이나 화물의 적재불량이 가장 큰 원인인데 세월호 사고도 바로 이와 같은 원인으로 발생한 것이다. 상기와 같이 25개 요인으로 분류한 데이터를 바탕으로 해양사고에 대한 주성분을 구하고, 각 주성분의 기여율을 계산한 결과를 누적기여율로서 <Table 3>에 나타내었다. 전체 25개의 요인 중 8개의 요

인의 기여율이 전체 기여율의 50% 이상을 차지하고 있다. 주성분을 몇 개까지 취할 것인지 정하는 방법은 몇 가지가 있으나 본 연구에서는 각 주성분의 고유치를 기준으로 그 값이 1.3 이상이 되는 제3주성분까지 취하였다.

<Table 4>는 제1주성분, 제2주성분 및 제3주성분의 인자부하량을 나타낸 것이다. 25개의 요인을 각각  $x_1, x_2, \dots, x_m$ 으로 표시한다면 제1주성분, 제2주성분 및 제3주성분은 각각 다음과 같이 나타낼 수 있다.

$$Z_1 = 0.06x_1 + 0.08x_2 + \dots + 0.16x_{25} \quad (9)$$

$$Z_2 = 0.04x_1 - 0.03x_2 + \dots + 0.23x_{25} \quad (10)$$

$$Z_3 = 0.03x_1 - 0.00x_2 + \dots - 0.10x_{25} \quad (11)$$

[Fig. 1]은 제1주성분과 제2주성분을 축으로 하여 각 인자부하량의 산포도를 나타낸 그래프이다. 상대적으로 그 절대값이 크지 않은 일부 요인들은 표기를 생략하였다. 모든 요인들의 인자부하량은  $Z_1$  축의 플러스(+) 방향에 위치하며, 따라서  $Z_1$  축은 각 요인들의 출현율을 의미한다. 즉, 사고원인으로서는 경계소홀이나 항행법규위반 등의 인적요인이, 사고유형으로서는 충돌, 좌초, 화재폭발 등이 해양사고의 큰 비중을 차지하고 있음을 알 수 있다.

[Fig. 2]는 제2주성분과 제3주성분을 축으로 하여 각 인자부하량의 산포도를 나타낸 그래프이다. 이 그림에서도 일부의 요인들은 표기를 생략하였다. 기관실비취급불량, 기관손상, 선체기관실비결함, 화기취급불량, 전선노후, 합선 등 모든 기관실 관련 해양사고 요인들은  $Z_2$  축의 플러스(+) 방향에, 항행법규소홀, 침로선정유지불량, 경계소홀 등의 조타실 관련 해양사고 요인들은  $Z_2$  축의 마이너스(-) 방향에 자리하고 있으므로  $Z_2$  축은 해양사고의 원인을 나타내는 것으로 해석된다. 또한  $Z_3$  축의 플러스(+) 방향에 인명사상, 좌초, 전복 등의 인자가 자리하고, 마이너스(-) 방향에는 화재폭발, 기관손상 등의 요인들이

<Table 1> Correlation Frequency of 1417 Marine Casualties

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	5																	1	1			3			
2		9															7			2					
3			118														3	20	48	8		6		4	
4				731													701	14	10			1		3	2
5					48												10	2	6	15		13		1	1
6						127											125	1		1					
7							18										8	2	4	2	1				1
8								39									11	9	4	4	1	5		4	1
9									78										1	4	2	2		67	2
10										41										6	21	6		1	6
11											80						3	3			26	2	37		9
12												24									23				1
13													18								8	3	2	1	4
14														10				1		2		4		3	
15															35		14	1	5	3	8	2		1	1
16																36	11		5	3	6	4		5	2
17		7	3	701	10	125	8	11		1	3				14	11	923								
18	1		20	14	2	1	2	9			3		1	1				54							
19	1		48	10	6		4	4	1						5	5			84						
20		2	8		15	1	2	4	4	6			8	2	3	3					58				
21							1	1	2	21	26	23	3		8	6						91			
22	3		6	1	13			5	2	6	2		2	4	2	4							50		
23												37												37	
24			4	3	1		1	4	67	1			1	3	1	5									91
25				2	1			1	2	6	9	1	4		1	2									29

- |  |  |
|--|--|
| 1 Bad preparation for sailing,<br>Insufficient hydrographical survey           | 14 Bad shipping control, Inadequate aid facilities                                       |
| 2 Bad track-keeping  | 15 Irresistible force (weather, etc)   |
| 3 Inadvertent positioning, Inadequate ship-handling                            | 16 Other sailing disturbance   |
| 4 Careless guard   | 17 Collision   |
| 5 Bad preparation for storms and rough seas,<br>Inadequate anchoring & mooring | 18 Contact   |
| 6 Violation of navigation law  | 19 Standing  |
| 7 Inadvertent supervision, Neglect of watch-keeping                            | 20 Capsize   |
| 8 Sailing mistake etc  | 21 Fire explosion  |
| 9 Disobeying the safety instruction  | 22 Sinking   |
| 10 Defect of hull and engine system  | 23 Engine damage   |
| 11 Bad handling of engine system   | 24 Casualty  |
| 12 Bad handling of fire arms, Worn-out wire,<br>Short circuit                  | 25 Etc* (shipping disturbance, propeller damage,<br>marine pollution, facilities damage) |
| 13 Inadequate loading of passenger & cargo                                     |  |

최근 5년간 국내 연근해에서 발생한 해양사고에 대한 주성분분석

<Table 2> Correlation Matrix of 25 factors

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1																	.06	.05		.19				
2		1															.08		.09						
3			1														.01	.25	.48	.10		.08		.04	
4				1													.85	.07	.04		.01		.01	.01	
5					1												.05	.04	.09	.28		.27		.02	.03
6						1											.37	.01	.01						
7							1										.06	.06	0.1	.06	.02			.02	
8								1									.06	0.2	.07	.08	.02	.11		.07	.03
9									1									.01	.06	.02	.03			.80	.04
10										1							.01		.12	.34	.13		.02	.17	
11											1						.01	.05		.30	.03	.68		.19	
12												1									.49			.04	
13													1							.25	.07	.07		.02	.18
14														1				.04	.08	.18		.10			
15															1		.08	.02	.09	.07	.14	.05	.02	.03	
16																1	.06	.09	.07	.10	.09		.09	.06	
17		.08	.01	.85	.05	.37	.06	.06		.01	.01			.08	.06	1									
18	.06		.25	.07	.04	.01	.06	.20		.05			.04	.02			1								
19	.05		.48	.04	.09		.10	.07	.01					.09	.09			1							
20		.09	.10		.28	.01	.06	.08	.06	.12			.25	.08	.07	.07				1					
21						.02	.02	.02	.34	.30	.49	.07	.14	.10							1				
22	.19		.08	.01	.27		.11	.03	.13	.03	.07	.18	.05	.09								1			
23											.68												1		
24			.04	.01	.02		.02	.07	.80	.02		.02	.10	.02	.09										1
25				.01	.03		.03	.04	.17	.19	.04	.18	.03	.06											1

<Table 3> Eigenvalue and Accumulated Proportion of 25 Principal Component

Principal Component No.	Eigenvalue	Accumulated Proportion (%)	Principal Component No.	Eigenvalue	Accumulated Proportion (%)
No. 1	1.41	0.08	No. 14	1.00	0.75
No. 2	1.38	0.15	No. 15	1.00	0.79
No. 3	1.35	0.23	No. 16	1.00	0.83
No. 4	1.29	0.29	No. 17	0.94	0.87
No. 5	1.24	0.35	No. 18	0.92	0.90
No. 6	1.19	0.41	No. 19	0.87	0.93
No. 7	1.11	0.46	No. 20	0.76	0.95
No. 8	1.07	0.51	No. 21	0.69	0.97
No. 9	1.06	0.55	No. 22	0.58	0.99
No. 10	1.00	0.59	No. 23	0.43	0.99
No. 11	1.00	0.63	No. 24	0.33	1.00
No. 12	1.00	0.67	No. 25	0.13	1.00
No. 13	1.00	0.71			

<Table 4> 25 Factor Loadings

Factor	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>
1 Bad preparation for sailing, Insufficient hydrographical survey	0.06	0.04	0.03
2 Bad track-keeping	0.08	-0.03	0.00
3 Inadvertent positioning, Inadequate ship-handling	0.20	0.09	0.13
4 Careless guard	0.72	-0.51	-0.08
5 Bad preparation for storms and rough seas, Inadequate anchoring & mooring	0.18	0.10	0.08
6 Violation of navigation law	0.30	-0.22	-0.04
7 Inadvertent supervision, Neglect of watch-keeping	0.11	0.01	0.03
8 Sailing mistake etc	0.16	0.06	0.09
9 Disobeying the safety instruction	0.25	0.36	0.73
10 Defect of hull and engine system	0.19	0.29	-0.12
11 Bad handling of engine system	0.27	0.52	-0.50
12 Bad handling of fire arms, Worn-out wire, Short circuit	0.15	0.28	-0.21
13 Inadequate loading of passenger & cargo	0.12	0.16	0.01
14 Bad shipping control, Inadequate aid facilities	0.09	0.09	0.12
15 Irresistible force (weather, etc)	0.16	0.08	-0.03
16 Other sailing disturbance	0.16	0.11	0.05
17 Collision	0.79	-0.55	-0.09
18 Contact	0.18	0.04	0.04
19 Standing	0.20	0.07	0.11
20 Capsize	0.20	0.17	0.11
21 Fire explosion	0.29	0.49	-0.34
22 Sinking	0.19	0.17	0.08
23 Engine damage	0.19	0.40	-0.42
24 Casualty	0.27	0.36	0.75
25 Etc* (shipping disturbance, propeller damage, marine pollution, facilities damage)	0.16	0.23	-0.10

자리하고 있으므로 Z<sub>3</sub> 축은 해양사고의 유형을 나타내는 것으로 해석된다. 주성분분석의 경우 많은 요인들이 갖는 데이터를 몇 개의 주성분으로 함축하여 설명할 수 있는가하는 것이 중요한데, 본 연구의 결과 해양사고의 원인과 유형을 설명하기 위해서는 제3주성분까지 취하면 충분한 것으로 판단된다. 다만, 제1주성분이 명확히 해양사고의 출현율을 나타내고 있는데 반해 제2주성분은 주로 해양사고의 원인을 나타내지만 그 원인으로 인한 해양사고의 유형이 일부 포함되어 있고, 제3주성분은 주로 해양사고의 유형을 나타내지만 그 유형의 해양사고를 초래하는 원인이 혼재되어 있다.

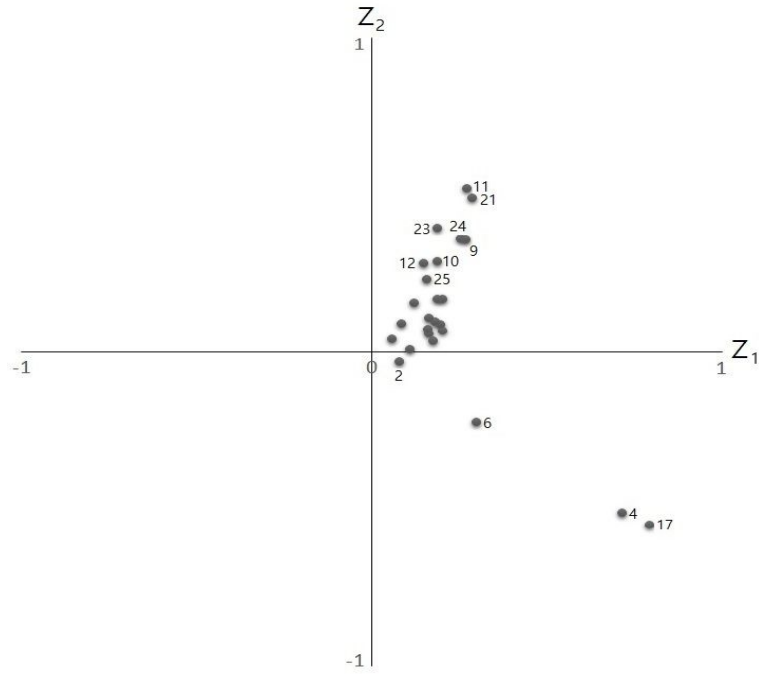
[Fig. 1, 2]를 종합적으로 살펴보면, 대부분의

해양사고는 기관설비취급불량, 선체기관설비결함, 화기취급불량, 전선노후합선, 항행법규소홀, 침로 선정유지불량, 경계소홀 등 기관실 및 조타실 관련 인적요인에 의해 발생한다. 또한 기관실 관련 인적요인 에 의해 발생하는 사고유형은 주로 기관손상이나 화재폭발 등이며, 조타실 관련 인적요인에 의해 발생하는 사고는 충돌과 좌초 등이 큰 비중을 차지하고 있음을 알 수 있다.

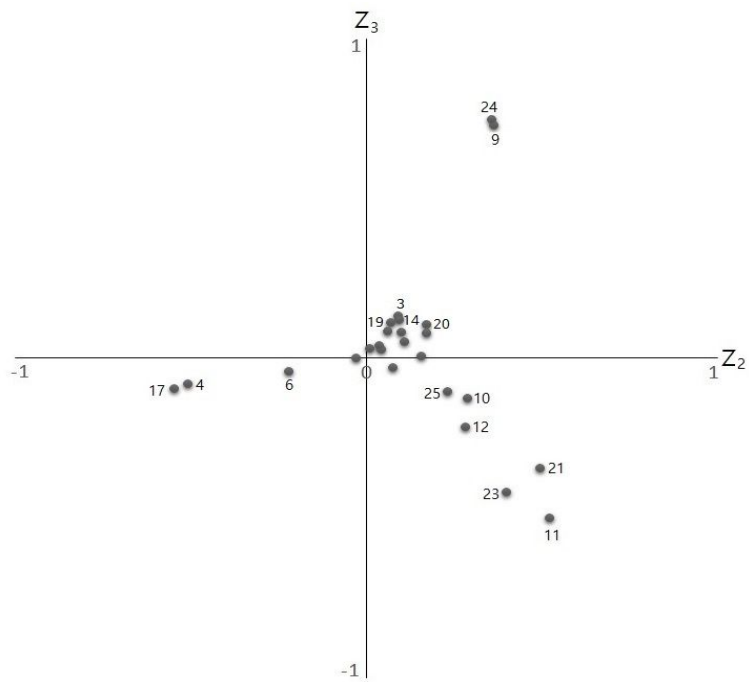
#### IV. 요약

본 연구에서는 2010년부터 2014년까지 최근 5년간 우리 나라 주변해역에서 발생하여 중앙해양안전심판원의 재결을 마친 1417건의 해양사고에

최근 5년간 국내 연근해에서 발생한 해양사고에 대한 주성분분석



[Fig. 1] Dispersion of 25 Factors as to  $Z_1$  and  $Z_2$  Axes



[Fig. 2] Dispersion of 25 Factors as to  $Z_2$  and  $Z_3$  Axes

대해 이를 25개 요인별로 분류하고, SPSS 통계 프로그램에 의한 주성분분석(Principal Component Analysis; PCA)을 행하여 이들 각 요인들의 상관성 및 주요 해양원인을 분석·고찰하였다. 얻어진 주요한 결과들을 요약하면 다음과 같다.

1. 해양사고의 주된 원인은 기관설비취급불량, 화기취급불량, 항행법규소홀, 침로선정유지불량, 경계소홀 등 기관실 및 조타실 관련 인적요인에 의해 발생한다.
2. 조타실 관련 인적요인에 의해 발생하는 사고는 충돌과 좌초 등이 큰 비중을 차지하며, 기관실 관련 인적요인에 의해 발생하는 사고유형은 주로 기관손상이나 화재폭발 등이다.
3. 주성분분석의 결과 제1주성분은 해양사고의 출현율을, 제2주성분은 해양사고의 원인을, 제3주성분은 해양사고의 유형을 나타낸다.

### References

Alvin, C. Rencher & William, F. Christensen(2012). *Methods of Multivariate Analysis*, 3rd Ed., WILEY.

Jeremy J Foster(2001). *Data Analysis Using SPSS for Windows Versions 8 - 10*, SAGE.

Kim, Su-Beom(2001). A Study on the reason and style of marine accident occurred at adjacent seas of the Korea, Jeju Nat. Univ. Master's Thesis.

Kim, Yeong-Sik & Kim, Jeong-Chang(1994). Multivariate data analysis on marine casualties, *Journal of Fishier and Marine Educational Research* 6(2), 190~197.

Kim, Yeong-Sik · Yoon, Suck-Hun & Koh, Dae-Kwon (1990). Principal component analysis on marine casualties, *Bull. Korean Fish. Soc.* 26(3), 303~307.

Koh, Dae-Kwon et al.(2001) A study on the analysis of engine accidents on fishing vessels, *Journal of the Korean Society for Power System Engineering* 5(3), 25~30.

Korea Maritime Safety Tribunal(2015). *Statistics, Marine accidents according to vessel type*.

Parinya, S.(2012). *Principal Component Analysis-Multidisciplinary Applications*, InTech.

Robin, Beaumont(2012). *An introduction to Principal Component Analysis & Factor Analysis Using SPSS 19 and R*.

Yoon, Suck-Hun(1987). Principal component analysis of marine engine troubles, *Bull. Korean Fish. Soc.* 23(1), 25~29.

- 
- Received : 09 November, 2015
  - Revised : 25 November, 2015
  - Accepted : 03 December, 2015