

# R을 활용한 이윤 최적화 시뮬레이션 모델링 및 결과 분석

조민호\* · 전용호\*\*

## Simulation Modeling of Profit Optimization and Output Analysis using R

Min-Ho Cho\* · Yong-Ho Jeon\*\*

### 요 약

시뮬레이션은 현대의 복잡한 환경에서 효율적인 의사 결정을 위한 도구로서 폭넓게 사용되고 있다. 하지만, 대부분은 결과의 분석 보다는 시뮬레이션 모델의 개발 및 수행에 중점을 두고 있다. 본 논문에서는 시뮬레이션 모델링은 물론 결과의 분석도 중요하고 체계적으로 진행되어야 한다는 점을 강조하고, 이를 위하여 통계 분석과 다양한 데이터 조작 및 그래픽 기능을 가진 R을 사용하여 이윤 최적화 시뮬레이션에 대한 모델링과 결과 데이터의 조작 및 다양한 그래픽을 사용한 분석 작업을 수행하여 유용성을 입증하였다.

### ABSTRACT

Simulation is now using in various area as an effective decision analysis tool in complex environment of today. But, There is a focus to the simulation model development and execution better than result analysis. This article will emphasis to the importance of result analysis apart from model development in simulation, and will use R package for profit optimization simulation. R has a various function in statistic analysis and data manipulation, graphic display. So this research can show the value of R as a tool for simulation.

### 키워드

Simulation, Statistic Analysis, Model, R package, Result Analysis  
시뮬레이션, 통계분석, 모델, 알 패키지, 결과분석

## 1. 서 론

시뮬레이션은 해석적인 처리가 곤란하거나, 실행 불가능한 경우, 위험 부담이 많거나 비용이 많이 드는 경우 그리고 복잡해서 처리가 곤란한 경우에 적용되는 기술이며[1], 새로운 시스템을 구축하기 전에 타당성 조사 및 대안의 비교 평가가 가능하여 투자에 따

른 위험을 줄일 수 있는 도구로 사용되고 있다[2].

최근에는 시스템이 대형화되고 복잡해져서 수리적 기법으로는 분석이 더욱 어려워지고, 동시에 시뮬레이션을 위한 소프트웨어가 지속적으로 개발됨에 따라 시스템의 분석을 위한 도구로서 시뮬레이션의 이용은 급속히 증가하고 있으며[3], Decomposition에 기반을 둔 분석 기법이 최적화를 위한 방안으로 도입되어 활

\* 중원대학교 컴퓨터시스템공학과 교수(chominhokr@jwu.ac.kr)

\*\* 교신저자(corresponding author) : 중원대학교 메카트로닉스학과 교수(waterjliar@jwu.ac.kr)

접수일자 : 2014. 06. 23

심사(수정)일자 : 2014. 07. 21

게재확정일자 : 2014. 08. 11

용되고 있다[4-6].

하지만, 앞에서 제시한 대부분의 연구는 모델의 구성 부분에 치우쳐서 수행되어 왔고, 결과 분석이나 요인 간의 상관관계 등 시뮬레이션을 구성하는 다른 부분에 대한 연구는 가치에 비해서 상대적으로 중요하게 다루어 지지 않았다.

이해를 위하여 일반적으로 수행되는 시뮬레이션의 단계를 정리하면 아래와 같다[7].

#### [시뮬레이션의 단계].

- 문제의 정의
- 시뮬레이션의 사용 여부 결정
- 자료의 수집 및 조사
- 모형의 작성
- 모형의 타당성 검토
- 시뮬레이션 수행
- 시뮬레이션 결과의 분석
- 판단 자료의 작성

본 논문은 시뮬레이션 단계를 충실히 수행한다는 측면에서 기존에 사용하던 도구들이 모델 구성에 중점을 두고 있다는 점을 지적하고자 한다.

시뮬레이션은 모델의 구성도 중요하지만, 결과의 분석을 통한 올바른 판단을 유도하는 것도 중요하다.

이러한 점에 착안하여 데이터의 분석 및 조작에 우수한 기능을 가지고 있는 R을 사용하여 시뮬레이션을 수행하고자 한다[8-9]. R은 통계 분석 기능을 중심으로 구성된 제품이지만, 시뮬레이션의 수행을 위해 필요한 프로그램 기능을 가지고 있다. 특히, 생성된 결과의 다양한 분석을 위한 데이터 조작, 그래픽 표현, 통계 분석 기능이 우수하다.

그러므로, 시뮬레이션의 전 과정을 R로 수행할 수 있다면, 다른 제품 보다 다양한 결과 분석을 통해서 올바른 결론을 유도하는데 많은 장점을 가질 수 있을 것으로 생각한다. 본 논문에서는 신문 판매 시뮬레이션을 통하여 R을 활용한 구현 사례를 제시하고 모델링과 결과분석을 수행할 것이다.

## II. 신문 판매 시뮬레이션 모델의 소개

시뮬레이션을 수행할 때, R 패키지의 적용이 가능 여부를 판단하기 위하여 다양한 시뮬레이션 모델 중에서 기본적인 환경을 두루 갖추고 있는 신문 판매 시뮬레이션을 개발한다

신문 판매 시뮬레이션 모델은 지하철역에서 신문을 파는 경우에 관련된 모델이다. 상세한 내용은 다음과 같다.

#### [신문 판매 시뮬레이션 모델]

- 매일 몇 부의 신문을 구매해서 판매하면 이익을 최대로 얻을 수 있는지에 대한 모델이다. 당일 몇 부의 신문이 판매될 지에 대해서는 과거 1년간의 통계가 있다.
- 신문의 판매 가격은 130원이다
- 신문의 원가는 개당 80원이다
- 팔리지 않은 신문은 반품 할 수 있는데, 이때 가격은 개당 60원이다
- 초기 모델의 단순화를 위하여 판매되는 신문은 1종류만 취급하는 것으로 가정 한다
- 과거 1년간 판매 부수에 대한 확률은 다음과 같다  
10부 - 0.08, 20부 - 0.14, 30부 - 0.17,  
40부 - 0.36, 50부 - 0.46, 60부 - 0.66,  
70부 - 0.77, 80부 - 0.81, 90부 - 0.95,  
100부 - 1.00

판매부수에 대한 확률의 의미는 매일 10부가 팔릴 확률은 0.08, 20부가 팔릴 확률은 0.06(0.14-0.08)와 같이 계산된다. 즉, 100일 동안 팔았다면 10부를 팔 확률이 5일이고 20부를 팔 확률이 7일이라는 의미이다.

신문 판매 시뮬레이션 모델을 통하여 상인은 매일 몇 부의 신문을 진열하였을 때, 최대의 이익을 얻게 되는 지에 대한 것을 알고자 한다

## III. R을 활용한 신문판매 시뮬레이션 모델의 개발

신문 판매 시뮬레이션의 수행을 위하여 실제 수행되는 시뮬레이션의 흐름을 정리하면 그림 1과 같다.

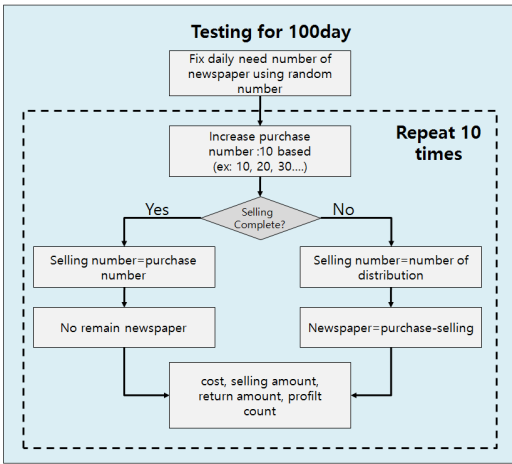


그림 1. 신문 판매 시뮬레이션 상세도  
Fig. 1 Detailed flowchart of simulation for the newspaper sale

그림 1의 상세 내용에 따라 실제 수행되는 프로그램을 제작하면 다음과 같다. 제공되는 소스는 R의 프로그래밍 기능으로 제작된 것으로서 프로그램 자체의 내용 설명은 프로그램의 주석 기능을 통하여 상세하게 설명하였다.

[R로 제작한 신문 판매 시뮬레이션 프로그램]

```

=====
# 내용 : 신문 판매 시뮬레이션
# 2014. 7
# 작성자 : 조민호
=====

#최종 보고할 내용 자료를 선언하는 부분
busu<-c(10,20,30,40,50,60,70,80,90,100) #부수범주
ic<-c(0,0,0,0,0,0,0,0,0,0) #원가
is<-c(0,0,0,0,0,0,0,0,0,0) #총판매금액
ir<-c(0,0,0,0,0,0,0,0,0,0) #이익

#남은 부수 : 보고서에 포함되지 않음. 중간계산
#을 위한 부분
iz<-c(0,0,0,0,0,0,0,0,0,0)

# 100일 동안 시뮬레이션을 수행
  
```

```

for( day in 1:100) {

# 난수 발생-특정일에 몇 부를 팔았는지 표현
k1<-runif(1)*100

# 발생한 난수를 확률변수에 따라 분류
if(k1<=8){ x<-1}
else if(k1>8 && k1<=14){ x<-2}
else if(k1>14 && k1<=17) {x<-3}
else if(k1>17 && k1<=36) {x<-4}
else if(k1>36 && k1<=46) {x<-5}
else if(k1>46 && k1<=66) {x<-6}
else if(k1>66 && k1<=77) {x<-7}
else if(k1>77 && k1<=81) {x<-8}
else if(k1>81 && k1<=95) {x<-9}
else {x<-10}

# 특정일의 수요부수를 결정
suyo<-x*10

# 수요에 따라 신문을 10~100까지 10가지
# 경우로 구비한 경우에 대한 계산 부분
# busu1 : 그날 비치한 신문의 숫자
# ic : 원가, is : 총판매금액, ir : 이익,
# iz : 남은 부수

for(i in 1:10){
#부수설정
busu1<-i*10

#부수에 따른 중간 계산
if(busu1<=suyo){
ud<-busu1
zd<-0
}
else {
ud<-suyo
zd<-busu1-suyo
}

# 최종결과의 계산
  
```

```

cd<-busu1*80
sd<-ud*130
hd<-zd*60
rd<-sd+hd-cd

# 결과를 보관하는 부분
ic[i]<-ic[i]+cd
is[i]<-is[i]+sd
iz[i]<-iz[i]+zd
ir[i]<-ir[i]+rd
}
}

```

# 시뮬레이션 결과를 데이터 프레임 형태로 보여주기 위해서 정리

```
mydata<-data.frame(busu,ic,is,ir)
```

# 시뮬레이션 결과를 데이터 프레임 형태로 화면에 보여줌

```
mydata
```

# 시뮬레이션 결과를 Plot 도표로 보여줌

```
plot(mydata)
```

# 시뮬레이션 결과를 Boxplot 도표로 보여줌

```
boxplot(mydata$ic, mydata$is, mydata$ir)
```

```

=====
# 완료 및 시뮬레이션 결과의 저장
=====

```

제공된 R 프로그램 소스를 실제 수행한 결과는 그림 2와 같다.

제공된 결과는 시뮬레이션 된 것을 실제 분석을 위한 형태인 데이터 프레임 형태로 만드는 과정과 실제 생성된 데이터를 화면에 보여주는 과정을 나타낸다.

일반 시뮬레이션 패키지의 경우에는 이 단계에서 데이터를 보여주는 것으로 실행 과정을 마치는 경우가 대부분이고[7], 필요한 경우에는 Excel이나 다른 제품과 연결하여 그래프로 표현하고자 하는 경우가 있다.

```
>mydata<-data.frame(busu,ic,is,ir)
```

```
>mydata
```

	Newspaper #	purchase money	selling money	profit amount
1	10	80000	130000	50000
2	20	160000	253500	96500
3	30	240000	369200	138800
4	40	320000	481000	179000
5	50	400000	565500	204500
6	60	480000	631800	220200
7	70	560000	677300	224700
8	80	640000	711100	222900
9	90	720000	739700	218300
10	100	800000	746200	201800

그림 2. 시뮬레이션 수행 결과  
Fig. 2 Result of simulation

R을 활용하게 되면, 기본적으로 데이터 프레임 형태의 데이터를 생성할 수 있고, 그 외에 필요한 경우에는 다양한 데이터 형으로 변형할 수 있다. R이 생성할 수 있는 데이터의 형태는 벡터, 행렬, 데이터프레임, 배열, 요인, 시계열, 리스트로 다양하며 각 데이터 타입마다 분석을 위한 도구가 준비되어 있다[9].

그림 2에서 제시된 시뮬레이션 결과를 살펴보면, 100일간 시뮬레이션을 수행했을 때, 원가와 판매금액, 이익간의 상호 관계에 대한 것이 쉽게 파악되지 않는다. 이러한 상황을 좀 더 쉽게 분석하기 위하여 R에서 제공하는 그래픽 기능을 활용할 수 있다.

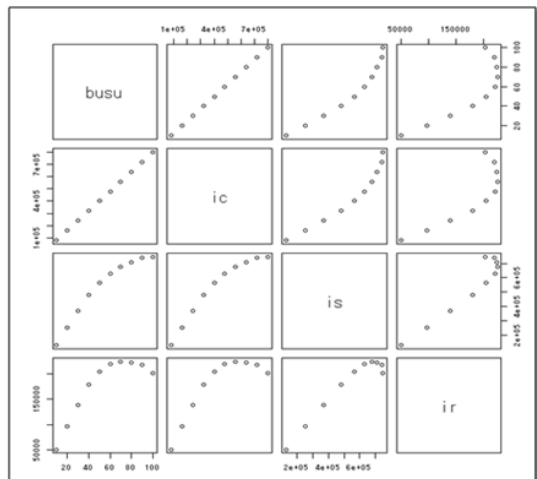


그림 3. plot 도표로 분석한 시뮬레이션 결과  
Fig. 3 Simulation result using plot diagram

그림 3을 보면 부수, 원가, 판매금액, 이익이라는 4가지 요인에 대한 상호 연관성을 한 개의 그래프로 표현하고 있다. 이것을 활용하면 시뮬레이션에 관련되는 4가지 요인간의 연관성에 대해 손쉽게 파악할 수 있으며, `plot(mydata)`라는 간단한 명령으로 수행 가능하다.

이외에도 시뮬레이션에 관련된 요인들의 산포와 최대값을 보여주는 시점에 대한 분석을 한다고 가정하는 경우에 그림 2의 결과만으로는 손쉽게 발견하기가 어렵다.

그림 4를 살펴보면 이익 요인의 경우 상위 부분에서 최대의 값을 보인다는 것을 그래프로 확인

할 수 있고 이것을 기반으로 의사결정에 도움이 되는 부수 정보를 추적할 수 있다.

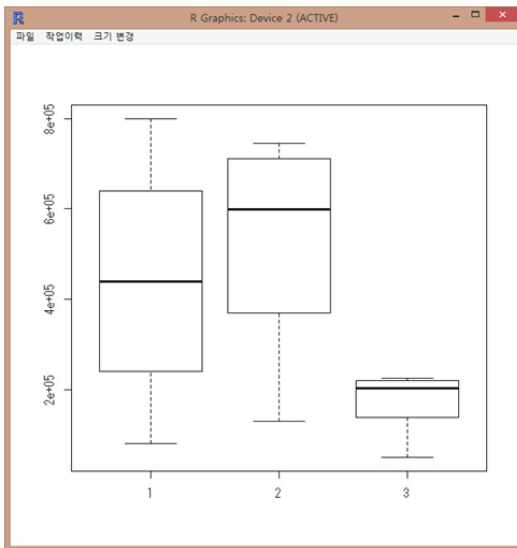


그림 4. Boxplot 도표로 분석한 시뮬레이션 결과  
Fig. 4 Simulation result using boxplot diagram

#### IV. 평가 및 향후 연구 방향

R은 통계 분석을 위한 패키지이다. 하지만, SAS, SPSS와는 다르게 시뮬레이션의 수행을 위한 프로그래밍 기능과 수행 결과를 다양한 관점에서 분석할 수 있는 기능을 가지고 있다[8-9].

#### [R이 제공하는 시뮬레이션 관련 기능]

- 시뮬레이션 모델을 프로그램 할 수 있는 프로그램 기능(예 : 순환, 반복, 자료형 선언)
- 시뮬레이션 수행을 위한 수준 높은 난수 발생 기능
- 시뮬레이션 결과를 다양한 분석에 적용할 수 있도록 변경할 수 있는 기능  
: 벡터, 행렬, 데이터 프레임의 다양한 자료형을 구성할 수 있도록 하는 기능
- 그래픽 기능  
: 고수준/저수준 그래픽 함수를 활용하여 시뮬레이션 결과를 다양한 관점에서 시각화할 수 있는 기능과 외부에서 제공하는 다양한 그래픽 기능 활용
- 통계 분석 기능  
: 시뮬레이션 결과를 가설 검증과 시계열 분석, Cluster/Factor를 포함한 다양한 통계 기법에 적용할 수 있는 기능

R을 사용하여 시뮬레이션을 수행하는 것이 가능하다는 것을 이번 논문의 결과로서 실증해 보였다. 단순한 몬테카를로 시뮬레이션을 떠나서 좀 더 복잡한 형태의 시뮬레이션에도 R을 적용하면 정확하고 다양한 분석이 가능한 시뮬레이션을 수행할 수 있다.

이번 논문의 결과를 기반으로 하면, 대기 행렬이나 의사 결정과 같은 시뮬레이션 모델의 개발도 가능하다. 향후, 추가적인 연구를 통하여 실제 사례를 제시할 예정이다.

마지막으로 ARENA와 같은 도구를 이용하여 시뮬레이션을 수행하더라도, 생성된 결과에 대한 정확한 분석에 R의 기능을 활용하는 것은 시뮬레이션의 가치와 정확성을 높이는데 크게 기여할 것으로 생각한다.

#### Reference

- [1] S. Robinson, *Simulation : The Practical of Model Development and Use*. New York : John Wiley & Sons, 2004.
- [2] A. Seila, V. Ceric, and P. Tadikamalla, *Applied Simulation Modeling*. Kentucky : Cengage Learning, 2003.

- [3] C. Lee and Y. Lim, "Simulation Analysis of a Underground Rail Haulage System Using SIMAN," *The Korean Society of Mineral and Energy Resources Engineers*, vol. 28, no. 5, 1991, pp. 339-347.
- [4] J. A. Buzacott and J. G. Shanthikumar, *Stochastic models of manufacturing systems*. New Jersey : Prentice Hall, 1993.
- [5] S. Kim, "Approximation of multi-class queueing networks with highly variable arrivals under deterministic routing," *Naval Research Logistics*, vol. 52, 2005, pp. 339-408
- [6] A. Heindl, "Decomposition of general tandem queueing networks with MMPP input," *Performance Evaluation*, vol. 44, 2001, pp. 5-23.
- [7] W. D. Kelton, R. P. Sadowski, and D. T. Sturrock, *Simulation with ARENA*. New York : McGraw Hill, 2012.
- [8] H. Nobuho, *Data Analysis and Visualization by using R*. Seoul : Hanbit Media, 2014.
- [9] K. Yang and M. Kim, *R Introduction and basic programming*. Paju : Freedom Academy, 2007.

### 저자 소개



#### 조민호(Min-Ho Cho)

1982년 인하대학교 컴퓨터공학과 졸업(공학사)

1998년 숭실대학교 대학원 컴퓨터공학과 졸업(공학석사)

2003년 숭실대학교 대학원 컴퓨터공과 졸업(공학박사)

2013년~현재 중원대학교 컴퓨터시스템공학과 교수

※ 관심분야 : 소프트웨어공학, 운영체제 및 소프트웨어 아키텍처



#### 전용호(Yong-Ho Jeon)

1996년 광운대학교 제어계측공학과 졸업(공학사)

1998년 광운대학교 대학원 제어계측공학과 졸업(공학석사)

2008년 광운대학교 대학원 정보제어공학과 졸업(공학박사)

2013년~현재 중원대학교 메카트로닉스학과 교수

※ 관심분야 : 가변속 시스템, 지능로봇 시스템