

FOUR PROJECTS FOR STK-MAT2011

EMIL A. STOLTENBERG

1. CLTS WITHOUT TAYLOR

The central limit theorem says that if $X_1, \dots, X_n$ are independent and identically distributed (i.i.d.) random variables with expectation zero and unit variance, then $n^{-1/2} \sum_{i=1}^n X_i$ converges in distribution to a standard normal random variable. There are perhaps two standard proofs of this: one involves characteristic functions; the other uses a clever technique sometimes called Lindeberg swapping. What both these proof techniques have in common is that they involve second order Taylor expansions. In this project we will explore proofs that do not involve Taylor expansions, in fact, where no differentiation takes place at all. Rather these proofs relies on clever ways of measuring the distance between distributions (see [Ott \(2023\)](#) and [Wooyoung Chin \(2024\)](#)). Ambitiously, we might try to generalise these proofs to situations more involved than the i.i.d. case, and potentially using other distances.

2. CONVERSION OF HIGH-SCHOOL GRADES

Many students at the the Lycée français d'Oslo want to pursue higher education in Norway, and some Norwegian students that have attende Lycée in France want to return home for their studies. To be able to apply to Norwegian higher education in Samordna opptak, the grade point averages (GPAs) of these students are converted from the French system (0-20) to the Norwegian grading system (1-6). To get this conversion right, one ought to use data on all French GPAs and all Norwegian GPAs. I have these data for a few graduating cohorts. The standard way of doing such a conversion is to match percentiles (i.e, the GPA of the French top one percent is mapped to the GPA of the Norwegian top one percent, etc.) There several challenges with such an approach, one of which is that we only see the grades of those that pass (10 and higher in France, 2 and up in Norway), do not drop out, enter Lycée/Videregående in the first place, etc.. In other words, there might be a certain amount of [survivorship bias](#) here. In this project the aim is to develop methods that takes this challenge, as well as other challenges, into account, in order to get the 'correct' grade conversion formula.

3. INFERENCE WITH NOISY AND PRECISE OBSERVATIONS

Suppose that we have noisy observations Y_i of a signal X_i , think $Y_i = X_i + \varepsilon_i$ for some noise terms ε_i . If we want to, and can afford the cost, however, we can make precise measurements of X_i . Suppose that it is only when the probability (in the i.i.d. case) $\Pr(X_i \in B \mid Y_i)$ is deemed sufficiently high that we care to precisely

measure X_i . The data we observe are then

$$Z_i = \begin{cases} Y_i, & \text{if } \Pr(X_i \in B \mid Y_i) \geq c, \\ (Y_i, X_i) & \text{if } \Pr(X_i \in B \mid Y_i) < c, \end{cases}$$

for some threshold c set by the statistician. Based on the data $Z_1, \dots, Z_n$ the goal is to make inference on the parameters governing the distribution of the $X_1, \dots, X_n$. To do so, more model building is called for. Dependent process $(X_i)_{1 \leq i \leq n}$ can also be looked at, in which case $\Pr(X_i \in B \mid Y_i, Y_{i-1}, \dots, Y_1)$ might be the relevant probability deciding if a precise measurement is made, and so on.

4. SEMIPARAMETRIC PREDICTIVE DISTRIBUTIONS

Let $Y_1, Y_2, \dots$ be a sequence of random variables, and write $\alpha_n(A) = \Pr(Y_{n+1} \in A \mid Y_1, \dots, Y_n)$ for $n \geq 1$ and $\alpha_0(A) = \Pr(Y_1 \in A)$ for the predictive distributions of this sequence. If $(\alpha_n)_{n \geq 1}$ converges (in a suitable sense) to some limiting random measure α_∞ , then this limit can be interpreted as a draw from something resembling a Bayesian posterior distribution (see for example [Fong et al. \(2023\)](#) and [Battiston and Cappello \(2025\)](#)). As shown in [Battiston and Cappello \(2025, Sect. 4.3\)](#), sequences $(\alpha_n)_n$ converging to some α_∞ can be constructed via kernel density estimators. Ingrid and Nils, who are professors here ([Hjort and Glad, 1995](#)), studied kernel density estimators with a ‘parametric start’. In this project we will look at kernel density estimators with parametric starts in the context of predictive distributions. The main question then pertains to conditions under which predictive distributions so constructed converge in an appropriate sense.

REFERENCES

- Battiston, M. and Cappello, L. (2025). New (and old) predictive schemes with acid sequences. *arXiv preprint arXiv:2507.21874*.
- Fong, E., Holmes, C., and Walker, S. G. (2023). Martingale posterior distributions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85:1357–1391.
- Hjort, N. L. and Glad, I. K. (1995). Nonparametric density estimation with a parametric start. *The Annals of Statistics*, 23:882–904.
- Ott, S. (2023). A note on the renormalization group approach to the central limit theorem. *arXiv preprint arXiv:2303.13905*.
- Wooyoung Chin, C. (2024). A proof of the central limit theorem using the 2-wasserstein metric. *arXiv:2402.17085v1*.