R講義整理

林品寬

Outline

1. 基本程式碼整理
   +自創數列
   +基本運算
   +Vector組合
   +抽取元素
   +Vector運算
   +Simple patterned vectors
   .
2. Matrix,Array,Data frame
   +創造矩陣
   +Matrix inverse +創造Array
   +創造Data frame
   .
3. 繪圖整理
   +Bar charts and dot chart
   +Pie chart
   +Histograms
   +Boxplot
   +Scatterplots
   +QQplot
   +增加內容豐富度
   .
4. 統計模擬
   +Uniform
   +Set seed
   +Binomial random variables
   +Poisson
   +Exponential
   +Normal
   +Monte Carlo integration
   .
   .

Content

1.基本程式碼整理

a.自創數列

```
1:20
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```

.
.

b.基本運算

```
x=1:11
mean(x)
```

```
## [1] 6
```

```
var(x)
```

```
## [1] 11
```

```
sum(x)
```

```
## [1] 66
```

.
.

c.Vector組合

```
x=c(0,7,8)
number5to20=5:20
c(x,number5to20)
```

```
##  [1]  0  7  8  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```

.
.

d.抽取元素 抽取Vector中的元素

```
number5to20[1]
```

```
## [1] 5
```

```
number5to20[-2]
```

```
##  [1]  5  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```

```
number5to20[-(3:10)]
```

```
## [1]  5  6 15 16 17 18 19 20
```

```
number5to20[c(0,3:6)]
```

```
## [1]  7  8  9 10
```

.
.

e.Vector運算

```
number5to20*3
```

```
##  [1] 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60
```

```
number5to20%%2:3 #第一個除2，第二個除3，類推
```

```
##  [1] 1 0 1 2 1 1 1 0 1 2 1 1 1 0 1 2
```

.
.

f.Simple patterned vectors

```
seq(1,21,by=2)
```

```
##  [1]  1  3  5  7  9 11 13 15 17 19 21
```

```
rep(3,12) #repeat the value 3,12times
```

```
##  [1] 3 3 3 3 3 3 3 3 3 3 3 3
```

```
rep(seq(2,20,by=2),2) #repeat the pattern 2 4 ..20, twice
```

```
##  [1]  2  4  6  8 10 12 14 16 18 20  2  4  6  8 10 12 14 16 18 20
```

```
rep(c(1,4),c(3,2)) #repeat 1,three times and 4, twice
```

```
## [1] 1 1 1 4 4
```

```
rep(c(1,4),each=3) #repeat each value three times
```

```
## [1] 1 1 1 4 4 4
```

```
rep(seq(2,20,2),rep(2,10)) #repeat each value twice
```

```
##  [1]  2  2  4  4  6  6  8  8 10 10 12 12 14 14 16 16 18 18 20 20
```

.
/
.
.

2.Matrix,Array,Data frame

a.創造矩陣

```
m=matrix(1:6,nrow=2,ncol=3) #先塞行
m
```

```
##      [,1] [,2] [,3]
## [1,]    1    3    5
## [2,]    2    4    6
```

.
.
b.Matrix 計算
反矩陣:solve(h)
eigen(h) .
.
c.創造Array

```
a=array(1:24,c(3,4,2))
a
```

```
## , , 1
##
##      [,1] [,2] [,3] [,4]
## [1,]    1    4    7   10
## [2,]    2    5    8   11
## [3,]    3    6    9   12
##
## , , 2
##
##      [,1] [,2] [,3] [,4]
## [1,]   13   16   19   22
## [2,]   14   17   20   23
## [3,]   15   18   21   24
```

.
.
d.創造Data frame

```
color=c('red','yellow','blue')
number=c(1,2,3)
color.number=data.frame(color,number,more.number=c(4,5,6))
color.number
```

```
##    color number more.number
## 1    red      1           4
## 2 yellow      2           5
## 3   blue      3           6
```

.
/
.
.

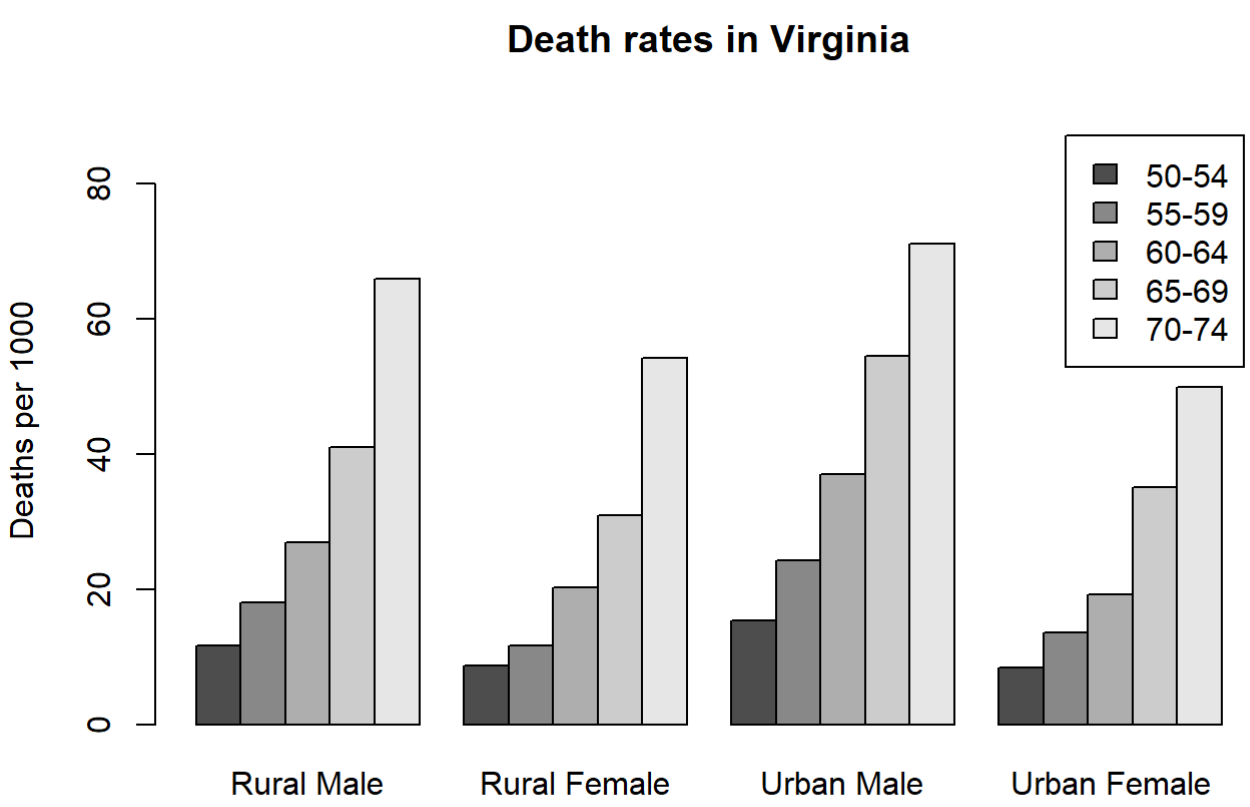
3.繪圖整理

a.Bar charts and dot chart 繪製簡單圖形

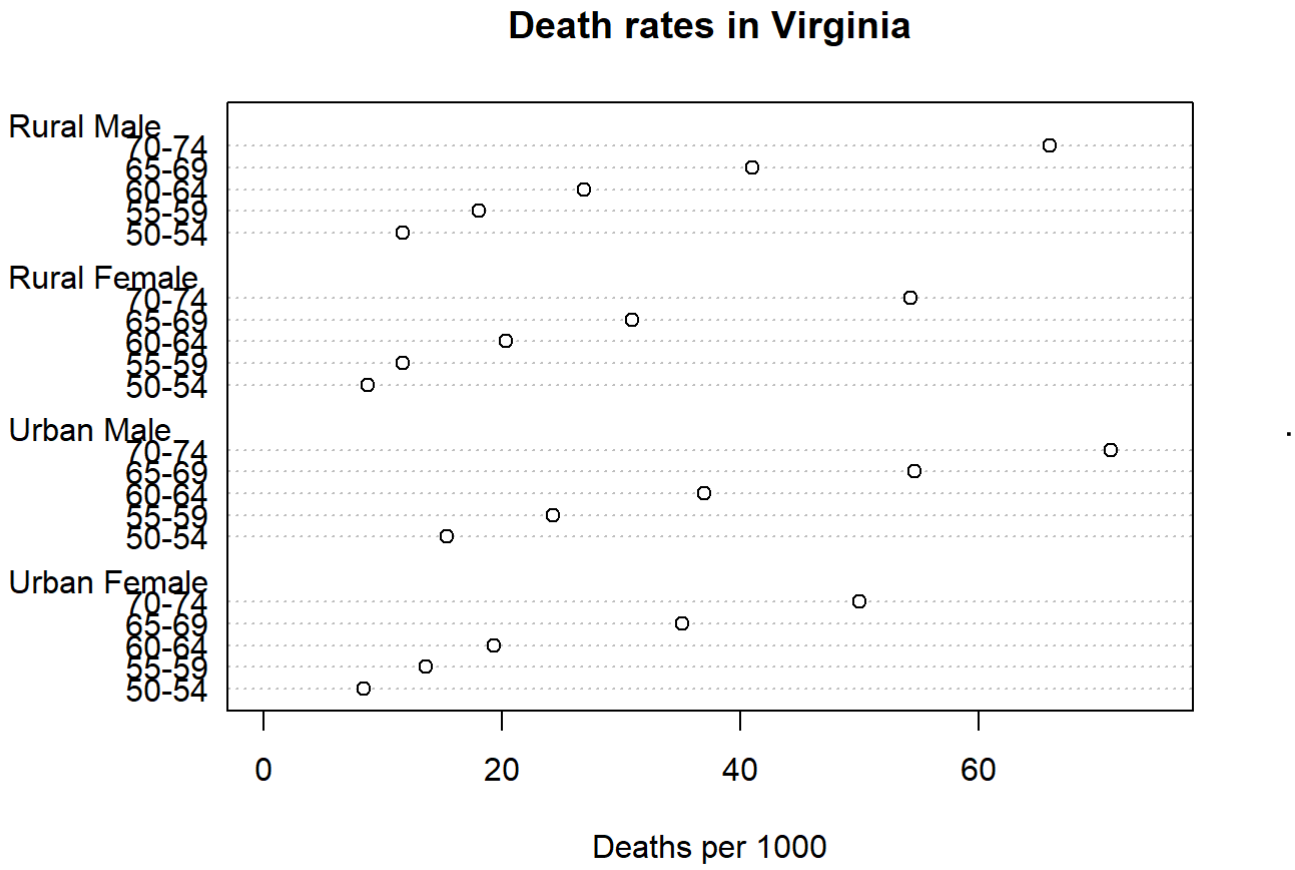
```
VADeaths
```

```
##       Rural Male Rural Female Urban Male Urban Female
## 50-54       11.7          8.7       15.4          8.4
## 55-59       18.1         11.7       24.3         13.6
## 60-64       26.9         20.3       37.0         19.3
## 65-69       41.0         30.9       54.6         35.1
## 70-74       66.0         54.3       71.1         50.0
```

```
barplot(VADeaths,beside=TRUE,legend=TRUE,ylim=c(0,90),
        ylab='Deaths per 1000',
        main='Death rates in Virginia')
```



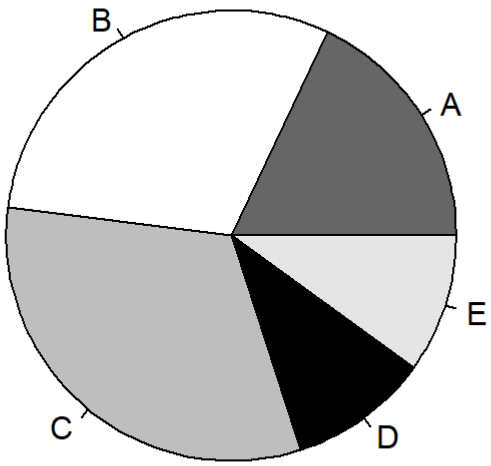

```
dotchart(VADeaths,xlim=c(0,75),
         xlab='Deaths per 1000',
         main='Death rates in Virginia')
```



.

b.Pie chart

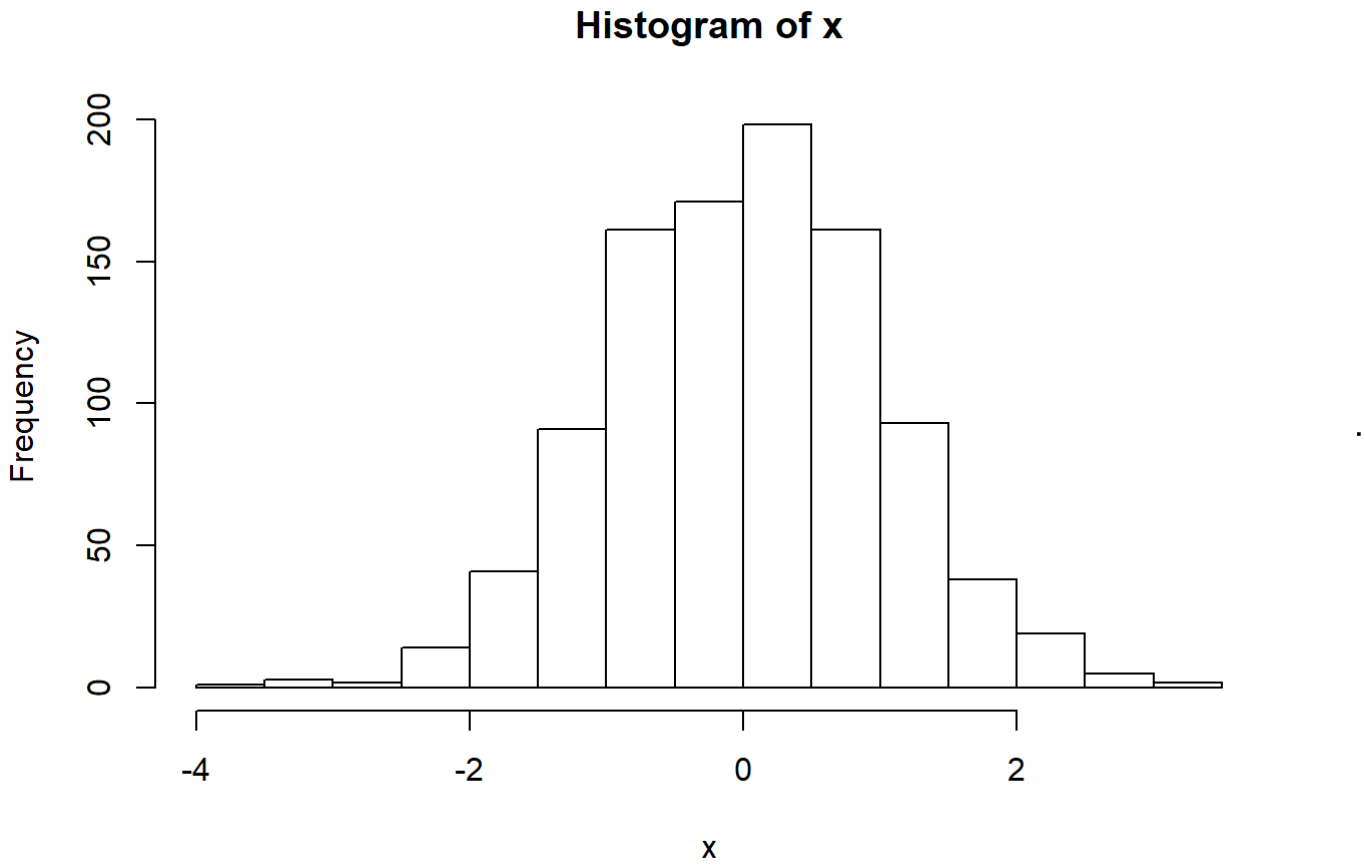
```
groupsizes=c(18,30,32,10,10)
labels=c('A','B','C','D','E')
pie(groupsizes,labels,col=c('grey40','white','grey','black','grey90'))
```



.

c.Histograms

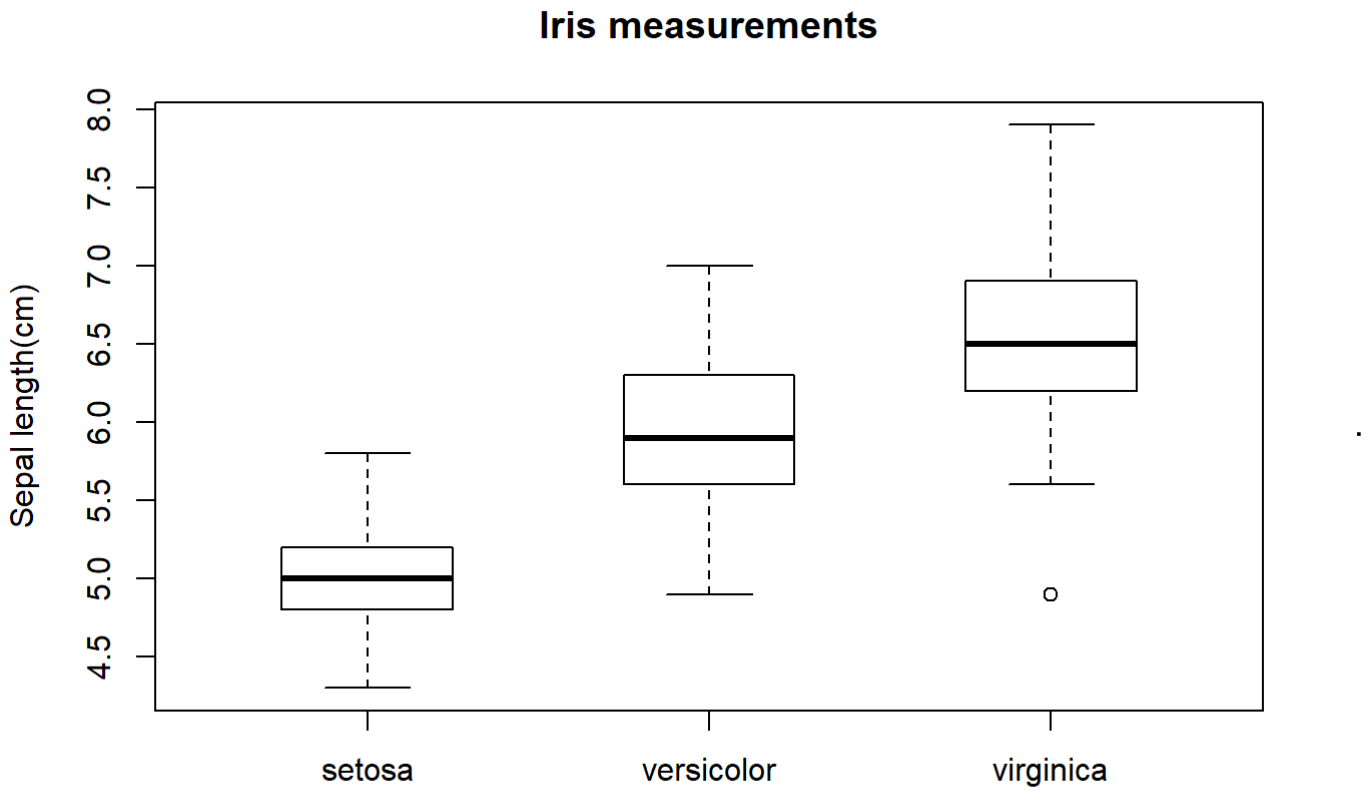
```
x=rnorm(1000)
hist(x)
```



.

d.Boxplot

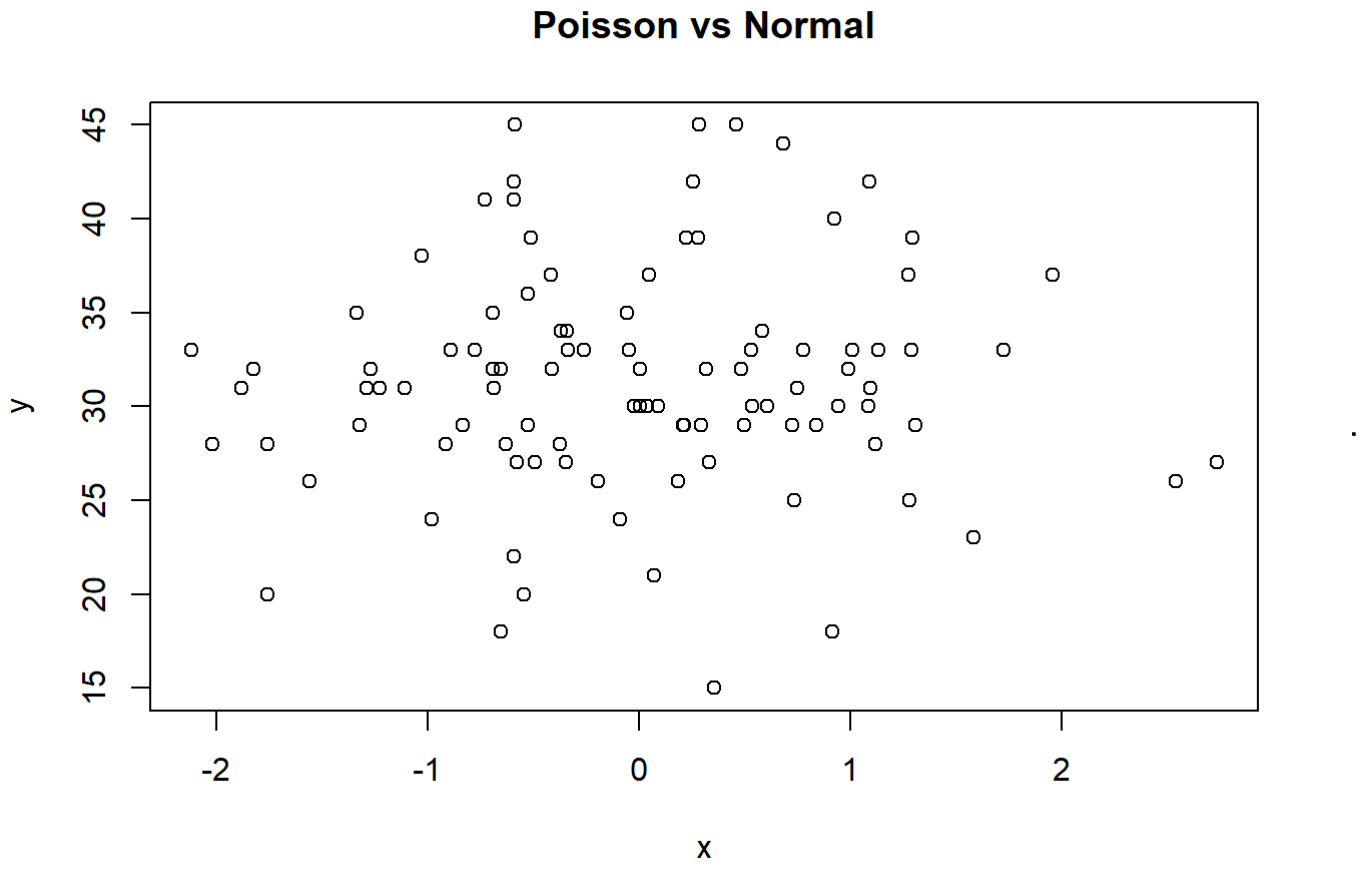
```
boxplot(Sepal.Length~Species,data=iris,ylab='Sepal length(cm)',main='Iris measurements',boxwe
x=0.5)
```



.

e.Scatterplots

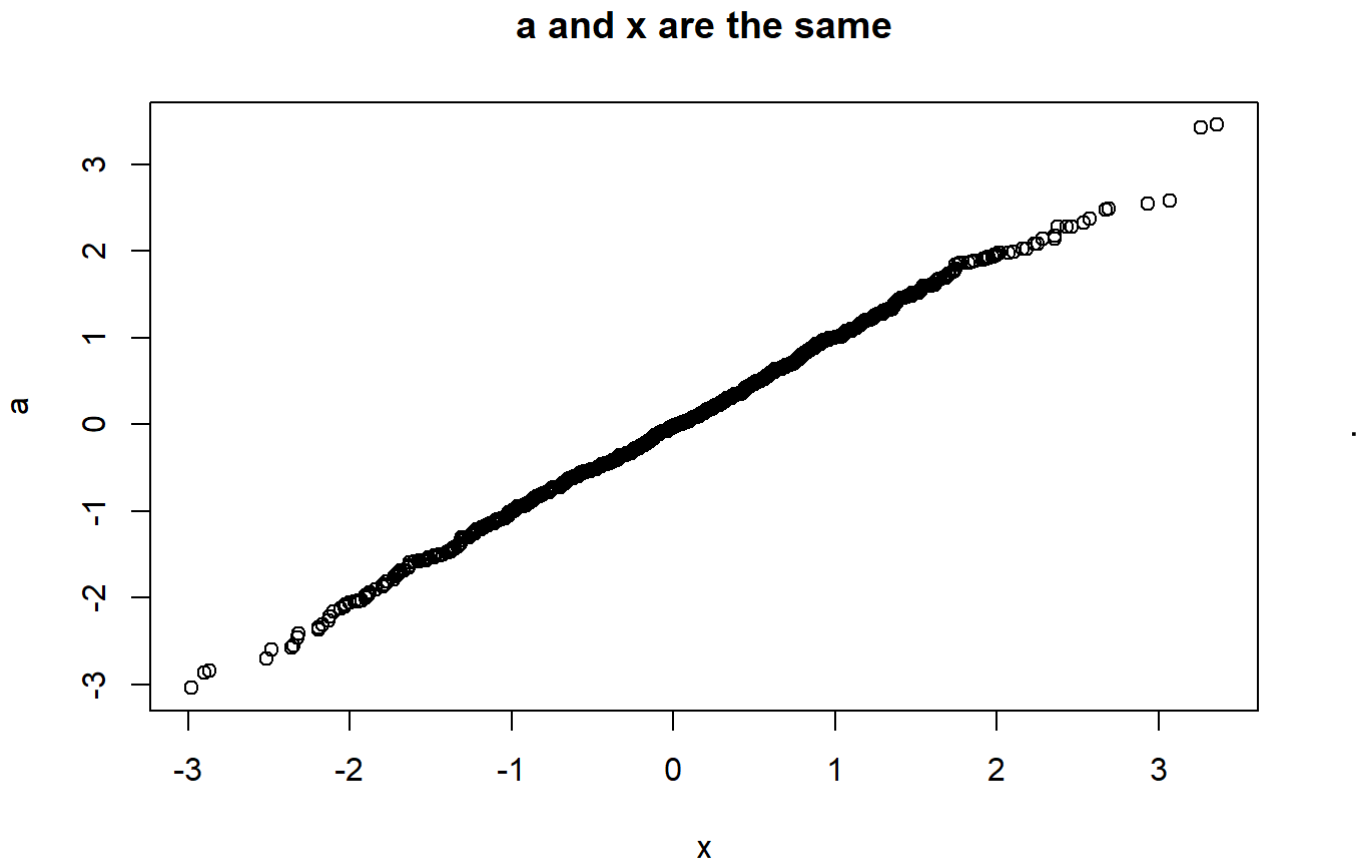
```
x=rnorm(100)
y=rpois(100,30)
plot(x,y,main='Poisson vs Normal')
```



.

f.QQplot

```
x=rnorm(1000)
a=rnorm(1000)
qqplot(x,a,main='a and x are the same')
```



.

g.增加內容豐富度
p(x,y,…)
lines(x,y,…) adds line segments
text(x,y,labels,…) adds text into the graph
abline(a,b,…) adds the line y=a+bx
abline(h=y,…) adds a horizontal line
abline(v=x,…) adds a vertical line
polygon(x,y…) adds a closed and possibly filled polgon
segments(x0,y0,x1,y1,…) draws line segments
arrows(x0,y0,x1,y1,…) draws arrows
symbols(x,y,…) draws circles, squares, thermometers, etc.
legeng(x,y,legend,…) draws a legend.

.
/
.
.

4.統計模擬

a.uniform
創造一個uniform分配
runif(n,min=,max=b)

```
runif(10,min=1,max=5)
```

```
##  [1] 1.912863 1.455830 2.196594 3.603267 1.298026 1.286819 4.325935
##  [8] 3.630143 4.267262 1.257159
```

.
.
b.set seed
透過設定seed可以取得固定結果

```
set.seed(32789)
runif(5)
```

```
## [1] 0.3575211 0.3537589 0.2672321 0.9969302 0.1317401
```

.
.
c.Binomial random variables 使用dbinom(x,size,prob)函數可取得機率

```
dbinom(x=4,size=6,prob=0.4)
```

```
## [1] 0.13824
```

.
pbinom則是算cdf

```
pbinom(4,6,0.5)
```

```
## [1] 0.890625
```

.
qbinom算成功次數

```
qbinom(0.89,6,0.5)
```

```
## [1] 4
```

.
rbinom會給你n個結果，一一列出

```
rbinom(24,15,0.1)
```

```
##  [1] 1 2 1 4 2 1 0 1 1 3 5 0 2 3 2 2 0 4 1 2 1 0 2 3
```

.
.
c.Poisson dpois(x,lambda) rpois(n,lambda)

```
dpois(x=3,lambda=0.5)
```

```
## [1] 0.01263606
```

```
rpois(10,3.7)
```

```
##  [1] 4 4 5 4 6 2 8 5 7 2
```

.
.
d.Exponential pexp(q,rate)，P(T<=q) rexp(n,rate)

```
pexp(1,3)
```

```
## [1] 0.9502129
```

```
rexp(10,3)
```

```
##  [1] 0.25689508 0.16740990 0.42135995 0.00242532 0.24017326 0.22605584
##  [7] 0.55537507 0.27051260 0.55672199 0.08279920
```

.
.
e.Normal

```
qnorm(0.95,mean=2.7,sd=3.3)
```

```
## [1] 8.128017
```

```
rnorm(10,2.7,3.3)
```

```
##  [1]  1.7340755  9.1629996 -0.8278777  0.6072807  0.8504775  4.8598930
##  [7]  5.9906160  0.6894131 -0.9699724  1.1312293
```

.
.
f.Monte Carlo integration
to approximate

$$\int_0^1 x^4 \mathrm{d}x$$

```
u=runif(100000)
mean(u^4)
```

```
## [1] 0.1998083
```

.
to approximate

$$\int_2^5 sin(x) \mathrm{d}x$$

```
u=runif(10000,min=2,max=5)
mean(sin(u)*(5-2))
```

```
## [1] -0.6932975
```

.
to approximate

$$\int_3^9 \int_1^7 sin(x-y)\mathrm{d}x\mathrm{d}y$$

```
u=runif(10000,min=1,max=7)
y=runif(10000,min=3,max=9)
mean(sin(u-y))*42
```

```
## [1] -0.3927658
```

.
.