

СТАНОВИЩЕ

от доцент д-р Пламен Стоянов Матеев,
от Факултета по математика и информатика на
Софийски университет „Св. Климент Охридски“,

за дисертационен труд на **Боно Стойчев Нончев** на тема „MDL принцип за избор на модел при анализ на данни “ за присъждане на образователната и научна степен „доктор“ в професионално направление 4.5. *Математика (Теория на вероятностите и математическа статистика)*.

Научен ръководител е авторът на това становище – доц. П. Матеев.

Становището е подготвено в съответствие със заповед на Ректора на СУ № ВО 38-388 от 05.06.2015 г. и изискванията на ЗРАСРБ, Правилника за неговото прилагане, както и ПУРПНСЗАД на СУ и на ФМИ на СУ и приетите от ФС на ФМИ указания за изготвяне на рецензии и становища.

Дисертационният труд е посветен на задачи от съвременен клон на анализа на данни, основан на информационния аспект към стохастиката, а именно, измерването на количеството информация като вероятност в логаритмична скала. Начало на теорията на информацията поставя Клод Шенон (1948), като доказва, че оптималното кодиране на съобщенията се основава на вероятностното им разпределение. На Колмогоров (1963~) дължим доказателството, че по сложността или алгоритмично достигнато оптимално кодиране може да се съди за вероятностното разпределение на съобщенията. Рисанен формулира принципа MDL (1978), който предлага използване на Колмогоровската сложност в практиката при избор на подходящ вероятностен модел. Боно Нончев атакува нерешения проблем за сравняване на сложността на модели от фамилия непрекъснати разпределения, чиято сложност във вид на нормализирано правдоподобие е разходящ интеграл. Той предлага оригинален нов подход - сложност при ограничения. Подходът се прилага успешно за определяне сложността на многомерно нормално разпределение, след което кръгът на приложение се разширява последователно върху сферични и независими фамилии с параметри на локация и мащаб, както и ограничения върху параметри на форма на примера на многомерно некорелирано T-разпределение.

В изброените случаи сложността на разпределенията е получена в експлицитна форма като се използва нетривиална техника, основана на обобщени функции – по-специално делта-функция на Дирак.

Аналитичните резултати са апробирани върху изкуствено генерирани данни. При решаването на този проблем са демонстрирани програмистки сръчности и професионални умения. Накрая е показана и възможността за използване на MDL

принципа и конкретните резултати за избор на модел/разпределение при анализ на реални данни.

Резултатите в дисертацията са представени на седем научни форума – три редовни пролетни научни сесии във ФМИ и четири международни конференции у нас (2) и в чужбина (2). Подготвени са пет самостоятелни научни публикации, от които три са излезли от печат в авторитетни реферирани издания през 2013 и 2014 година.

Дисертацията е на общо 173 страници включително титулните. Основното съдържание е върху 110 страници, от които четири страници заключение (глава 5) и двадесетина уводни (глава 1) за формулиране на задачата и уточняване на основни понятия и факти от стохастиката и теорията на информацията. Тук трябва да се отбележи представената вероятностна интерпретация (теорема 1.20) на факторизацията на Нейман-Фишер.

Втората глава е посветена на въвеждане на MDL принципа. На 35 страници са описани историята на идеите, философията и основни класически резултати. Главата завършва със структуриран обзор на публикации в областта от последните години.

Третата глава представя на 25 страници получените от дисертанта аналитични резултати. Още толкова страници заемат числените (глава 4). Двете глави съдържат основните приноси на дисертанта, изброени по-горе. Трябва да подчертая, че предложеният *оригинален метод* може да се приложи при определяне на стохастичната сложност и на други фамилии разпределения.

Библиографията е от 69 заглавия на цитираните в текста статии и други източници. Три са авторските статии, 31 са публикувани след 2006, 15 са от периода 1990-2005 и останалите 10 са от 1971-1990, с изключение на основополагаща статия на Колмогоров от 1963.

Описанието на дисертацията няма да е пълно, ако не бъдат споменати приложенията. Освен стандартното „Съдържание“, дисертацията е снабдена със „Списък на означенията с речник (англо-български) на основните понятия“ и показалец (индекс) с хипер-връзки. Три самостоятелни текста са отделени като спомагателни приложения. Първото може да определим като учебно помагало по теория на информацията от над 40 страници. Второто и третото са по-кратки – общо 10 страници и съдържат съответно: първото - сведения за работа с обобщени функции; второто - описание на използвания алгоритъм за оптимизация и имплементацията му с паралелизация на изчислителния процес.

Авторефератът е самостоятелна творба. В ограничения обем са представени резултатите на дисертанта и необходимите за възприемането им предварителни сведения, формулирани са претенциите за научен принос, намеренията за следващи

изследвания и всички останали атрибути в съответствие с изискванията. Авторефератът е на български език. Не е споменат фактът, че дисертацията е на английски.

Познавам Боно Нончев от десет години. Първо като студент в бакалавърската програма „Приложна математика“ и след това и в магистърската програма „Вероятности и статистика“. Като научен ръководител на дипломната му работа и след това на дисертацията, съм бил свидетел на търсенията, трудностите и успехите на Боно Нончев. Без съмнение всички резултати са негово собствено дело. Моята заслуга като ръководител се ограничаваше в насърчване и напътствия. Препоръката ми към него е да не губи набраната в последната година инерция и максимално бързо да реализира публикации на непубликуваните си резултати.

ЗАКЛЮЧЕНИЕ

Всичко изложено дотук свидетелства в полза на тезата, че докторантът **Боно Стойчев Нончев** притежава задълбочени теоретични знания, вероятно мислене и професионални компютърни умения, демонстрира качества и възможности за самостоятелни научни изследвания в областта на статистическия анализ на данни, вероятностите и теорията на информацията. Представени са **достатъчно теоретични и научно-приложни резултати, които представляват оригинален принос в науката**, отговарящи на изискванията на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ, съответния Правилник на СУ „Св. Климент Охридски“ и специфичните изисквания на ФМИ.

Без колебание потвърждавам своята **положителна оценка** на представеното в дисертационен труд и автореферат изследване, научни резултати и приноси, и ще се солидаризирам с *решение на почитаемото научно жури да присъди на*

Боно Стойчев Нончев

образователната и научна степен „доктор“

по научна специалност Теория на вероятностите и математическа статистика, професионално направление 4.5 Математика.

27 юли 2015 г.

Подпис:

(доц. д-р Пламен Матеев)