Ankara Üniversitesi-BLM bölümü

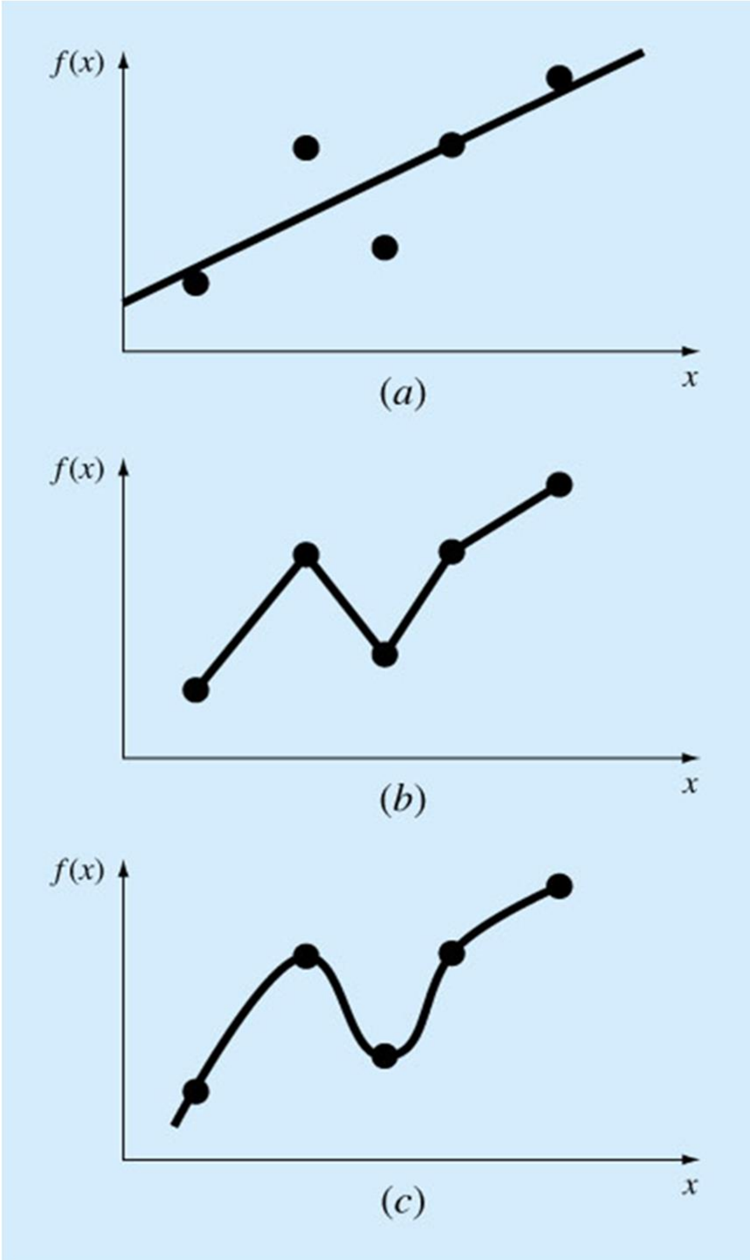
BLM433 Sayısal Analiz Teknikleri

CURVE FITTING

Describes techniques to fit curves (curve fitting) to discrete data to obtain intermediate estimates.

There are two general approaches two curve fitting:
Data exhibit a significant degree of scatter. The strategy is to derive a single curve that represents the general trend of the data.
Data is very precise. The strategy is to pass a curve or a series of curves through each of the points.

In engineering two types of applications are encountered:
Trend analysis. Predicting values of dependent variable, may include extrapolation beyond data points or interpolation between data points.
Hypothesis testing. Comparing existing mathematical model with measured data.


f(x)
x
(a)
f(x)
x
(b)
f(x)
x
(c)

In course of engineering study,

if several measurements are made of a particular quantity, additional insight can be gained by summarizing the data in

one or more well chosen statistics that convey as much information as possible about specific characteristics of the data set.

These descriptive statistics are most often selected to represent

The location of the center of the distribution of the data,
The degree of spread of the data.

Arithmetic mean. The sum of the individual data points (yi) divided by the number of points (n).

$$\bar{y} = \frac{\sum y_i}{n}$$

$$i = 1, \ldots, n$$

Standard deviation. The most common measure of a spread for a sample.

$$S_y = \sqrt{\frac{S_t}{n-1}}$$

$$S_t = \sum (y_i - \bar{y})^2$$

or

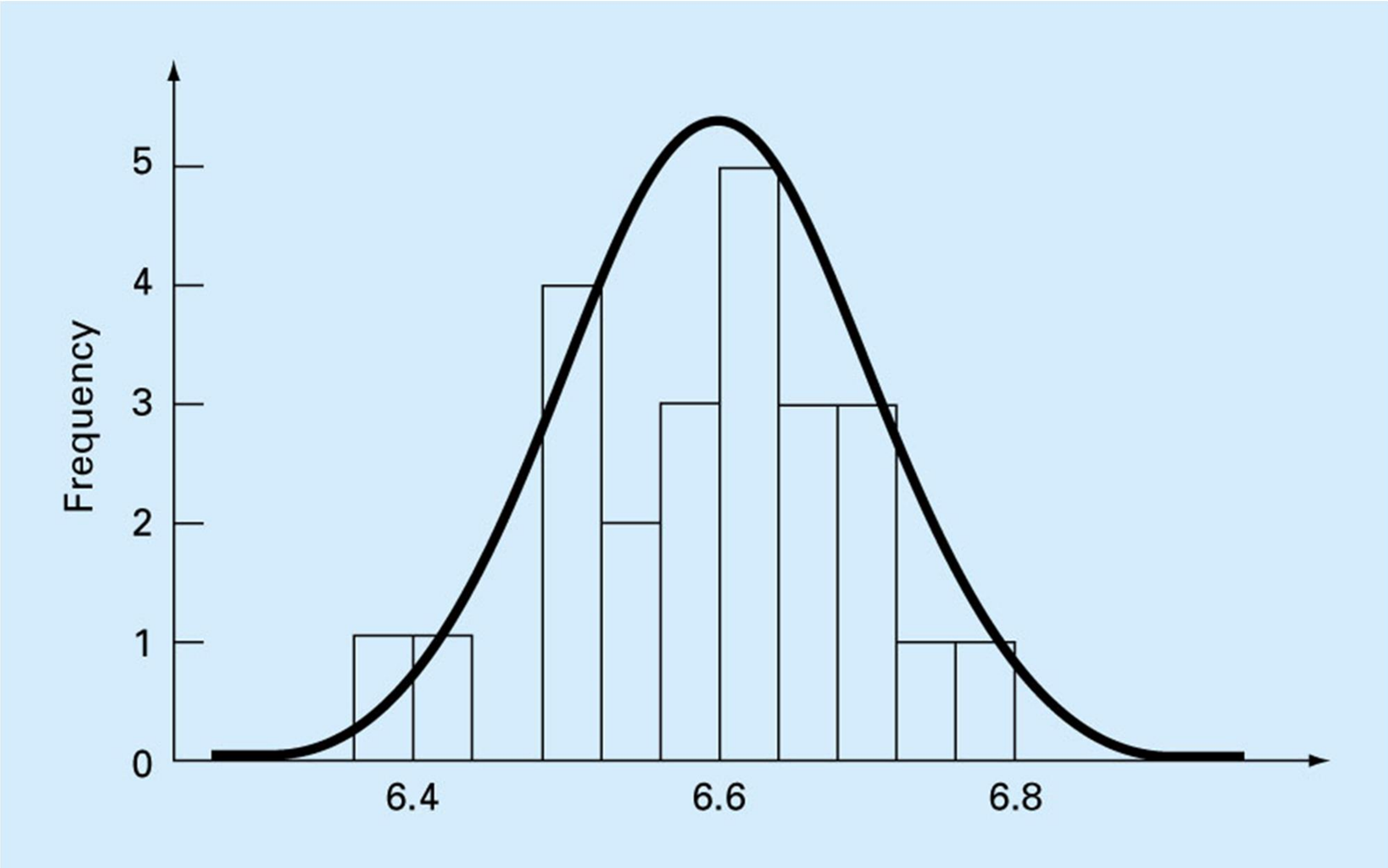
$$S_y^2 = \frac{\sum y_i^2 - \left(\sum y_i\right)^2 / n}{n-1}$$

Variance

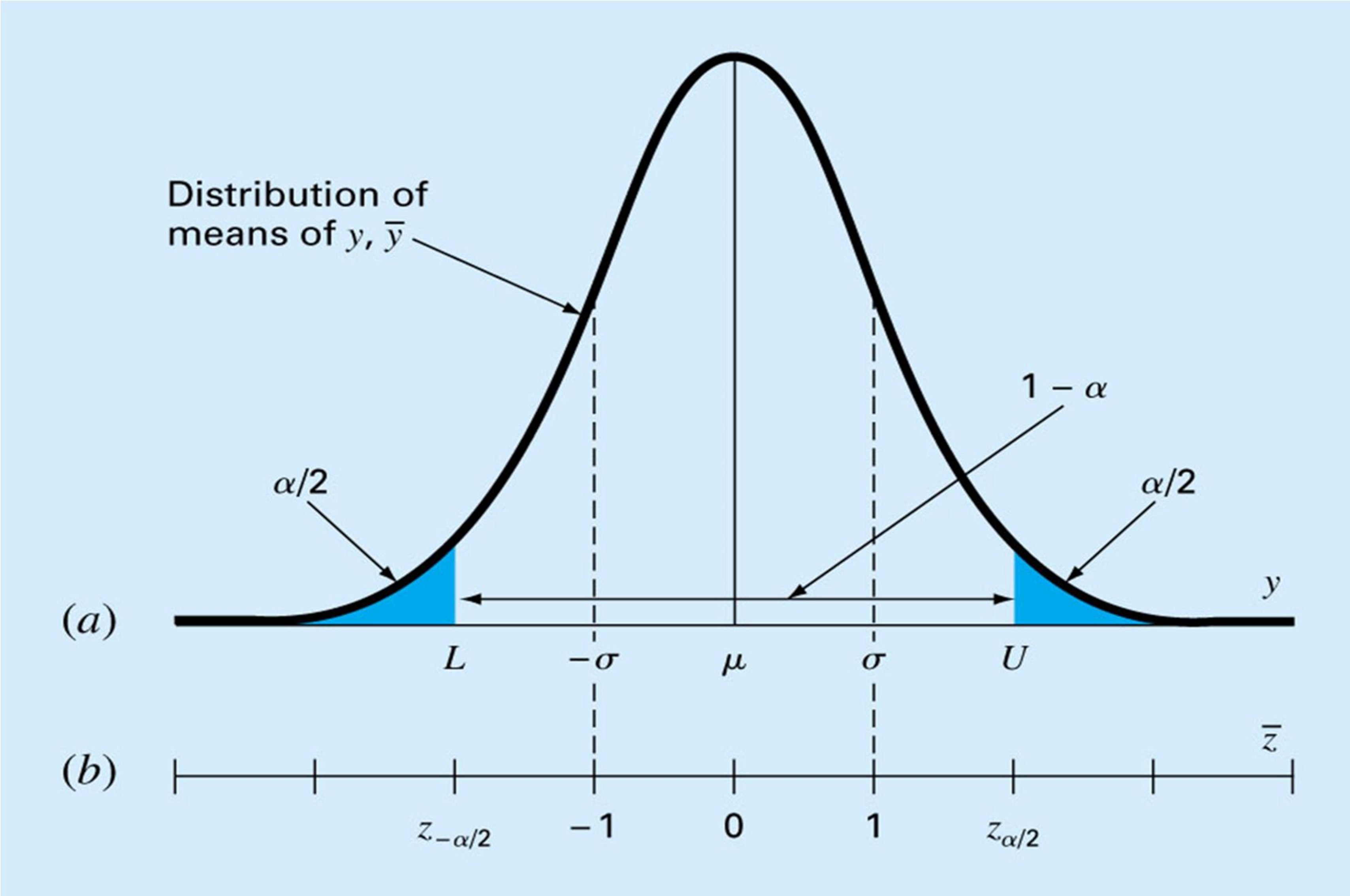
$$S_y^2 = \frac{\sum (y_i - \bar{y})^2}{n-1}$$

Coefficient of variance

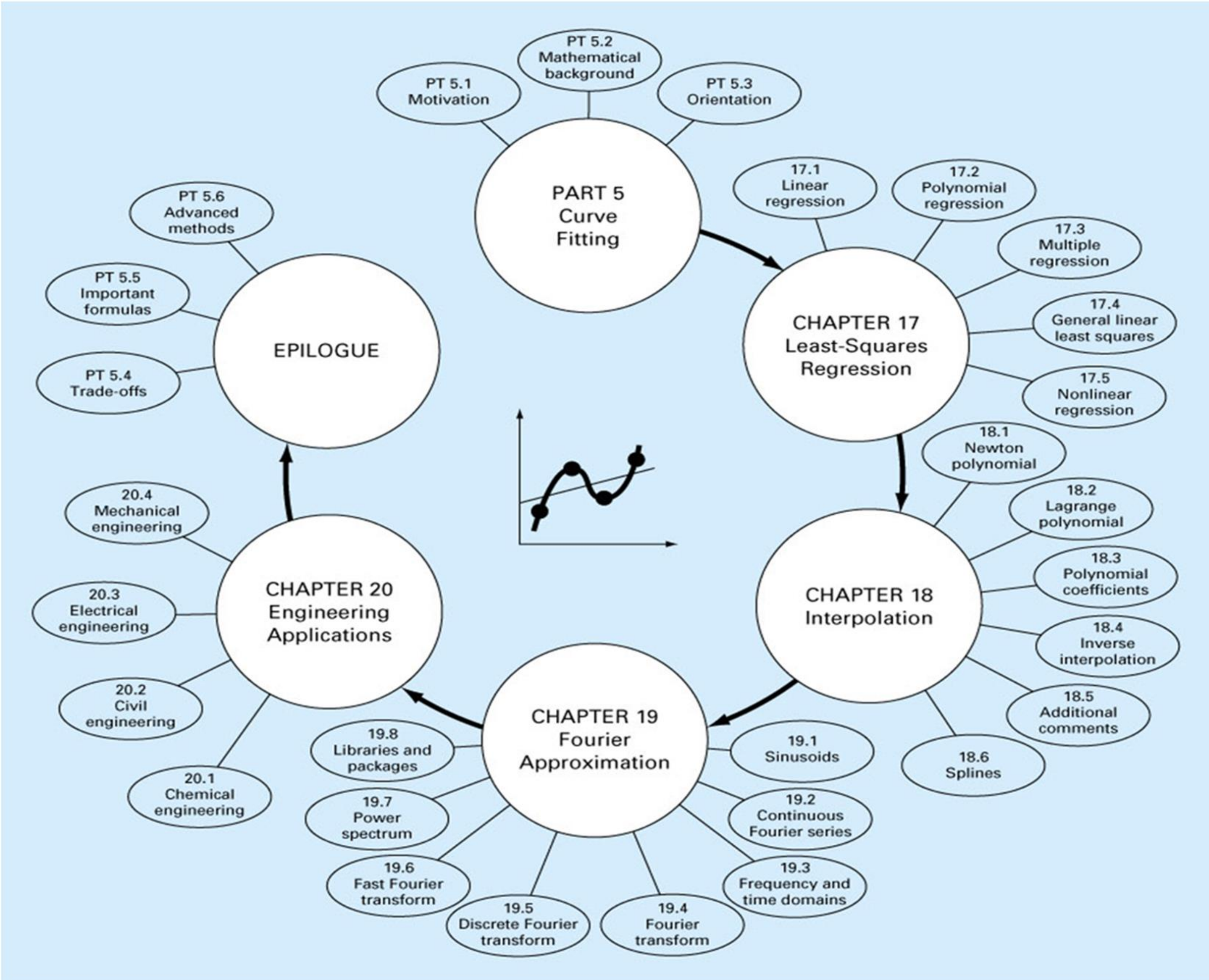
$$c.v. = \frac{S_y}{\bar{y}} 100\%$$



Frequency
0
1
2
3
4
5
6.4
6.6
6.8



Distribution of
means of y, ȳ
1 − α
α/2
α/2
y
(a)
L
−σ
μ
σ
U
z̄
(b)
z−α/2
−1
0
1
zα/2


PT 5.1
Motivation
PT 5.2
Mathematical background
PT 5.3
Orientation
PART 5
Curve Fitting
17.1
Linear regression
17.2
Polynomial regression
17.3
Multiple regression
17.4
General linear least squares
17.5
Nonlinear regression
CHAPTER 17
Least-Squares Regression
18.1
Newton polynomial
18.2
Lagrange polynomial
18.3
Polynomial coefficients
18.4
Inverse interpolation
18.5
Additional comments
18.6
Splines
CHAPTER 18
Interpolation
CHAPTER 19
Fourier Approximation
19.1
Sinusoids
19.2
Continuous Fourier series
19.3
Frequency and time domains
19.4
Fourier transform
19.5
Discrete Fourier transform
19.6
Fast Fourier transform
19.7
Power spectrum
19.8
Libraries and packages
CHAPTER 20
Engineering Applications
20.1
Chemical engineering
20.2
Civil engineering
20.3
Electrical engineering
20.4
Mechanical engineering
EPILOGUE
PT 5.4
Trade-offs
PT 5.5
Important formulas
PT 5.6
Advanced methods

Least square regression

Linear Regression

Fitting a straight line to a set of paired observations: $(x_1, y_1), (x_2, y_2),\ldots,(x_n, y_n)$.

$y=a_0+a_1x+e$

a_1- slope

a_0- intercept

e- error, or residual, between the model and the observations

Minimize the sum of the residual errors for all available data:

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (y_i - a_o - a_1 x_i)$$

n = total number of points

However, this is an inadequate criterion, so is the sum of the absolute values

$$\sum_{i=1}^{n} |e_i| = \sum_{i=1}^{n} |y_i - a_0 - a_1 x_i|$$

$$\frac{\partial S_r}{\partial a_o} = -2\sum (y_i - a_o - a_1 x_i) = 0$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum \left[(y_i - a_o - a_1 x_i) x_i\right] = 0$$

$$0 = \sum y_i - \sum a_0 - \sum a_1 x_i$$

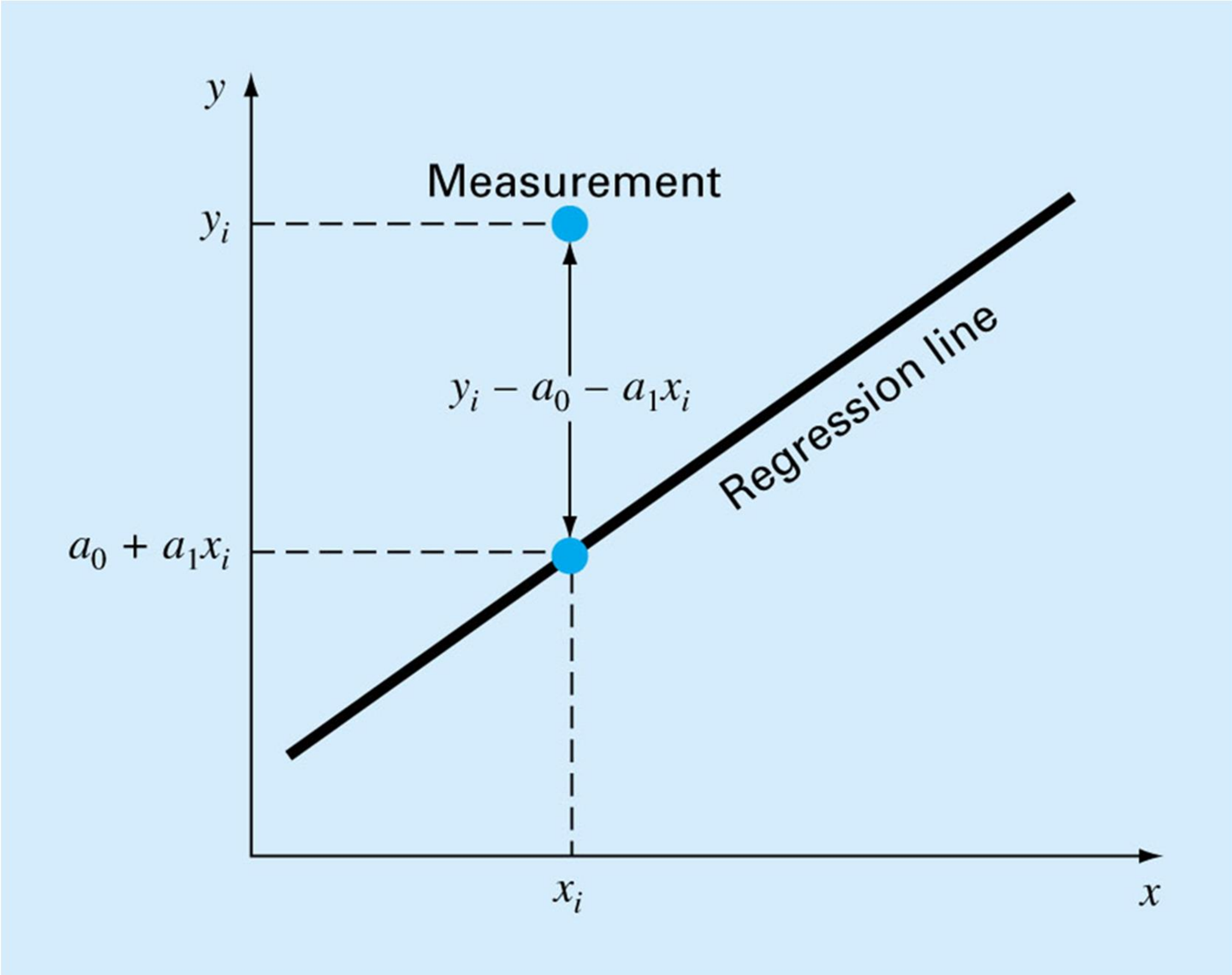
$$0 = \sum y_i x_i - \sum a_0 x_i - \sum a_1 x_i^2$$

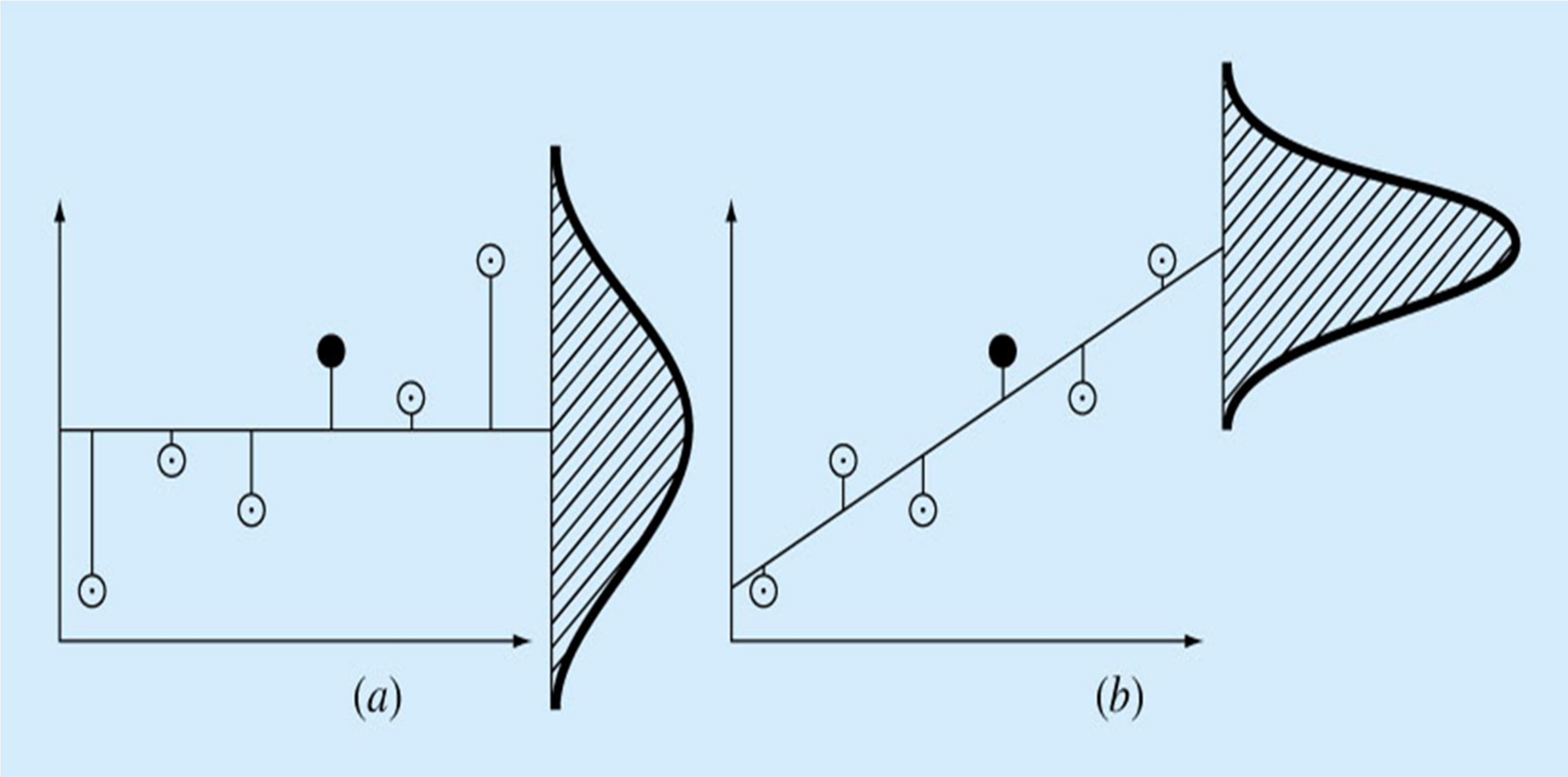
$$\sum a_0 = n a_0$$

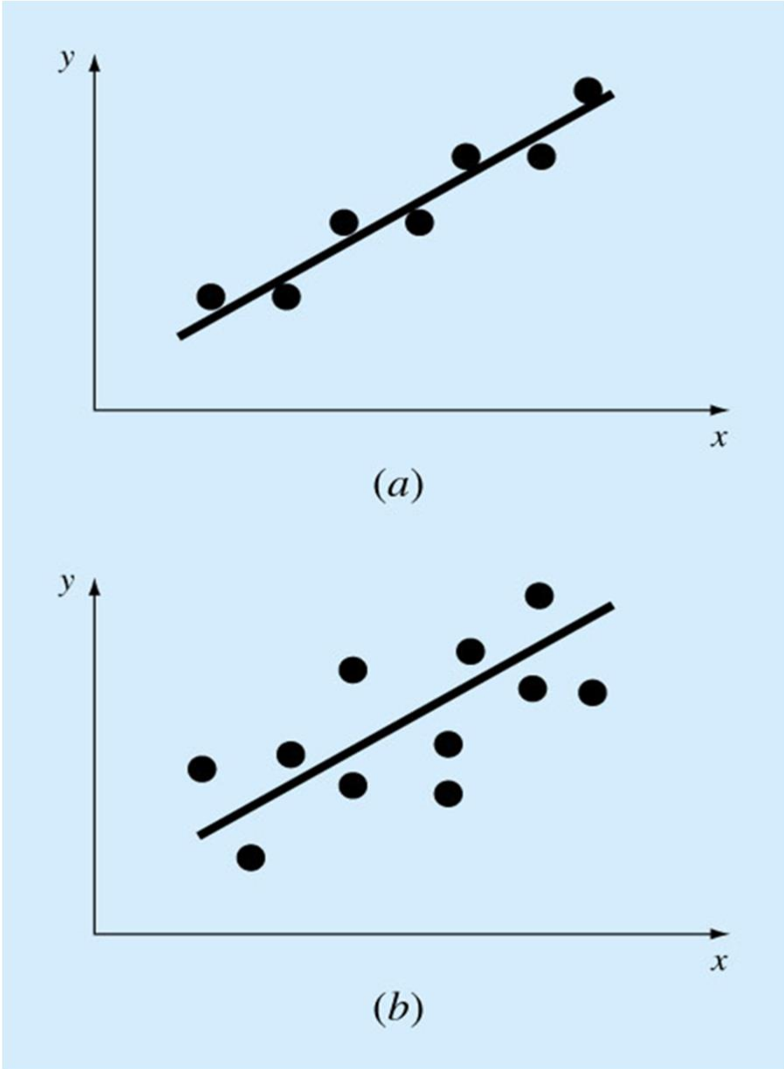
$$n a_0 + \left(\sum x_i\right) a_1 = \sum y_i$$

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \left(\sum x_i\right)^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$


y
Measurement
y_i
$y_i - a_0 - a_1x_i$
Regression line
$a_0 + a_1x_i$
x_i
x


(a)
(b)


y
x
(a)
y
x
(b)

$$r^2 = \frac{S_t - S_r}{S_t}$$

For a perfect fit

Sr=0 and r=r2=1, signifying that the line explains 100 percent of the variability of the data.

For r=r2=0, Sr=St, the fit represents no improvement.

Some engineering data is poorly represented by a straight line. For these cases a curve is better suited to fit the data. The least squares method can readily be extended to fit the data to higher order polynomials.