MASTER LEVEL INTERNSHIP PROPOSAL: LAMN FOR MIXTURES OF LAN MODELS

TRISTAN BENOIST, FABRICE GAMBOA AND CLÉMENT PELLEGRINI
INSTITUT DE MATHÉMATIQUES DE TOULOUSE

In [1] we consider parameter estimation for a finite mixture of multinomial statistical models. More precisely for any $\alpha \in \{1, \ldots, \ell\}$ and $\theta \in \Theta \subset \mathbb{R}^d$ with Θ compact, $j \mapsto p_\theta(j|\alpha)$ defines a probability distribution on $\{1, \ldots, J\}$. We denote $\mathbb{P}_\theta^\alpha$ the probability measure of a sequence of iid random variables distributed according to $j \mapsto p_\theta(j|\alpha)$. Then for each $\alpha \in \{1, \ldots, \ell\}$, $(\mathbb{P}_\alpha^\theta, \Theta)$ is a statistical model. It is standard that if for each $j \in \{1, \ldots, J\}$, $\theta \mapsto p_\theta(j|\alpha)$ is C^3, then the statistical model corresponding to α is Locally Asymptotically Normal (LAN).

In [1], motivated by quantum mechanic experiments, we show under appropriate assumptions that if we consider not each model individually but a mixture of them, then this new statistical model is Locally Asymptotically Mixed Normal (LAMN). More precisely for a probability distribution q over $\{1, \ldots, \ell\}$ and any $\theta \in \Theta$ we set

$$\mathbb{P}^\theta = \sum_{\alpha=1}^{\ell} q(\alpha) \mathbb{P}_\alpha^\theta.$$

Then the model $(\mathbb{P}^\theta, \Theta)$ is LAMN. Namely, asymptotically it is not a Gaussian model but a mixture of Gaussian models.

From de Finetti's theorem we learn that any sequence of random variable whose law is exchangeable (*i.e.* such that it is invariant if we do any permutation in the sequence) has a law that is actually a mixture of iid laws. Generalizing our previous definitions, the model $(\mathbb{P}^\theta, \Theta)$ is a model for an exchangeable set of data iff there exists a probability space $(\mathcal{A}, q)$ such that

$$\mathbb{P}^\theta = \int_{\mathcal{A}} \mathbb{P}_\alpha^\theta dq(\alpha).$$

In [1] we limit ourselves to a finite set $\mathcal{A}$. One of the aims of this internship is to generalize to different $\mathcal{A}$. These models have many applications beyond the initial quantum mechanic motivating example (see [2]).

A first step would be to take $\mathcal{A} = \mathbb{N}$ or $\mathcal{A} = \mathbb{R}$. Another direction of research is to go beyond iid model for each $\mathbb{P}_\alpha^\theta$. The goal would be to find the least assumptions needed directly on the models $(\mathbb{P}_\alpha^\theta, \Theta)$ such that we can conclude on the model $(\mathbb{P}^\theta, \Theta)$.

This work could eventually be continued during a Ph.D. Thesis in statistics.

References

[1] T. Benoist, F. Gamboa and C. Pellegrini, *Quantum non demolition measurements: parameter estimation for mixtures of multinomials*, Electron. J. Stat. **12**(1) (2018) 555–571

[2] F. Camerlenghi, A. Lijoi, P. Orbanz and I. Prünster, *Distribution theory for hierarchical processes*, Ann. Statist. **47**(1) (2019) 67–92.

Email address: tristan.benoist@math.univ-toulouse.fr, fabrice.gamboa@math.univ-toulouse.fr and clement.pellegrini@math.univ-toulouse.fr

Key words and phrases. Parameter Estimation, Local Asymtotic Normality.