means that an underlying missing mechanism is not stochastic, and in the book we are considering only random missing mechanisms.

We will also use several inequalities for sums. The first one is the analog of (1.3.33) and it is also called the Cauchy–Schwarz inequality,

$$|\sum_{l=1}^{n} u_l v_l|^2 \leq \sum_{l=1}^{n} u_l^2 \sum_{l=1}^{n} v_l^2. \tag{1.3.43}$$

Here $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are numbers. The Cauchy inequality

$$(u+v)^2 \leq (1+\gamma)u^2 + (1+\gamma^{-1})v^2, \quad \gamma > 0 \tag{1.3.44}$$

is another useful inequality, and here $\gamma^{-1} := 1/\gamma$. The Minkowski inequality for numbers is

$$\Big[\sum_{k=1}^{n} |u_k + v_k|^r\Big]^{1/r} \leq \Big[\sum_{k=1}^{n} |u_k|^r\Big]^{1/r} + \Big[\sum_{k=1}^{n} |v_k|^r\Big]^{1/r}, \quad r \geq 1. \tag{1.3.45}$$

There is also a useful Minkowski inequality for the sum of two random variables,

$$\Big[\mathbb{E}\{|X+Y|^r\}\Big]^{1/r} \leq \Big[\mathbb{E}\{|X|^r\}\Big]^{1/r} + \Big[\mathbb{E}\{|Y|^r\}\Big]^{1/r}, \quad r \geq 1. \tag{1.3.46}$$

The inequality for the sum of numbers

$$|\sum_{k=1}^{n} u_k|^r \leq n^{r-1} \sum_{k=1}^{n} |u_k|^r, \quad r \geq 1 \tag{1.3.47}$$

implies the following inequality for not necessarily independent random variables $X_1, \ldots, X_n$,

$$\mathbb{E}\{|\sum_{k=1}^{n} X_k|^r\} \leq n^{r-1} \sum_{k=1}^{n} \mathbb{E}\{|X_k|^r\}, \quad r \geq 1. \tag{1.3.48}$$

If additionally the random variables $X_1, \ldots, X_n$ are independent and zero mean, that is $\mathbb{E}\{X_k\} = 0$ for $k = 1, \ldots, n$, then

$$\mathbb{E}\{|\sum_{k=1}^{n} X_k|^r\} \leq C(r)\Big[\sum_{k=1}^{n} \mathbb{E}\{|X_k|^r\} + \Big(\sum_{k=1}^{n} \mathbb{E}\{X_k^2\}\Big)^{r/2}\Big], \quad r \geq 2, \tag{1.3.49}$$

and

$$\mathbb{E}\{|\sum_{k=1}^{n} X_k|^r\} \leq C(r) n^{\frac{r}{2}-1} \sum_{k=1}^{n} \mathbb{E}\{|X_k|^r\}, \quad r \geq 2, \tag{1.3.50}$$

where $C(r)$ is a positive and finite function in r.

There are four corner (test) functions that are used throughout the book. They are the *Uniform*, the *Normal*, the *Bimodal* and the *Strata*. These functions are defined in Section 2.1 and Figure 2.1 exhibits them.

1.4 Software

R is a commonly used free statistics software. R allows you to carry out statistical analyses in an interactive mode, as well as allowing simple programming. You can find detailed installation instructions in the R Installation and Administration manual on CRAN (www.r-project.org, http://cran.r-project.org). After installing R (or if it was already installed), you need to choose your working directory for the book's software package. For instance, for

MAC you can use ″/Users/Me/Book2″. Download to this directory file book2.r from the author's web site www.utdallas.edu/∼efrom. The web site also contains relevant information about the book. By downloading the package, the user agrees to consider it as a "black-box" and employ it for educational purposes only.

Now you need to start the R on your computer. This should bring up a new window, which is the R console. In the R console you will see:

>

This is the R prompt. Now, only once while you are working with the book, you need to install several standard R packages. Type the following command in the R console to install required packages

> install.packages(″MASS″)
> install.packages(″mvtnorm″)
> install.packages(″survival″)
> install.packages(″scatterplot3d″)

These packages are installed just once, you do not need to repeat this step when you start your next R session.

Next you need to source (install) the book's package, and you do this with R operating in the chosen working directory. To do this, first type

> getwd()

This R command will allow you to see a working directory in which R is currently operating. If it is not your chosen working directory ″/Users/Me/Book2″, type

> setwd(″/Users/Me/Book2″)

and then R will operate in the chosen directory for the book. Then type

> source(″book2.r″)

and the book's R software will be installed in the chosen directory and you are ready to use it. This sourcing should be done every time when you start a new R session.

For the novice, it is important to stress that no knowledge of R is needed to use the package and repeat/modify figures. Nonetheless, a bit of information may be useful to understand the semantics of calling a figure. What you are typing is the name of an R function and then in its parentheses you assign values to its arguments. For instance,

> ch1(fig=1, n=300)

is the call to R function ch1 that will be run with arguments fig=1 and n=300, and all other arguments of the function will be equal to their default values (here a=0 and b=0.7) indicated in the caption of the corresponding Figure 1.1. Note that ch1 indicates that the figure is from Chapter 1 while its argument fig=1 indicates that it is the first figure in Chapter 1. In other words, ch1(fig=1,n=300) runs Figure 1.1 with the sample size $n = 300$. In R, arguments of a function may be either scalars (for instance, a=5 implies that a function will use value 5 for argument a), or vectors (for instance, vec = c(2,4) implies that a function will use vector vec with the first element equal to 2 and the second element equal to 4, and note that c() is a special R function called "combine" that creates vectors), or a string (say den= ″normal″ implies that a function will use name *normal* for its argument den). R functions are smart in terms of using shorter names for its arguments, for instance both fig=5, fi=5 and f=5 will be correctly recognized and imply the same value 5 for the argument fig unless there is a confusion with other arguments.

In general, to repeat Figure k.j, which is jth figure in Chapter k, and to use the default values of its arguments outlined in square brackets of the caption of Figure k.j, type

> chk(f=j)

If you want to use new values for two arguments, say n=130 and a=5, type

> chk(f=j, n=130, a=5)

Note that values of other arguments will be equal to default ones. Further, recall that typing f=j and fig=j implies the same outcome.

To finish the R session, type

```
> q()
```

Finally a short comment about useful captions and figures. Figure 2.1 exhibits four corner (test) functions: the Uniform, the Normal, the Bimodal and the Strata. Figure 2.2 shows first eight elements of the cosine basis. The caption to Figure 2.3 explains how to create a custom-made corner function. Figure 2.4 shows the sequence of curves and their colors on the monitor. The caption to Figure 2.5 explains arguments used by the E-estimator.

1.5 Inside the Book

There are always two parts in nonparametric estimation based on missing and modified data. The first one is how to construct a feasible nonparametric estimator for the case of direct data, and the second is how to address the issue of missing and modified data. The book takes the following approach used for all considered problems. The same orthogonal series estimator, called an E-estimator and defined in Section 2.2 for the case of direct data, is also used without any "nonparametric" modification for indirectly observed data. The attractiveness of this approach is that the book is less about how to estimate a nonparametric function (curve) and more about how the E-estimator may be used in a variety of settings with missing and modified data. This methodology is supported by the asymptotic theory which states that for considered settings no other estimator can outperform the E-estimation methodology. As a result, other known methods for dealing with missing and modified data are not considered, but a relevant literature can be found at the end of each chapter in the section Notes. Further, the interested reader can find a book-length treatment of nonparametric series estimators for direct data as well as a discussion of other popular nonparametric estimators in the author's book Efromovich (1999a), *Nonparametric Curve Estimation: Methods, Theory and Applications*.

Let us explain why and how a single (and we may say universal) E-estimator may be used for the totality of considered problems. Suppose that our aim is to estimate a bounded function $m(x)$ over an interval $[0,1]$. Then, according to (1.3.36) and (1.3.37) it can be written as a Fourier series

$$m(x) = \sum_{j=0}^{\infty} \theta_j \varphi_j(x), \tag{1.5.1}$$

where $\varphi_j(x)$ are elements of the cosine basis and

$$\theta_j := \int_0^1 m(x)\varphi_j(x)dx \tag{1.5.2}$$

are Fourier coefficients of $m(x)$.

The main step of E-estimation is to propose a sample mean estimator $\hat{\theta}_j$ of Fourier coefficients. If this step is done, then there is a straightforward and universal (it depends on neither the estimand, nor the setting, nor the data) procedure of plugging $\hat{\theta}_j$ in (1.5.1). This procedure, called the E-estimator, is defined in Section 2.2. As a result, a more complicated nonparametric problem is converted into a simpler parametric problem of proposing a sample mean estimator for Fourier coefficients.

Let us present several examples that shed light on construction of sample mean estimators. Consider the problem of estimation of the probability density $f^X(x)$ of a random variable X supported on $[0,1]$. Suppose that a sample $X_1, \ldots, X_n$ from X is available. Then the sample mean estimator of Fourier coefficient $\theta_j := \int_0^1 f^X(x)\varphi_j(x)dx = \mathbb{E}\{\varphi_j(X)\}$ is

$$\hat{\theta}_j := n^{-1} \sum_{l=1}^{n} \varphi_j(X_l). \tag{1.5.3}$$