Optical Metrology

Lecture 2: Random Data and Characterization of Measurement Systems

Content of the Lecture

- Deterministic Data.
- Random Data.
- Characteristics of Random Data.
- Characterization of measurement systems.
- Static and Dynamic characterization.

Deterministic versus Random Data

Deterministic Data

- Any observed data representing a physical phenomenon can be broadly classified as being either deterministic or nondeterministic.
- Deterministic data are those that can be described by an explicit mathematical relationship.

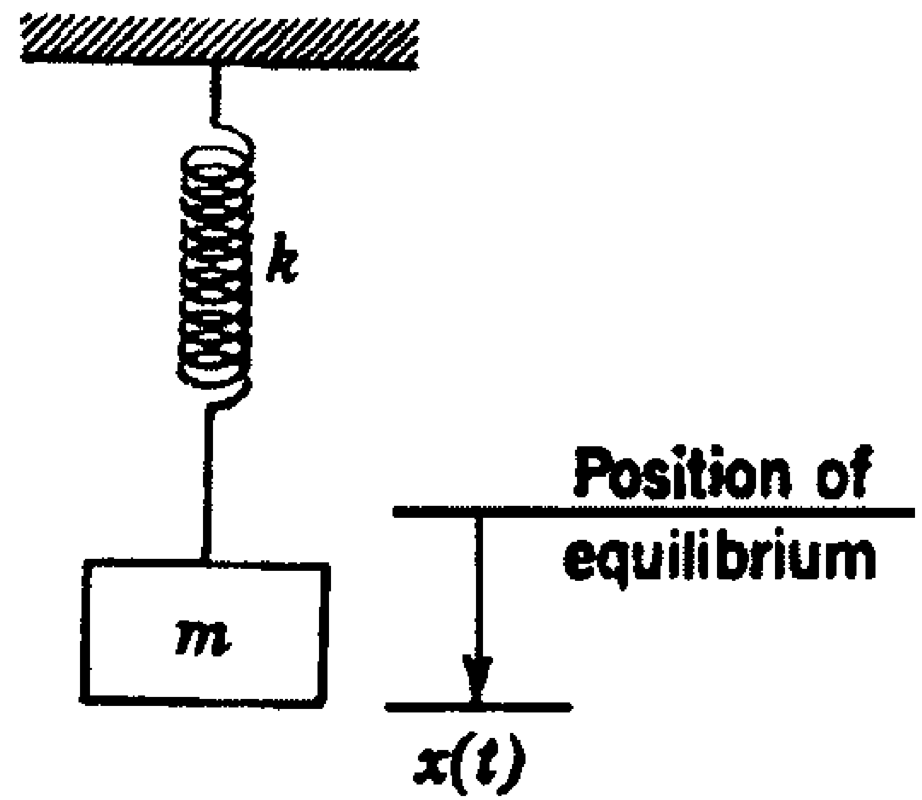
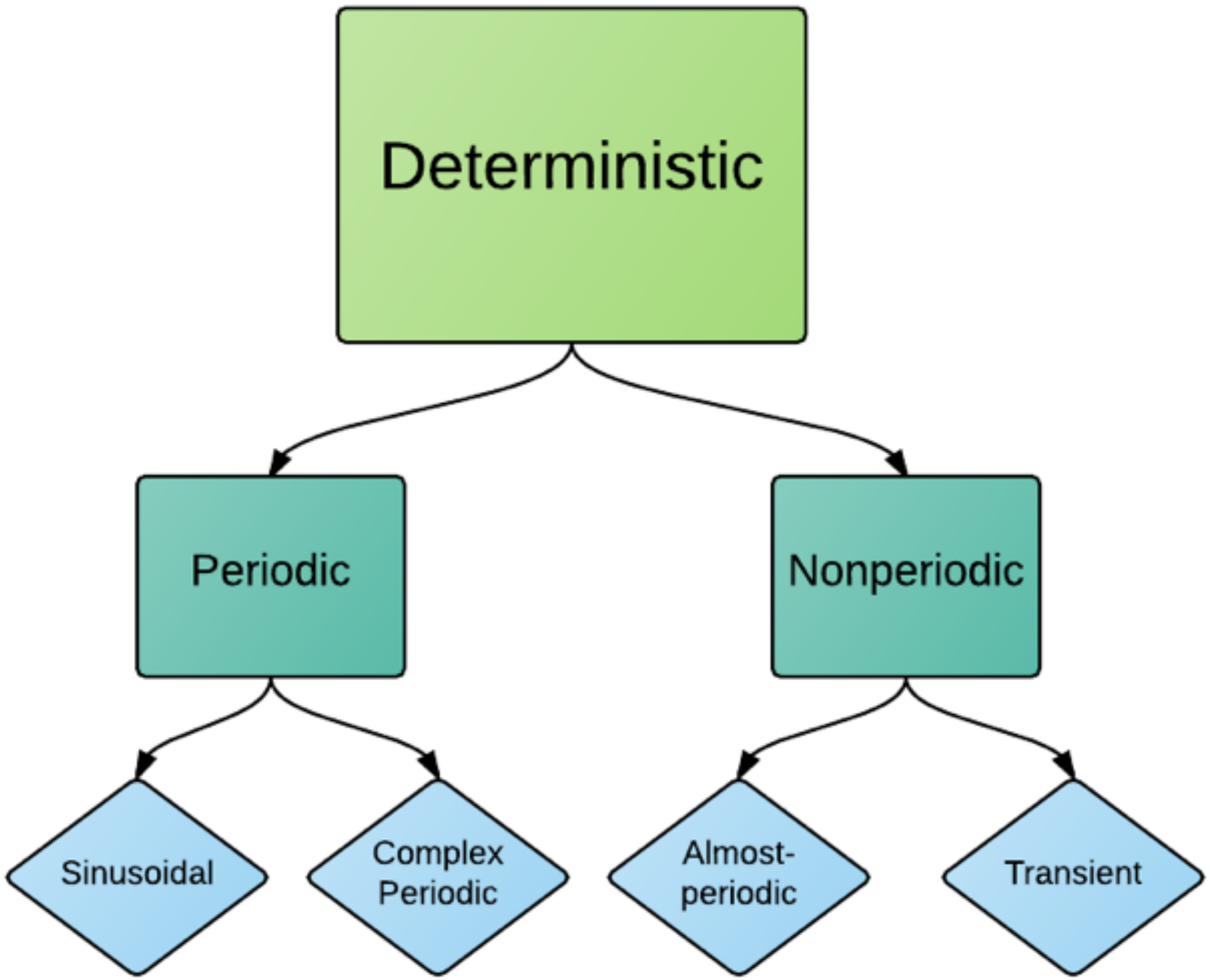


Figure 1.1 Simple spring mass system.

$$x(t) = X \cos \sqrt{\frac{k}{m}}t \quad t \geq 0$$

Classification of Deterministic Data



(Periódico arbitrario)

Sinusoidal

$$x(t) = X \sin(2\pi f_0 t + \phi)$$

Discrete spectra

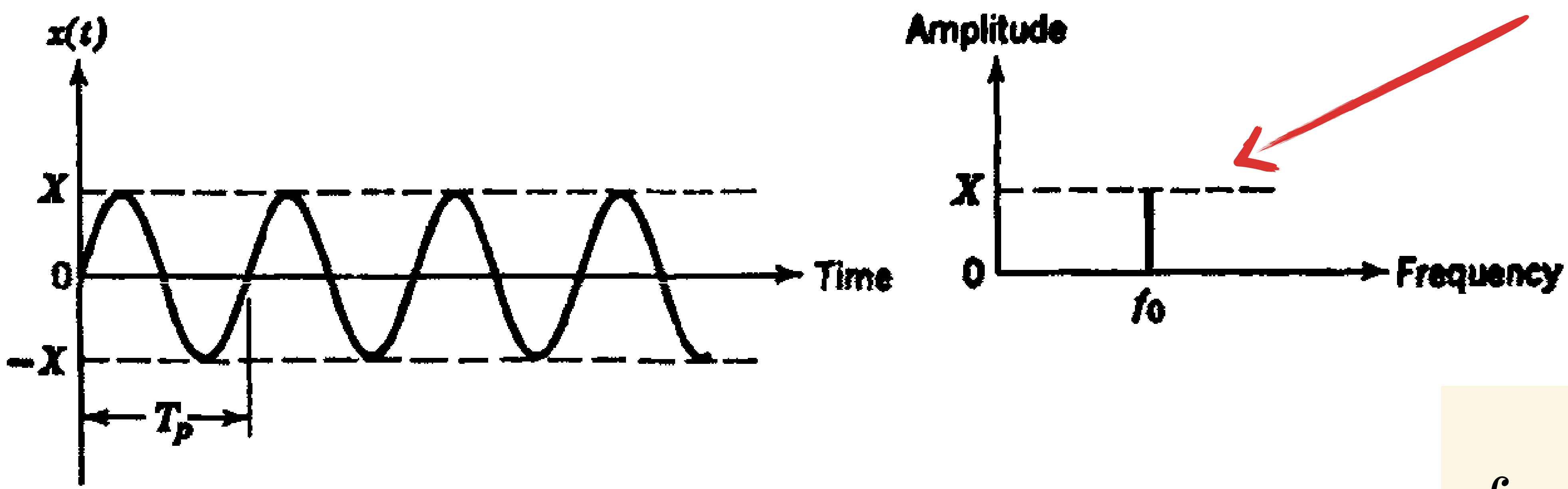


Figure 1.3 Time history and spectrum of sinusoidal data.

$$f_0 = \frac{1}{T_p}$$

Complex Periodic

(Arbitrario)

$$x(t) = X(t \pm nT_p), \quad n = 1, 2, 3, \ldots$$

$$f_1 = \frac{1}{T_p}$$

Data consists of a static component X_0 and an infinite number of sinusoidal components called harmonics. integral multiples of f_1.

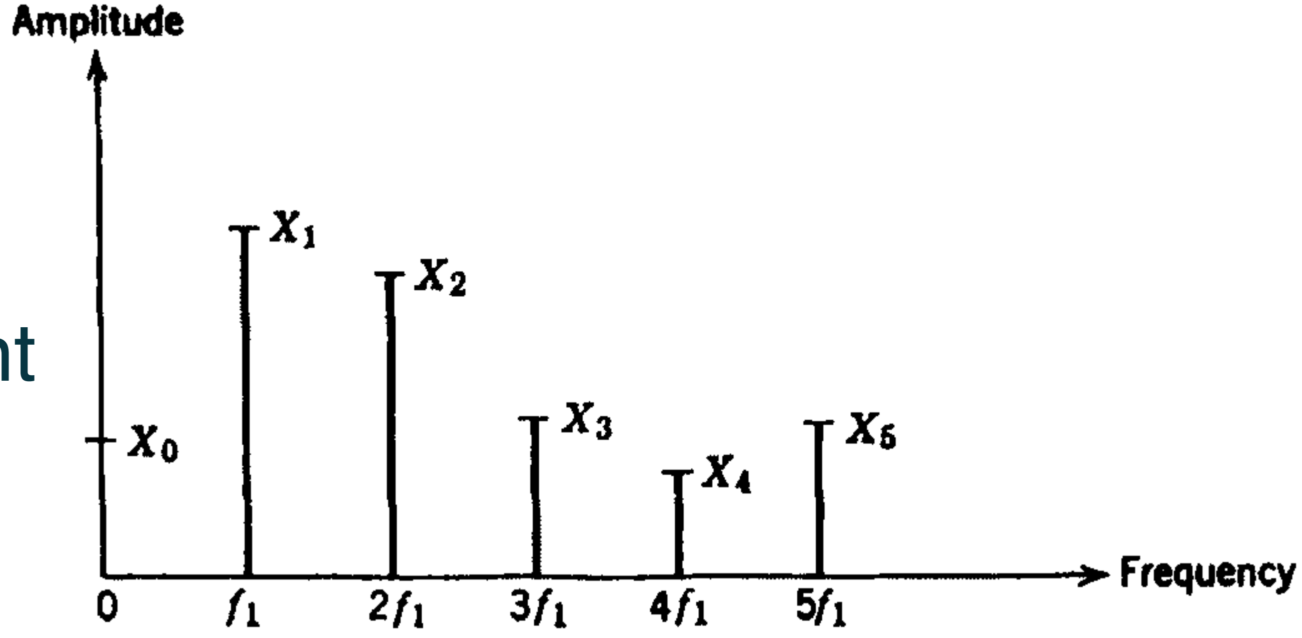


Figure 1.4 Spectrum of complex periodic data.

Almost-periodic

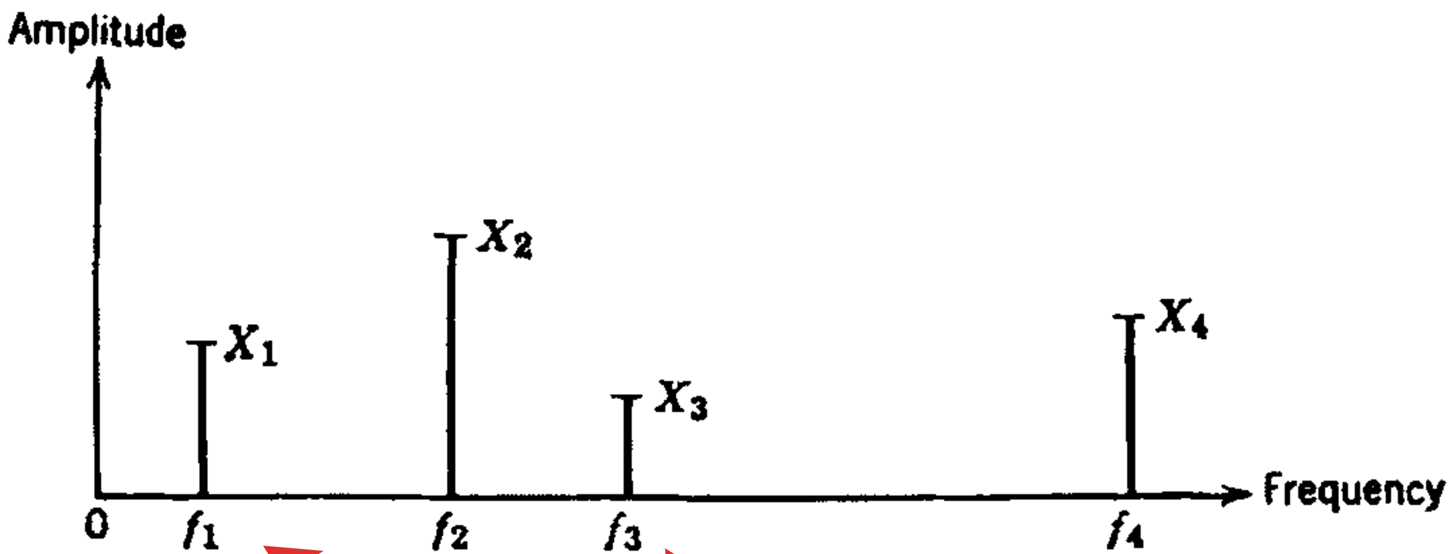


Figure 1.5 Spectrum of almost-periodic data.

No relation

Transient Nonperiodic Data

$$x(t) = \begin{cases} A\,e^{-at} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

$$x(t) = \begin{cases} A\,e^{-at}\cos bt & t \geq 0 \\ 0 & t < 0 \end{cases}$$

$$x(t) = \begin{cases} A & c \geq t \geq 0 \\ 0 & c < t < 0 \end{cases}$$

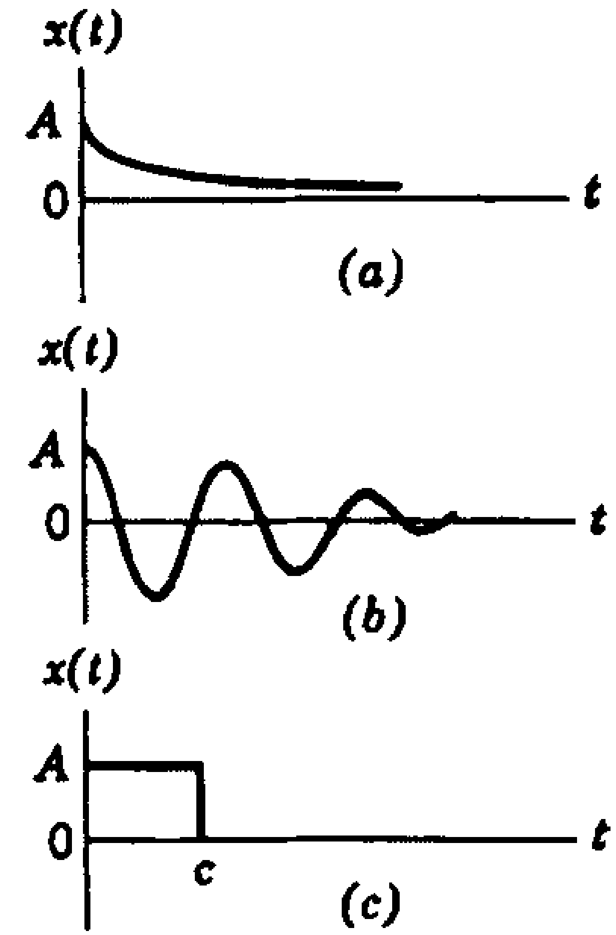


Figure 1.6 Illustrations of transient data.

Continuous spectral representation.

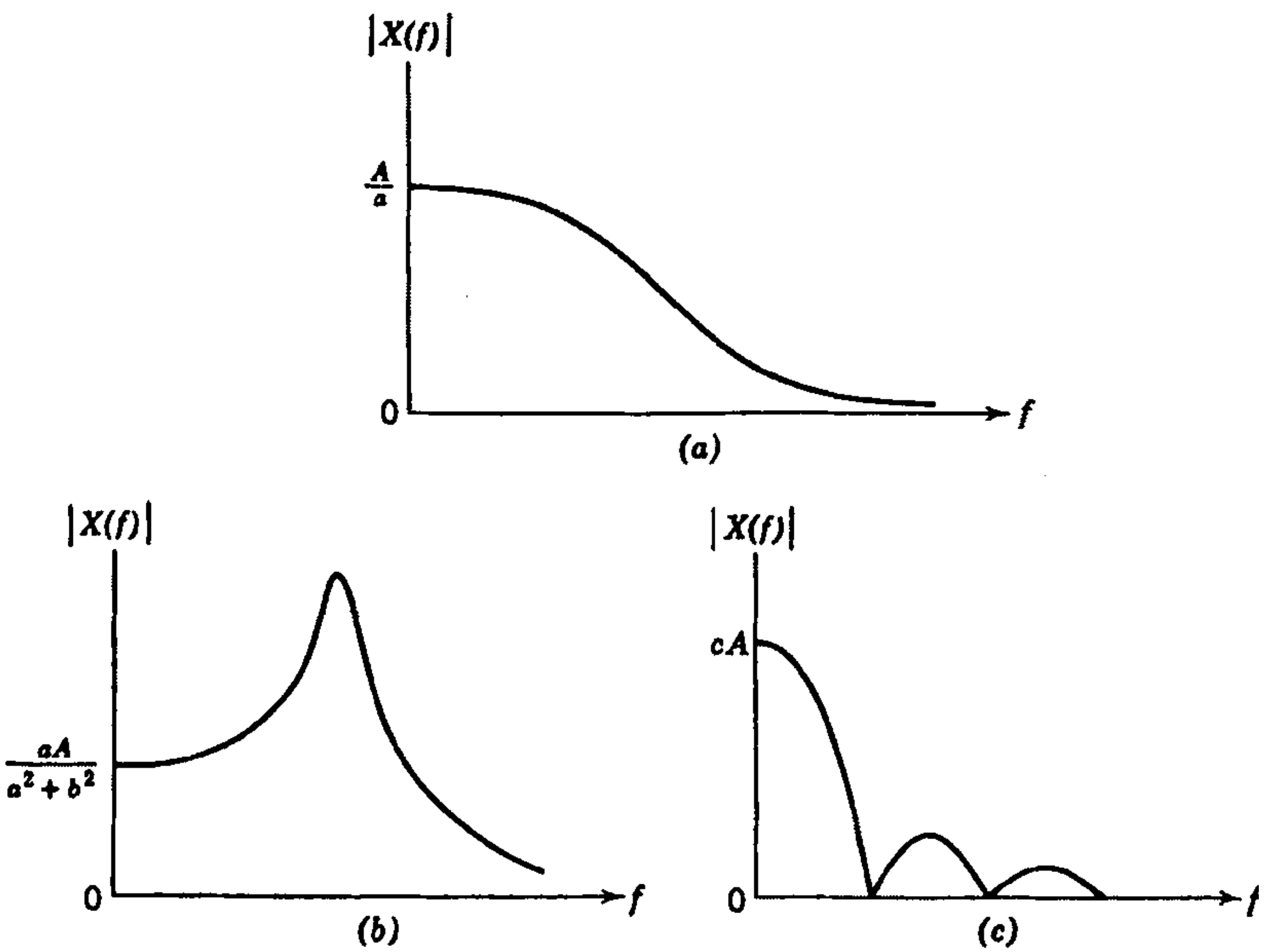
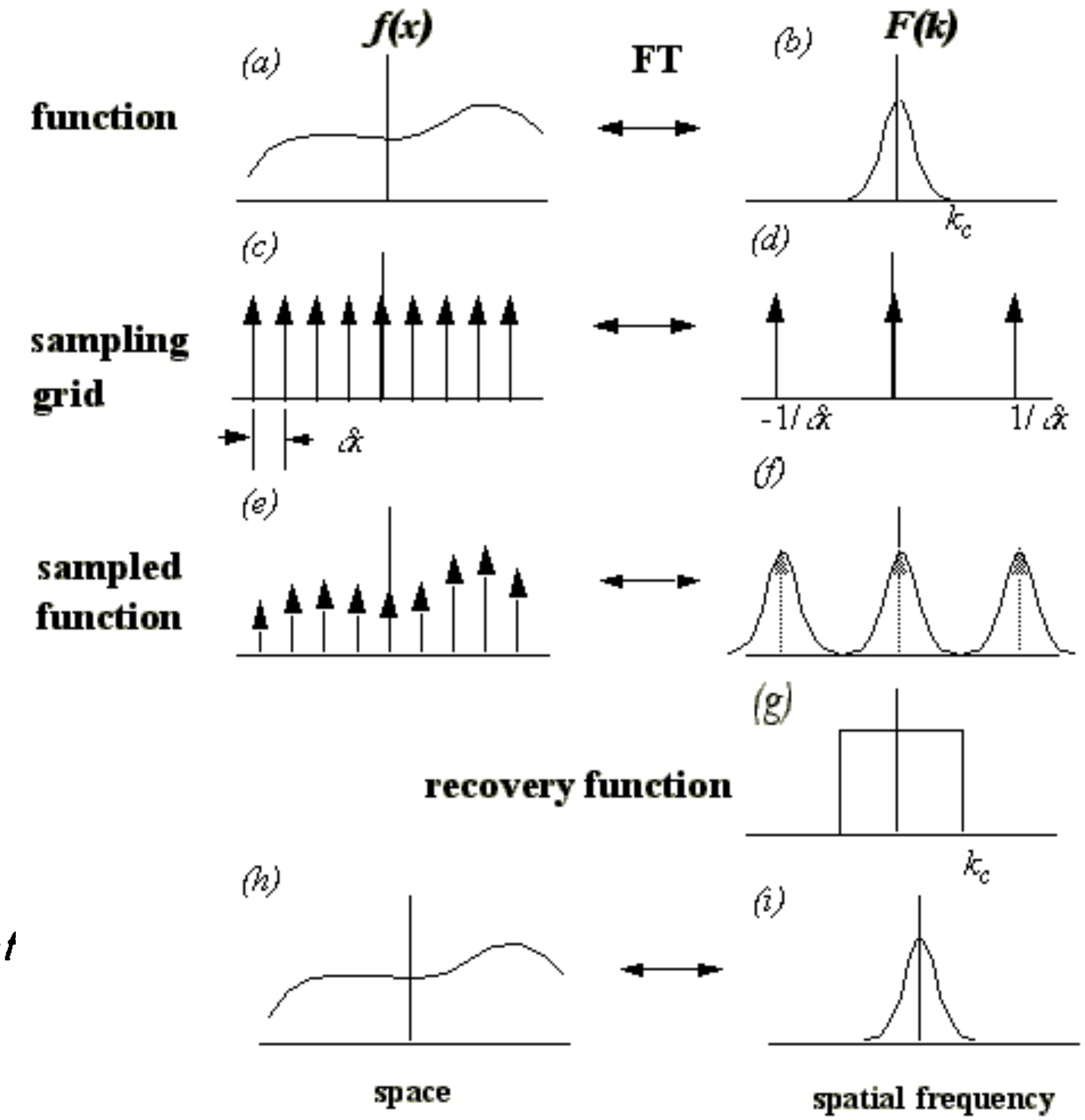
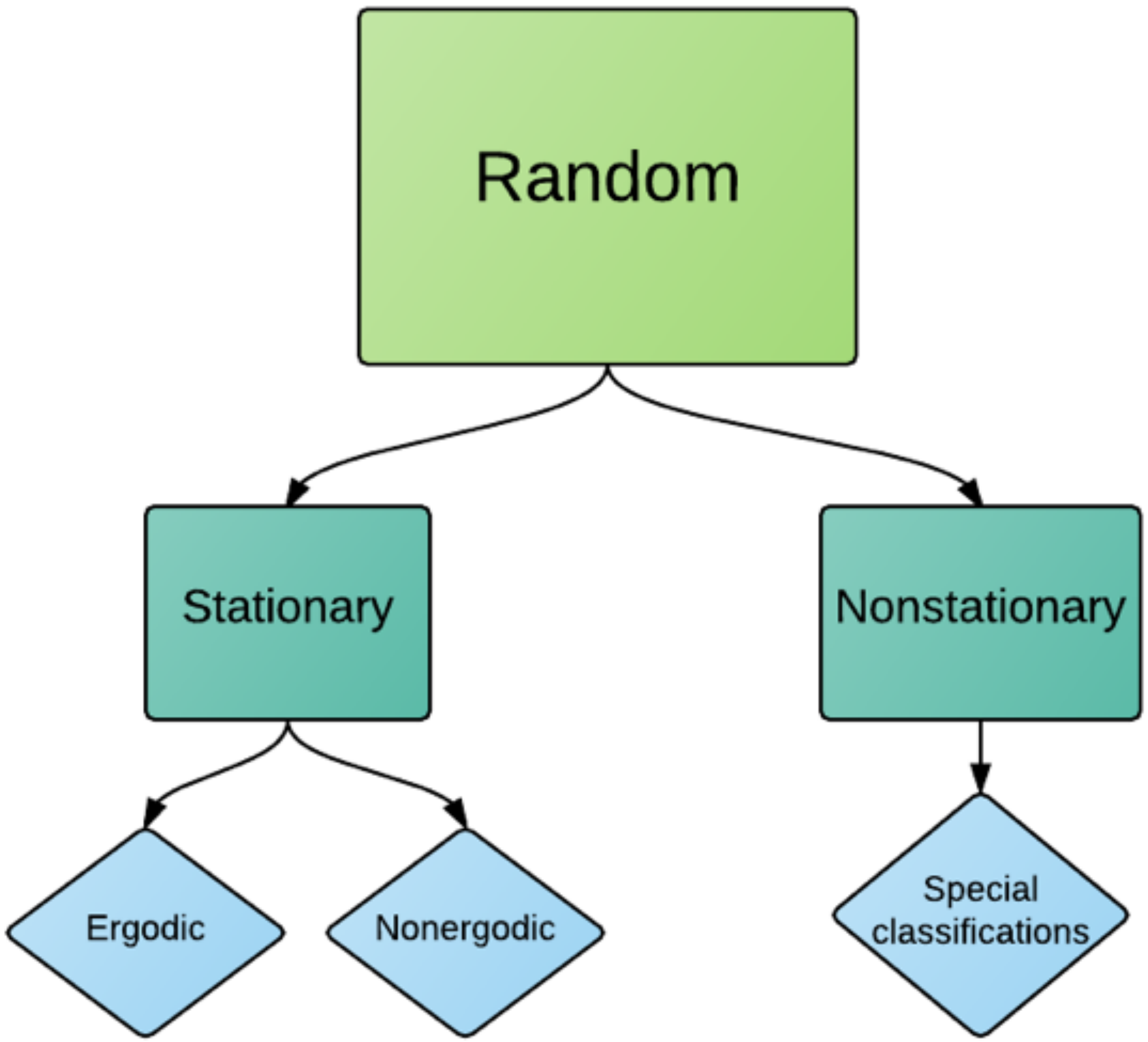


Figure 1.7 Spectra of transient data.

How do you approximate sampling?



Classification of Random Data



Random Data

- A single time history representing a random phenomenon is called a sample function (or a sample record when observed over a finite time interval).
- The collection of all possible sample functions that the random phenomenon might have produced is called a random process or a stochastic process.

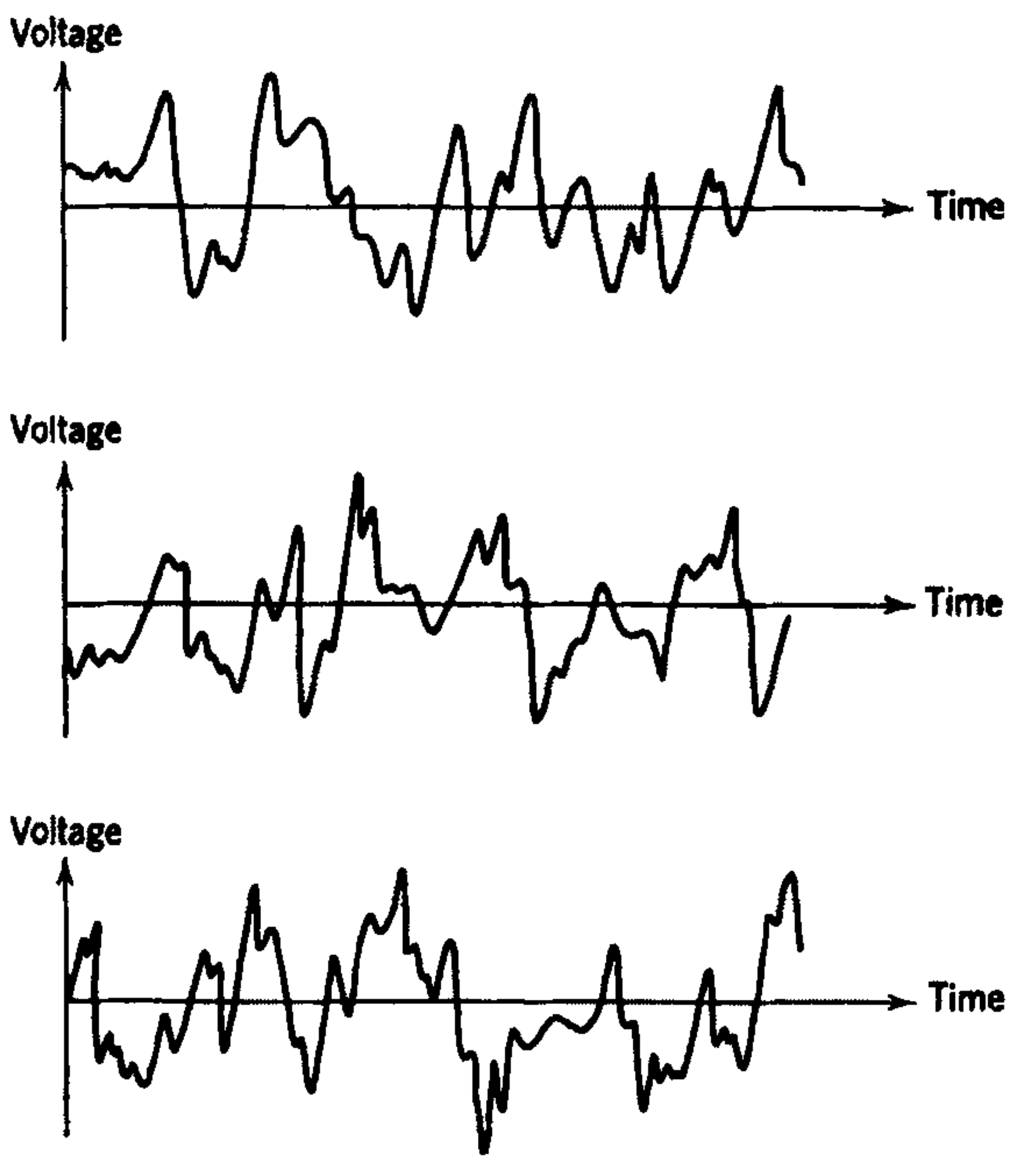


Figure 1.8 Sample records of thermal noise generator outputs.

Stationary Random Data

- A random process can be described by computing average values over the collection of sample functions

$$\mu_x(t_1) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1)$$

$$R_{xx}(t_1, t_1 + \tau) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} x_k(t_1) x_k(t_1 + \tau)$$

- If $\mu_x(t_1)$ and $R_{xx}(t_1, t_1 + \tau)$ vary with t_1, the process is non-stationary.

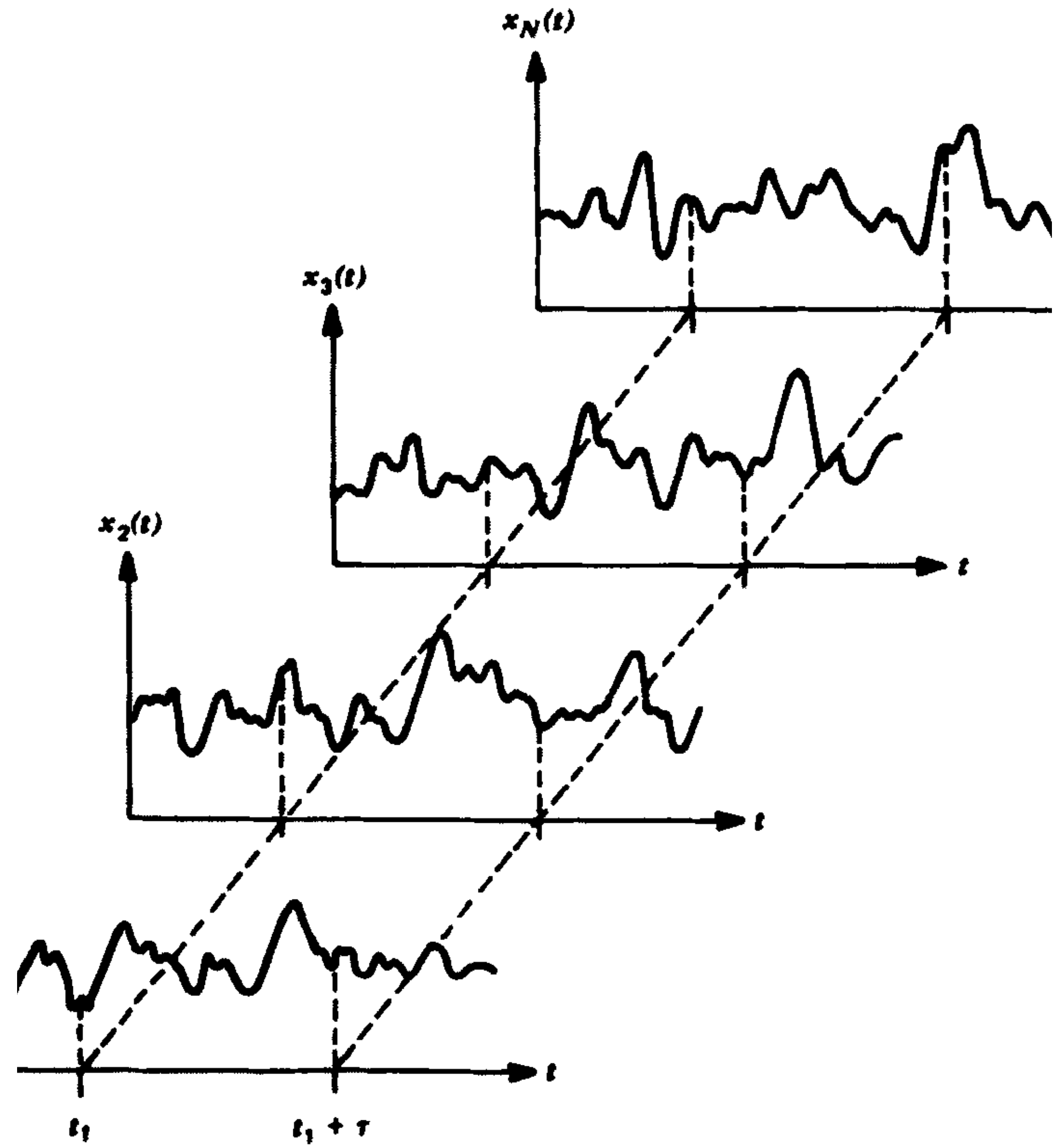
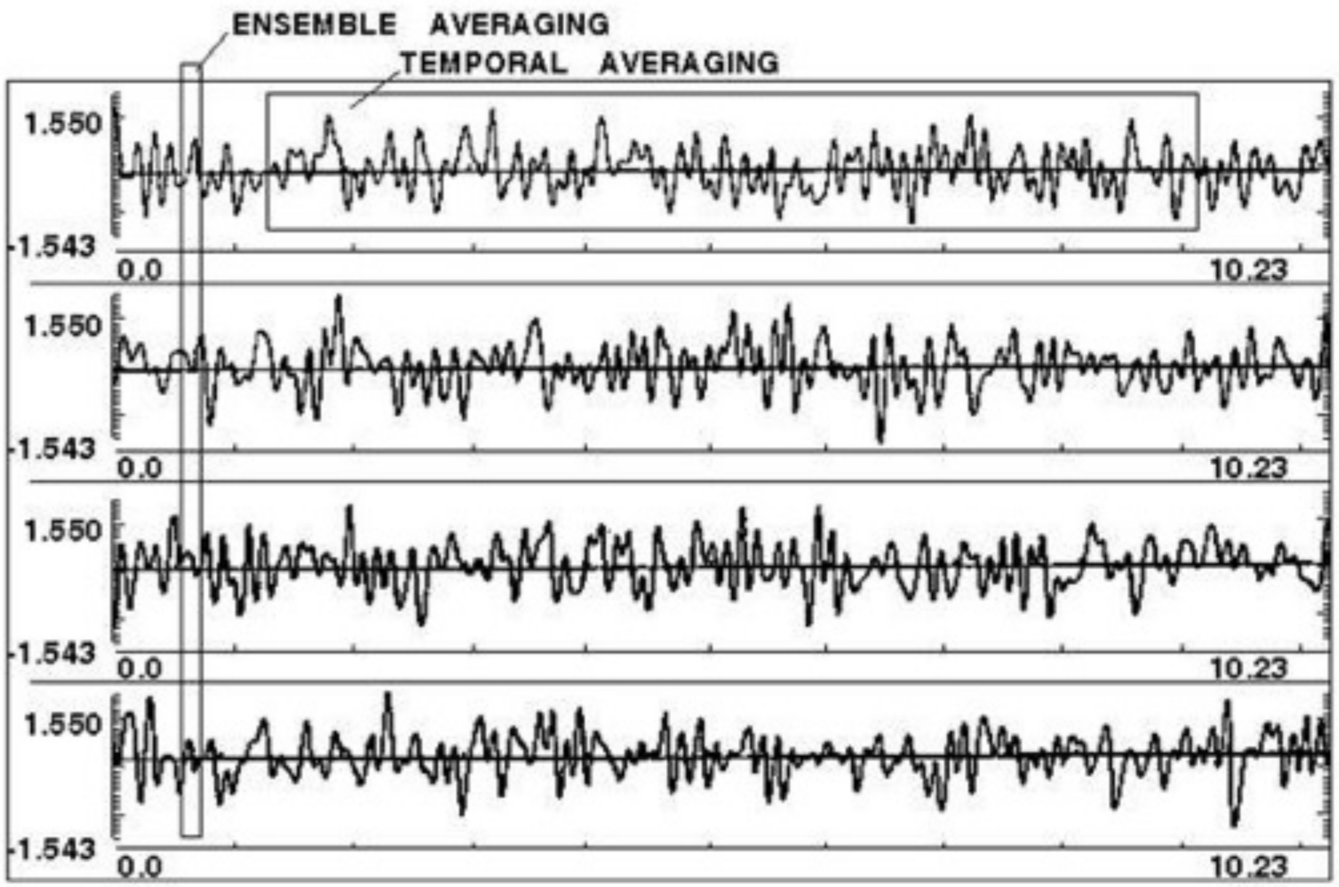


Figure 1.10 Ensemble of time history records defining a random process.

Ergodic Random Data

- A sample can be taken out of any signal, or across a signal and it will be representative of the event.

- This example could be turbulence across 4 flights in similar conditions with similar aircraft.



Analysis of Random Data

- Basic statistical properties of importance for describing single stationary random records are:
  - Mean, mean square values, and moments of order n
  - Probability density functions
  - Autocorrelation functions
  - Autospectral density functions
  - Joint probability density functions
  - Cross-correlation functions

Probability density functions

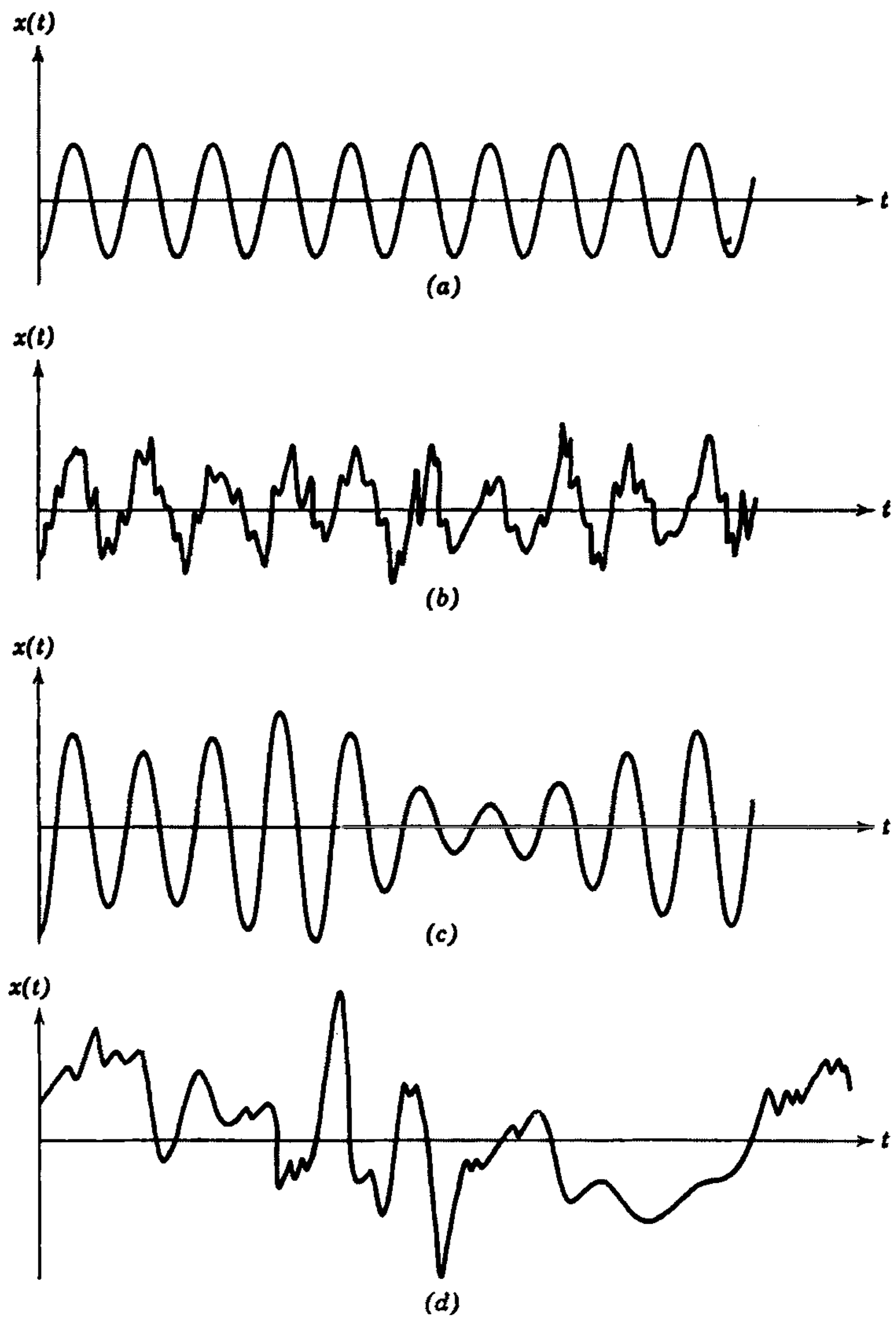


Figure 1.11 Four special time histories. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

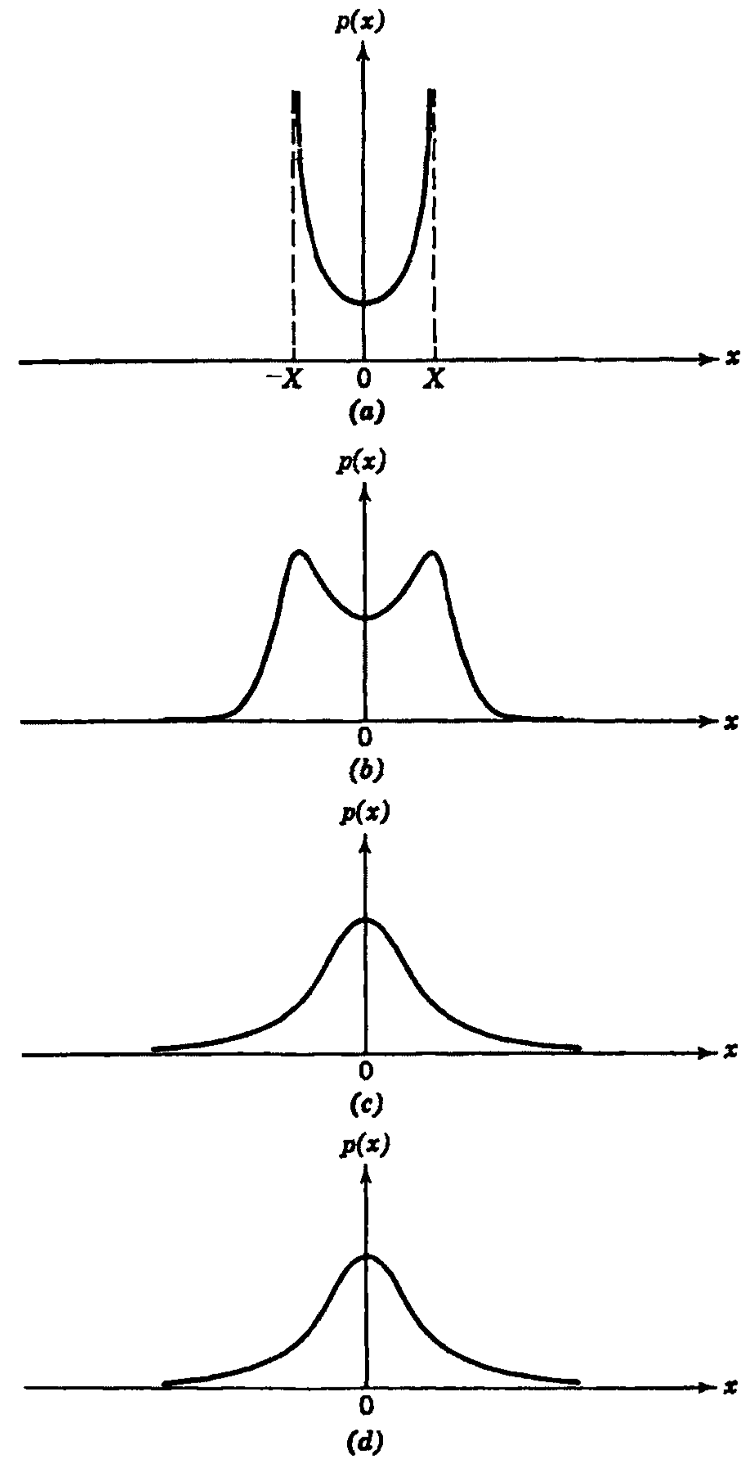


Figure 1.12 Probability density function plots. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

Autocorrelation functions

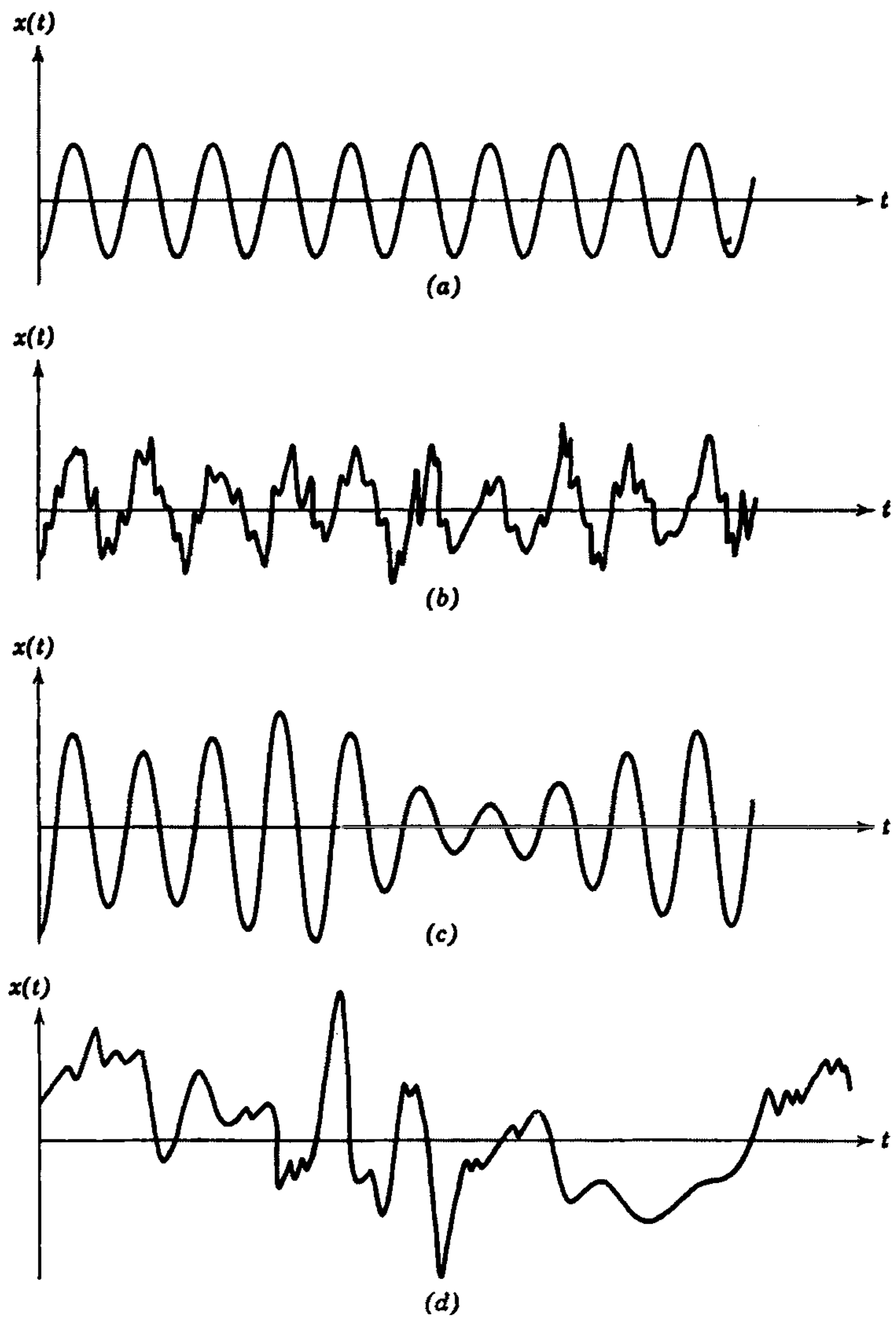


Figure 1.11 Four special time histories. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

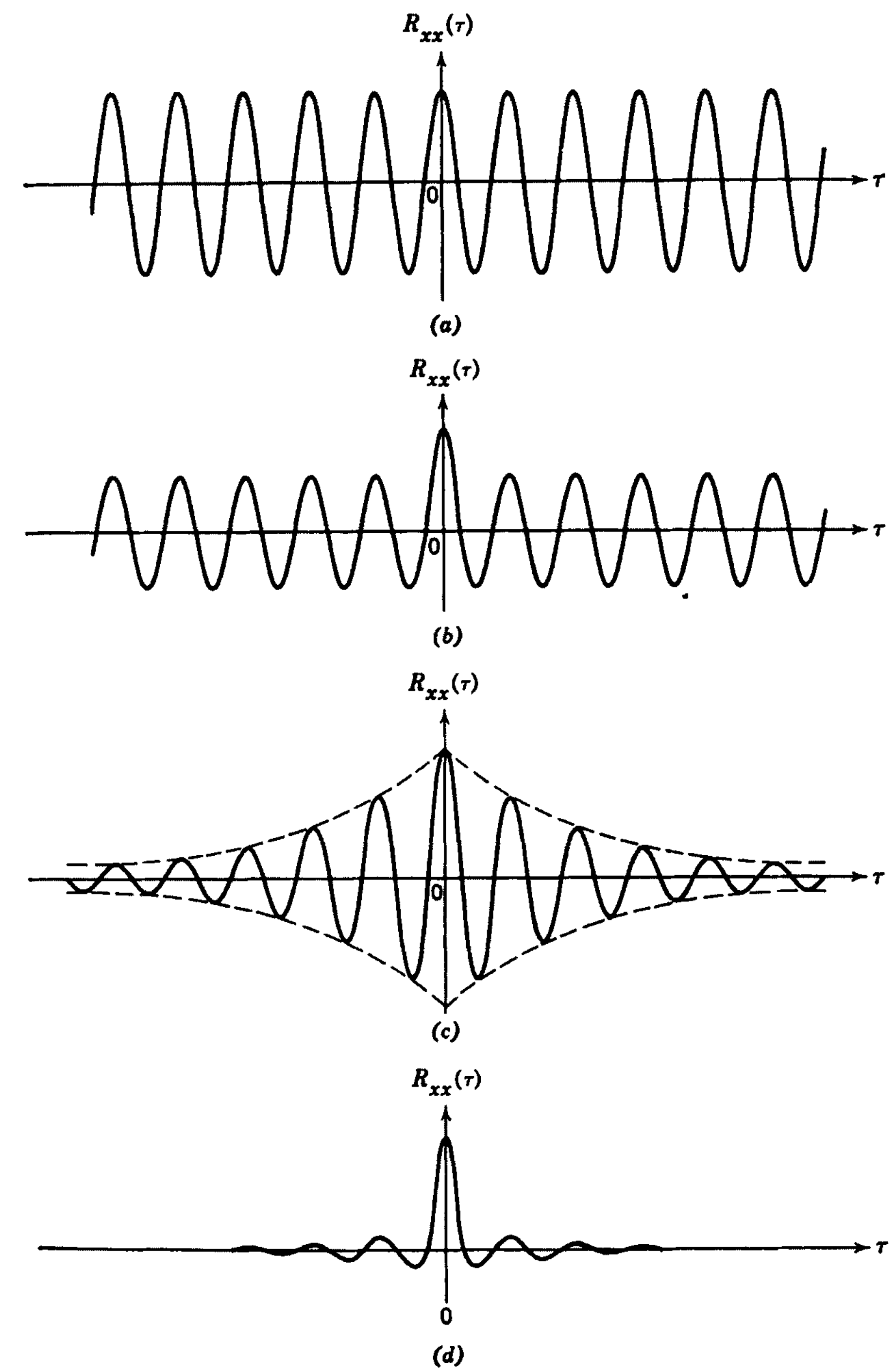


Figure 1.13 Autocorrelation function plots. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

Autospectral density functions

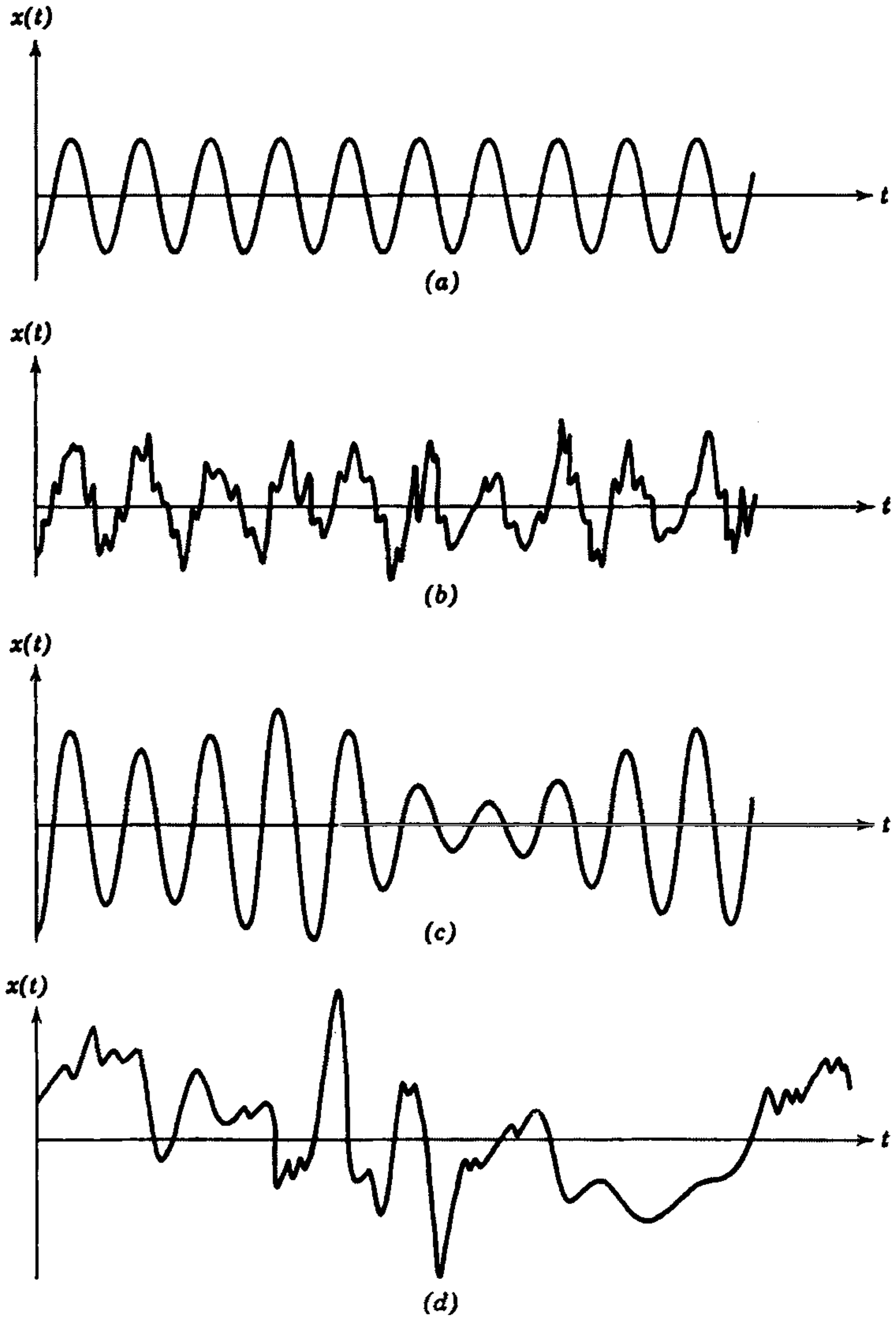


Figure 1.11 Four special time histories. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

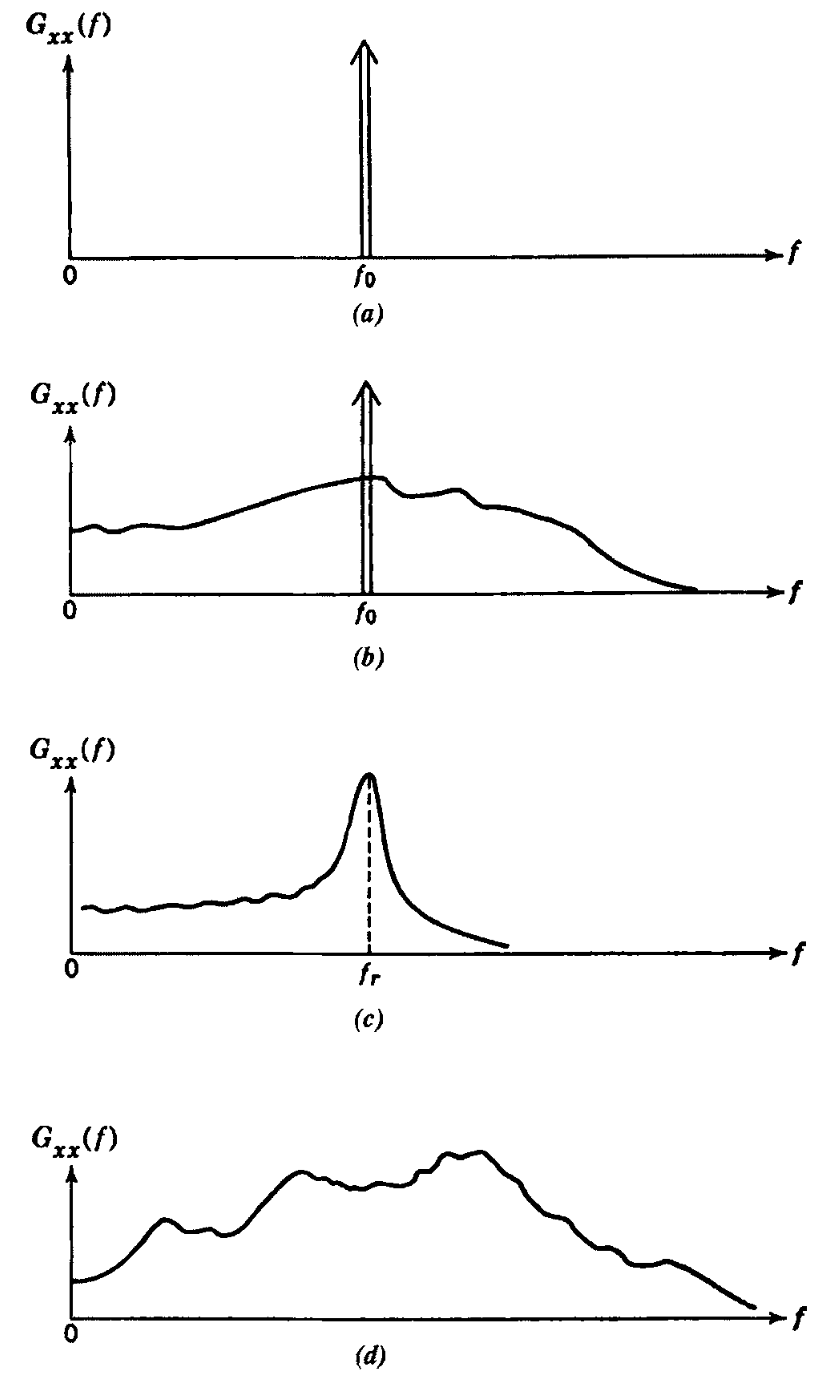
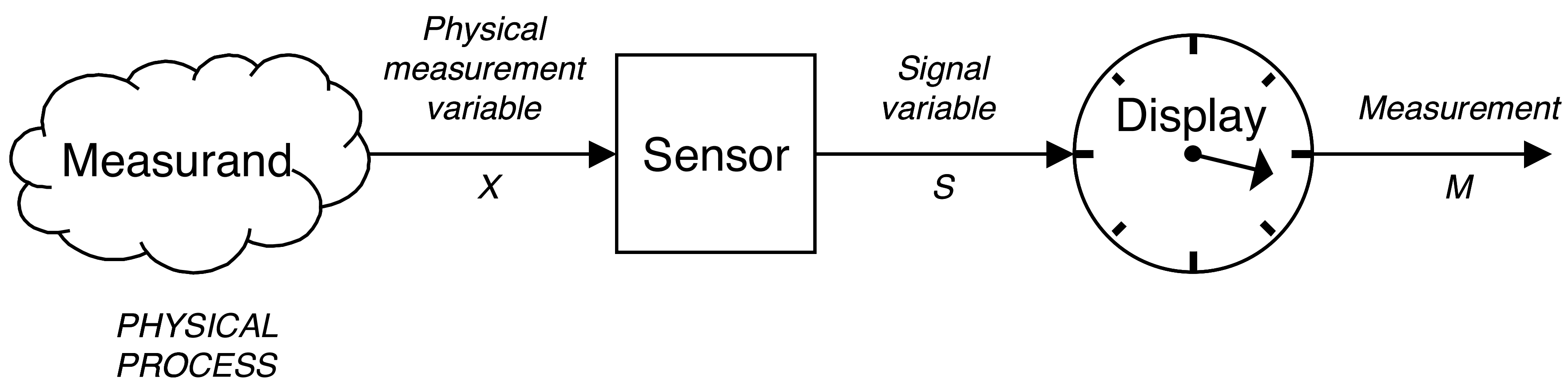


Figure 1.14 Autospectral density function plots. (*a*) Sine wave. (*b*) Sine wave plus random noise. (*c*) Narrow bandwidth random noise. (*d*) Wide bandwidth random noise.

Characterization of Measurement Systems

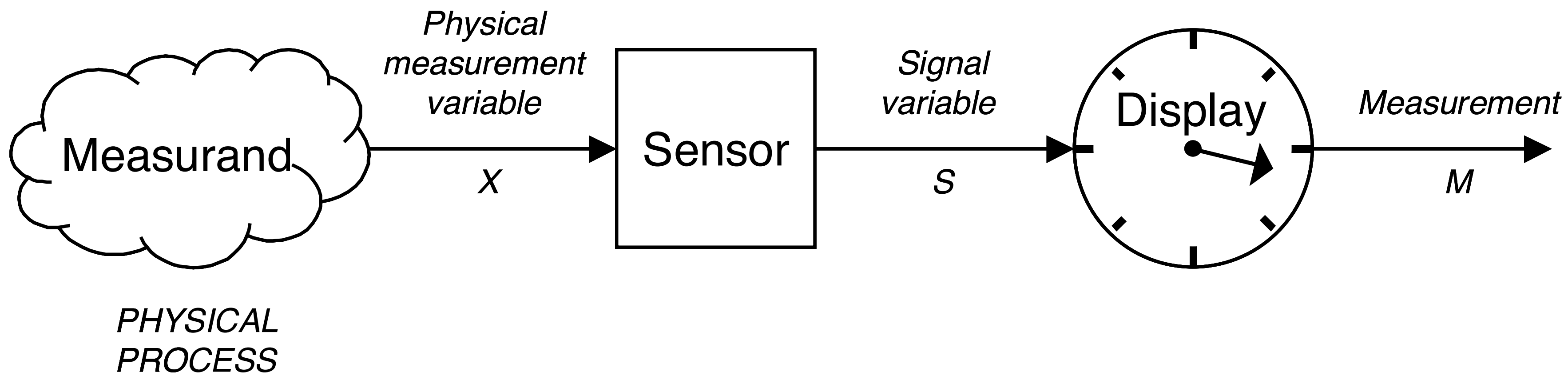
A simple instrument model



- An observable variable X is obtained from the measurand.
  - X is related to the measurand in some KNOWN way (i.e., measuring mass)
- The sensor generates a signal variable that can be manipulated:
  - Processed, transmitted or displayed
- In the example above the signal is passed to a display, where a measurement can be taken

Characterization of Measurement Systems

A simple instrument model



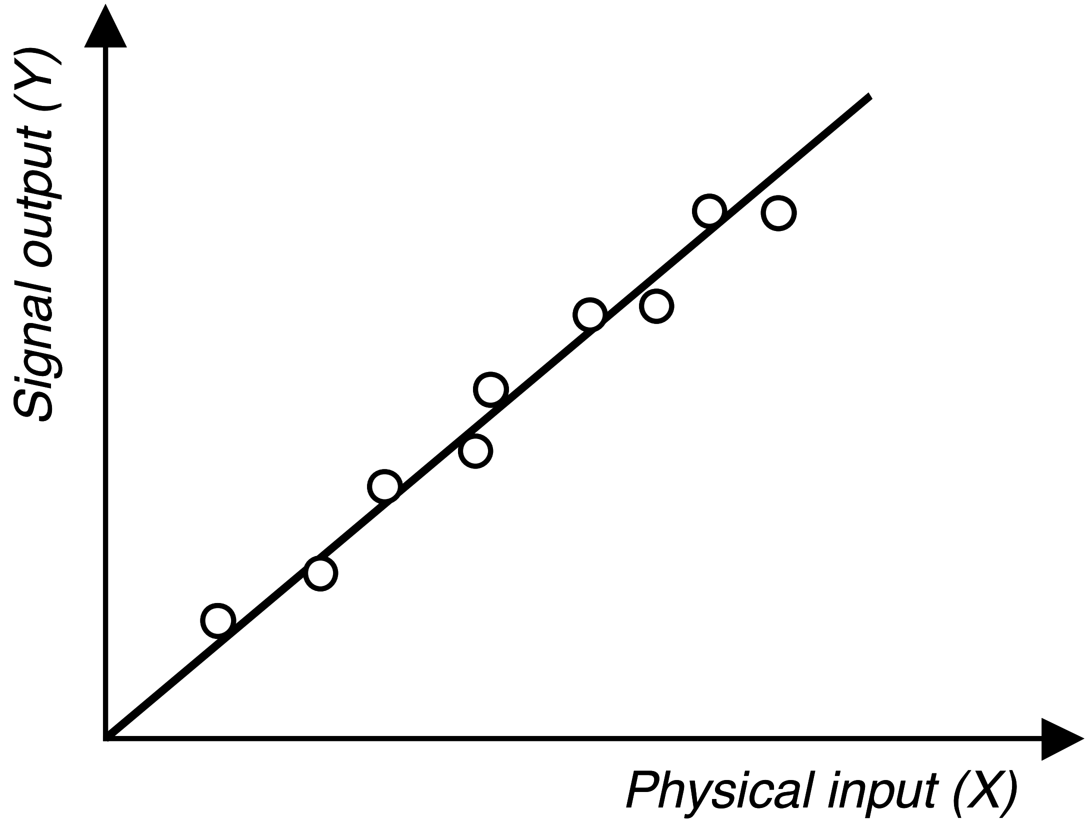
Measurement

- The process of comparing an unknown quantity with a standard of the same quantity (measuring length) or standards of two or more related quantities (measuring velocity)

Characterization of Measurement Systems

The relationship between the physical measurement variable (X) and the signal variable (S)

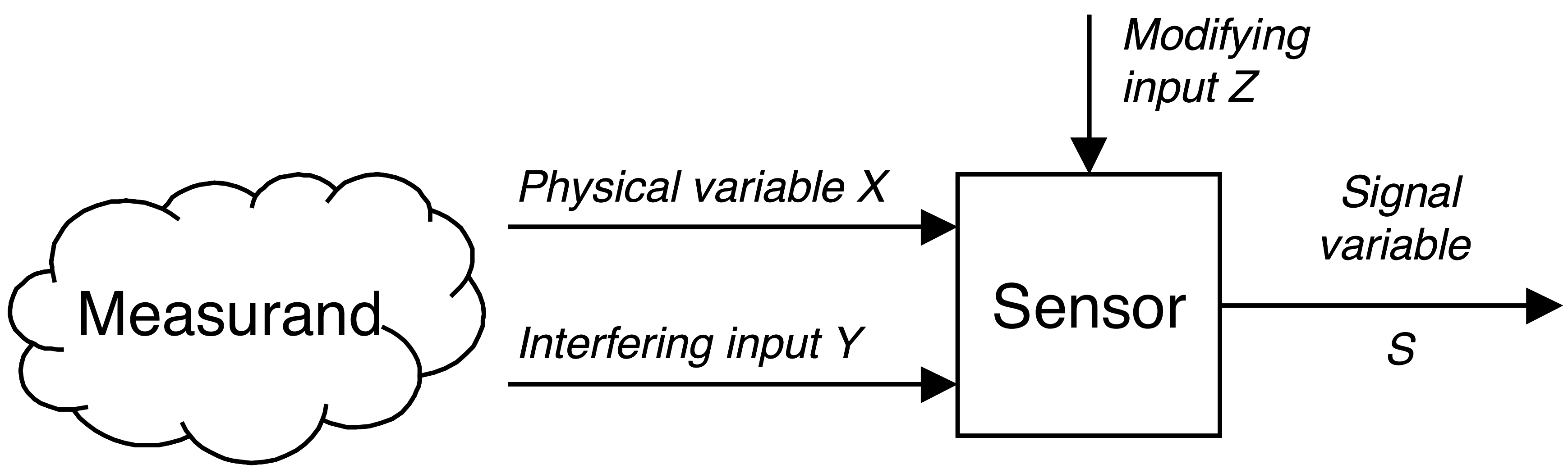
- A sensor or instrument is calibrated by applying a number of KNOWN physical inputs and recording the response of the system.



Characterization of Measurement Systems

Interfering inputs (Y)

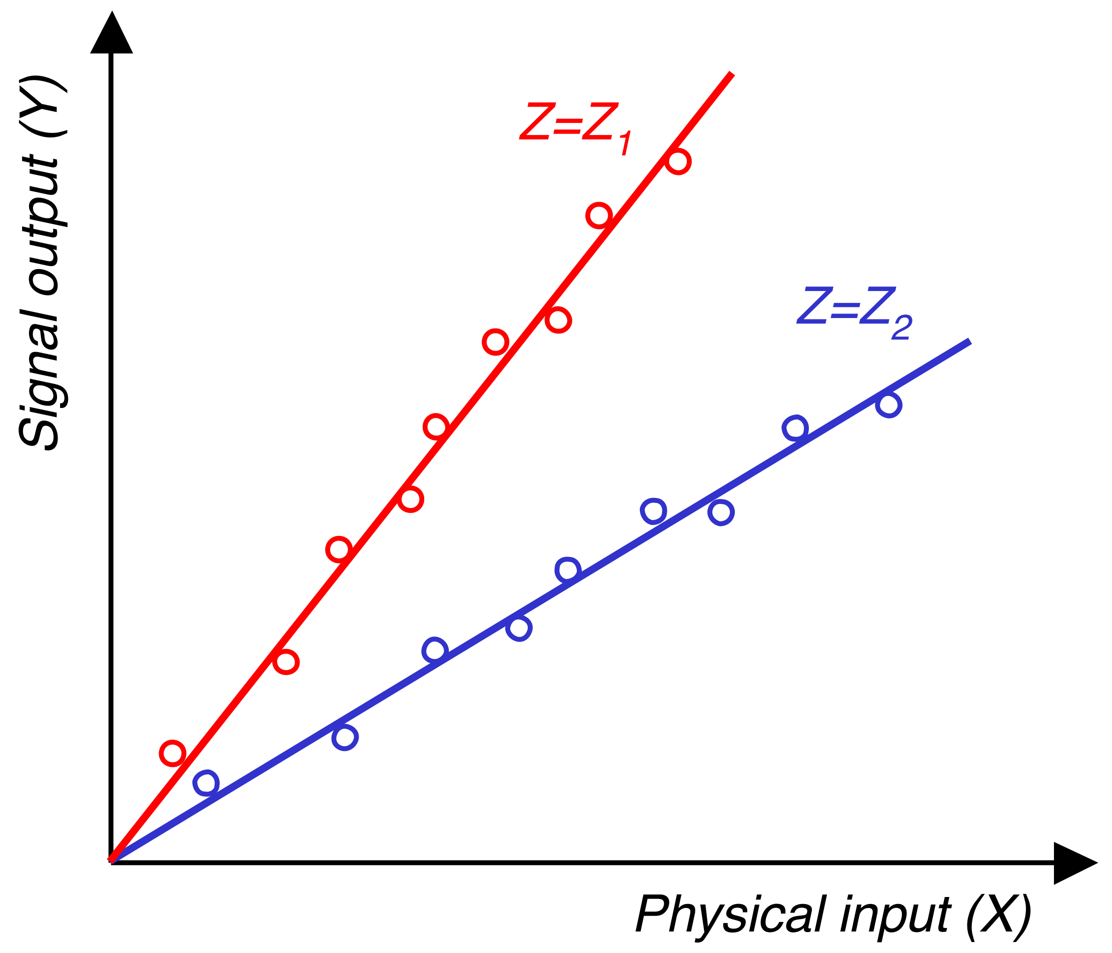
- Those that the sensor to respond as the linear superposition with the measurand variable X.
  - Linear superposition assumption: S(aX +bY)=aS(X)+bS(Y)



Characterization of Measurement Systems

Modifying inputs (Z)

- Those that change the behavior of the sensor and, hence, the calibration curve
  - Temperature is a typical modifying input.



Characterization of Measurement Systems

Static characteristics

- The properties of the system after all transient effects have settled to their final or steady state.
    - Accuracy
    - Discrimination
    - Precision
    - Errors
    - Drift
    - Sensitivity
    - Linearity
    - Hystheresis

Characterization of Measurement Systems

Dynamic characteristics

- The properties of the system transient response to an input.
  - Zero order systems.
  - First order systems.
  - Second order systems.

Characterization of Measurement Systems

Static quantities can be classified as general or specific.

- **Specific quantities** are related to unique variables related to the measurement instrument.
- **General quantities** are common to all measurement instruments.

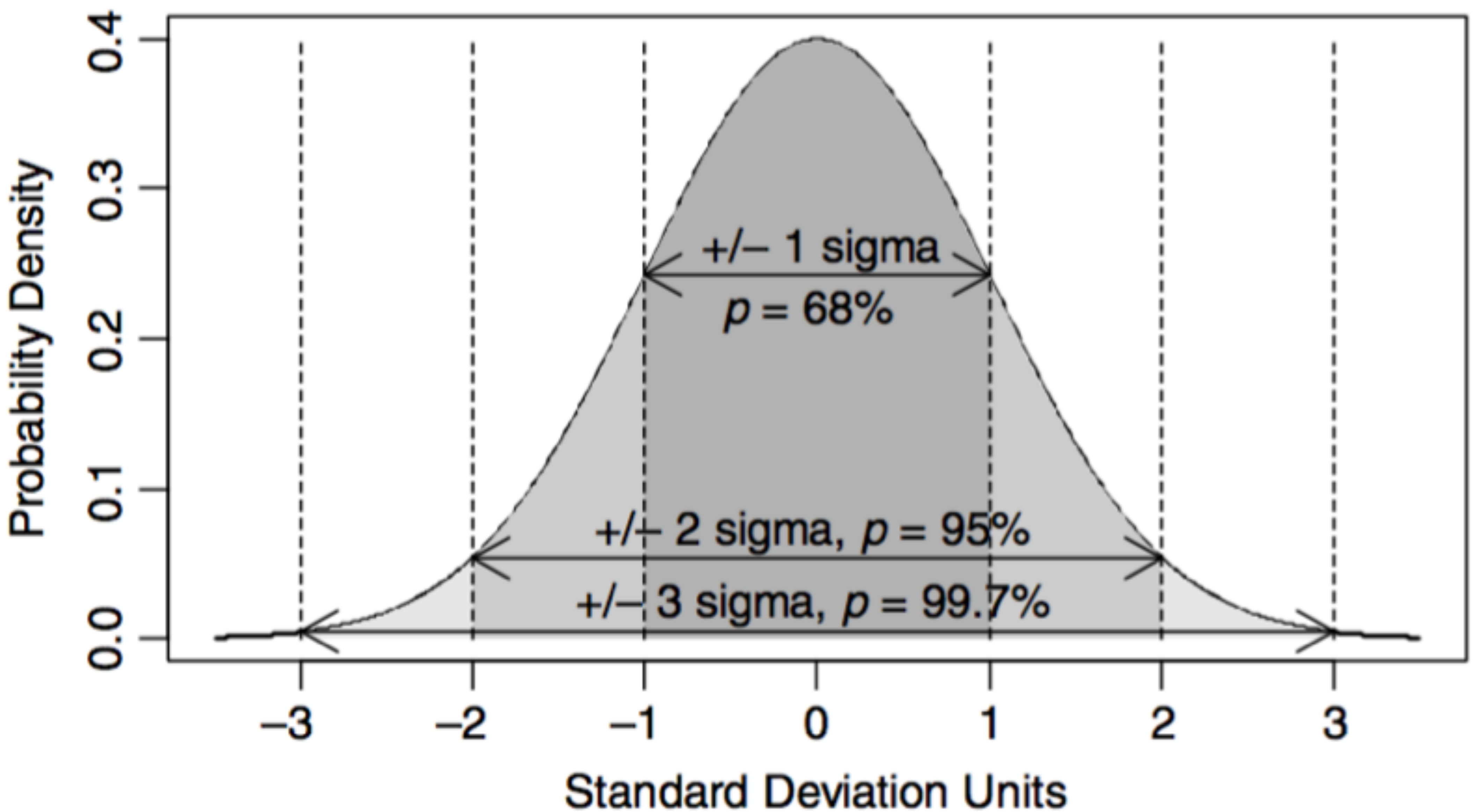


Figure 2.22 One-, two-, and three-sigma rules shown as areas under the normal density curve.

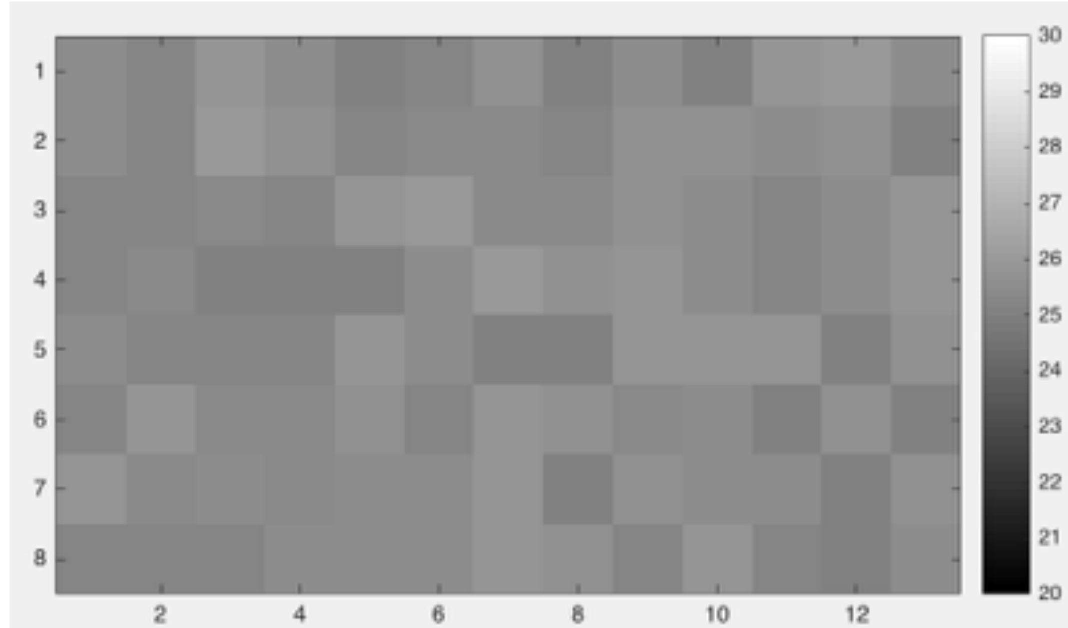
Confidence interval

more general notation with the classic percentile. The resulting confidence interval, called the *one-sample t confidence interval*, is of the form analogous to (3.22), that is, it can be written as

$$\overline{X} \pm t_{n-1}(\alpha/2) \cdot s/\sqrt{n}. \tag{3.25}$$

An implementation of this confidence interval is demonstrated in the following example.

Example 3.1 Consider Small Image data describing an 8 by 13 pixel image of a monochromatic, highly uniform tile in three wide spectral bands (in reflectance units). Our goal is to estimate the "true" tile reflectances μ_1, μ_2, and μ_3 in the three spectral bands, respectively. First, we are going to concentrate on the reflectances in Band 1. Since the tile surface is highly uniform, it makes sense to assume that reflectances in Band 1 for all pixels are independent random variables X_i, $i = 1, \ldots, n = 104$, all having the same distribution with the mean $E(X_i) = \mu_1$ (assuming that the measurements are unbiased). We expect the data to follow the normal distribution because the variability is largely due to the measurement error. For $\alpha = 0.05$, we obtain $t_{n-1}(\alpha/2) = 1.98$. For Band 1 data, we have $\overline{x} = 25.0245$, $s = 0.2586$, and the resulting half of the length of the confidence interval is equal to $h = t_{n-1}(\alpha/2) \cdot s/\sqrt{n} = 0.0503$. The confidence interval can now be written as 25.0245 ± 0.0503 or $(24.9742, 25.0748)$. □



Example 3.1 Consider Small Image data describing an 8 by 13 pixel image of a monochromatic, highly uniform tile in three wide spectral bands (in reflectance units). Our goal is to estimate the "true" tile reflectances μ_1, μ_2, and μ_3 in the three spectral bands, respectively. First, we are going to concentrate on the reflectances in Band 1. Since the tile surface is highly uniform, it makes sense to assume that reflectances in Band 1 for all pixels are independent random variables X_i, $i = 1, \ldots, n = 104$, all having the same distribution with the mean $E(X_i) = \mu_1$ (assuming that the measurements are unbiased). We expect the data to follow the normal distribution because the variability is largely due to the measurement error. For $\alpha = 0.05$, we obtain $t_{n-1}(\alpha/2) = 1.98$. For Band 1 data, we have $\overline{x} = 25.0245$, $s = 0.2586$, and the resulting half of the length of the confidence interval is equal to $h = t_{n-1}(\alpha/2) \cdot s/\sqrt{n} = 0.0503$. The confidence interval can now be written as 25.0245 ± 0.0503 or $(24.9742,\ 25.0748)$. $\square$

Calibration: Simple Linear Regression Model

Example 4.1 The Landsat Program is a series of Earth-observing satellite missions jointly managed by NASA and the U.S. Geological Survey since 1972. Due to the long-term nature of the program, there is a significant interest in the long-term calibration of the results, so that measurements taken at different times can be meaningfully compared. One approach to this calibration problem is discussed by Anderson (2010). As part of the approach, Landsat measurements of a fixed desert site were collected. The desert site was confirmed to be sufficiently stable over time, so that the changes in measurements can be attributed to a drift of the measuring instrument, except for some factors such as the Sun position in the sky. **In this example, we consider the surface reflectance measurements of the desert site performed at 76 different times (different days and times of the day). The reflectance measurements are from one spectral band (Band 2) of the instrument.** For each time of the measurement, we also know the solar elevation angle.

Calibration: Simple Linear Regression Model

In order to investigate a relationship between reflectance in Band 2 and the solar elevation angle, we can create a scatter plot of the two variables as shown in Figure 4.1. Based on the pattern in the scatter plot, we expect a linear relationship between the two variables.

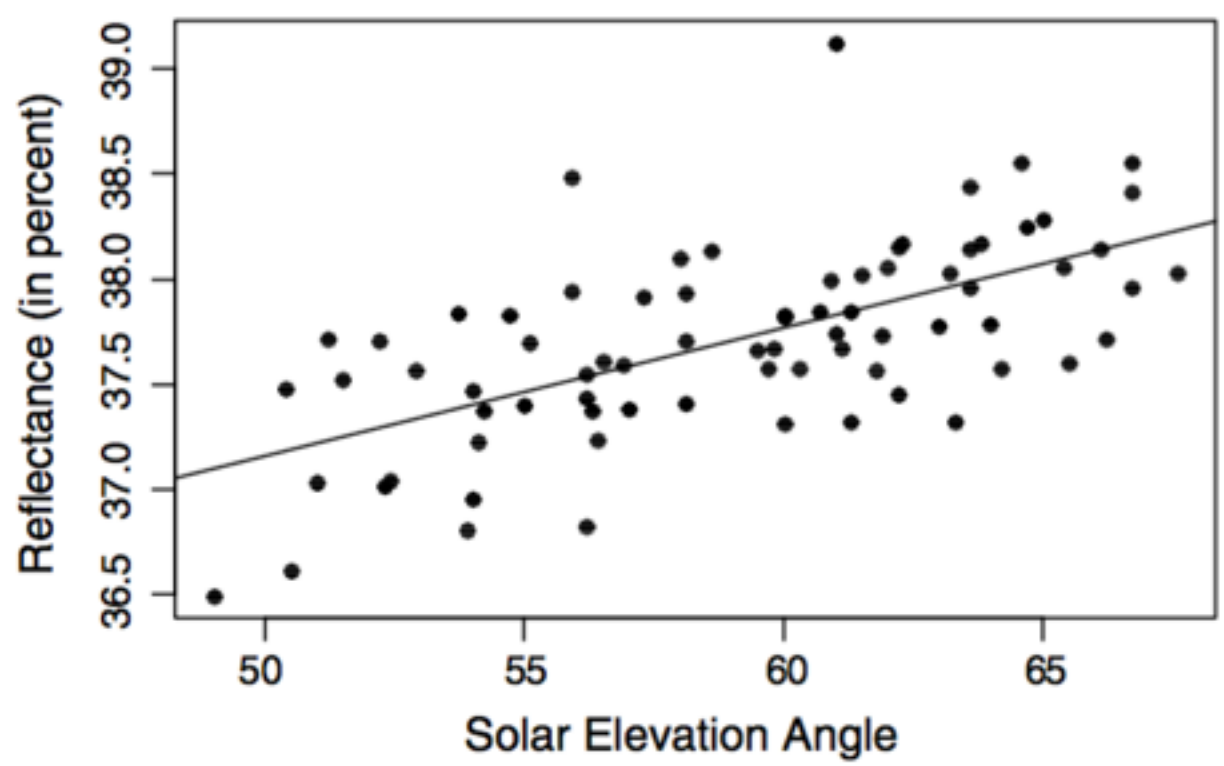


Figure 4.1 A scatter plot of reflectance in Band 2 versus the solar elevation angle for Landsat data discussed in Example 4.1.

Calibration: Simple Linear Regression Model

In the simplest scenario of a linear relationship between the response Y and a single predictor x, as seen in Figure 4.1, we can describe this relationship using a population linear regression model written as

$$Y = \beta_0 + \beta_1 x + \varepsilon,$$

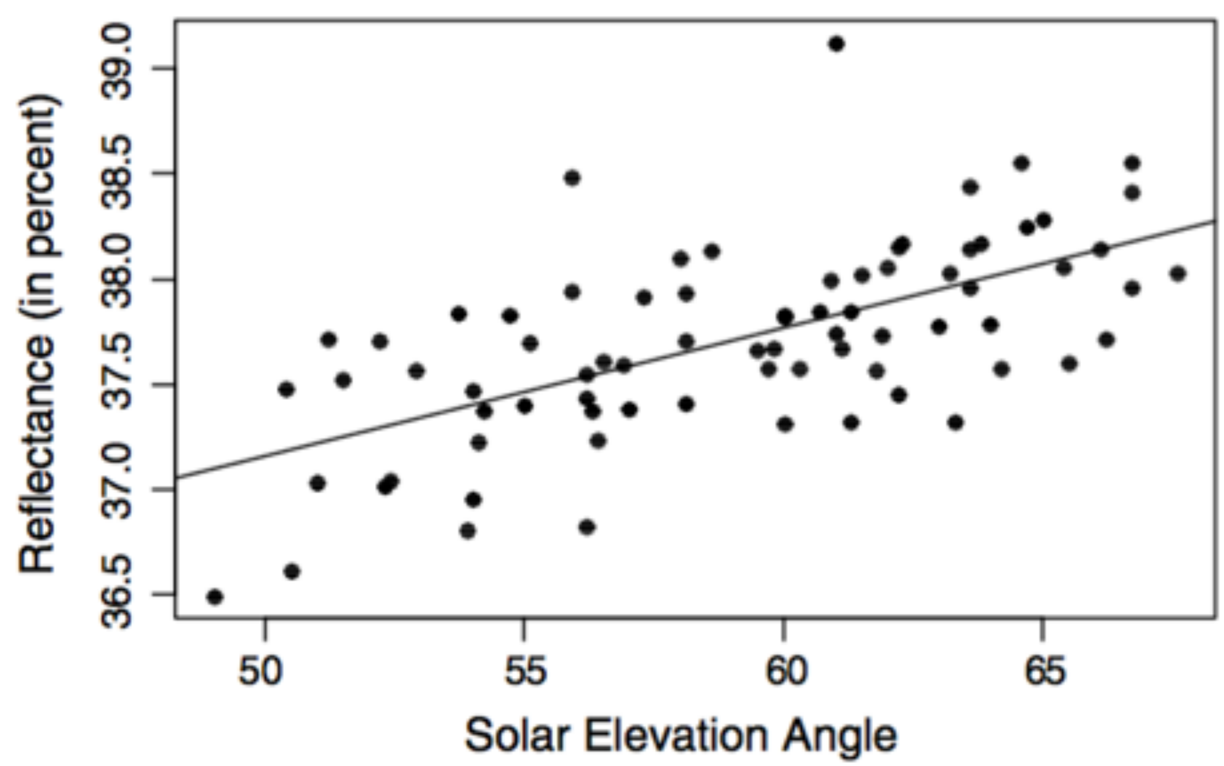


Figure 4.1 A scatter plot of reflectance in Band 2 versus the solar elevation angle for Landsat data discussed in Example 4.1.

Calibration: Simple Linear Regression Model

We usually assume that $E(\varepsilon) = 0$, which means that $E(Y) = \beta_0 + \beta_1 x$, that is, the population average of Y is a linear function of x. This function is called a *regression function*. The line $y = \beta_0 + \beta_1 x$ is a *regression line*. The regression function $\beta_0 + \beta_1 x$ can be regarded as the deterministic part of the model.

We often make the assumption that the error term ε follows a specific distribution, often a normal distribution with the mean zero. Under this assumption, the distribution of Y is also normal and centered at its expected value $E(Y) = \beta_0 + \beta_1 x$. Figure 4.2 illustrates the normal distribution of Y by drawing a normal density curve

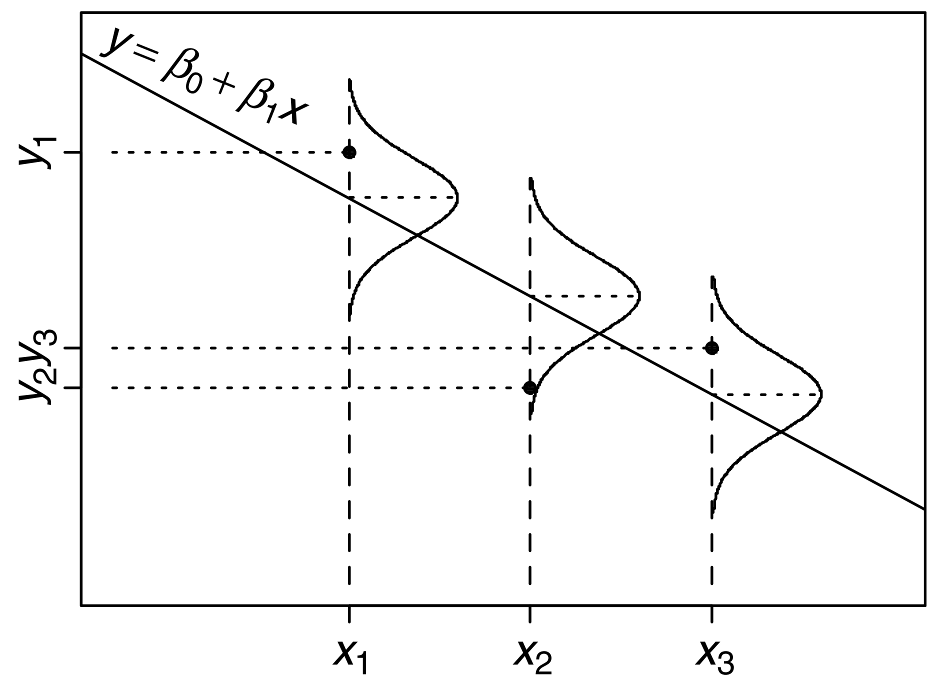


Figure 4.2 Conditional distributions of Y given x are shown here as normal distributions centered at their expected values $E(Y) = \beta_0 + \beta_1 x$, which depend on x in a linear fashion.

Calibration: another example

Example 4.2 An experiment was performed in order to find out how much power is lost when sending signals through optical fiber. This was similar to the experiment described in Example 2.1, except that only one piece of optical fiber was tested this time. The input power of a laser light signal sent from one end of the fiber was set at four different levels: 80, 82, 84, and 86 mW, and the corresponding output power was measured at the other end of the fiber. The purpose was to see how the power loss might depend on the power input. The advantages of using only one piece of fiber are that fewer measurements need to be taken and we do not need to deal with fiber-to-fiber variability. An important disadvantage is that we would not know if our findings apply to other pieces of optical fiber as well.

Five repeated runs were performed at each input power level. The resulting 20 runs were done in a random order. Figure 4.3 shows a scatter plot of the output power (Y) versus the input power (x). The straight line in the plot shows the estimated regression line. For the two cases of x equal to 84 and 86 mW, the line goes almost perfectly through the middle of the group of five data points. The other two cases of x are not as perfect, but

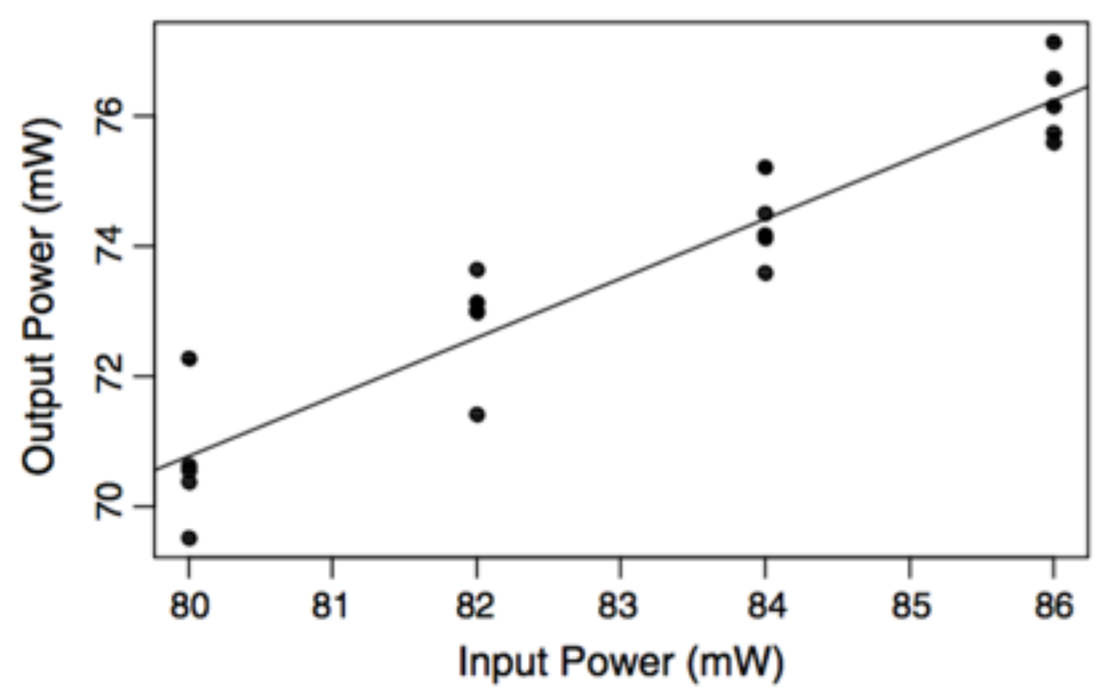


Figure 4.3 The input and output power in a laser light experiment as described in Example 4.2.

Calibration: linear fit

- Least squares estimates

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2.$$

called the *least-squares normal equations*. The solutions to equations (4.6), called the *least-squares estimates*, are given as

$$b_1 = \frac{S_{xy}}{S_{xx}}, \qquad b_0 = \bar{y} - b_1 \bar{x}, \tag{4.7}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of the x and y values and

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2, \qquad S_{xy} = \sum_{i=1}^{n} y_i (x_i - \bar{x}). \tag{4.8}$$

Calibration: linear fit

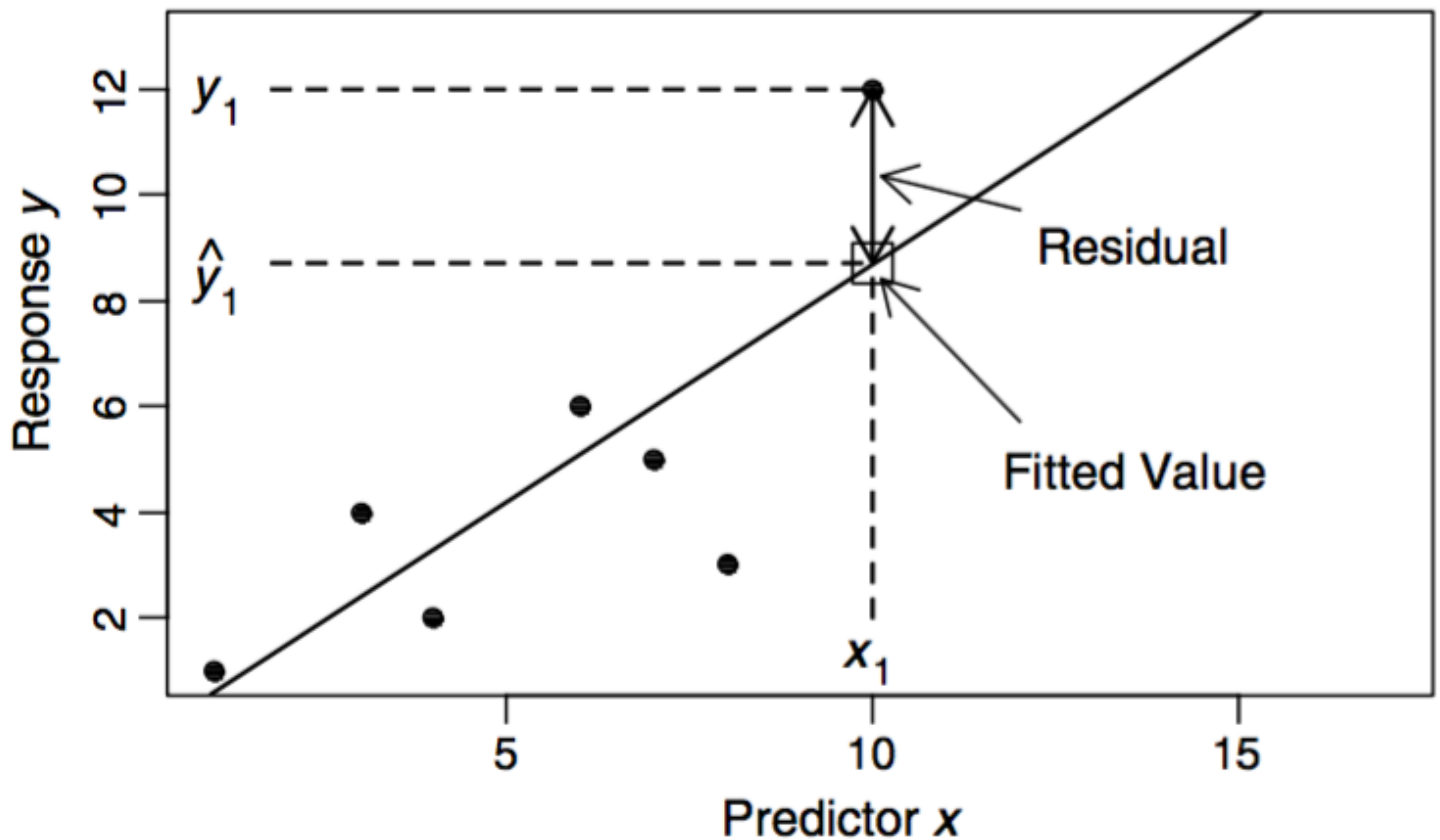


Figure 4.5 The fitted value and the residual for the first observation pair (x_1, y_1).

Calibration: linear fit

calculating the residuals (approximating e_i^*'s), we say that we lose two degrees of freedom for estimation of the regression coefficients, and an unbiased estimator of σ^2 turns out to be

$$\widehat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n} e_i^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2. \tag{4.9}$$

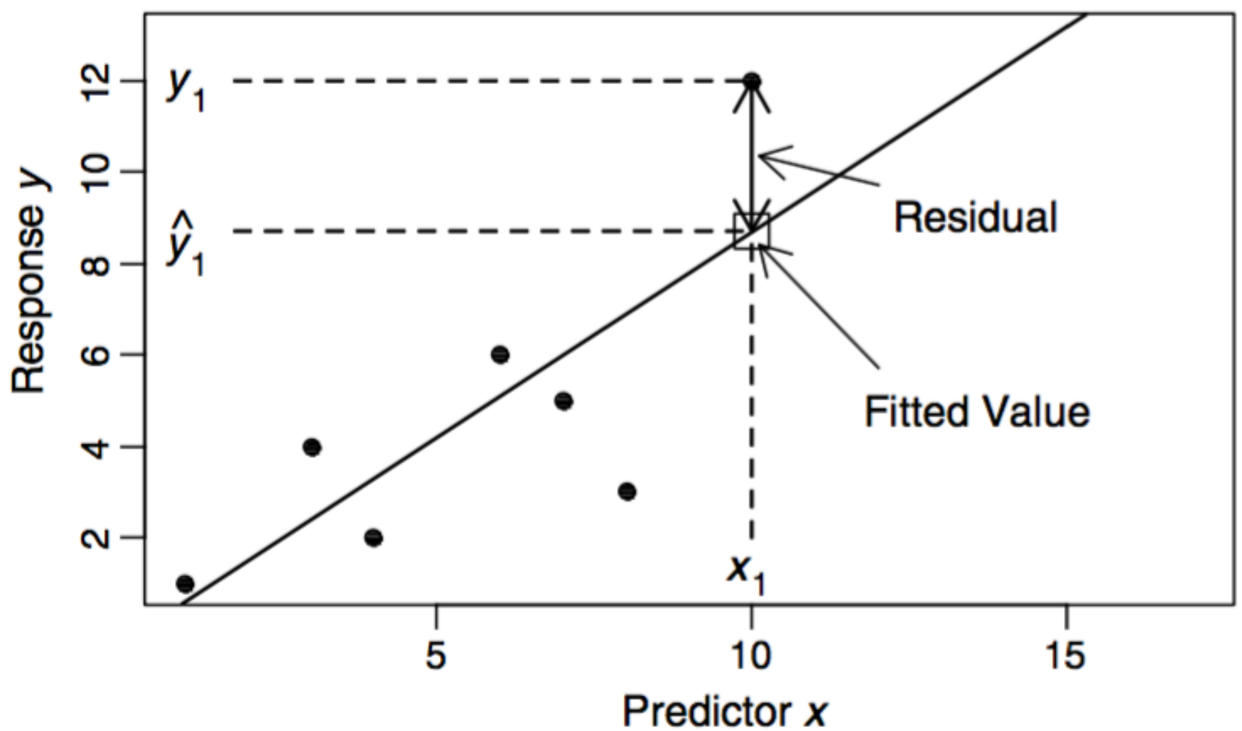


Figure 4.5 The fitted value and the residual for the first observation pair (x_1, y_1).

Calibration: linear fit

An overall variability in all residuals can be measured by the residual sum of squares $\mathrm{SS}_{\mathrm{Res}} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2$, which is the variability of the response values around the regression line. The total variability of the response values (around their mean) can be measured by the total sum of squares $\mathrm{SS}_{\mathrm{Total}} = \sum_{i=1}^{n} (y_i - \bar{y})^2$. In the so-called *analysis of variance* (ANOVA), we can partition the total sum of squares into $\mathrm{SS}_{\mathrm{Res}}$ and the regression sum of squares defined as $\mathrm{SS}_{\mathrm{Regr}} = \sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2$. That is,

$$\mathrm{SS}_{\mathrm{Total}} = \mathrm{SS}_{\mathrm{Regr}} + \mathrm{SS}_{\mathrm{Res}}. \tag{4.10}$$

Calibration: linear fit

- The fraction of the total variability explained by the model is measured by the *coefficient of determination* defined by

$$R^2 = \frac{\mathrm{SS}_{\mathrm{Regr}}}{\mathrm{SS}_{\mathrm{Total}}} = 1 - \frac{\mathrm{SS}_{\mathrm{Res}}}{\mathrm{SS}_{\mathrm{Total}}}. \tag{4.11}$$

We always have $0 \le R^2 \le 1$. The R^2 coefficient may serve as a general indicator by how much a given model can be potentially improved. For example, if $R^2 = 0.7$, we may try to find additional predictors that would explain the remaining 30% of variability. On the other hand, when $R^2 = 0.95$, we know that almost all variability has been explained, and not much more can be explained by finding a better model. At the same time, explaining an additional 3% of variability might be important in some applications.

Calibration: linear fit

Example 4.1 (cont.). As a continuation of the Landsat data example, we find the estimated least-squares regression line as $y = 0.3412 + 0.00061x$. The intercept is the value of y for $x = 0$, but the solar elevation angle never gets close to zero in our data set, and it would not be reasonable to extrapolate our model to such values. Hence, the intercept has no particular interpretation in this case. The slope of 0.00061 means that for each degree of the solar elevation angle, the average reflectance increases by 0.061% of reflectance. The variance σ^2 was estimated as $\widehat{\sigma}^2 = 0.0000132$. It is easier to interpret the estimated standard deviation $\widehat{\sigma} = \sqrt{\widehat{\sigma}^2} = 0.00363$ or 0.363% of reflectance. As an approximate calculation assuming the normal distribution of the error term, we can use the rule of two sigma from Section 2.6 and conclude that 95% of reflectance values in Band 2 will be within $\pm 2 \times 0.363 = \pm 0.726\%$ of reflectance from the regression line $y = 0.3412 + 0.00061x$ drawn in Figure 4.1. This calculation does not take into account the uncertainty in the parameters that were estimated. More precise calculations will be performed in Section 4.2.6.

The sums of squares were calculated as $\mathrm{SS}_{\mathrm{Regr}} = 0.000630$ and $\mathrm{SS}_{\mathrm{Res}} = 0.000977$ for the total of $\mathrm{SS}_{\mathrm{Total}} = 0.001607$. Hence, the fraction of variability explained by the model is $R^2 = 0.392$ or 39.2%. From a statistical point of view, there is still room for model improvement (by using other predictors), although it might be difficult or impossible in practice. □

Calibration: residual analysis

The most important part of Assumption 4.1 introduced in the previous subsection was that $E(\varepsilon_i) = 0$ or equivalently $E(Y_i) = \beta_0 + \beta_1 x_i$, that is, the relationship between the two variables is linear. In Figure 4.3, we were checking this assumption by observing the distribution of points around the estimated regression line for a fixed value x. This was made possible by the presence of repeated observations. In Figure 4.4, we considered a different example that did not have repeats. In that case, we identified

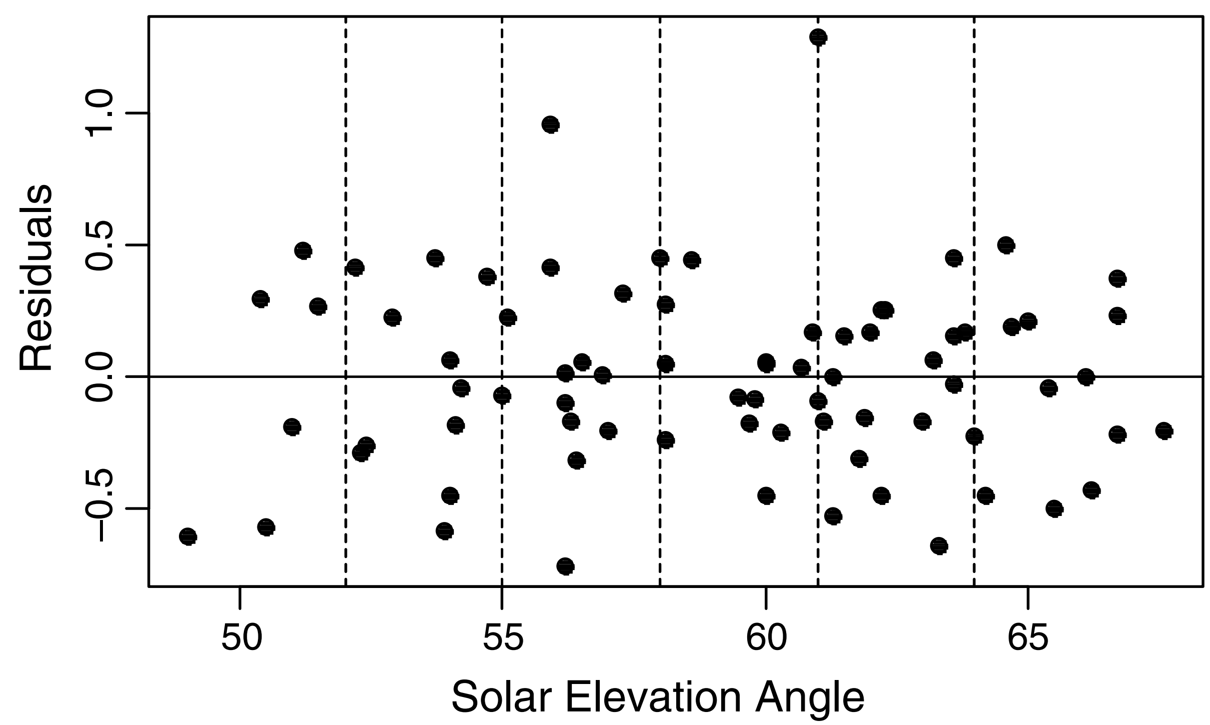


Figure 4.8 The residuals plotted versus the solar elevation angle (the x predictor) for the model fitted in Figure 4.4.

Calibration: residual analysis

The most important part of Assumption 4.1 introduced in the previous subsection was that $E(\varepsilon_i) = 0$ or equivalently $E(Y_i) = \beta_0 + \beta_1 x_i$, that is, the relationship between the two variables is linear. In Figure 4.3, we were checking this assumption by observing the distribution of points around the estimated regression line for a fixed value x. This was made possible by the presence of repeated observations. In Figure 4.4, we considered a different example that did not have repeats. In that case, we identified

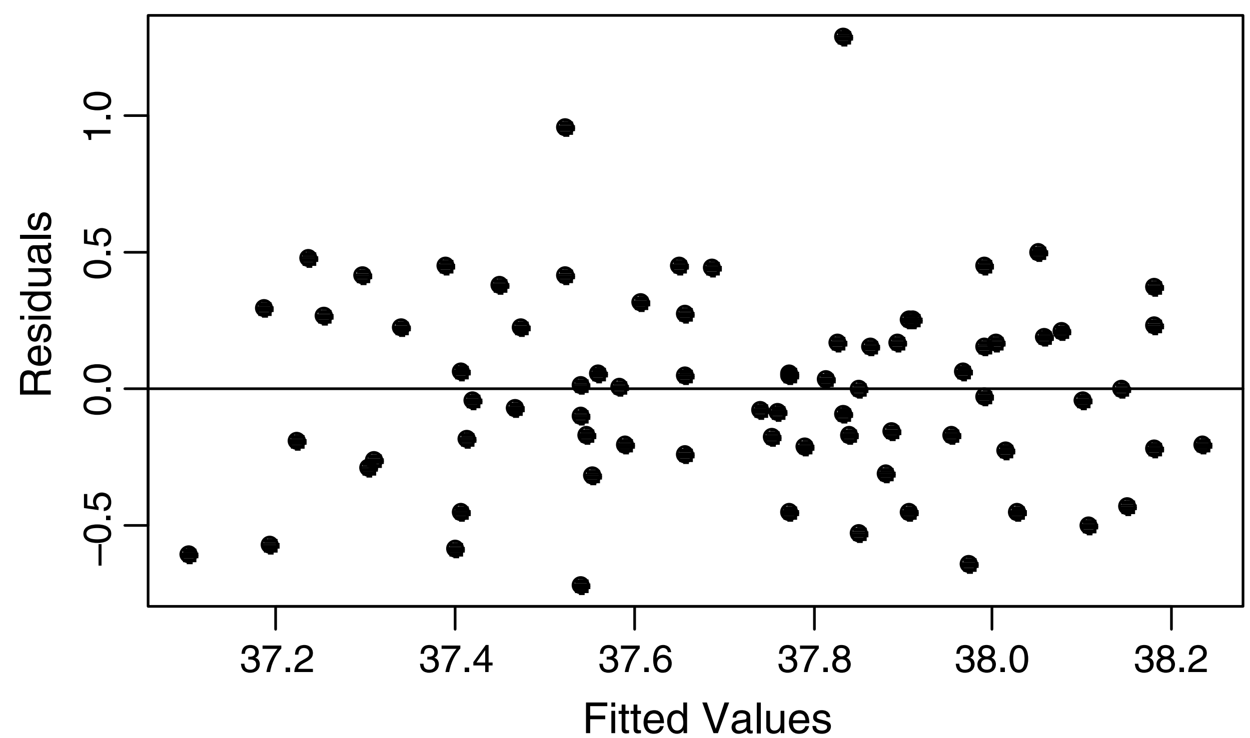
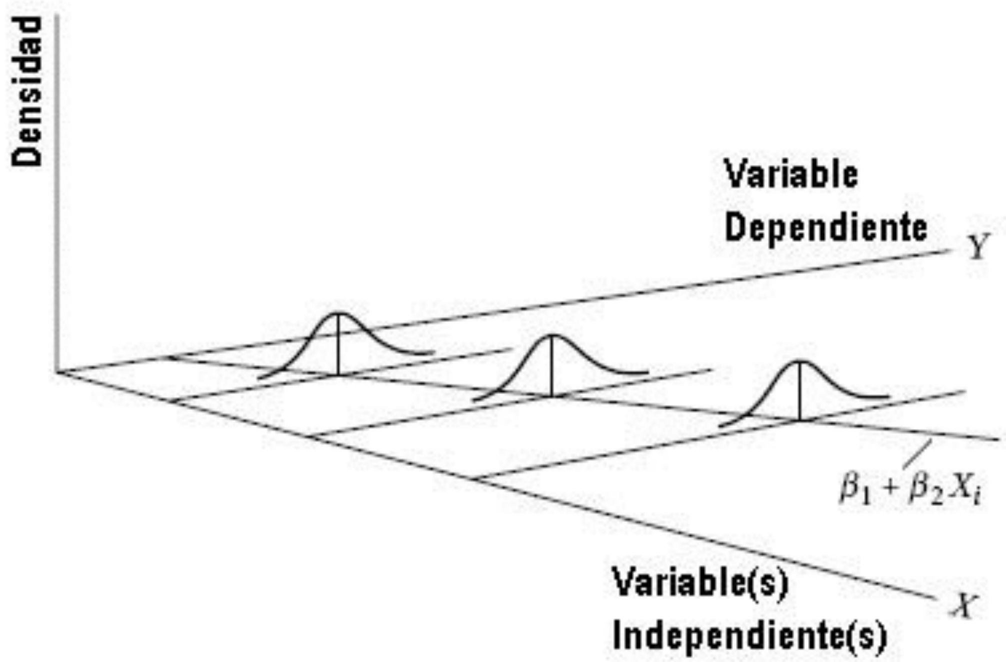


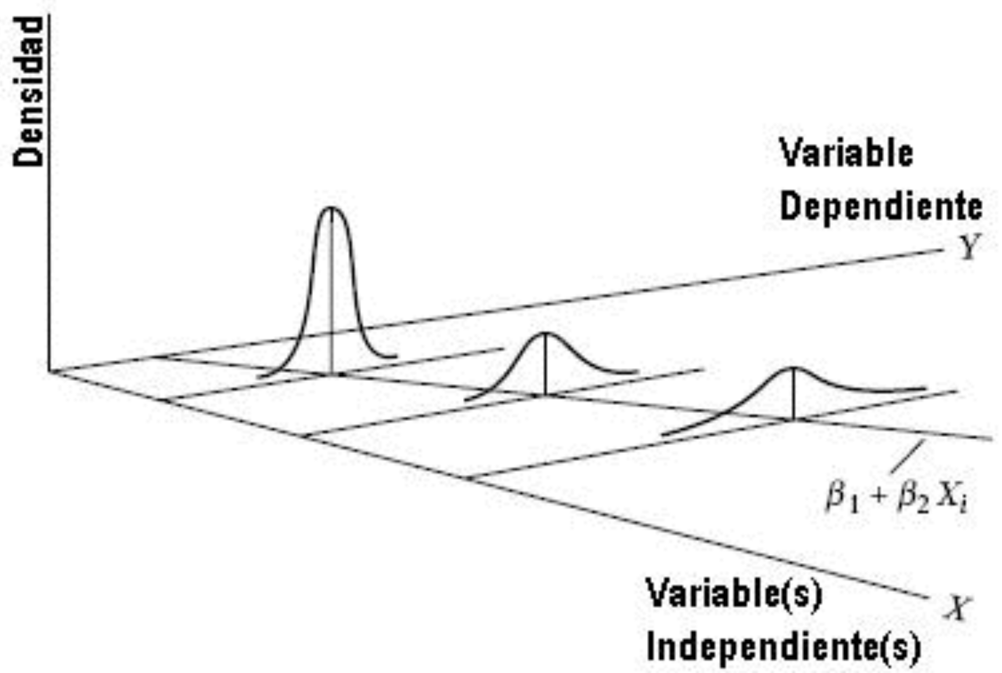
Figure 4.9 The residuals plotted versus fitted values for the model fitted in Figure 4.4.

Calibration. Homoscedastic and Heteroscedastic Data.

Homoscedastic



Heteroscedastic

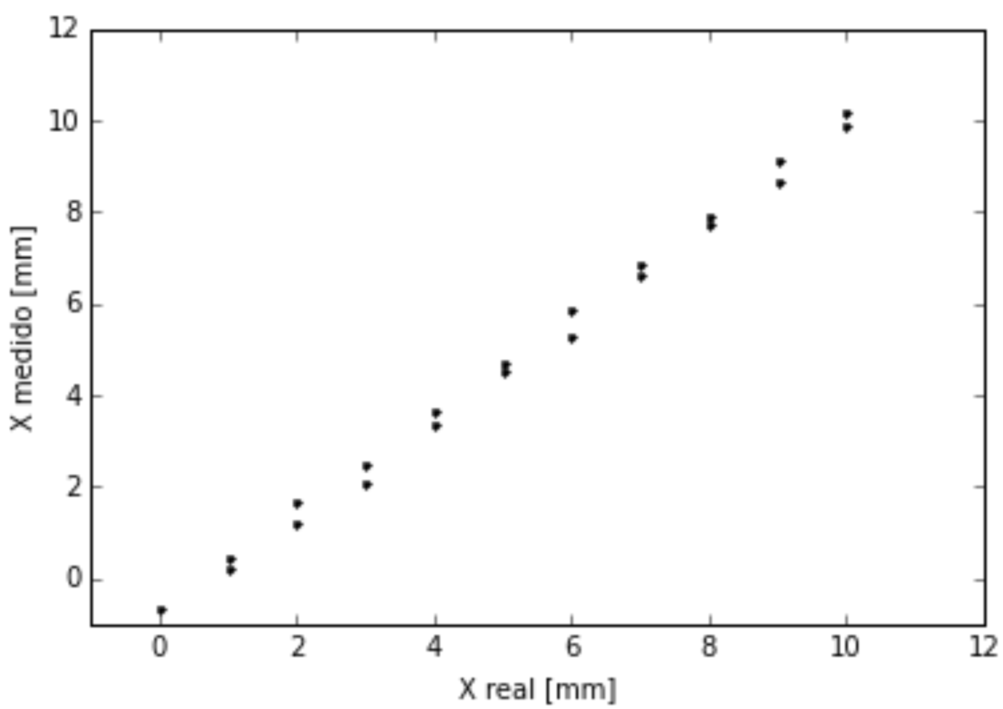


Calibration example

- **Ejemplo.** Un sistema de medida de altura usando pulsos de luz. La tabla muestra los valores reales y los medidos (con error) cuando se incrementa la distancia y cuando se disminuye.

X real(mm)	X medido (Inc.)	X medido (Dism.)
0	-1.12	-0.69
1	0.21	0.42
2	1.18	1.65
3	2.09	2.48
4	3.33	3.62
5	4.50	4.71
6	5.26	5.87
7	6.59	6.86
8	7.73	7.92
9	8.68	9.10
10	9.88	10.20

Calibration example



X real(mm)	X medido (Inc.)	X medido (Dism.)
0	-1.12	-0.69
1	0.21	0.42
2	1.18	1.65
3	2.09	2.48
4	3.33	3.62
5	4.50	4.71
6	5.26	5.87
7	6.59	6.86
8	7.73	7.92
9	8.68	9.10
10	9.88	10.20

Combination of errors

- In general when *f* is a function of *x,y,z,*

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + \cdots$$

Table 3.1 *Propagation of standard uncertainties in combined quantities or functions.*

$f = x + y$ or $f = x - y$	$\sigma_f^2 = \sigma_x^2 + \sigma_y^2$
$f = xy$ or $f = x/y$	$(\sigma_f/f)^2 = (\sigma_x/x)^2 + (\sigma_y/y)^2$
$f = xy^n$ or $f = x/y^n$	$(\sigma_f/f)^2 = (\sigma_x/x)^2 + n^2(\sigma_y/y)^2$
$f = \ln x$	$\sigma_f = \sigma_x/x$
$f = e^x$	$\sigma_f = f\sigma_x$