The Korean Communications
in Statistics Vol. 7, No. 3, 2000
pp. 913-926

Implementation of Estimation and Inference on the Web

Heemo Kang[1] and Songyong Sim[2]

Abstract

An electronic statistics text on the web is implemented. The introduced text provide interactive instructions on the statistical estimation and inference. As a by-product, we also provide a calculation of quantiles and p-value of t-distribution and standard normal distribution. This program was written in JAVA programing language.

Keywords : Dynamic graphics, Confidence interval, Quantiles, p-value.

1. 서론

통계에 대한 활용은 거의 모든 분야에서 중요한 도구로서 이용되고 있으며, 다양한 통계 패키지가 자료분석을 실행하는데 사용되고 있다. 그러나 대부분 패키지의 출력이 통계량의 계산결과만 주고, 출력에 대한 설명 및 해설이 없기 때문에 통계에 대한 기초적인 개념이 없는 통계학 비전공자 또는 초심자들이 출력결과를 이해하고 분석하는데 여러 가지 어려움이 있다 (조신섭(1998)).

이러한 원인을 세 가지 측면에서 살펴보면 첫째 대부분 사람들이 사용하는 통계프로그램이 한글화가 되어 있지 않다. 즉 통계전공 영어를 모르는 사람은 분석하는데 어려움이 있다. 둘째 통계학을 전파하는 문제로 통계의 기본개념이 추상적이고 통계 이론상 복합적인 개념을 내포하고 있고, 통계를 분석하는 기술, 즉 공식의 사용, 변수의 사용, 모델의 설정 등을 이해시키는 것이 어렵다. 셋째 통계의 기술면을 배우려면 약간의 수학적인 기초가 있어야 하는데 초보자들은 이러한 것들이 부족하다 (Yilmaz(1996)).

이러한 것들을 초보자들이 짧은 기간에 배운다는 것은 어려운 일이다. 그리고 대부분의 통계초보자들은 수학적으로 통계개념을 이해하는 것이 쉽지 않다. 어떻게 통계를 가르치는 것이 통계를 이해하는데 효과가 있는지 학습하는 방법을 보면, 기존의 교육방법으로 기초적인 수학 이론을 학습하고 이것을 통계이론에 적용하여 통계의 개념을 이해하는 방법과 인터넷으로 전자교재를 사용하여 통계를 배우는 방법이 있다. 전자교재란 학습자가 스스로 학습할 수 있도록 웹상에 올려진 자료로 사용자와 전자교재 사이에 상호작용을 할 수 있는 것에서 강의노트에 이르기까지 다양한 것이 있다. 전자교재의 사용은 여러 가지 면에서 유용하다. 강의실이나 특별한 교재가 준비되어 있지 않아도 사용할 수 있고 인터넷이 연결되어 있는 장소이면 누구나 언제든지 사용할 수 있는

1) (200-702) Department of Statistics, Hallym Universiy, Kangwondo, Korea
E-Mail : hmkang@fisher.hallym.ac.kr
2) (200-702) Assistant Professor, Department of Statistics, Hallym University, Kangwondo, Korea
E-Mail : sysim@sun.hallym.ac.kr

면에서 시간적 공간적 비용을 절감할 수 있다. 학습하는 면에서는 여러 가지 다양한 자료를 구할 수 있고, 특히 동적으로 구현된 전자교재는 사용자와 상호작용이 가능하기 때문에 시각적으로 쉽게 이해할 수 있는 점에서 유용하다(심송용(1997), 한경수(1998)).

본 고에서는 통계이론의 하나인 '추정 및 검정'을 웹상에 전자교재로 구현하여 통계 비전공자들이 쉽게 개념을 이해할 수 있도록 하였다. 통계전자교재는 웹상에 많이 존재하는데, 그 중에서 동적 그래프로 제작된 전자교재의 장소는

http://anova.inha.ac.kr/~dhuhm
http://compstat.chonbuk.ac.kr/
http://statpots.chonbuk.ac.kr/
http://168.126.199.80/
http://www.stat.sc.edu/~west/javahtml
http://www.ruf.rice.edu/~lane/stat_sim

등이 있다.

본 고에서 제작한 전자교재의 구현은 객체지향언어인 JAVA(Horton(1997))를 사용하여 구현하였다. 구현한 전자교재가 있는 장소는

http://fisher.hallym.ac.kr/~hmkang

에 있다.

2. 웹상에서 구현한 추정 및 검정

2.1 모의실험을 통한 신뢰구간

기초 통계를 수강한 사람이더라도 신뢰 구간의 개념을 잘못 이해하는 경우가 있는데, 그 예로 확률변수가 서로 독립이고 동일한 정규분포를 따를 때 모 평균 μ에 대한 95%의 신뢰구간(Confidence Interval)을 구하면 그 신뢰구간 안에 모 평균 μ가 포함될 확률이 95%라고 알고 있는 사람들이 있다. 그 원인은 하나의 신뢰구간에 모 평균 μ가 포함되어 있을 확률이 95%라고 신뢰구간의 의미를 잘못 해석하는 것이다.

이 소절에서는 동적 그래프를 이용하여 신뢰구간의 의미를 이해하는 방법을 제시하는데, 확률변수가 서로 독립이고 동일한 정규분포를 따를 때 모 평균 μ에 대한 $100(1-p)\%$의 신뢰구간을 구한 뒤 그 신뢰구간에 모수가 포함되는 경우의 수를 파악하여 몇 개의 신뢰구간이 모 평균 μ를 포함하는지 확인하고 계속 반복하면 모 평균 μ를 포함하는 누적된 신뢰구간의 개수의 비율이 $100(1-p)\%$에 근사하는 것을 알 수 있다.

모평균의 신뢰구간은 2가지로 구현하였는데, "분산을 알고 있는 경우"와 "분산을 모르는 경우"로 나누었다. 구현된 전자교재의 신뢰구간 사용법 및 출력결과를 살펴보면 중앙에 있는 검은색 직

선은 모 평균 μ를 나타내고 빨강색으로 되어 있는 수직선은 모 평균 μ를 포함하는 신뢰구간이며 파랑색으로 되어 있는 수직선은 모 평균을 포함하지 않는 신뢰구간이다. 왼쪽 하단의 글씨는 모 평균 μ를 포함하지 않는 신뢰구간의 개수와 신뢰구간을 포함하는 개수를 나타낸다. 하단의 '확률' 우측의 입력 창에는 신뢰구간의 신뢰도를 입력하는 곳으로, 본래는 0에서 1사이의 값을 입력해야 되지만 통계 초보자들에게 친숙한 %의 개념을 사용하여 0에서 100까지의 표본의 크기를 입력 할 수 있도록 하였고 변경도 가능하도록 하였다. '표본의 크기' 좌측의 입력창의 숫자는 모집단에서 추출한 표본의 크기를 나타내는데, 출력을 디지털 화면에 표현하기 때문에 표본의 크기를 너무 크게 하면 신뢰구간의 폭이 0에 가까워서 변화하는 정도의 차이를 알 수 없으므로 표본의 크기를 1에서 999까지 제한하여 입력할 수 있도록 하였다. 실행 방법은 신뢰도와 표본의 크기에 입력을 하고 '다시 그리기' 버튼을 누르면 그 값을 입력받아서 노랑색 바탕화면에 모 평균 μ에 대한 신뢰구간 100개가 구현된다.

모비율에 대한 신뢰구간은 모 평균의 신뢰구간에서 모의 실험한 동일한 방법으로 구현하였고, 사용방법도 동일하다. 전자교재로 신뢰구간을 구현한 장소는

http://anova.inha.ac.kr/~dhuhm
http://168.126.199.80/

에서 볼 수 있다. 위 장소에서는 모 평균 μ를 포함하는 신뢰구간의 개수만 보여주고, 제한적인 신뢰도를 사용하고 있다. 이 소절에서는 분산과 표본의 개수에 따라 신뢰구간의 길이가 어떻게 변하는지 보여주고 신뢰도는 0에서 100사이의 값은 모두 입력이 되도록 하였다.

2.1.1 모 평균의 신뢰구간(분산을 알고 있을 때)

신뢰구간을 구현하기 위하여 자료가 필요한데 이 소절에서는 표준정규분포에서 표본의 개수만큼 난수를 생성하여 자료를 만들고, 이 자료로 신뢰구간을 구현하였다. 이 소절에서는 분산을 알고 있는 경우(σ=1)를 구현하였는데 신뢰구간을 100개 생성하여 이 중에서 모 평균 μ를 포함하는 신뢰구간이 몇개가 있는지 알 수 있도록 하였다. 모 평균 μ에 대한 100(1- p)% 신뢰구간의 출력결과를 그림 1에 구현하였다. 이 그림을 보고 알 수 있는 것은 분산과 표본의 개수가 같으므로 모든 신뢰구간의 길이가 일정한 것과 표본개수의 증가, 감소에 따라 모 평균 μ의 신뢰구간 변화를 알 수 있다. 신뢰구간을 구하는데 필요한 표준정규분포의 제 p백분위수는 유리분수근사(rational fraction approximation)를 이용하였다(Kennedy(1980)). 제 p백분위수를 구하는 원리 및 방법은 제 p백분위수의 참값 x_p에 대응하는 누적확률밀도함수(Cumulative Distribution Function)를 $\phi(x_p)$, 근사값 x_0에 대응하는 누적확률밀도함수를 $\phi(x_0)$, 확률을 p라 할때, $Z(\widehat{x_p})=|\phi(x_p)-\phi(x_0)|$라 하면 $Z(\widehat{x_p})$의 역함수 $Z^{-1}(\widehat{x_p})$가 $\widehat{x_p}$의 제 p백분위수가 되는 것이다. $p\in(0,0.5)$인 p에 대하여 $p=1-\phi(\widehat{x_p})$를 구하고 이 식을 $\widehat{x_p}$에 대하여 정리하면, $\widehat{x_p}=1-\phi^{-1}(p)$인 식을 구할 수 있다. 이 $\widehat{x_p}$가 표준정규분포 제 p백분위수로 이것을 구하는 식은 Kennedy(1980)에 주어진 알고리즘을 사용하였다.

이 알고리즘은 $p \in (10^{-20}, 0.5)$인 p에 대하여 $y = [\log(1/p^2)]^{\frac{1}{2}}$를 구하여 x_p의 근사값

$$\widehat{x_p} = y + \frac{(((yp_4 + p_3)y + p_2)y + p_1)y + p_0}{(((yq_{4+q_3})y + q_2)y + q_1)y + q_0} \tag{1}$$

를 계산하는 것이다. 이 알고리즘의 정밀도는 $|x_p - \widehat{x_p}| < 10^{-7}$인 것으로 알려져 있다. 여기서 p_i와 q_i는 $(i = 0, 1, \cdots, 4)$ Kennedy(1980)에서 얻을 수 있다. 만일 $p \in (0.5, 1)$인 p에 대해서는 $1-p$를 위의 식 1에 적용하여 얻는다.

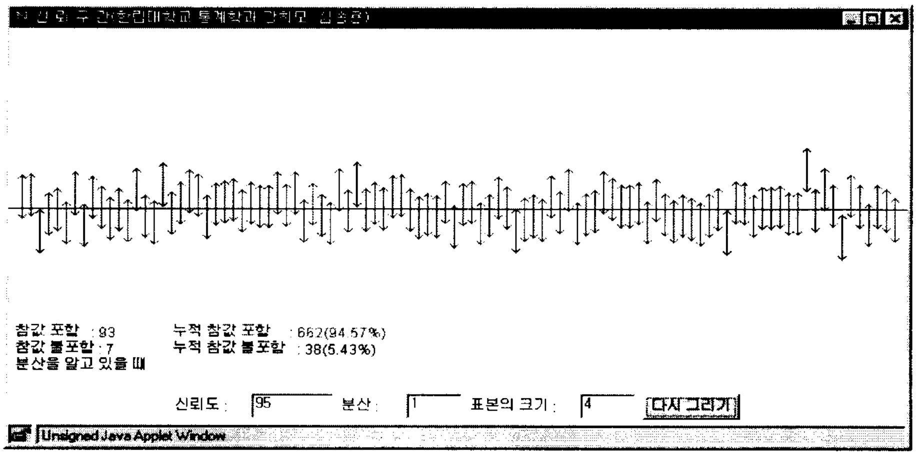

그림 1 모 평균의 신뢰구간(분산을 알고 있을 때)

2.1.2 모 평균의 신뢰구간(분산을 모를 때)

이 소절에서는 σ를 모르는 경우에 대한 신뢰구간을 구현하였다. 자료의 생성은 2.1.1소절에서 사용한 동일한 방법으로 표준정규분포에서의 난수를 구하였다. 이 소절에서 모의실험을 실행하면서 알 수 있는 것은 분산을 모르기 때문에 분산 σ^2의 추정치 $\widehat{\sigma^2}$을 사용하여 신뢰구간을 구하였기에 각각의 신뢰구간 길이가 다르게 나타나는 것이다. 2.1.1소절과 동일하게 모 평균 μ를 포함하는 신뢰구간의 개수를 확인하고, 모의실험을 계속 반복하여 실행하면 모 평균 μ를 포함하는 신뢰구간의 개수의 비율이 신뢰도에 근접함을 알 수 있다. 실행 결과는 그림 2에 나타내었다.

이 소절에서 사용한 모 평균 μ에 대한 신뢰구간은 t-분포의 제 p백분위수가 필요한데, 이 계산은 연속된 분수 근사(continued fraction approximation)를 이용하였다. 계산 방법은 t-분포의 누적확률밀도함수를 베타분포의 누적확률밀도함수로 전환하여 William(1992)에서 제시한 방법을 사용하여 구하였다. t-분포의 제 p의 백분위수의 정밀도는 $|x_p - \widehat{x_p}| < 10^{-4}$인 것으로 알려져 있다.

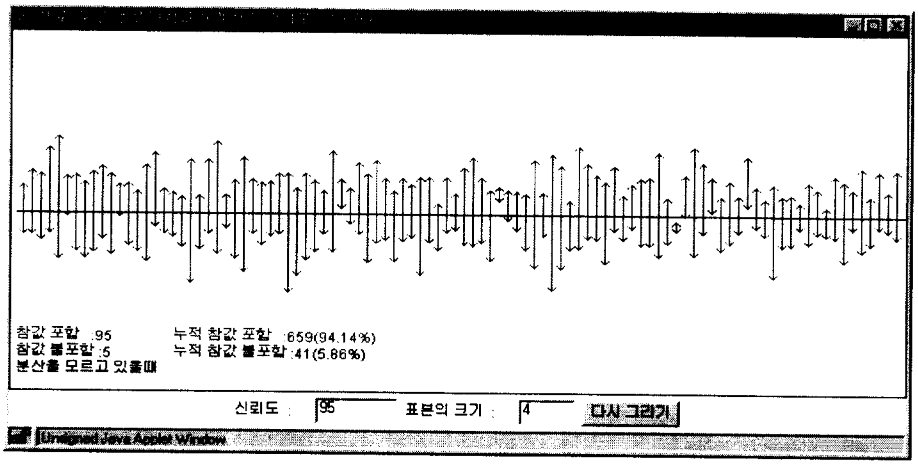

그림 2 모 평균의 신뢰구간(분산을 모를 때)

2.1.3 모 비율에 대한 신뢰구간

이 소절에서는 모 평균과는 다른 개념의 중요한 모수인 모 비율 p에 대한 신뢰구간을 구현하였다. 모 비율 p의 신뢰구간을 구하려면 모비율의 추정량 $\hat{p}$ 필요한데, 이것은 다음과 같은 과정으로 생성하였다.

1. 모 비율 p를 정한다.
2. 일양분포(uniform distribution)에서 난수를 생성한다.
3. 생성된 난수는

$$\begin{cases} 1 & \text{if} \quad 0 < u \le p \\ 0 & else \end{cases}$$

의 베르누이(bernoulli) 시행을 한다.

4. 3번 과정을 n번 반복하여 시행하여, 모비율의 추정량 $\hat{p}$ (성공회수/시행회수)를 100개를 구한다.

위의 과정으로 생성된 자료는 이항분포이므로 이항분포의 정규근사를 이용하여 신뢰구간을 구한다. 그런데 이 신뢰구간이 이항분포의 정규근사를 이용하므로 n과 p가 정규근사에 적합한 경우, 즉 n, p가 $\min\{np, n(1-p)\}>5$ (통계학, 구자홍)인 경우에만 주어진 값(시행회수, 확률, 신뢰도)으로 모의 실험을 하면 신뢰도에 가까운 값을 보이고(그림 3 n=30, p=0.5) 그렇지 않으면 모 비율 p를 포함하는 신뢰구간의 개수에 대한 비율이 신뢰도에 근접하게 되지 않는 것을 알 수 있다(그림 4 n=10, p=0.1).입력 창에 입력되는 값은 다음과 같은 조건을 만족해야 한다. 시행회수는 자연수, 확률은 0에서1사이의 값, 신뢰도는 0과 100사이의 값만 입력이 가능하도록 하였다.

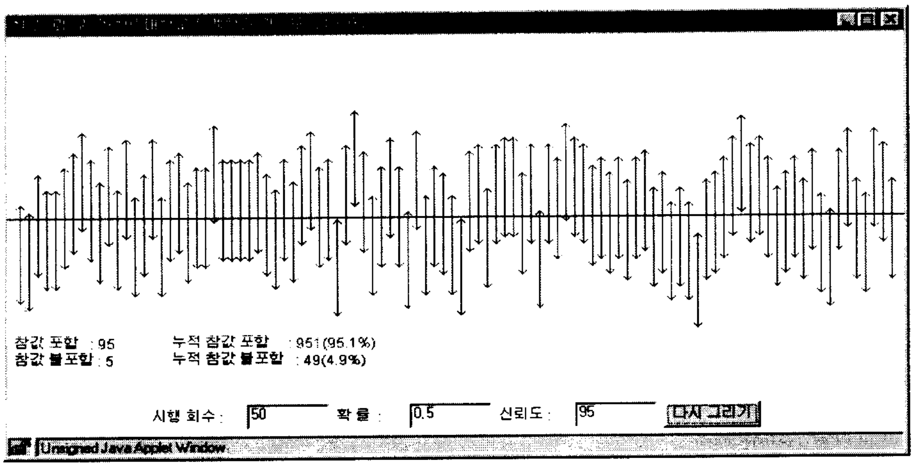

그림 3 모 비율의 신뢰구간(n=30, p=0.5)

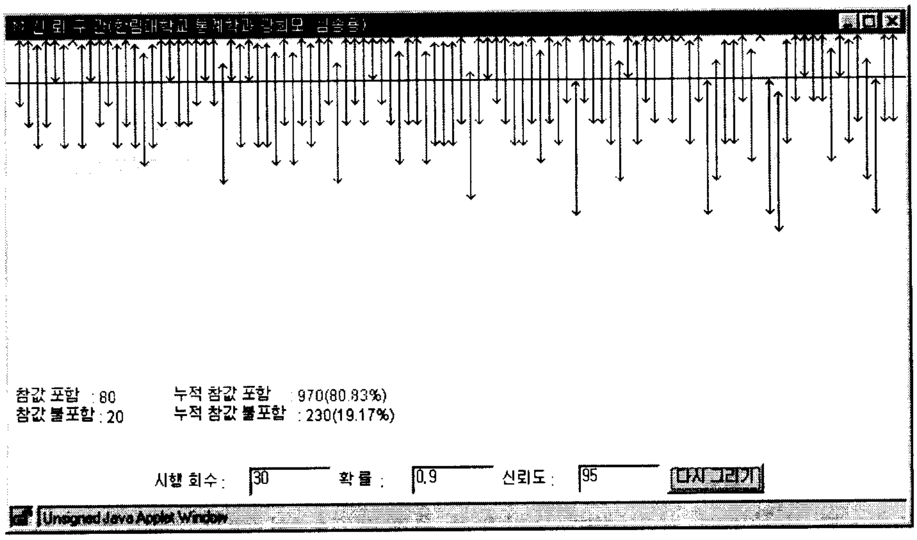

그림 4 모 비율에 대한 신뢰구간(n=30, p=0.9)

2.2 분위수(Quantile) 및 유의확률(p-value) 구하기

기존의 통계 프로그램은 검정결과를 유의확률값으로 출력하는데, 대부분의 통계 초보자들은 분포의 모양이나 성질을 모르기 때문에 출력결과를 이해하기가 쉽지 않다고 한다. 웹상에 전자교재로 확률밀도함수의 특성을 구현한 장소는

http://anova.inha.ac.kr/~dhuhm
http://stat.chonbuk.ac.kr/

등이 있는데, 이러한 것들은 신뢰도 100(1- p)%에 대하여 분위수만 구하는 것이거나 분위수에 대하여 누적확률밀도함수 $F_X(x)=P[X\le x]$의 값만 구할 수 있게 구현되어 있다. 이 소절에서는 분위수와 유의확률 중 원하는 값을 구할 수 있도록 구현하여 위의 장소에 있는 것보다 조금 더 사용자들에게 편리함을 제공하고 있다.

정규분포는 모수의 개수가 2개이므로 평균과 분산을 입력하여 분포가 변하는 것을 보여주도록 하였고, 신뢰도 100(1- p)%를 입력하면 제 p/2백분위수가 출력되도록 구현하였다. p/2백분위수에서 극단값(extreme value)까지 색칠하여 100(1- p)% 신뢰구간이나 양측검정에서 사용하기 편리하도록 하였다. 또한 분위수를 입력하면 신뢰도에서 입력한 방법과 동일하게 그 값을 구하면 유의확률은 $p\in(0,0.5)$인 p에 대하여

$$P[X\le x_{p/2}]+P[X\ge x_{1-p/2}]=1-2\Phi(x_{p/2})$$

값이 출력되도록 하였다.

평균이 μ, 분산이 σ^2, 신뢰도가 100(1- p)%인 정규분포에 관한 동적그래프가 그림 3에 있는데, 정규분포의 그래프 형태도 설명할 수 있을 뿐만 아니라 분산이 커지면 모 평균 μ에 대한 신뢰구간의 길이가 커지는 것과 평균은 신뢰구간의 길이에는 영향을 주지 않지만 위치가 변하는 것을 그림으로 이해할 수 있다. 또한 신뢰도가 변함에 따라 모 평균 μ에 관한 정규분포의 신뢰구간 길이가 변하는 것을 확인할 수 있다. 정규분포의 제 p백분위수와 유의확률은 2.1소절의 것을 이용하였다.

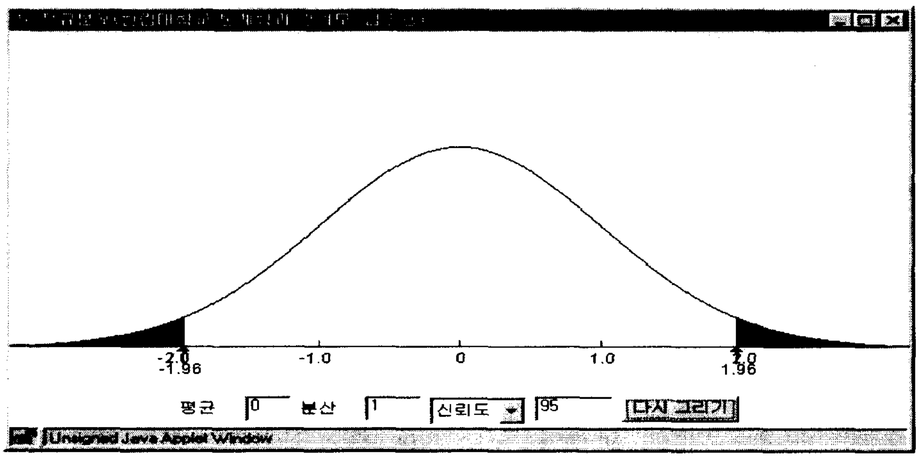

그림 5 정규분포에서 분위수 및 유의확률 구하기

t-분포는 정규분포에서 구현한 것과 동일한 방법으로 자유도가 v, 신뢰도 100(1- p)%를 입력하여 실행하면 자유도가 v인 t-분포의 그래프를 그리도록 하였다. 그림 3에서 분위수에 관한 그림을 보여주었으므로 여기에서는 유의확률에 관한 그림을 출력하였다. 자유도가 v이고, $p\in(0,0.5)$인 p에 대하여 유의확률이

$$P[T \le t_{p/2}] + P[T \ge t_{1-p/2}] = 1 - 2F[t_{p/2}]$$

값을 가지는 t-분포가 그림 4에 있다. 여기에서 사용자들이 알 수 있는 것은 자유도를 작게하면 꼬리부분이 두텁게 되고, 자유도를 크게 하면 꼬리부분이 작아지면서 정규분포에 가깝게 되는 것을 알 수 있다. 자유도가 6인 t-분포의 그림 4와 표준정규분포의 그림 3을 비교하면서 두 분포의 차이점을 사용자들이 확인할 수 있다. 또한 자유도와 신뢰도는 입력과 변경이 가능케하여 두 값을 변경하여 입력을 하면 그래프의 변화를 볼 수 있다. 자유도가 v인 t-분포의 제 p백분위수와 유의확률은 2.1.2소절에서 제시한 방법을 이용하여 구하였다.

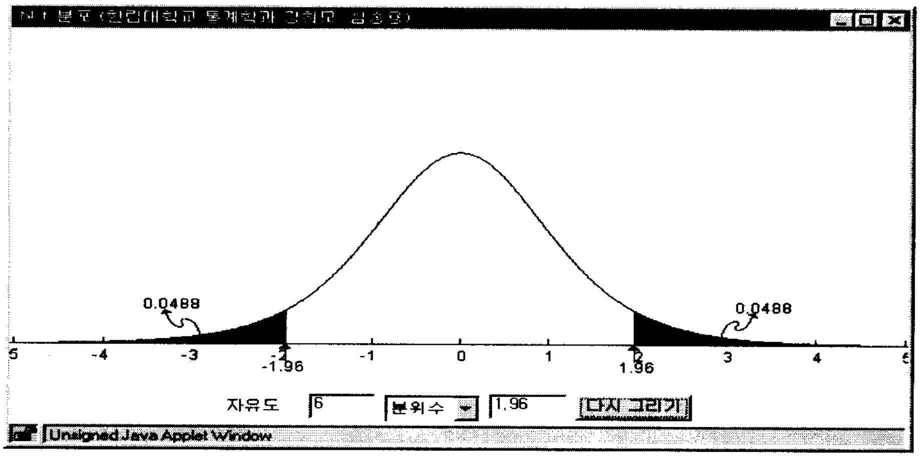

그림 6 t-분포에서 분위수 및 유의확률 구하기

이 소절에서 사용한 애플릿(Applet) 대한 사용방법은

1. 정규분포
   (a) 평균과 분산(σ)을 입력한다.
   (b) 신뢰도와 분위수 중 입력할 값을 선택한다.
   (c) 신뢰도의 값은 0 ~ 100의 값 사용(100(1- p))하고 분위수는 모든 값을 입력할 수 있다.
2. t-분포
   (a) 자유도는 자연수만 입력해야 한다.
   (b) 신뢰도와 분위수 중 입력할 값을 선택한다.
   (c) 신뢰도의 값은 0 ~ 100의 값 사용(100(1- p))하고 분위수는 모든 값을 입력할 수 있다.

이다. 위의 조건에 틀리는 경우에 애플릿이 실행되지 않고 입력이 틀린 원인을 보여 준다.

2.3 입력된 자료의 추정 및 검정

이 소절에서는 사용자가 실험결과의 자료를 직접 입력하여 자료분석을 할 수 있도록 하였다. 자료를 입력하여 통계분석을 하는 전자교재는

http://anova.inha.ac.kr/~dhuhm
에 있는데, 독립인 두 표본일 경우만 결과가 출력 되도록 구현되어 있다. 본 고에서 제작한 전자교재는 입력된 자료를 가지고 평균 분산 등의 기초통계량, 신뢰구간, 검정통계량을 출력하고 다음 창에는 t-분포의 그래프로 자료의 유의확률, 유의수준을 출력하도록 구현하여 분석 결과를 쉽게 이해할 수 있도록 하였다. 구성된 면을 살펴 보면 3개의 창으로 되어 있는데, 자료를 입력하는 창, 통계량을 출력하는 창, 출력된 통계량을 그림으로 출력하는 창으로 구성되어 있다.

기초 통계량과 신뢰구간은 기존에 많이 사용되고 있는 통계프로그램(SAS, MINITAB, S-Plus 등)의 출력결과가 영문으로 출력되지만 본 고에서 제작한 전자교재는 한글로 출력이 되도록 하였다. 이렇게 한글화함으로서 통계 초보자들도 출력결과를 쉽게 이해할 수 있도록 하였다. 또한 출력된 통계량과 분포의 그래프를 보고 비교하면서 결과을 명확하게 이해하고 분석할 수 있도록 구현하였다.

이 소절의 사용방법은

1. 자료 입력창
   (a) 단일표본(one sample), 독립인 두표본(two sample), 대응표본(paired comparison sample) 중 선택
   (b) 자료를 입력
      I. 자료입력시 집단의 구분은 ";"으로 한다.
      ii. 자료와 자료의 구분은 space로 한다.
   (c) 귀무가설, 대립가설을 설정한다.
   (d) 신뢰도를 입력한다.(0 ~ 100사이의 실수)
   (e) "계산하기" 버튼을 누른다.
2. 통계량 출력창
   (a) 통계량의 출력을 본 다음 "그림으로 출력" 버튼을 누른다.
3. 그래프 출력창
   (a) 파랑색은 유의수준(α)을 나타낸다.
   (b) 빨강색은 유의확률(p-value)을 나타낸다.

이다. 만일 위의 조건에 맞지 않게 입력을 하고 실행하는 경우는 틀린 원인을 보여 주었다.예를 들어 단일표본에서 ";"을 2개 입력한 경우나 두 표본이나 대응표본인데 ";"이 2개가 아닌경우, 대응표본이면서 두 집단의 자료의 개수가 일치하지 않는 경우, 자료 입력시 문자가 포함되어 있는 경우 등이 있다.

2.3.1 자료입력창

자료입력창은 실험결과로 얻은 자료를 입력하고, 사용자가 분석하려는 자료에 알맞게 여러가지를 설정해야 한다. 프로그램을 시작하면 사용설명서를 보여주어 사용자가 분석할 내용에 맞게 설정을 할 수 있도록 하였다. 입력할 수 있는 자료의 개수는 사용자가 원하는 만큼 입력이 가능하도록 하였다. 입력창에 대한 사용방법은 그림 5, 그림 6, 그림 7에 있다.

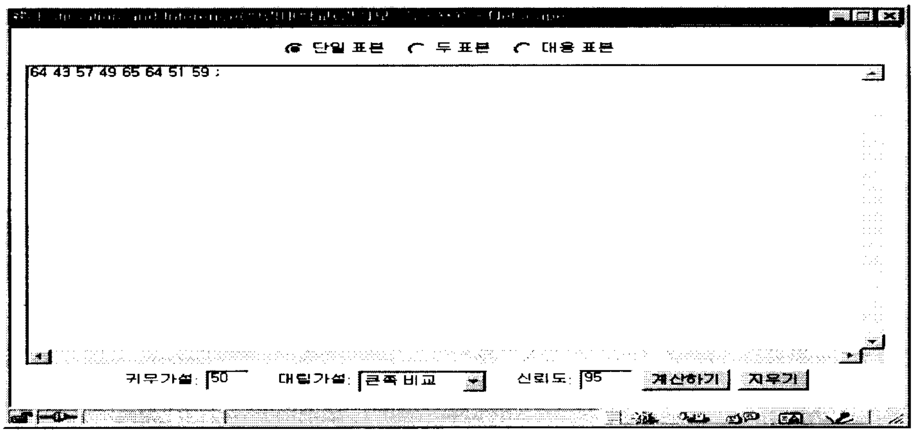

그림 7 입력된 자료의 추정 및 검정(자료 입력창-단일 표본)

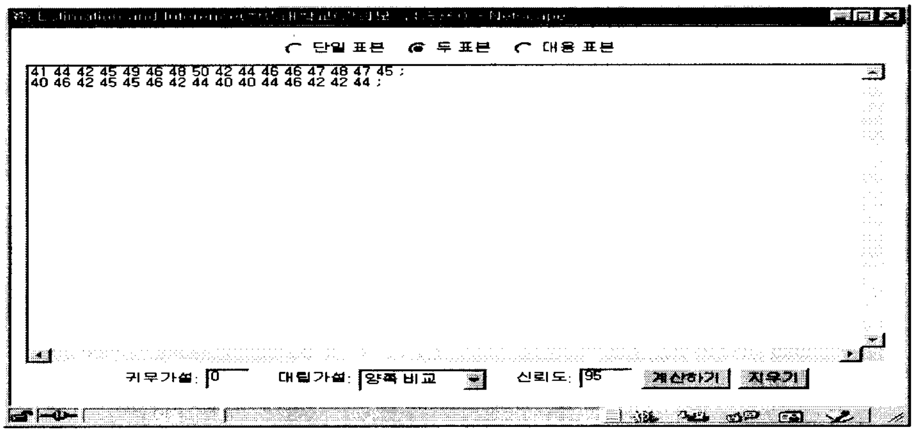

그림 8 입력된 자료의 추정 및 검정(자료 입력창-두 표본)

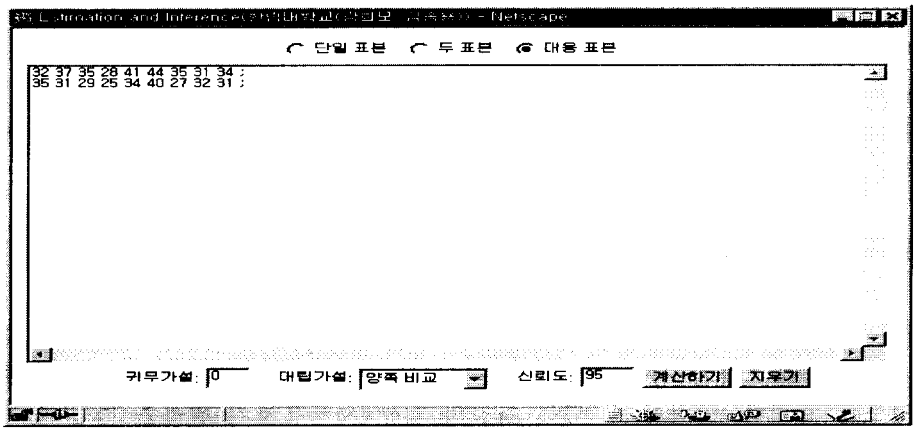

그림 9 입력된 자료의 추정 및 검정(자료 입력창-대응 표본)

2.3.2 통계량 출력창

이 소절에서는 자료를 입력하여 모 평균 μ에 대한 100(1- p)% 신뢰구간과 통계량을

1. 기초통계량 : 자료수 ,최소값, 최대값, 평균, 표준편차
2. 신뢰구간 : 자료입력창에서 입력된 신뢰도로 계산한다.
3. 검정통계량
   (a) 단일표본, 대응표본 : 모 평균 μ에 대한 t검정을 한다.
   (b) 두표본 : 등분산($\sigma_1 = \sigma_2$)에 대한 F검정과 모 평균 μ에 대한 t검정을 한다.

출력한다. 검정통계량은 입력창에서 설정한 귀무가설과 대립가설을 직접 문자로 출력하였으며 t-분포의 제 p백분위수와 유의확률을 출력하였다. 단일표본과 대응표본의 경우에 유의확률과 그 점에서의 제 p백분위수는 2.1.2소절에서 사용한 동일한 것이고, 두 표본의 경우에 등분산성($\sigma_1 = \sigma_2$)에 대한 검정은 F-분포의 검정통계량 2.1.2소절 t-분포에서 사용하였던 것과 동일한 방법으로 F-분포와 베타분포의 연관성을 가지고 베타분포의 함수로 나타내었다. F-분포의 제 p백분위수 및 유의확률에 대한 자세한 내용은 William(1992)에 있다. 지금까지 사용한 t-분포와 F-분포의 제 p백분위수 및 유의확률의 출력을 SAS Release 6.12, Mathematica Version 2.2로 비교하였는데, 2.1.2소절에서 제시한 정밀도를 초과하지 않는 범위에서 대부분이 같았으나 두 표본의 출력 중 이분산일 경우 유의확률값이 SAS에 있어서 잘못된 계산결과를 발견하였다. SAS에 문의해본 결과 SAS Release 7.0부터는 올바른 계산을 하도록 수정을 하였다고 한다. 이것에 대한 자세한 내용이 있는 장소는

```
http://www.sas.com/service/techsup/intro.html
```

에 있다.

통계량을 출력결과는 그림 8, 그림 9, 그림 10 형태로 출력된다.

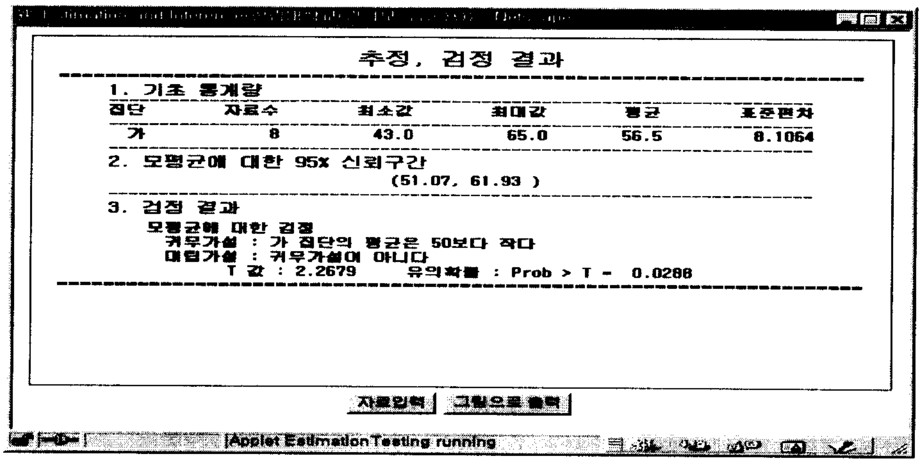

그림 10 입력된 자료의 추정 및 검정(통계량 출력창-단일표본)

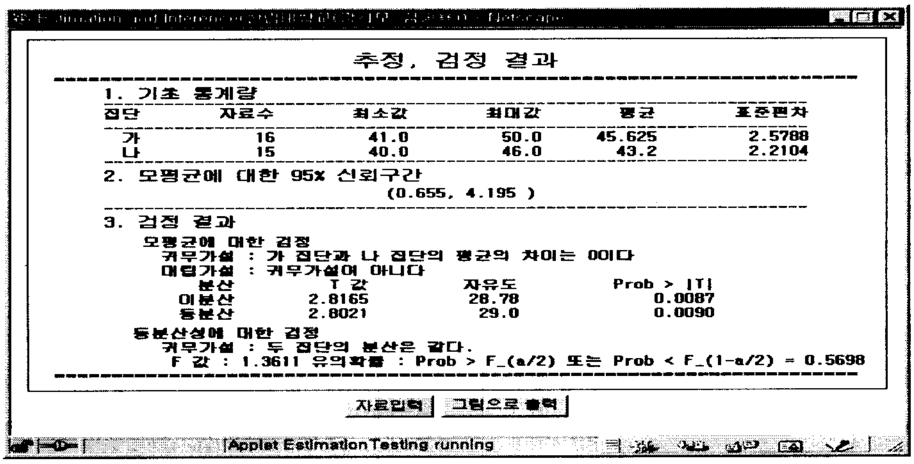

그림 11 입력된 자료의 추정 및 검정(통계량 출력창-두 표본)

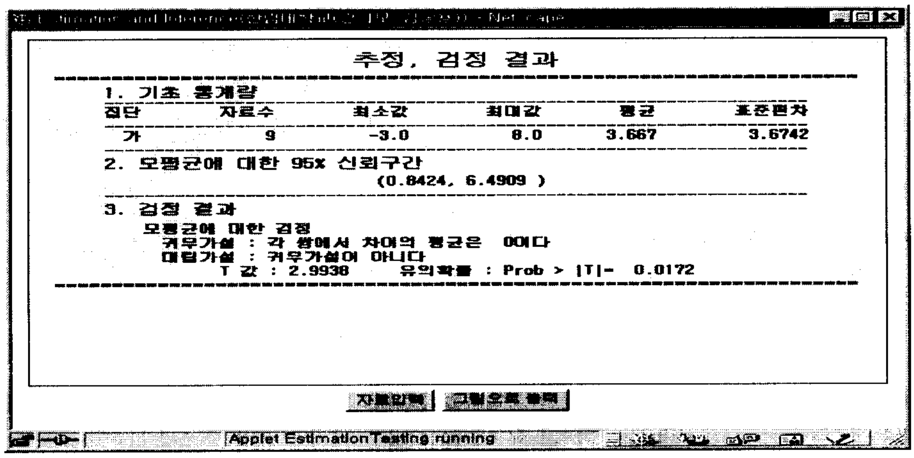

그림 12 입력된 자료의 추정 및 검정(통계량 출력창-대응 표본)

2.3.3 그래픽 출력창

그래픽출력창에서는 출력을 기각역과 유의확률을 비교하여 보여준다. 유의수준(α)을 파랑색 유의확률(p-value)을 빨강색으로 영역을 나타내어 첫화면에서 입력한 유의수준의 면적과 유의확률의 면적을 눈으로 쉽게 구별하여 귀무가설의 기각 채택을 쉽게 할 수 있고, 또한 유의확률의 면적을 숫자로 출력함으로서 유의수준 α와 차이를 한번 더 확인할 수 있도록 하였다. 그래프로 출력한 결과는 그림 11, 그림 12, 그림 13에 있다.

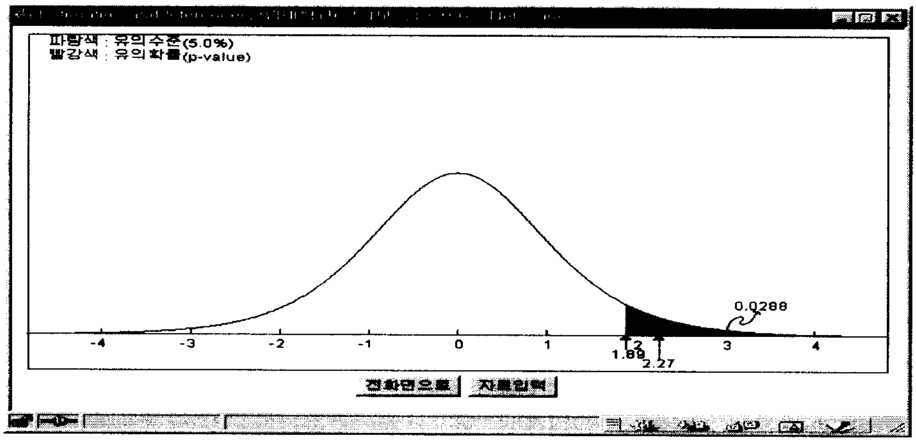

그림 13 입력된 자료의 추정 및 검정(그래픽 출력창-단일 표본)

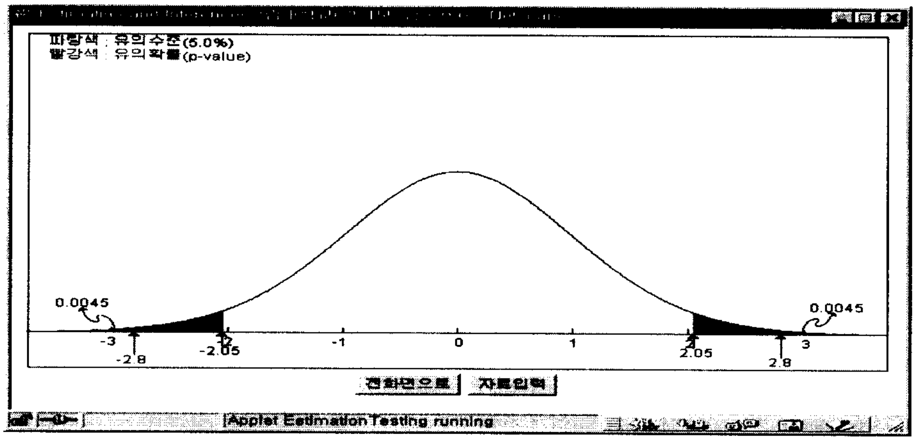

그림 14 입력된 자료의 추정 및 검정(그래픽 출력창-두 표본)

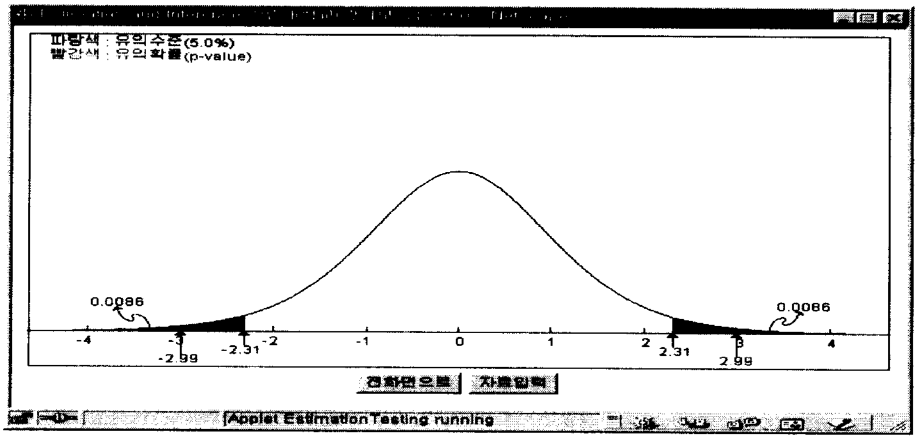

그림 15 입력된 자료의 추정 및 검정(그래픽 출력창-대응 표본)

3. 전자교재를 사용하기 위한 조건

앞에서 구현한 전자교재를 브라우저나 Appletviewer를 사용하여 볼 수 있는데 사용한 애플릿은 JDK1.1버전이므로 JDK1.0을 지원하는 오래된 브라우저나 Appletviewer로는 이것을 볼 수 없다. JDK1.1이상을 지원하는 브라우저는 Internet Explorer 4.0 이상인 버전이나 Netscape 4.05에 JDK1.1이 이식된 버전 이후의 것이어야 되고 Appletviewer도 JDK1.1이상의 버전이 있어야 구현된 전자교재를 볼 수 있다. 이러한 것들은 각각 아래의 URL에서

http://home.netscape.com/computing/download
http://www.microsoft.com/ie/ie50
http://java.sun.com/products/jdk/1.2/index.html

구할 수 있다.

4. 결론

추정과 검정은 통계학에 있어서 기본이 되며 중요한 이론이다. 통계학을 비 전공한 사람이 이론을 교재를 보면서 이해하는 것이 쉽지 않을 것이라 생각된다. 통계 초보자 또는 통계 비전공자들이 모의 실험을 실행하면서 동적으로 작동되는 전자교재를 보고 통계학을 습득하고 이해하는 것이 정적인 교재를 사용하는 것보다는 쉽다. 또한 주어진 환경을 보면 이 시대는 컴퓨터가 많이 보급되어 있고 인터넷을 자유스럽게 사용 할 수 있는 여건이 확립되어 있다. 전자교재의 활용은 많은 사람이 시간과 공간에 제약 없이 볼 수 있다는 점에서 통계학의 전파를 기존의 교육 방법보다 빠르게 할 수 있다..

참고문헌

[1] 구자홍, 김진경, 박헌진, 이재준, 전홍석, 최지훈, 황진수(1999), 「통계학」, 자유아카데미, 서울.
[2] 심송용 (1997), 인터넷을 이용한 원격수업, 1997년 한국통계학회 추계 학술 발표회 논문집, 1-5.
[3] 조신섭, 송문섭, 이윤모, 성병찬, 윤영주, 이현부 (1998), 기초통계교육을 위한 통계패키지의비교 연구 및 엑셀을 이용한 한글 통계패키지의 구현, 1998년 한국통계학회 춘계 학술 발표회 논문집, 75-79.
[4] 한경수, 안정용, 강윤비 (1998), 통계학 교육을 위한 전자 교재의 활용, 「응용통계연구」, 제 11권 1호, 5-12.
[5] Horton, Ivor(1997) *Beginning JAVA*, Wrox Press, Birmingham, UK.
[6] Kennedy W. J. and Gentile J. E. (1980), *Statistical Computing*, Marcel Dekker, Inc.
[7] Yilmaz M. R. (1996), *The Challenge of Teaching Statistics to Non-Specialists*, Journal of Statistics Education. Vol. 4. No. 1
http://www.anstat.org/publications/jse/v4n1/yilmaz.html
[8] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery(1992), *Numerical Recipes in C Second Edition*, Cambridge University Press