

Regression and Correlation Analysis via Dynamic Graphs¹⁾

Hee Mo Kang²⁾, Songyong Sim³⁾

Abstract

In this article, we propose a regression and correlation analysis via dynamic graphs and implement them in Java Web Start. For the polynomial relations between dependent and independent variables, dynamic graphics are implemented for both polynomial regression and spline estimates for an instant model selection. The results include basic statistics. They are available both as a web-based service and an application.

Keyword : Spline, Correlation Coefficient, Java Web Start, Studentized Residual, Cook's Distance.

1. 서론

회귀분석은 1900년 초에 도입되었으며 컴퓨터가 개발되면서 그 적용영역이 확장되어 사회과학, 공학, 의학, 언어학 등 여러 분야에 적용되고 있고 자료분석 과정에서 자주 사용되는 통계기법이다. 이 분석방법은 확률변수, 선형대수, 추정, 검정, 상관계수, 그리고 몇 가지 분포 등 여러 가지 기본지식을 바탕으로 독립변수와 종속변수간의 관계를 잘 설명되는 모형을 찾아내는 통계기법으로 통계 입문자나 비전공자가 회귀분석을 사용하기 위하여 이러한 기본지식을 습득하려면 적지 않은 시간이 소요된다. 회귀나 상관분석에 관련된 계산이나 그림은 기존의 통계 패키지에도 잘 구현되고 있으나 사용자와 대화식으로 구현하는 경우는 많지 않다. 웹과 자바(JAVA)언어의 등장 이후 통계적 이론을 설명하기 위한 대화식 응용프로그램이 많이 구현되고 있다(심송용(1997), 한경수, 안정용, 강윤비(1998), 강희모, 심송용(2000)). 본 논문에서는 여러 회귀모형을 한 화면에 출력하여 어떤 모형이 변수간의 관계를 잘 설명하는지 비교하는 프로그램과 상관계수 개념을 이해할 수 있는 프로그램을 웹을 기반으로 구현하였다. 이러한 프로그램은 모두 동적 그림을 기반으로 구현되었으며 사용자가 마우스나 키보드를 이용하여 프로그램의 설정 내용을 바꿀 수 있도록 설계되었다. 따라서 산점도에서 한 점의 위치나 자료의 값 변화에 따른 회귀나 상관분석의 결과가 어

1) This work was supported by Korea Research Foundation(KRF-2002-013-D00023).

2) Instructor, Department of Statistics, Hallym University, 1 Okchon-dong, Chunchon, Kangwon-do, 200-702.
Email: hmkang@hallym.ac.kr.

3) Associate professor, Department of Statistics, Hallym University, 1 Okchon-dong, Chunchon, Kangwon-do, 200-702.
Email: sysim@hallym.ac.kr.

떻게 바뀌는지 사용자가 쉽게 이해할 수 있도록 하였다. 이러한 결과의 변화는 추정치나 분산분석 표 등의 통계량의 변화뿐만 아니라 회귀직선, 곡선 스플라인 회귀선 등의 변화를 동적 그림으로 보여준다. 회귀모형을 비교할 수 있는 프로그램에서 제시한 모형은 모수 추정인 단순회귀, 다항회귀와 스플라인(spline)을 이용한 함수추정(Hastie and Tibshirani (1990), 구자용(2000))이 있고 상관계수에 관한 프로그램은 상관계수와 단순회귀의 관계를 사용자가 자료를 변경하면서 비교할 수 있게 되어 있다. 스플라인은 기저함수 방법론(Basis Function Methods)의 한 종류로 Stone의 함수추정 방식을 의미하는데 Wahba의 평활 스플라인 방식에 비하여 이론 정립이나 실자료에 대하여 성능이 우수하다는 장점이 있다고 알려져 있다(구자용(2000)).

본 논문에서 사용한 프로그램은 썬(Sun)사에서 제작한 Java Web Start를 이용하였다(Java 언어는 Arnold, Gosling and Holmes(2001)를 Java Web Start는 참고문헌의 Java Web Start 홈페이지 참고). Java Web Start는 기존의 Java Applet이 실행할 때마다 서버에서 클래스 파일을 내려 받는 비효율성을 제거하여 훨씬 효율적으로 웹기반에서 응용될 수 있는 새로운 기술로 Java Applet과 일반 응용프로그램의 장점을 취한 방식으로 웹을 통하여 설치, 실행 및 업그레이드를 제공할 수 있다. Java Web Start가 설치되지 않은 상태에서 저자들이 만든 URL에 접속하면 자동으로 설치화면이 연결되도록 설계하였고, 이미 설치된 경우 자동으로 최신 버전인지 확인하고 필요하면 업그레이드하도록 구현하였다. 최초 설치시 약간의 설치시간이 걸릴 수 있으나 한번 설치한 후에는 업그레이드된 부분만 확인하도록 프로그램이 구현되어 저자들이 프로그램을 업그레이드 한 경우에는 업그레이드 된 부분만 새로 설치하고 바로 실행할 수 있다. 자동 업그레이드 부분을 제외한 나머지 기능은 오프라인에서도 실행이 가능하다. 여기서 사용한 각종 분포에 대한 p 값은 Kennedy and Gentle(1980)과 Press, Teukolsky, Vetterling and Flannery(1992)에 소개된 알고리즘으로 직접 구현하였다.

2. 상관과 회귀 및 산점도

산점도, 상관계수, 단순회귀는 두 변수 사이의 관계에 대한 추론 방법으로 사용하며 이들은 서로 밀접한 관계가 있다. Pearson 표본상관계수는 두 변수에 대한 선형성 척도이고, Spearman 순위상관계수는 두 변수가 선형이 아닌 단조증가 또는 단조감소 하는 관계일 때 Pearson 계수보다 더 좋은 척도이다. 그러나 두 개의 상관계수는 두 변수의 관계가 3차 곡선, 원주 위의 점 등과 같은 비선형 관계를 측정해주지 못 한다. 하지만 실제 자료의 분석에서 두 변수의 상관계수나 분산분석 표만 보고 두 변수의 관계를 단정하는 경우가 많다. 본 논문에서는 이러한 상관계수가 선형 또는 단조성만을 측정하여 계수값이 실제 상관관계를 정확히 나타내 주지 못하는 경우를 동적 그림으로 구현하여 보았다. 또한 본 연구에서는 회귀분석과 상관계수의 관계 및 회귀분석과 함수추정식의 관계 등도 동적 그림으로 제공하여 입문자 뿐 아니라 연구자들에게도 기초자료를 제공하고자 하였다. 여기서 소개할 프로그램에서는 자료의 추가 또는 삭제가 마우스 클릭, 마우스 드래그 또는 자료값을 입력할 때도 가능하고 이러한 추가 삭제가 이루어진 즉시 Pearson 상관계수와 Spearman 상관계수, 분산분석표의 변화는 물론이며 단순회귀직선, 신뢰대(Confidence Band), 예측구간(Prediction Band)의 그림 변화량도 즉시 확인이 가능하도록 하였다. 신뢰대와 예측구간의 신뢰도는 사용자가 직접입력 할 수 있도록 하였는데

1. 0에서 1사이의 값인 경우 100을 곱하여 백분율의 값으로 변환한 신뢰도를 사용하고
2. 1에서 100사이의 값인 경우 그대로 신뢰도 값으로 사용하여

0에서 100사이의 모든 값을 입력받을 수 있도록 설계하였다. 사용자와의 대화성이 중요하므로 자료의 입력은

1. 마우스 클릭으로 클릭된 곳의 좌표를 산점으로 만들거나
2. 자료 파일에서 산점의 좌표를 읽어들이 수 있다.

마우스 클릭으로 생성된 산점의 좌표를 파일로 저장할 수 있도록 하여 이 자료를 다른 목적으로 사용하거나 애플리케이션 자료 파일로 읽어 사용할 수 있도록 하였다. 자료 파일에서 읽은 자료는 최초로 생성된 창 크기의 95%에 산점이 모두 찍히도록 하였고 자료 편집창에 자료가 모두 나타난다. 마우스 클릭으로 저장된 파일을 자료 파일로 읽어 사용하면 창 크기의 95%안에 모든 점이 찍히게 설계되었기 때문에 자료를 좌표로 읽어들이 후 마우스 클릭으로 점을 추가하거나 그 반대의 경우 산점도 창의 크기가 다를 수 있다. 실제로 창에 찍히는 산점의 픽셀 좌표 (x_i^{win}, y_i^{win}) 는

$$x_i^{win} = \left[\frac{0.05 + 0.95(x_i^{real} - x_{min}^{real})}{x_{max}^{real} - x_{min}^{real}} x_{win_range} \right] \quad i=1, 2, \dots, n$$

$$y_i^{win} = \left[\frac{0.05 + 0.95(y_{max}^{real} - y_i^{real})}{y_{max}^{real} - y_{min}^{real}} y_{win_range} \right] \quad i=1, 2, \dots, n$$

로 계산된다. 여기서 $[\cdot]$ 은 정수부분을 취하는 함수이며 (x_i^{real}, y_i^{real}) 는 실제 자료의 (x, y) 값이며 x_{win_range} 는 x 축 창 크기, y_{win_range} 는 y 축 창 크기이다. 자료 파일을 읽어들이 다음 자료를 추가하려면

1. 마우스를 클릭하거나
2. 자료 편집창에 자료를 입력한다.

마우스를 클릭하면 동적 그래픽 창에 산점이 찍히고 자료 편집창에 자료가 추가되어 어떠한 값이 추가되었는지 확인할 수 있다. 또 자료 편집창에 자료를 추가하고 "Redraw"버튼을 누르면 자료 편집창에 있는 자료로 그래픽 창에 산점을 그린다. 이때 추가된 자료가 창 크기 범위를 벗어나면 창 크기에 맞게 범위를 조정할 것인지 묻는 대화상자가 나타난다.

자료의 일부가 삭제되었을 경우 회귀식의 모양, 신뢰대, 예측구간, 피어슨 상관계수, 스피어만

순위상관계수, 통계량 등의 변화를 보고 싶으면,

1. 삭제하고자 하는 점을 포함하도록 마우스를 드래그하면 사각형이 생기고 사각형 안의 점이 자료에서 삭제되며,

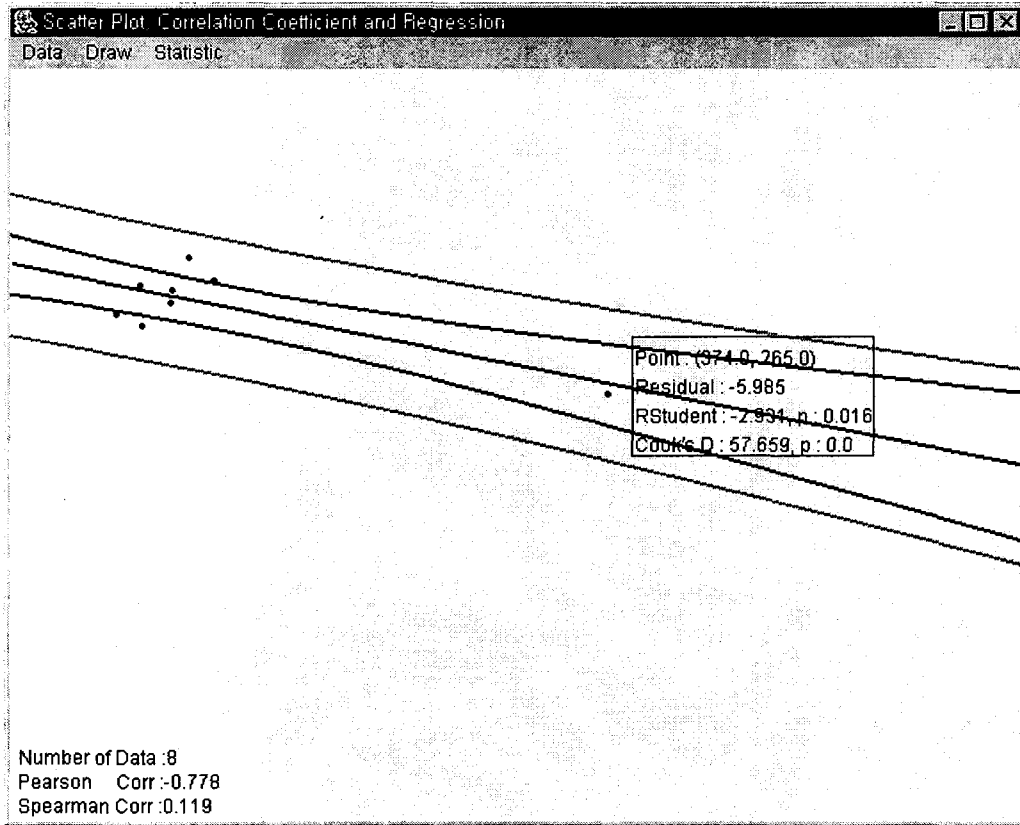


그림 1 상관과 회귀 및 산점도

2. 자료 편집장에 있는 자료를 삭제한 후 "Redraw"버튼을 누른다.

이 점들이 삭제된 후의 자료들에서의 회귀식, 신뢰대, 예측구간, 상관계수, 산점도 등을 자료에서 얻게 된다. 자료 편집창에서 자료를 삭제한 후 "Redraw"버튼을 누르면 마우스를 사용하여 삭제하였을 경우와 동일한 형태를 얻게 된다.

동적 그림에 찍힌 각 산점에 대한 정보는 마우스가 특정한 산점 위에 위치하면 그 산점의 좌표, 잔차, 스튜던트화 잔차(studentized residual)와 유의확률(p-value), 쿡 통계량(Cook's distance)과 유의확률을 작은 창을 만들어 알려주게 구현하여 이상치(outlier)와 영향력 관측치(influential observation)를 확인할 수 있고 마우스가 그 산점을 벗어나면 그 산점에 대한 정보와 통계량이 없어진다(그림 1 참조). 동적 그림에서 스튜던트화 잔차는

$$r_i^* = \frac{e_i}{s_{(i)}\sqrt{1-h_{ii}}} \quad (1)$$

로 계산한다. 여기서 e_i 는 잔차이고, $s_{(i)}^2$ 는 i 번째 관측치를 제외시킨 후 $(n-1)$ 개의 관측치로 회귀모형을 적합하여 얻어지는 $\sigma_{(i)}^2$ 의 불편추정량이고, h_{ii} 는 모자행렬(hat matrix)의 (i, i) 번째 원소이다. r_i^* 는 자유도가 $(n-p-1)$ 인 t분포를 따른다. 또한 영향력 관측치를 판별하는데 사용하는 쿡 통계량은

$$C_i = \frac{1}{p} \cdot \frac{r_i^2 h_{ii}}{(1-h_{ii})} \quad (2)$$

로 계산된다. 쿡 통계량 C_i 가 $C_i \geq F_{0.05}(p, n-p)$ 를 만족하면 영향력 관측치로 판단한다. 여기서 r_i 는 표준화 잔차(standardized residuals)로 식 (1)에서 $s_{(i)}^2$ 를 관측치 n 개에 대한 σ^2 의 불편추정량으로 구한 후 식 (1)을 계산한 것이다. p 는 회귀식에서 모수의 개수이다(Belsley, Kuh and Welsch(1980)).

동적 그림에 찍힌 산점들에 대한 통계량은 “Statistic→Anova” 메뉴를 선택하면 분산분석표, 결정계수, 회귀분석 모수 추정치가 새로운 창에 나타난다. 그림 1은 마우스 클릭으로 입력된 자료에 대하여 산점도, 단순회귀직선, 피어슨 표본상관계수, 스피어만 순위상관계수, 신뢰대, 예측구간, 좌표, 잔차, 스튜던트화 잔차, 쿡 통계량에 대한 결과이다. 본 논문에서 구현한 프로그램은

<http://pluto.hallym.ac.kr/plot/>

에서 웹을 통하여 실행할 수 있다.

3. 회귀분석 모형 비교

본 절에서는 회귀함수 추정에 대하여 고전적 회귀모형에서 사용하는 모수추정 방법과 스플라인 함수를 이용한 함수추정 방법에 대하여 동적 그래프와 통계량으로 비교하였다. 회귀추정식을 동적 그래픽에 표현하기 용이하도록 독립변수의 개수를 한 개로 하였다. 독립변수가 한 개인 경우 p 차 다항회귀함수 추정은 모형

$$E[Y] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x^p \quad (3)$$

을 추정하는 것이다. 다항회귀함수 추정에서 차수가 커지면 분산이 지수적으로 증가하여 큰 차수에 대해서는 추정량이 불안정해지기 때문에(구자용(2000)) 차수 p 는 9까지만 사용할 수 있도록 제한하였다. 본 논문에서 구현한 다항회귀는 동적 그래프와 통계량이 출력되도록 설계하였다. 이

두 가지 중 하나를 실행하면 사용자는 차수를 입력하도록 하였는데 차수 p 를 입력하면 그 즉시 다항회귀 추정식에 대한 동적 그래프와 분산분석표, 결정계수, 회귀계수 추정치와 유의확률 등 통계량을 동시에 볼 수 있고 차수를 변경하면 동적 그래프와 통계량이 즉시 변경된다. 함수추정방법은 스플라인 함수(spline function)를 사용하였다. 스플라인 함수는 일종의 조각(piecewise) 다항식으로 여러 구간에서 다항식인 함수이다. 스플라인 함수는

$$(x_i - t_{ik})_+ = \begin{cases} x_i - t_k & \text{if } x_i \geq t_k \text{ 일 때} \\ 0 & \text{if 그 외의 경우} \end{cases} \quad (4)$$

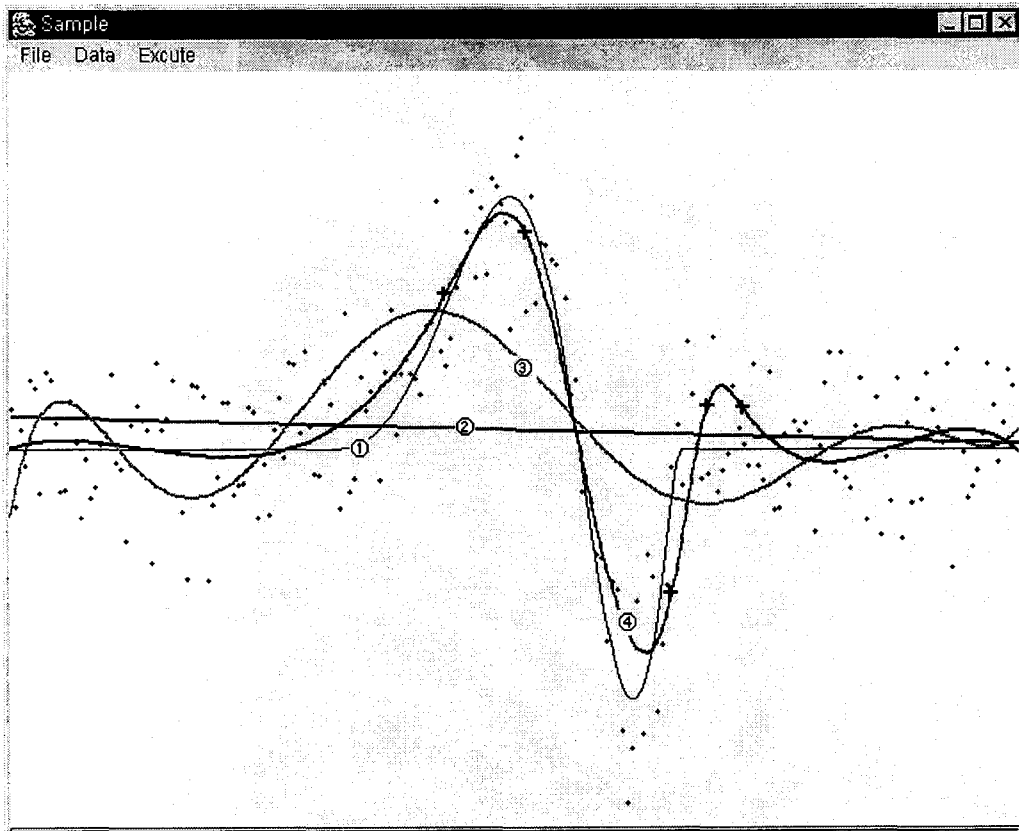


그림 2 회귀분석 모형의 성능비교

로 정의하는데 t_k 들을 절단점(break point) 또는 매듭점(knot point)이라 한다. 본 논문에서 구현한 스플라인 함수추정식은 3차 스플라인(Cubic Spline)으로

$$y_i = \beta_{-3} + \beta_{-2}x_i + \beta_{-1}x_i^2 + \beta_0x_i^3 + \sum_{j=1}^k \sum_{i=1}^n \beta_j(x_i - t_j)_+^3 + \varepsilon$$

급격히 변하는 자료에 대응할 수 있는 식으로 알려져 있다(구자용(2000)).

식 (5)에서 매듭점 t_j 의 선택은 $k=1, 2, \dots$ 와 $\alpha=0.05$ 및

$$m = \frac{-\log\left(-\frac{\log(1-\alpha)}{n}\right)}{\log(2) \times 5}$$

인(Friedman and Silverman(1989)) k, α, m 에 대해서 $1+km$ 번째 x 들 중 BIC를 기준으로 한 자료 선택단계와 삭제단계를 거쳐 최종 매듭점을 선택하여 다중공선성이 발생하지 않도록 하였다. 본 논문에서 구현한 함수추정식을 실행하면 앞의 다항회귀에서 출력된 동적 그래프와 통계량이 출력되고 추가적으로 매듭점에 대한 정보도 제공한다. 특히 동적 그래프에서 추가된 매듭점을 '+'로 표시하여 사용자가 눈으로 확인할 수 있도록 하였고 통계량도 출력 창에 출력되도록 구현하였다.

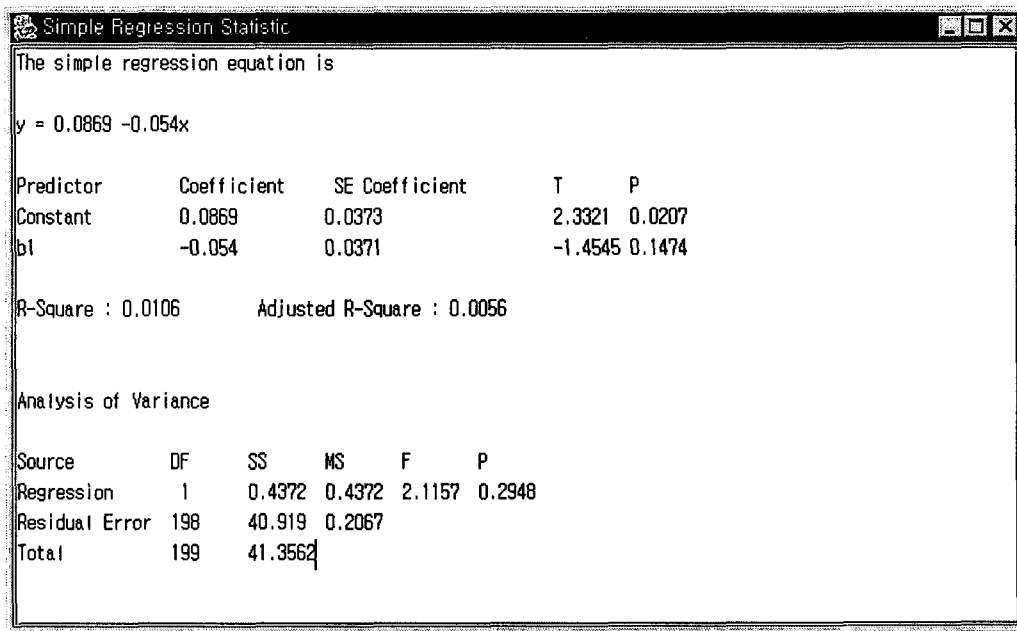


그림 3 단순회귀추정식 통계량

자료의 개수가 적어 매듭점이 추가되지 않는 경우 함수추정식과 다항회귀는 동일한 추정식을 보여 주어 두 모델을 비교할 수 없게 된다. 따라서 여기서 사용할 프로그램의 경우 자료의 크기가 상대적으로 큰 경우에 적용하는 분석방법이기 때문에 최초 자료의 입력은 ASCII 파일을 읽어서 실행되도록 하였다. 자료 파일은 "File" 메뉴에서 열 수 있으며 자료 파일을 열면 자료 편집창이 생긴다. 편집창에 있는 자료는 수정, 삭제 및 추가가 가능하도록 하였고 자료 변경 후 "Redraw" 버튼을 누르면 변경된 자료에 대한 동적 그래프와 통계량을 얻는다. 또한 편집창에 뜬 자료는 자료의 변경여부에 관계없이 파일로 저장할 수 있도록 하여 추후에 다시 사용할 수 있도록 하였다.

여기서 제공하는 분석방법은 모수가 많아 상대적으로 많은 자료를 필요로 하기 때문에 적절한 자료가 없는 사용자를 위하여 자료를 생성하는 기능을 제공하였다. 이 기능은 "Demo Data→

Random data from special function” 메뉴를 선택하면 200개 자료가 얻어진다.

자료 (x_i, y_i) 는

$$f(x) = \begin{cases} \sin(2\pi x^2) & \text{if } x \in [0, 1] \text{ 일 때} \\ 0 & \text{if } x \in [-1, 0] \text{ 또는 } x \in [1, 2] \end{cases} \quad (7)$$

에서 난수를 발생한 함수

$$y_i = f\left(-1 + \frac{3}{n}(i-1)\right) + \varepsilon_i, \quad \varepsilon_i \stackrel{iid}{\sim} N(0, 1/16), \quad i=1, \dots, n \quad (8)$$

으로 얻은 자료이다. 난수를 사용하기 때문에 자료생성을 할 때마다 새로운 자료가 만들어지며 새 자료에 대한 분석을 얻는다. 식 (7)과 (8)을 이용하여 프로그램 내부에서 생성된 자료로 단순회귀, 다항회귀, 함수추정식에 대한 그래프를 비교한 결과는 그림 2에 있으며, 이 그림에서의 각 그래프는

- 그림의 ①은 식(5)에 있는 그림
- 그림의 ②는 식(6)에 대한 단순회귀추정식
- 그림의 ③은 식(7)에 대한 차수가 7인 다항회귀추정식
- 그림의 ④는 식(8)에 대한 함수추정식

이다. 그림 2는 단순회귀 < 다항회귀 < 함수추정식 순으로 추정회귀식이 잘 적합된 것을 보여준다. 그림 3은 단순회귀추정식에 대한 통계량으로 수정 결정계수가 0.0056로 이 자료를 잘 설명해 주지 못한다.

4. 프로그램 사용방법

이 프로그램은 시작할 때 그림을 출력하는 곳과 자료를 입력할 수 있는 부분이 하나로 되어 있는 창이 뜬다. 자료는 반응변수와 설명변수만 입력이 되도록 하였다. 첫 번째 입력하는 값이 설명변수이고 두 번째 입력하는 값이 반응변수이다. 자료를 화면에서 직접 입력을 할 수도 있고, 저장된 파일로부터 ASCII 파일을 읽어 올 수 있도록 제작되었다. “File→Open” 메뉴를 선택하면 파일 탐색기가 떠서 데이터 파일을 찾을 수 있다. 이때 주의할 점은 클라이언트의 리소스를 이용하기 위해서는 반드시 서명 인증을 해 주어야 하며 그렇지 않은 경우에는 프로그램이 클라이언트 컴퓨터에 접근할 수 없기 때문에 파일 읽기, 파일저장, 프린트 사용 등과 같은 기능을 수행할 수 없게 된다.

프로그램 사용 방법에 대하여 알아보자.

4.1 상관과 회귀 및 산점도 사용방법

- Data : 자료 입력의 시작 방법과 인쇄, 저장의 서브메뉴가 존재한다.
 - Mouse : 마우스 클릭으로 자료를 입력하기 시작한다. 자료 편집창에서 자료의 수정, 추가, 삭제가 가능하다.
 - Input : 마우스 클릭으로 자료를 생성하는 메뉴이다.
 - Save Points : 마우스 클릭으로 생성된 자료를 저장하는 메뉴이다.
 - File : 자료 파일에서 자료를 읽거나 저장하는 메뉴이다. 파일을 읽은 후 마우스를 이용하여 자료의 추가, 삭제, 수정이 가능하다.
 - Open : 읽을 자료 파일을 지정하는 메뉴이다.
 - Save : 자료를 파일로 저장하는 메뉴이다.
 - Print : 동적 그림 화면을 프린터로 출력하는 메뉴이다.
 - Exit : 프로그램을 종료한다.
- Draw : 회귀직선, 신뢰대, 예측구간 관련 메뉴이다.
 - Line : 회귀직선을 그리는 메뉴이다.
 - Confidence Band : 신뢰대 $\hat{E}(Y|X=x)$ 를 그리는 메뉴이다.
 - Prediction Band : 예측구간 $(\hat{Y}|X=x)$ 를 그리는 메뉴이다.
- Statistic : 관련 통계량을 출력하는 메뉴이다.
 - Anova : 분산분석표, 계수의 추정 및 검정 p 값 등을 출력하는 메뉴이다.

4.2 회귀분석 모형비교 사용방법

- File : 자료의 읽기, 저장, 종료 등을 제공하는 메뉴로 다음의 서브메뉴가 있다.
 - New : 새 문서로 시작한다. 전에 내용이 존재하면 모든 내용을 지운다.
 - Open : 기존에 존재하는 파일을 열기. 기존에 파일이 존재하면 모든 내용을 지운다.
 - Save : 새 문서에 추가한 내용이나 파일 열기 생성된 문서를 저장한다.
 - Exit : 모든 프로그램을 종료한다.
- Demo Data : 자료를 생성하는 메뉴이다.
 - Random data from special function : 이 메뉴는 자료가 없는 경우 또는 데모용으로 만들어진 메뉴로 난수의 생성은 식 (8)을 이용하였다. 여기서 생성되는 자료의 개수는 200개이며 난수를 사용하므로 이 메뉴를 선택할 때마다 새로운 자료가 만들어진다.
- Execute : 여기에는 자료 편집창에 있는 자료를 이용하는 경우는 네 개, 자료를 생성한 경우에는 다섯 개의 서브메뉴를 가진다. 문자를 입력하면 실행하지 않고 에러가 발생한다.
 - Scatter : 산점도를 그린다.
 - Draw Real Function : 자료를 생성한 경우에만 나오는 메뉴로 추정의 대상이 되는 함수의 그림을 그린다. 식 (7)이 그려진다.
 - Simple Regression : 단순회귀의 그림 및 통계량을 출력한다.

- Draw Simple Regression : 단순회귀의 그림 출력한다.
- Simple Regression Statistic : 단순회귀 통계량 출력한다.
- Polynomial Regression : 다항회귀의 그림 및 통계량을 출력한다. 프로그램 실행은 동적 그림과 통계량 중 한 메뉴를 선택하면 디자인 행렬 차수를 입력하여 실행 되도록 하였다. 차수를 입력하면 선택한 메뉴에 대한 결과가 출력된다. 다른 차수를 입력하려면 두 메뉴 모두 선택을 취소 한 후 다시 두 메뉴중 한 메뉴를 선택하면 차수를 입력할 수 있다.
- Draw Polynomial Regression : 다항회귀 그림 출력.
- Polynomial Regression Statistic : 다항회귀 그림 통계량 출력.
- Cubic Spline : 함수추정식의 그림, 매듭점 및 통계량을 출력한다.
- Draw Cubic Spline : 함수추정식의 그림 출력한다.
- Cubic Spline Statistic : 함수추정식의 통계량 출력한다.
- Check the knot point : 매듭점을 그린다. 함수추정식 그림이 출력되어야만 실행할 수 있다.

4. 결론

본 고에서는 상관관계, 회귀분석 또는 스플라인을 처음 접하는 사람들이 보다 쉽게 그 내용을 이해할 수 있도록 동적 그림을 위주로 분석 결과를 제공하였다. 이러한 연구는 추후 이미 제공된 이상치와 영향치에 대한 통계량을 포함한 더 많은 통계량을 제시하고 이러한 분석을 다변량일 경우까지 확대할 수 있을 것이다.

참고문헌

- [1] 강희모, 심송용 (2000), *Implementation of Estimation and Inference on the Web*, 「한국통계학회논문집」, 제 7권 제3호, pp. 913-926.
- [2] 구자용 (2000), 기저함수 방법론, 「2000년 한국통계학회 추계 학술논문발표회 초청강좌」.
- [3] 심송용 (1997), 인터넷을 이용한 원격수업, 「1997년 한국통계학회 추계 학술 발표회 논문집」, pp. 1-5.
- [4] 한경수, 안정용, 강윤비 (1998), 통계학 교육을 위한 전자교재의 활용, 「응용통계연구」 제 11권 1호, pp. 5-12.
- [5] Arnold, K., Gosling, J. and Holmes, D. (2000), *The Java Programming Language Third Edition*, Addison-Wesley Pub Co.
- [6] Belsley, D. A., Kuh, E., Welsch, R. E. (1980), *Regression Diagnostics*, Wiley, New York.
- [7] Friedman, J.H. and Silverman, B.W. (1989), Flexible Parsimonious Smoothing and Additive Modeling(with discussion), *Techometrics*, Vol. 31, 3-39.
- [8] Hastie, T.J. and Tibshirani, R.J. (1990), *Generalized Additive Models*, Chapman and Hall.

- [9] Java Web Start 홈페이지, <http://java.sun.com/products/javawebstart/>.
- [10] Kennedy, W.J. and Gentle, J.E. (1980), *Statistical Computing*, Marcel Dekker Inc.
- [11] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992), *Numerical Recipes in C*, Cambridge Univ. Press.

[2003년 5월 접수, 2003년 9월 채택]