

# 확률분포에 의한 리스크 빈도수와 손실규모 추정 프로세스 연구

이 영 재\* · 이 성 일\*\*

## A Research on Process of Estimation about Frequency and Loss of Risk by distribution of Probability

Youngjai, Lee\* · Seongil, Lee\*\*

### Abstract

Risk that breed large size disaster is happening variously for cause at social, natural a management. Incidence and damage scale are trend that increase rapidly than past. In these circumstance, to keep operational continuity of organization, area, society, risk management action that establish systematic counter measure estimating and analyze occurrence possibility and expectation damage of risk is essential indispensable issue and the best countermeasure. Risk management action does by main purpose establish optimum disaster reduction countermeasure.

To deduce various countermeasure, process that estimate and analyze occurrence possibility and expectation damage of risk is essential indispensable issue. Therefore, this paper studies process design that can presume risk occurrence frequency and damage scale through distribution of probability.

Keywords : Risk, Risk Management, Countermeasure, Frequency, Loss, Probability

## 1. 서론

지구 온난화로 재해가 대형화되는 추세이며 특히 현대인의 생활권 정보화, 도시화와 고령화 등 사회 구조 변화로 취약성이 노출되어 재난에 의한 대규모 손실은 지역사회의 생존을 좌우하는 수준에 이르렀다. 대형 재해를 야기하는 리스크는 사회적(예, 1997년 한국의 IMF 사태), 자연적(예, 미얀마의 사이클론, 중국 스촨성에서 발생한 지진), 경영적(예, IT 버블 붕괴에 따른 벤처기업 도산) 원인으로 다양하게 나타나고 있으며 과거에 비하여 발생 빈도와 손실 규모가 급증하는 추세이다. 이러한 상황에서 조직(기업, 지역사회 등)의 운영 연속성을 유지하기 위해 리스크의 발생 가능성과 예상 손실을 예측·분석하여 체계적인 대응책을 수립하는 리스크 관리 활동은 필수 불가결한 사항이며 최선의 대책이다[1].

국내의 경우 IMF 이후 금융권을 중심으로 체계적인 리스크 관리 방법론을 도입하고 있다. 선행 연구에 따르면 금융 분야에서 운영리스크의 빈도수와 손실규모를 예측하는데 있어 포아송분포와 로그정규분포가 적합한 것으로 기술되어 있으며 고장 빈도수 혹은 장애 발생 빈도수 등에 대한 확률 분포는 포아송 분포를 따르는 것으로 제시되어 있다[2, 3].

정보기술 분야에서는 정보시스템의 이중화 같은 초보적인 수준의 리스크 관리 체계를 도입하고 있다. 그러나 현재 도입 중인 리스크 관리 방법론은 과거의 리스크 발생 빈도수와 손실 규모 같은 실증 데이터에 기반한 리스크 분석 결과를 활용하는 것이 아니라 조직의 자산 가치, 취약점 등 현재의 상태 데이터를 활용하는 방법이므로 다음과 같은 현실적인 문제점이 존재한다. 첫째, 분석 결과의 객관성이나 분석 방법의 과학성이 상대적으로 미흡하다. 현행 방법론은

빈도수나 손실 규모를 과거의 실증 데이터로 직접 분석하는 것이 아니라 자산의 가치를 산정하여 손실에 따른 대체 비용을 다시 계산하는 방식이므로 조직에서 자산의 유·무형 가치를 어떻게 판단하느냐에 따라서 상당히 다른 결과가 산출될 수 있다. 둘째, 조직의 규모, 사업 전략 등은 조직 고유의 특성에 매우 종속적인 데이터이므로 공통적인 방법론에 적용하여 리스크를 산출하기에는 한계가 있다.

리스크 관리 활동은 최적의 재난경감대책을 수립하는 것이 핵심이다. 다양한 대책을 도출하기 위해서 리스크의 발생 가능성과 예상 손실을 예측·분석하는 프로세스가 필수 불가결한 사항이다. 따라서 조사된 리스크분석 문제점을 근거로 설계된 연구 질문은 다음과 같다. 첫째, 리스크분석과 관련된 연구에서 기술하고 있는 리스크의 빈도수와 손실규모를 예측하는 데 있어 포아송분포와 로그정규분포가 다른 재난 유형에도 적용 가능성이 있는가? 둘째, 리스크는 빈도수와 손실규모로 표현되는데, 이를 계량화할 수 있는 프로세스를 기술할 수 있는가? 셋째, 확률분포에 의해 산출된 결과가 리스크 의사결정을 지원할 수 있는 정보로서 유용한가?

그러므로 본 논문은 확률 분포를 활용하여 리스크 발생 빈도수와 손실 규모를 추정할 수 있는 프로세스 설계를 연구한다. 연구 목표를 달성하기 위해 본 논문에서는 소방방재청에서 제공하는 우심 피해사례와 인천국제공항공사에서 제공하는 결항 피해사례의 통계 분포 통해 미래의 손실 규모를 예측하는 사례를 들어 리스크 의사결정을 지원하는 프로세스를 설계하였다.

본 논문의 구성은 다음과 같다. 선행연구에서는 본 논문의 주요 실험 대상인 포아송 분포, 로그정규 분포, 시뮬레이션, 적합도 검정 방법에 대한 기존 연구의 결과를 소개하였다. 연구의 설계에서는 연구의 방법, 가설을 정의하고 실험

에 활용될 리스크의 빈도수와 손실 규모에 대한 실증데이터를 제시하였다. 이후의 시뮬레이션의 실행에서 구체적인 실험 결과를 정리하였고 마지막으로 결론을 통해 실험 결과 시사점, 연구의 한계, 향후 연구의 방향 등을 정리하였다.

## 2. 선행연구

### 2.1 리스크 분석 방법

리스크는 위협 요소의 발생 빈도와 해당 리스크로 인해 나타날 수 있는 손실 규모의 조합을 통해 표현할 수 있다[4]. 리스크 평가 결과를 통해 조직은 특정 리스크가 발생할 수 있는 확률과 발생할 경우 조직에 나타날 수 있는 손실 규모를 파악할 수 있게 되므로 해당 리스크를 관리하기 위한 대책을 수립할 수 있는 것이다. 발생할 수 있는 확률을 산정하기 위해서는 평가 대상 리스크의 과거 발생 빈도를 조사해야 하고 손실 규모를 예측하기 위해서는 과거의 손실 규모를 파악 해야 한다. 과거의 데이터를 조사, 분석하는 것이 미래의 리스크를 예측하기 위한 원천 데이터가 되는 것이다. 과거의 데이터를 조사, 분석하는 가장 기본적인 접근방법은 통계적 분포를 확인하는 것이다.

### 2.2 리스크 확률 분포

안종찬(2003)은 선행 연구에서 금융기관의 위험 탐지 모형을 제시하였는데 제시된 위험 탐지 모형에 따르면 발생 빈도수는 포아송 분포를 손실 규모는 로그정규 분포를 따르는 것으로 나타나 있다[2].

김영곤(2003)의 연구결과 또한 금융기관의 정보시스템 고장 빈도수 혹은 IT 장애 발생 빈도수 등에 대한 확률분포는 포아송 분포를 나타내

고 있음을 제시하고 있다[3].

황명수(2005)는 바젤II의 운영리스크 산정 방법 중에 하나인 금융기관의 운영 리스크 분석 기법인 손실분포법(LDA : Loss Distribution Approach)을 이용하여 손실규모를 추정 하였다. 손실분포법에서는 손실사건 빈도는 포아송분포를, 손실규모는 로그정규 분포를, 추정모델은 몬테카를로 시뮬레이션을 이용 하였다[5].

포아송 분포를 리스크 발생 빈도수 분석에 적용한 근거는 같다.

- 개념 : 주어진 시간, 생산량 또는 길이 등과 같은 단위 구간에서 어떤 특정 사건이 발생하는 수에 관련된 분포
- 적용 분야 : 연속(Continuous) 시간축 상에서 임의(Random)로 발생하는 이산(Discrete) 사건을 묘사할 때 활용(예, 안전사고의 발생, 전자 부품의 고장, 고객의 도착 등)

로그정규분포를 리스크 손실 규모 추정에 적용한 근거는 다음과 같다.

- 개념 : 어떠한 확률 변수  $Y = \ln X$ 가 정규분포를 나타낼 때,  $X$ 가 로그정규분포를 나타내며,  $Y$ 가 갖는 범위가 넓은 경우, 로그 함수를 이용하여 이를 유도하는 분포
- 적용 분야 : 로그정규분포는 손해액분포의 특성을 가지고 있으며, 대규모 손해의 발생 가능성이 있는 경우에 유효함

이전 언급된 포아송 분포와 로그정규 분포의 적용 근거를 보면 연구자들의 연구 결과가 타당함을 일정 부분 확인할 수는 있지만 기존 연구는 공통적으로 금융 기관에 한정적인 모형을 제시하고 있으므로 해당 연구결과가 다른 리스크 유형과 다른 유형의 조직에 적용 가능한지는 실험한 연구가 부재한 상태이다.

## 2.3 몬테카를로 시뮬레이션

황명수(2005)의 연구에서 금융기관의 운영 리스크를 분석할 때 몬테카를로 시뮬레이션 기법은 특정 영업 부문의 운영리스크 기대손실(Expected Loss : EL)에 대한 정보가 없어도 계산이 가능하지만, 사고 발생 가능성이 낮으나 손실강도가 막대한 경우를 반영하지 못할 가능성도 있는 것으로 언급되어 있다[5].

보험개발원(2003)의 연구에 따르면 몬테카를로 시뮬레이션 기법은 불확실성을 갖는 시스템 문제의 각 구성요소들의 변수에 대해 확률적 특성을 고려할 수 있는 알고리즘이다. 통계적인 데이터로부터 정량적으로 위험도를 추정하기 어려운 경우 관련 과거 자료, 전문가의 경험과 판단 같은 정성적인 자료를 바탕으로 추정된 통계적 불확실량과 확률적 위험도를 제한적인 현장조사, 시험결과를 바탕으로 갱신하고 추정하는 기법이다[6].

따라서 통계데이터가 유의성을 확인할 수 없을 정도로 불충분하거나 해당되는 확률 분포를 따르지 않을 경우 몬테카를로 시뮬레이션을 실행하여 기초 데이터를 재 생성하여 연구자료로 이용할 수 있는데 도움이 되는 기법이다.

## 2.4 적합도 검정

적합도 검정(Goodness of fit test)을 통해 어떤 모집단을 검정하려면 해당 모집단의 확률분포를 알아야 하며 정규분포가 아닌 경우에는 비모수 통계방법을 이용하여 모집단이 어떤 분포를 갖고 있는가를 검정해야 한다. 비모수 통계방법을 이용한 검정은  $\chi^2$  검정과 Kolmogorov-Smirnov 검정(이하 K-S 검정) 방법이 대표적이다. 다음과 같이 데이터의 유형에 따라 검정 방법을 선택해야 한다.

- 명목 척도 : 데이터가 분류만의 의미를 갖고 있는 경우(남자와 여자 등) →  $\chi^2$  검정
- 순서 척도 : 순위를 통해 비교할 수 있는 데이터(프로야구팀의 순위 등) → K-S 검정
- 등간 척도 : 일정한 간격으로 배열이 가능한 데이터(연간/월간 온도 등) → K-S 검정
- 비율 척도 : 데이터 간 비율을 계산할 수 있는 경우(체중, 키 등) → K-S 검정

재난과 관련해서 리스크 발생 빈도와 손실 규모는 등간 척도에 해당되므로 K-S 검정 방법을 선택하여 적합도 검정을 하는 것이 타당하다.

## 3. 연구의 설계

### 3.1 연구 방법의 설계

본 논문에서는 빈도수와 손실규모에 관한 확률분포를 검정하기 위하여 필요 가설을 수립하였으며, 빈도는 포아송 분포를, 손실규모는 로그정규 분포를 따르는지 유의수준 95% 범위 내에서 검정하였다. 수립된 가설을 기반으로 리스크 관련 기초 데이터를 적용하여 통계 검정하였고, 통계 데이터가 불충분하거나 해당되는 확률분포를 따르지 않는 경우 시뮬레이션을 통해 기초데이터를 재생성하여 검정작업을 실시하였다. 조직이 보유한 기초 데이터가 검정하고자 하는 확률분포를 따른다면 다양한 기법으로 향후 발생할 리스크의 빈도와 손실액을 추정할 수 있다. 그러나 시뮬레이션을 통해서 기초 데이터를 생성한 경우에는 발생빈도의 확률을 산정할 수 있지만 시간 경과에 따른 발생빈도 추이를 검정하는 데는 한계가 있으므로 유의수준에서 발생 빈도와 손실액을 추정하였다.

본 연구의 기존 데이터는 산업과 업종에 구별 없이 유효한 데이터가 제공 가능한 조직에서 확

보되었다. 데이터의 유효성이 미흡한 경우에는 통계검정 과정에서 확보한 모수( $\lambda$ )를 기초로 시뮬레이션을 실시하여 확률분포검정과 추정을 실시하였다.

### 3.2 연구 가설 설정

#### (1) 빈도수 유효성 검정을 위한 가설

포아송 분포는 연속(continuous) 시간축 상에서 임의로 발생하는 이산(discrete) 사건을 묘사하는데 사용되므로 리스크 발생의 빈도가 포아송 분포를 따르는지를 검정해야 한다[5].

#### 포아송 분포의 공식

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2,$$

- $\lambda$ 는 평균사건발생건수
- $x$ 는 사건 발생건수
- $e$ 는 자연로그의 밑수

가설 1 : 리스크발생 빈도수는 포아송 분포를 형성한다.

(귀무가설)  $H_0$  : 리스크발생 빈도수가 포아송 분포를 나타낸다.

(대립가설)  $H_1$  : 리스크발생 빈도수가 포아송 분포를 나타내지 않는다.

#### (2) 손실액 유효성 검정을 위한 가설

대규모 손실이 예상되는 경우 발생 손실액을 보다 용이하게 추정할 수 있는 확률분포로는 로그정규분포가 대표적이므로 손실액이 로그정규 분포를 따르는지 검정해야 한다[5].

#### 로그정규 분포의 공식

$$f(x) = \frac{1}{x \sqrt{2\pi\sigma^2}} \theta x p \left( -\frac{(\ln x - \mu)^2}{2\sigma^2} \right), \quad x \geq 0$$

$x$ 는 사건발생금액

$\mu$ 는 평균

$\sigma^2$ 는 표준편차

가설 2 : 리스크발생으로 인한 손실액은 로그정규 분포를 형성한다.

(귀무가설)  $H_0$  : 리스크발생으로 인한 추정 손실액은 로그정규 분포를 나타낸다.

(대립가설)  $H_1$  : 리스크발생으로 인한 추정 손실액은 로그정규 분포를 나타내지 않는다.

### 3.3 연구 대상기관 선정 및 연구대상 설계

기초 데이터를 수집해야 하는 연구대상 조직을 선정함에 있어 첫째, 사회적 파급효과가 높은 리스크에 대한 기초 데이터를 공개할 수 있는지 여부와 둘째, 신뢰할 수 있는 데이터를 제공할 수 있는 조직인지를 고려하였다. 이러한 고려사항을 기초로 소방방재청의 우심피해사례 및 인천국제공항공사의 결항피해사례를 본 논문의 연구대상으로 선정하였다.

#### (1) 우심피해사례

소방방재청의 경우 매년 소방, 방재 크게 두 가지 측면에서 국가차원의 재난현황을 종합 통계치 형태의 연감으로 발간하고 있다[7]. 우심지역이란 매년 풍수해에 따라 일정규모 이상의 피해를 입은 시군구를 칭하는 용어이며 통상적으로 매년 풍수해에 따라 100~130여개 지역이 선정되고 있다. 본 사례에서 분석할 재난 유형은 호우이며 피해액 산출범위는 국가 차원의 재난피해액을 기준으로 한다.

#### (2) 결항피해사례

인천국제공항공사는 장애 현황을 항공기 지

연과 결항으로 구분하여 2001년부터 월별로 인터넷을 통해 기상, 장치 불량, 여객처리, 복합원인 등 유형별로 공개하고 있다[8].

〈표 1〉 최근 10년간 연도별 우심피해 발생률(1995년~2004년)  
(단위 : 천원)

년 도	발생회수	피해액	평균피해액
1995	108	574,270,847	5,317,323
1996	36	465,873,197	12,940,922
1997	69	152,509,044	2,210,276
1998	177	1,516,133,495	8,565,726
1999	162	1,178,854,678	7,276,881
2000	133	605,325,387	4,551,319
2001	162	1,195,999,356	7,382,712
2002	238	6,073,745,368	25,519,939
2003	136	4,370,987,940	32,139,617
2004	133	1,190,238,820	8,949,164
합계	1,354	17,323,938,132	
평균	135.4	1,732,393,813	

일반적인 재해 유형의 구분 방법에 따라 기상으로 인한 지연 및 결항을 “자연적 재해”, 장치

〈표 2〉 출발기준의 운항횟수 및 결항 통계

2005년			2006년		
구분 (월별)	운항 (회)	결항 (회)	구분 (월별)	운항 (회)	결항 (회)
1월	10,003	7	1월	10,750	17
2월	9,036	6	2월	9,755	11
3월	9,864	7	3월	10,666	9
4월	10,017	9	4월	10,685	9
5월	10,500	4	5월	10,984	13
6월	10,229	12	6월	10,746	9
7월	10,811	16	7월	11,448	22
8월	10,749	14	8월	12,039	23
9월	10,253	38	9월	11,578	22
10월	10,578	18	10월	12,498	13
11월	10,153	8	11월	13,230	17
12월	10,372	29	12월	13,893	54
합계	122,565	168	합계	112,193	219

불량 및 여객처리는 “인위적 재해”, 복합원인 및 기타는 “사회적 재해”로 구분할 수 있는데 장치불량이 주요 원인인 결항피해사례는 인위적인 재해에 포함되는 것으로 볼 수 있다. 인천국제공항공사의 항공기 지연 및 결항원인을 살펴본 결과, 장치불량(A/C 접속, A/C 정비)이 전체 장애 원인 중에 50% 이상을 차지하므로 본 연구의 실험 대상으로 결정하였다. 2005년과 2006년에 발생한 장치불량에 의한 결항 건수는 <표 2>와 같다.

장애에 따른 손실금액은 공항공사의 경영실적 자료를 통해 유추할 수 있다. 공항공사의 매출액은 출발과 이에 따르는 운항관리수수료에 의해 산정되므로 <표 3>의 경영실적 자료에서 매 1회 운항 당 매출액을 추정할 수 있고, 그 결과 1회 결항 당 평균 710만원의 매출 손실을 가져온 것으로 볼 수 있다.

〈표 3〉 공항 공사의 경영 실적

(단위 : 백만원)

구 분	2005년	2006년
매출액	875,256	792,549
매출총이익	467,569	391,890
영업이익	406,406	343,792
경상이익	202,013	175,551
당기순이익	145,128	123,893
운항횟수(회)	122,565	112,193

### 3.4 기초 데이터에 대한 통계검정

<표 1>~<표 3>의 기초 데이터를 토대로 가설 1을 실험(Kolmogorow-Smirnov 검정 활용, 이하 K-S 검정)한 결과, 포아송 분포를 따르지 않음을 확인하였다. 따라서, 본 논문 “연구 방법의 설계”에서 언급한 바와 같이 몬테카를로 시뮬레이션을 통해 기초데이터를 재생성하여 검정작업을 실시하였다. 통계검정 과정에서 산출된 모수  $\lambda$ 를 활용해서 빈도수에 대한 포아송 분



포를 만들었고 난수를 발생시켜 발생 빈도수를 계산하였으며 이를 토대로 로그정규 분포를 따르는 손실액을 산출하였다.

## 4. 시뮬레이션과 실행

### 4.1 우심피해사례

<표 1>에서 발생 회수의 총합은 1,354회이고, 피해액은 17,323,938,132(원)이므로 평균 발생회수는 135.4이고 평균 피해액은 1,732,393,813(원)이다. 여기서 발생 1회 당 피해금액은 12,794,637(원)이다. 이러한 결과를 포아송 분포 공식에 입력하기 위해 포아송 분포 공식에 발생건수를 0건에서부터 구하고자 하는 회수만큼 입력하고 평균 발생 회수는 <표 1>을 통해 산정한 135.4회를 적용한다. 이렇게 공식에 적용한 결과는 <표 4>와 같다.

<표 4>의 결과를 통해 난수구간을 설정할 수 있는데, 난수구간은 누적 포아송 분포를 통해 산출할 수 있다(<표 5> 참조). 손실액은 발생회수와 이전 산출된 발생 1회당 피해금액을 곱하여 산정한다.

<표 4> 우심 피해 사례의 포아송 분포 계산결과

발생회수	포아송분포	누적포아송분포
0	0	0
...	...	...
100	0.000244563	0.000876149
105	0.00096051	0.003916711
110	0.002976203	0.01403674
115	0.007354172	0.040915385
120	0.014634336	0.09843668
125	0.023663792	0.198504065
130	0.031351647	0.341175778
...	...	...
190	1.65E-06	0.99999613
200	4.20E-08	0.999999916
210	6.48E-10	0.999999999

발생 회수를 95에서 180까지로 결정하고 난수표를 통해 각각 난수를 발생시킨 후 해당 구간의 피해 발생회수와 함께 발생회수 당 피해액을 함께 입력한다(<표 6> 참조).

손실금액은 로그정규분포 공식에 대입하기 위해 자연로그(ln) 값으로 변환하여 사건 발생 수별로 함께 기록한다(<표 6> 참조).

<표 5> 우심 피해 사례의 난수 구간표

발생 회수	포아송 분포	포아송분포(누적)	난수 구간	손실액(단위 : 원)
95	4.86E-05	0.000154836	0~2	1,215,490,515
100	0.000244563	0.000876149	3~9	1,279,463,700
...	...	...	...	...
120	0.014634336	0.09843668	410~984	1,535,356,440
125	0.023663792	0.198504065	985~1985	1,599,329,625
130	0.031351647	0.341175778	1986~3412	1,663,302,810
135	0.034294008	0.50914566	3413~5091	1,727,275,995
140	0.031191063	0.673582335	5092~6736	1,791,249,180
145	0.023743761	0.808307943	6737~8083	1,855,222,365
150	0.015220657	0.901248192	8084~9012	1,919,195,550
...	...	...	...	...
175	0.000150957	0.999529923	9983~9995	2,239,061,475
180	3.85E-05	0.999892043	9996~9999	2,303,034,660

〈표 6〉 우심 피해 사례의 난수 발생과 시뮬레이션

난 수	발생회수	손실금액 (단위:원)	로그지수 (손실금액)
6139	140	1791249180	21.30617908
6262	140	1791249180	21.30617908
3848	135	1727275995	21.26981144
871	120	1535356440	21.1520284
8418	150	1919195550	21.37517195
2398	130	1663302810	21.23207111
1959	125	1599329625	21.19285039
...	...	...	...
648	120	1535356440	21.1520284
6981	145	1855222365	21.3412704
8853	150	1919195550	21.37517195
8853	150	1919195550	21.37517195
합계	6830	87387370710	1063.878266

〈표 7〉 우심 피해사례의 로그정규분포값

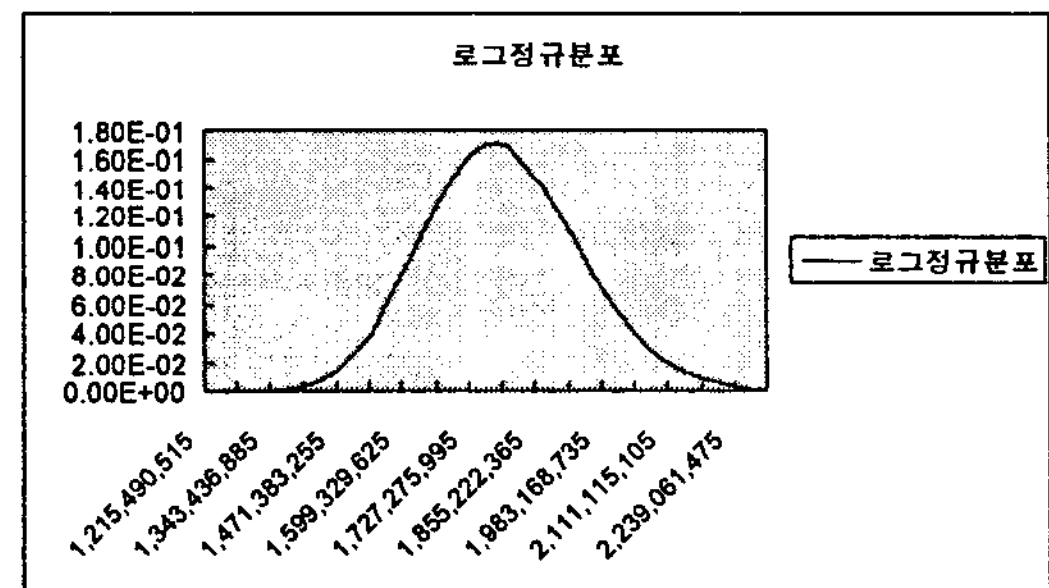
발생회수	손실금액(단위 : 원)	로그정규분포
95	1,215,490,515	6.64045E-06
100	1,279,463,700	8.25538E-05
105	1,343,436,885	0.000673425
110	1,407,410,070	0.003649846
115	1,471,383,255	0.013798391
120	1,535,356,440	0.037929216
...	...	...
150	1,919,195,550	0.11100502
155	1,983,168,735	0.071233869
160	2,047,141,920	0.040012987
165	2,111,115,105	0.019923425
170	2,175,088,290	0.008893255
175	2,239,061,475	0.00359467
180	2,303,034,660	0.00132762

시뮬레이션을 총 100회 실시한 분석결과는 다음과 같다.

- 총 발생횟수 : 13770회  
평균발생횟수 : 137.7회
- 총 피해액 : 176,182,151,490원  
평균피해액 : 1,761,821,514원
- 로그지수 총합 : 2128.61
- 로그지수 평균 : 21.2861
- 로그지수 표준편차 : 0.084

포아송 분포는 도수분포를 근거로 하고 로그 정규 분포는 연속형 변수를 기반으로 하기 때문에 로그 정규 분포를 통한 손실액 분석이 필요하다. 로그정규분포 공식에서  $x$ 는 실제 사건발생 금액 또는 손실금액을 의미한다(〈표 6〉 참조). 로그정규분포의 공식에 따라 분석한 결과는 〈표 7〉과 같다.

〈표 7〉의 로그정규분포값을 그래프로 표현하면 다음의 〈그림 1〉과 같으며 그래프의 형태가 정규분포를 따르고 있음을 확인할 수 있다.



〈그림 1〉 우심 피해사례의 로그정규분포 그래프

〈표 7〉에 의하여 우심피해사례가 로그정규 분포를 따르고 있음이 확인되었으므로 〈표 6〉을 통해 분석한 로그지수 평균과 로그지수 표준편차, 실제 손실금액을 95% 신뢰구간에 적용하여 비예상피해액 2,017,122,590(원)을 산정하였다.

#### 4.2 결항피해사례

〈표 2〉의 기초 데이터에서 발생회수는 24개월 동안 387회(평균 발생회수 : 16.12회)이고 손실액은 2,747,700(천원)(평균 손실액 : 7,100천원)



이다. 이러한 결과를 포아송 분포 공식에 입력한 결과는 <표 8>과 같다.

<표 8> 결항피해사례의 포아송 분포 계산결과

장애 발생수 x	포아송 분포	누적 포아송 분포
...	...	...
4	0.000278712	0.00036238
5	0.000899126	0.00126150
6	0.002417151	0.00367865
7	0.005569806	0.00924846
8	0.011230121	0.02047858
9	0.020126872	0.04060545
10	0.032464644	0.07307010
...	...	...
21	0.044341238	0.90512195
22	0.032510189	0.93763214
23	0.022799537	0.96043167
24	0.015323189	0.97575486
25	0.009886521	0.98564138
26	0.006133446	0.99177483
27	0.003664166	0.99543900
28	0.002110821	0.99754982
29	0.001174053	0.99872387
30	0.000631249	0.99935512
31	0.000328453	0.99968357

여기서 유효 구간은 발생회수가 4에서 30까지로 결정할 수 있다. 이를 토대로 난수표를 통해 각각 난수를 발생시키고 각 난수 값 해당 구간의 피해발생회수를 입력하여 발생회수 당 피해액을 함께 입력한 결과는 <표 9>와 같다.

장애회수가 4회인 경우부터 30회인 경우로 난수를 입력하고 난수표를 통해 난수를 발생시켜 해당되는 장애회수별 빈도수와 이에 따르는

손실액을 입력한 결과는 <표 10>과 같다.

시뮬레이션을 100회 실행한 분석결과는 다음과 같다.

- 총 발생회수 : 1,610 회
- 평균 발생 회수 : 16.10회
- 총 피해액 : 11,431,000원
- 평균 피해액 : 7,100,000원
- 로그지수 총합 : 1,161.0080
- 로그지수 평균 : 11.6100803
- 로그지수 표준편차 : 0.2761

<표 9> 결항 피해 사례의 난수 구간표

발생수	누적 포아송 분포	난수 구간	손실금액(천원)
4	0.0003624	0 · 4	28,400
...	...	...	...
9	0.0406055	206 ~ 407	63,900
10	0.0730701	407 ~ 731	71,000
11	0.1206751	732 ~ 1207	78,100
12	0.1846641	1208 ~ 1847	85,200
13	0.2640597	1848 ~ 2641	92,300
14	0.3555348	2642 ~ 3556	99,400
15	0.453901	3556 ~ 4540	106,500
16	0.5530664	4540 ~ 5531	113,600
17	0.6471569	5531 ~ 6472	120,700
18	0.7314724	6473 ~ 7315	127,800
19	0.8030519	7316 ~ 8031	134,900
20	0.8607807	8032 ~ 8608	142,000
21	0.9051219	8609 ~ 9052	149,100
22	0.9376321	9052 ~ 9377	156,200
23	0.9604317	9377 ~ 9605	163,300
...	...	...	...
30	0.9993551	9989 ~ 9999	213,000

이전 분석 방법과 동일하게 로그정규분포를 통해 손실액을 분석한 결과는 <표 11>과 같다.

〈표 10〉 결항 피해 사례의 난수 발생과 시뮬레이션 시행

난 수	장애발생수	손실금액(천원)	로그지수(손실금액)	난 수	장애발생수	손실금액(천원)	로그지수(손실금액)
8059	20	142,000	11.86358	4290	15	106,500	11.57590
4106	15	106,500	11.57590	3869	14	99,400	11.50691
8060	20	142,000	11.86358	6174	17	120,700	11.70106
5092	16	113,600	11.64044	9211	22	156,200	11.95889
6329	17	120,700	11.70106	5576	17	120,700	11.70106
2216	13	92,300	11.43280	6472	17	120,700	11.70106
8998	21	149,100	11.91237	1223	12	85,200	11.35276
8674	21	149,100	11.91237	2627	13	92,300	11.43280
1835	12	85,200	11.35276	9591	23	163,300	12.00334
360	9	63,900	11.06507	6633	18	127,800	11.75822
449	10	71,000	11.17044	7783	19	134,900	11.81229
3930	15	106,500	11.57590	5190	16	113,600	11.64044
4809	16	113,600	11.64044	9189	22	156,200	11.95889
9832	23	163,300	12.00334	1506	12	85,200	11.35276
1889	13	92,300	11.43280	2778	13	92,300	11.43280
7777	19	134,900	11.81229	2636	13	92,300	11.43280
...	...	...	...	...	...	...	...
2425	13	92,300	11.43280	1842	12	85,200	11.35276
4480	15	106,500	11.57590	415	10	71,000	11.17044
합계	797	5,658,700	579.95002	합계	813	5,772,300	581.0580

〈표 11〉 결항 피해 사례의 로그정규분포

장애 발생수	손실금액	로그정규 분포	누적로그 정규분포
7	49,700	0.001965682	0.001965682
8	56,800	0.006230148	0.008195830
9	63,900	0.016019910	0.024215740
10	71,000	0.031472844	0.055688584
...	...	...	...
18	127,800	0.075076852	0.704175135
19	134,900	0.063811973	0.767987107
20	142,000	0.052697210	0.820684318
21	149,100	0.042481563	0.863165880
22	156,200	0.033559528	0.896725408
23	163,300	0.026062799	0.922788208
24	170,400	0.019951387	0.942739594
25	177,500	0.015088387	0.957827981
26	184,600	0.011294047	0.969122028

## 5. 통계분포의 검정(K-S 검정)

몬테카를로 시뮬레이션을 이용하여 예상 피해액을 산정하였지만 발생 회수가 포아송 분포를 따르고, 피해액이 로그정규 분포를 따른다고 가정하였으므로 K-S 검정을 사용하여 가정에 대해 검정하려 한다.

가설 1 : 리스크발생빈도수가  $\alpha = 0.05$ 에서 포아송 분포를 형성한다.

(귀무가설)  $H_0$  : 리스크발생빈도수가  $\alpha = 0.05$ 인 포아송 분포를 나타낸다.

(대립가설)  $H_1$  : 리스크발생빈도수가  $\alpha = 0.05$ 인 포아송 분포를 나타내지 않는다.

가설 2 : 리스크발생으로 인한 손실액이  $\alpha = 0.05$ 에서 로그정규 분포를 형성한다.  
 (귀무가설)  $H_0$  : 리스크발생으로 인한 추정 손실액은 로그정규 분포를 나타낸다.  
 (대립가설)  $H_1$  : 리스크발생으로 인한 추정 손실액은 로그정규 분포를 나타내지 않는다.

통계 분포를 검정하기 위해 <표 4>의 누적 포아송분포값과의 차이의 절대값을 산정하였다 (<표 12> 참조).

<표 12>에서 “차이”의 최대값인 0.09110이 검정통계량이 되는데 K-S 검정표에서  $n = 125$ 에 해당하는  $D_{0.05}$ 는  $(\frac{1.36}{\sqrt{125}} = 0.122)$ 이므로  $D(= 0.0911) < D_{0.05}(= 0.122)$ 이다. 따라서 귀무가설을 기각할 수 없다. 다시 말해서 발생회수의 분포가 포아송 분포를 따른다.

5.1 우심피해사례

K-S 적합도 검정 방법으로 우심피해 사례의

손실액은 연속확률 변수이고 K-S 검정은 이

<표 12> 우심피해 사례의 포아송 분포 적합도 검정

발생회수	빈도	상대빈도	누적상대 빈도( $S_i$ )	포아송 분포값(누적)( $F_i$ )	차이 = ( $S_i - F_i$ )의 절대값
115	2	0.0200	0.02	0.0266917329	0.00669173
120	6	0.0600	0.08	0.0689779046	0.01102209
125	16	0.1600	0.24	0.1488967409	0.09110325
130	12	0.1200	0.36	0.2726904418	0.08730955
135	15	0.1500	0.51	0.4310391040	0.07896089
140	15	0.1500	0.66	0.5994704635	0.06052953
145	10	0.1000	0.76	0.7494165579	0.01058344
150	13	0.1300	0.89	0.8618173224	0.02818267
155	6	0.0600	0.95	0.9331634451	0.01683655
160	4	0.0400	0.99	0.9717137134	0.01828628
165	1	0.0100	1.00	0.9895329001	0.01046709

<표 13> 우심피해 사례의 로그정규 분포 적합도 검정

발생회수	손실금액	누적빈도( $S_i$ )	누적로그 정규분포( $L_i$ )	차이 = 절대값( $S_i - L_i$ )
115	1,471,383,255	0.0200	0.018210857	0.0018
120	1,535,356,440	0.0800	0.056140073	0.0239
125	1,599,329,625	0.2400	0.134698834	0.1053
130	1,663,302,810	0.3600	0.261137398	0.0989
135	1,727,275,995	0.5100	0.423568954	0.0864
140	1,791,249,180	0.6600	0.594074571	0.0659
145	1,855,222,365	0.7600	0.743357169	0.0166
150	1,919,195,550	0.8900	0.85436219	0.0356
155	1,983,168,735	0.9500	0.925596059	0.0244
160	2,047,141,920	0.9900	0.965609046	0.0244
165	2,111,115,105	1.0000	0.985532471	0.0145

산확률변수뿐 아니라 연속확률 변수에서도 검정이 가능하므로 손실액이  $\alpha = 0.05$ 에서 로그정규 분포를 따르는지에 대한 적합도 검정을 위해 <표 7>의 결과값을 토대로 누적로그정규 분포값을 산정하여 차이 분석표를 작성하였다(<표 13> 참조).

검정 통계량  $D$ 는 “차이” 열에서 가장 큰 값인 0.1053이 된다. K-S 검정표에서  $D_{0.05} = 0.122$ 이므로  $D(= 0.1053) < D_{0.05}(= 0.122)$ 이다. 따라서 귀무가설을 기각할 수 없으므로 우심피해액이 로그정규 분포를 따른다.

## 5.2 결항피해사례

이전 사례와 동일한 방법으로 결항피해 사례에 대한 적합도를 검정하였다. 장애발생빈도가

포아송 분포를 따르는지  $\alpha = 0.05$ 에서 검정하기 위해 <표 8>의 누적 포아송 분포값과의 차이의 절대값을 산정하였다(<표 14> 참조).

<표 14>에서 마지막 열인 “차이”의 최대값 0.055194가 검정통계량이 되는데 K-S 검정표에서  $n = 19$ 에 해당되는  $D_{0.05}$ 는 0.311이므로  $D(= 0.055194) < D_{0.05}(= 0.311)$ 이다. 따라서 귀무가설을 기각할 수 없다. 다시 말해서 발생회수의 분포는 포아송 분포를 따른다.

또한 손실액은 연속확률 변수이므로 이전 사례와 동일한 방법으로 검정을 실시하기 위해 <표 11>의 누적 로그정규분포값을 토대로 차이의 절대값을 산정하였다(<표 15> 참조).

검정 통계량  $D$ 는 “차이” 열에서 최대값인 0.05623이다. K-S 검정표에서  $n = 16$ 일 때의

<표 14> 결항피해 사례의 포아송 분포 적합도 검정

장 애	빈 도	누적상대빈도(Si)	포아송분포값( $\lambda = 16.10$ )	포아송 누적분포(Fi)	차이 = 절대값(Fi-Si)
7	1	0.01	0.0056651	0.009417	0.000583
8	2	0.03	0.0114011	0.020818	0.009182
9	1	0.04	0.0203952	0.0412133	0.001213
10	4	0.08	0.0328363	0.0740496	0.005950
11	5	0.13	0.0480604	0.12211	0.007890
12	6	0.19	0.0644811	0.1865911	0.003409
13	12	0.31	0.0798574	0.2664485	0.043552
14	8	0.39	0.091836	0.3582845	0.031716
15	10	0.49	0.0985706	0.456855	0.033145
16	11	0.6	0.0991867	0.5560417	0.043958
17	8	0.68	0.0939356	0.6499773	0.030023
18	4	0.72	0.0840202	0.7339975	0.013997
19	3	0.75	0.071196	0.8051935	0.055194
20	7	0.82	0.0573128	0.8625063	0.042506
21	5	0.87	0.0439398	0.9064462	0.036446
22	3	0.9	0.032156	0.9386021	0.038602
23	3	0.93	0.0225092	0.9611113	0.031111
24	4	0.97	0.0150999	0.9762112	0.006211
25	2	0.99	0.0097243	0.9859355	0.004064
26	1	1	0.0060216	0.9919572	0.008043

〈표 15〉 결항피해 사례의 로그정규 분포 적합도 검정

장애 발생수	손실금액	누적빈도(Si)	누적 로그정규분포(Li)	차이 = 절대값(Li-Si)
7	49,700	0.01	0.0019657	0.008034319
8	56,800	0.03	0.0081958	0.021804172
9	63,900	0.04	0.0242157	0.015784266
10	71,000	0.08	0.0556886	0.024311429
11	78,100	0.13	0.1062174	0.023782641
12	85,200	0.19	0.1757161	0.014283872
13	92,300	0.31	0.2604477	0.049552324
14	99,400	0.39	0.3543481	0.035651886
15	106,500	0.49	0.4507478	0.039252168
16	113,600	0.60	0.5437689	0.056231127
...	...	...	...	...
23	163,300	0.93	0.9227882	0.007211823
24	170,400	0.97	0.9427396	0.027260431
25	177,500	0.99	0.957828	0.032172039
26	184,600	1.00	0.969122	0.030877988

$D_{0.05} = 0.34$ 이므로  $D(= 0.05823) < D_{0.05}(= 0.34)$ 이다. 따라서 귀무가설을 기각할 수 없으므로 결항손실액이 로그정규 분포를 따른다.

## 6. 추정손실액 계산

### 6.1 우심 피해사례

이전 검정결과를 기반으로 차기년도의 우심 발생주기와 피해액을 예측하여 본다. 예측을 위해 몬테카를로 시뮬레이션을 통해 산출된 우심 평균 발생주기( $\lambda$ ) 값 137.7를 엑셀의 포아송분포 함수에 적용하면 다음 <표 16>과 같이 각 회당 확률과 그 회수 이하의 확률을 추정할 수 있다. 상기 표의 결과를 분석하여 볼 때 차기년도의 재난발생회수는 119회 이상이고 155회 이하일 가능성이 95%라고 할 수 있다. 아울러 <표 16>을 참조하면 135회의 발생 가능성이 제일 높다는 것을 나타내고 있다. 이때 이로 인한 손실규모는 약 17억 2천만원 정도를 추정할 수 있다.

또한, 로그정규 분포를 활용하여 손실금액을 추정하기 위해 1회당 평균 손실액 12,794,637원과 각 발생회수를 곱하여 손실금액을 산정하고 상기 결과에 대한 로그지수평균(21.2860459)과 로그표준편차(0.0844168)를 엑셀의 로그정규분포 함수에 입력하면 아래와 같이 산출된다. 이러한 결과를 신뢰구간 95%와 99%로 구분하여 적용하면 피해액은 아래와 같다.

- 95% 신뢰구간 : 2,017,122,584(원)
- 99% 신뢰구간 : 2,136,569,694(원)

따라서 비예상 피해액은 95% 신뢰구간에서 2,017,122,584(원)이라고 할 수 있다.

### 6.2 결항피해사례

이전 검정된 결과를 근거로 2007년도의 발생주기와 피해액을 예측하기 위해 다른 사례와 동일하게 엑셀의 'GROWTH(포아송 분포값, 장애 발생수, 장애발생수 : 포아송 분포값, TRUE)' 함수를 이용하였다[12].

〈표 16〉 차년도 재난발생횟수 및 손실금액 추정

(단위 : 원)

회 수	P(X = 회수)	P(X ≤ 회수)	손실금액	로그정규분포
90	0.0000	0.0000	1,151,517,330	0.0000
95	0.0000	0.0001	1,215,490,515	0.0000
100	0.0001	0.0005	1,279,463,700	0.0001
105	0.0006	0.0022	1,343,436,885	0.0008
110	0.0019	0.0085	1,407,410,070	0.0044
115	0.0051	0.0267	1,471,383,255	0.0182
120	0.0111	0.0690	1,535,356,440	0.0561
125	0.0195	0.1489	1,599,329,625	0.1347
130	0.0281	0.2727	1,663,302,810	0.2611
135	<b>0.0334</b>	<b>0.4310</b>	<b>1,727,275,995</b>	<b>0.4236</b>
140	0.0331	0.5995	1,791,249,180	0.5941
145	0.0274	0.7494	1,855,222,365	0.7434
150	0.0191	0.8618	1,919,195,550	0.8544
155	0.0113	0.9332	1,983,168,735	0.9256
160	0.0057	0.9717	2,047,141,920	0.9656
165	0.0024	0.9895	2,111,115,105	0.9855
170	0.0009	0.9966	2,175,088,290	0.9944
175	0.0003	0.9990	2,239,061,475	0.9980
180	0.0001	0.9998	2,303,034,660	0.9993

〈표 17〉의 2007년도 추정손실계산을 살펴보면 결항피해사례 중 발생확률이 가장 큰 값은

'0.099186700'이며, 이때 장애 발생건수는 16건이며 손실금액은 약 1억 1천 3백 만원 정도로 예측된다.

〈표 17〉 결항피해사례 손실 추정

(손실금액 단위 : 천원)

장애발생수	손실금액	포아송 분포	예측값
7	49700	0.005665100	0.041516678
8	56800	0.011401100	0.040806004
9	63900	0.020395200	0.040107495
10	71000	0.032836300	0.039420943
11	78100	0.048060400	0.038746143
12	85200	0.064481100	0.038082894
13	92300	0.079857400	0.037430998
14	99400	0.091836000	0.036790262
15	106500	0.098570600	0.036160493
16	113600	<b>0.099186700</b>	<b>0.035541505</b>
17	120700	0.093935600	0.034933112
18	127800	0.084020200	0.034335134
19	134900	0.071196000	0.033747392
20	142000	0.057312800	0.033169711
21	149100	0.043939800	0.032601918
22	156200	0.032156000	0.032043845
23	163300	0.022509200	0.031495325
24	170400	0.015099900	0.030956194
25	177500	0.009724300	0.030426292
26	184600	0.006021600	0.029905460



## 7. 연구 결론

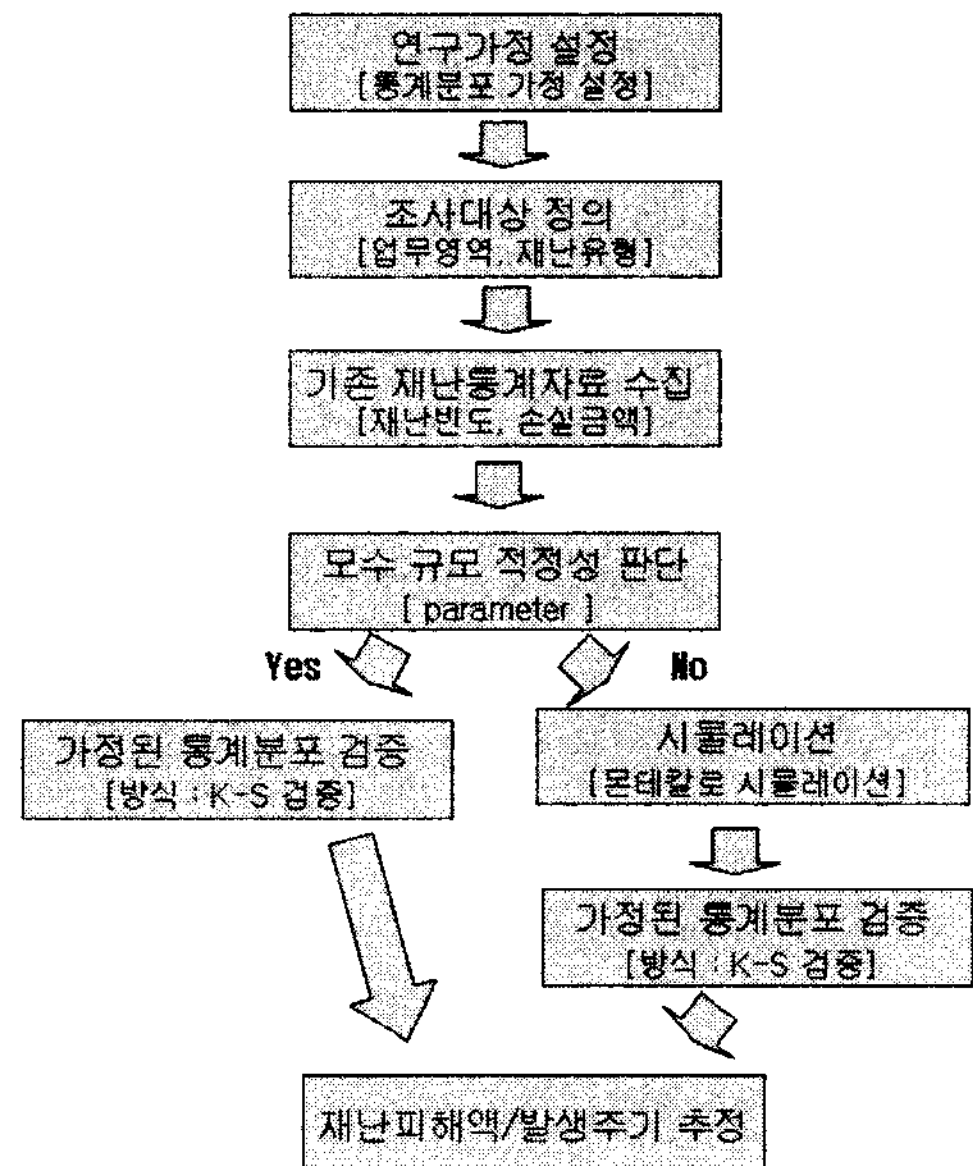
피해발생 건수와 손실액의 확률분포를 통해 예상되는 손실을 추정하는 본 연구는 과거의 실증 데이터에 근거하여 계량적으로 리스크를 평가할 수 있는 방법을 제시하고 있다. 그러나 기초 자료가 불충분할 경우 향후 발생할 확률을 추정하는 데 한계가 있어 몬테카를로 시뮬레이션을 활용한다. 그러나 초기 데이터에서 충분한 피해발생 건수가 조사되어 포아송 분포를 이루었다면 회귀분석 등을 통해 보다 정확한 추정을 할 수 있었을 것이다.

연구의 결론은 연구 질문에 대한 답을 함으로써 기술한다.

첫째, 리스크분석과 관련된 연구에서 기술하고 있는 리스크의 빈도수와 손실규모를 예측하는데 있어 포아송 분포와 로그정규 분포가 자연재해 그리고 인적재난 유형에도 적용 가능성이 있다는 사실이 연구를 통해서 증명되었다. 그러므로 리스크가 발생할 빈도수 추정에 관한 확률분포는 포아송 분포를 활용하기를 추천하며 리스크로 인한 손실액 추정 확률분포는 로그정규 분포를 이용할 것을 추천한다. 리스크에 대한 피해 규모를 산출하는 관련 연구가 지속적으로 진행되고 있으나, 리스크의 특성상 통계적 유의성을 증명할 만한 표본의 수를 확보하는 것은 매우 어렵기 때문에 피해액 산출에 가장 큰 제약사항이 된다. 본 논문에서는 이러한 제약사항을 극복하기 위한 대안으로 몬테카를로 시뮬레이션을 활용하는 방법을 제시하였으며 향후 제약사항을 극복할 수 있는 보다 다양한 방법들이 연구되어야 할 것이다.

둘째, 리스크는 빈도수와 손실규모로 표현되는데, 각 변수를 확률분포에 의하여 계량화할 수 있는 프로세스를 다음과 같이 기술한다. 기술된 프로세스에 따라 대부분의 재난유형에서

리스크 발생 횟수와 리스크 발생에 따른 손실 금액을 계량적으로 추정할 수 있다.



셋째, 확률분포에 의해 산출된 재난 발생 횟수와 손실 금액 추정은 리스크 관리 즉 재난경감 대책을 수립하는 데 유용한 정보로 활용될 것이다. 추정된 손실 금액이 조직의 현금흐름(cash flow)을 초과한다면 조직의 운영 연속성에 심대한 영향을 주기 때문에 조직은 리스크 발생에 대비할 수 있는 대안을 구현할 것이다. 이러한 수치적인 데이터는 리스크 의사결정에서 의사결정자의 생산성을 높여주게 된다.

## 참고 문헌

- [1] 이영재, *위기관리*, 생능출판사, 2006.
- [2] 안종찬, “금융기관의 위험탐지모형 신뢰성 검증 및 활용방안”, 한양대학교, 2003.
- [3] 김영곤, “국내 금융기관의 운영위험에 관한 연구: 전산위험을 중심으로”, 건국대학교, 2003.

- [4] Micheal E. Whitman and Herbert J. Mattord, "Management of Information Security", TH-OMSON, 2004, pp. 308-312.
- [5] 황명수, 이영재, "시뮬레이션 기법을 이용한 운영리스크 평가", 한국SI학회지, 제4권 제1호, 2005.
- [6] 몬테카를로 기법을 이용한 손해액 추정의 손해액분포추정, 보험개발연구, 제11권 제3호, 2003.
- [7] 천일영, 주민정, 신동현, "보험회사의 경영리스크 관리방안", 보험개발원 보험연수원, 2003.
- [8] 소방방재청, 재해연보 2004, 2005.
- [9] 인천국제공항공사홈페이지(www.airport.or.kr) 내의 항공통계, 2007.
- [10] Erwin Kreyszig, Introductory Mathematical Statistics(Principles and Methods), John Wiley and Sons, Inc., New York, 1970.
- [11] 박태선, 김광목, 윤양수, 이승복, "홍수피해 특성 분석 및 홍수피해지표 개발에 관한 연구", 국토연구원, 2005.
- [12] 김양렬, 민충기, 경영경제통계학(excel을 이용한 분석), 2005.

#### ■ 저자소개



#### 이영재

Florida Institute of Technology, Computer Science, Master of Science, George Washington Univ. Administrative Science, Master of Arts.

George Washington Univ. Information Mgmt. Doctor of Science 학위를 취득하고, 현재 동국대학교 경영정보학과 교수로 재직 중이다. 관심분야는 재난관리, 리스크 분석 및 평가, 의사결정지원시스템 등이다.



#### 이성일

중앙대학교 정보시스템학과에서 석사학위를 취득하고, 2000년부터 6년간 퓨처시스템, 이니텍 등 정보통신부 지정 정보보호컨설팅전문업체 재직 후,

현재 동국대학교 경영정보학과 박사과정에 재학 중이다. 관심분야는 IT 정보보안, BCP, 리스크 분석 및 평가 등이다.