บทที่ 4 การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

88520159

PROBABILITY AND STATISTICS FOR COMPUTING

การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

- บทเรียนที่ผ่านมาศึกษาถึง การหาค่าความน่าจะเป็นของเหตุการณ์ที่สนใจเพียงบางเหตุการณ์
- บทเรียนนี้เราจะศึกษาความน่าจะเป็นของผลลัพธ์ที่มีค่าเป็นตัวเลขที่เป็นไปได้ทั้งหมดทุกกรณี
- เรียกความน่าจะเป็นของผลลัพธ์ที่เป็นไปได้ทั้งหมดทุกกรณีนี้ว่าเป็นการแจกแจงความน่าจะเป็น (Probability Distributions)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

- ในชีวิตประจำวันนั้น จะต้องหาค่าความน่าจะเป็นของผลลัพธ์ทั้งหมดที่อาจเกิดขึ้นเพื่อนำข้อมูลเหล่านั้นมาใช้ประกอบในการตัดสินใจ
- ตัวอย่าง พนักงานขายของสามารถคำนวณหาความน่าจะเป็นที่เขาจะขายของได้ 0 1 2 ชิ้นหรือมากกว่านั้น ต่อวันได้
- คำนวณหาค่าเฉลี่ย ความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของเหตุการณ์

1. ตัวแปรสุ่ม (Random Variables)

การทดลองสุ่มหรือการทดลองใดๆ ที่เราไม่ทราบผลการทดลองล่วงหน้า

<u>ตัวอย่าง</u>

- การทอดลูกเต๋า 1 ลูก 1 ครั้ง กำหนดให้ X แทนผลลัพธ์ที่เกิดขึ้น ดังนั้น X อาจจะมีค่าเท่ากับ 1 , 2 , 3 , 4 , 5 , 6 ก็ได้
- การโยนเหรียญ 2 เหรียญ 1 ครั้ง กำหนดให้ Y แทนจำนวนเหรียญที่ขึ้นหัว ดังนั้น Y อาจจะมีค่าเท่ากับ 0 , 1 , 2 ก็ได้
- เรียกตัวแปร X หรือ Y ในที่นี้ว่า *ตัวแปรสุ่ม (random Variable)*

ตัวแปรสุ่ม คือ ตัวแปรที่มีค่าเป็นจำนวนจริง ที่ได้จากการทดลองสุ่ม หรือความบังเอิญ หรือไม่สามารถทราบค่าล่วงหน้าได้

1. ตัวแปรสุ่ม (Random Variables)

<u>ตัวอย่าง</u> ตัวแปรสุ่ม

- กล่องใบหนึ่งมีลูกบอลสีแดง 4 ลูก สีดำ 3 ลูก สุ่มหยิบลูกบอลมา 2 ลูก โดยหยิบทีละลูกแล้วไม่ใส่คืน
  กำหนดให้ Y คือ จำนวนลูกบอลสีแดงที่หยิบได้
  ดังนั้นค่าของ Y ที่เป็นไปได้คือ 0,1 หรือ 2

- ในการชั่งน้ำหนักของเด็กแรกเกิดที่คลอด ณ โรงพยาบาลแห่งหนึ่ง
  กำหนดให้ Z แทนน้ำหนักของเด็กแรกเกิด
  ดังนั้น Z มีค่าที่เป็นไปได้คือ $0 < Z < 8$ กิโลกรัม

1. ตัวแปรสุ่ม (Random Variables)

ชนิดของตัวแปรสุ่ม

1. ตัวแปรสุ่มชนิดไม่ต่อเนื่อง (Discrete random variable) ถ้า X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ค่าของ X จะเป็นเลขจำนวนเต็มบวกเท่านั้น และไม่สามารถมีค่าเป็นทศนิยมหรือเศษส่วนในระบบจำนวนจริงได้ ดังนั้น ตัวแปรสุ่มชนิดไม่ต่อเนื่องมักจะเป็นจำนวนนับ เช่น จำนวนบุตรของประชากร

2. ตัวแปรสุ่มชนิดต่อเนื่อง (Continuous random variable) ถ้า X เป็นตัวแปรสุ่มแบบต่อเนื่อง ค่าของ X จะเป็นจำนวนจริงที่ต่อเนื่องกัน และไม่สามารถนับได้ เช่น ความสูงของคนไทย

1. ตัวแปรสุ่ม (Random Variables)

ชนิดของตัวแปรสุ่ม

จงบอกว่าตัวแปรสุ่มในแต่ละข้อเป็นตัวแปรสุ่มชนิดต่อเนื่องหรือชนิดไม่ต่อเนื่อง

1. จำนวนนิสิตที่ขาดเรียน
2. เวลาที่นิสิตใช้ในการสอบ
3. อุณหภูมิในแต่ละวัน
4. จำนวนวันที่มีอุณหภูมิต่ำกว่า 20 องศาเซลเซียส
5. ความสูงของตึกที่สร้างขึ้นในเขตกรุงเทพมหานคร
6. ความหนาของชั้นหิน
7. จำนวนชั้นหินที่พบในภูเขาแห่งหนึ่ง

2. การแจกแจงความน่าจะเป็น (Probability Distribution)

- ถ้า X เป็นตัวแปรสุ่ม แล้วค่าของตัวแปรสุ่ม X อาจเกิดขึ้นด้วยความน่าจะเป็นที่แตกต่างกัน
- ความน่าจะเป็นของตัวแปรสุ่มเขียนแทนด้วยสัญลักษณ์

$$P(X = x)$$

หมายถึงค่าความน่าจะเป็นของตัวแปรสุ่ม X เมื่อ X มีค่าเท่ากับ x

- นำมาแสดงการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

2. การแจกแจงความน่าจะเป็น (Probability Distribution)

<u>ตัวอย่าง</u> หากโยนเหรียญ 1 อัน 3 ครั้ง ให้ X เป็นจำนวนครั้งที่ขึ้นหัว

$S = \{HHH, HHT, HTH, THH, TTH, HTT, THT, TTT\}$

จะได้ x = 0, 1, 2, 3

ความน่าจะเป็นที่เหรียญไม่ขึ้นหัวเลย

ความน่าจะเป็นที่เหรียญขึ้นหัว 1 ครั้ง

ความน่าจะเป็นที่เหรียญขึ้นหัว 2 ครั้ง

ความน่าจะเป็นที่เหรียญขึ้นหัว 3 ครั้ง

2. การแจกแจงความน่าจะเป็น (Probability Distribution)

สามารถแสดงค่าความน่าจะเป็นของตัวแปรสุ่ม X ทุกๆ ค่าที่เป็นไปได้ ดังนี้

$\boldsymbol{x}$	0	1	2	3
$\boldsymbol{P(X = x)}$	$\mathbf{\frac{1}{8}}$	$\mathbf{\frac{3}{8}}$	$\mathbf{\frac{3}{8}}$	$\mathbf{\frac{1}{8}}$

เรียกว่าตารางแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

กล่าวได้ ว่า $P(X = x)$ เป็นฟังก์ชันของ x จึงสามารถเขียนแทนได้ด้วย

$f(x)$ เช่น $f(0) = \frac{1}{8}$, $f(1) = \frac{3}{8}$, $f(2) = \frac{3}{8}$ และ $f(3) = \frac{1}{8}$

2. การแจกแจงความน่าจะเป็น (Probability Distribution)

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$P(X < 3)$ หมายถึง ความน่าจะเป็นที่เหรียญขึ้นหัวน้อยกว่า 3 ครั้ง

$P(X \geq 2)$ หมายถึง ความน่าจะเป็นที่เหรียญขึ้นหัวมากกว่าหรือเท่ากับ 2 ครั้ง

$P(1 < X < 3)$ หมายถึง ความน่าจะเป็นที่เหรียญขึ้นระหว่าง 1 เหรียญ กับ 3 ครั้ง

2. การแจกแจงความน่าจะเป็น (Probability Distribution)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม คือ ฟังก์ชันที่แสดงค่าความน่าจะเป็นที่จะเกิดค่าของตัวแปรสุ่ม X ในการทดลอง

การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม แบ่งออกเป็น 2 ชนิด ตามชนิดของตัวแปร

1. การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่อง

2. การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง

2.1 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่อง

- การทดลองสุ่มแบบทั่วไป ลักษณะของตัวแปรสุ่มจะมีลักษณะคล้ายๆ กัน
- การโยนเหรียญ 1 อัน 3 ครั้ง สนใจจำนวนครั้งที่ขึ้นหัว จำนวนครั้งที่ขึ้นหัว จะมีค่าเป็นไปได้คือ 0,1,2 หรือ 3
- ครอบครัวที่มีบุตร 3 คน จะได้ว่าจำนวนบุตรหญิงที่เป็นไปได้คือ 0,1,2 หรือ 3
- การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่องที่สำคัญ

  2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

  2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

การทดลองแบบทวินาม

1. การทดลองประกอบด้วยการกระทำซ้ำๆ กัน n ครั้ง แต่ละครั้งเป็นอิสระต่อกัน

2. ในการทดลองแต่ละครั้งจะมีผลลัพธ์ 2 อย่าง

ความสำเร็จ

ความไม่สำเร็จ

3. ความน่าจะเป็นของความสำเร็จ ที่เกิดขึ้นจากการกระทำแต่ละครั้งมีค่าคงที่เท่ากับ p และความน่าจะเป็นของความไม่สำเร็จเท่ากับ $q = 1 - p$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

ฟังก์ชันการแจกแจงความน่าจะเป็นแบบทวินาม

$$P(X = x) = f(x;\, n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{(n - x)!\, x!} p^x q^{n-x}$$

x คือ จำนวนครั้งของความสำเร็จ $(0, 1, 2, \dots, n)$

n คือ จำนวนครั้งของการกระทำซ้ำ

p คือ ความน่าจะเป็นของความสำเร็จในแต่ละครั้ง

q คือ ความน่าจะเป็นของความไม่สำเร็จในแต่ละครั้ง (1 - p)

ค่าเฉลี่ยและความแปรปรวนของการแจกแจงแบบทวินาม

ค่าเฉลี่ย (ความคาดหวัง) $\mu = E(X) = np$

ความแปรปรวน $\sigma^2 = V(X) = npq$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

ตัวอย่าง การทดลองแบบทวินาม ด้วยการโยนเหรียญหนึ่งอัน 6 ครั้ง

- กระทำที่ซ้ำๆ กัน 6 ครั้ง ซึ่งแต่ละครั้งเป็นอิสระต่อกัน $\boxed{n = 6}$
- หากสนใจผลลัพธ์ขึ้นหัว → ขึ้นหัวเป็นความสำเร็จ
  → ขึ้นก้อยเป็นความไม่สำเร็จ
- ความน่าจะเป็นของความสำเร็จ และไม่สำเร็จ ในแต่ละครั้ง

  ความน่าจะเป็นของการขึ้นหัว เท่ากับ 1/2 $\boxed{p = 0.5}$

  ความน่าจะเป็นที่จะขึ้นก้อย เท่ากับ 1/2 $\boxed{q = 0.5}$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

<u>ตัวอย่าง</u> ในการโยนเหรียญ 1 อัน 6 ครั้ง

จากการทดลอง $n = 6$ $p = 0.5$

X คือจำนวนครั้งที่ขึ้นหัว ซึ่งมีค่าที่เป็นไปได้คือ x = 0,1,2,3,4,5,6

พิจารณาการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ขึ้นหัว

```
> x=0:6
> pr=dbinom(x,6,0.5) #คำนวณค่าความน่าจะเป็นของ x ทุกค่าเก็บในตัวแปร pr
> plot(x, pr, type = "h",ylab="probability",main="Binomial(n=6,p=0.5)")
#สร้างกราฟความน่าจะเป็นของ x
> points(0:6,pr,pch=16) #กำหนดจุดที่ปลายเส้นกราฟความน่าจะเป็น
```

คำสั่งใน R สร้างกราฟแจกแจงความน่าจะเป็น

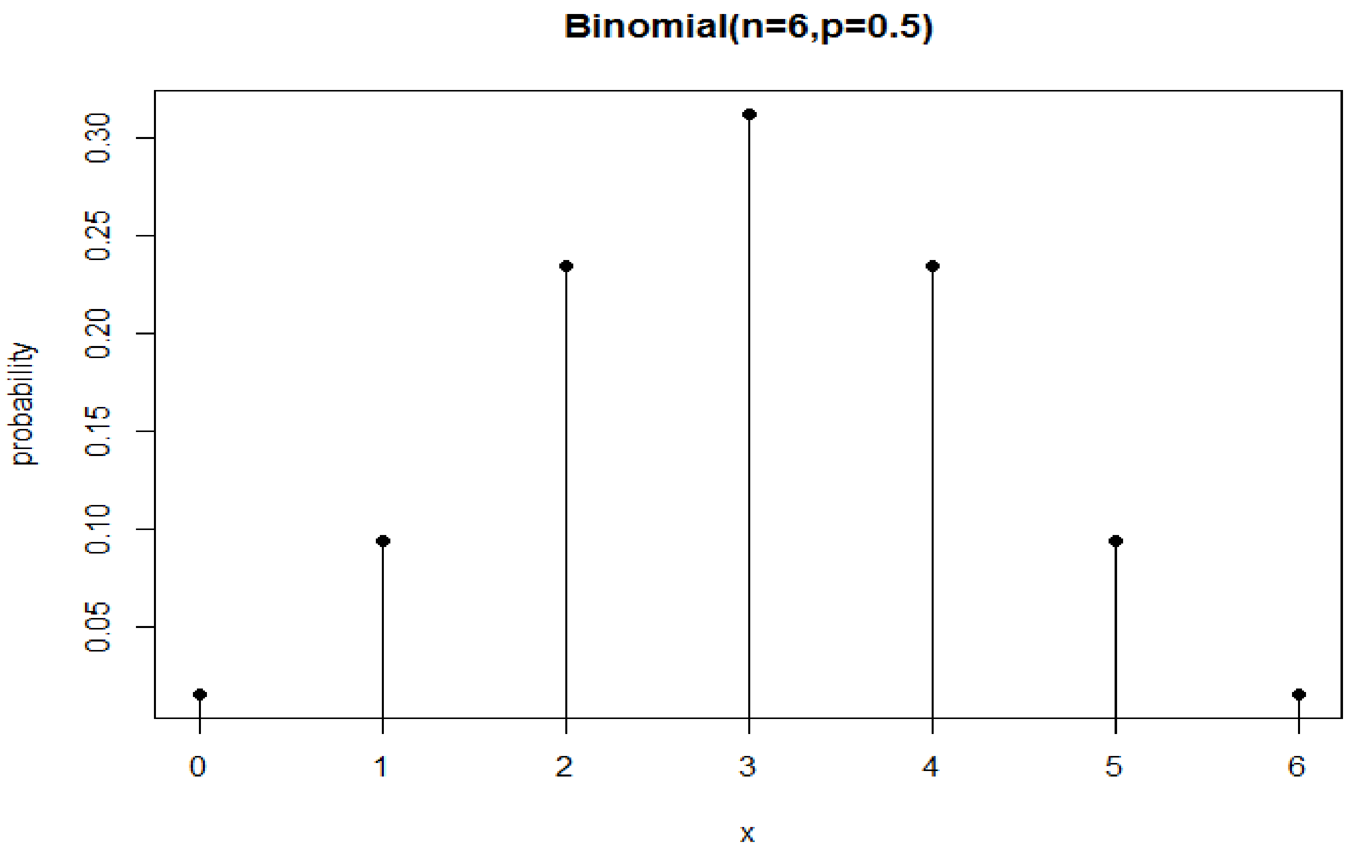
ฟังก์ชัน dbinom() คำนวณค่าความน่าจะเป็นของตัวแปรสุ่ม x ที่มีการแจกแจงทวินาม เมื่อกำหนดค่า x n และ p

ฟังก์ชัน plot() สร้างกราฟ โดยกำหนดตัวแปรสำหรับแกน x และ y ตามลำดับ type= เป็นการกำหนดชนิดของกราฟ ”h” หมายถึง กราฟในแนวตั้ง (horizontal)

ฟังก์ชัน points() เป็นการกำหนดจุดที่ coordinate x, y และ pch คือการกำหนดรูปแบบของจุด

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

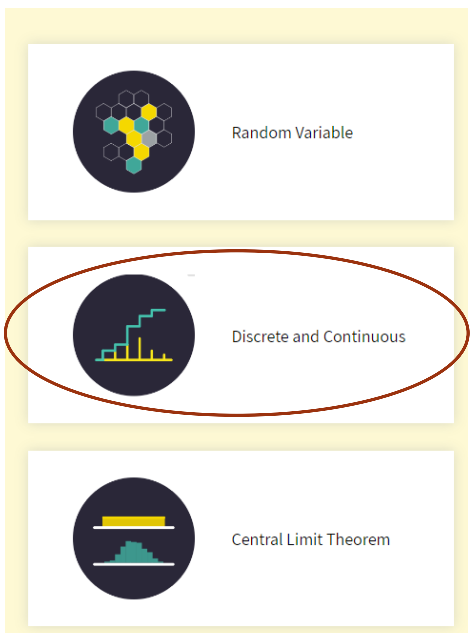
การแจกแจงความน่าจะเป็นของจำนวนครั้งที่ขึ้นหัวจากการโยนเหรียญ 6 ครั้ง

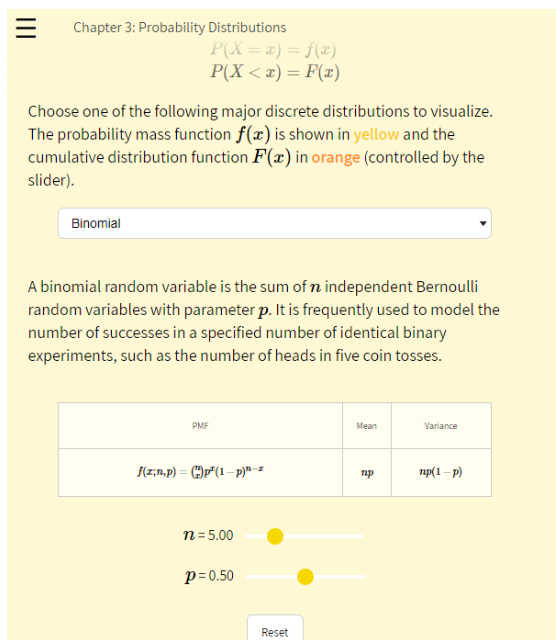


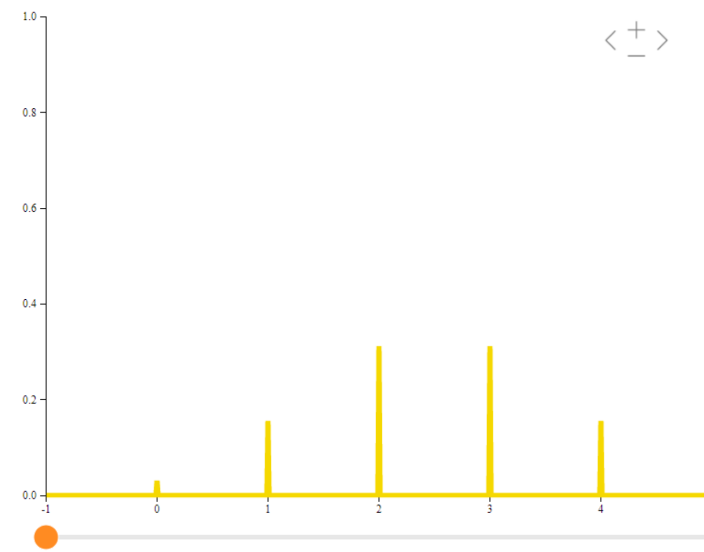
รูปร่างการแจกแจงของการแจกแจงต่างๆ

ดูรูปร่างการแจกแจงของการแจกแจงต่างๆ เมื่อเปลี่ยนค่าของพารามิเตอร์ตามที่ต้องการได้ที่

https://students.brown.edu/seeing-theory/probability-distributions/index.html







คำสั่งใน R การแจกแจงแบบทวินาม

Binomial {stats} R Documentation

The Binomial Distribution

Description

Density, distribution function, quantile function and random generation for the binomial distribution with parameters `size` and `prob`.

This is conventionally interpreted as the number of 'successes' in `size` trials.

Usage

```
dbinom(x, size, prob, log = FALSE)
pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
rbinom(n, size, prob)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`size`	number of trials (zero or more).
`prob`	probability of success on each trial.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).

คำสั่งใน R การแจกแจงแบบทวินาม

- dbinom คือการหาค่าความน่าจะเป็นของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม x ขนาดของการทดลอง size และ ความน่าจะเป็นที่จะเกิดความสำเร็จ prob สัญลักษณ์คือ $P(X = x)$
- pbinom คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q ขนาดของการทดลอง size และความน่าจะเป็นของการเกิดความสำเร็จ prob โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสมเมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์คือ $P(X \leq q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ์ $P(X > q)$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ขึ้นหัวจากการโยนเหรียญ 6 ครั้ง

1. จงหาความน่าจะเป็นที่จะขึ้นหัว 4 ครั้ง

   ต้องการหาค่า $P(X = 4)$

```
> dbinom(4,size=6,prob=0.5)
[1] 0.234375
```

ดังนั้น $P(X = 4) = 0.234375$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ขึ้นหัวจากการโยนเหรียญ 6 ครั้ง

2. จงหาความน่าจะเป็นที่จะขึ้นหัวน้อยกว่า 4 ครั้ง

ต้องการหาค่า $P(X < 4)$ *ซึ่งเท่ากับ* $P(X \leq 3)$

หรือ $P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

```
> pbinom(3,6,0.5)
[1] 0.65625
```

ดังนั้น $P(X < 4) = 0.65625$

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

ตัวอย่าง นักบาสเกตบอลคนหนึ่งมีความสามารถในการยิงลูกบอลลงห่วง 70% ถ้าเขายิงลูกบอลทั้งหมด 10 ครั้ง

จากการทดลอง n = 10 p = 0.7

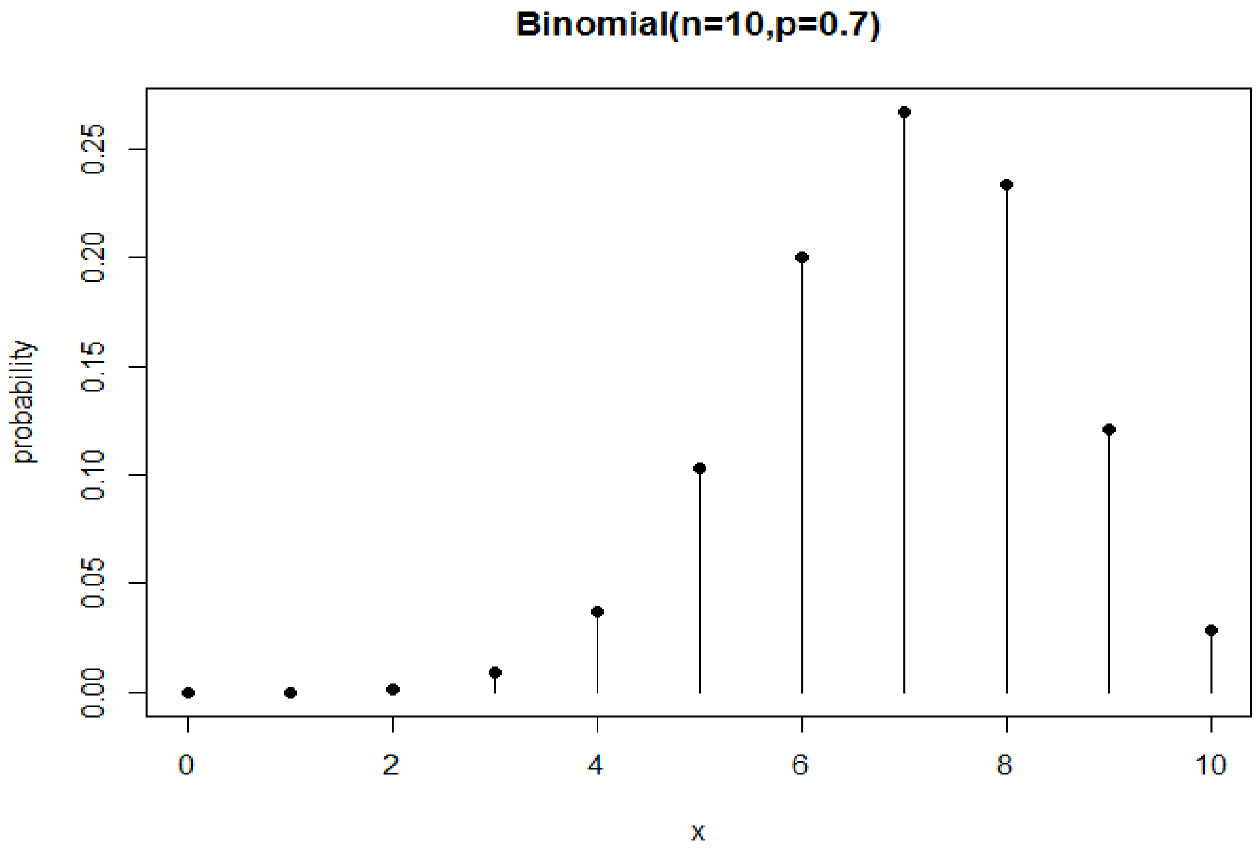
X คือจำนวนครั้งที่เขายิงลูกบอลลงห่วง ซึ่งมีค่าที่เป็นไปได้คือ 0,1,2,3,4,5,6,7,8,9,10

พิจารณาการแจกแจงความน่าจะเป็นของจำนวนครั้งที่เขายิงลูกบอลลงห่วง

```
> x=0:10
> pr=dbinom(x,10,0.7)
> plot(x,pr,type = "h",ylab="probability",main="Binomial(n=10,p=0.7)")
> points(0:10,pr,pch=16)
```

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

การแจกแจงความน่าจะเป็นของจำนวนครั้งที่ยิงบอลลงห่วงจากการยิงลูกบอล 10 ครั้ง



2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ยิงบอลลงห่วงจากการยิงลูกบอล 10 ครั้ง

1. จงหาความน่าจะเป็นที่เขายิงลูกบอลลงห่วงอย่างน้อย 6 ครั้ง

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ยิงบอลลงห่วงจากการยิงลูกบอล 10 ครั้ง

2. เขายิงลูกบอลลงห่วงไม่เกิน 4 ครั้ง

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ยิงบอลลงห่วงจากการยิงลูกบอล 10 ครั้ง

3. เขายิงลูกบอลลงห่วงตั้งแต่ 2 ถึง 5 ครั้ง

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

จากการแจกแจงความน่าจะเป็นของจำนวนครั้งที่ยิงบอลลงห่วงจากการยิงลูกบอล 10 ครั้ง

4. เขายิงลูกบอลลงห่วงระหว่าง 2 และ 5 ครั้ง

2.1.1 การแจกแจงแบบทวินาม (Binomial distribution)

<u>ตัวอย่าง</u> ถ้าความน่าจะเป็นของกลอนประตูที่ชำรุดเท่ากับ 0.01 จงหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของจำนวนกลอนประตูที่ชำรุด ถ้ามีกลอนประตู 4,000 อัน

ค่าเฉลี่ย (ความคาดหวัง)

ส่วนเบี่ยงเบนมาตรฐาน

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

การทดลองแบบปัวซง

- การทดลองที่ค่าของตัวแปรสุ่ม X แสดงจำนวนครั้งของความสำเร็จในช่วงเวลาหนึ่งที่กำหนดหรือภายในบริเวณที่กำหนดให้

  ช่วงเวลา หนึ่งนาที หนึ่งชั่วโมง หนั่งวัน หนึ่งเดือน หนึ่งปี

  บริเวณ ส่วนหนึ่งของพื้นที่ ส่วนหนึ่งของปริมาตร

- เราทราบค่าเฉลี่ย (λ) ของจำนวนความสำเร็จ (success) ที่ปรากฏขึ้นในช่วงเวลาหนึ่งหรือบริเวณหนึ่งที่กำหนด

- X เป็นตัวแปรสุ่มปัวซง โดย x = 0, 1, 2, 3 ..... ไม่สามารถระบุ n ได้

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

สมการของการแจกแจงปัวซง

$$P(X = x) = f(x; \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x = 0, 1, 2, \ldots$$

เมื่อ

λ คือ ค่าเฉลี่ยของจำนวนครั้งของความสำเร็จที่เกิดขึ้นในช่วงเวลาหรืออาณาบริเวณที่กำหนด

e คือ ค่าเอกซ์โปแนนเชียลฟังก์ชัน ซึ่งมีค่าโดยประมาณ 2.71828...

ค่าเฉลี่ยและความแปรปรวนของการแจกแจงแบบปัวซง

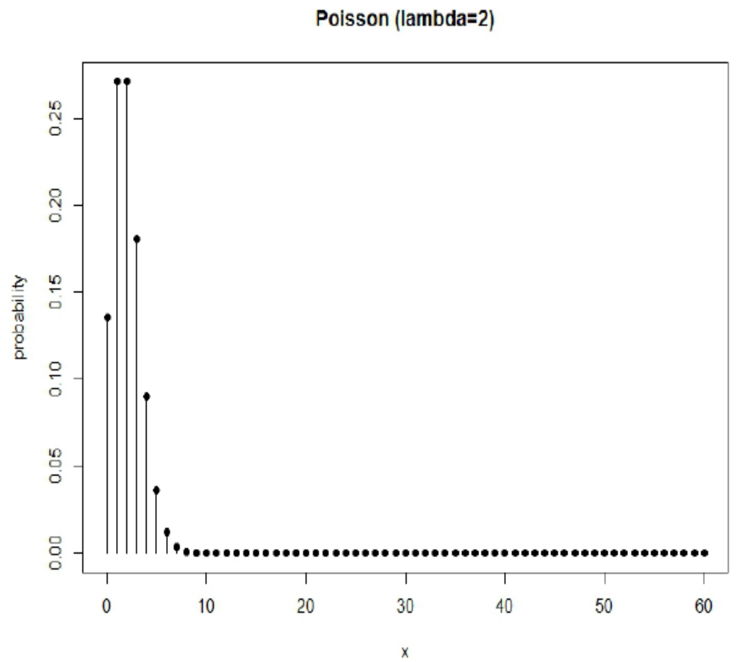
ค่าเฉลี่ย (ความคาดหวัง) $\mu = E(X) = \lambda$

ความแปรปรวน $\sigma^2 = V(X) = \lambda$

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

สร้างกราฟการแจกแจงความน่าจะเป็นแบบปัวซง ซึ่งจะขึ้นอยู่กับ λ

```
> x=0:60
> pr=dpois(x,2) # กำหนดให้ λ = 2
> plot(x,pr,type = "h",ylab="probability",main="Poisson (lambda=2)")
> points(0:60,pr,pch=16)
```



คำสั่งใน R การแจกแจงแบบปัวซง

Poisson {stats} R Documentation

The Poisson Distribution

Description

Density, distribution function, quantile function and random generation for the Poisson distribution with parameter `lambda`.

Usage

```
dpois(x, lambda, log = FALSE)
ppois(q, lambda, lower.tail = TRUE, log.p = FALSE)
qpois(p, lambda, lower.tail = TRUE, log.p = FALSE)
rpois(n, lambda)
```

Arguments

`x`	vector of (non-negative integer) quantiles.
`q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of random values to return.
`lambda`	vector of (non-negative) means.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

คำสั่งใน R การแจกแจงแบบปัวซง

- dpois คือการหาค่าความน่าจะเป็นของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม x และค่าเฉลี่ย λ สัญลักษณ์คือ $P(X = x)$

- ppois คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q และค่าเฉลี่ย λ โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสมเมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์คือ $P(X \leq q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ์ $P(X > q)$

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

<u>ตัวอย่าง</u> จากการสำรวจอุบัติเหตุที่เกิดขึ้นบนถนนสายหนึ่งพบว่า จะมีอุบัติเหตุเกิดขึ้นโดยเฉลี่ย 5 ครั้งต่อหนึ่งวัน

X คือจำนวนครั้งในการเกิดอุบัติเหตุในช่วงเวลา 1 วัน มีค่าที่เป็นไปได้คือ 0, 1, 2, 3, ...

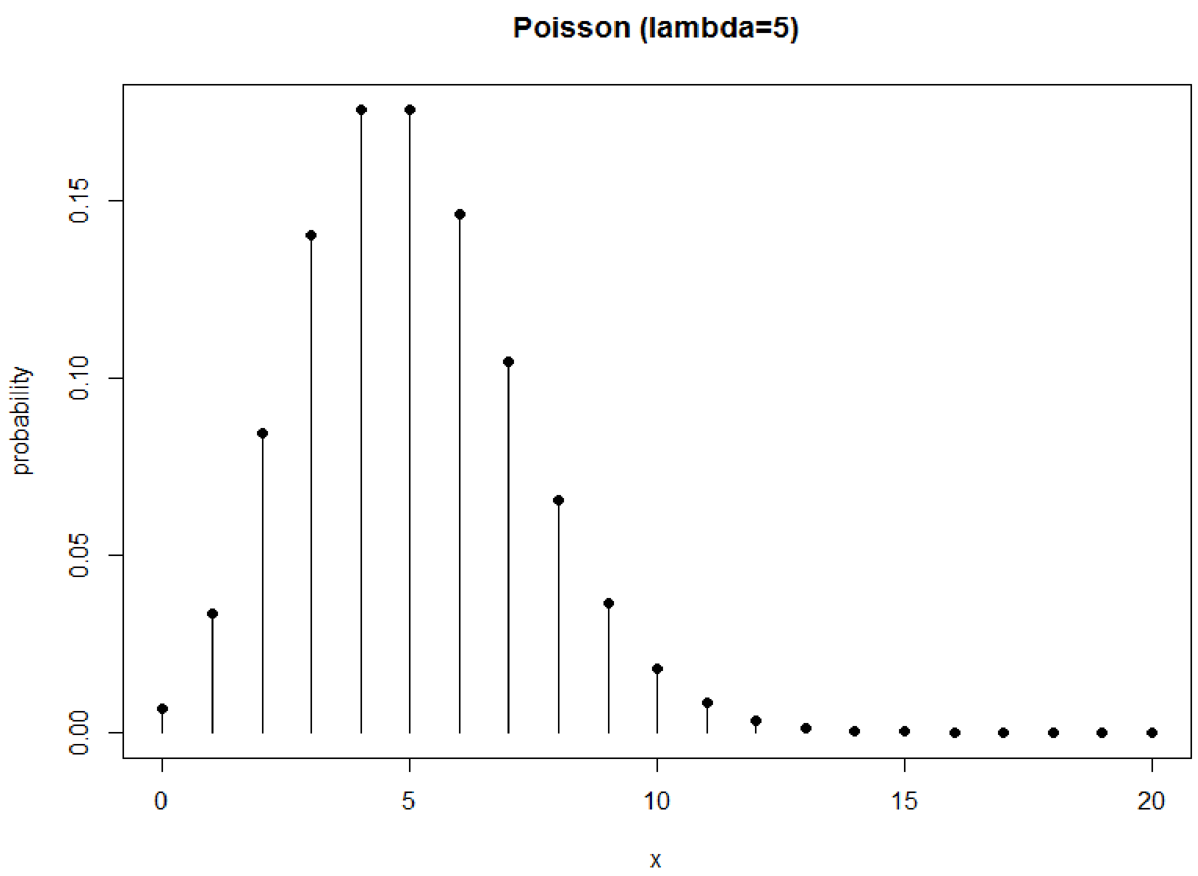
λ = 5 ครั้งต่อหนึ่งวัน

สร้างกราฟการแจกแจงแบบปัวซง

```
> x=0:20
> pr=dpois(x,5)
> plot(x,pr,type = "h",ylab="probability",main="Poisson (lambda=5)")
> points(0:20,pr,pch=16)
```

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

กราฟการแจกแจงปัวซง ที่มีค่าเฉลี่ยเป็น 5



2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

จงหาความน่าจะเป็นที่จะเกิดอุบัติเหตุบนถนนสายนี้

1. 6 ครั้งต่อวัน

2. ตั้งแต่ 6 ครั้งขึ้นไปต่อวัน

2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

จงหาความน่าจะเป็นที่จะเกิดอุบัติเหตุบนถนนสายนี้

3. ไม่ต่ำกว่า 3 ครั้งต่อวัน

4. น้อยกว่า 4 ครั้งต่อวัน

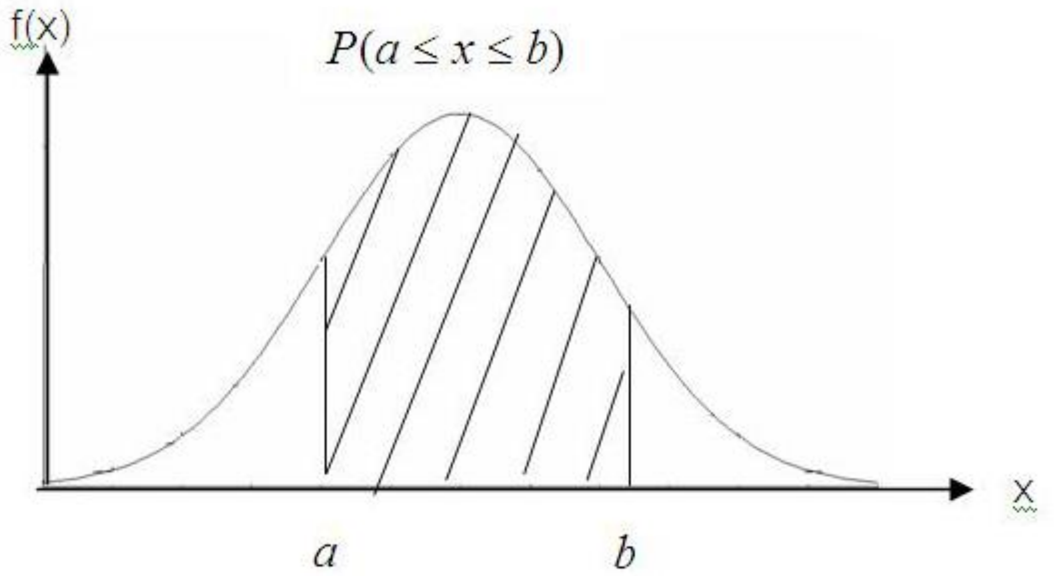
2.1.2 การแจกแจงแบบปัวซง (Poisson distribution)

จงหาความน่าจะเป็นที่จะเกิดอุบัติเหตุบนถนนสายนี้

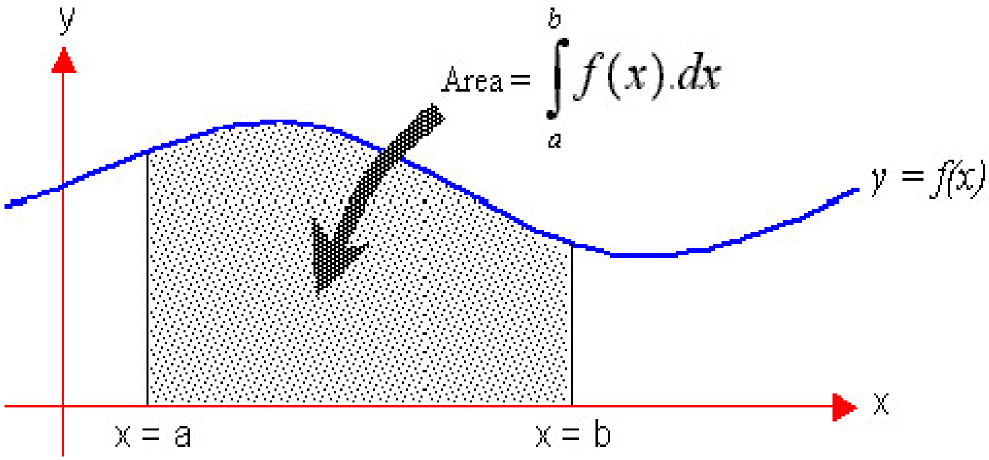
5. มากกว่า 12 ครั้งต่อ 2 วัน

3. การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง

- กรณีที่ X เป็นตัวแปรสุ่มแบบต่อเนื่อง
- ค่าของตัวแปรสุ่ม X จะมีค่าอยู่ในช่วง ๆ หนึ่ง และมีค่ามากมายไม่จำกัด
- ไม่สามารถหาค่าความน่าจะเป็นที่จะเกิดค่าของตัวแปรสุ่มแต่ละค่าได้
- เมื่อ X เป็นตัวแปรสุ่มแบบต่อเนื่องจะหาความน่าจะเป็นที่ x จะมีค่าในช่วงใดช่วงหนึ่ง เท่านั้น



3. การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง



การอินทิเกรทฟังก์ชันในช่วงของ X ที่กำหนด เป็นการหาพื้นที่ใต้โค้ง โดยพื้นที่ใต้โค้งนั้นจะแสดงถึงความน่าจะเป็นในช่วงของ [a, b]

$$P(a < x < b) = \int_a^b f(x)dx$$

$$P(X = c) = \int_c^c f(x)dx = 0$$

ความน่าจะเป็นที่ x จะมีค่าเท่ากับค่าใดค่าหนึ่งจะเป็นศูนย์

$$P(a < x < b) = P(a < x \le b) = P(a \le x < b) = P(a \le x \le b)$$

3. การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่องที่สำคัญ ได้แก่

3.1 การแจกแจงปกติ (Normal Distribution)

3.2 การแจกแจงที (t Distribution)

3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

3.4 การแจกแจงเอฟ (F Distribution)

3.1 การแจกแจงปกติ (Normal Distribution)

- กราฟของการแจกแจงจะมีลักษณะโค้งเป็นรูประฆังคว่ำ เรียกว่าโค้งปกติ (normal curve)

1. เส้นโค้งปกติจะมีลักษณะสมมาตรรอบค่าเฉลี่ย

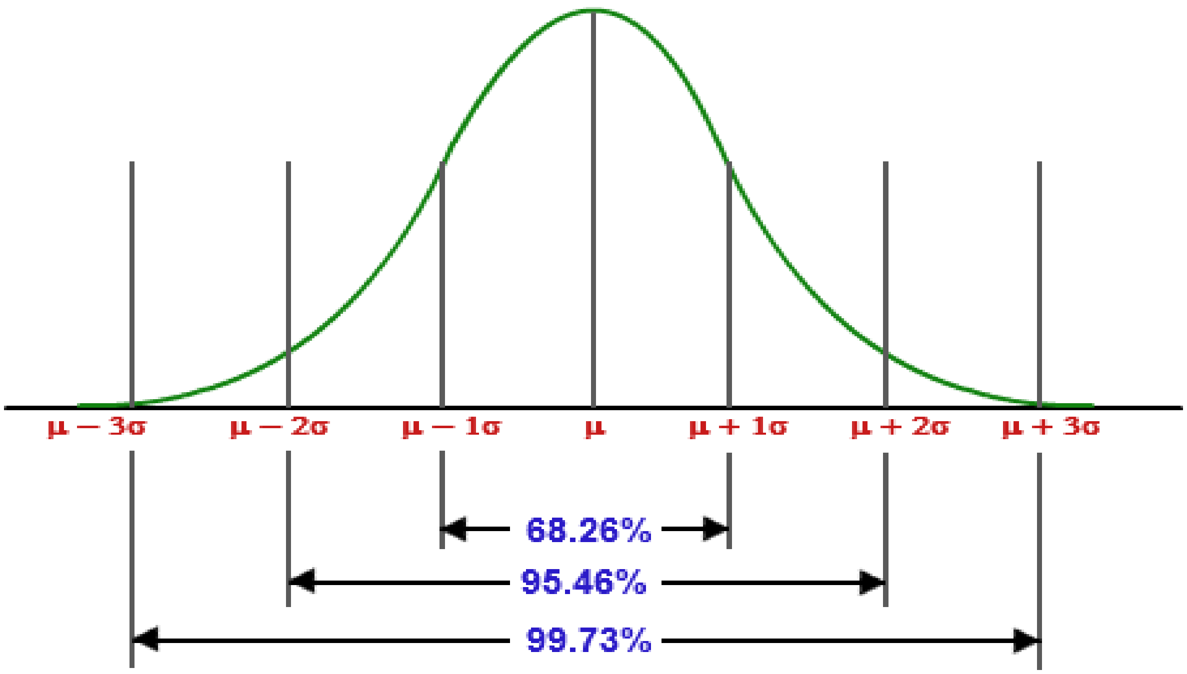
2. ค่าเฉลี่ย มัธยฐาน และฐานนิยม จะมีค่าเท่ากันและมีค่าอยู่ ณ ตำแหน่งกึ่งกลางของข้อมูล

3. พื้นที่ใต้โค้งปกติทั้งหมด คือความน่าจะเป็นซึ่งมีค่าเท่ากับ 1 หรือ 100%

4. พื้นที่ใต้โค้งปกติระหว่างค่าเฉลี่ย $\pm$1 ของค่าเบี่ยงเบนมาตรฐาน ประมาณ 68% พื้นที่ระหว่าง $\pm$2 ของค่าเบี่ยงเบนมาตรฐาน ประมาณ 95% และพื้นที่ระหว่าง $\pm$3 ของค่าเบี่ยงเบนมาตรฐาน ประมาณ 99%

3.1 การแจกแจงปกติ (Normal Distribution)

พื้นที่ใต้โค้งการแจกแจงปกติ



3.1 การแจกแจงปกติ (Normal Distribution)

- เนื่องจากข้อมูลส่วนใหญ่ที่ปรากฎอยู่ตามธรรมชาติมีการแจกแจงแบบปกติ เช่น ส่วนสูงหรือน้ำหนักคน ผลผลิตทางการเกษตร และการอุตสาหกรรมต่างๆ
- ตัวแปรสุ่มที่มีการแจกแจงปกติเรียกว่า ตัวแปรสุ่มปกติ (normal random variable)

μ เป็นค่าเฉลี่ยของตัวแปรสุ่ม

σ เป็นส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม

เขียนแทนด้วยสัญลักษณ์ $N(X; \mu, \sigma^2)$

- กราฟการแจกแจงปกติจะมีรูปร่างขึ้นอยู่กับ μ และ σ

3.1 การแจกแจงปกติ (Normal Distribution)

ตัวแปรสุ่ม X มีฟังก์ชันการแจกแจงปกติ ความน่าจะเป็นคือ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad , -\infty < x < \infty$$

π = 3.14159... , e = 2.71828...

μ คือ ค่าเฉลี่ย เมื่อ $-\infty < \mu < \infty$

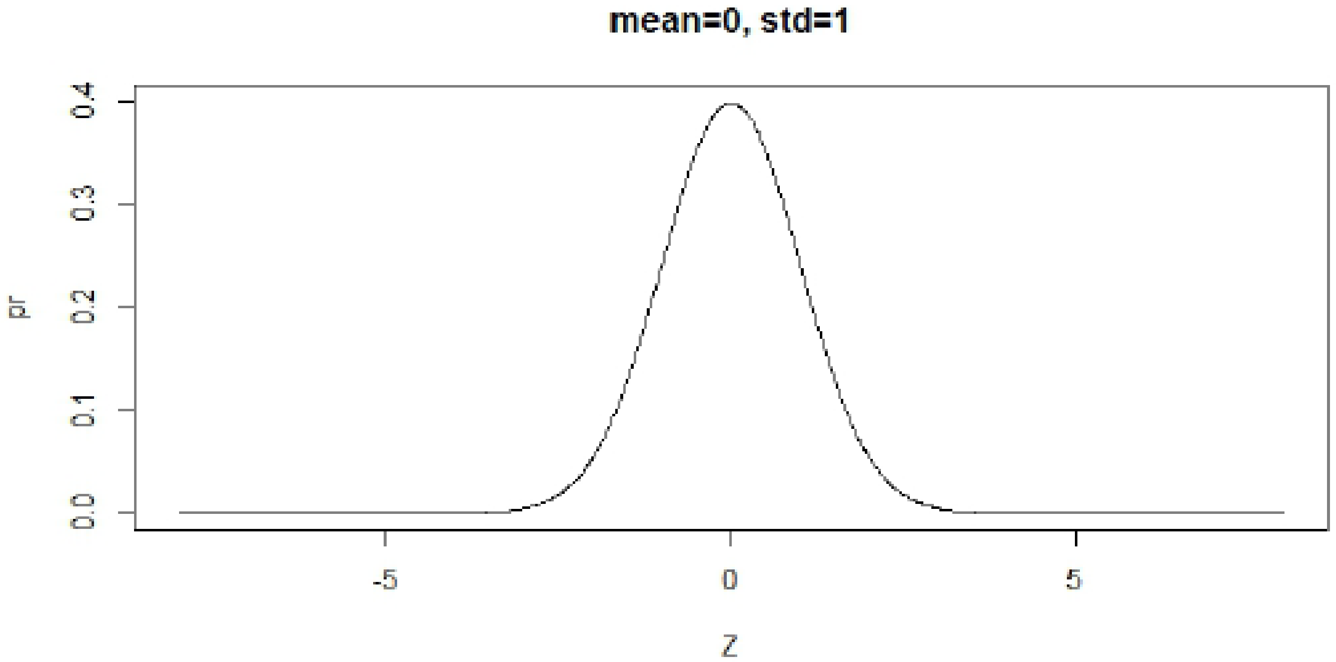
σ คือ ส่วนบี่ยงเบนมาตรฐาน เมื่อ $\sigma > 0$

ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่มปกติ

$$E(X) = \mu \text{ และ } V(X) = \sigma^2$$

3.1 การแจกแจงปกติ (Normal Distribution)

- การแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 เรียกว่า การแจกแจงแบบปกติมาตรฐาน (Standard Normal Distribution)
- ใช้อักษร Z แทนตัวแปรสุ่มของการแจกแจงแบบปกติมาตรฐาน



3.1 การแจกแจงปกติ (Normal Distribution)

ตัวแปรสุ่ม Z มีฟังก์ชันการแจกแจงปกติมาตรฐาน ความน่าจะเป็นคือ

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad , -\infty < z < \infty$$

π = 3.14159... , e = 2.71828...

ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่มปกติมาตรฐาน

$E(Z) = 0$ และ $V(Z) = 1$

3.1 การแจกแจงปกติ (Normal Distribution)

- ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น μ และความแปรปรวนเป็น σ^2 แล้ว ตัวแปรสุ่ม Z จะมีค่าเป็น

$$Z = \frac{X-\mu}{\sigma}$$

- การหาค่าความน่าจะเป็นของการแจกแจงแบบปกติคำนวณได้โดยการอินทิเกรตฟังก์ชันการแจกแจงความน่าจะเป็น คือ

$$P(Z < a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz$$

คำสั่งใน R การแจกแจงปกติ

Normal {stats} R Documentation

The Normal Distribution

Description

Density, distribution function, quantile function and random generation for the normal distribution with mean equal to `mean` and standard deviation equal to `sd`.

Usage

```
dnorm(x, mean = 0, sd = 1, log = FALSE)
pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
rnorm(n, mean = 0, sd = 1)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`mean`	vector of means.
`sd`	vector of standard deviations.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$ otherwise, $P[X > x]$.

คำสั่งใน R การแจกแจงปกติ

- pnorm คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q ค่าเฉลี่ย mean และ ส่วนเบี่ยงเบนมาตรฐาน sd โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสมเมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์ คือ $P(X \leq q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ์ $P(X > q)$

- qnorm คือการหาค่าของตัวแปรสุ่ม a เมื่อกำหนดค่าเฉลี่ย mean และส่วนเบี่ยงเบนมาตรฐาน sd สัญลักษณ์ คือ $P(X \leq a) = p$ เมื่อ lower.tail=TRUE และ สัญลักษณ์ คือ $P(X > a) = p$ เมื่อ lower.tail=FALSE

Default: mean = 0, sd =1 (Standard Normal Distribution)

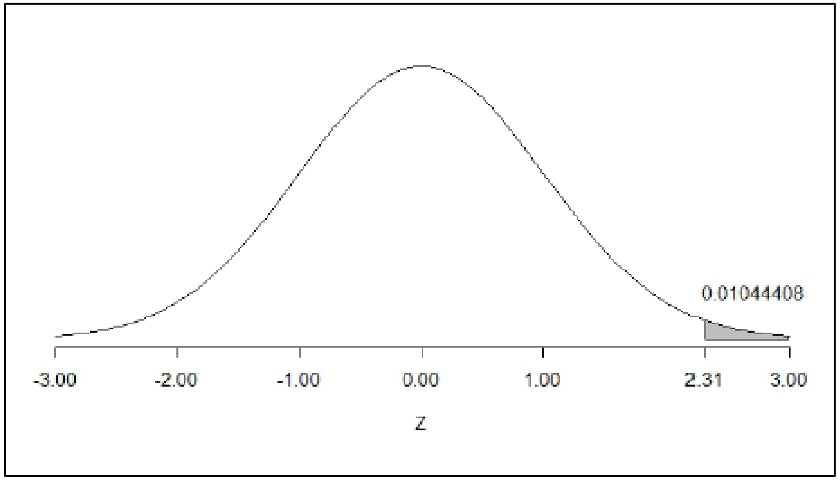
3.1 การแจกแจงปกติ (Normal Distribution)

ในการหาค่าความน่าจะเป็นของการแจกแจงแบบปกติเราจะสามารถคำนวณได้โดยการอินทิเกรตฟังก์ชันการแจกแจงความน่าจะเป็น

จงหาความน่าจะเป็นของตัวแปรสุ่ม Z ต่อไปนี้

1. $P(Z \geq 2.31)$

```
> 1-pnorm(2.31)
[1] 0.01044408
```



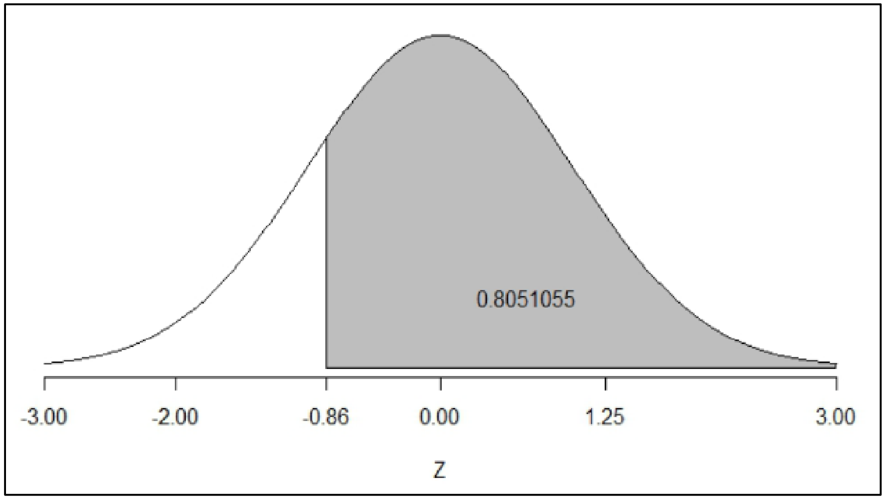
ดังนั้น $P(Z \geq 2.31) = 0.01044408$

3.1 การแจกแจงปกติ (Normal Distribution)

2. *P(Z > -0.86)*

```
> 1-pnorm(-0.86)
[1] 0.8051055
```

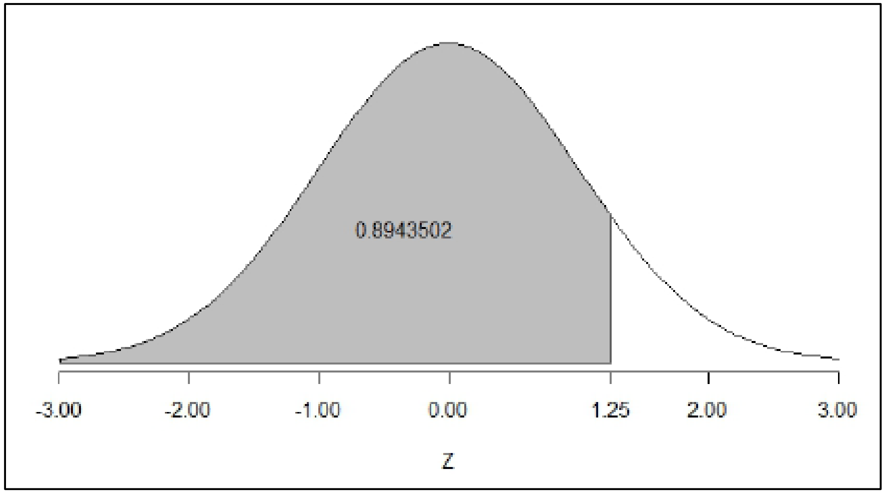
ดังนั้น P(Z > -0.86) = 0.8051055



3. *P*(*Z* < 1.25)

```
> pnorm(1.25)
[1] 0.8943502
```

ดังนั้น P(Z < 1.25) = 0.8943502

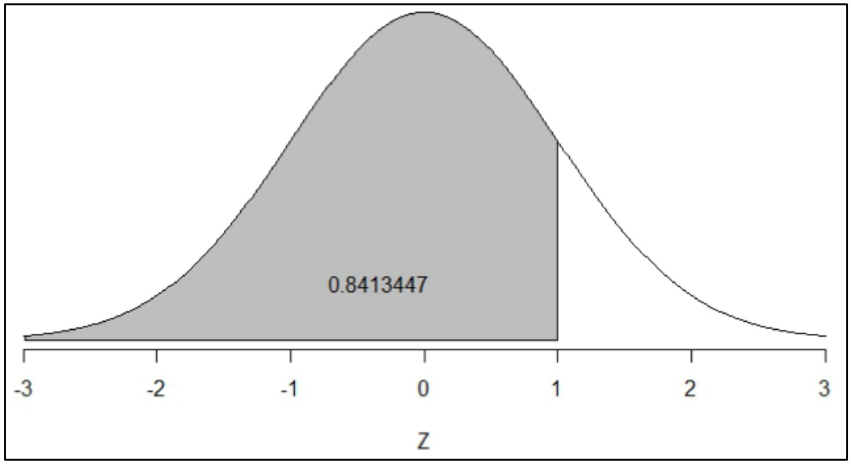


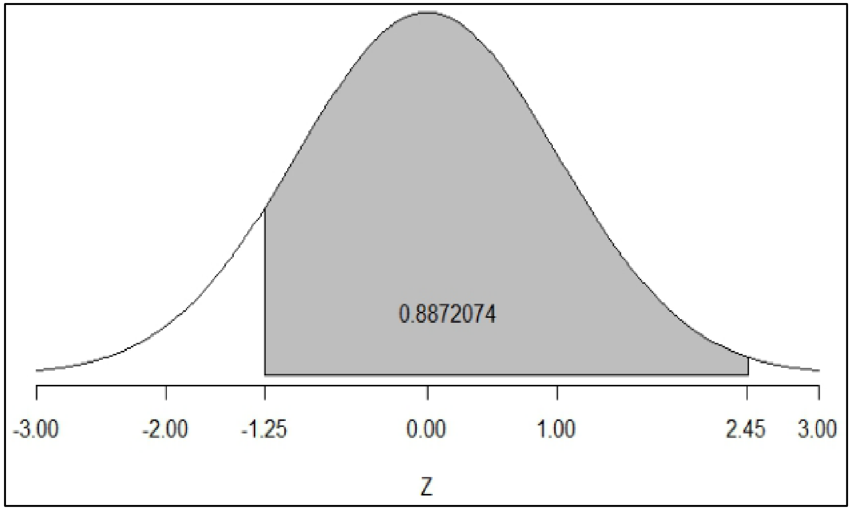
3.1 การแจกแจงปกติ (Normal Distribution)

4. $P(Z \leq 1.00)$

```
> pnorm(1)
[1] 0.8413447
```

ดังนั้น P(Z ≤ 1.00) = 0.8413447





5. $P(-1.25 < Z \leq 2.45)$

```
> pnorm(2.45)-pnorm(-1.25)
[1] 0.8872074
```

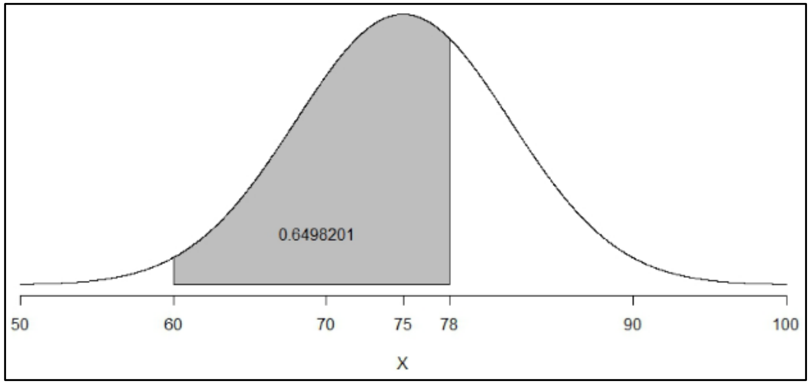
ดังนั้น P(-1.25 < Z ≤ 2.45) = 0.8872074

3.1 การแจกแจงปกติ (Normal Distribution)

ตัวอย่าง นักศึกษาชายของวิทยาลัยแห่งหนึ่งมีน้ำหนักเฉลี่ย 75 กิโลกรัม และส่วนเบี่ยงเบนมาตรฐาน 7 กิโลกรัม สมมติว่าน้ำหนักมีการแจกแจงแบบปกติ

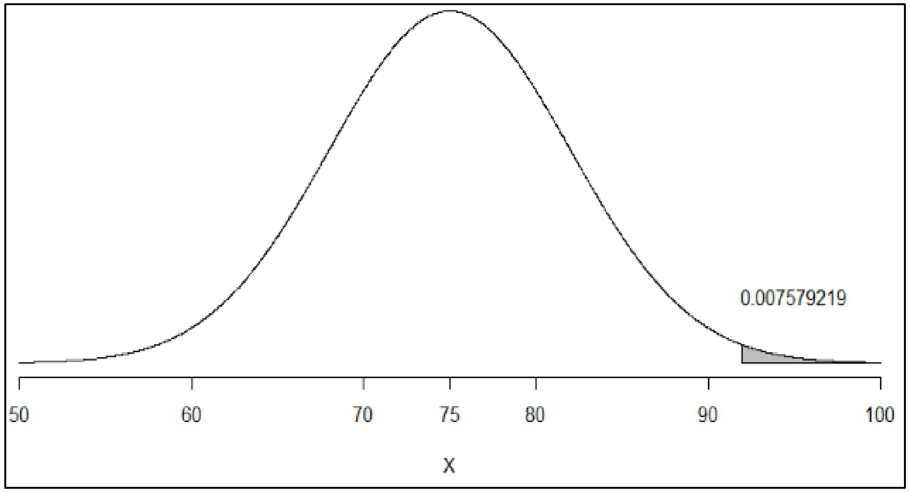
กำหนดให้ μ = 75 และ σ = 7 จงหาความน่าจะเป็นที่นักศึกษามีน้ำหนัก

1. ระหว่าง 60 และ 78 กิโลกรัม



3.1 การแจกแจงปกติ (Normal Distribution)

2. มากกว่า 92 กิโลกรัม



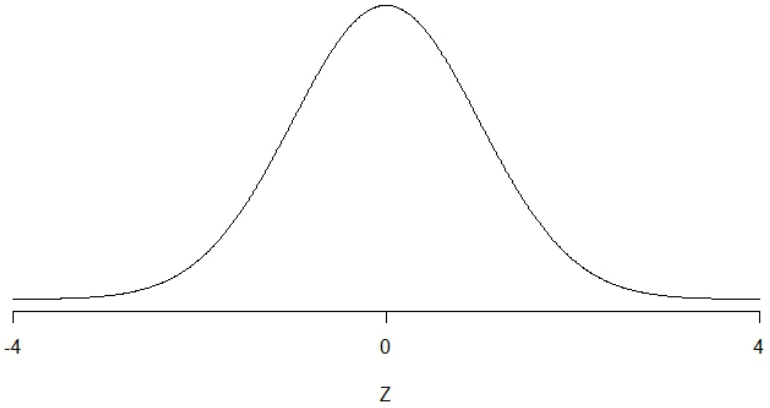
3.1 การแจกแจงปกติ (Normal Distribution)

จงหาค่า a ที่ทำให้

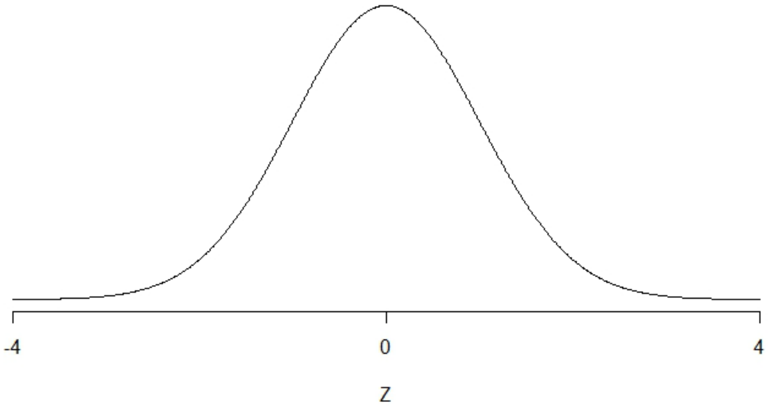
1. $P(Z < a) = 0.0025$

```
> qnorm(0.0025)
[1] -2.807034
```

ดังนั้น a = -2.807034

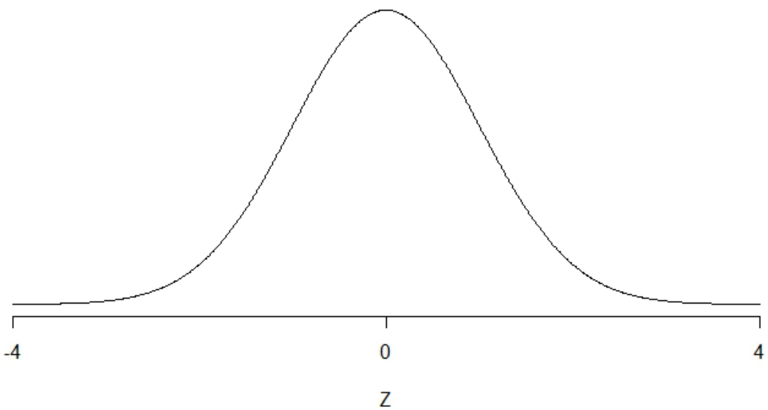


2. $P(Z \geq a) = 0.0244$

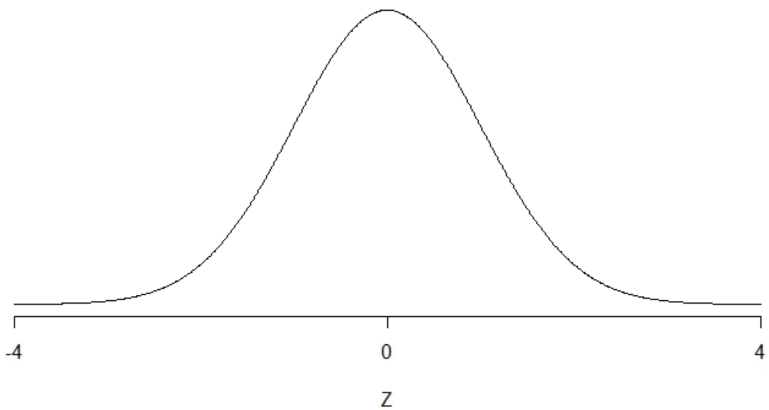


3.1 การแจกแจงปกติ (Normal Distribution)

3. $P(0 < Z < a) = 0.4808$



4. $P(-a < Z < a) = 0.995$



3.1 การแจกแจงปกติ (Normal Distribution)

<u>ตัวอย่าง</u> ถ้าคะแนนสอบวิชาสถิติของนิสิตกลุ่มหนึ่งมีการแจกแจงปกติด้วยคะแนนเฉลี่ย 60 คะแนน ค่าส่วนเบี่ยงเบนมาตรฐาน 15 คะแนน ถ้าอาจารย์ให้เกรด F แก่นิสิตที่ได้คะแนนต่ำสุด 10% อยากทราบว่านิสิตจะต้องได้คะแนนอย่างน้อยเท่าใดจึงจะไม่ได้เกรด F

กำหนดให้ μ = 60 และ σ = 15

จากสิ่งที่กำหนดให้คือ ถ้าอาจารย์ให้เกรด F แก่นิสิตที่ได้คะแนนต่ำสุด 10 %
แสดงว่า P(X < F) = 0.1

```
> qnorm(0.1,60,15)
[1] 40.77673
```

ดังนั้น *F* = 40.77673 คะแนน
นั่นคือนิสิตต้องได้คะแนนอย่างน้อย 40.78 คะแนนจึงจะไม่ได้เกรด F

3.2 การแจกแจงที (t Distribution)

ถ้า $z \sim N(0,1)$ และ $V \sim \chi^2_{(v)}$ ให้ T เป็นตัวแปรสุ่มที่มีค่าดังนี้

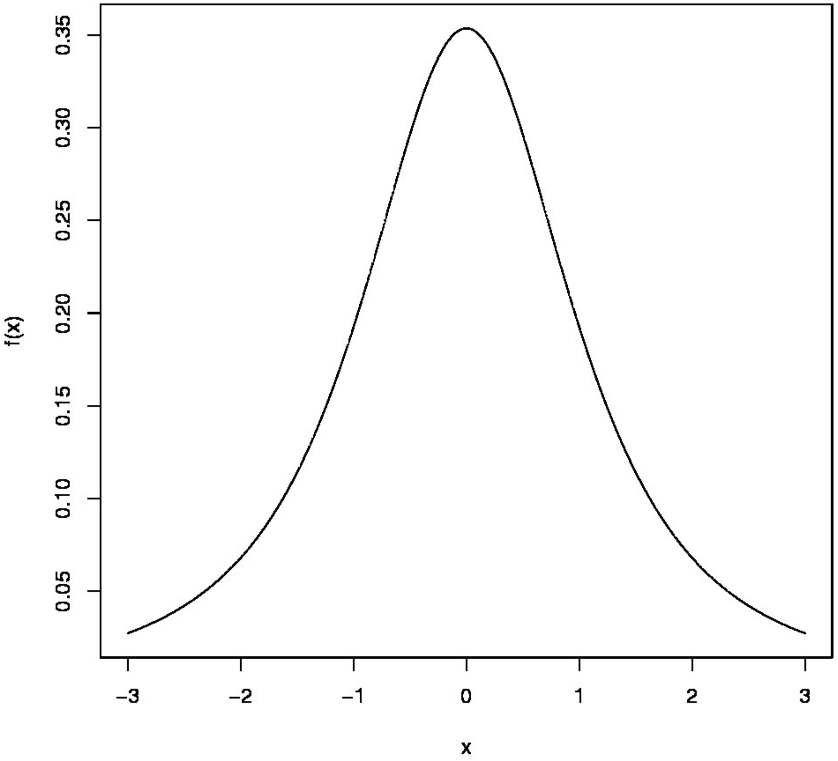
$$T = \frac{Z}{\sqrt{V/v}}$$

T จะมีการแจกแจงที (Student-t) ที่มีองศาเสรี = v และมีฟังก์ชันการแจกแจงความน่าจะเป็นคือ

$$f(t)\frac{\Gamma\left[\frac{\mathrm{v}+1}{2}\right]}{\sqrt{v\pi}\,\Gamma\left(\frac{\mathrm{v}}{2}\right)}\ [1+\frac{t^2}{v}]^{-(v+1)/2} \quad ,-\infty < t < \infty$$

3.2 การแจกแจงที (t Distribution)

กราฟการแจกแจงความน่าจะเป็น



คุณสมบัติ

1. กราฟของฟังก์ชันเป็นรูประฆังคว่ำ

2. สมการ $t = 0$ เป็นแกนสมมาตร

3. มีจุดสูงสุดอยู่จุดเดียว (unimodel) ที่ $t = 0$

4. พื้นที่ใต้กราฟรวมกันทั้งหมด มีค่าเท่ากับ 1

5. $P(t = t_0) = 0$

คำสั่งใน R การแจกแจงที

TDist {stats} R Documentation

The Student t Distribution

Description

Density, distribution function, quantile function and random generation for the t distribution with `df` degrees of freedom (and optional non-centrality parameter `ncp`).

Usage

```
dt(x, df, ncp, log = FALSE)
pt(q, df, ncp, lower.tail = TRUE, log.p = FALSE)
qt(p, df, ncp, lower.tail = TRUE, log.p = FALSE)
rt(n, df, ncp)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`df`	degrees of freedom (> *0*, maybe non-integer). `df = Inf` is allowed.
`ncp`	non-centrality parameter *delta*; currently except for `rt()`, only for `abs(ncp) <= 37.62`. If omitted, use the central t distribution.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

คำสั่งใน R การแจกแจงที

- pt คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q องศาเสรี df โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสมเมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์ คือ $P(X \leq q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ์ $P(X > q)$

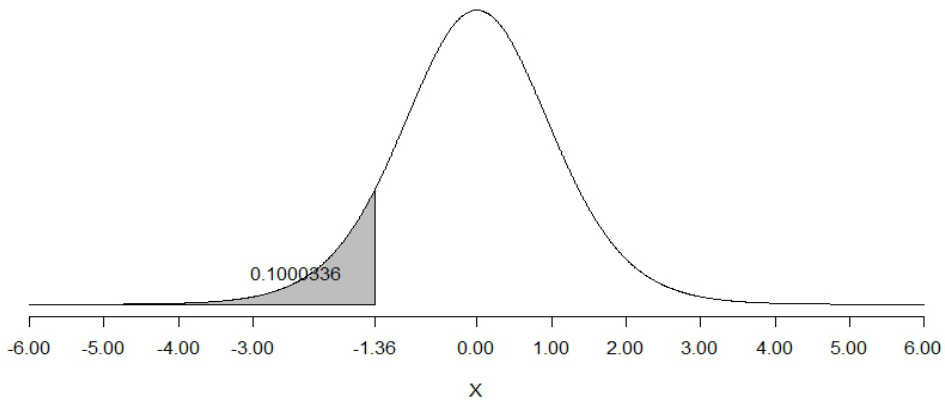
- qt คือการหาค่าของตัวแปรสุ่ม a เมื่อกำหนดค่า องศาเสรี df สัญลักษณ์คือ $P(X \leq a) = p$ เมื่อ lower.tail=TRUE และ สัญลักษณ์คือ $P(X > a) = p$ เมื่อ lower.tail=FALSE

3.2 การแจกแจงที (t Distribution)

จงหาค่าความน่าจะเป็นต่อไปนี้

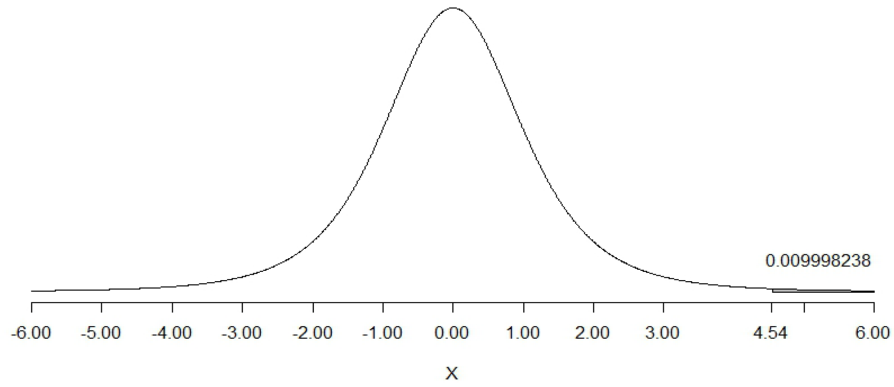
1. *df = 12, P(T < -1.356)*

```
> pt(-1.356,12)
[1] 0.1000336
```



2. *df = 3, P(T > 4.541)*

```
> 1-pt(4.541,3)
[1] 0.009998238
```

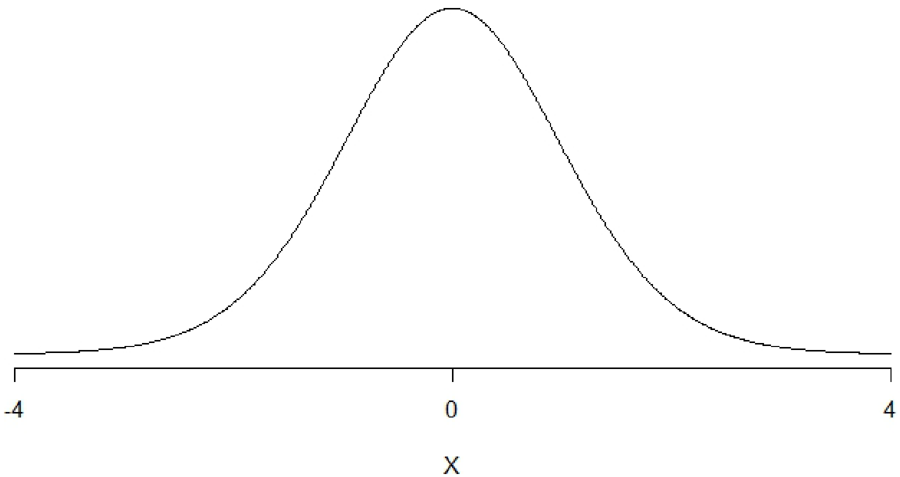


3.2 การแจกแจงที (t Distribution)

จงหาค่า a ต่อไปนี้

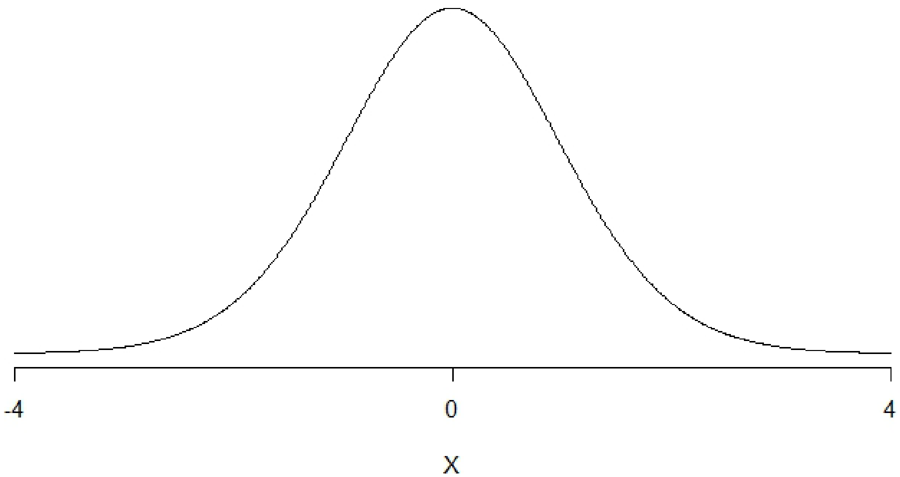
1. $t_{0.025,3} = a$

```
> qt(0.025,3)
[1] -3.182446
```



2. $t_{0.975,10} = a$

```
> qt(0.975,10)
[1] 2.228139
```

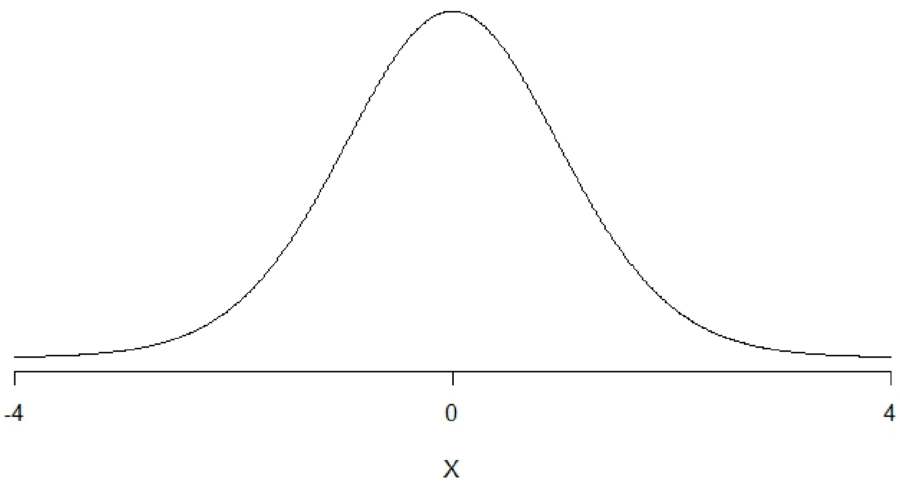


3.2 การแจกแจงที (t Distribution)

จงหาค่า a ต่อไปนี้

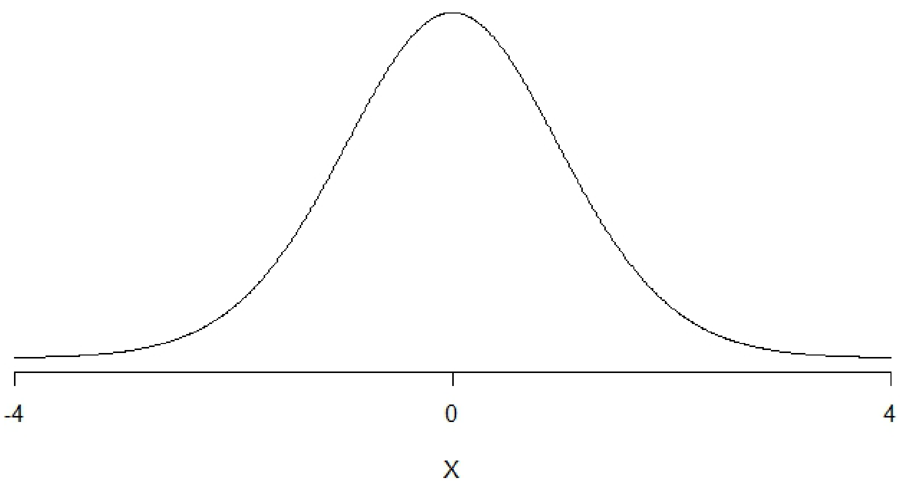
3. $df = 11$, $P(T > a) = 0.1$

```
> qt(0.9,11)
[1] 1.36343
```



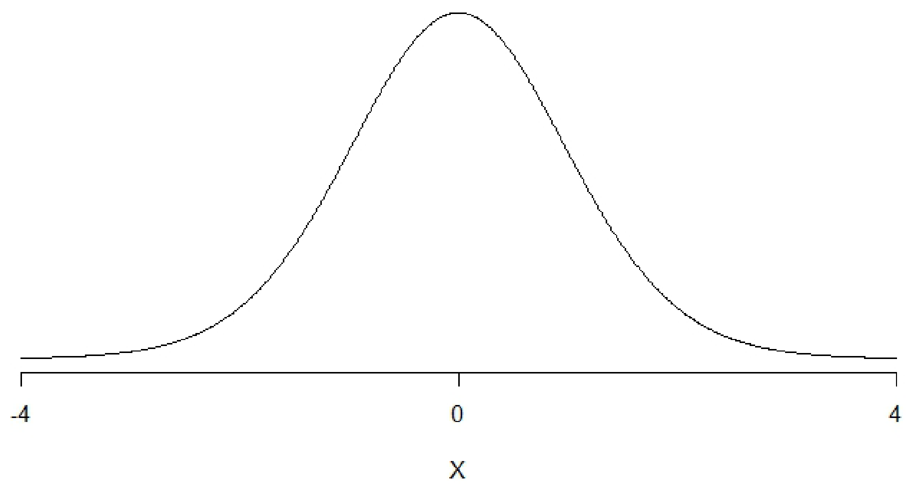
4. $df = 30$, $P(T < a) = 0.005$

```
> qt(0.005,30)
[1] -2.749996
```



3.2 การแจกแจงที (t Distribution)

จงหาค่า a ที่ทำให้ $P(-a \leq T \leq a) = 0.99$ เมื่อ $df = 23$



3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

ตัวแปรสุ่มไคกำลังสองมีฟังก์ชันการแจกแจงความน่าจะเป็นคือ

$$f(x) = \frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)} e^{-\frac{x}{2}} x^{\frac{v}{2}-1} \quad , 0 < x < \infty$$

และมีค่าเฉลี่ย $E(X) = v$ และ ความแปรปรวน $V(X) = 2v$

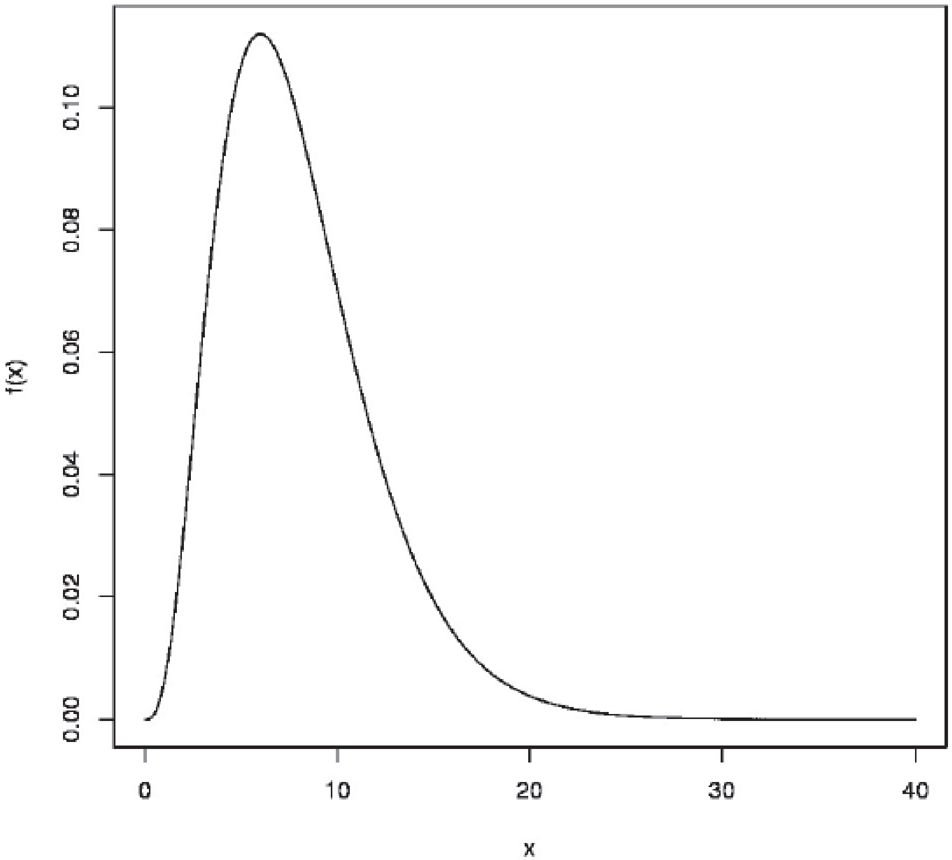
ถ้าตัวแปรสุ่ม $Z_1, Z_2, \ldots, Z_n$ จำนวน n ตัวเป็นอิสระต่อกันและมีการแจกแจงปกติมาตรฐาน (standard normal distribution) N(0,1) จะได้

$$Z_1^2, Z_2^2, \ldots, Z_n^2 \sim \chi^2_{(n)}$$

เมื่อ χ^2 อ่านว่า ไคกำลังสอง

3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

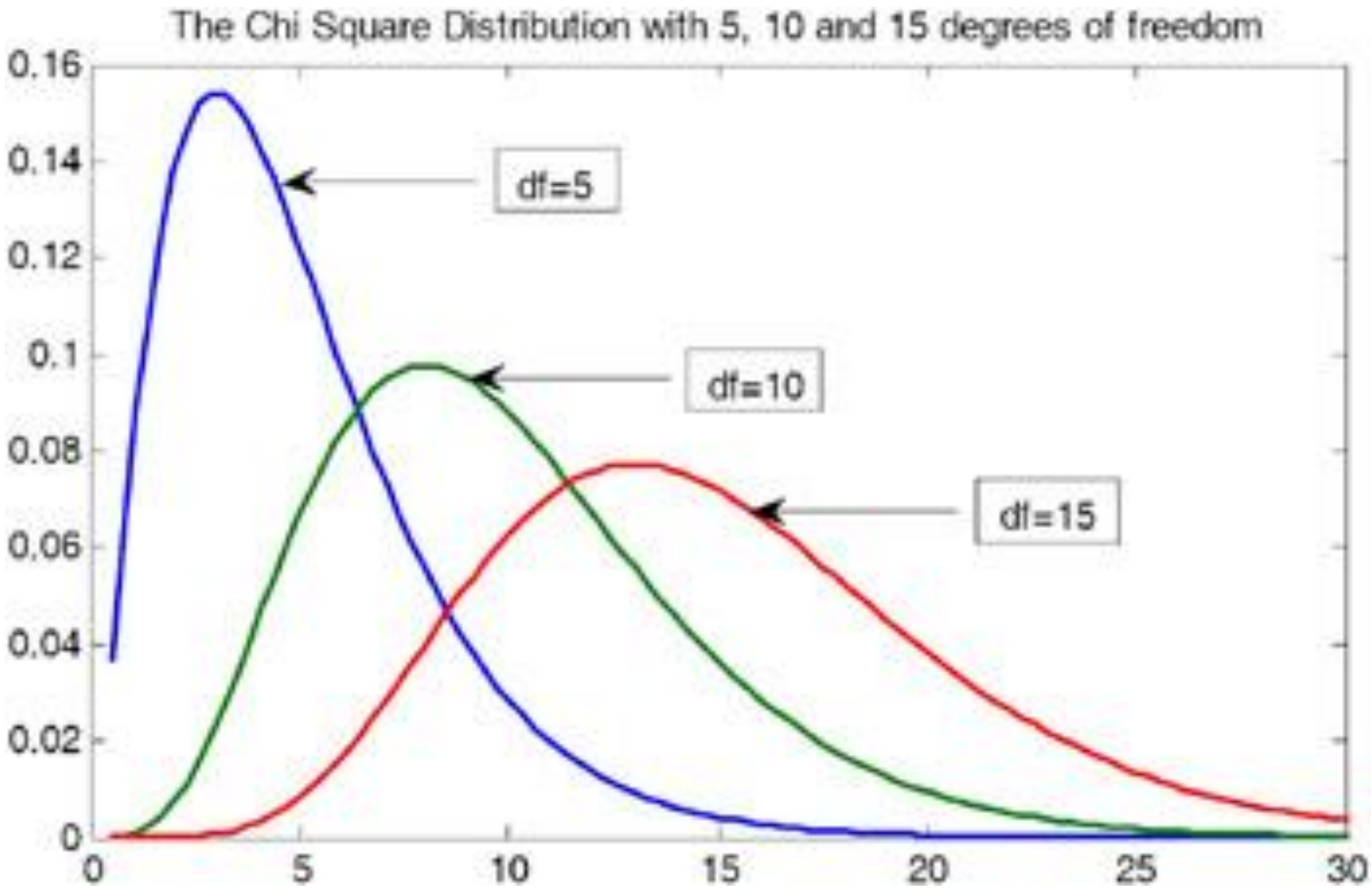
กราฟการแจกแจงความน่าจะเป็นไคกำลังสอง



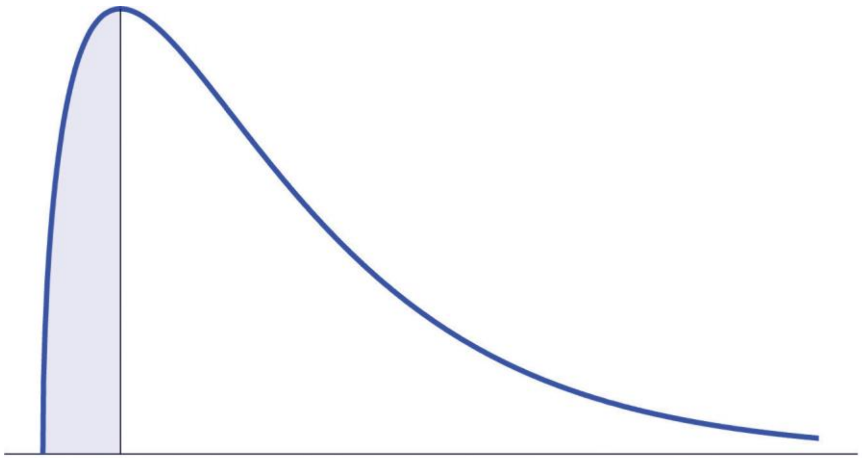
คุณสมบัติ

1. $\chi^2 \geq \mathbf{0}$
2. พื้นที่ใต้กราฟรวมกันทั้งหมดมีค่าเท่ากับ 1
3. กราฟมีลักษณะเบ้ขวา แตกแต่งกันตามค่าของ df
4. $\boldsymbol{P}(\chi^2 = \chi^2) = \mathbf{0}$

3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)



3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)



$\chi^2_{0.025,4}$ = a

$P(\chi^2_6 <$ a$) = 0.025$

df	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21

คำสั่งใน R การแจกแจงไคสแควร์

Chisquare {stats} R Documentation

The (non-central) Chi-Squared Distribution

Description

Density, distribution function, quantile function and random generation for the chi-squared (*chi*^2) distribution with `df` degrees of freedom and optional non-centrality parameter `ncp`.

Usage

```
dchisq(x, df, ncp = 0, log = FALSE)
pchisq(q, df, ncp = 0, lower.tail = TRUE, log.p = FALSE)
qchisq(p, df, ncp = 0, lower.tail = TRUE, log.p = FALSE)
rchisq(n, df, ncp = 0)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`df`	degrees of freedom (non-negative, but can be non-integer).
`ncp`	non-centrality parameter (non-negative).
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

คำสั่งใน R การแจกแจงไคสแควร์

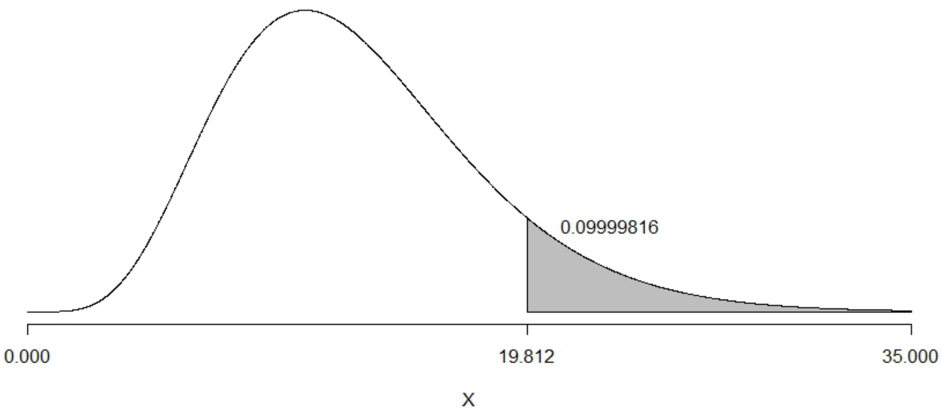
- pchisq คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q องศาเสรี df โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสม เมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์ คือ $P(X \le q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ $P(X > q)$
- qchisq คือการหาค่าของตัวแปรสุ่ม a เมื่อกำหนดค่า องศาเสรี df สัญลักษณ์คือ $P(X \le a) = p$ เมื่อ lower.tail=TRUE และ สัญลักษณ์คือ $P(X > a) = p$ เมื่อ lower.tail=FALSE

3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

จงหาความน่าจะเป็นต่อไปนี้

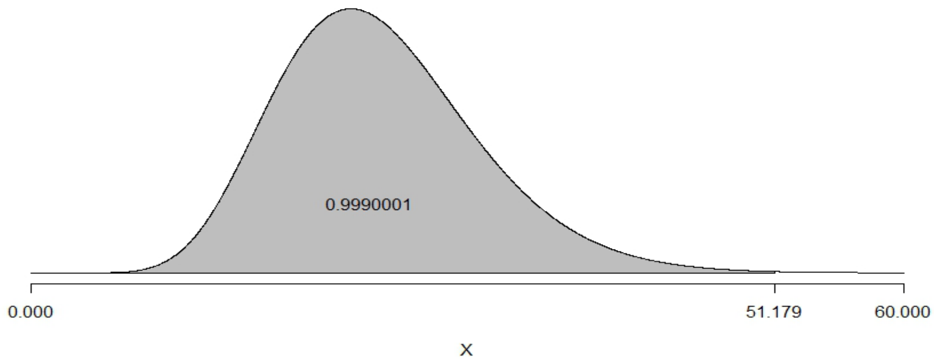
1. $P(\chi^2_{13} > 19.812)$

```
> 1-pchisq(19.812,13)
[1] 0.09999816
```



2. $P(\chi^2_{24} < 51.179)$

```
> pchisq(51.179,24)
[1] 0.9990001
```

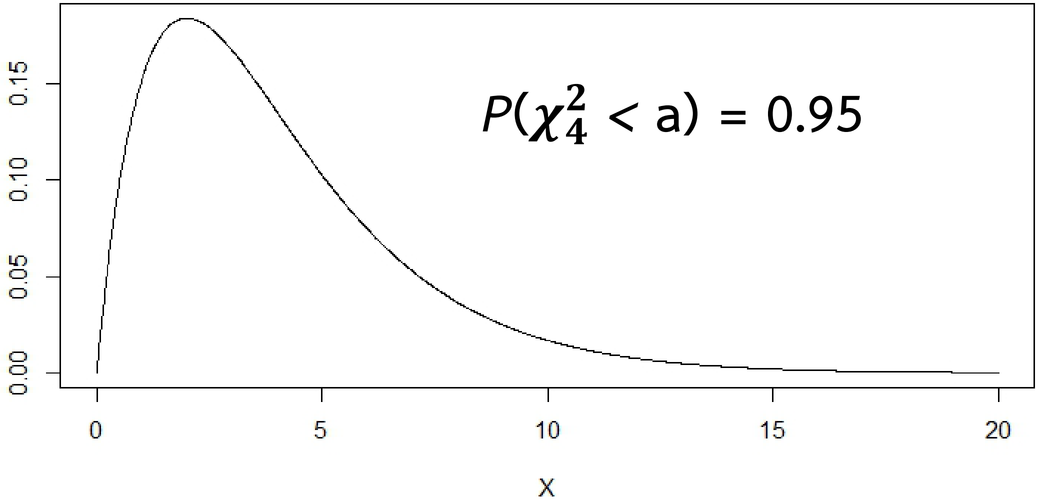


3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

จงหาค่า *a* ต่อไปนี้

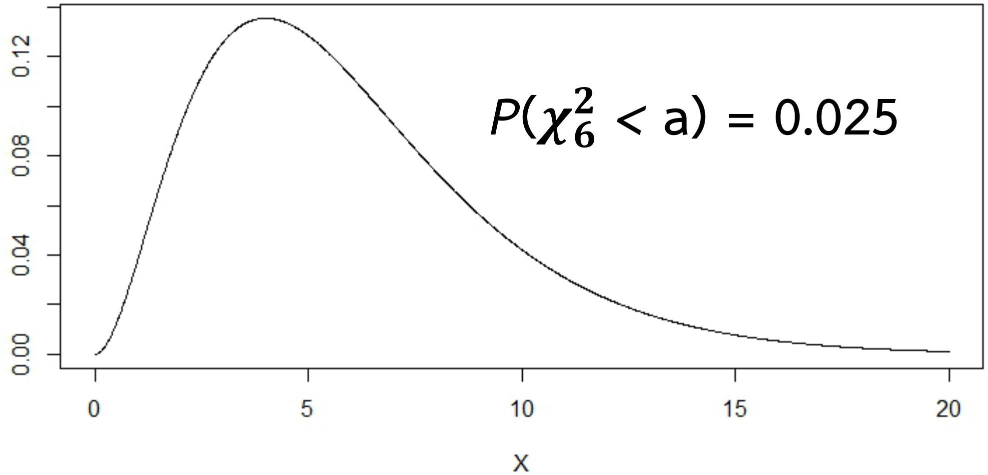
1. $\chi^2_{0.95,4}$ = a

```
> qchisq(0.95,4)
[1] 9.487729
```



2. $\chi^2_{0.025,6}$ = a

```
> qchisq(0.025,6)
[1] 1.237344
```

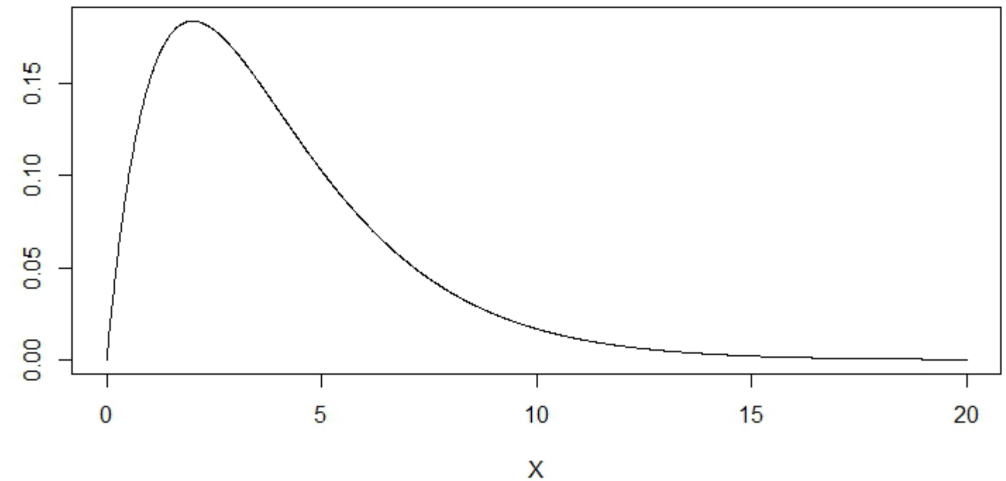


3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

จงหาค่า a ต่อไปนี้

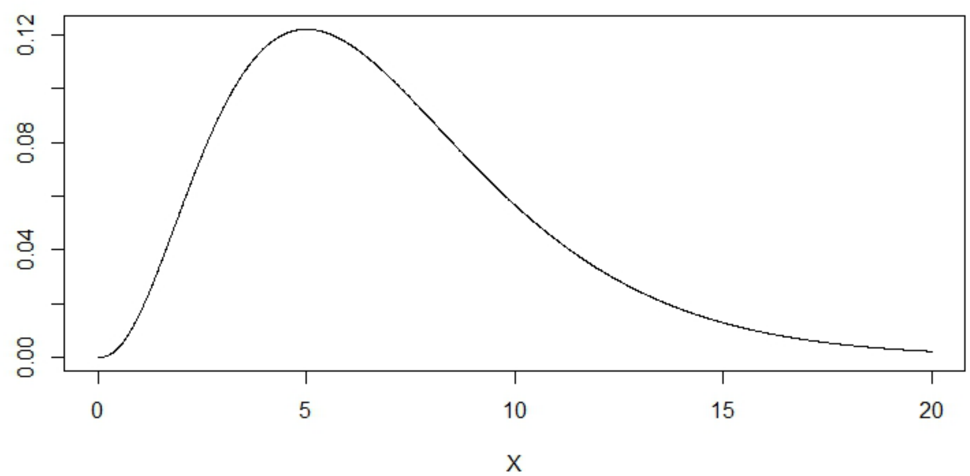
3. $P(\chi^2_5 > a)$ = 0.025

```
> qchisq(0.975,5)
[1] 12.8325
```



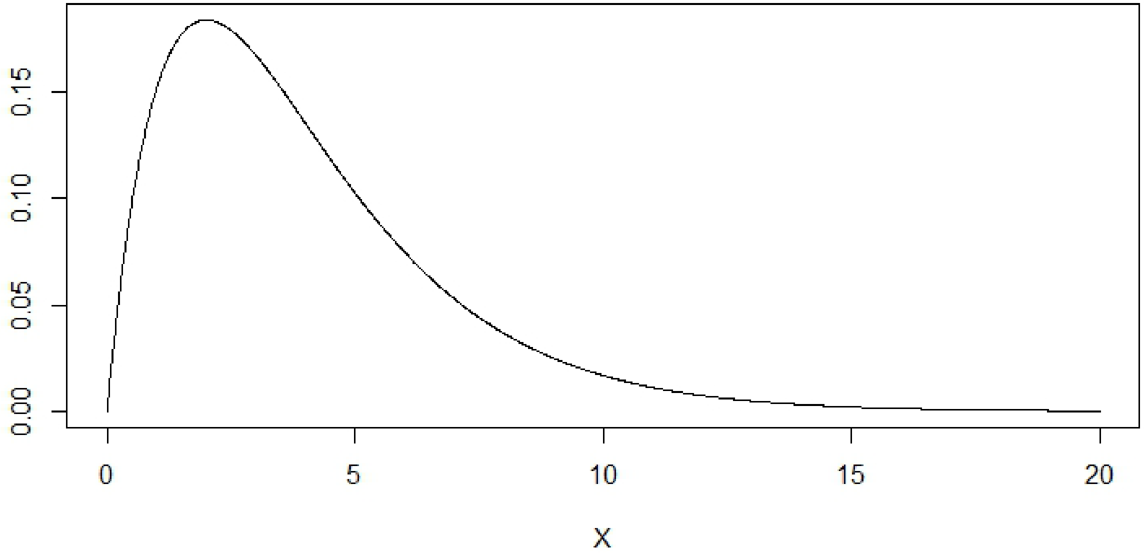
4. $P(\chi^2_7 < a)$ = 0.995

```
> qchisq(0.995,7)
[1] 20.27774
```



3.3 การแจกแจงไคสแควร์ (Chi-Square Distribution)

จงหาค่า a และ b ที่ทำให้ $P(a \leq \chi^2_9 \leq b) = 0.90$ และ $P(\chi^2_9 \leq a) = 0.05$



3.4 การแจกแจงเอฟ (F Distribution)

การแจกแจงเอฟเป็นอัตราส่วนของการแจกแจงไคสแควร์ที่เป็นอิสระจากกัน 2 ชุด

ถ้า U และ V เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยที่ $U \sim \chi^2_{(n)}$ และ $V \sim \chi^2_{(m)}$ จะได้ว่า

$$F = \frac{U/n}{V/m}$$

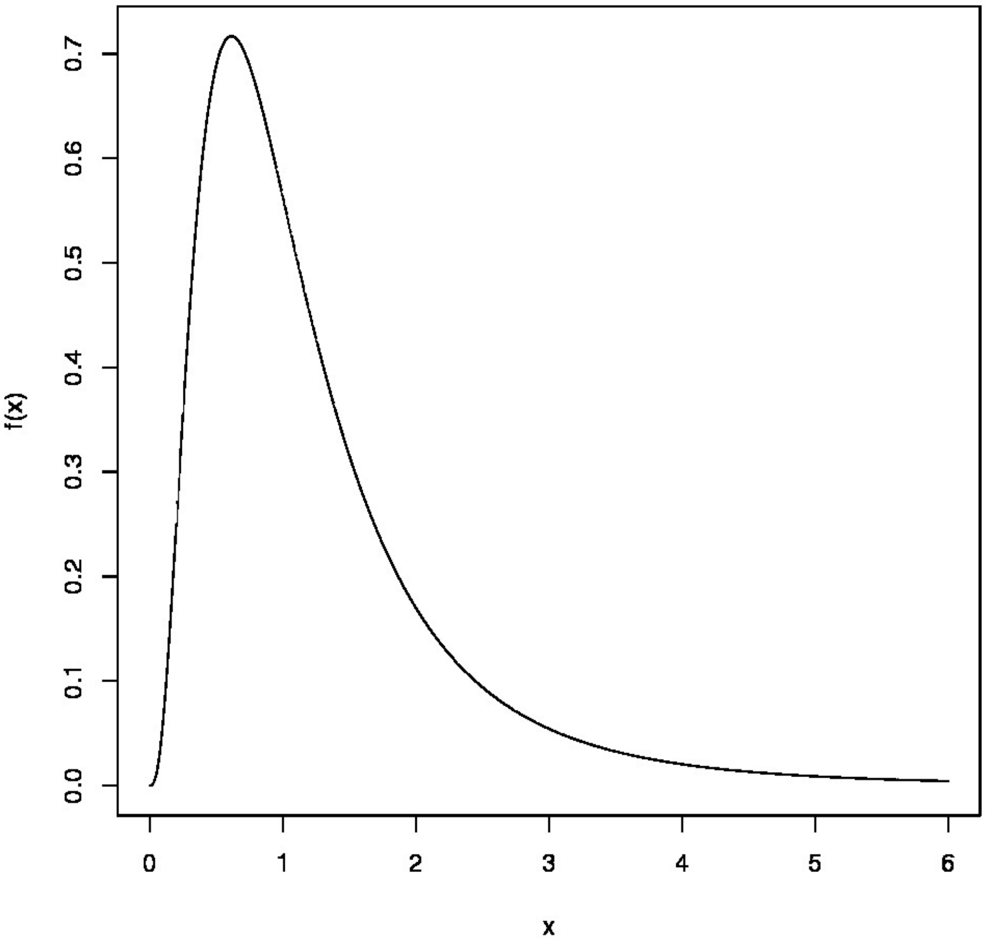
เป็นตัวแปรสุ่มที่มีการแจกแจงแบบ F มี $df1 = n$ และ $df2 = m$ และฟังก์ชันการแจกแจงความน่าจะเป็น

ของ F คือ

$$f(x) = \frac{\Gamma\left(\frac{n+m}{2}\right)\left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)\left(1+\frac{nx}{m}\right)^{\frac{n+m}{2}}}\ , x > 0$$

3.4 การแจกแจงเอฟ (F Distribution)

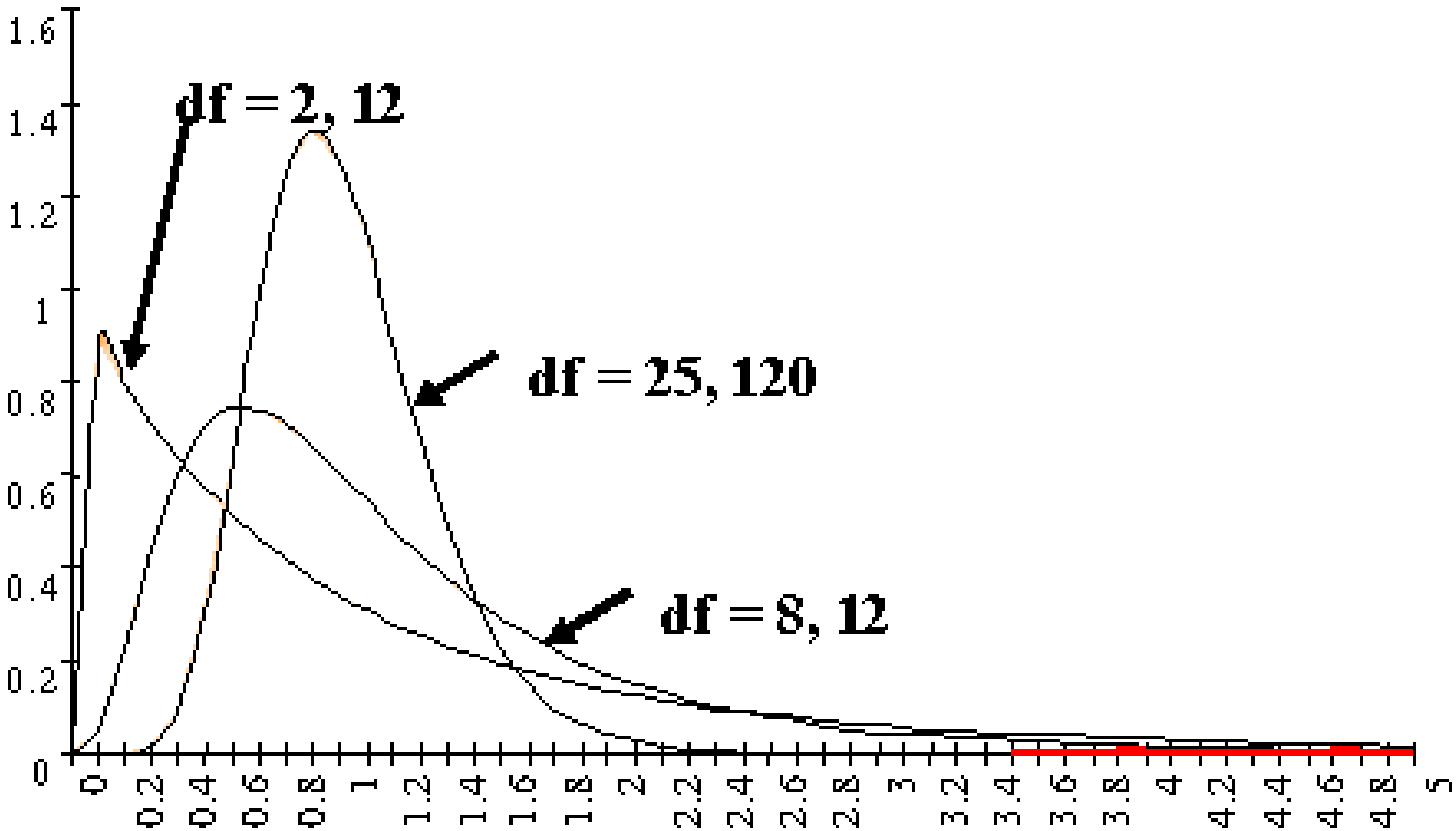
กราฟการแจกแจงความน่าจะเป็นของเอฟ



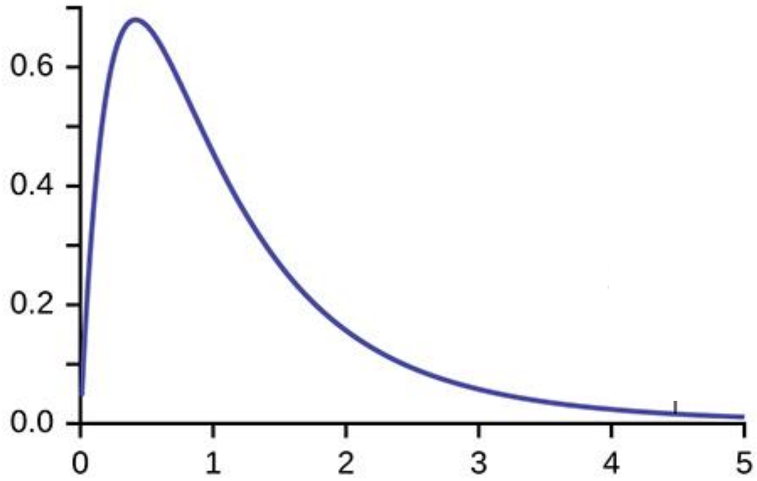
คุณสมบัติ

1. กราฟของฟังก์ชั่นมีลักษณะเบ้ขวา
2. พื้นที่ใต้กราฟรวมกันทั้งหมดมีค่าเท่ากับ 1
3. $F \geq 0$
4. $P(F = f) = 0$

3.4 การแจกแจงเอฟ (F Distribution)



3.4 การแจกแจงเอฟ (F Distribution)



$$F_{0.95,(1,4)} = a$$

$$P(F_{1,4} < a) = 0.95$$

$P(F \le f)$	Den. d.f. r_2	Numerator Degrees of Freedom, r_1							
		1	2	3	4	5	6	7	8
0.95	1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9
0.975		647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66
0.99		4052	4999.5	5403	5625	5764	5859	5928	5981
0.95	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37
0.975		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37
0.99		98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37
0.95	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85
0.975		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54
0.99		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49
0.95	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04
0.975		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98
0.99		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80

คำสั่งใน R การแจกแจงเอฟ

FDist {stats} R Documentation

The F Distribution

Description

Density, distribution function, quantile function and random generation for the F distribution with `df1` and `df2` degrees of freedom (and optional non-centrality parameter `ncp`).

Usage

```
df(x, df1, df2, ncp, log = FALSE)
pf(q, df1, df2, ncp, lower.tail = TRUE, log.p = FALSE)
qf(p, df1, df2, ncp, lower.tail = TRUE, log.p = FALSE)
rf(n, df1, df2, ncp)
```

Arguments

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`df1, df2`	degrees of freedom. `Inf` is allowed.
`ncp`	non-centrality parameter. If omitted the central F is assumed.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

คำสั่งใน R การแจกแจงเอฟ

- pf คือความน่าจะเป็นสะสมของตัวแปรสุ่ม เมื่อกำหนดค่าตัวแปรสุ่ม q องศาเสรี1 df1 และองศาเสรี2 df2 โดย lower.tail=TRUE หมายถึงความน่าจะเป็นสะสมเมื่อตัวแปรสุ่มมีค่าตั้งแต่ 0 ถึง q สัญลักษณ์ คือ $P(X \leq q)$ ถ้ากำหนด lower.tail=FALSE จะมีความหมายเป็นสัญลักษณ์ $P(X > q)$

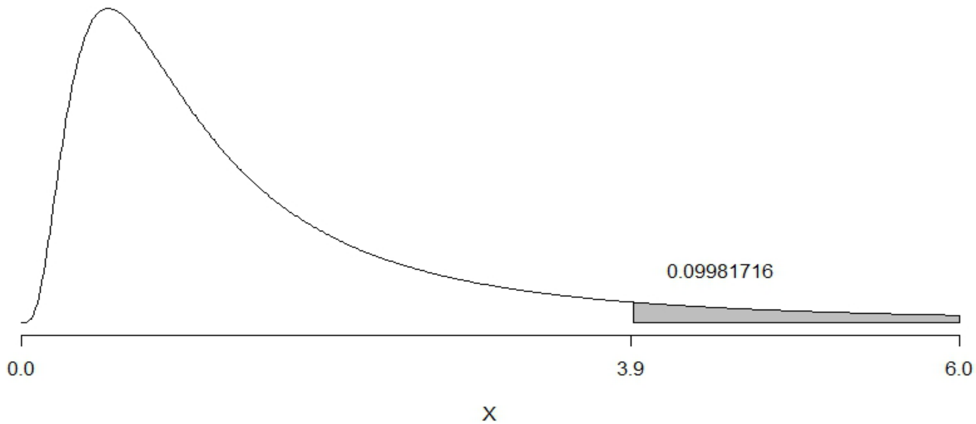
- qf คือการหาค่าของตัวแปรสุ่ม a เมื่อกำหนดค่า องศาเสรี1 df1 และองศาเสรี2 df2 สัญลักษณ์คือ $P(X \leq a) = p$ เมื่อ lower.tail=TRUE และ สัญลักษณ์ คือ $P(X > a) = p$ เมื่อ lower.tail=FALSE

3.4 การแจกแจงเอฟ (F Distribution)

จงหาค่าความน่าจะเป็นต่อไปนี้

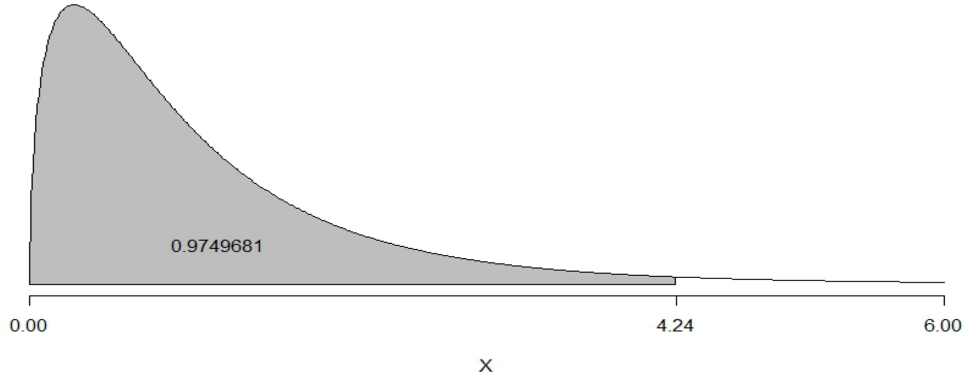
1. $P(F_{12,4} > 3.90) =$

```
> 1-pf(3.9,12,4)
[1] 0.09981716
```



2. $P(F_{3,14} < 4.24) =$

```
> pf(4.24,3,14)
[1] 0.9749681
```

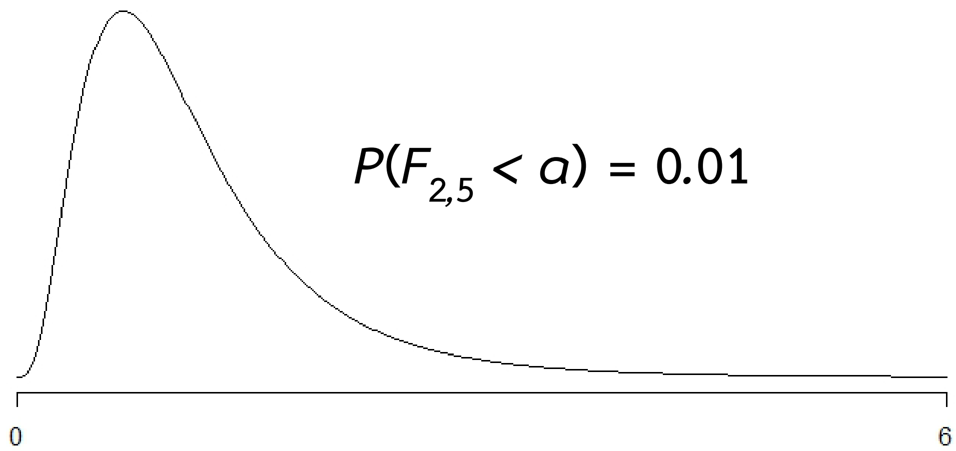


3.4 การแจกแจงเอฟ (F Distribution)

จงหาค่า a ต่อไปนี้

1. $F_{0.01,(2,5)} = a$

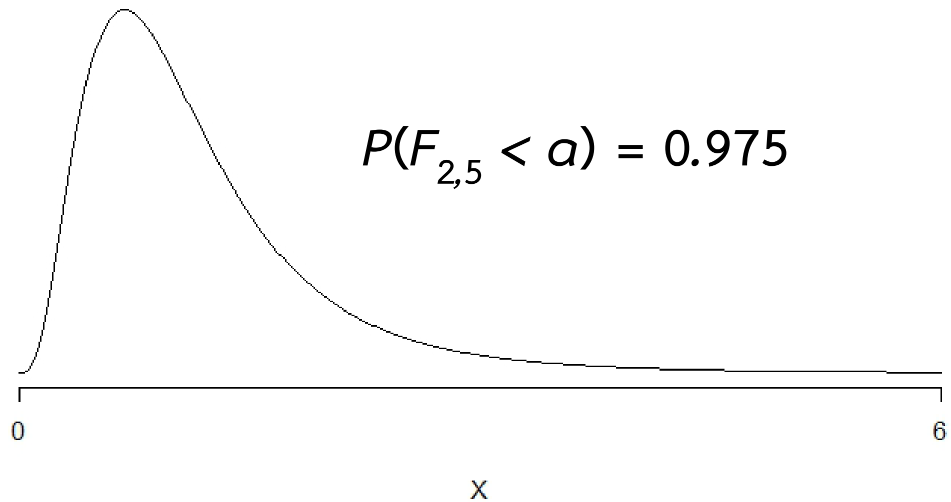
```
> qf(0.01,2,5)
[1] 0.01007056
```



$P(F_{2,5} < a) = 0.01$

2. $F_{0.975,(2,5)} = a$

```
> qf(0.975,2,5)
[1] 8.433621
```



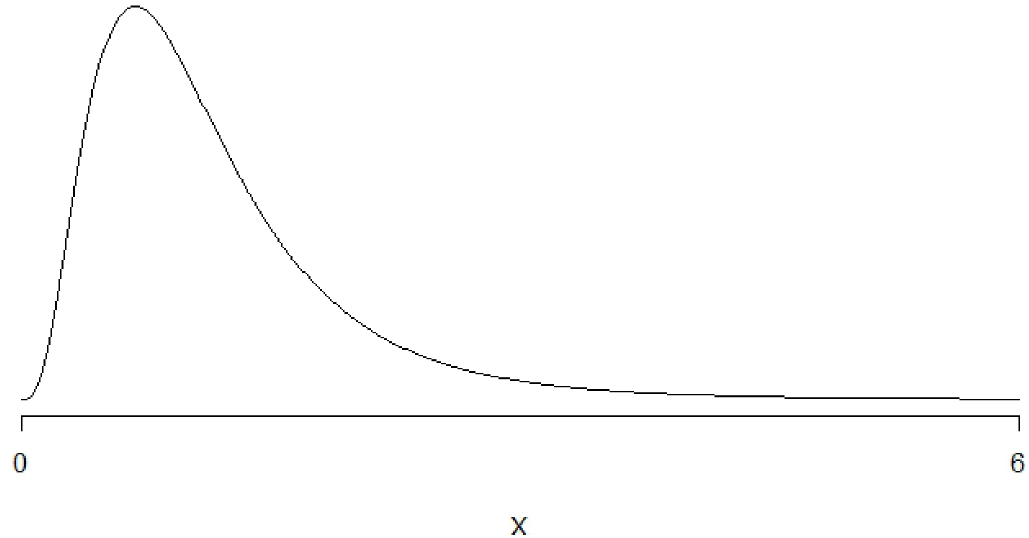
$P(F_{2,5} < a) = 0.975$

3.4 การแจกแจงเอฟ (F Distribution)

จงหาค่า *a* ต่อไปนี้

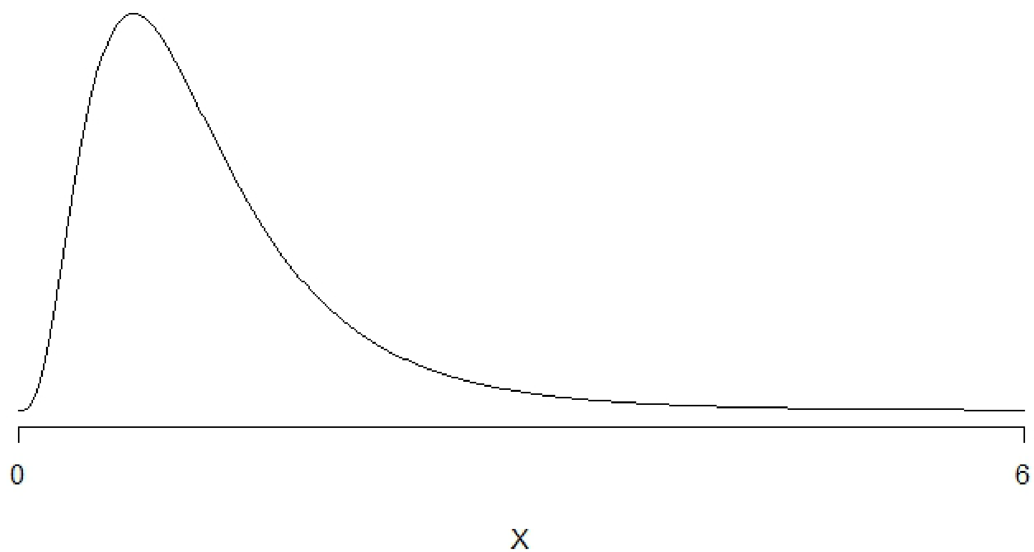
1. $P(F_{3,4} > a) = 0.05$

```
> qf(0.95,3,4)
[1] 6.591382
```



2. $P(F_{2,10} < a) = 0.025$

```
> qf(0.025,2,10)
[1] 0.02538202
```



3.4 การแจกแจงเอฟ (F Distribution)

ตัวอย่าง จงหาค่า a และ b ที่ทำให้ $P(a < F < b) = 0.90$

และ $P(F \leq a) = 0.05$ ที่ df เท่ากับ 9 และ 15