

R PROGRAMMING

Part 2



ผู้ช่วยศาสตราจารย์ ดร. อัจฉาณัท รัตนเลิศนุสรณ์
สาขาสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

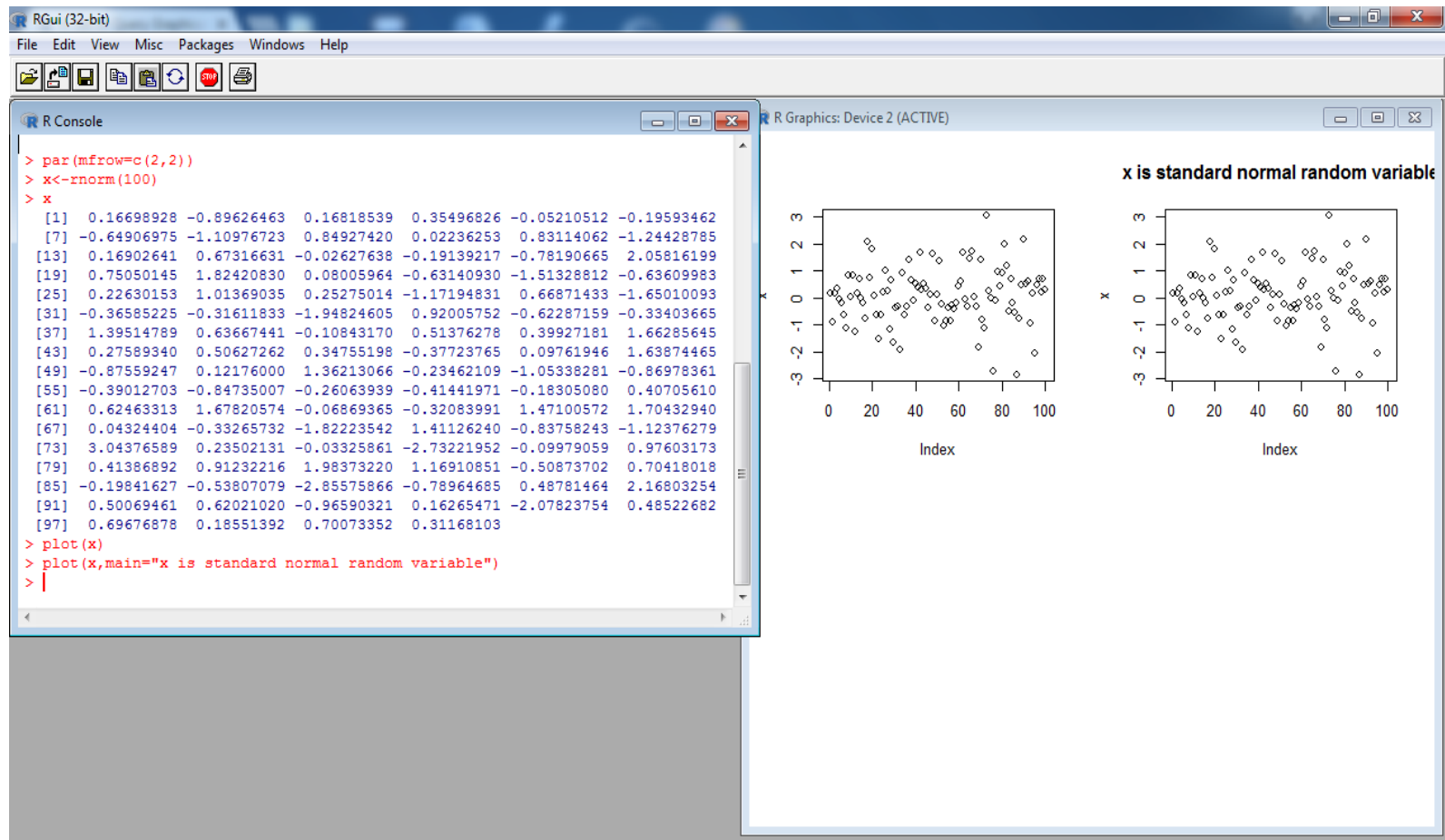
สารบัญ

Contents

- การสร้างกราฟด้วยคำสั่ง plot
- การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง
- การหาความน่าจะเป็นของตัวแปรสุ่ม
- การหาความน่าจะเป็นสะสมของตัวแปรสุ่ม
- การหาตำแหน่งของตัวแปรสุ่มเมื่อทราบความน่าจะเป็นสะสม
- การหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม
- การจำลองค่าทางสถิติ

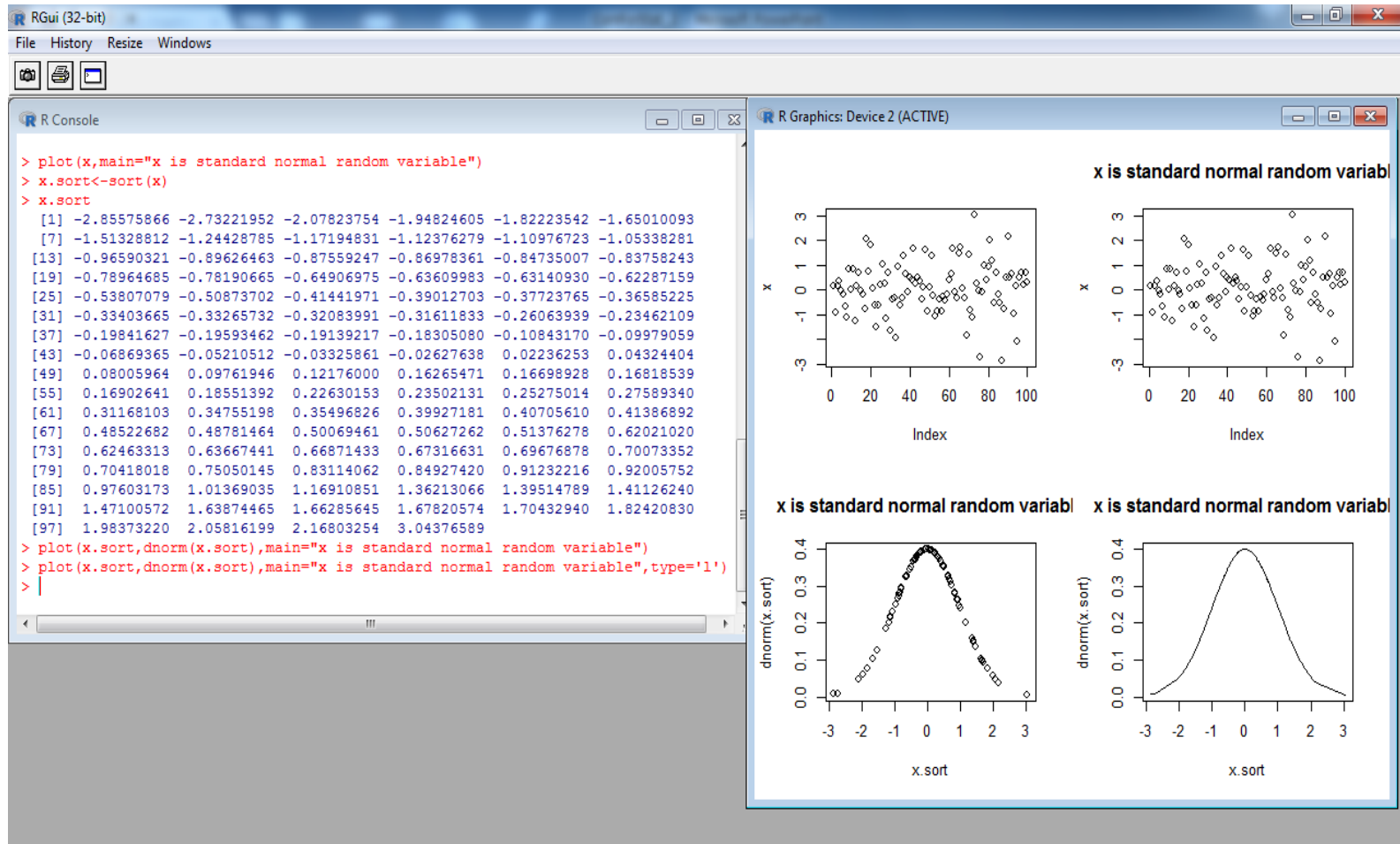
การสร้างกราฟด้วยคำสั่ง plot

- ฟังก์ชัน plot() สามารถใช้สร้างกราฟดังนี้



การสร้างกราฟด้วยคำสั่ง plot

- ฟังก์ชัน plot() สามารถใช้สร้างกราฟดังนี้



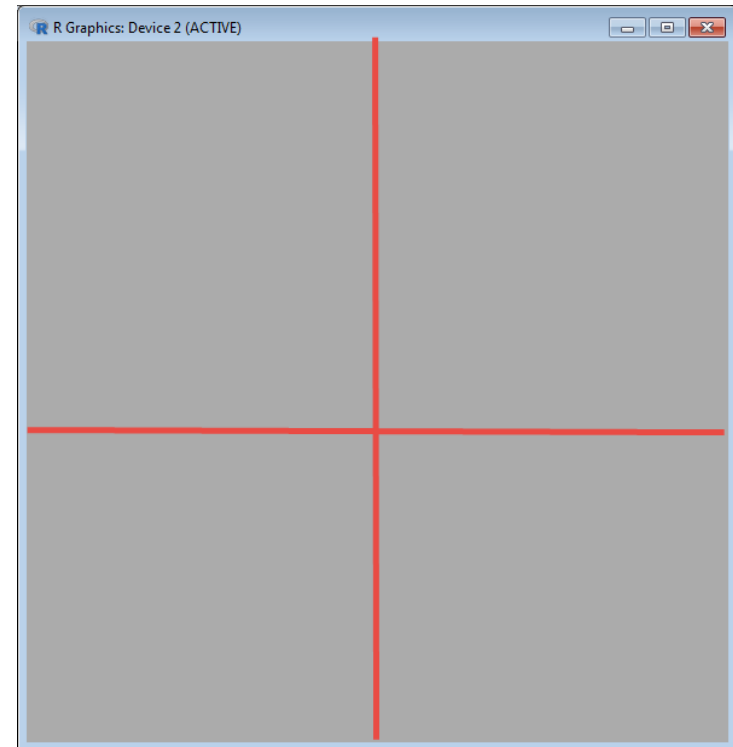
การสร้างกราฟด้วยคำสั่ง plot

- ฟังก์ชัน `par()` ใช้สำหรับกำหนดค่าพารามิเตอร์ของรูปภาพ

ตัวอย่างการใช้ฟังก์ชัน

```
> par(mfrow=c(2,2))
```

```
# กำหนดให้แสดงรูปภาพจำนวน 4 รูป
```



การสร้างกราฟด้วยคำสั่ง plot

- ตัวอย่างการใช้ฟังก์ชัน `par()` และ `plot()`

```
> par(mfrow=c(2,2))
```

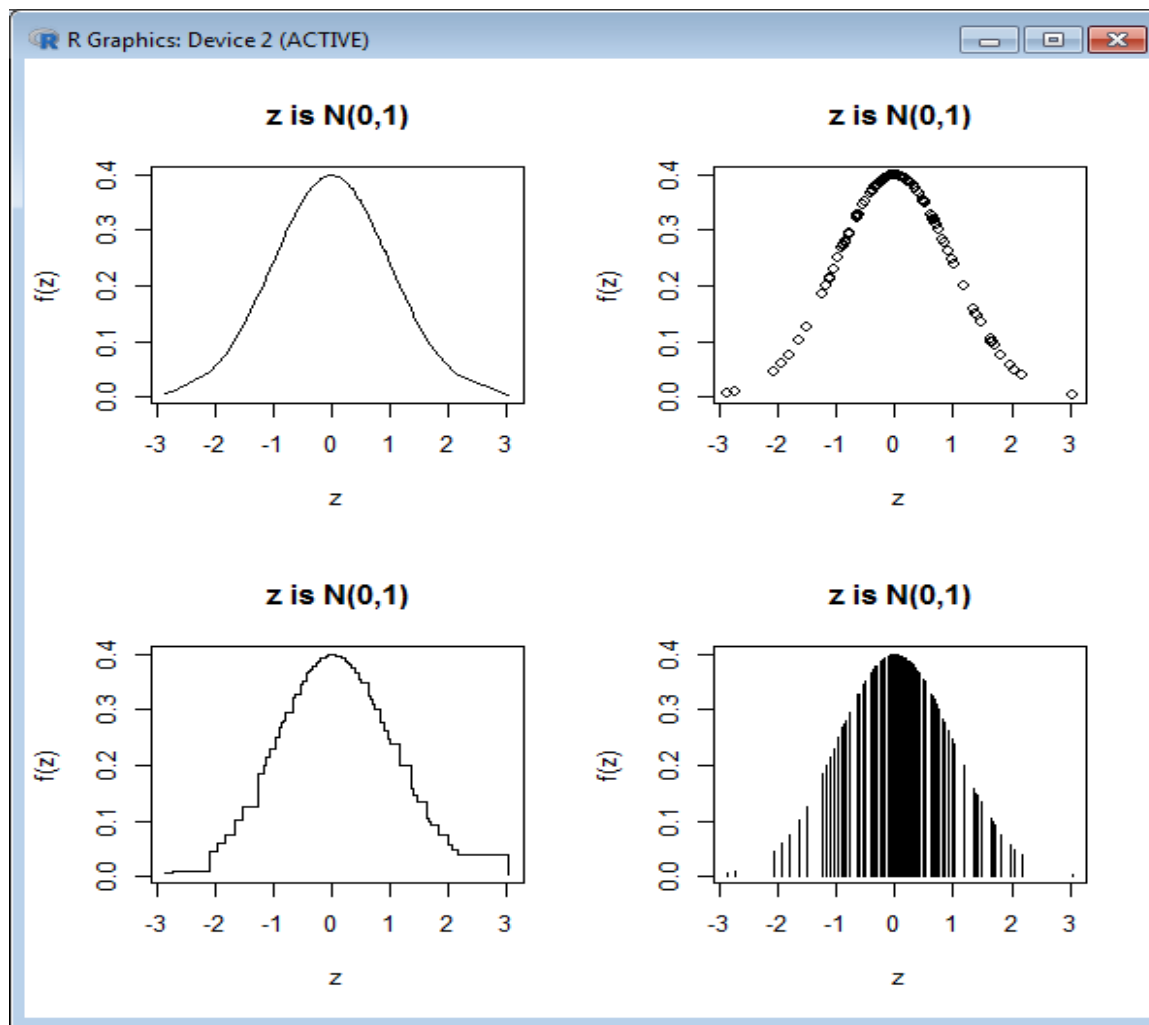
```
> plot(x.sort, dnorm(x.sort), type='l', xlab='z',  
ylab='f(z)', main='z is N(0,1)')
```

```
> plot(x.sort, dnorm(x.sort), type='p', xlab='z',  
ylab='f(z)', main='z is N(0,1)')
```

```
> plot(x.sort, dnorm(x.sort), type='s', xlab='z',  
ylab='f(z)', main='z is N(0,1)')
```

```
> plot(x.sort, dnorm(x.sort), type='h', xlab='z',  
ylab='f(z)', main='z is N(0,1)')
```

การสร้างกราฟด้วยคำสั่ง plot

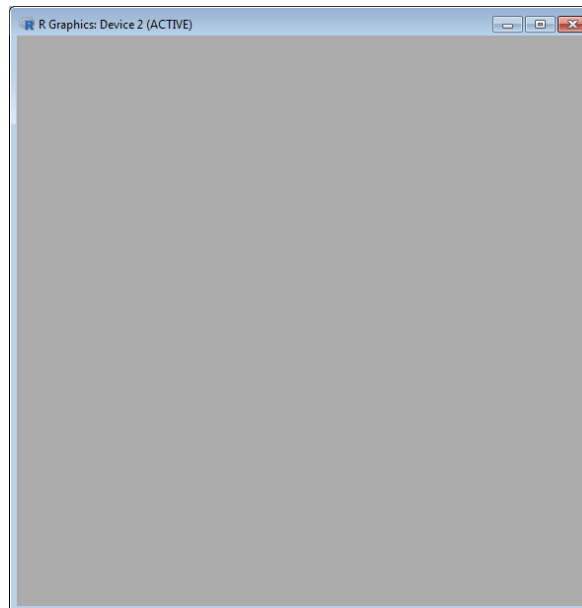


การสร้างกราฟด้วยคำสั่ง plot

- เมื่อต้องการใช้ค่าพารามิเตอร์ของรูปภาพที่โปรแกรม R กำหนดให้ คือวาดรูปภาพได้ 1 รูป ให้ใช้คำสั่งดังนี้

```
> par()
```

```
# ให้แสดงรูปภาพ 1 รูปตามที่โปรแกรม R กำหนดมาให้
```



การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

- โปรแกรม R มีการแจกแจงความน่าจะเป็นพื้นฐานดังนี้

Distribution	Functions			
Beta	pbeta	qbeta	dbeta	rbeta
Binomial	pbinom	qbinom	dbinom	rbinom
Cauchy	pcauchy	qcauchy	dcauchy	rcauchy
Chi-Square	pchisq	qchisq	dchisq	rchisq
Exponential	pexp	qexp	dexp	rexp
F	pf	qf	df	rf
Gamma	pgamma	qgamma	dgamma	rgamma
Geometric	pgeom	qgeom	dgeom	rgeom
Hypergeometric	phyper	qhyper	dhyper	rhyper
Logistic	plogis	qlogis	dlogis	rlogis
Log Normal	plnorm	qlnorm	dlnorm	rlnorm
Negative Binomial	pnbinom	qnbinom	dnbinom	rnbinom
Normal	pnorm	qnorm	dnorm	rnorm
Poisson	ppois	qpois	dpois	rpois
Student t	pt	qt	dt	rt
Studentized Range	ptukey	qtukey	dtukey	rtukey
Uniform	punif	qunif	dunif	runif
Weibull	pweibull	qweibull	dweibull	rweibull
Wilcoxon Rank Sum Statistic	pwilcox	qwilcox	dwilcox	rwilcox
Wilcoxon Signed Rank Statistic	psignrank	qsignrank	dsignrank	rsignrank

การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

- ทุกการแจกแจงจะมี prefix นำหน้าดังนี้

p สำหรับหาความน่าจะเป็น และความน่าจะเป็นสะสม

q สำหรับหาดำแหน่งที่ตรงกับความน่าจะเป็นสะสม

d สำหรับหาฟังก์ชันมวลความน่าจะเป็น (p.m.f) หรือฟังก์ชันความหนาแน่น (p. d. f.)

r สำหรับสร้างตัวเลขสุ่มที่มีการแจกแจงความน่าจะเป็นตามที่ระบุ

ตัวอย่างการแจกแจงปกติ(Normal distribution: **norm**)

เราสามารถใชฟังก์ชัน **pnorm()** , **qnorm()** , **dnorm()** และ **rnorm()** ได้

การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

- การสร้างตัวเลขสุ่มซึ่งมาจากการแจกแจงปกติมาตรฐานทำได้ดังนี้

```
> set.seed(13579) #กำหนดค่า seed number สำหรับสร้างตัวเลขสุ่ม
> x<-rnorm(50)     #สร้างตัวเลขสุ่ม 50 ค่าจากการแจกแจงปกติมาตรฐาน
> x
[1] -1.23471549 -1.25283387 -0.25477803 -1.52664663  1.09711469  2.48874422
[7]  0.77948026  0.18837501 -1.02644594 -0.25670760  0.74605016  0.47122077
[13] -0.76666991 -0.43728978 -0.47155099 -0.56138418 -1.21516541  0.24681796
[19] -0.86169804  0.65991091 -0.03684157 -0.95252445 -1.69970971  0.89985099
[25]  0.72111984 -0.16349925 -0.76652896  1.81144293  0.23834234 -0.09482171
[31]  0.90522523  0.03273290  0.24664125 -1.24344864 -1.47698088 -0.84283895
[37]  1.54991801  1.04955308  1.35743691  1.33925300  0.77756672  0.39454066
[43]  0.15138266 -0.38780454  0.46815376  1.67414620 -1.47819242 -1.63289605
[49] -0.97490915 -0.76499014
>
```

ถ้าไม่ใช้คำสั่ง `set.seed(13579)` ตัวเลขสุ่มที่สร้างขึ้นจะเปลี่ยนไปทุกครั้ง

การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

ตัวเลข
สุ่มที่ได้
แต่ละ
ครั้งจะ
เหมือน
กัน

```
> set.seed(13579)
> x<-rnorm(50)
> x
[1] -1.23471549 -1.25283387 -0.25477803 -1.52664663 1.09711469 2.48874422
[7] 0.77948026 0.18837501 -1.02644594 -0.25670760 0.74605016 0.47122077
[13] -0.76666991 -0.43728978 -0.47155099 -0.56138418 -1.21516541 0.24681796
[19] -0.86169804 0.65991091 -0.03684157 -0.95252445 -1.69970971 0.89985099
[25] 0.72111984 -0.16349925 -0.76652896 1.81144293 0.23834234 -0.09482171
[31] 0.90522523 0.03273290 0.24664125 -1.24344864 -1.47698088 -0.84283895
[37] 1.54991801 1.04955308 1.35743691 1.33925300 0.77756672 0.39454066
[43] 0.15138266 -0.38780454 0.46815376 1.67414620 -1.47819242 -1.63289605
[49] -0.97490915 -0.76499014
> set.seed(13579)
> x<-rnorm(50)
> x
[1] -1.23471549 -1.25283387 -0.25477803 -1.52664663 1.09711469 2.48874422
[7] 0.77948026 0.18837501 -1.02644594 -0.25670760 0.74605016 0.47122077
[13] -0.76666991 -0.43728978 -0.47155099 -0.56138418 -1.21516541 0.24681796
[19] -0.86169804 0.65991091 -0.03684157 -0.95252445 -1.69970971 0.89985099
[25] 0.72111984 -0.16349925 -0.76652896 1.81144293 0.23834234 -0.09482171
[31] 0.90522523 0.03273290 0.24664125 -1.24344864 -1.47698088 -0.84283895
[37] 1.54991801 1.04955308 1.35743691 1.33925300 0.77756672 0.39454066
[43] 0.15138266 -0.38780454 0.46815376 1.67414620 -1.47819242 -1.63289605
[49] -0.97490915 -0.76499014
> |
```

การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

ตัวเลข
สุ่มที่ได้
แต่ละ
ครั้งจะ
ไม่ซ้ำ
กัน

```
R Console
> x<-rnorm(50)
> x
 [1]  1.05593992  0.83720459 -0.76507999 -1.33710296  0.67734515  0.58997711
 [7]  0.17173553 -0.41383270 -1.35033709  2.31043656 -0.05148653 -0.49715707
[13] -1.11295238  1.07269558  0.24449772 -0.45434493  0.78932293  0.92857081
[19] -1.06433315  0.95224360 -0.37142803 -1.85265198 -0.67365112  0.71962465
[25]  1.09859148 -0.94668767  0.43521283 -0.24918222  0.33325341 -1.58470676
[31]  0.77971351  0.28699072  1.31212469 -1.27254580  1.36997865  2.15192641
[37] -0.35454729 -1.12620100 -0.67639614 -0.81877628 -0.70365591 -0.89596182
[43] -0.17585945 -0.82183502 -0.26703823 -0.88387729 -0.05398451  0.14806838
[49] -2.01581210  0.34718645
> x<-rnorm(50)
> x
 [1] -0.13764099  1.09241173  0.43113405  0.79425202  1.26068511 -1.37422825
 [7]  0.79276804 -0.52904247 -1.24438360 -0.15015589  0.94465655 -1.04213054
[13] -0.27244232  0.48184018  0.31009487  1.14992348 -0.74946209 -0.37528074
[19] -0.34469871  1.04420633  2.27650082  0.55522111 -2.36337470  0.53274532
[25] -0.28146206 -0.26838092  0.26701486 -1.82985560  0.84319147  0.92312410
[31] -1.05297273 -0.30523659  1.42338064 -0.41706923 -0.08694590 -0.39963520
[37] -1.88808285 -0.61910253  0.91990615  1.09220967  0.05430448  0.12730472
[43]  0.54374982 -0.01587695  0.57361550  0.53439017 -0.83432596 -1.27939038
[49] -0.72256106  0.42767339
> x<-rnorm(50)
> x
 [1] -1.17500082  0.15427902 -0.17487836 -0.10676279  1.36739079  0.30324650
 [7] -1.02133004  0.59874733  0.05747929  0.34912963  0.19976471  0.67965406
[13]  1.11905222 -0.52886082  0.55465798  0.34788752 -1.64783270 -0.08899103
[19] -1.00427935 -0.36835576 -0.23169142  1.59137178 -0.61786402  0.48375918
[25] -0.22361784  1.08749341  0.89918594 -1.13198514 -1.17993792  0.66062703
[31] -0.17371948 -0.75563239 -0.63359278 -1.96457351 -0.33491473 -1.59771944
[37] -0.13552619  1.21634750 -0.64809421 -1.65856242  0.24717205  0.50860914
[43]  0.12948233 -0.77359322 -0.20411979  1.15464631 -0.05999662  1.08372938
[49] -0.92415065 -0.93112504
> |
```

การสร้างตัวเลขสุ่มและการสุ่มตัวอย่าง

- การสร้างประชากรของตัวเลขสุ่ม และการสุ่มตัวอย่าง

```
> set.seed(2468) #กำหนดค่า seed number สำหรับสร้างตัวเลขสุ่ม
> x<-rnorm(1000) #สร้างตัวเลขสุ่ม 1000 ค่าจากการแจกแจงปกติมาตรฐาน
> x[1:50]          #แสดงค่าในตัวแปร x ค่าที่ x[1] ถึง x[50] จากทั้งหมด 1000 ค่า
[1] -0.09395560 -0.31919244 -0.75563464  0.45898188  0.70171265 -1.62431161
[7] -1.14884439 -1.74360972 -0.74711849 -0.74462565 -0.54400272 -1.10078367
[13] -0.58245788  1.80760280 -0.93897607 -1.17734832 -1.72956190 -0.20133922
[19]  0.51832385 -1.03687173  0.01705797  1.13115293  0.24080185  0.99965895
[25] -0.43234243  1.44652241  1.69692793 -0.93354111  1.07762234 -0.43344772
[31] -0.73239837 -0.37598091  0.98957785 -1.37770913  1.02991511 -0.25660521
[37]  0.29537741 -0.04973895  1.50391988 -0.75560704 -0.99590324 -0.16765170
[43]  0.06547842 -0.79418678 -1.17924156  0.22089260  0.53382259  1.06754656
[49]  0.67298018 -0.62562458
> x.samp<-sample(x,size=25) #สุ่มค่าจากตัวแปร x จำนวน 25 ค่า มาเก็บที่ตัวแปร x.samp
> x.samp          #แสดงค่าในตัวแปร x.samp
[1]  0.67298018  0.91312443  1.56640655 -0.46609980  1.61159727 -1.01259553
[7] -0.03132962 -2.68698078  1.28057434  0.16678177  0.06077329  0.07951142
[13] -0.14906921 -0.40838009 -0.75629301  0.30550037  0.05706630 -0.55805311
[19]  1.23389352  0.34069837 -1.09602144  1.30199218  1.80760280 -0.22051550
[25]  0.04752172
```

ตัวแปรสุ่ม

- ตัวแปรสุ่ม(random variable) แบ่งออกได้ 2 ชนิด
 - ❖ ตัวแปรสุ่มไม่ต่อเนื่อง (discrete random variable)
 - ❖ ตัวแปรสุ่มต่อเนื่อง (continuous random variable)
- ตัวแปรสุ่มไม่ต่อเนื่อง คือ ตัวแปรสุ่มที่ให้ค่าเป็นจำนวนนับ อาทิ จำนวนรถยนต์ จำนวนนักศึกษา จำนวนผู้ใช้งาน อินเทอร์เน็ตของมหาวิทยาลัยฯ เป็นต้น
- ตัวแปรสุ่มต่อเนื่อง คือ ตัวแปรสุ่มที่ให้ค่าเป็นจำนวนจริงในช่วงที่ระบุเฉพาะ อาทิ ความดันโลหิต เวลา อุณหภูมิ น้ำหนัก ส่วนสูง เป็นต้น

สัญลักษณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม

- ชื่อตัวแปรสุ่ม อาทิ X, Y, Z, \dots
- ค่าของตัวแปรสุ่ม อาทิ x, y, z, \dots
- ฟังก์ชันความน่าจะเป็น $f(x) = P(X = x)$
- ฟังก์ชันความน่าจะเป็นสะสม $F(x) = P(X \leq x)$

ตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

$$X \sim N(\mu, \sigma^2)$$

- ตัวแปรสุ่มที่มีการแจกแจงแบบปกติ มีฟังก์ชันความน่าจะเป็นดังนี้

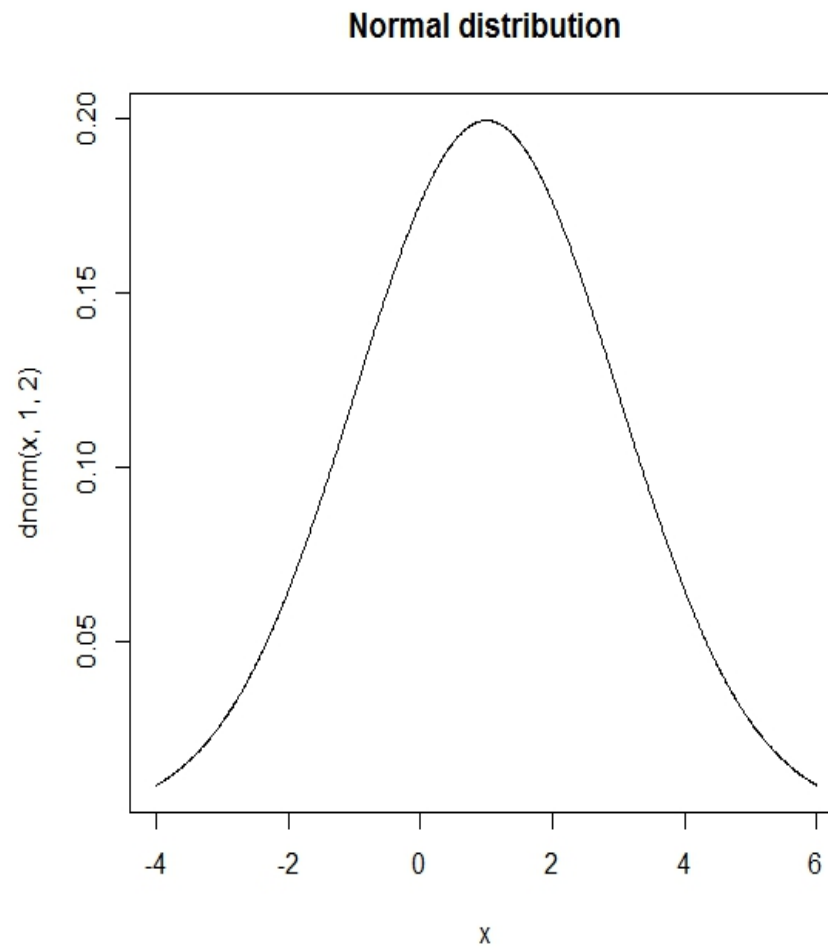
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\},$$

$$-\infty \leq x \leq \infty, -\infty \leq \mu \leq \infty \text{ และ } \sigma > 0.$$

- ตัวแปรสุ่มแบบปกติ มีฟังก์ชันความน่าจะเป็นสะสมดังนี้

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx$$

กราฟแสดงการแจกแจงแบบปกติ



การหาความน่าจะเป็นของตัวแปรสุ่ม

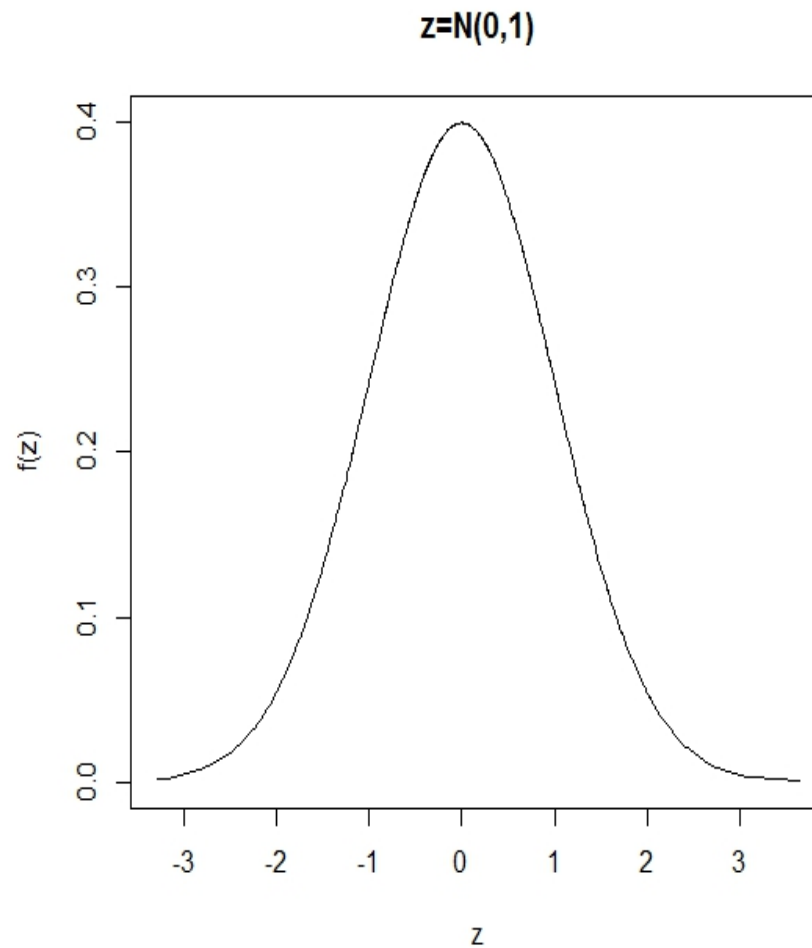
- การหาความน่าจะเป็นของตัวแปรสุ่มแบบปกติมาตรฐาน
- ตัวแปรสุ่มแบบปกติมาตรฐาน มีฟังก์ชันความหนาแน่นดังนี้

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right)$$

- ตัวแปรสุ่มแบบปกติมาตรฐาน มีฟังก์ชันความน่าจะเป็นสะสมดังนี้

$$F_Z(z) = P(Z \leq z) = \int_{-\infty}^z f(z) dz$$

กราฟแสดงการแจกแจงแบบปกติมาตรฐาน



การหาความน่าจะเป็นของตัวแปรสุ่ม

- รูปแบบการใช้ฟังก์ชัน `pnorm()`

```
pnorm(q, mean = 0, sd = 1, lower.tail = TRUE,  
      log.p = FALSE)
```

- ตัวอย่างการหาความน่าจะเป็นของตัวแปรสุ่มแบบปกติมาตรฐาน

#ความน่าจะเป็นสะสมถึงตำแหน่ง 1.96

```
> pnorm(1.96)
```

```
[1] 0.9750021
```

#ความน่าจะเป็นสะสมตั้งแต่ตำแหน่ง 1.96 ขึ้นไป

```
> pnorm(1.96, lower.tail=FALSE)
```

```
[1] 0.0249979
```

#ความน่าจะเป็นจาก 0 ถึง 1.96

```
> pnorm(1.96) - pnorm(0.00)
```

```
[1] 0.4750021
```

การหาตำแหน่งของตัวแปรสุ่มเมื่อทราบความน่าจะเป็น

- รูปแบบการใช้ฟังก์ชัน `qnorm()`

```
qnorm(p, mean = 0, sd = 1, lower.tail = TRUE,  
      log.p = FALSE)
```

- ตัวอย่างการหาตำแหน่งของตัวแปรสุ่มแบบปกติมาตรฐาน

```
> qnorm(0.975)
```

```
[1] 1.959964
```

```
> qnorm(0.975, lower.tail=FALSE)
```

```
[1] -1.959964
```

การหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม

- รูปแบบการใช้ฟังก์ชัน `dnorm()`

```
dnorm(x, mean = 0, sd = 1, log = FALSE)
```

- ตัวอย่างการหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบปกติ
มาตรฐาน

```
> x
[1] 1.0026529 0.3245132 0.2836438 1.4798500 0.3665132 -0.1947419 2.0266552
[8] 0.2346057 -0.2767158 0.7069725
> fx<-dnorm(x)
> fx
[1] 0.24132879 0.37847966 0.38321257 0.13346492 0.37302700 0.39144872 0.05116988
[8] 0.38811311 0.38395714 0.31072606
```

แบบฝึกหัดข้อที่ 1

กำหนดให้ $X \sim N(\mu = 5, \sigma = 1)$ จงหาค่าดังต่อไปนี้ด้วยโปรแกรมอาร์

1) $P(X \leq 6) = \dots ? \dots$

2) $P(X \leq 6) = \dots ? \dots$

3) $P(4 \leq X \leq 6) = \dots ? \dots$

4) $P(X \leq ?) = 0.95$

5) $P(X > ?) = 0.95$

6) $P(X = 5), P(X = 6), P(X = 7)$

7) สร้างตัวเลขสุ่ม 10 ค่าจากการแจกแจงของตัวแปรสุ่ม X

แบบฝึกหัดข้อที่ 2

กำหนดให้ $Z \sim N(\mu = 0, \sigma = 1)$ จงหาค่าดังต่อไปนี้ด้วยโปรแกรมอาร์

- 1) $P(Z \leq 1.96) = \dots ? \dots$
- 2) $P(Z > 1.96) = \dots ? \dots$
- 3) $P(-1.96 \leq Z \leq 1.96) = \dots ? \dots$
- 4) $P(Z \leq ?) = 0.975$
- 5) $P(Z > ?) = 0.975$
- 6) $P(Z = -1), P(Z = 0), P(Z = 2)$
- 7) สร้างตัวเลขสุ่ม 10 ค่าจากการแจกแจงของตัวแปรสุ่ม Z

แบบทดสอบการจำลองค่าทางสถิติ

- การจำลองค่าทางสถิติ (Statistical Simulation)
- งานที่มอบหมายให้นักศึกษาทำการทดลองดังนี้
 1. สร้างตัวแปรสุ่ม x จากการแจกแจงปกติ ที่มีค่าเฉลี่ยเท่ากับ 10 ความแปรปรวนเท่ากับ 2 จำนวน 1000 ค่า
 2. สุ่มตัวอย่างค่าตัวแปรสุ่ม x มาจำนวน 15, 25, 35, 45, 60, 75, 90, 105, 120, 135 และ 150
 3. หาค่าเฉลี่ยของตัวอย่างสุ่มแต่ละขนาดตัวอย่าง
 4. สังเกตผลว่า ค่าเฉลี่ยของตัวอย่างสุ่มว่ามีแนวโน้มอย่างไร
 5. อภิปรายผล