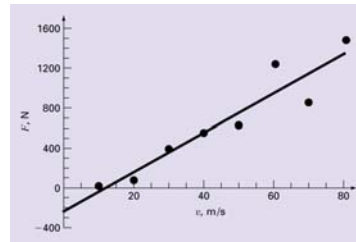


Chap. 14 Curve Fitting: 선형 회귀분석

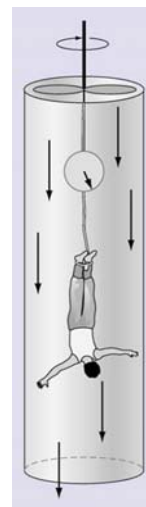
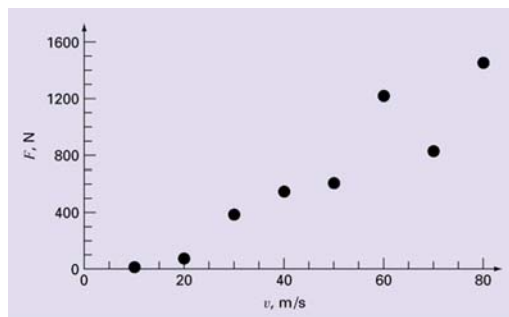
- ❖ 곡선 맞춤 개요
- ❖ 통계학 기초
- ❖ 선형 회귀분석
- ❖ 비선형 방정식의 선형화



곡선맞춤의 필요성

□ 사례: 실험에 의한 항력계수 결정

- 번지점프하는 사람에 작용하는 항력: $F = cv^2$
- 풍동실험에 의한 항력계수(c) 결정



곡선맞춤 (Curve Fitting)

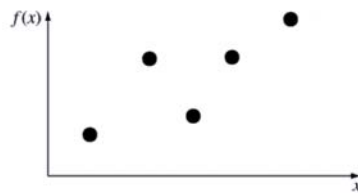
□ 곡선맞춤

- 불연속적인 데이터 사이에 있는 점에서의 값 추정시 필요
- 주어진 데이터를 가장 적절히 표현할 수 있는 함수식 계산
- 곡선맞춤 종류: 최소제곱 회귀분석 vs. 보간법

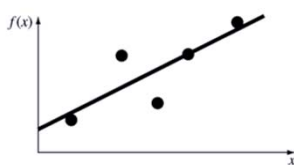
□ 최소제곱 회귀분석 (Least-Square Regression)

□ 보간법 (Interpolation)

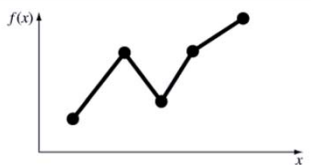
곡선맞춤 (Curve Fitting)



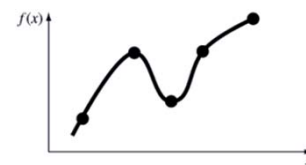
최소제곱법 (선형)



선형보간법



곡선보간법



For n data points y_1, y_2, \dots, y_n

P. 351 예제 14.1

평균 (mean)

표준편차 (standard deviation)

분산 (variance)

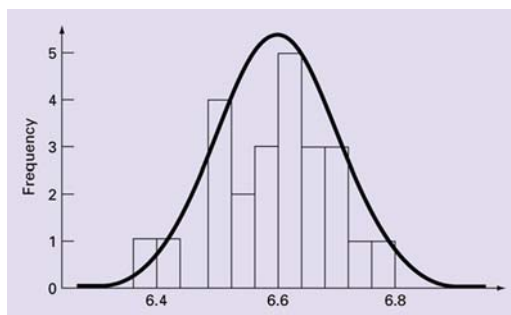
분산계수 (coefficient of variation)

: 평균값에 대한 표준편차의 비

정규분포 (Normal Distribution)

□ 히스토그램 (histogram)

- 데이터가 평균값 주변에 어떻게 분포되어있는지를 시각적으로 표현
- 데이터가 충분할 경우 히스토그램은 완만한 종 형태로 근사화됨



68% of total measurements in

95% of total measurements in

MATLAB 내장함수

mean(x), median(x), hist(x)

Linear Least Squares Regression

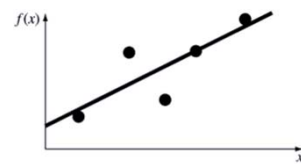
$f(x)$ is in a linear form : $f(x) = ax + b$

the error :

Is minimized when :

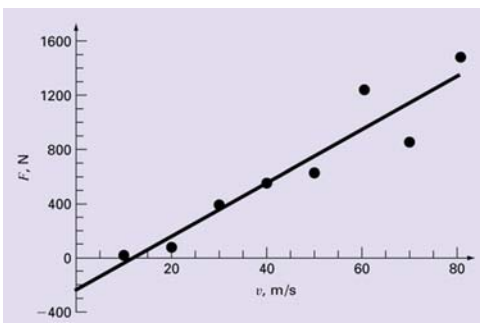
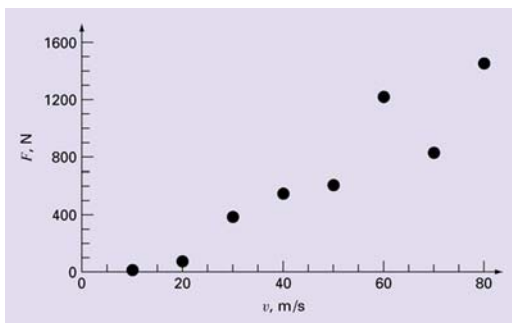
→

→

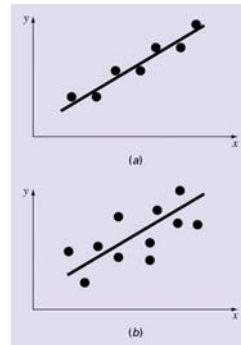
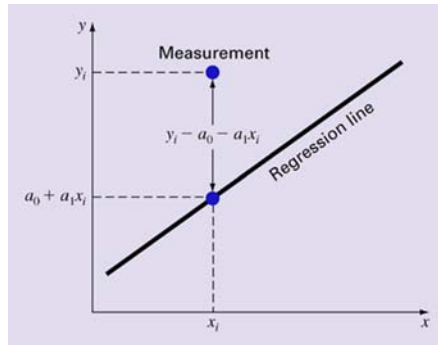


선형 회귀분석 사례

P. 364 예제 14.4

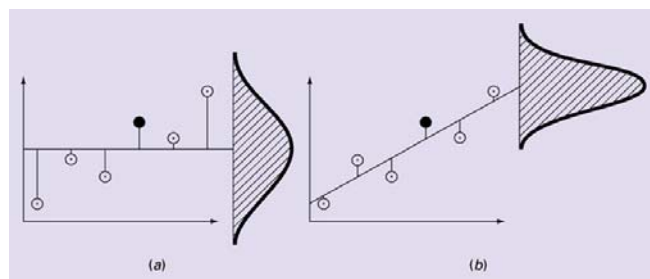


잔차 (Residual)



데이터 평균값 기준 표준편차

선형 회귀곡선 기준 표준편차

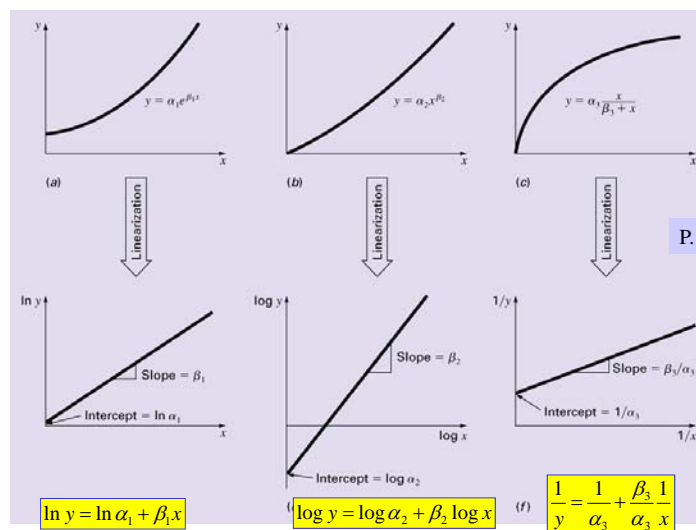


□ 결정계수 (Coefficient of determination)

□ 상관계수 (Correlation coefficient)

P. 367 예제 14.5

- $r = 1$:
- $r = 0$:



P. 371 예제 14.6

```
function [a, r2] = linregr(x,y)
% [a, r2] = linregr(x,y):
% Least squares fit of a straight line to data
% by solving the normal equations.
% input:
% x = independent variable
% y = dependent variable
% output:
% a = vector of slope, a(1), and intercept, a(2)
% r2 = coefficient of determination

n = length(x);
if length(y)~=n, error('x and y must be same length'); end
x = x(:); y = y(:); % convert to column vectors
sx = sum(x); sy = sum(y);
sx2 = sum(x.*x); sxy = sum(x.*y); sy2 = sum(y.*y);
a(1) = (n*sxy-sx*sy)/(n*sx2-sx^2);
a(2) = sy/n-a(1)*sx/n;
r2 = ((n*sxy-sx*sy)/sqrt(n*sx2-sx^2)/sqrt(n*sy2-sy^2))^2;
```

```
% create plot of data and best fit line
xp = linspace(min(x),max(x),2);
yp = a(1)*xp+a(2);
plot(x,y,'o',xp,yp)
grid on
```

Michaelis-Menten 모델 : 기질의 농도에 따른 효소의 반응속도 관계식

v : 초기반응속도, v_m : 최대초기반응속도
 $[S]$: 기질의 농도, k_s : 반포화 상수

상호작용을 고려한 2차 모델

Matlab Command

```
S = [1.3 1.8 3 4.5 6 8 9];
v = [0.07 0.13 0.22 0.275 0.335 0.35 0.36];
```