**MR2284824 (2010e:60002)** 60-02 60-06 60B11 60B12 60F17 62F03 62G30

**del Barrio, Eustasio** (E-VALD-OR); **Deheuvels, Paul** (F-PARIS6-AS);
**van de Geer, Sara** [**van de Geer, Sara A.**] (CH-ETHZ-S)

★**Lectures on empirical processes.**
Theory and statistical applications.
With a preface by Juan A. Cuesta Albertos and Carlos Matrán.
EMS Series of Lectures in Mathematics.

*European Mathematical Society (EMS), Zürich,* 2007. $x$+254 *pp.*
*ISBN* 978-3-03719-027-2

**Contents:**

Eustasio del Barrio, "Empirical and quantile processes in the asymptotic theory of goodness-of-fit tests", 1–92. MR2347003

Paul Deheuvels, "Topics on empirical processes", 93–190. MR2347004

Sara van de Geer, "Oracle inequalities and regularization", 191–252. MR2347005.

This book consists of the three expository papers, as well as a preface by Juan A. Cuesta Albertos and Carlos Matrán, from whom we learn that the idea of this book was born in Laredo (Spain), where the two of them organized an EMS (European Mathematical Society) Summer School on Empirical Processes in September 2004. The aim of this School was to disseminate the power of the theory of empirical processes via courses that were also to provide a vision of the present perspectives of its development, as well as that of its more significant statistical applications.

The three respective papers of the three authors in this book are expanded versions of their lectures given at the School. The preface concludes that "this book constitutes an excellent overview of a broad scope of the theory of empirical processes that will be an invaluable aid to those interested in an advanced and well-documented approach to the topics selected by the authors". By and large, this reviewer agrees with this conclusion and will mainly call attention to, and comment on, further references only when it will be deemed to be desirable for the sake of additional documentation and clarifications. Unfortunately this will occur far too many times in the case of del Barrio's paper in this collection. Hence also the unusual length of this review.

The first paper, "Empirical and quantile processes in the asymptotic theory of goodness-of-fit tests", by Eustasio del Barrio, consists of six sections (and will hereafter be referred to as [del B]):

1. Introduction.
2. Testing fit to a fixed distribution.
3. Some tools from empirical process theory.
4. Testing fit to a family of distributions.
5. Tests based on Wasserstein distance.
6. Asymptotics for $L_2$ functionals of the quantile process.

Section 1 consists of a short historical overview that includes pointers to some relevant literature. The author pays special attention to stressing the parallel evolution of the theory of empirical processes and the asymptotic theory of goodness-of-fit tests. This inevitably leads to dealing with invariance principles that constitute some of the main tools for studying the asymptotic theory of empirical and quantile processes.

Section 2 provides an initial scenario via discussing Pearson's asymptotic $\chi^2$ test of fit, the asymptotic distribution of Cramér-von Mises type statistics, and also that of

Kolmogorov-Smirnov type statistics. The latter is viewed via approximations (strong and weak invariance principles) that yield necessary and sufficient conditions for the weak convergence of the empirical process to a Brownian bridge in weighted uniform metric à la D. M. Chibisov [Trudy Mat. Inst. Steklov. **71** (1964), 104–112; MR0172324] ([11] in [del B]) and N. E. O'Reilly [Ann. Probability **2** (1974), 642–651; MR0383486] ([70] in [del B]), and also via approximations as in [M. Csörgő et al., Ann. Probab. **14** (1986), no. 1, 31–85; MR0815960] (hereafter referred to as [CCHM86]) ([15] in [del B]) that also yield the best possible characterization of convergence in distribution of the weighted sup-functional of the uniform empirical process (cf. Theorems 2.5 and 3.31 in [del B]). On the other hand, an appropriate Hilbert space setting combined with the theory of probability in Banach spaces is shown to provide a natural setting for concluding optimal results for quadratic statistics, like for example that of the convergence in distribution of the weighted $L_2$-functional of the empirical process as in Theorem 2.4 in [del B]. I note in passing though that, using weighted approximations in $L_p$, $0 < p \leq 2$, of the uniform empirical process by Brownian bridges in probability, in combination with a dichotomy theorem for a Brownian bridge of L. A. Shepp [Ann. Math. Statist. **37** (1966), 321–354; MR0190999], G. R. Shorack and J. A. Wellner [*Empirical processes with applications to statistics*, Wiley, New York, 1986; MR0838963] ([88] in [del B]) also concluded the latter theorem. Moreover, M. Csörgő, L. Horváth and Q. M. Shao [Stochastic Process. Appl. **45** (1993), no. 2, 283–294; MR1208874] found a necessary and sufficient condition for the weak convergence of the uniform empirical and quantile processes to a Brownian bridge in weighted $L_p$-distances, $0 < p < \infty$. Naturally, with $p = 2$, the latter condition coincides with that of Theorem 2.4 in [del B]).

In Section 3 the author presents some of the most important tools from empirical and partial sum processes theory. He deals with some inequalities in Subsection 3.1 and with the central limit theorem (both on the real line and in separable Banach spaces) in Subsection 3.2. The introductory part of Subsection 3.3 on strong approximations of partial sums and uniform empirical and quantile processes presents well-known coupling inequalities for these processes that were achieved via methodologies of the Hungarian construction school. Up to this point the author provides a fair amount of pointers to the relevant references. However, his style changes dramatically in Subsection 3.3.1, Weighted approximations of empirical and quantile processes, where the absence of relevant citations has warranted most of the detailed criticism in this lengthy review. Some slips of the pen and mathematical oversights will be pointed out as well.

I will now detail my comments on Subsection 3.3, Strong approximations. Let $\{\alpha_n(t),\ 0 \leq t \leq 1\}$ and $\{u_n(t),\ 0 \leq t \leq 1\}$ respectively be the uniform empirical and quantile processes. As regards Theorem 3.18 for $\alpha_n(\cdot)$ and Theorem 3.19 for $u_n(\cdot)$ in [del B], I note that the latter for $u_n(\cdot)$ was proved in [CCHM86 (cf. Theorem 1.1)], but Theorem 3.18 for $\alpha_n(\cdot)$ was only conjectured there (cf. Remark 2.1 of [CCHM86])—del Barrio states that both results come from the latter paper. It was, in turn, established by D. M. Mason and W. R. van Zwet [Ann. Probab. **15** (1987), no. 3, 871–884; MR0893903 (cf. Theorem 1)]. Based on their Theorem 1.1 (Theorem 3.19 in [del B]), M. Csörgő et al. [CCHM86] proved their Theorem 2.1 for the uniform quantile process $u_n(\cdot)$ (cf. Theorem 3.23 in [del B]) and, based on their Theorem 1 (Theorem 3.18 in [del B]), Mason and van Zwet [op. cit.] concluded their Theorem 2 for the uniform empirical process $\alpha_n(\cdot)$ that is stated as Theorem 3.20 in [del B]. The Mason and van Zwet paper is not cited in [del B]. The way their just-mentioned Theorem 2 is spelled out as Theorem 3.20 in [del B] coincides with Theorem 4.3.1 in the book by M. Csörgő and Horváth [*Weighted approximations in probability and statistics*, Wiley, Chichester, 1993; MR1215046] (hereafter referred to as [CH93]) ([17] in [del B]). The proof of Theorem 3.20 in [del B] also coincides with that of Theorem 4.3.1 in the latter book (cf. pp. 213–

217, where the two corollaries that are respectively restated as Corollaries 3.21 and 3.22 in [del B] are also proved), but del Barrio makes no mention of it being borrowed from there. It (the proof at hand) goes along the lines of the original proof of the first such theorem, i.e., that of Theorem 2.1 for the uniform quantile process $u_n(\cdot)$ of [CCHM86]. This proof is reproduced in [CH93] as the first proof of Theorem 4.2.1 there. The latter theorem is restated as Theorem 3.23 in [del B] as the corresponding result for the uniform quantile process $u_n(\cdot)$, and is preceded by a statement by del Barrio that its proof is similar (no wonder!) to the proof of Theorem 3.20 and hence (*sic*) omitted. M. Csörgő and L. Horváth [Statist. Probab. Lett. **4** (1986), no. 6, 275–280; MR0858316] and Mason ["A note on weighted approximations to the uniform empirical and quantile processes", tech. report, Univ. München, Munich, 1986; also in *Sums, trimmed sums and extremes*, 269–283, Birkhäuser Boston, Boston, MA, 1991; MR1117273] gave mathematically equivalent short and simple proofs of Theorem 2.1 of [CCHM86] based directly on the Komlós, Major and Tusnády (KMT) [J. Komlós, P. Major and G. Tusnády, Z. Wahrsch. Verw. Gebiete **34** (1976), no. 1, 33–58; MR0402883] ([56] in [del B]) approximation of partial sums (cf. Theorem 3.14 in [del B]). The second proof of Theorem 4.2.1 in [CH93] (Theorem 3.23 in [del B]) goes along these lines [cf. also D. M. Mason, op. cit.]. In connection with Theorem 3.24 in [del B], in which a Brownian bridge $B$ should be replaced by a sequence of Brownian bridges, one should note that it is due to R. Pyke and G. R. Shorack [Ann. Math. Statist. **39** (1968), 755–771; MR0226770] (cf. also Lemma 2.4 in [M. Csörgő, S. Csörgő and L. Horváth, *An asymptotic theory for empirical reliability and concentration processes*, Lecture Notes in Statist., 33, Springer, Berlin, 1986; MR0856407]). Furthermore, the statement of Lemma 3.27 in [del B] is a special case of Lemma 2.4 in the just-mentioned monograph of M. Csörgő, S. Csörgő and Horváth, as it is summarized there right after the proof of this lemma on page 27. Lemma 3.25 in [del B] is borrowed from [CH93 (cf. Lemma 1.1 and its proof on pp. 178–179)], and so also is the class of functions $FC_0$ together with the two integrals denoted by $E(q,c)$ and $I(q,c)$ respectively. Connecting the finiteness of these two integrals played an important role in [CCHM86] (cf. Proposition 3.1, which is Lemma 1.3 on page 180 in [CH93]). Lemma 3.26 and its proof in [del B] are lifted from page 180 of [CH93]. The integrals $E(q,c)$ and $I(q,c)$ provide necessary and sufficient conditions for the characterization of the tail behaviour of a Wiener process. For details we refer to Theorem 4.1.1 and its proof in [CH93], pp. 181–188. Theorem 3.28 and its unfinished proof in [del B] are correspondingly lifted from pp. 181–top of 183 of [CH93], the proof part with a few slips of the pen. This is the only time that [CH93] is mentioned in [del B], and only fleetingly. Furthermore, as is also noted on page 249 of [CH93] but not by del Barrio, this theorem in its present form, i.e., Theorem 4.1.1 in [CH93], is due to M. Csörgő, Q. M. Shao and B. Szyszkowicz [Studia Sci. Math. Hungar. **26** (1991), no. 2-3, 239–259; MR1180493]. For a discussion of this theorem, as well as that of the Chibisov (1964)–O'Reilly (1974) theorem in this context, we refer to pp. 249–250 of [CH93]. In (iii) of Corollary 3.29 in [del B] (cf. Corollary 4.1.1 in [CH93]), the indicated limit as $t \to 0$, or $t \to 1$, should be equal to $\infty$ instead of 1. The presentation and proof of the Chibisov-O'Reilly theorem as that of Theorem 4.3.2 in [CH93] is based on Theorem 4.2.1 and Remark 4.2.1 of [CCHM86]. Theorem 3.30 (the Chibisov-O'Reilly theorem) and its unfinished proof in [del B] are borrowed from pp. 217–220 of [CH93]. Theorem 4.2.3 of [CCHM86], an extension of the Chibisov-O'Reilly theorem, is one of the highlights of that paper, and it is also properly mentioned in [del B], and quoted as well as Theorem 2.5 in del Barrio's Section 2. As it is restated by him as Theorem 3.31 at the end of his Section 3.3.1, it is borrowed from [CH93] (cf. Theorem 4.3.3 on page 220). In his exposition del Barrio does not say that these two theorems at hand are the same. It is regrettable that the material of Subsection 3.3.1 in del Barrio's paper is presented in the way illustrated above. A more conscientious approach would have saved this reviewer the embarrassment of having to call attention to appropriate references for many of the primary and subsequent items of information that are freely used in his presentation of these important results.

In Section 4, del Barrio also deals with the empirical process with estimated parameters à la J. Durbin [Ann. Statist. **1** (1973), 279–290; MR0359131] ([37] in [del B]), suggesting Durbin's results via strong approximations as in Theorem 4.1. Under appropriate conditions the latter concludes the optimal almost sure KMT [J. Komlós, P. Major and G. Tusnády, Z. Wahrsch. Verw. Gebiete **32** (1975), 111–131; MR0375412; also op. cit., 1976] ([55] and [56] in [del B]) type (cf. (3.10) in [del B]) approximation of what del Barrio calls the uniform estimated empirical process. It, however, does not yield Durbin's results in [op. cit.] as easy corollaries, as claimed by del Barrio on page 39 of his paper. Indeed, his example on pages 38–39 in the Gaussian case ends up being a counterexample to this claim. Nevertheless, it is true that methods of strong approximations can be, and have already been, used in studying the empirical process when parameters are estimated (cf. e.g. [M. D. Burke et al., Ann. Probab. **7** (1979), no. 5, 790–810; MR0542130] and Section 5.7 in the book [M. Csörgő and P. Révész, *Strong approximations in probability and statistics*, Academic Press, New York, 1981; MR0666546]). In the just-mentioned paper and book passage, the authors studied the problem of obtaining asymptotic-in-probability and almost sure representations, in terms of Gaussian processes, of the empirical process when parameters are estimated (cf. Remark 5.7.1 of the just-mentioned Csörgő-Révész book for relating these representations to the weak convergence results of Durbin [op. cit.]). The Csörgő-Révész book, however, is not cited by del Barrio. On page 39 del Barrio notes that a survey of results connected to Theorem 4.1 as well as a simple derivation of it based on the Skorokhod embedding can be found in the Shorack-Wellner book [op. cit.]. While it is true that in the latter book there is an excellent survey of tests of fit that are based on the empirical process when parameters are estimated, as well as that of stochastic integrals via integration by parts, there is no simple derivation of Theorem 4.1 with the KMT (1975, 1976) type almost sure rate of convergence based on Skorokhod embedding, for that of course is impossible.

Sections 5 and 6, dealing with asymptotics of quadratic forms of the quantile process via $L_2$ and weighted $L_2$-Wasserstein distances, are mainly based on the two insightful del Barrio et al. papers [E. del Barrio et al., Ann. Statist. **27** (1999), no. 4, 1230–1239; MR1740113; E. del Barrio, E. Giné and F. Utzet, Bernoulli **11** (2005), no. 1, 131–189; MR2121458] ([26], [28] in [del B]). Parts of those papers and of Sections 5 and 6 overlap with some parts of the just-mentioned Csörgő-Révész book (cf., e.g., Lemma 5.5.5, Theorem 5.5.3, and Sections 5.8 and 5.9) and Section 7.1 of [M. Csörgő, *Quantile processes with statistical applications*, SIAM, Philadelphia, PA, 1983; MR0745130]). For example, the just-mentioned parts of these two research books, together with further related references therein, already deal with the Shapiro-Wilk-Francia-de Wet-Venter approach to testing for normality and their extensions to testing for other composite hypotheses as well, in the context of weighted approximations of the general quantile process. To say the least, mentioning this would have strengthened the historical context of the course given by del Barrio. Moreover, calling attention to Section 6.4, $L_p$-functionals of weighted quantile processes, of [CH93], where $1 \leq p < \infty$, and its discussion in Section 6.6 of the same book, would have provided an appropriate backdrop for studying and appreciating the related comprehensive results in $L_2$ of the del Barrio et al. papers cited above. The same holds true for mentioning also Section 5.4, $L_p$-functionals of weighted uniform quantile processes, of [CH93] (cf., e.g., Theorem 5.4.3 in the latter book vis-à-vis Theorem 6.16 in [del B]). I also note in passing that not only condition (6.7), but also the idea of the limited Taylor expansion as in (6.10) in [del B], comes from [M. Csörgő and P. Révész, Ann. Statist. **6** (1978), no. 4, 882–894; MR0501290] ([18] in [del B]). Also, when writing about quantile processes along these lines, one should at least mention the pathbreaking contributions of R. R. Bahadur [Ann. Math. Statist. **37** (1966), 577–580; MR0189095] and J. Kiefer [Ann. Math. Statist. **38** (1967), 1323–1342; MR0217844; in *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)*, 299–319, Cambridge Univ. Press, London, 1970; MR0277071] that initiated the modern theory of sample quantiles and quantile processes. All this notwithstanding, Sections 5 and 6 deal with impressive contributions of del Barrio et al.'s two papers to the asymptotics for $L_2$-functionals of quantile processes.

Deheuvels' paper concentrates mainly on strong functional limit laws generated by partial sums and empirical processes via strong approximations, and studies such laws on their own as well, beyond the latter approach. To further the reader's interest, exercises are included in the text, along with references. The first section provides an excellent introduction of some of the preliminary tools and results.

Section 2, Fluctuation of partial sums, deals with some classical large deviations theory, martingale inequalities, functional large deviations and limit theorems for the Wiener process and for processes with independent increments, the KMT (1975, 1976), Major [P. Major, Z. Wahrsch. Verw. Gebiete **35** (1976), no. 3, 213–220; MR0415743], Strassen [V. Strassen, Z. Wahrsch. Verw. Gebiete **3** (1964), 211–226 (1964); MR0175194] and Major [Z. Wahrsch. Verw. Gebiete **35** (1976), no. 3, 221–229; MR0415744; Ann. Probab. **7** (1979), no. 1, 55–61; MR0515812] approximations of partial sums of i.i.d. random variables, the classical Erdős-Rényi [P. Erdős and A. Rényi, J. Analyse Math. **23** (1970), 103–111; MR0272026] theorem that plays an important role in the theory of strong invariance principles in providing a lower bound to the best possible rates of strong approximations of partial sums by a Wiener process, and the functional Erdős-Rényi law of Deheuvels [Studia Sci. Math. Hungar. **26** (1991), no. 2-3, 261–295; MR1180494].

Section 3, Empirical processes, studies some classical distribution theory of the uniform empirical distribution and quantile functions and processes, martingale inequalities for empirical processes and Brownian bridges, relations between empirical and Poisson processes, the KMT (1975) inequality for the approximation of the uniform empirical process by Brownian bridges, and the similarly optimal CsR [M. Csörgő and P. Révész, op. cit., 1978] variant for the approximation of the uniform quantile process. I note in passing that the latter is used to give a glimpse, via quantiles, of the preceding KMT (1975) approximation by the reviewer [Acta Sci. Math. (Szeged) **73** (2007), no. 1-2, 349–366; MR2339870]. On pages 350 and 351 of the latter paper there is also a short summary of the history of the proof of this celebrated KMT (1975) approximation inequality. Back to Section 3 of Deheuvels' paper at hand, in Subsection 3.5 it also deals with the KMT (1975) approximation of the uniform empirical process by a Kiefer process, viewing it also via [N. Castelle and F. Laurent-Bonvalot, Ann. Inst. H. Poincaré Probab. Statist. **34** (1998), no. 4, 425–480; MR1632841] and [N. Castelle, Canad. J. Math. **54** (2002), no. 3, 533–553; MR1900762]. Combining the latter KMT (1975) approximation with the Bahadur-Kiefer representation (cf., e.g., Theorem 3.21 in the paper under review), one arrives at a strong approximation of the uniform quantile process by a Kiefer process (cf. (3.54)) that was proved to be optimal by Deheuvels [J. Theoret. Probab. **11** (1998), no. 4, 997–1018; MR1660907] (cf. (3.55)). The work of M. Csörgő et al. [CCHM86], M. Csörgő and L. Horváth [Statist. Probab. Lett. **4** (1986), no. 6, 275–280; MR0858316] and Mason and van Zwet [op. cit.] is also explored in this section. I note in passing that with references to [CCHM86] and [CH93], in Theorem

3.11 one should have $0 \leq \nu < 1/4$, $0 \leq \tau < 1/2$. As it is stated, with $0 \leq \nu < 1/2$, $0 \leq \tau < 1/4$, (3.56) of the latter theorem follows from Theorem 1 of [D. M. Mason and W. R. van Zwet, op. cit.] (cf. Theorem 3.7 in the paper at hand). Some results on the standardized empirical distribution, like, e.g., its asymptotic distribution as in [F. Eicker, Ann. Statist. **7** (1979), no. 1, 116–138; MR0515688] and [D. Jaeschke, Ann. Statist. **7** (1979), no. 1, 108–115; MR0515687] and its strong laws as in [E. Csáki, in *Nonparametric statistical inference, Vol. I, II (Budapest, 1980)*, 123–138, Colloq. Math. Soc. János Bolyai, 32, North-Holland, Amsterdam, 1982; MR0719704], are studied in Subsection 3.6. Finkelstein's functional law of the iterated logarithm (LIL) [H. Finkelstein, Ann. Math. Statist. **42** (1971), 607–615; MR0287600] is established via invariance principles in Subsection 3.7. Local and tail empirical processes are studied in Subsection 3.8, and moduli of continuity of the uniform empirical and quantile processes in Subsection 3.9.

Subsection 3.10 deals with the Bahadur [op. cit.] and Kiefer [op. cit.] representation via [G. R. Shorack, Z. Wahrsch. Verw. Gebiete **61** (1982), no. 3, 369–373; MR0679681] concerning its strong law (cf. Theorem 3.21) and via [P. Deheuvels and D. M. Mason, in *Probability in Banach spaces, 8 (Brunswick, ME, 1991)*, 255–266, Birkhäuser Boston, Boston, MA, 1992; MR1227623; Ann. Probab. **22** (1994), no. 3, 1619–1661; MR1303659] when studying its pointwise behaviour (cf. Theorem 3.22). Subsection 3.11 demonstrates how functional laws of the iterated logarithm for local empirical processes as in [P. Deheuvels and D. M. Mason, Ann. Probab. **20** (1992), no. 3, 1248–1287; MR1175262] can be applied to studying the almost sure asymptotic behaviour of Akaike-Parzen-Rosenblatt type [cf. H. Akaike, Ann. Inst. Statist. Math., Tokyo **6** (1954), 127–132; MR0067412; E. Parzen, Ann. Math. Statist. **33** (1962), 1065–1076; MR0143282; M. Rosenblatt, Ann. Math. Statist. **27** (1956), 832–837; MR0079873] kernel density estimators.

Section 4, Auxiliary results, contains six very useful informative subsections on some fundamental results and references. Subsection 4.1 introduces the reproducing kernel Hilbert space (RKHS), and a generalized version of the Cameron-Martin formula for a centered Gaussian process taking values in a separable Banach space, the isoperimetric inequality of C. Borell [Invent. Math. **30** (1975), no. 2, 207–216; MR0399402], and some further useful inequalities as well, for further use in the sequel. Subsection 4.2 deals with a functional LIL for superpositions of independent Gaussian processes.

Subsection 4.3 is on Karhunen-Loève (KL) expansions for centered Gaussian processes on $(0,1)$ with a positive integrable variance. Subsection 4.4 studies the RKHS of the Wiener and Brownian bridge processes on $[0,1]$. Based on [P. Deheuvels and G. V. Martynov, in *High dimensional probability, III (Sandjberg, 2002)*, 57–93, Birkhäuser, Basel, 2003; MR2033881] in the spirit of [M. Kac and A. J. F. Siegert, J. Appl. Phys. **18** (1947), 383–397; MR0020229; Ann. Math. Statistics **18** (1947), 438–442; MR0021672; M. Kac, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, 189–215, Univ. California Press, Berkeley and Los Angeles, 1951; MR0045333; *Integration in function spaces and some of its applications*, Accad. Naz. Lincei, Pisa, 1980; MR0660839], Subsection 4.5 provides an excellent summary of KL expansions for weighted Wiener processes and Brownian bridges.

Subsection 4.6 gives and discusses the definition of Bessel functions and studies some of their special cases for the sake of illustrating the use of the KL expansions of a Brownian bridge in two exercises that deal with known results, mentioning also their references.

Van de Geer's paper does indeed read like carefully written lecture notes for a graduate course on oracle inequalities and regularization. The best way of describing what it is all about is to quote the first part of Chapter 1 of her exposition, Statistical models, which reads as follows:

"In this chapter, the construction of a statistical model is discussed. We contemplate deviations from the model on the one hand, and simplicity of a model on the other. We introduce the concepts of approximation error and estimation error. The idea of complexity regularization is illustrated in two situations: histograms in density estimation and smoothing splines in regression.

"Here is a brief sketch of the contents of the other chapters. In Chapter 2, we introduce penalized $M$-estimators—or penalized empirical risk estimators. These are obtained by minimizing a loss function (e.g., least squares loss, minus maximum likelihood, or, in classification, support vector machine loss). A roughness penalty is added to the loss function to avoid overfitting. We study the behavior of the estimators in a general context. The excess risk of an estimator is a global measure for its performance. We consider so-called oracle inequalities for the excess risk. These inequalities relate the performance of the estimator to the procedure that chooses the optimal model by trading off bias and variance (or, more generally, approximation error and estimation error). In Chapter 2, we highlight the role of empirical process theory in this context.

"As an important particular case, we investigate high-dimensional linear spaces. The approximation error then comes from approximating curves or images by elements of a high-dimensional parameter space. Chapter 3 studies oracle inequalities in a regression framework, using the least squares estimators of the coefficients. That setup has the advantage that everything can be calculated explicitly. It serves as a preparation for more complicated situations.

"Chapter 4 considers general least squares estimators. In Chapter 5, we look at robust regression estimators, density estimation, and binary classification. We consider there a penalty of $l_1$-type. Chapter 6 summarizes some tools from empirical process theory. Each chapter ends with bibliographical remarks.

"Most of the work will be on estimation theory and on where this theory can make use of inequalities for empirical processes. Approximation theory (for example the approximating properties of truncated series expansions) will be touched upon only briefly.

"We consider a data set consisting of $n$ observations on a variable, say $X$, with values in some space $\mathcal{X}$. These observations are denoted by $X_1, \ldots, X_n$. We assume that the observations are independent, and that each observation follows the same probability law as $X$, say $P$ (independent, non-identically distributed observations will also be considered).

"The probability distribution $P$ is in whole or in part unknown. Our aim is to estimate $P$ or certain aspects of it. A statistical model is a set of candidate distributions $\mathcal{P}$ for $P$. If 'nothing' is known, one might want to choose $\mathcal{P}$ as the set of 'all' distributions on $\mathcal{X}$. However, if one has some idea about the form of $P$, one may want to incorporate this information into the model set $\mathcal{P}$. In that way the estimation problem is made easier, i.e., the accuracy is greater. If $P \notin \mathcal{P}$, the model is misspecified. In that case one usually has a systematic error (bias) in the estimator."

Van de Geer's paper provides an impressive set of essays on topics of estimation theory in the context of inequalities for empirical processes. *M. Csörgő*