Curriculum STK4011 Autumn 2019

Emil Aas Stoltenberg

November 20, 2019

All the homework exercises, the mandatory assignment, and nearly everything that we've covered in class, make up the curriculum. The book we use in the course is Casella and Berger (2002) (referred to as CB from now on), but we have not always followed it that closely. Here is a list of the chapters from the CB book that is assumed known, or that we have covered more or less in detail during lectures. Following it is an outline of what we have covered in the lectures.

- Chapters 1, 2, 3.1–3.3, 4, 5.1, 5.2 and 5.3.1 are assumed known from previous courses.
- Ch. 3.4 Exponential families.
- Ch. 5.4 Order statistics. All of Ch. 5.5.
- Ch. 6.2
- Ch. 7 Point estimation. Everything, except ch. 7.2.4.
- Ch. 8.2.1, 8.2.2, 8.3.1, 8.3.2, and 8.3.5. Also read ch. 10.3

Lecture 1 22. Aug. Probability background. **Ch. 1** of CB, including Fatou's lemma, monotone convergence theorem, dominated convergence theorem. This is background material assumed known from other courses, so proofs of things in ch. 1 is not part of the curriculum.

Ch. 2.3 in CB on moment generating functions. We proved a CLT for rv's with moment generating functions. See exercise 16 in Nils' lecture notes from 2014. We showed that existence of mgf implies all moments finite. To do so we used (and proved) Jensen's inequality.

Lecture 2 29. Aug. Ch. 3.4 in CB on exponential families. In addition, we saw that the joint distribution of the natural statistics $t_1(X),\dots,t_k(X)$ is an exponential family with natural parameters $w_1(\theta),\dots,w_k(\theta)$; and that for any $S \subset 1,\dots,k$, the joint distribution of $t_j(X)$, $j \in S$ given $t_j(X)$, $j \notin S$, is an exponential family with a distribution depending only on $w_j(\theta)$, $j \in S$.

Lecture 3 25. Sep. Ch. 6.2 in CB on sufficiency. We also covered **Ch. 5.4** on order statistics.

Lecture 4 12. Sep. More **Ch. 6.2** in CB; Sufficiency, completeness, and ancillarity. Some of the theorems that are stated without proofs in the book, were proved in class.

Lecture 5 19. Sep. Ch. 5.5 in CB, convergence concepts. Chebyshev's inequality with proof. Borel–Cantelli lemma, 'both ways', with proof. Also an inequality from p. 123 in CB.

Lecture 6 26. Sep. Ch. 5.5 continued. Slutsky–Cramér rules. Mentioned the Portmanteau theorem. See Exercise 13 Nils' lecture notes (Homework). Proved that (a) if $X_n \to_d X$ and $d(Y_n, X) \to_p 0$, then $Y_n \to_d X$; and (b) if $X_n \to_d X$ and $Y_n \to_p c$, with c a constant, then $(X_n, Y_n) \to_d (X, c)$. Delta-method, with a proof for the univariate case. Example with variance stabilising transformation. Examples of delta-method in higher dimensions.

Lecture 7 3. Oct. Ch. 7.1 Finding estimators. Method of moments. Maximum likelihood, likelihood function, score function, observed information, and Fisher information. Invariance property of $\hat{\theta}_{\mathrm{ML}}$. Introduced Kullback–Leibler divergence, and proved the Shannon–Kolmogorov information inequality. Consistency of MLE. Proved that if $\hat{\theta}_n$ is unique maximiser of $\ell_n(\theta)$; θ_0 unique maximiser of $\ell(\theta) = \mathrm{E}_{\theta_0} \log f(X;\theta_0)$, and $\sup_\theta |\ell_n(\theta) - \ell(\theta)| \to_p 0$, then $\hat{\theta}_n$ is consistent for θ_0. Mentioned sufficient conditions for the uniformity assumption. Limit distribution of MLE. We sketched a proof of $\sqrt{n}(\hat{\theta}_n - \theta_0) \to_d \mathrm{N}(0, J(\theta_0)^{-1})$, with $J(\theta_0)$ the Fisher information, in one dimension. Mentioned generalisation to several dimensions, and talked about behaviour outside model conditions.

Lecture 8 10. Oct. Ch. 7.3.2 and Ch. 7.3.3 Best unbiased estimators (UMVUE), Cramér–Rao lower bound, Sufficiency and unbiasedness, Rao–Blackwell inequality.

Lecture 9 17. Oct. Example of Rao–Blackwellisation when estimating of θ^2 with $X_1, \ldots, X_n$ independent Bernoulli(θ) data. Started **Ch. 7.2.3, 7.3.1 and 7.3.3**. Bayesian methods; posterior $\propto$ likelihood $\times$ prior. Decision theory concepts, loss, risk, Bayes risk, and Bayes solutions. Definitions of admissibility and minimaxity.

Lecture 10 24. Oct. More **ch. 7**, Bayes and decision theory. How to prove (or disprove) that an estimator is admissible; how to prove it is minimax. Sequences of priors and limiting Bayes estimators; Blyth's method. Shrinkage and the James–Stein estimator.

Lecture 11 31. Oct. Wrap up **ch. 7**. We covered all of this chapter except Sec. 7.2.4. Started **ch. 8** on hypothesis testing. Type I and and Type II errors. Power functions. Likelihood ratio tests, Neyman–Pearson lemma.

Lecture 12 7. Nov. Continued **ch. 8**. Monotone likelihood ratio families; connection to exponential families. The Karlin–Rubin theorem. UMP-test. Unbiased tests.

Lecture 13 14. Nov. The Lindeberg condition and the central limit theorem for independent random variables.

Lecture 14 21. Nov. We'll look at applications of the central limit theorem from the last lecture. Cover some left over material from ch. 8; applications of likelihood ratio tests; the Wilks theorem, see **ch. 10.3**. Exercises from Sections 6 and 11 of the exercise set.

References

Casella, G. and Berger, R. L. (2002). *Statistical Inference. Second Edition.* Duxbury Pacific Grove, CA.