Machine Learning

Mathmatics

김선녕(ksycafe@gmail.com)

로그함수

행렬

미분

확률분포/정규분포

로그함수의 정의

$$y = log_a x \,,\; x = a^y$$

$$log_a a^x = x$$

$$log_a 1 = 0$$

$$log_a a = 1$$

$$log_a pq = log_a p + log_a q$$

$$log_a \frac{p}{q} = log_a p - log_a q$$

$$log_a p^q = q log_a p$$

$$a^{log_a p + log_a q} = a^{log_a p} \times a^{log_a q} = pq$$

$$log_a a^{log_a p + log_a q} = log_a p + log_a q$$

$$log_a b = \frac{1}{log_b a}$$

$$a^{log_a b} = b$$

로그함수
행렬
미분
확률분포/정규분포

행렬의 곱셈

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{ms} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1m}b_{m1} & a_{11}b_{12} + a_{12}b_{22} + \dots + a_{1m}b_{m2} & \dots & a_{11}b_{1s} + a_{12}b_{2s} + \dots + a_{1m}b_{ms} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2m}b_{m1} & a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2m}b_{m2} & \dots & a_{21}b_{1s} + a_{22}b_{2s} + \dots + a_{2m}b_{ms} \\ \dots & \dots & \dots & \dots \\ a_{n1}b_{11} + a_{n2}b_{21} + \dots + a_{nm}b_{m1} & a_{n1}b_{12} + a_{n2}b_{22} + \dots + a_{nm}b_{m2} & \dots & a_{n1}b_{1s} + a_{n2}b_{2s} + \dots + a_{nm}b_{ms} \end{bmatrix}$$

행렬A의 열 크기와 행렬 B의 행 크기가 같아야 함

$$A \quad \times \quad B \quad = \quad C$$
$$[\mathbf{3} \times \mathbf{2}] \quad [\mathbf{2} \times \mathbf{4}] \quad [\mathbf{3} \times \mathbf{4}]$$

(일치)

$$\boldsymbol{A} \quad \times \quad \boldsymbol{B} \quad = \quad \boldsymbol{C}$$
$$[\mathbf{3} \times \mathbf{2}] \quad [\mathbf{2}] \quad [\mathbf{3}]$$

(일치)

대각행렬/단위행렬(항등행렬)

- 영행렬(Zero Matrix)

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- 대각행렬(Diagonal Matrix)
  - n차 정사각행렬에서 대각원소 $a_{11}, a_{22}, \cdots, a_{nn}$이외의 모든 원소가 0인 행렬

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

- 단위행렬(항등행렬, Unit Matrix, Identity Matrix: I)
  - 대각행렬에서 대각원소가 모두 1인 행렬
  - $AI = IA = A$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

전치행렬/대칭행렬

- 전치행렬$(Transpose\ Matrix)$
  - $n \times m$행렬$A = [a_{ij}]$가 있을 때, 행과 열을 바꾼 $m \times n$행렬

$$\mathrm{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

- 대칭행렬$(Symmetric\ Matrix)$
  - n차 정방행렬 $A = [a_{ij}]$가 있을 때, $\boldsymbol{A^T} = \boldsymbol{A}$인 행렬

$$\mathrm{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}, \qquad \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

- 반대칭행렬$(Skew - Symmetric\ Matrix)$
  - n차 정방행렬 $A = [a_{ij}]$가 있을 때, $\boldsymbol{A^T} = -\boldsymbol{A}$인 행렬

$$\mathrm{A} = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 6 \\ -3 & -5 & 0 \end{bmatrix}, \qquad \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 0 & -2 & -3 \\ 2 & 0 & -5 \\ 3 & 6 & 0 \end{bmatrix} = -\mathrm{A}$$

전치행렬의 성질

행렬 A, B와 스칼라 α에 대해 다음 성질을 만족한다

$(A^T)^T = A$

$(A+B)^T = A^T + B^T$

$(AB)^T = B^T A^T$

$(\alpha A)^T = \alpha A^T$

A가 가역이면, $(A^T)^{-1} = (A^{-1})^T$

역행렬(Inverse Matrix)

- 정사각행렬 A에 대해 $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$를 만족하게 하는 행렬 $\boldsymbol{B}$

$$\boldsymbol{AA^{-1}} = \boldsymbol{A^{-1}A} = \boldsymbol{I}$$

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} w + 2y & x + 2z \\ w + 3y & x + 3z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

역행렬의 성질

n차 정방행렬 A, B가 가역일 때 다음 성질을 만족한다

A^{-1}는 가역이고, $(A^{-1})^{-1} = A$ 이다.

AB는 가역이고, $(AB)^{-1} = B^{-1}\ A^{-1}$이다

αA는가역이고, $(\alpha A)^{-1} = \dfrac{1}{\alpha} A^{-1}$이다($\alpha$는 0이 아닌 스칼라)

A^k 는 가역이고, $(A^{-1})^k = \left(A^k\right)^{-1}$ 이다($k \geq 0$ 정수)

※ 역행렬이 존재하는 행렬을 가역$(invertible)$행렬이라 한다

직교행렬(Orthogonal Matrix)

- 두 행렬은 직교. 직교행렬의 행렬식은 +1 혹은 -1이다
  $\boldsymbol{A^T A = A A^T = I}$
- 직교행렬의 전치행렬은 역행렬과 같다
  $\boldsymbol{A^T = A^{-1}}$

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

$$A^T A = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

- 2차 정사각행렬

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

- 3차 정사각행렬

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$\det(B)$$

$$= (b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{32}b_{21}) - (b_{13}b_{22}b_{31} + b_{23}b_{32}b_{11} + b_{33}b_{21}b_{12})$$

행렬의 성질

$A + B = B + A$
$A + (B + C) = (A + B) + C$
$A + O = O + A = A$
$A + (-A) = (-A) + A = O$
$(-1)A = -A$
$k(A + B) = kA + kB$
$(k + l)A = kA + lA$
$(kl)A = k(lA)$
$k(AB) = (kA)B = A(kB)$
$IA = A = AI$

O: 영행렬
I: 단위행렬

n-차원 데이터를 위한 학습 모델

$$\boldsymbol{y} = \boldsymbol{w_1 x_1} + \boldsymbol{w_2 x_2} + \boldsymbol{w_3 x_3}$$

$$y = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = (w_1 x_1 + w_2 x_2 + w_3 x_3)$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \cdots x_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ . \\ . \\ w_n \end{pmatrix} = (w_1 x_1 + w_2 x_2 + w_3 x_3 + \cdots + w_n x_n)$$

Iris 데이터셋

Sepal length	Sepal width	Petal length	Petal width	Species
5.2	3.5	1.4	0.2	*I. setosa*
4.9	3.0	1.4	0.2	*I. setosa*
4.7	3.2	1.3	0.2	*I. setosa*
4.6	3.1	1.5	0.2	*I. setosa*

$$x_1 = (5.2 \;\; 3.5 \;\; 1.4 \;\; 0.2), \qquad x_2 = (4.9 \;\; 3.0 \;\; 1.4 \;\; 0.2),$$
$$x_3 = (4.7 \;\; 3.2 \;\; 1.3 \;\; 0.2), \; \cdots \; , \; x_{150} = (5.9 \;\; 3.0 \;\; 5.1 \;\; 1.8)$$

$$X = \begin{pmatrix} 5.2 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 1.4 & 1.3 & 0.2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 6.2 & 3.4 & 5.4 & 2.3 \\ 5.9 & 3.0 & 5.1 & 1.8 \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{149,1} & x_{149,2} & x_{149,3} & x_{149,4} \\ x_{150,1} & x_{150,2} & x_{150,3} & x_{150,4} \end{pmatrix}$$

예제


```
import numpy as np
X = np.array([[1,2,3],[4,5,6],[7,8,9]])
X.shape
```

(3, 3)

```
W = np.array([[1,3,5,2,4,6,1,3,5,2],[4,6,1,3,5,2,4,6,1,1],[2,4,6,1,
3,5,2,4,6,5]])
W.shape
```

(3, 10)

```
print(W)
```

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

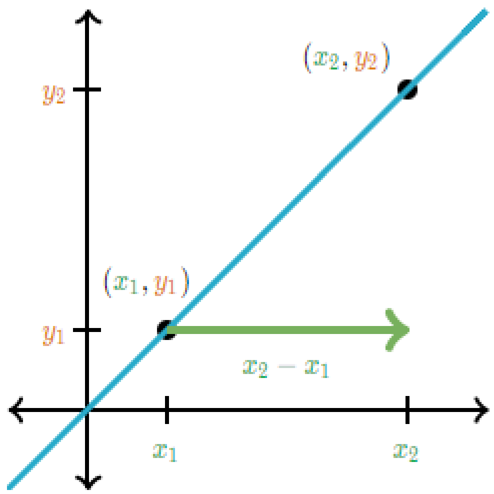
```
Y = np.dot(X,W) # (3,3) * (3,10) = (3,10)
print(Y)
```

```
[[ 15  27  25  11  23  25  15  27  25  19]
 [ 36  66  61  29  59  64  36  66  61  43]
 [ 57 105  97  47  95 103  57 105  97  67]]
```

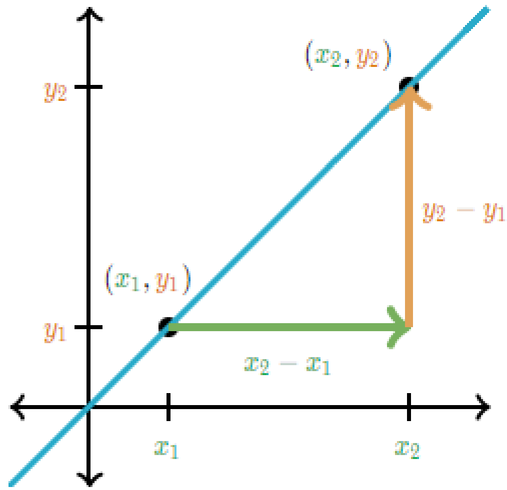
로그함수
행렬
미분
확률분포/정규분포

기울기(Slope, Gradient)

x 값의 증가량을 식으로 나타내면 $x_2 - x_1$



y 값의 증가량을 식으로 나타내면 $y_2 - y_1$

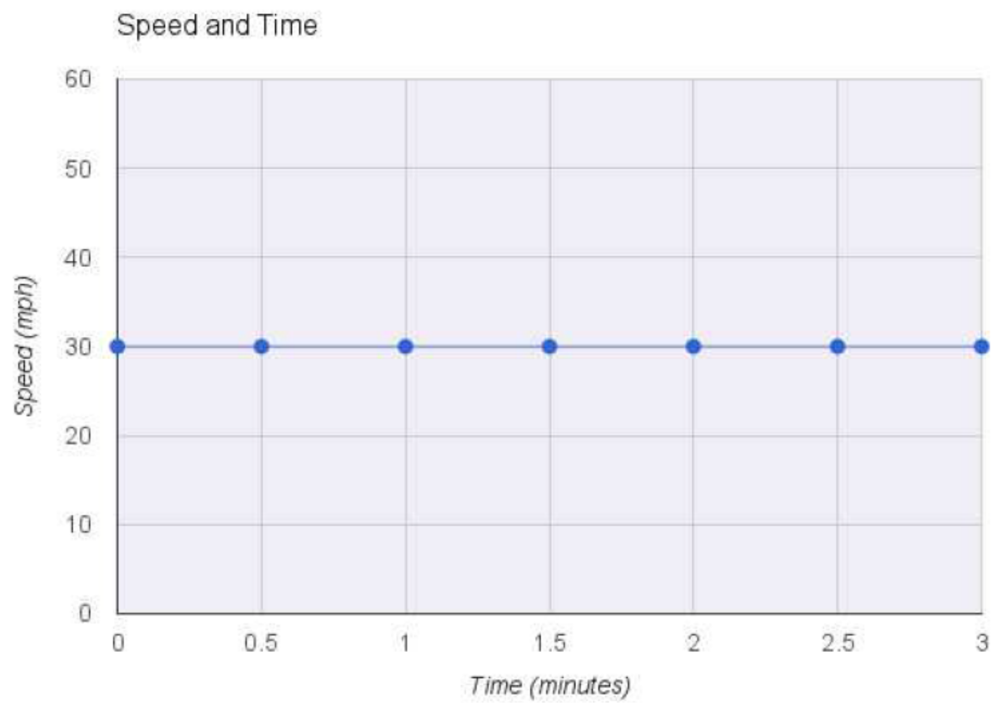


$$기울기 = \frac{\text{y 값의 증가량}}{\text{x 값의 증가량}} = \frac{y_2 - y_1}{x_2 - x_1}$$

https://ko.khanacademy.org/math

직선(A Flat Line)

Time (mins)	Speed (mph)
0.0	30
0.5	30
1.0	30
1.5	30
2.0	30
2.5	30
3.0	30



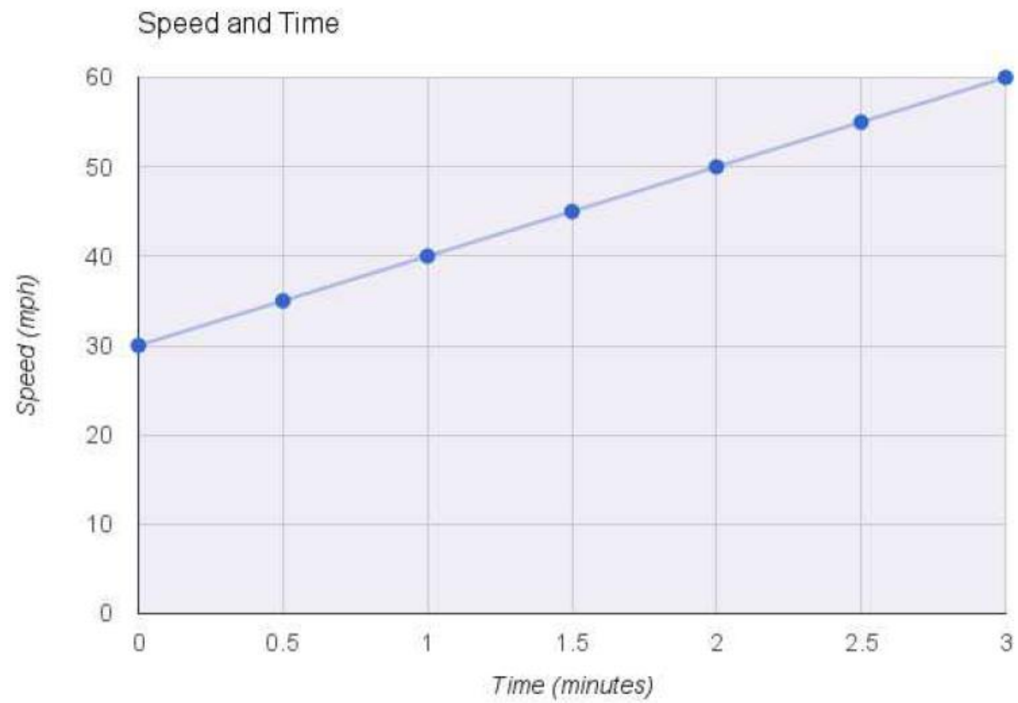
- speed : s = 30
- 변화율 : 0 (시간의 흐름에 따른 속도 변화)

$$\frac{\delta s}{\delta t} = 0$$

ref : Make Your Own Neural Network, Tariq Rashid

경사직선(A Sloped Straight Line)

Time (mins)	Speed (mph)
0.0	30
0.5	35
1.0	40
1.5	45
2.0	50
2.5	55
3.0	60



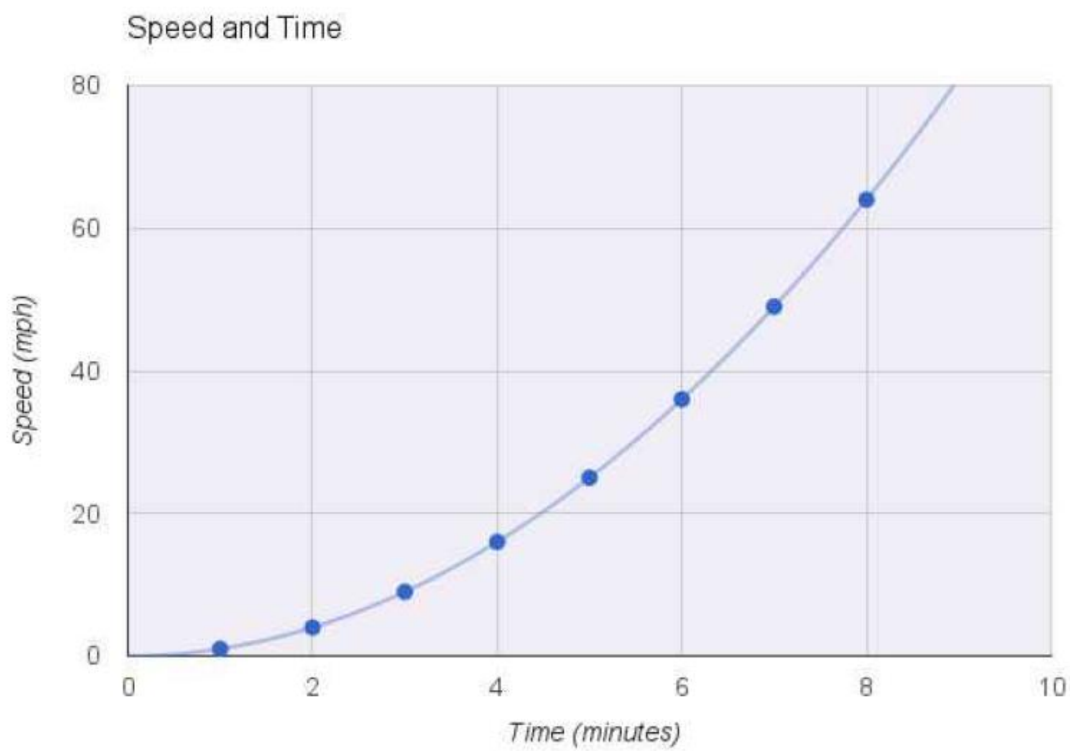
- 일정한 비율로 증가 (1분 경과할 때 마다 시속 10마일씩 가속)
- speed = 30 + (10 * time)
  - s = 30 + 10t
- 변화율 : 10

$$\frac{\delta s}{\delta t} = 10$$

ref : Make Your Own Neural Network, Tariq Rashid

곡선(A Curved Line)

Time (mins)	Speed (mph)
0	0
1	1
2	4
3	9
4	16
5	25
6	36
7	49
8	64

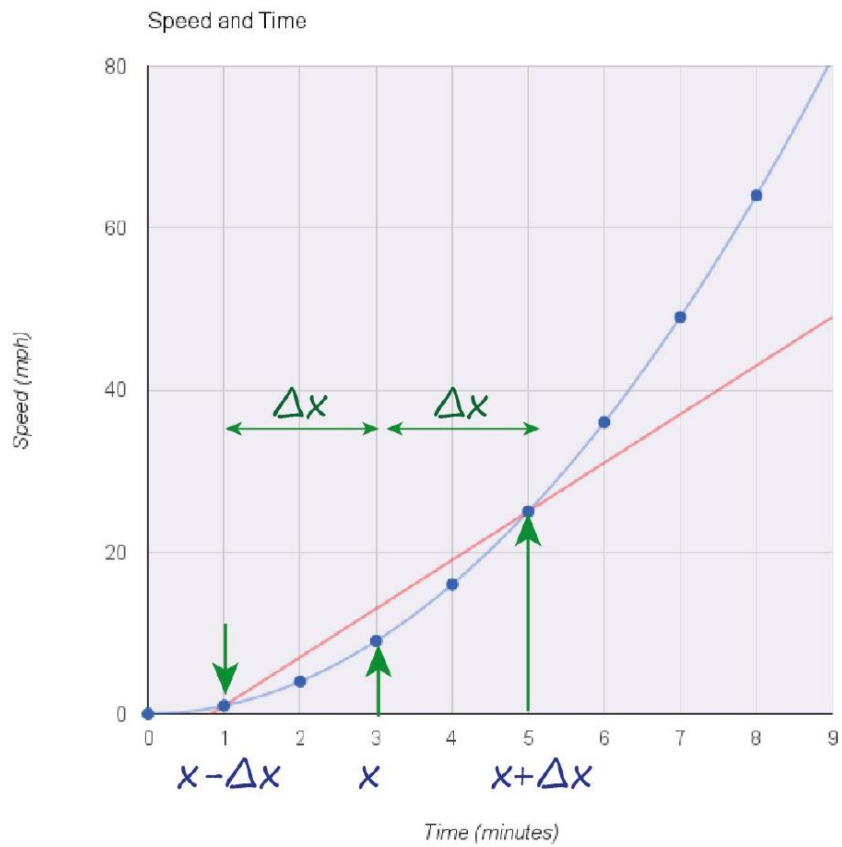


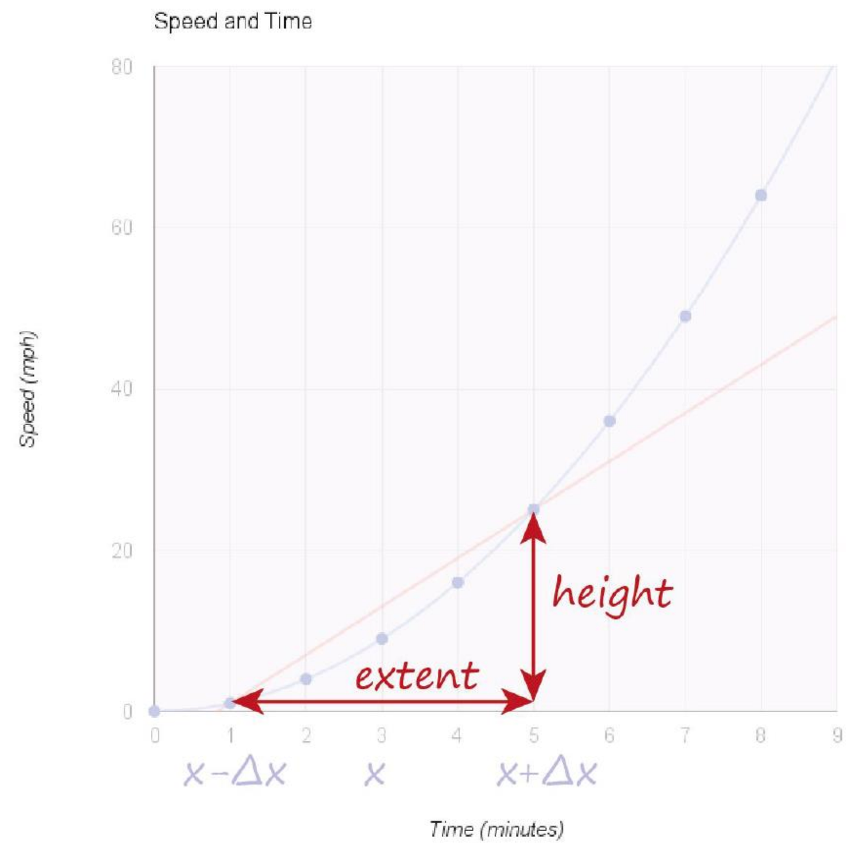
- 매분의 속도가 시간의 제곱
- $s = t^2$
- 시간의 어떤 순간에서의 속도의 변화율은?

ref : Make Your Own Neural Network, Tariq Rashid

3분(minute)을 중심으로 기울기 구하기

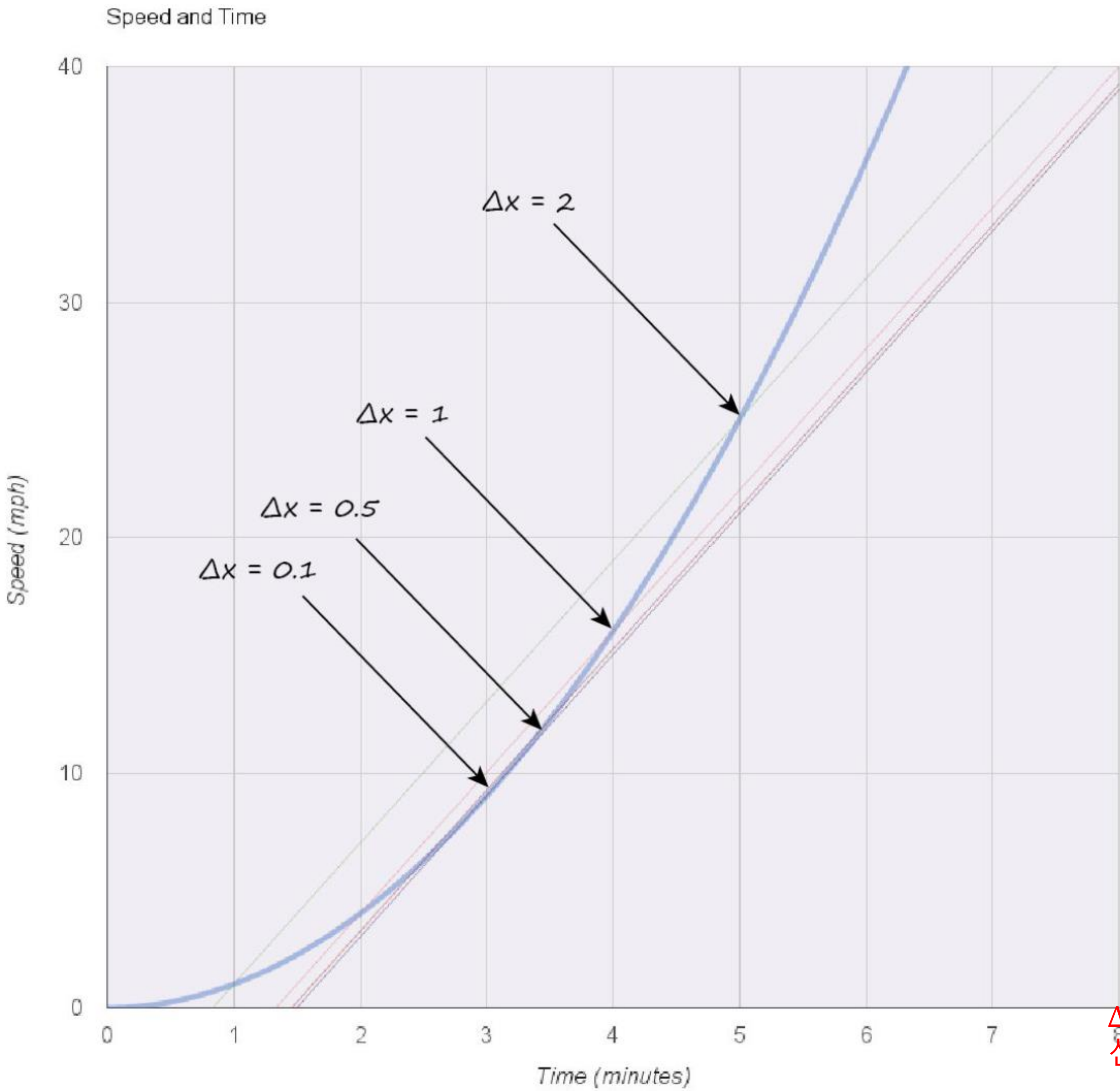
$$gradient = \frac{height}{extent} = \frac{24}{4} = 6$$





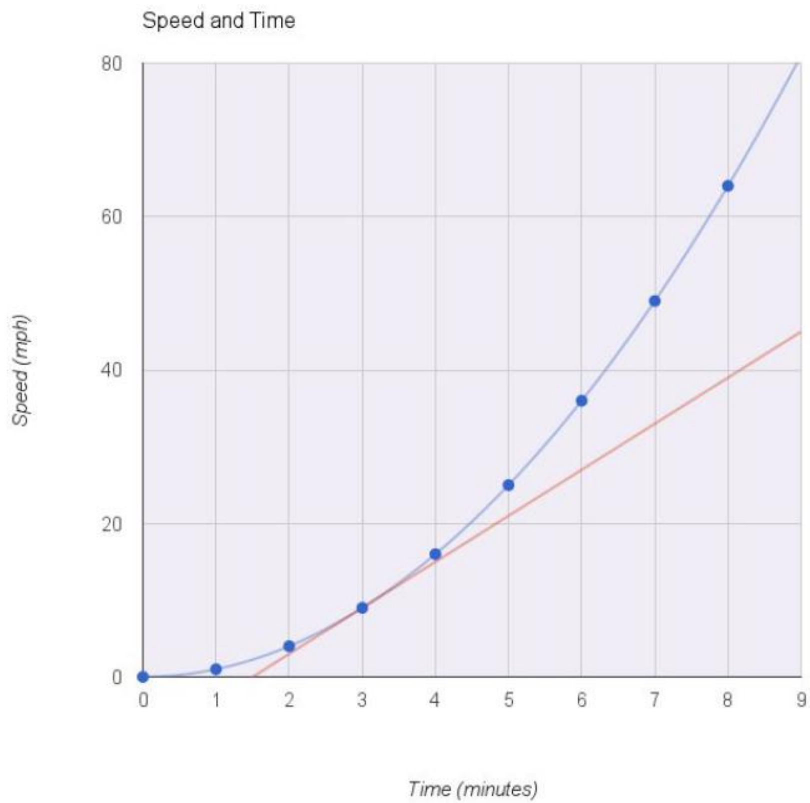
ref : Make Your Own Neural Network, Tariq Rashid

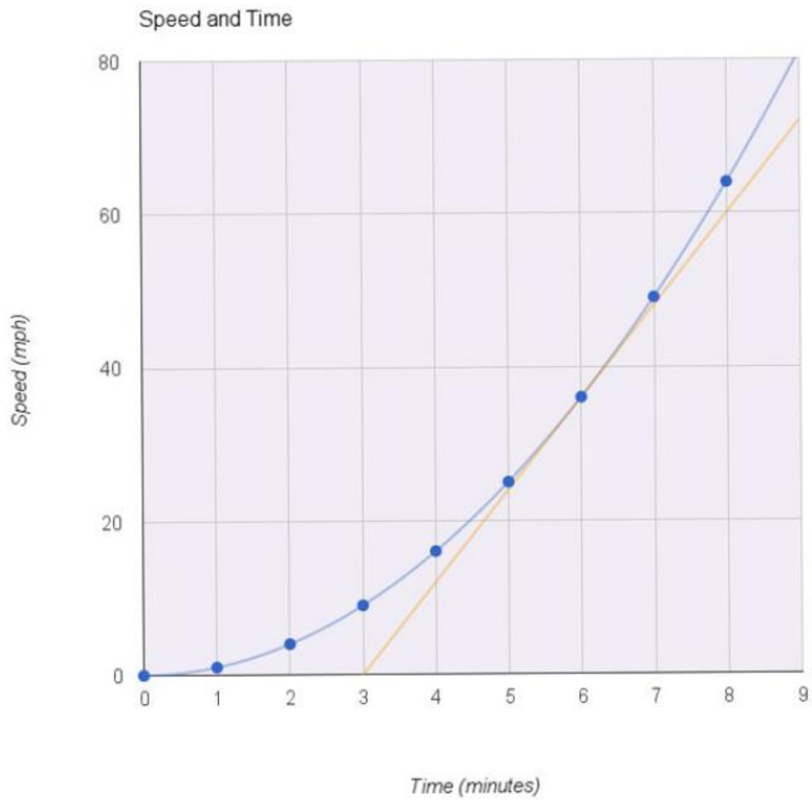
Δx가 작아지면 작아질수록



Δx가 작아지면 작아질수록 직선이 3분 지점으로 근접해 진다.

시간에 따른 기울기변화





- 3분(minute) 경과 시 속도 : 9 마일(mile)
- 6분(minute) 경과 시 속도 : 36 마일(mile)

- **3분에서의 기울기 < 6분에서의 기울기**

ref : Make Your Own Neural Network, Tariq Rashid

시간에 따른 기울기계산

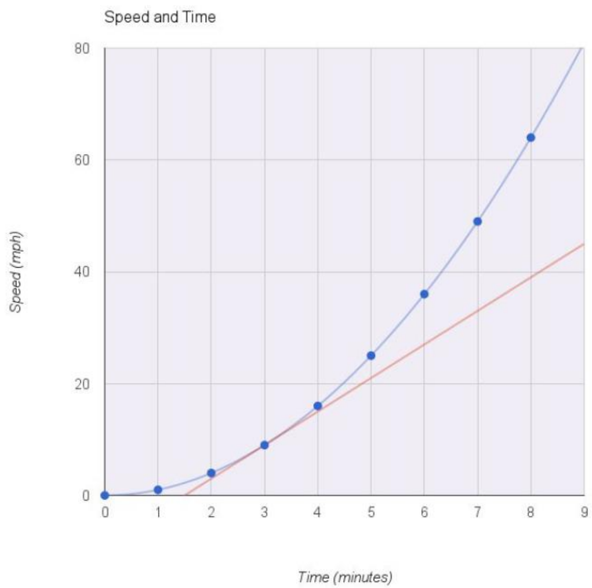
$$\frac{\delta s}{\delta t} = \frac{height}{extent} = \frac{(t+\Delta x)^2 - (t-\Delta x)^2}{2\Delta x}$$
$$= \frac{t^2 + \Delta x^2 + 2t\Delta x - t^2 - \Delta x^2 + 2t\Delta x}{2\Delta x}$$
$$= \frac{4t\Delta x}{2\Delta x}$$

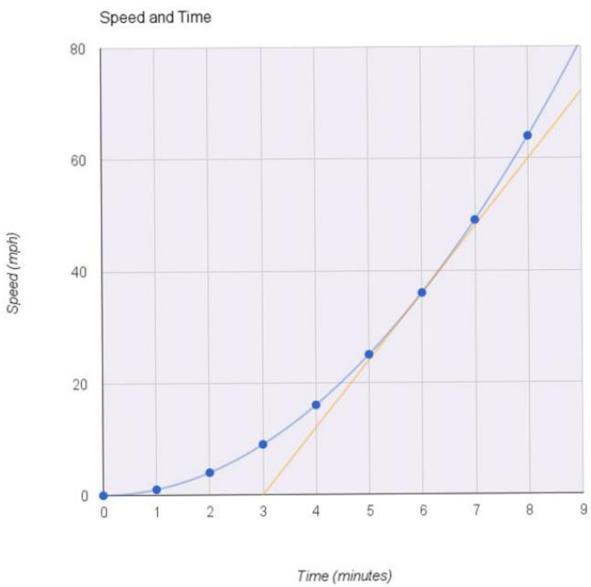
$$\boldsymbol{\frac{\delta s}{\delta t} = 2t}$$

$t = 3,\qquad 2t = 6$

$t = 6,\qquad 2t = 12$

$t = 100,\qquad 2t = 200$

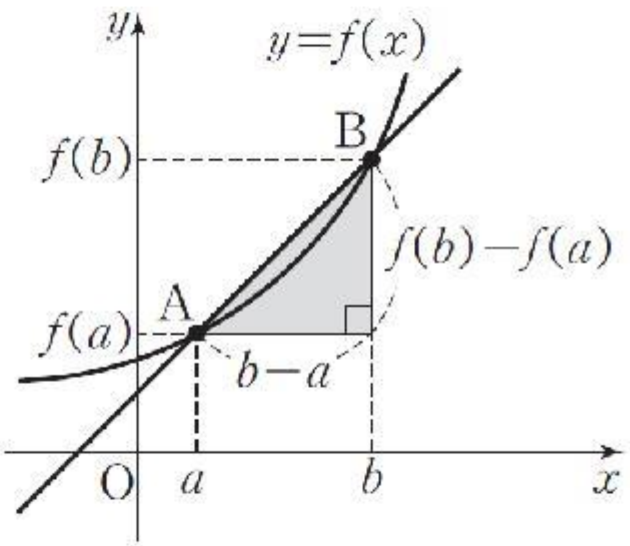




평균변화율/순간변화율

- 평균변화율(기울기)
  A, B 곡선의 디테일은 고려하지 않고 A,B사이의 변화율. 즉, 시작점과 끝점의 변화율

$$\frac{\Delta y}{\Delta x} = \frac{f(b) - f(a)}{b - a}$$



- 순간변화율(접선의 기울기)
  매 순간마다의 변화율. 곡선을 따라가면서 변화의 양상을 추정
  $(x = a)$ 에서의 순간변화율

$$\lim_{\Delta x \to 0} \frac{f(b) - f(a)}{b - a} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(a)$$

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = = \lim_{\Delta x \to 0} \frac{f(b) - f(a)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

$(\Delta x = b - a \Rightarrow b = a + \Delta x)$

($\Delta x = h$ 라하면)

$$y' = f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$ (도함수의 정의)

- 미분계수 = 순간변화율 = 접선의 기울기
- 미분한다 = 미분함수를 구한다 = 도함수를 구한다.
- 미분 표현하는 다양한 기호

$$y' = f'(x) = \frac{dy}{dx} = \frac{d}{dx}y = \frac{df(x)}{dx} = \frac{d}{dx}f(x) = \frac{\partial f}{\partial x} = \frac{\delta f}{\delta x}$$

미분공식

$$(f(x)+g(x))' = f'(x)+g'(x)$$

$$(f(x)-g(x))' = f'(x)-g'(x)$$

$$(f(x)g(x))' = f'(x)g(x)+f(x)g'(x)$$

$$(f(x)g(x)h(x))' = f(x)'g(x)h(x)+f(x)g(x)'h(x)+f(x)g(x)h'(x)$$

$$\left(\frac{g(x)}{f(x)}\right)' = \frac{f'(x)g(x)+f(x)g'(x)}{[f(x)]^2} \text{ (단, } f(x)\neq 0\text{)}$$

$$(x^n)' = nx^{n-1}$$

$$(x^{-n})' = -nx^{-n-1}, \qquad (x^{-1})' = -1x^{-1-1} = -x^{-2} = -\frac{1}{x^2}$$

합성함수(Chain Rule)

$y = f\big(g(x)\big),\quad g(x) = u$ 라 하면 $f(u)$

$$y' = f'(u)\cdot u' = f'\big(g(x)\big)\cdot g'(x) = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x}$$

(ex)

$y = (3x^3 + 2x + 1)^7$

$3x^3 + 2x + 1 = u$ 이라 하면 $y = f(u) \Rightarrow y' = f'(u)\cdot u'$

$f'(u) = 7(3x^3 + 2x + 1)^6$

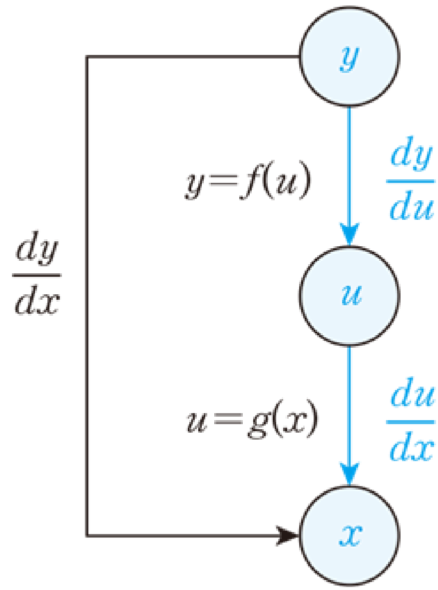
$u' = 9x^2 + 2$

$$y' = f'(u)\cdot u' = 7(3x^3 + 2x + 1)^6\cdot(9x^2+2) = \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x}$$

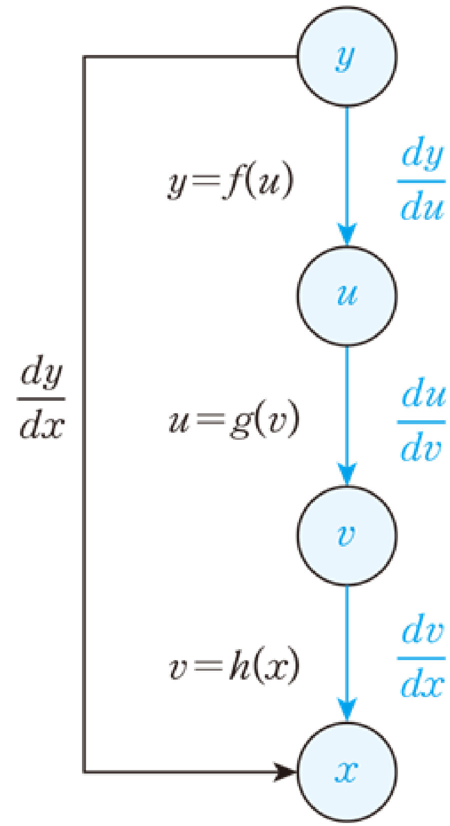
합성함수(Chain Rule) - 함수의 확장

$$y = f(u),\ \ u = g(v),\ \ v = h(x)$$

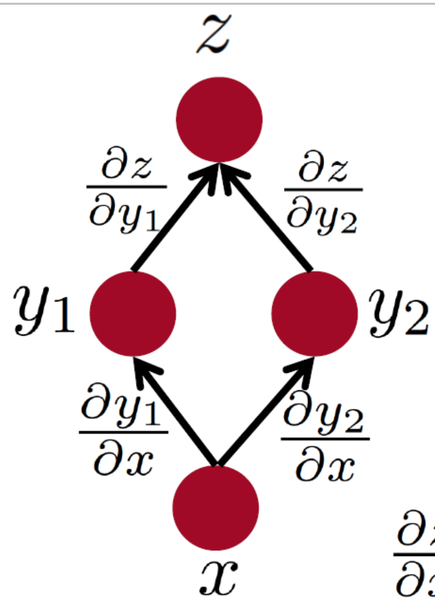
$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx}$$



(a) $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$



(b) $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dv}\frac{dv}{dx}$



$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial x} + \frac{\partial z}{\partial y_2}\frac{\partial y_2}{\partial x}$$

Example

$$f(x) = x\,, \qquad \frac{\partial f}{\partial x} = 1$$

$$f(x) = 2x, \qquad \frac{\partial f}{\partial x} = 2$$

$$f(x) = x^7, \qquad \frac{\partial f}{\partial x} = 7x^6$$

$$f(x) = 3x^3 + 2x + 1 = 9x^2 + 2$$

$$\boldsymbol{f(x, y) = xy}, \qquad \frac{\boldsymbol{\partial f}}{\boldsymbol{\partial x}} = \boldsymbol{y}$$

$$\boldsymbol{f(x, y) = xy}, \qquad \frac{\boldsymbol{\partial f}}{\boldsymbol{\partial y}} = \boldsymbol{x}$$

$$\boldsymbol{f(x, y) = x + y}, \qquad \frac{\boldsymbol{\partial f}}{\boldsymbol{\partial x}} = \mathbf{1}$$

$$\boldsymbol{f(x, y) = x + y}, \qquad \frac{\boldsymbol{\partial f}}{\boldsymbol{\partial y}} = \mathbf{1}$$

로그함수
행렬
미분
확률분포/정규분포

평균,분산,표준편차

- 평균

$$\mu = \frac{X_1 + X_2 + \cdots + X_N}{N} = \frac{\sum X_i}{N}$$

- 분산

$$\sigma^2 = \frac{\sum_i^N (X_i - \mu)^2}{N}$$

- 표준편차

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_i^N (X_i - \mu)^2}{N}}$$

확률분포

- 확률변수(random variable)
  - 표본공간에서 나타나는 수치 또는 수치로 표현된 함수
- 확률분포(Probability distribution)
  - 확률로서 자료집단의 모습을 표준화 한 것

$$\sum_i P(X_i) = 1$$

- (예)

완치기간 (일)	빈도수(명)
4	50
5	200
6	450
7	200
8	100
계	1000

확률변수(X)	4	5	6	7	8	계
확률분포$P(X)$	0.05	0.20	0.45	0.20	0.10	1.0

기대값$(Expectation)$, 분산, 표준편차

- 기대값(Expectation) : 확률변수의 평균(mean) 값
  - 확률값 p_i와 값x_i 을 취하는 확률변수 X 의 기대값

$$E(X) = \sum_i x_i \cdot P(x_i)$$

- 확률변수 X의 분산

$$\sigma^2(X) = E[X - E(X)]^2 = \sum_{i=1}^{n} [x_i - E(X)]^2 \cdot P(x_i)$$

- 확률변수 X의 표준편차

$$\sigma(X) = \sqrt{\sigma^2(X)} = \sqrt{E[X - E(X)]^2} = \sqrt{\sum_{i=1}^{n} [x_i - E(X)]^2 \cdot P(x_i)}$$

(예) 독서량에 대한 확률분포

확률변수(X)	2	4	6	8	10	계
확률분포$P(X)$	0.45	0.25	0.15	0.10	0.05	1.0

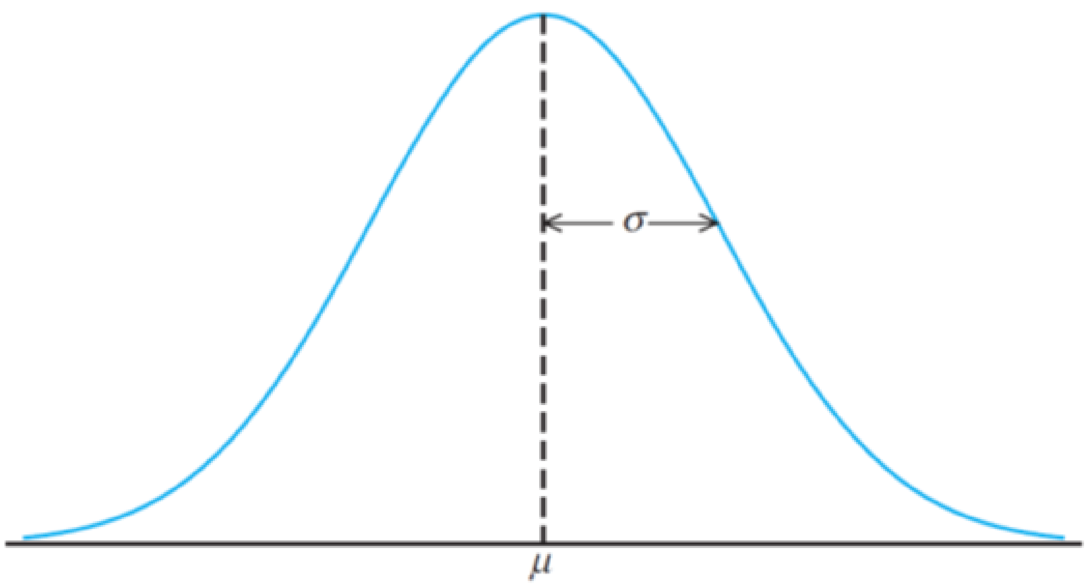
$$E(X) = \sum_i x_i \cdot P(x_i)$$
$$= x_1 \cdot P(x_1) + x_2 \cdot P(x_2) + x_3 \cdot P(x_3) + x_4 \cdot P(x_4) + x_5 \cdot P(x_5)$$
$$= 2 \cdot 0.45 + 4 \cdot 0.25 + 6 \cdot 0.15 + 8 \cdot 0.10 + 10 \cdot 0.05 = 4.1$$

$$\boldsymbol{\sigma^2(X)} = E[X - E(X)]^2 = \sum_{i=1}^{n} [x_i - E(X)]^2 \cdot P(x_i)$$
$$= (2-4.1)^2(0.45) + (4-4.1)^2(0.25) + (6-4.1)^2(0.15)$$
$$+(8-4.1)^2(0.10) + (10-4.1)^2(0.05)$$
$$= 1.9845 + 0.0025 + 0.5415 + 1.5210 + 1.7405 = 5.79$$

$$\boldsymbol{\sigma(X)} = \sqrt{\sigma^2(X)} = \sqrt{5.79} = 2.406$$

정규분포

- 평균이 μ이며 분산이 σ^2인 정규분포를 따는 확률변수 X에 대해 $X \sim N(\mu, \sigma^2)$ 로 표기
- 가우스분포(Gaussian distribution)라고도 한다.
- 정규분포의 특징
  - 모든 정규분포는 종 모양을 나타내고 좌우 대칭이다.
  - 정규곡선의 위치는 평균 μ에 의해서 정해지고, 모양은 표준편차 σ크기에 의해서 결정된다.
  - 자료집단에 따라 평균과 표준편차 크기는 다르다.



정규분포

- μ : 위치모수($location\ parameter$), σ : 형상모수($shape\ parameter$)
- 각각 평균이 μ_1, μ_2이고, 표준편차가 σ_1, σ_2인 두 정규분포의 비교
  - **$\mu_1 \neq \mu_2,\ \sigma_1 = \sigma_2$인 경우**
    두 정규분포곡선은 중심위치(평균값, 최빈값, 중앙값)는 다르지만, 동일한 종 모양이다
  - **$\mu_1 = \mu_2,\ \sigma_1 \neq \sigma_2$인 경우**
    두 정규분포곡선의 중심위치는 동일하지만, 자료의 흩어진 정도(분산, 표준편차)는 다르다
  - **$\mu_1 \neq \mu_2,\ \sigma_1 \neq \sigma_2$인 경우**
    두 정규분포곡선의 중심위치와 자료의 흩어진 정도가 완전히 다르다

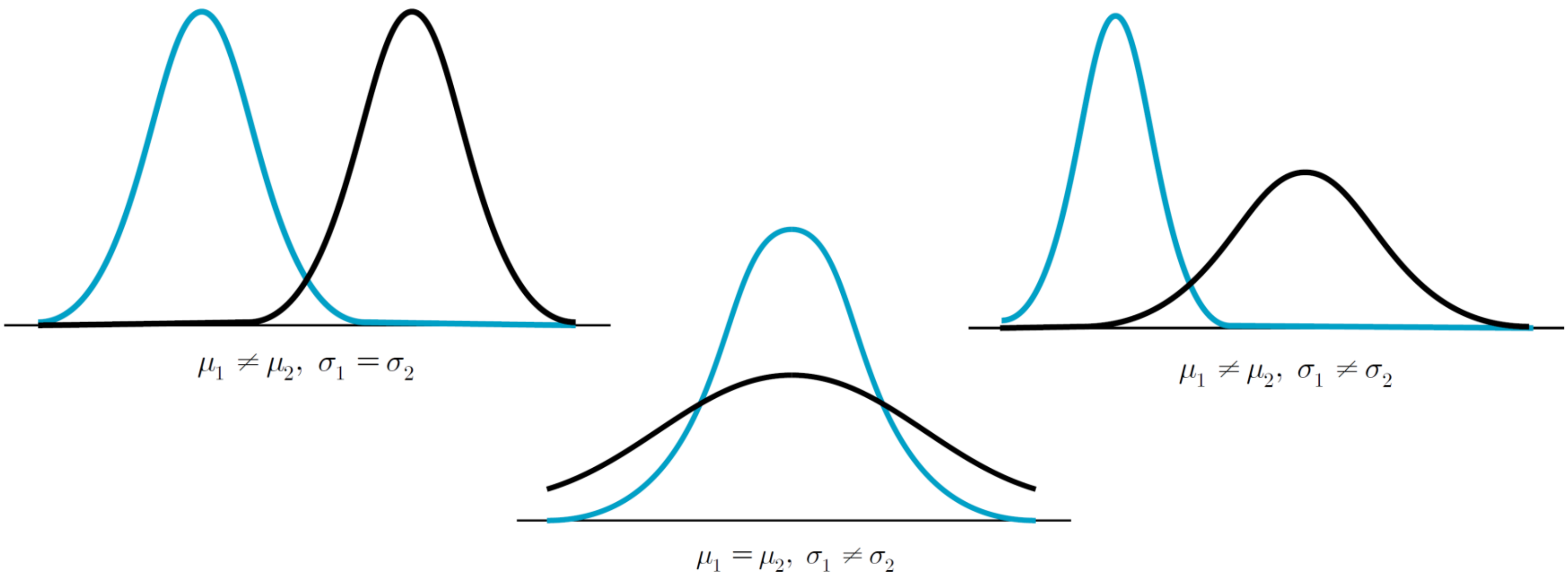


FIGURE 5.1

The effect of changing the mean of a normal distribution

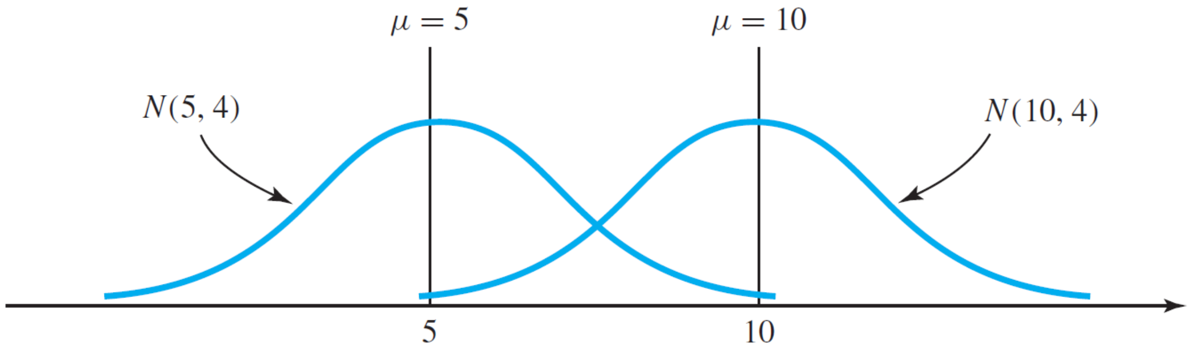
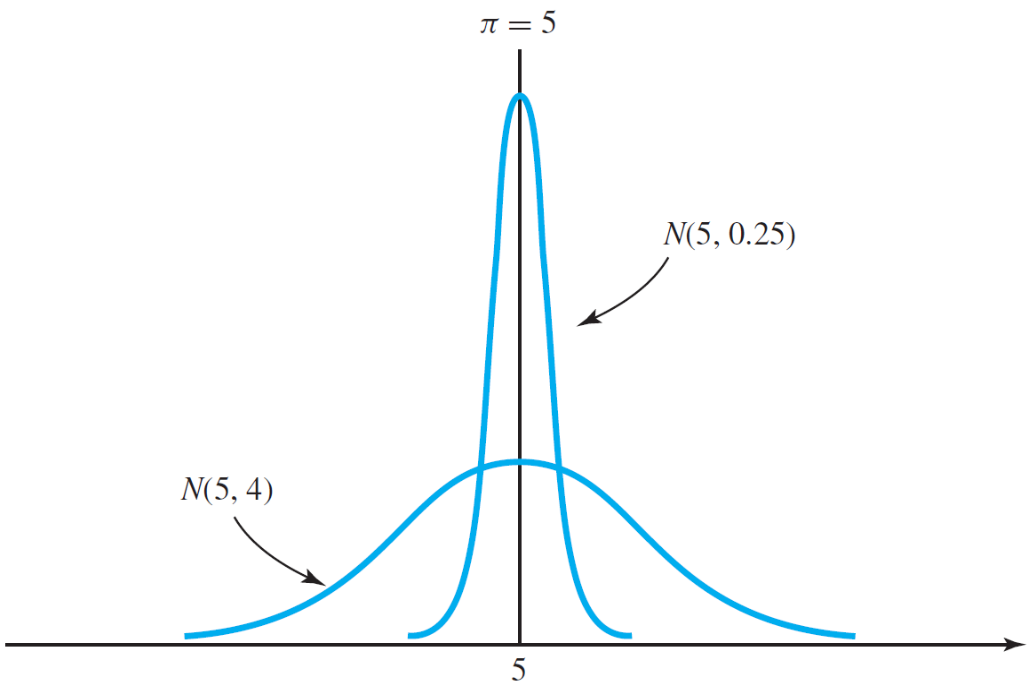


FIGURE 5.2

The effect of changing the variance of a normal distribution



표준정규분포(standard normal distribution)

- 확률변수 X가 정규분포 $N(\mu, \sigma^2)$을 따를 때, 표준화된 확률변수 Z의 확률밀도 함수

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad (-\infty < x < \infty), \qquad Z = \frac{X - \mu}{\sigma}$$

이때 확률변수 Z가 표준정규분포를 따른다고 하고 $\boldsymbol{N(0,1)}$로 표기한다

- 확률변수 Z의 평균과 분산

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\{E(X) - \mu\} = \mathbf{0}$$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(X) = \mathbf{1}$$

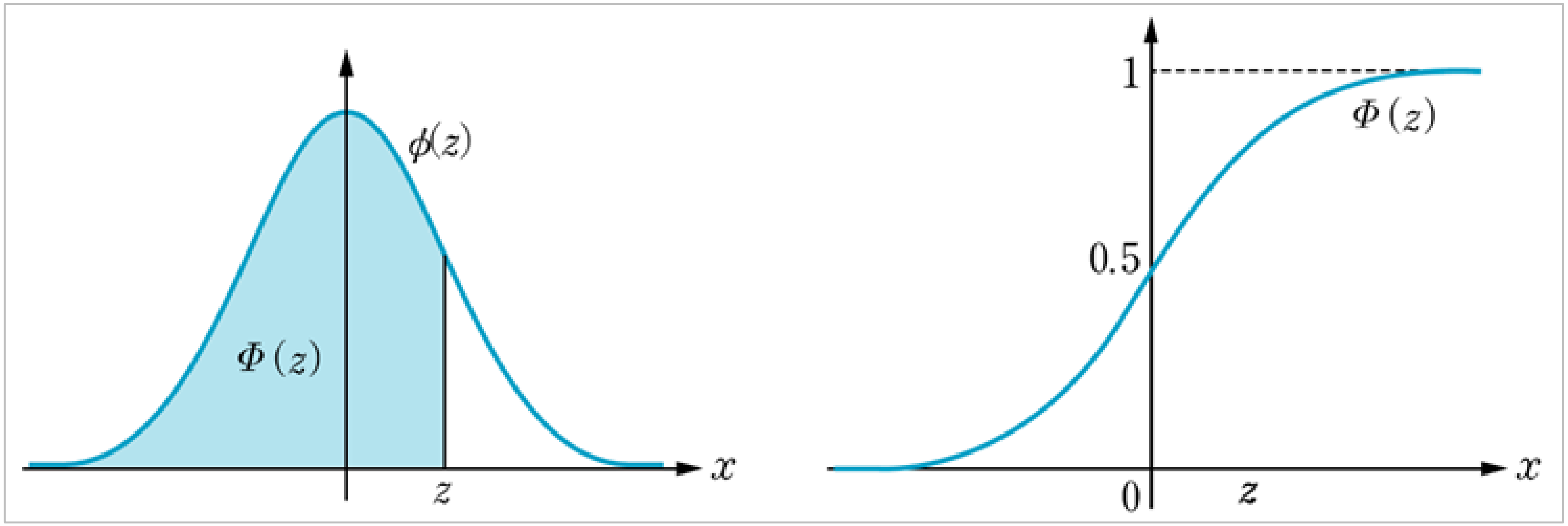
표준정규분포(standard normal distribution)

- 확률변수 Z가 표준정규분포 $N(0, 1)$을 따를 때, 그 분포함수 $\phi(z)$로 나타내며, 다음과 같이 정의된다

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$\Phi(z)$를 표준정규분포함수라 한다

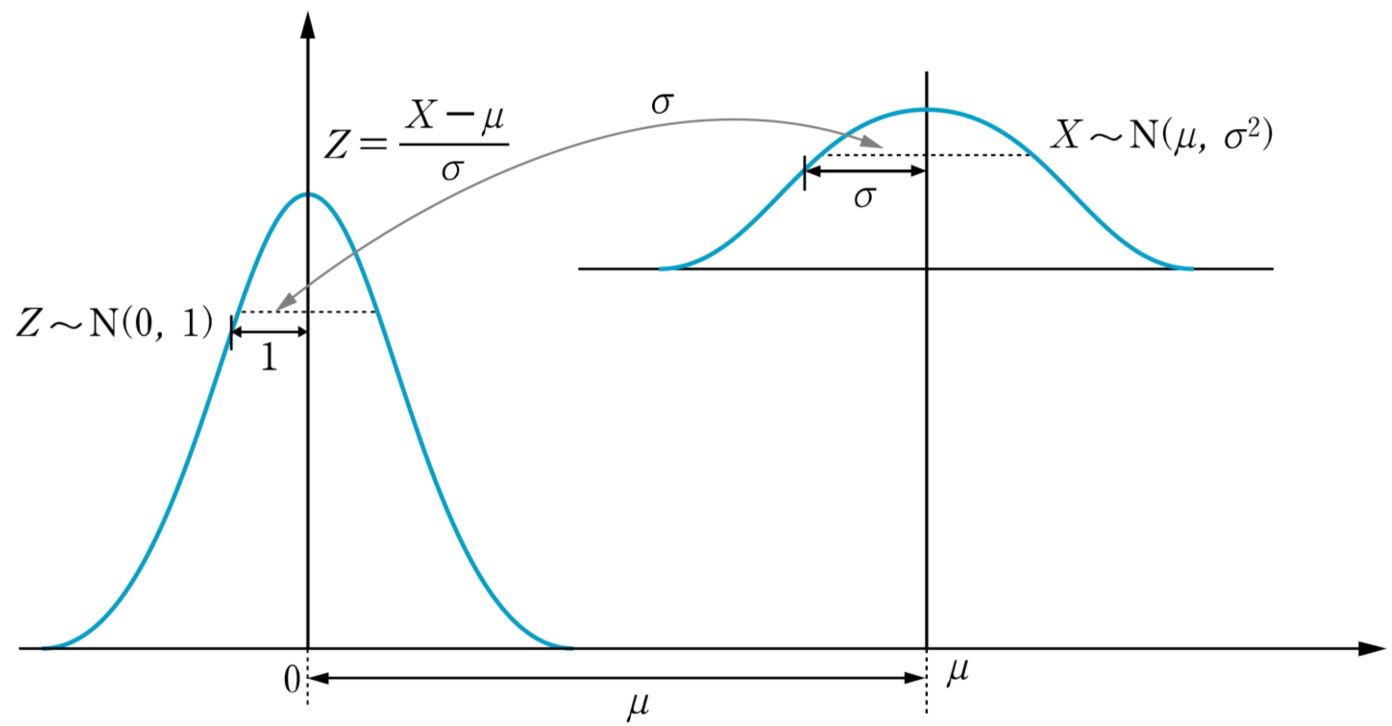
- $\Phi(z)$가 $z = 0$에 대해 좌우 대칭이므로 $\Phi(z) = 0.5$이다
- $\Phi(z)$는 표준정규분포의 누적분포함수를 의미함



표준정규분포

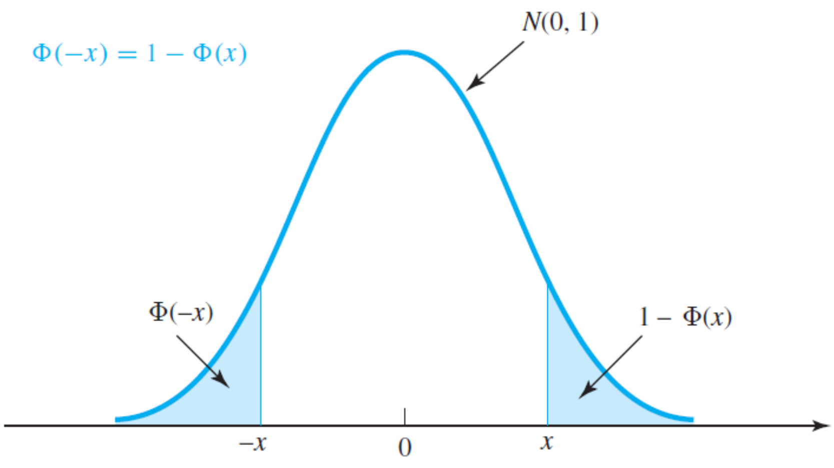
- 확률변수 X가 정규분포 $N(\mu, \sigma^2)$을 따를 때, 확률 $P(a \le X \le b)$ 는 다음과 같이 표준정규분포 $N(0, 1)$을 이용하면 편리하게 구할 수 있다

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = P(z_a \le Z \le z_b)$$

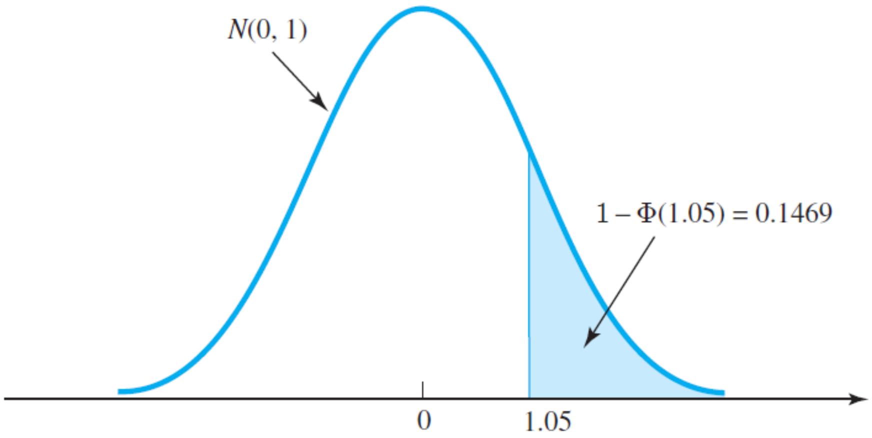


[그림 4-13] 정규분포와 표준정규분포의 관계

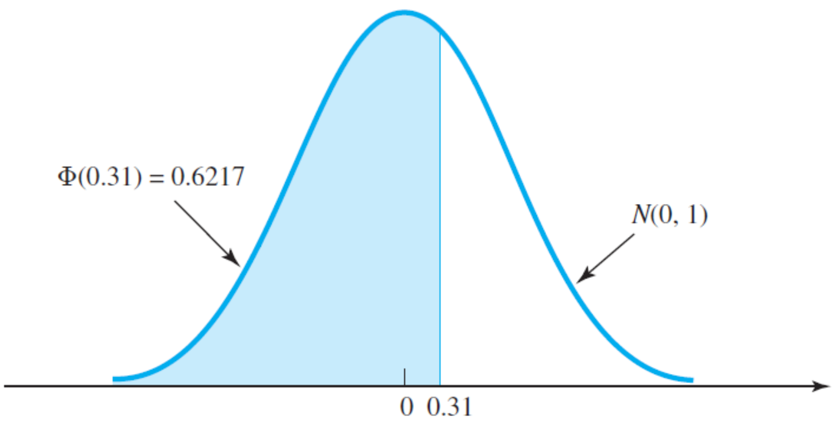
표준정규분포



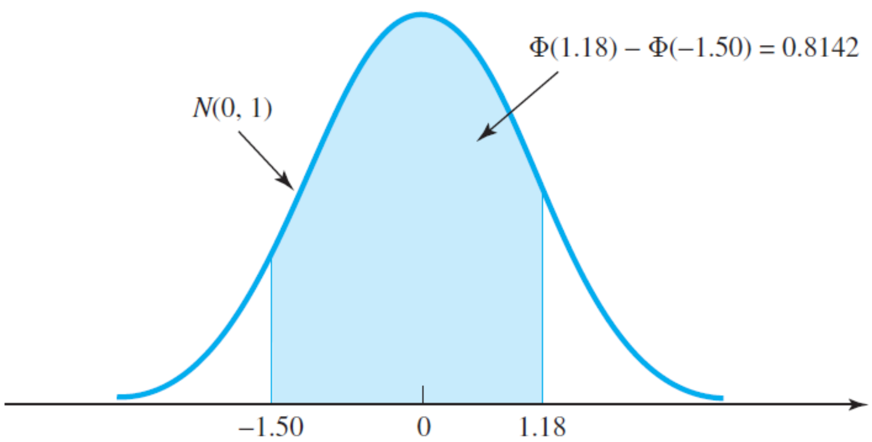
Φ(−x) = 1 − Φ(x)
N(0, 1)
Φ(−x)
1 − Φ(x)
−x
0
x



N(0, 1)
1 − Φ(1.05) = 0.1469
0
1.05



Φ(0.31) = 0.6217
N(0, 1)
0
0.31



Φ(1.18) − Φ(−1.50) = 0.8142
N(0, 1)
−1.50
0
1.18