四．S-Plus Language

本章簡介 S-Plus 的基本指令，語法，程式設計及其所提供的 programming 環境，對於前述幾章提到的視窗式資料處理，統計分析及繪圖功能，均可利用 S-Plus Language 來做更進一步的控制及編輯，對於 S-Plus 未提供的視窗式功能及繪圖，使用者可依自已的需求設計程式。

1. Commands Window/指令及語法

Command Window 提供您一個互動式(interactive)的程式設計環境，對於資料分析，統計分析及繪圖具有強大的處理能力。

1.1 開啓 Commands Window

a. Window/Commands Window 或

b. Commands Window 按鈕

1.2 從 Commands Window 離開 S-Plus

> q( )

1.3 一些性質

Properties	例子
S-Plus is Case sensitive	Splus, splus, SPLUS
S-Plus ignores most space	>3* 6
Continuation	>3* + 6
中止執行	Esc 鍵
Assign operator	> x<- 3 或 > x_3
註解	#註解
錯誤訊息	> 1 2*1 Error in parse(text = txt): Syntax error: literal ("2") used illegally at this point : 1 2
指令記錄	↓ ↑鍵
Help	>?lm >help(lm) >help( ) >?methods(summary)
Data Object Names	Periods “.” : data.1, data.2 (data_1, data_2 錯誤命名)
Components 選擇子	$, person$name, person$salary
NULL	Null
NA	Missing values

1.4 S-Plus data 物件基本屬性(mode)

Numeric	12 c(2.1,3.4,1.0) salary.month
Complex	sqrt(as.complex(-2)) 8.659275e-017+1.414214i
Logical	c(T,T,F,F) salary.month>20000
Character	c(“month”,”salary”)

1.5 Data objects 資料物件

Vector	> x<- c(1,4,3)
Matrix	>mymatrix<- matrix(scan(“mydata.txt”),nrow=3, byrow=T)
	>as.matrix(x)
	>matrix(1:12,nrow=3,ncol=4)
Data Frame 一個 data frame 物件是由有欄資料和列資料組成，很像 matrix object，但 data frame 的欄位資料可允許不同 modes	> var1<-c("F","B","E","D") > var2<-c(5,11,10,12) > var3<-c(11,5,6,4) >myframe1<-data.frame(var1,var2,var3, + row.names=c("name1","name2","name3","name4")
	> state<-data.frame(state.x77) #export state to ascii file to see its format. >state$Life.Exp
	例:frame.txt 內容如下 5 11 small 11 5 small 10 6 meddle 8 12 large >myframe2<- data.frame(scan("frame.txt",what= + list(a=0,b=0,c=""))) #0:numeric, "":character >myframe$a
List 每一 list component 可以是任何 data object	>mylist<- list(item1=1:10,item2=c(“name1”,”name2”)) > mylist$item1 > mylist$item2
Time series	>myseries<- ts(1:12,start=c(1992,2),frequency=12)
Table	例:table.txt 內容如下 Names x y labels John 2 3 first Bill 3 7 second >mytable<- read.table(“table.txt”,header=T,row.names=

	+ ”labels”)

1.6 運算元

+ 加 - 減 * 乘 / 除	>2+1 >3-2 >3*4 >4/5
^ 指數	>3^2
: 整數生成序列運算子	>2:12
%% 餘數 %/% 相除取商的整數 %*% 矩陣相乘	>5%%2 >7%/%2 >matrix1%*%matrix2
Logical and comparison operators	= = equal to != not equal to <= less than or equal to >= greater than or equal to > greater than | vectorized or & vectorized and < less than ! not && control and || control or

1.7 匯入資料

a. >import.data(FileName=”test.sd2”, FileType=”SAS”, DataFrame=”Test”)
   >help(import.data)
b. >mydata1 <-scan()
c. >mydata2 <-scan(“c:\\test\\test.txt”)
d. >mydata3<-scan(what=””,sep=”\n”,n=1)

1.8 條件式

if (condition) expression1 else expression2
例: if (x>0) cat(“condition holds. \n”) else cat(“condition not holds. \n”)

1.9 Loop 迴圈

a. for(i in 1:100){ expression}
   例:> for(i in 1:5){cat("ok \n")}
b. while(condition) {expression}
   例: > i<-0
   > while(i<5){cat("ok\n")
   + i<-i+1}
c. repeat{ condition-expression break}
   例: >i<- 0
   > repeat{ cat("ok \n")
   + if(i>5) break
   + i<-i+1}

1.10 函數

>fix(FunctionName)　　#利用編輯器寫程式
>FunctionName()　　#執行程式

數學函數	abs 絕對值 exp 指數 log 或 log10 logarithm sqrt 平方根	>abs(-2) >exp(2) >log10(2) >sqrt(2)
統計函數	cor 相關係數 cumsum 向量元素累積和 mean 平均數 median 中位數 min 最小值 max 最大值 prod 向量元素積 quantile 分位數 range 最大，最小值 sample 抽樣 sum 和 var 變異數	>cor(x,y) >cumsum(x) >mean(x) >median(x) >min(x) >max(x) >prod(x) >quantile(x) >range(x) >sample(x) >sum(x) >var(x)
機率分配函數	Probability density (d) pdf 值 Probabilities (p) 機率值 quantiles (q) 機率分位值 random samples (r) 隨機樣本	>dnorm(x,mean=0,sd=1) >pnorm(x,mean=0,sd=1) >qnorm(p,mean=0.sd=1) >rnorm(5,mean=0,sd=1)
其它常用函數	length 長度 paste 附加上 dimnames 各維度欄位名稱 rev 反順序	>length(x) > paste("no.",1:10) >dimnames(iris) >rev(x)
繪圖 routines	abline 在圖上加線 identify 用滑鼠指出點 legend 圖線說明 lines 線圖 points 點圖 qqline qqplot 上加直線 text 圖上加文字 title 加 title par 分割區域	> abline(v = c(0, 10), lty = 2) >identify(x,y,z,cex=1) #help(identify) (在 x-y plot 上，用滑鼠指出 z 值) >legend(2,3,c(“Small”,”Medium”, + Large”), pch=”SML”) >lines(x,y) >point(x,y) >qqline(x) >text(x,y) >title(“test plot”) >par(mfg=c(1,1,3,2))

圖形參數	cex: 字元大小 lty: 線條型式 1=solid; >1=dots and/or dash pch: 繪圖的字元 pty: 點的型式 type: values of "p", "l", "b", "o", "n", "s" and "h" produce points, lines, both, both (overlaid), nothing, stairsteps and high-density lines. xlab, ylab : x,y 座標的說明

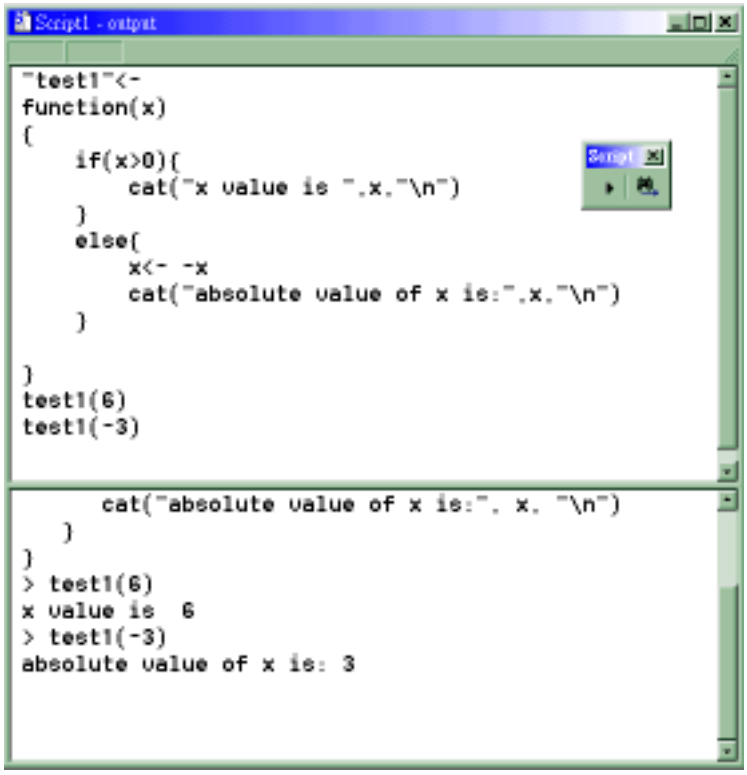
2. Script Window

Script Window 提供使用者設計及編輯程式，script 檔案副檔名爲*.ssc

例:

a. File/New/Script File

b.



3. Report Window

Report Window 是程式執行或統計分析後的 output 報表，可以直接在其上做編輯。

開啓:File/New/Report File

儲存:File/Save/可存成文字格式(.txt)或 Rich text format (.rtf 或.srp)

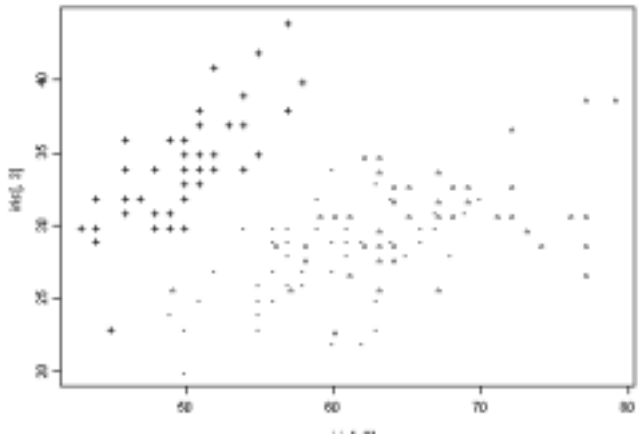
列印:File/Print Report

4. Advanced Graphics

4.1 a. File/Import Data/From files/exiris.sdd
b.Copy “Species”,”Sepal.L”,”Sepal.W” 存成 iris.txt
c.

```
>iris<-read.table(“iris.txt”)
>lab<-c(rep(“+”,50),rep(“-“,50),rep(“^”,50))
>plot(iris[,2],iris[,3],type=’n’)
>text(iris[,2],iris[,3],label=lab)
```

d. output



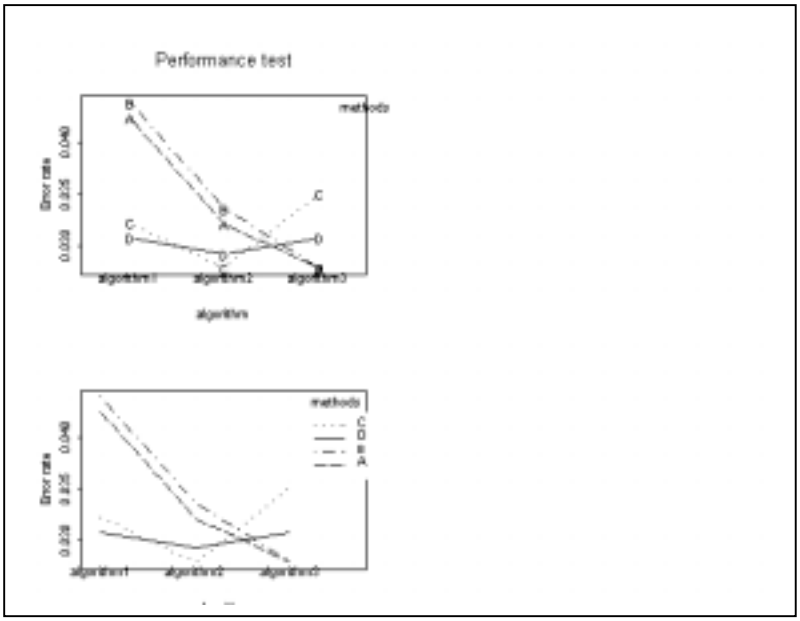
4.2 Two-Way Interaction Plots

a. File/New/Script File

b.

```
"interplot1"<-function()
{       qname <- list(algorithm = c("algorithm1", "algorithm2", "algorithm3"),
                        methods= c("A", "B", "C", "D"))
        options.design <- fac.design(c(3, 4), qname, rep = 1)
        error.rate <- c(0.0424, 0.0321, 0.0278, 0.0439, 0.0336, 0.0278,
                        0.0322, 0.0278, 0.0350, 0.0307, 0.0292, 0.0307)
        options.df <- data.frame(options.design, error.rate)
        par(mfcol = c(2,2), mai = c(0.8, 0.8, 0.8, 0.6))
        attach(options.df)
        interaction.plot(algorithm, methods, error.rate, main = "Performance test" ,
                        pch = "ABCD",type ="bbbb", trace.label = deparse(substitute(methods) ),
                        xlim = c(0.5, 3.5), ylab = "Error rate")
        interaction.plot(algorithm, methods, error.rate, ylab = "Error rate")
}
interplot1()
```

c. output


Performance test
methods
Error rate
algorithm
methods
C
D
B
A

5. Case Study

Youden Diagram

```
function(object=NULL, take.log=F, code.letter=F)
{
       if(length(object)==0){
              cat("please enter the data (three columns)\n Lab Codes,
                   First Sample & Second Sample\n")
              data<- matrix(scan(),ncol=3,byrow=T)
              cat("\n First sample name? \n")
              name1<- paste("Sample", scan(what="",sep="\n",n=1))
              cat("\n Second sample name? \n")
              name2<- paste("Sample", scan(what="",sep="\n",n=1))
       }
       else{
              data<- object
              name1<-dimnames(object)[[2]][2]
              name2<-dimnames(object)[[2]][3]}
       if(take.log){
              x1<-log10(data[,2])
              x2<-log10(data[,3])}
       else{
              x1<-data[,2]
              x2<-data[,3]}

       cat("\n Test title? \n")
       test<-scan(what="",sep="\n",n=1)

       if(code.letter){
              cat("\n Please enter the lab code letters: \n")
              lab <- scan(what="")}
       else lab<- data[,1]

       cat("\n Thank you. Now I'll perform the analysis ...\n \n")

              std.sum<-(x1+x2)/sqrt(2)
              if(median(x1)<median(x2))
                     std.diff<-(x2-x1)/sqrt(2)
              else
                     std.diff<-(x1-x2)/sqrt(2)

                     m.s<-median(std.sum)
                     m.d<-median(std.diff)
                     iqr.d<-iqr(std.diff)
                     iqr.s<-iqr(std.sum)

                     ff<-qnorm(0.75)-qnorm(0.25)
                     z.b<-((std.sum-m.s)*ff)/iqr.s
                     z.w<-((std.diff-m.d)*ff)/iqr.d

                     r<-sqrt(-2*log(2*qnorm(-2)))
                     x<--r+((2*r)/1000)*(0:1000)
                     yneg<- -sqrt(r^2-x^2)
                     ypos<-sqrt(r^2-x^2)

                     x<-c(x,rev(x)[2:1001])
                     y<-c(yneg,rev(ypos)[2:1001])

 if(median(x1)<median(x2)){
  e1<-((x*iqr.s)/ff+m.s-(y*iqr.d)/ff-m.d)/sqrt(2)
  e2<-((x*iqr.s)/ff+m.s+(y*iqr.d)/ff-m.d)/sqrt(2)
  }
else{
  e2<-((x*iqr.s)/ff+m.s-(y*iqr.d)/ff-m.d)/sqrt(2)
  e1<-((x*iqr.s)/ff+m.s+(y*iqr.d)/ff-m.d)/sqrt(2)
                     }
repeat{
```

```
 65           plot(x1,x2,xlab=name1,ylab=name2,xlim=c(min(c(x1,e1))
 66                ,max(c(x1,e1))),ylim=c(min(c(x1,e1)),max(c(x1,e1))))
 67           title(main=test)
 68           points(e1,e2,type="l")
 69           abline(v=median(x1),lty=5)
 70           abline(h=median(x2),lty=5)
 71           identify(x1,x2,lab,cex=0.75)
 72           cat("Re-Draw the Youden diagram?\n(y-yes, n-no)\n")
 73
 74           if(scan(what="",n=1) =="n") break
 75         }
 76
 77          cat("\n Print this Youden diagram\n(y-yes, n-no)\n")
 78
 79          if(scan(what="",n=1)=="y") dev.print()
 80
 81          cat("\n Extreme results: \n")
 82          output<- cbind(x1,x2,z.b,z.w)
 83          dimnames(output)<- list(lab,c(name1,name2,"between z-score",
 84                                  "within z-score"))
 85          return(output[(abs(z.b)>3 | abs(z.w)>3),])
 86
 87  }
```