

## РЕЦЕНЗИЯ

от доц. д-р Мариян Неделчев Милев, УХТ – ПЛОВДИВ,  
гр. Пловдив 4002, бул. Марица 26, тел. 032 603 744

върху дисертационен труд „**Robust Statistical Modeling Through Trimming**” с автор **Нейко Матеев Нейков** за придобиване на научна степен „доктор на науките” в Институт по математика и информатика при БАН - София (ИМИ – БАН); Област на висшето образование: 4. Природни науки, математика и информатика; Професионално направление: 4.5. Математика  
Научна специалност: 01.01.10 „Теория на вероятностите и математическата статистика”

Представям становището си по този конкурс като член на Научното жури, определено със Заповед №99/26.05.2016г. на директора на ИМИ – БАН, съгласно Решение на НС на ИМИ - БАН (Протокол №5/20.05.2016). Становището е изготвено според изискванията на: Закона за развитието на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ, Правилника за условията и реда за придобиване на научни степени и заемане на научни длъжности в БАН и Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности в Института по математика и информатика на БАН.

### **1. Обща характеристика на научната и научно-приложната дейност на кандидата, свързани с придобиването на научна степен „доктор на науките” в научна специалност: „Теория на вероятностите и математическата статистика”.**

Нейко Матеев Нейков завършва СУ “Св. Кл. Охридски”, Факултет по математика и механика, специалност: „Теория на вероятностите и математическа статистика” и след магистратура по математика, доктор по същата специалност с дисертационен труд: “Робастни методи в многомерния статистически анализ”. Понастоящем заема длъжността доцент в департамент “Прогнози на времето”, НИМХ-БАН. За научните звания и степени са представени документи, удостоверяващи валидността им. Участвал е в изследователски проекти от научния план на НИМХ и редица международни научни проекти, девет от които са COST през последните шест години, както и проекти, свързани с робастно моделиране и оценки: *робастно статистическо моделиране в задачи с големи размерности* [6], *робастна множествена квантилна регресия* [7], *робастни квази-правдоподобни регресионни оценки* [9], *робастни оценки с висока прагова точка за обобщени адитивни модели* [11], *робастно статистическо моделиране със смеси от разпределения* [15], *робастни статистически методи с висока прагова точка* [20]. Доцент д-р Нейко Нейков е изнасял лекции за докторанти по робастни статистически методи и приложения, стохастични модели, изчислителни аспекти на робастната статистика и др., ръководител на редица дипломанти от ФМИ, СУ “Кл. Охридски”. Д-р Нейко Нейков е участвал в множество международни конференции и работни семинари с доклади.

### **2.Общо описание на представените материали.**

Доц. д-р Н. Нейков е представил дисертационен труд, написан на английски език, състоящ се от 200 страници, автореферат (94 стр.) на български и 11 публикации, които са на английски и са излезли от печат като две от статиите [9,10] са публикувани в списание Плиска, една статия [4] в Доклади на БАН; три публикации [2,3,11] са в елитните международни издания Physica-Verlag, Heidelberg и Birkhauser, Basel; статиите [1, 5, 6, 7, 8] са публикувани в международни списания с импакт фактор, съответно – [1] *Journal of Statistical Planning and Inference* (0.307), [5,6,7] в *Computational Statistics & Data Analysis* (1,029), [8] в *Statistical Papers* (0,68);

Всичките представени единадесет статии са в съавторство - пет са с един съавтор [1,2,3,9,10], три с два съавтори [4,6,8], три с трима съавтори [5,6,11]. Приемам, че участието на авторите в тези статии е равностойно. Общият импакт фактор е 4,624, среден импакт фактор на статия е 0,923, а личният принос е 1,501. От представеният списък с цитирания, общият брой цитирания на всички публикации е 214, като Н-факторът е 6, съставен от [1,2,3,5,6,11].

Всичките 11 публикации приемам за научни постижения в научната специалност „Теория на вероятностите и математическата статистика”.

От представените документи се вижда, че са изпълнени формалните изисквания на ЗРСАРБ, както и тези на от ПБАН, гл. 4, чл. 6 (1). Ще продължа становището си, съгласно указанията за изготвянето на становища за придобиване на научна степен „доктор на науките” в ИМИ-БАН.

### 3. Дисертационен труд и автореферат, преценка на приносите.

Представеният дисертационен труд е в обем от 200 страници и се състои от увод, 8 глави и цитирана литература, посветени на статистически оценки с висока прагова точка. Въведени са класове робастни оценки като  $S$ -оценките, основани на минориране – мажориране от  $k$ -ти нареден елемент на  $d$ -пълно множество от неотрицателни функции, тримиращи квази-правдоподобни оценки, тримиращи квантилни оценки, тримиращи (правдоподобни) оценки с пенализация в задачи с висока размерност. Предложения са обобщения на техниката на  $d$ -пълнота на Vandev (1993), за да се характеризират праговите точки на някои от изброените класове оценки, основани на тримиране. Предложени са FAST-TLE алгоритъм за приближено пресмятане на тримиращи правдоподобни оценки и FAST-GTE алгоритъм за приближено пресмятане на обобщени тримиращи оценки. Тяхното приложение е демонстрирано за пресмятане на въведените рабастни оценки при експерименти с извадки от реални данни.

В глава 1 на дисертацията, бих искал да отлича един от приносите на автора, където са характеризирани праговите точки на WTLE( $k$ ) оценки в обобщените линейни модели от линейната експоненциална фамилия. Тези резултати са предварително публикувани съответно като Теорема 1 и Теорема 6 в статия [1] от представения списък, която е многократно цитирана в престижни статистически списания. Определен е индексът  $d$ -пълнота чрез параметъра  $N(x)$  за регресионни модели с Лапласово разпределение от ред  $q$  на грешката, което има широко практическо приложение за редица проблеми, освен посочените в дисертацията. Например, за моделиране на икономически процеси и оценяване на финансови деривати, където се използват leptokurtic свойствата на това разпределение при  $q \in (1, 2)$ .

В Глава 2 е предложен FAST-TLE алгоритъм за приближено пресмятане на TLE( $k$ ) оценките, както е доказана и локалната сходимост на този алгоритъм. Изследвано е поведението на тези оценки върху оригинални извадки от данни, (пример 1 от ваксиниране, пример 2 от реален токсикологичен експеримент, пример 3, лого-линеен модел при злополуки).

В глава 3 е предложена обобщена техника на  $d$ -пълнота за изследване на праговата точка на wGTE( $k$ ) оценките като са дадени необходими условия за съществуване на решение на wGTE( $k$ ) оптимизационна задача и долна граница на праговата точка в Предложение 3.2. В предложение 3 чрез доказване пълнотата на множество от специален клас функции, дисертантът обобщава предишни свои резултати, в следствие на което е характеризирана праговата точка на WTLE( $k$ ) оценки за линейната логистична регресия с обобщена свързваща функция. В глава 4 са проведени експерименти чрез Монте Карло симулации, показващи, че оценките по метода на максималното правдоподобие (ММП) са ненадеждни при наличие на несъгласувани наблюдения в данните, докато TLK( $k$ ) са надеждни оценки при процент на тримиране по-висок от процента на замърсяване на данните с несъгласувани наблюдения. В апендикс 4.6 е характеризирана праговата точка на TLE( $k$ ) оценките за регресионни модели с Гумбелово разпределение на грешката.

Глава 5 е посветена на рабастно оценяване на параметрите на крайна смеси от вероятностни разпределения по данни чрез WTLE( $k$ ). Показано е предимството на FAST-TLE алгоритъма, в сравнение с оценките по ММП. Предложен е адаптивен подход за определяне на параметъра на тримиране  $k$ , основан на тримиращия информационен критерий на Бейс, както и процедура на робастно клъстериране на данни. Поведението на оценките е изследвано чрез Монте Карло симулационни експерименти. Глава 5 отразява до голяма степен приносите, значимостта и научно приложния характер на изследванията на дисертанта Нейков. Резултатите от нея са публикувани в статия [5], която е цитирана над 100 пъти в престижни

международни статистически списания, монографии и доклади. Праговата точка на  $WTLE(k)$  оценките е характеризирани с Твърдение 5.1, изнесено в апендикса към тази глава.

В глави 6 и 7 дисертантът изследва два типа задачи чрез тримиране. Въвежда се нов клас робастни оценки  $ETQL(k)$  като е демонстрирано чрез примери предимството им в сравнение с квази-правдоподобните оценки. В глава 7 са въведени клас квантилни регресионни оценки  $LTQR(k)$ . Теорема 6.1 и 7.1 характеризират съответно праговата точка на тези класове оценки, докато Теорема 7.2 характеризира асимптотичното поведение на  $LTQR(k)$ . Поведението на тези класове оценки в сравнение с класически оценки е изследвано чрез Монте Карло симулации.

В глава 8 дисертантът изследва задачи с големи размерности чрез тримиране и пенализация като предлага робастни двустъпкови процедури  $TSIS$  и  $TISIS$  за наблюдение, чрез които се избират значими предикторни променливи, основани на  $TLE(k)$  оценките и  $TLE(k)$  оценките с пенализация. Така се преодолява проблемът с неприложимостта на методите на робастната статистика в обобщени линейни модели с изключително висока размерност, където броят на предикторите е по-голям от броя на наблюденията на извадката.

В заключение ще добавя, че дисертацията има завършен вид, като твърденията в нея са ясно формулирани и доказани. Представеният автореферат правилно отразява приносите на дисертацията. От направеният анализ на дисертационния труд на доц. д-р Нейко Нейков мога убедено да заключа следните изводи:

- Актуалността на разработените методи за робастно статическо моделиране чрез тримиране в дисертационния труд говорят за неговата научно-приложна значимост.
- В дисертационния труд ясно и коректно са дефинирани целите и методите за постигането им. Дисертантът е показал високата степен на познаване на състоянието на проблема и актуалните литературни източници.
- Избраните класически методи за статическо моделиране са в съответствие с поставената цел и задачи на дисертационния труд

#### **4. Критични бележки и препоръки.**

Нямам критични бележки и препокръпки, но бих искал да отбележа, че в представените материали всички публикации са в съавторство. Приемам, че необходимостта от работа в колектив се дължи на широтата и обхвата на научно-приложните изследвания. Значимостта на тези научно-приложни разработки се вижда и от нивото на цитируемост, в случая общо 214, H-фактор 6, което говори за качеството на научните изследвания на доц. Нейков и за високата им оценка от международната научна общност.

#### **6. Заключение.**

Представеният от доц. д-р Нейко Матеев Нейков материали показват, че той удовлетворява основните изисквания на ЗРАСРБ, на Правилника за прилагането му, на съответните правилници на БАН и на ИМИ - БАН за придобиване на научна степен „доктор на науките”. Поради това убедено предлагам на Научното жури да присъди на доц. д-р Нейко Матеев Нейков научната степен „доктор на науките” в областта на висшето образование: 4. Природни науки, математика и информатика; професионално направление: 4.5. Математика; научна специалност: 01.01.10 „Теория на вероятностите и математическата статистика”

Дата:

23.08.2016г.

Член на журито

/доц. д-р М. Милев/