

## 엑셀 매크로기능을 이용한 베イズ 정리 교육도구 개발

최현석<sup>1</sup> · 하정철<sup>2</sup>

<sup>12</sup>계명대학교 통계학과

접수 2012년 6월 29일, 수정 2012년 7월 23일, 게재확정 2012년 7월 30일

### 요약

본 논문에서는 엑셀매크로로 베イズ 정리 교육도구를 개발하여 사용의 예를 소개한다. 주어진 어느 한 사건이 발생했을 때 그 사건이 특정조건하에서 발생되었는지 여부에 관심이 있다. 이런 경우의 확률계산에 사용할 수 있는 것이 베イズ 정리이다. 베イズ 정리는 새롭게 얻어진 부가적인 정보를 기초로 통계적 의사결정을 하는데 매우 유용한 정리이다. 베イズ 정리를 중간과정과 설명을 통해 학습자 스스로 효율적으로 학습할 수 있도록 개발한 교육도구를 소개한다. 조건부확률, 곱셈 법칙, 전확률 공식, 사전확률, 사후확률 등에 대한 설명과 활용 예를 단계적 학습을 통해 이해할 수 있도록 하였다. 결과가 나오기까지의 과정을 단계적인 개념설명과 그림으로 표현하여 단계적, 시각적인 학습이 되도록 하였다. 한 화면상에서 계산과정과 결과를 나타내도록 하기 위하여 분할 2개와 3개에 대하여 엑셀 자체에서 제공되는 분석기능과 비주얼베이직으로 작성된 프로그램을 연결하여 명령단추를 누르면 매크로가 실행되게 하였다.

주요용어: 베イズ 정리, 사전확률, 사후확률, 양식도구, 엑셀 매크로.

### 1. 서론

조건부확률의 개념에서 특정사건의 확률은 그 사건과 연관된 새로운 정보가 입수되면 개선하거나 수정할 수 있다. 새로운 실험결과에서 나온 정보를 이용하여 어떤 사건의 처음 확률을 개선시킬 수 있는데 처음 확률을 사전확률 (prior probability) 이라 하고, 개선된 확률을 사후확률 (posterior probability)이라한다. 이러한 확률들을 계산하는 방법을 제공하는 것이 베イズ 정리 (Bayes' rule)이다. 베イズ 정리는 Thomas Bayes가 'Essay towards solving a problem in the doctrine of chances'에 발표한 이론이다. 베イズ 정리는 의사결정에 광범위하게 이용된다. 사전확률은 종종 의사결정자가 생각하는 주관적 확률추정치로 주어진다. 표본으로부터 정보가 얻어지면 최선의 의사결정을 하기 위해 사후확률을 계산한다. 자동화시스템과 통합시스템 등의 제조 시스템에서 공정의 이상 유무를 민감하게 탐지하거나 주가동향에 대한 예측 등 많은 분야에서 그 응용이 증가하는 추세이다.

사후확률을 계산하는 것은 여러 단계를 거치기 때문에 번거롭고 복잡하다. 본 연구에서는 엑셀의 VBA (visual basic for application)를 이용하여 개념과 원리에 따른 단계적인 계산과정을 제공하여 학습자 스스로 단계적 절차에 따라 개념을 이해할 수 있는 교육용도구를 개발한다.

베イズ 법칙의 활용에 대한 연구는 미니탭 매크로를 이용한 연구 (Albert, 1996; Berry, 1996)와 엑셀 매크로를 이용한 연구가 있다 (Kim과 Baek, 2011). Kim과 Baek (2011)의 연구는 자료입력 후

<sup>1</sup> (704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 통계학과, 초빙교수.

<sup>2</sup> 교신저자: (704-701) 대구광역시 달서구 신당동 1000번지, 계명대학교 통계학과, 부교수,  
E-mail: jeicy@kmu.ac.kr

개념에 대한 설명 없이 결과만 제공하는 등 분석기능에 초점을 두고 있으나 본 연구는 결과가 나오기까지의 과정을 단계적인 개념설명, 동적으로 변화하는 그래프의 움직임으로 표현함으로써 단계적, 시각적, 동적인 학습이 되도록 한 것이 다른 연구와 구별된다고 할 수 있다.

본 연구의 목적은 엑셀 매크로를 이용하여 베이즈 정리를 중간과정, 설명들과 함께 나타내어 학습자 스스로 효율적으로 학습할 수 있도록 개발된 교육용도구를 소개한다. 즉, 조건부확률, 곱셈법칙, 전확률 공식, 사전확률, 사후확률 등에 대한 설명과 활용 예를 단계적인 학습을 통하여 이해하도록 하는 것이다. 2절에서는 연구방법에 대하여 설명하며, 3절에서는 프로그램의 구성 및 설명, 마지막으로 4절에서 본 연구의 결론을 맺고자 한다.

## 2. 연구방법

기초통계학을 배우는 학생에게 사전분포가 여러 가지 값을 갖는 형태로 주어졌을 때 사전분포를 설정하고 그에 따른 실험결과를 얻어 사후분포를 계산하는 것은 복잡하다. 또한 여러 단계를 거쳐 독립적으로 자료를 얻었을 때, 매 단계마다 사후분포를 계산하기가 매우 번거롭다 (Kim과 Baek, 2011). 이와 같이 사후확률의 매 단계 복잡하고 번거로운 계산을 엑셀의 매크로 기능을 이용하여 단계적 학습이 될 수 있도록 한다.

엑셀의 분석도구는 제한적이어서 메뉴를 이용한 일반통계분석에는 한계가 있으나 엑셀 매크로 (Excel Macro)로 프로그램을 작성하면 고급기법들을 사용할 수 있다 (Jacobson, 2002; Walkenbach, 2004). 엑셀 기능의 차원을 높일 수 있는 양식도구 (forms)와 VBA에 기초한 매크로를 이용한 통계 교육도구는 통계학의 개념과 이론을 이해하는데 많은 도움을 줄 수 있다. 이미 엑셀의 비주얼베이직 매크로를 이용한 다양한 통계교육 혹은 분석도구가 개발되어 활용되고 있다 (Choi와 Ha, 2011; Lee, 2008). 또한 웹사이트 [www.unistat.com](http://www.unistat.com), [www.xlstat.com](http://www.xlstat.com)에서도 엑셀을 이용한 다양한 통계 분석방법을 제공하고 있다.

본 교육용 도구 개발을 위해 VBA Project의 모듈 창에 다음과 같이 코드를 작성하였다.

첫째, Sub 문으로 지정된 셀에서 입력받은 값을 수식을 이용하여 기록하는 프로시저를 작성하였다.

둘째, 엑셀 자체에서 제공되는 분석기능과 VBA 등으로 작성된 프로그램을 연결하였다.

셋째, 명령단추 (Command Button)를 사용하여 단추를 누르면 바로 매크로가 실행되게 하였다.

## 3. 프로그램의 구성 및 설명

베이즈 이론은 사전적 확률을 추가적인 정보를 토대로 수정하여 사후적 확률을 얻는 과정이다. 즉 미지의 현상에 대하여 실험과 관찰 전에 어떤 확률 (사전확률)을 할당하기 보다는 실험과 관찰을 통하여 얻은 결과를 추가적인 정보로 하여, 실험과 관찰 이전의 확률을 수정하여 다시 확률 (사후확률)을 구하는 것이다. 베이즈 정리는 다음과 같이 정의한다.

사건  $A_1, A_2, \dots, A_n$ 이 표본공간  $S$ 의 분할이라고 하자. 사건  $B$ 가 주어졌을 때 사건  $A_i$ 가 발생할 확률  $P(A_i|B)$ 는 다음과 같다.

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

단,  $P(A_i) > 0$ ,  $P(B) > 0$

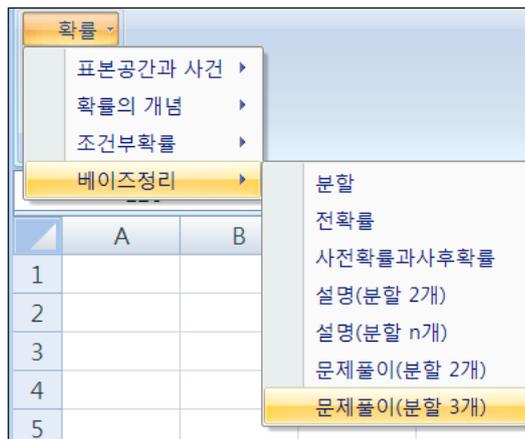
사건  $A_i$ 에 대하여  $P(A_i)$ 를 사전확률이라 하고  $P(A_i|B)$ 를 사후확률이라고 한다.

베이즈 정리에 활용되는 조건부확률, 곱셈법칙, 전확률 공식 등을 정리하면 Table 3.1과 같다. 본 교육용도구는 Table 3.1을 단계적으로 나타내어 최종적인 사후확률을 계산할 수 있도록 하였다.

**Table 3.1** Terms used in calculating Bayes' rule

terms	formula
partition	$\cdot A_1 \cup A_2 \cup \dots \cup A_n = S$ $\cdot A_i \cap A_j = \phi, (i \neq j)$
exclusive events	$\cdot P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$
multiplication rule	$\cdot P(A \cap B) = P(A)P(B A)$
conditional probability	$\cdot P(B A) = \frac{P(A \cap B)}{P(A)}$
law of total probability	$\cdot P(B) = \sum_{i=1}^n P(A_i)P(B A_i)$
Bayes' rule	$\cdot P(A_i B) = \frac{P(A_i)P(B A_i)}{\sum_{i=1}^n P(A_i)P(B A_i)}$

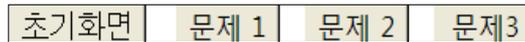
엑셀을 실행하면 ADD IN 프로그램으로 구성되어 있어 메뉴표시줄에 ‘확률’이 추가 된다 (Figure 3.1). 각각의 목차를 클릭하면 VBA Project의 모듈 창에서 비주얼베이직으로 작성한 프로그램의 해당 메뉴로 이동한다.



**Figure 3.1** Initial display

일부의 기능을 소개하면 다음과 같다.

단계 1: Figure 3.1에서 ‘문제풀이(분할 3개)’를 클릭하면 Figure 3.2와 같이 나타나 베이지 정리에 관한 문제 3개를 활용하도록 하였다. ‘문제1’을 클릭하면 Figure 3.3과 같이 문제가 나타난다. Figure 3.4의 ①을 클릭하면 사건 A, B, C에 사전확률 입력 셀이 나타나는데 확률을 입력하면 분할 (partition) 된 화면이 나타난다.



**Figure 3.2** Problem display

[문제 1] 우리나라 근로자들의 학력을 조사한 결과 중졸이 40%, 고졸이 50%, 대졸이 10%이었다. 중졸자 중에서 실업자는 10%이었고, 고졸자 중에서는 5%, 대졸자 중에서는 2%이었다. 임의로 한명의 실업자를 선택하여 학력을 조사하였을 때 다음의 확률을 구하시오.

- ① 실업자일 확률은?
- ② 실업자라는 조건하에 그 사람이 중졸자일 확률은?
- ③ 실업자라는 조건하에 그 사람이 고졸자일 확률은?
- ④ 실업자라는 조건하에 그 사람이 대졸자일 확률은?

Figure 3.3 Problem

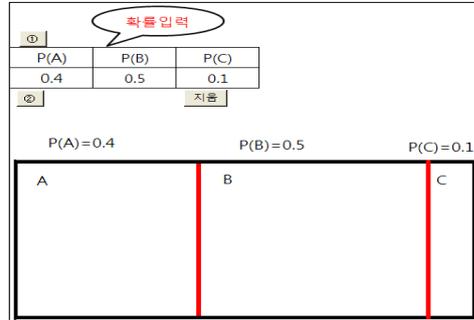


Figure 3.4 Prior probability

단계 2: ②를 클릭하여 추가사건에 대한 확률을 입력하면 Figure 3.5와 같은 화면이 나타난다. 화면에 나타나는 명령단추를 클릭하면 나무그림 (tree diagram)으로 사전확률, 조건부확률 및 결합확률에 대한 값이 계산되고 정리된 표가 제공된다 (Figure 3.6).

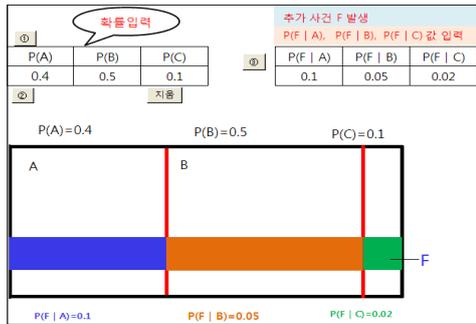


Figure 3.5 Additional event F

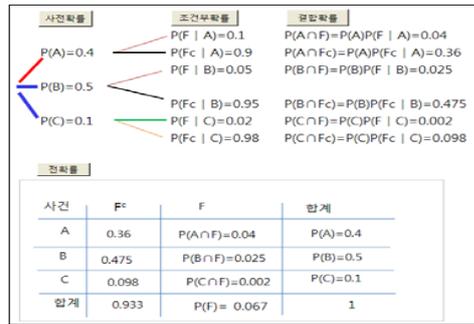


Figure 3.6 Tree diagram

단계 3: 전확률과 추가사건(F)이 주어졌을 때 사후확률이 계산된다 (Figure 3.7). 사건, 사전확률, 조건부확률, 결합확률, 사후확률에 대한 값과 설명을 제공하여 이해하기 쉽도록 하였다 (Figure 3.8).

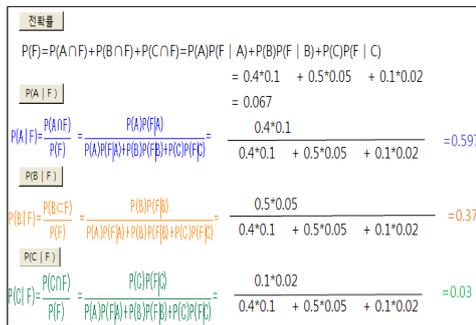


Figure 3.7 Posterior probability

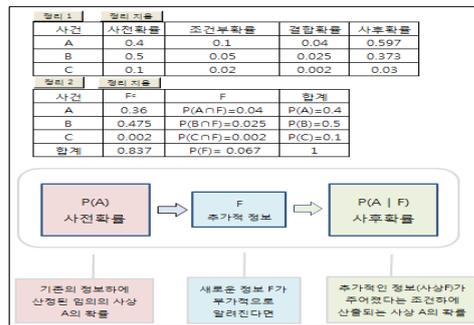


Figure 3.8 Summary

단계 4: Figure 3.9는 사전확률 입력에서 사후확률계산까지의 전 과정을 보여주는 화면이다.

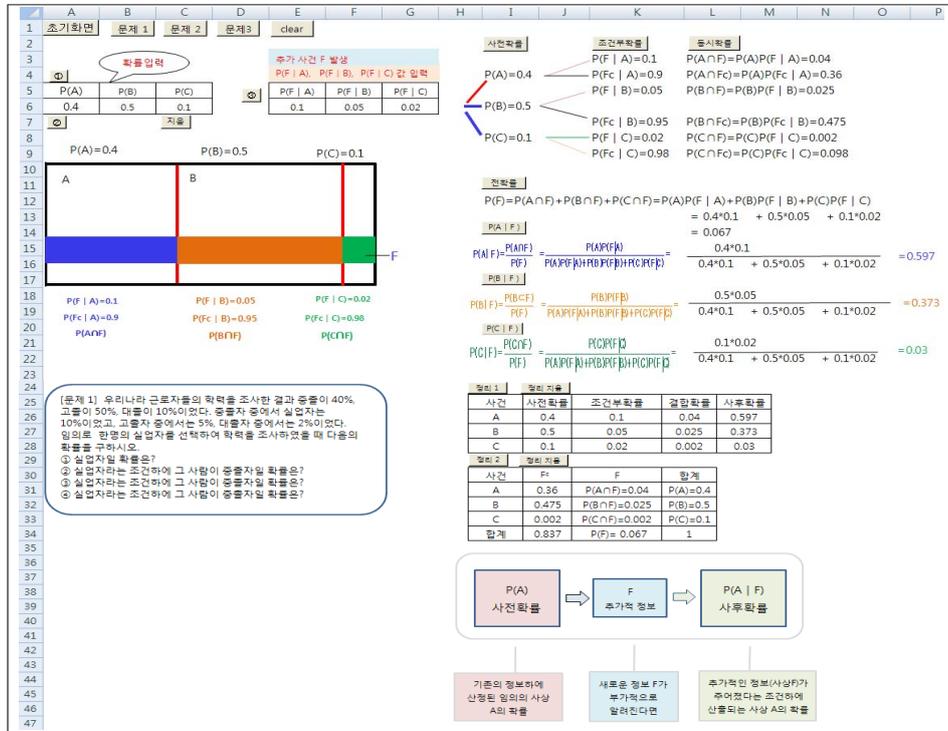


Figure 3.9 Full screen

문제1에 대한 학습이 끝난 후에는 Figure 3.2에서 문제2, 문제3을 클릭하여 다른 문제를 풀 수도 있고 사전확률과 추가확률을 변화시켜 사전확률과 추가확률의 관계를 이해할 수 있다.

Figure 3.1의 초기화면과 별도로 메뉴방식 (Figure 3.10)으로도 제공하여 표본공간과 사건, 확률의 법칙, 조건부 확률, 분할 설명, 베이즈 정리 설명 등 확률의 전반에 대하여 학습할 수 있도록 하였다.



Figure 3.10 Main menu

Figure 3.11은 조건부확률 곱셈정리에 대한 계산화면으로 자료를 입력하여 단계적으로 확률의 계산과정이 나타나게 하여 이해를 돕도록 하였다.

문제
clear

자료입력

안경 \ 성별	남(M)	여(F)	합계
안경(G)	20	30	50
미착용(N)	10	40	50
합계	30	70	100

안경 \ 성별	남(M)	여(F)	합계
안경(G)	0.2	0.3	0.5
미착용(N)	0.1	0.4	0.5
합계	0.3	0.7	1

<p><b>P(M)</b> <math>P(M) = 30 / 100 = 0.3</math></p> <p><b>P(F)</b> <math>P(F) = 70 / 100 = 0.7</math></p> <p><b>P(G)</b> <math>P(G) = 50 / 100 = 0.5</math></p> <p><b>P(N)</b> <math>P(N) = 50 / 100 = 0.5</math></p>	<p><b>P(M∩G)</b> <math>P(M∩G) = 20 / 100 = 0.2</math></p> <p><b>P(F∩G)</b> <math>P(F∩G) = 30 / 100 = 0.3</math></p> <p><b>P(M∩N)</b> <math>P(M∩N) = 10 / 100 = 0.1</math></p> <p><b>P(F∩N)</b> <math>P(F∩N) = 40 / 100 = 0.4</math></p>
---	---

한 학생을 선택했을 때

<p><b>P(M G)</b> 안경을 끼고 있다는 조건하에서 그 학생이 남학생일 확률 <math>P(M G) = P(M∩G) / P(G) = 0.2 / 0.5 = 0.4</math></p> <p><b>P(F G)</b> 안경을 끼고 있다는 조건하에서 그 학생이 여학생일 확률 <math>P(F G) = P(F∩G) / P(G) = 0.3 / 0.5 = 0.6</math></p> <p><b>P(G M)</b> 남학생이라는 조건하에서 그 학생이 안경을 착용할 확률 <math>P(G M) = P(G∩M) / P(M) = 0.2 / 0.3 = 0.667</math></p> <p><b>P(N M)</b> 남학생이라는 조건하에서 그 학생이 안경을 미착용할 확률 <math>P(N M) = P(N∩M) / P(M) = 0.1 / 0.3 = 0.333</math></p> <p><b>P(M∩G)</b> <math>P(M∩G) = P(M) P(G M) = 0.3 \times 0.667 = 0.2</math> <math>P(M∩G) = P(G) P(M G) = 0.5 \times 0.4 = 0.2</math></p> <p><b>P(F∩G)</b> <math>P(F∩G) = P(F) P(G F) = 0.7 \times 0.429 = 0.3</math> <math>P(F∩G) = P(G) P(F G) = 0.5 \times 0.6 = 0.3</math></p>	<p><b>P(M N)</b> 안경 미착용이라는 조건하에서 그 학생이 남학생일 확률 <math>P(M N) = P(M∩N) / P(N) = 0.1 / 0.5 = 0.2</math></p> <p><b>P(F N)</b> 안경 미착용이라는 조건하에서 그 학생이 여학생일 확률 <math>P(F N) = P(F∩N) / P(N) = 0.4 / 0.5 = 0.8</math></p> <p><b>P(G F)</b> 여학생이라는 조건하에서 그 학생이 안경을 착용할 확률 <math>P(G F) = P(G∩F) / P(F) = 0.3 / 0.7 = 0.429</math></p> <p><b>P(N F)</b> 여학생이라는 조건하에서 그 학생이 안경을 미착용할 확률 <math>P(N F) = P(N∩F) / P(F) = 0.4 / 0.7 = 0.571</math></p> <p><b>P(M∩N)</b> <math>P(M∩N) = P(M) P(N M) = 0.3 \times 0.333 = 0.1</math> <math>P(M∩N) = P(N) P(M N) = 0.5 \times 0.2 = 0.1</math></p> <p><b>P(F∩N)</b> <math>P(F∩N) = P(F) P(N F) = 0.7 \times 0.571 = 0.4</math> <math>P(F∩N) = P(N) P(F N) = 0.5 \times 0.8 = 0.4</math></p>
--	--

Figure 3.11 Conditional probability and multiplication rule

Figure 3.12~3.14는 ‘설명(분할 2개)’ 단추를 클릭하면 나타나는데 베이스 정리에 대한 이론 설명으로 단추를 클릭 할 때마다 관련된 설명을 나타나게 하여 단계적으로 이해하는데 도움이 되게 하였다.

초기화면
베이스 정리 설명

Figure 3.12 Explanation 1



Figure 3.13 Explanation 2

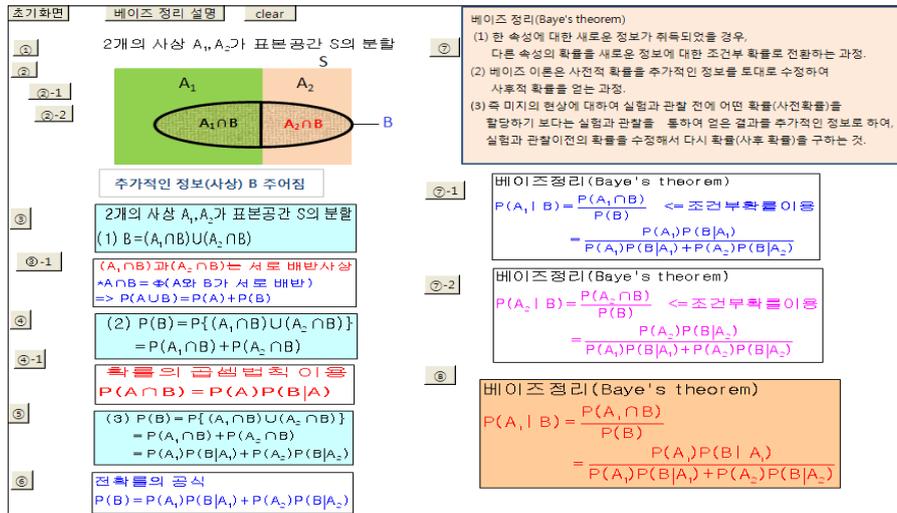


Figure 3.14 Explanation 3

#### 4. 결론

본 연구는 사후확률에 대한 이해를 돕고 계산과정과 계산의 번거로움을 해결하고자 엑셀매크로를 이용하여 단계적인 학습을 하도록 만든 교육용도구에 대한 소개이다. 사후확률이 나오기까지의 과정을 단계적인 설명과 동적으로 변화하는 그림으로 표현하여 단계적, 동적인 학습이 되도록 하였다. 엑셀을 실행하면 ADD IN 프로그램으로 구성되어 있어 메뉴표시줄에 '확률'이 추가 되어 각각의 목차를 클릭하면 VBA Project의 모듈 창에서 비주얼베이직으로 작성한 프로그램의 해당 메뉴로 이동한다. 즉, 엑셀 자체에서 제공되는 분석기능과 비주얼베이직으로 작성된 프로그램을 연결하여 명령단추를 누르면 매크로가 실행되게 하였다. 프로그램에 제시된 6문항의 예제 이외에 다른 예제들도 사전확률과 추가확률을 입력하면 반복된 계산을 할 수 있다.

이 프로그램은 학습자 스스로 통계학습을 할 수 있을 뿐만 아니라 기초통계학 강의에서 보조 자료로도 사용가능하다. 결과가 나오기까지 설명, 과정 등이 주어지므로 학습효과를 최대화 할 수 있으며, 엑셀 프로그램만 있으면 바로 실행하여 원리와 과정을 학습할 수 있는 장점이 있다. 한 화면상에서 계산과정과 결과를 나타내도록 하기 위하여 분할 2개와 3개에 대하여 프로그램을 개발하였으나 차후에 분할 개수에 관계없이 사용할 수 있도록 할 것이다.

#### 참고문헌

Albert, J. H. (1996). *Bayesian computation using minitab*, Duxbury Press, Belmont, CA.  
 Berry, D. A. (1996). *Statistics : A Bayesian perspective*, Duxbury Press, Belmont, CA.  
 Choi, H. S. and Ha, J. (2011). Development of process-oriented education tool for Statistics with Excel Macro. *Journal of the Korean Data & Information Science Society*, **22**, 643-650.  
 Jacobson, R. (2002). *Microsoft excel 2002 visual basic step by step*, Microsoft Press, Redmond, WA.  
 Kim, J. H. and Baek, H. Y. (2011). Excel macro for applying Bayes' rule. *Journal of the Korean Data & Information Science Society*, **22**, 1183- 1197.  
 Lee, J. Y. (2008). An example of participatory statistics class using Excel Macro. *The Korean Journal of Applied Statistics*, **21**, 355-359.  
 Walkenbach, J. (2004). *Excel 2003 power programming with VBA*, Wiley Publishing, New York, NY.

## Development of Bayes' rule education tool with Excel Macro

Hyun Seok Choi<sup>1</sup> · Jeongcheol Ha<sup>2</sup>

<sup>12</sup>Department of Statistics, Keimyung University

Received 29 June 2012, revised 23 July 2012, accepted 30 July 2012

### Abstract

We are dealing with the Bayes' rule education tool with Excel Macro and its usage example. When an event occurs, we are interested in whether it does under certain conditions or not. In this case, we use the Bayes' rule to calculate the probability. Bayes' rule is very useful in making decision based on newly obtained statistical information. We introduce an efficient self-teaching educational tool developed to help the learners understand the Bayes' rule through intermediate steps and descriptions. The concept and examples of intermediate steps such as conditional probability, multiplication rule, law of total probability, prior probability and posterior probability could be acquired through step-by-step learning. All the processes leading to result are given with diagrams and detailed descriptions. By just clicking the execution button, users could get the results in one screen.

*Keywords:* Bayes' rule, Excel Macro, form, posterior distribution, prior distribution.

---

<sup>1</sup> Assistant professor, Department of Statistics, Keimyung University, Daegu 704-701, Korea.

<sup>2</sup> Corresponding author: Associate professor, Department of Statistics, Keimyung University, Daegu 704-701, Korea. E-mail: jeicy@kmu.ac.kr