



# Table of Contents

<b>1</b>	<b>Lecture: Introduction, Tangent Sets</b>	<b>336</b>
1.1	Introduction	336
1.2	Tangent Spaces and Information	339
<b>2</b>	<b>Lecture: Lower Bounds</b>	<b>346</b>
2.1	Lower Bounds	346
2.2	Efficient Score Functions	353
<b>3</b>	<b>Lecture: Calculus of Scores</b>	<b>357</b>
3.1	Score and Information Operators	357
3.1.1	Information Loss Models	361
3.2	Semiparametric Models	365
<b>4</b>	<b>Lecture: Gaussian Approximations</b>	<b>370</b>
4.1	Contiguity	370
4.2	Gaussian Representations	373
<b>5</b>	<b>Lecture: Empirical Processes and Consistency of Z-Estimators</b>	<b>383</b>
5.1	Empirical Measures and Entropy Numbers	383
5.2	Glivenko–Cantelli Classes	384
5.3	Consistency of M- and Z-estimators	386
5.3.1	Trick 1: Errors-in-variables	388
5.3.2	Trick 2: Cox model	389
5.3.3	Trick 3: Mixture models	392
5.4	Nuisance Parameters	394
<b>6</b>	<b>Lecture: Empirical Processes and Normality of Z-Estimators</b>	<b>395</b>
6.1	Weak Convergence in Metric Spaces	395
6.2	Donsker Classes	399
6.3	Maximal Inequalities	402
6.4	Random Functions	402
6.5	Asymptotic Normality of Z-Estimators	406
6.6	Nuisance parameters	407

<b>7</b>	<b>Lecture: Efficient Score and One-step Estimators</b>	<b>412</b>
7.1	Efficient Score Estimators	412
7.2	One-step Estimators	413
7.3	Symmetric location	418
7.4	Errors-in-Variables	420
<b>8</b>	<b>Lecture: Rates of Convergence</b>	<b>424</b>
8.1	A General Result	424
8.2	Nuisance Parameters	427
8.3	Cox Regression with Current Status Data	428
<b>9</b>	<b>Lecture: Maximum and Profile Likelihood</b>	<b>433</b>
9.1	Examples	433
9.2	Asymptotic Normality	436
9.3	Cox Regression with Current Status Data	439
9.4	Profile Likelihood	441
<b>10</b>	<b>Lecture: Infinite-dimensional Z-Estimators</b>	<b>446</b>
10.1	General Result	446
10.2	Maximum Likelihood	448
	References	454

## Preface

These notes provide extended versions of my lectures in the St Flour meeting of 1999. The general subject are semiparametric models for replicated experiments, in particular the theory for functionals that are estimable at the rate equal to the square root of the number of replications. We discuss bounds on the efficiency of estimators and tests, and methods of constructing efficient or inefficient estimators and tests, with particular attention for maximum likelihood estimators. Furthermore, we discuss abstract empirical processes, which play an important role in the analysis of the estimators.

The ten lectures have a certain overlap with material earlier published in the books [41] and [42]. A number of proofs have been omitted, because they can be found in these works. On the other hand, these notes are an attempt to give a consistent and reasonably self-contained overview of (a part of) semiparametric statistics, including digressions into empirical process theory, new examples, and a number of more recent developments.

This area is certainly not complete. To illustrate this point, scattered through the text we pose some problems whose solutions are presently unknown (to me).

Our list of references is restricted to the references that are directly relevant to the lectures. In beginning 2000 the Mathematical Reviews gave 415 responses to a query on semiparametric models, so our list does not do justice to the great amount of work having been done. A general work covering the subject of semiparametric models, but from a somewhat different point of view with relatively little attention for the subject of Lectures 5–10, is the book [3] by Bickel, Klaassen, Ritov and Wellner. This book also has an extensive list of references.

## Notation

We use the wiggly arrow  $\rightsquigarrow$  for weak convergence, also for nonmeasurable maps: if  $X_n$  and  $X$  are maps defined on some probability spaces  $(\Omega_n, \mathcal{U}_n, P_n)$  with values in a metric space  $\mathbb{D}$ , then we say that  $X_n \rightsquigarrow X$  if  $E^*f(X_n) \rightarrow Ef(X)$  for all bounded, continuous functions  $f: \mathbb{D} \rightarrow \mathbb{R}$ . Here the limit  $X$  is always assumed Borel measurable, but the  $X_n$  may be arbitrary maps. The  $*$  in  $E^*f(X_n)$  is for *outer expectation* on  $(\Omega_n, \mathcal{U}_n, P_n)$ .

Given a measure space  $(\mathcal{X}, \mathcal{A}, P)$  the set  $L_r(P)$  (for  $r \geq 1$ ) is the collection of all measurable functions  $f: \mathcal{X} \rightarrow \mathbb{R}$  with  $\|f\|_{P,r}^r := \int |f|^r dP < \infty$ .

The wiggly inequality  $\lesssim$  means “less than equal up to a constant”. The range and kernel of an operator  $A$  are denoted by  $R(A)$  and  $N(A)$ . The space of all bounded functions  $z: T \rightarrow \mathbb{R}$  on a set  $T$  is denoted by  $\ell^\infty(T)$  and  $\|z\|_T$  is the uniform norm. The set  $UC(T, \rho)$  is the set of all  $\rho$ -uniformly continuous functions on  $T$ .