

마이크로 컴퓨터 상에서의 시계열 분석 통계 패키지의 활용에 관한 연구

왕 속 희 백 두 권
고려대학교 전산학과

A Study On The Application Of Time Series Analysis
Statistical Package Based On Micro-Computer

< 요약 >

본 논문에서는 시계열 자료 분석을 위한 마이크로 컴퓨터 용 통계 패키지인 TIMESLAB 을 소개하였다. TIMESLAB은 종래의 통계 분석용 패키지인 SPSS, SAS, TSP 등이 대형 컴퓨터 시스템에 맞도록 설계되어 있어 PC에서 활용하기에는 많은 디스크 용량을 필요로 하는 등의 문제점을 모두 극복한 PC용 시계열 자료 분석 프로그램이다.

TIMESLAB은 대화용으로 설계되어 있고 아주 적은 하드 디스크 용량만 가지고서도 쉽게 접근이 가능한 편리한 시스템으로 그 구성, 활용 및 명령어들을 제시하였고, 예제 프로그램과 그 결과를 제시하였다.

1. 서론

TIMESLAB(Time Series Analysis Laboratory)프로그램은 널리 보급된 퍼스널 컴퓨터의 활용을 높이고 대화식으로 통계 분석의 결과를 받아 볼 수 있으며 높은 질의 그림(high resolution graphics)을 저장, 출력 할 수 있도록 설계된 시계열 분석, 통계 패키지이다.

어떤 IBM 호환 가능 PC에서도 사용할 수 있으며 다른 어떤 소프트웨어도 요구되지 않으며 내부명령어를 독자적으로 갖고 있다. 사용자는 매크로와 자료화일을 마음대로 생성해서 쓸 수 있게 설계되어 있으며 TIMESLAB 명령어는 시간과 빈도를 도메인으로 사용한 많은 분석을 시도 할 수 있고 수학적 통계학적 다양성을 모두 수용하고 있다.

TIMESLAB 은 그 자신의 EDITOR를 갖고 있으며 고해상도 그래픽 명령어와 화일로 부터 읽고 쓸 수 있는 능력, 또한 프로그램에서 빠져 나오지 않은 상태에서 DOS 명령어로 돌아 올 수 있는 능력을 갖춘 거의 완전한 통합 시스템이다. 텍스트와 그래픽 출력이 모두 쉽게 프린트에서 merge 될 수 있으며 그래픽 스크린 전체를 저장했다가 다시 화일로 불러올 수 있는 기능을 갖추고 있다. 또한 매크로(macro) 기능을 이용해서 여러 명령어를 묶어 사용하는 화일처리 능력도 갖추고 있다.

본 논문에서는 TIMESLAB 의 HARDWARE요구사항, 구성, 기능 및 많이 사용되는 명령어와 사용 실례를 소개하기로 한다.

2. 하드웨어와 소프트웨어의 요구사항

- IBM PC (호환 가능 기종 포함), PC/XT, PC/AT.
- 512K 이상의 RAM

- CGA(Color Graphic Adaptor) 혹은 EGA(Enhanced Graphic Adaptor)
- 흑백 혹은 칼라 모니터
- 1 개 이상의 floppy 디스크 드라이브나 하드 디스크
- 출력을 위한 9핀 도트 매트릭스 프린터 혹은 24핀 프린터, 레이저 프린터 등
- DOS 2.0 이나 그 후 버전

3. TIMESLAB 의 구성 및 기능

두개의 환영 스크린이 있으며 TIMESLAB 프롬프트(prompt)와 140여개의 명령어 리스트가 있고 모든 명령어들에 대한 사용방법을 설명하는 도움말 (HELP)기능을 갖추고 있다.

- (1) 데이터 타입은 사용자가 정의하여 사용할 수 있다.
상수, 문자, 정수변수, 실수변수, 실수배열이 있고 실수는 10 에서 10 까지 사용이 가능하다.
- (2) 명령어 타입
명령어는 형태와 용도에 따라 분류할 수 있는데 형태에 의해 분류해 보면 5가지가 있다.

제 1 타입 명령어 : 명령어 이름 (인수의 나열)
`PLOT(x,n)`
`TEXTCOLOR(0,2)`
`READ(air.dat,x,n)`

제 2 타입 명령어 : 할당문과 같은 명령어들로
`n = 100`
`Z = X+Y`
 등

제 3 타입 명령어 : 변수명 = 명령어 (인수의 나열)
`x = wn(sead, n)`
`y = sort (x,n)`

제 4 타입 명령어 `x = <1,2,3,4,5,>` 와 같이 숫자나 배열을 연결 시킬 수 있다.
`x = <x,5,43,z,y,6.2>`

예) `A = <1,2,3,4,5,6>`
`B = <11,12,13,14>`
`C = < A, B >`

`A = TRANS (A,3,2)`
`B = TRANS (B,2,2)`
`C = < A, B >`
`D = TRANS (C,2,5)`

$A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$ $B = \begin{bmatrix} 11 & 13 \\ 12 & 14 \end{bmatrix}$

$C = \begin{bmatrix} 1 & 3 & 5 & 11 & 13 \\ 2 & 4 & 6 & 12 & 14 \end{bmatrix}$

제 5 타입 명령어 : 인수가 없이 이름에 의해서만 주어진다.

예> HELP

(3) 에러체킹

TIMESLAB은 명령어를 받아 들이면 타입을 결정한다. 명령어의 형태가 위에 열거한 5가지 중의 하나가 아니면 'unrecognizable command type' 메시지를 보여준다.

일단 입력 인수의 정의가 빠져 있으면 메시지를 보내준다. 매크로의 실행동안 에러를 만나면 매크로는 끝나게 된다.

(4) 결과의 출력

높은 해상도의 결과가 스크린에 보여지면 프린터에 각 픽셀 이미지를 출력 시킬 수 있다.

TIMESLAB의 출력을 위해서는 텍스트 모드에서 그래픽 모드로 전환시켜야 한다.

메뉴로서 그림을 출력할 것인지 이미지를 저장 할 것인지 선택할 수 있다.

그림의 크기 조정은 PLOT, PLOT2, PLOTK, HIST, PLOT CSP 명령어들을 사용하여 두가지 형태로 주어질 수 있다.

크기 조정 정보없이 명령어를 사용하면 원래 데이터의 최소치와 최대치를 가지고 그림을 그리고 크기 조정인수와 함께 사용하면 원하는 최소치와 최대치를 인수의 끝에 삽입 시켜서 그림의 크기를 확대, 축소시킬 수 있다.

```
X = WN (12345,100)
LABEL(X) = 'PLOT (X,100)'
PLOT(X,100)
LABEL(X) = 'PLOT (X,100,0,100,-4,4)'
PLOT(X,100,0,100,-4,4)
```

(5) TIMESLAB FULL SCREEN TEXT EDITOR

IBM의 전문 Editor system의 한 부분으로 사용할 수 있는 full screen editor로서 상대적으로 짧은 data나 매크로 화일을 위해 설계되었다.

(6) 텍스트와 그래픽 모드에서의 칼라

칼라 모니터를 사용할 경우 그래픽 모드에서 배경은 검은색을, 전면은 15가지 색 중의 하나를 선택할 수 있다. 시작은 Yellow-brown으로 된다.

텍스트모드에서 배경과 전면은 모두 8가지 색 중 선택할 수 있도록 설계되어 있다.

기본은 흰색과 청색으로 되어 있고 TEXTCOLOR 명령어를 사용해서 변경시킬 수 있다.

(가) 난수 생성 기능 (random number generator)

난수 생성 기능을 갖고 있으며 초기치를 임의로 줄 수 있다.
기본 초기치는 마지막 작업의 시뮬레이션 명령어에서 사용하였던 출력치를 사용한다.

(나) 각 장별 기록 유지

RECORD(file name) 에 의해, 사용하는 명령어의 기록을 유지하기 시작하고 RECORD(close) 에 의해 기록을 끝낸다.
이 화일을 기록했다가 인쇄 할 수도 있다.

4. TIMESLAB 의 주요 명령어들

다음에 주요 명령어 그룹과 명령어들을 소개함으로써 TIMESLAB의 사용범위를 이해하는 것을 도울 수 있을 것으로 생각된다.

COLOR : TEXTCOLOR 에 의해 화면의 배경과 전면의 색을 사용자 마음대로 선택할 수 있다.

CORRELATIONS: 자료의 상관관계를 나타내는 명령어들

ARCORR - AR 매개변수에 대응하는 상관계수를 찾을 수 있다.

ARCORR - bivariat AR 매개변수에 대응하는 상관계수를 찾을 수 있다.

ARPART - AR 매개변수에 대응하는 부분 상관

ARMACORR - ARMA 매개변수에 대응하는 상관계수를 찾을 수 있다.

CORR - 자기상관계수를 찾을 수 있다.

PARTAR - AR 매개변수로 부터 부분 상관을 찾을 수 있다.

DISTRIBUTION : 데이터의 분포를 연구할 수 있는 여러 명령어들이다.

BINOM - 이항 분포의 확률

DENSITY - 확률밀도함수를 측정

DIST-확률분포함수, 누적확률분포함수

Z-Test, T-Test, chi-square, F- 분포

HIST - 히스토그램을 그림

INFORT - 원타일 함수를 그림

DOS : TIMESLAB 에서 DOS로 빠져 나올 수 있다.

EDITOR : 내부의 built-in 텍스트 에디터를 사용해서 매크로화일이나 데이터를 수정할 수 있다.

ENVIRONMENT: TIMESLAB 시스템의 환경으로서 제공되는 명령어들

CLEAN : 변수들을 지움

CLS : 화면을 클리어 시킴

ECHO : 스크린을프린터로 출력시킴

PRINTER : 프린터에 명령어를 만듦.

RECORD : 세션(session)을 화일에 기록하여 보유함

TEXTCOLOR : 텍스트모드에서 색깔을 고정 시킴

ESTIMATION : 추정할 수 있는 명령어들

ARMASEL : ARMA 에 피팅(fitting)시키거나 추정할 수 있음.

ARSPCB : AR 스펙트라를 위한 신뢰구간을 찾을 수 있음.

DTARMA - ARMA 매개변수를 위한 최우추정

GRAPHICS: 고해상도 (high resolution) 그래프를 출력할 수 있는 명령어들

FIND : interactive location labeler utility 를 부를 수 있음.

LABEL : 그림의 제목을 붙일 때 사용

HIST : 히스토그램을 그림

PLOT : X 축대 Y 축의 그래프를 그림

PLOT2 : 2Y 축대 X 축의 그림을 그림 (Y축을 두배로 확대)

PLOTK : KY 축대 X 축의 그림을 그림(Y축을 K 배로 확대)

RESCREEN : 저장된 그래픽 스크린을 검색한다.

SAVESC : 그래픽 스크린을 파일로 저장 시킨다.

MACROS : 매크로 명령어에 의해 여러 명령어들을 파일로 만들어 실행시킬 수 있는 명령어들

PAUSE, PROMPTOFF, PROMPTON, SPEAKOFF, SINGLEOFF, TIME 등이 있다.

MATH : 수학적 작업을 수행하기 위한 명령어들

TYPE2 : 가감승제와 같은 유너리나 바이너리 오퍼레이션을 수행한다.

ABS : 절대치 함수

ARDT : difference 방정식의 값을 생성

COS : COSINE 함수

CUM : 누적합계나 배열의 평균을 구한다.

DIST : Pdf, Cdf, 쿼타일, Z, T, Chi-square, F 분포

EXP : 지수함수

POLY : 폴리노미알의 분석

SORT : 배열의 정렬

MATRICES : 벡터나 행렬의 오퍼레이션을 수행하는 명령어들

DOUBLE: 배열의 확장

EIG: Eigenvalue 와 Eigenvector 를 찾음.

EXTRACT : 배열의 일부를 추출

GS : 그람 - 슈미트 행렬의 분할

LISTM : 3 차원 배열의 원소를 나열

MCHOL : Cholesky 행렬 분할

MINV : 역 행렬

MMULT: 행렬의 곱 계산

REVERSE : 행렬의 sweep

TOEPL : Toeplitz 행렬을 생성

TRANS : 전치 행렬을 구함

TYPE4 : 행렬의 연결 (concatenation)

MODELS : 상관관계, 추정, 예측, 시뮬레이션, 스펙트라 등의 모델을 연구할 수 있는 명령어들이 있다.

CORRELATIONS, ESTIMATION, PREDICTION, SIMULATION, SPECTRA 등

POLYNOMIALS : 다항식에 관한 명령어들

INVPLOY : 다항식에 고나한 연구 다항식의 역수의 계수를 찾음

MULTPOLY : 두 다항식의 곱의 계수를 찾음

POLY : 다항식의 평가

POLROOTS: 주어진 계수를 갖고 다항식의 근을 구함

PREDICTION : 예측을 하기 위한 명령어

ARMAPRED : ARMA 모델을 위한 예측

SEAPRED : BOX-JENKINS 모델 예측

SIMULATION : 시뮬레이션을 이용하는 명령어들

AROT : AR 프로세스부터의 자료 시뮬레이션

WN : White Noise 프로세스로부터의 자료 시뮬레이션

SPECTRA : 스펙트랄 분석을 위한 명령어

ARSP : AR 프로세스를 위한 스펙트라

CORR : 유니베리에이트 자료의 상관관계의 피리오도그램

FFT : Fast Fourier transform

WINDOW : 유니베리에이트 자료를 위한 스펙트랄 추정.

TRANSFORMS : 자료의 전환을 위한 명령어들

DIFF : 자료의 difference 를 구한다.

DIVSDS : 계절적 표준 편차를 구한다.

REG : 회귀분석

SORT : 배열의 정렬

EXP : 지수 함수 계산

WNTTESTS : White noise 를 테스트 하는 명령어

BARTEST : Bartlett's test

QTEST : Portmanteau test

5. TIMESLAB 의 사용 실례

명령어 리스트

ARDT2	CLEAN	DIST	HELP	MACRO	PLOT2	RECORD	SPEAKEROFF
ARFILT	CLS	DIVSDS	HELP	WADT	PLOTCSF	REG	SPEAKERON
ARNACORR	COEFFCSD	DOS	HELP1	WASP	PLOTK	REPLACE	SUBMNS
ARNADT	COLOR	DOT	HIST	MAXMIN	PLOTSF	RESCREEN	SWEEP
ARNAPRED	CORR	DOUBLE	IF	WCHOL	POLAR	RESTART	TEXTCOLOR
ARNASBL	CORR2	DTAR	INFO	WDEL	POLY	REVERSE	TIME
ARNASP	CORRAR	DTARMA	INQNT	WINV	POLYROOTS	ROOTSPOLY	TOEPL
ARPART	CORRAR2	DTFORE	INVPLY	WMULT	PRINT	SAVE	TRANS
ARSP	CORRARMA	ECHO	LABEL	MULTPOLY	PRINTER	SAVBSC	WINDOW
ARSP2	CORRMA	EDIT	LENGTH	NOTBS	PRINTSEL	SBASEST	WN
ARSPCB	COS	BIG	LINE	OVEROFF	PROMPTOFF		

For help on a command,type HBLP(name)

For a list of general help areas,type HBLP(areas)

?;macro sample1

?;generate and superimpose the plots of length n

?;period p1 & p2 for user specified values p1 & p2

?n=100

?a=1

?p1=10

?p2=10

?y1=cos(100,1,10)

?y2=cos(100,1,20)

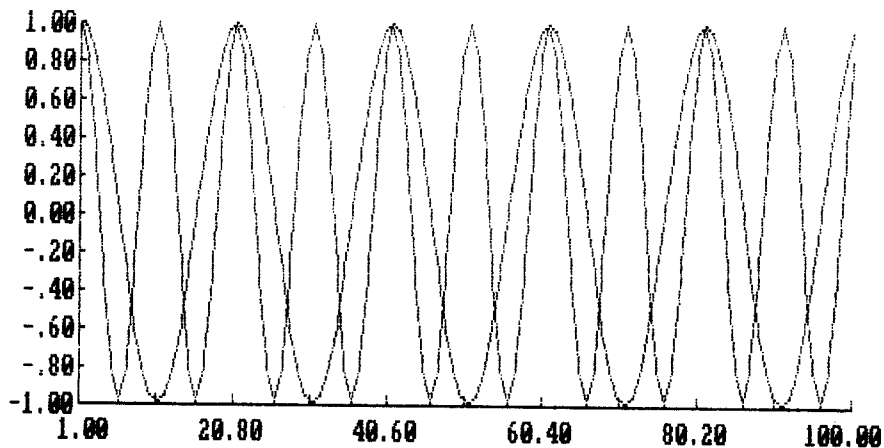
?label(y1)='plot the cosine of len.=100,per.=10'

?label(y2)='plot the cosine of len.=100,per.=20'

?plot2(y1,y2,100,100,1,1)

그림

plot the cosine of len.=100,per.=20
plot the cosine of len.=100,per.=10



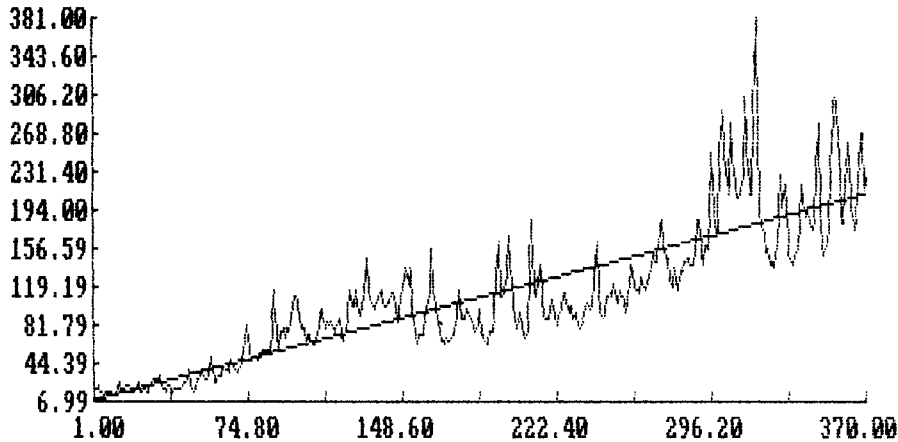
```

beta[1]      =      6.44450900 beta[2]      =      .54680300
rss          = 404676.80000000
?list(beta,2)
Regression Coefficients
   1 |      6.444509      .546803
?yhat=line(n,beta[1],beta[2])
?plot2(y,yhat,n,n,1,1)

```

회귀분석 실행

straight line
Beveridge Wheat Price Index, 1500-1869



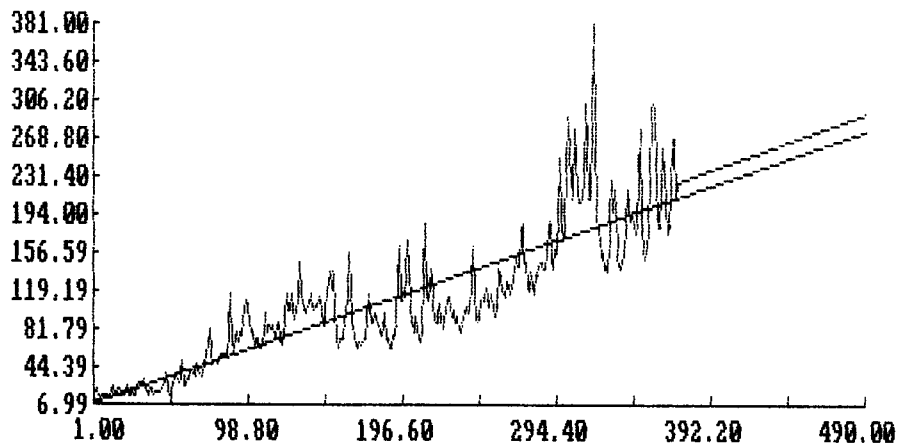
```

2J?yext=extend(y,n,120,1)
?n1=n+120
?YHAT=LINE(N1,BETA[1],BETA[2])
N1          not a valid integer
?n1=n+120
?yext=extend(y,n,120,1)
?yhat=line(n1,beta[1],beta[2])
?plot2(yext,yhat,n1,n1,1,1)

```

자료예측 실행

straight line
extended array



```

2J?read(bev,370,n)
Variable names must start with a-z or A-Z
?read(bev,n,n)
?lx=log(x)

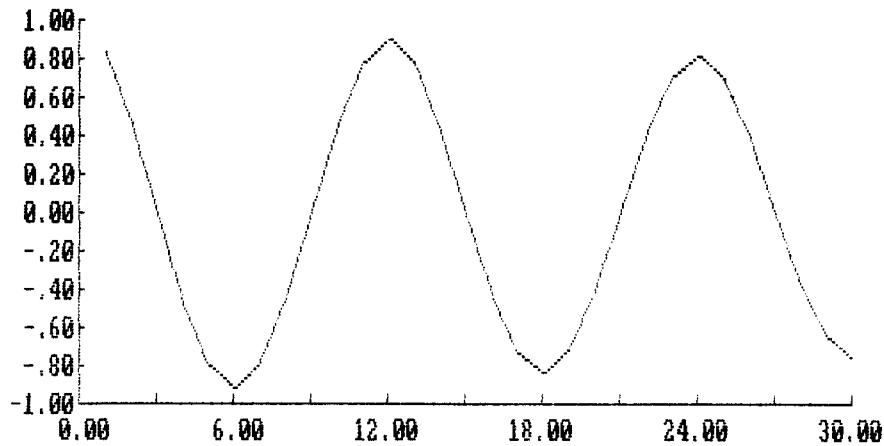
```



```
2J?rho=corr(x,n,30,0,1,r0,per)
```

```
?plot(rho,30,0,30,-1,1)
```

correlogram



```
2J?n=144
```

```
?m=30
```

```
?macro(wntest)
```

```
?;;
```

```
?;; WNTTEST.MAC: macro to find 95% simultaneous confidence
```

```
?;; bands for the correlogram and cumulative
```

```
?;; periodogram under the hypothesis of white noise.
```

```
?;;
```

```
?;; Input: x,n,m (data,sample size, and number of correlations)
```

```
?;;
```

```
?;; Output: rho,R0,per,cper
```

```
?;;
```

```
?PAUSE
```

```
PAUSE...STRIKE F9 TO BREAK, ANYTHING ELSE TO GO ON
```

```
?rho=CORR(x,n,m,0,1,r0,per)
```

```
?xx=LINE(m,0,1)
```

```
?xx=<xx,1,m,1,m>
```

```
?sqn:=n
```

```
?sqn=sqn^.5
```

```
?om:=m
```

```
?oom=1/om
```

```
?u=.95^oom
```

```
?u=1+u
```

```
?u=u/2
```

```
?zalph=DIST(z,3,1,u,u)
```

```
?lim2=zalph/sqn
```

```
?lim1=-lim2
```

```
?yy=<rho,lim1,lim1,lim2,lim2>
```

```
?nn=<m,2,2>
```

```
?type=<2,2,2>
```

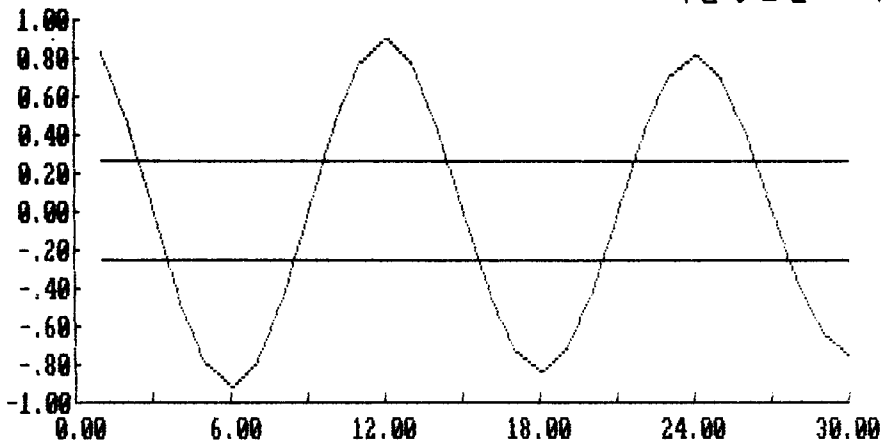
```
?LABEL(xx)=<x>
```

```
?LABEL(yy)='Correlogram and 95% Simultaneous Bands'
```

```
?PLOTK(xx,yy,nn,3,type,0,m,-1,1)
```

Correlogram and 95% Simultaneous Bands
Average Monthly Temperatures in Dubuq

계절평균을 고려한 자기상관 그래프

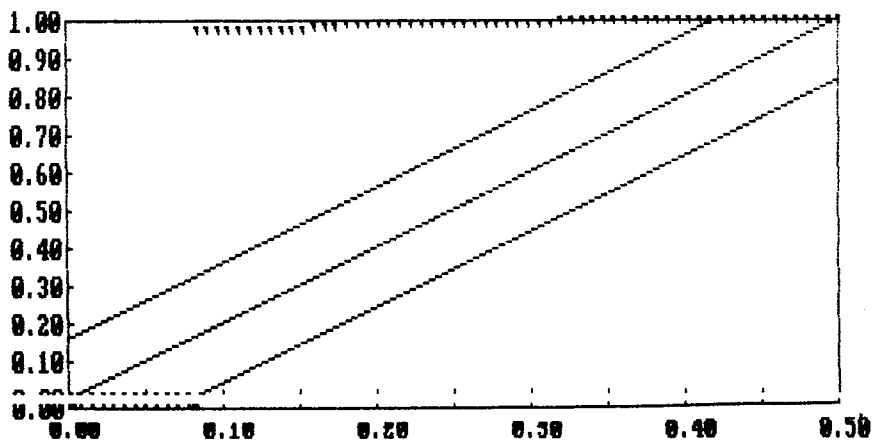


2J

```
?CLEAN(sq,om,oom,u,zalph,lim1,lim2,xx,yy,nn,type)
?rho=CORR(x,n,0,n,1,r0,per)
?cper=CUMSP(per,n)
?q=n/2
?q=q+1
?alpha=<0,1>
?freq=POLY(alpha,1,q,0,.5)
?sq:=q
?sq=sq^.5
?sq1=1.36/sq
?sq2=1-sq1
?sq3=.5*sq2
?sq4=.5*sq1
?xx=<freq,0,.5,0,sq3,sq4,.5,0,.5,.5,.5>
?yy=<cper,0,1,sq1,1,0,sq2,1,1,1,0>
?nn=<q,2,2,2,2,2>
?type=<34,2,2,2,2,2>
?LABEL(xx)=<x>
?LABEL(yy)='Cumulative Periodogram and 95% Bands'
?PLOTK(xx,yy,nn,6,type)
```

Cummlative Periodogram 실례

Cumulative Periodogram and 95% Bands
Average Monthly Temperatures in Dubuq



2J

6. 결론

1980년대 이후로 컴퓨터 기술의 혁신적인 발전으로 PC 와 Workstation의 가격이 많이 떨어져 그 보급이 급격히 확산되고 있다.

통계 패키지는 일반적으로 데이터의 편집, 데이터의 관리, 표 작성, 기술통계, 그리고 통계적 분석의 다섯가지 기본 요건을 갖추어야한다. 이러한 통계 패키지를 선택할 때에도 시스템에 대한 고려, 변수처리 능력, 편의성, 편집기능, 기본적 연산 기능, 그래픽 기능을 고려해야 할 것이며 이런 모든 면을 모두 만족시켜주는 TIMESLAB 패키지의 PC에서의 활용은 상당한 효과를 거둘 수 있을 것으로 본다. TIMESLAB은 특별히 한 장의 플로피 디스켓으로 PC에서 손쉽게 대화용으로 활용할 수 있으며 그래픽 기능이 강력하기 때문에 시계열 통계 분석을 하고자할 때 아주 성능이 좋으며 많은 활용이 있기를 바란다.

7. 참고문헌

김병천 (1989) "PC용 통계 패키지 활용과 전망"

양광민, 이장균 (1989) "마이크로 컴퓨터용 계량분석 소프트웨어의 개발"

Joseph Newton (1989) "TIMESLAB Application and Usage"

Kennedy, William J., Junior & Gentle, James E.(1980)
"Statiscal Computing"

J. S. Milton & Jesse C. Arnold (1986) "Probability and
Computing Sciences"

Lachenbruch, P. A.,(1983) "Statistical Programs for
Microcomputers", BYTE, pp.560-570

Neffendorf, H.,Cable, D., and Rowe, B. (1984) "Software for
Statistical Survey Analysis", Computational
Statistics and Data Analysis, Vol. 2, No. 1,
pp. 81-94