OPTIMAL AVERAGING ESTIMATION
FOR DENSITY FUNCTIONS

Peng Lin[#1], Jun Liao[#2], Zudi Lu[3,4], Kang You[5] and Guohua Zou[*5]

[1] *Shandong University of Technology*, [2]*Renmin University of China*,
[3]*University of Southampton*, [4]*City University of Hong Kong*
and [5]*Capital Normal University*

Abstract: Extraction of information from data is critical in the age of data science.
Probability density function theoretically provides comprehensive information on
the data. But, practically, different probability density models, either parametric
or non-parametric, can often characterize partial features on the data, e.g., owing
to model bias or less efficiency in estimation. In this paper we suggest a framework
to optimally combine different density models to catch the comprehensive data
features by a new information criterion (IC) based unsupervised learning approach.
Our optimal information extraction is in the sense that the resultant density
averaging or selected density minimises the Kullback–Leibler (KL) information loss
function. Differently from the usual supervised learning IC for model selection or
averaging, we first need to derive an estimator of the KL loss function in our setting,
which takes the Akaike and Takeuchi information criteria as two special cases.
A feasible density model averaging (DMA) procedure is accordingly suggested,
with the DMA estimation achieving the lowest possible KL loss asymptotically.
Further, the consistency of the weights of the DMA estimator tending to the optimal
averaging weights minimizing the KL distance is obtained, and the convergence rate
of our empirical weights is also derived. Simulation studies show that the DMA
performs overall better and more robustly than the commonly used parametric or
nonparametric density models, including kernel, finite mixture, logarithmic scoring
rule and selection methods for density estimation in the literature. The real data
analysis further demonstrates the performance of the proposed method.

Key words and phrases: Asymptotic optimality, density averaging, density estima-
tion, weight choice.

1. Introduction

Extracting knowledge from data is key to data science (c.f., Baimuratov
et al., 2019). As is well known, density function of a data-generating random
variable theoretically provides comprehensive information on the data, estimation
of which is hence a fundamental problem in Econometrics, Statistics and data
science. Usually there are two types of approaches that are widely applied to

---

*Corresponding author. E-mail: ghzou@amss.ac.cn
#Contributed equally to this work.