

НС на Институт по Математика и Информатика, БАН

РЕЦЕНЗИЯ

по конкурс за **професор** в област на висше образование **4. Природни науки, математика и информатика**, научна специалност 01.01.10 „Теория на вероятностите и математическа статистика”, обявен в „Държавен вестник” бр. 58/29.07.2011 г.

**Рецензент: професор дмн Николай Михайлов Янев,
Секция ‘Вероятности и Статистика’,
Институт по Математика и Информатика, БАН
Тел. 979-3861**

Единствен кандидат по гореописания конкурс е **доцент дмн Евгения Асенова Стоименова**, Секция ‘Вероятности и статистика’, Институт по Математика и Информатика, БАН. За участие в конкурса Е. Стоименова е представила един диск и една книга “Измерителни качества на тестове”. Дискът съдържа следните материали, представени като PDF файлове:

1. Молба
2. Професионална автобиография.
3. Диплома за завършено висше образование.
4. Диплома за придобита образователна и научна степен „доктор“.
- 4а. Диплома за придобита научна степен „доктор на науките“.
5. Пълен списък на научните трудове.
6. Списък на научните трудове за участие в конкурса.
7. Саморъчно подписана авторска справка за научните приноси на трудовете.
8. Списък цитирания.
9. Препис-извлечение от протокола на НС на ИМИ БАН за инициране на процедурата.
10. Държавен вестник с обявата за конкурса.
11. Документи за учебна работа.
- 11.1. Справки за четени лекции/упражнения.
- 11.2. Списък на издадени учебници и учебни помагала – на хартия и/или електронни.
12. Списък на научноизследователски проекти с ръководство или участие на кандидата.
13. Копия от трудовете по т.б.
- 14а. Документ, удостоверяващ заемането на академична длъжност “доцент” поне 2 години.
15. Свидетелство за академична длъжност (научно звание) „ст.н.с. II ст”.
16. Автореферат на дисертацията за доктор на математическите науки.

Документите са представени в 8 екземпляра – един на хартия и 7 на електронен носител. Има изразено съгласие само два от документите, Авторската справка и Списък на научните трудове за участие в конкурса, да бъдат поставени в интернет страницата на ИМИ. Последното обстоятелство затруднява много работата на рецензентите и журито, тъй като трябва да доведат до знанието на НС съдържанието на всички останали документи. Приемам за рецензиране всички представени материали.

1. Кратки биографични бележки.

В приложената автобиография не са дадени място и дата на раждане, като са посочени само следните четири дати:

1984 – диплома от ФМИ, СУ.

1994 – д-р по математика, н.с. 01.01.10. ТВ и МС.

2003 – ст.н.с. 2 ст. в секция В&С, ИМИ.

2010 – дмн по н.с. 01.01.10. ТВ и МС.

Посочени са различни аспекти от основните научни интереси: непараметрична статистика, статистическо моделиране и приложна статистика.

Отбелязано е, че са публикувани над 50 научни статии, 4 учебника и 10 рецензии на книги. Всъщност от списъка на публикациите се вижда, че научните статии са 56.

Дадени са подробно участие в конференции, покани от чужбина, участие в международни проекти, редколегии, програмни комитети, международни организации, експертна дейност, рецензиране и рефериране.

Е. Стоименова е била н. консултант на трима защитили докторанти и трима специализанти. В момента е н. ръководител на един задочен аспирант.

Посочена е сериозна преподавателска дейност в съвместни магистърски програми на ИМИ с други университети, както и четене на различни курсове в СУ, НБУ, Американски университет, Германия и др.

2. Анализ на представените по конкурса публикации.

Списъкът на представените в конкурса трудове съдържа 23 заглавия, от които 22 научни статии и една книга (спомената по-горе), всички излезли от печат.

Не е отбелязано специално, но ако се сравни приложеният автореферат с този списък, то може да се отбележи, че 7 от статиите са след докторската дисертация. От тях 4 са публикувани в престижни международни списания, 1 в сборник от международна конференция и 2 в Доклади на БАН.

Въобщо от статиите - 11 са самостоятелни и 11 съвместни (4 – с един съавтор, 3 – с двама съавтора и по една статия с 3, 4 и 5 съавтора). Общо 12 от представените работи са в списания с импакт фактор, даден поотделно за всяко списание през 2010 г.). Така за общият импакт фактор се получава 9.159. От тези статии 7 са в престижни международни математически списания: *Communications in Statistics*, *J. of Statistical Planning and Inference*, *Australian J. of Statistics*, *Appl. in Math.*, *Ann. Ins. Stat. Math.*, *Stat.&Prob.Letters*, *Statistics*. Две от

статииите са с приложен характер и са публикувани в съответни международни списания: *J. Agricultural Sci.* (IF 1.418) и *J. Geophysical Research – Space Physics* (IF 3.303). Осем от статиите са публикувани в сборници от международни конференции – 7 в чужбина и една у нас. Три от работите са отпечатани в Доклади на БАН.

Представен е подробен списък на 66 цитата, получени от 18 работи, а също и 10 цитата на статистически анализи.

3. Научни приноси в работите [5-15] и [19-23], на които се основава дисертацията за дмн с заглавие “Метрични методи в непараметрични статистически задачи”. Във връзка с това прилагам част от рецензията, която съм написал по този повод (цитираните подолу резултати могат да се видят също в приложеният към документите автореферат):

“Дисертацията е представена на 243 страници текст и съдържа Увод, четири глави, 3 стр. „Заключение”, 5 стр. „Приложение”, 3 стр. ‘Публикации по дисертацията’ (общо 24 работи) и 12 страници Библиография от цитирана литература (за съжаление неномерирана).

Най-общо казано, дисертационният труд е в областта на непараметричната математическа статистика.

В параметричната МС се предполага, че функциите на разпределение съдържат неизвестни параметри, които трябва да се оценят или относно които да се направят някои статистически изводи, т.е. да се проверят статистически хипотези.

В непараметричната статистика не се предполага принадлежност към някакво параметрично семейство, но понякога се правят друг тип ограничения, като например монотонност или непрекъснатост. Особено важна и интересна част от непараметричната статистика представляват така наречените рангови критерии за проверка на статистически хипотези. Те се основават на рангови статистики, получени от наредбата на наблюденията в извадките и от техните позиции в тези наредби, т.е. от техните рангове. Първите фундаментални резултати в тази област са много добре представени в класическата монография на Хайек и Шидак (1967). Важна роля играят и т.н. наредени статистики, основни резултати за които могат да бъдат намерени в монографията на Г. Дейвид (1970). Основополагащи принципи и резултати в областта на проверка на статистически хипотези са представени във фундаменталната монография на Е. Леман (1969).

В настоящата дисертация (най-общо казано) се използва метричен подход при изследване на рангови критерии и непараметрично оценяване на плътности като са дадени и някои приложения.

Както вече беше отбелязано, дисертацията започва с Увод, където на 4 страници е представен кратък исторически и научен преглед на тематиката и са формулирани основните цели и задачи. Изложението е стегнато и ясно, но за съжаление тук липсва една дори кратка представа за най-интересните и съществени постижения. Затова след Увода най-добре е читателят да се прехвърли на стр.221, където е Заключението, за да се запознае в резюме с най-важните резултати.

Основните теоретични изследвания и резултати са представени в първите три глави, а в последната Глава 4 са дадени някои приложения в механиката и надеждността на тестови измервания.

Глава 1 заема централно място в дисертацията и е разположена на 84 страници. Тя започва с кратко въведение, където се обсъждат основните рангови критерии, представени в тази глава.

В Параграф 1.1. се изследва алгебричната структура на тези критерии, основана на симетричната група от пермутации и се конструират рангови критерии, породени от метрики.

Особено внимание се обръща на метриката на Чебишев $M(\alpha, \beta)$, на чиито свойства е посветен Параграф 1.2. Така в Теорема 1.2.1 са намерени математическото очакване $E\{M(\alpha, \beta)\}$ и дисперсията $Var\{M(\alpha, \beta)\}$. При подходяща центровка и нормировка в Теорема 1.2.2 е получено като гранично разпределение равномерното в интервала $(-, \text{корен от } 3, +, \text{корен от } 3)$. Един без съмнение интересен резултат. Теорема 1.2.3 дава статистиката на критерия при две извадки с едностранна алтернатива, Теорема 1.2.4 – при двустранна алтернатива, Теорема 1.2.5 – при ненаредена алтернатива, а Теорема 1.2.7 – при хипотеза за симетрия. Точното разпределение на т.н. M -статистика е намерено в Теорема 1.2.8, а асимптотичното – в Теорема 1.2.9. Последното е геометрично при m/n клонящо към λ . При $m=n$ се изследва асимптотиката на математическото очакване и дисперсията.

В Теорема 1.3.3. е получен НМРК (най-мошен рангов критерий) при алтернатива на отместване, близко до 1, в случай на равномерно разпределение в $(0, 1)$.

Параграф 1.4 изследва т.н. рангови статистики на предхождане и превъзхождане a, a', b и b' (Дефиниция 1.10 на стр. 51). Като следствие от (1.4.1) е дадено представянето им на стр.52, в което обаче има допусната грешка, която следва да бъде обяснена при защитата. Съвместното разпределение на съответните статистики a' и b' при произволни разпределения е дадено в Теорема 1.4.3, а Теорема 1.4.4 изследва случая при нулева хипотеза (т.е. когато $F(x)=G(x)$). В Теорема 1.4.5 се изследва асимптотичното поведение. Лема 1.4.1 и 1.4.2 дават разпределението на обобщените статистики.

В Параграф 1.5 се изследват критерии за предхождане и превъзхождане. За съжаление Параграф 1.5.1 започва с нова дефиниция на статистиките a, a', b и b' , която съществено се различава от тази в Параграф 1.4. По този начин става неясно разбирането и проследяването по-нататък на статистиката M . Това недоразумение следва да бъде обяснено от докторантката.

Съответни резултати за т.н. обобщени статистики са получени в Твърдение 1.5.1 (Защо „Твърдение“?) и Теорема 1.5.6. Накрая в Таблица 1.3 са дадени критичните стойности на ранговата статистика M_r при ниво на съгласие 0.05 и са приведени два примера чрез симулации.

Глава 2 съдържаща 46 страници и изследва въпроса за разпределение на рангови статистики при различни алтернативи.

Да отбележим, че изследванията на статистиките при дадени хипотези, предложени в Глава 1, са съществени за определяне на критичните области за отхвърляне на хипотезата, докато резултатите от Глава 2 са свързани с изследване мощта на критериите. Разглежданите в Глава 2 алтернативи са от т.н. Леманов тип., които са подклас от общата алтернатива са стохастична наредба на две разпределения $H_a: (x) \geq G(x)$, когато $G(x)=h(F(x))$, $h(0)=0$, $h(1)=1$, $h'(x) \geq 0$.

Така в Параграф 2.1 е разгледан случая, когато $h(x)=x^k$, $k>1$. При тази алтернатива в Теорема 2.1.1 е намерено разпределението на M-статистиката $M=\max(m-a, n-b)$, където a и b са определени чрез Означение 1.10 (Всъщност става въпрос за Дефиниция 1.10 на стр. 51).

В Параграф 2.2 се предполага, че $G(x)=1-(1-F(x))^\gamma$, $0<\gamma<1$. За M-статистиката $M=\max(m-a', n-b')$ е намерено разпределението (Твърдение 2.2.1) при равни извадки.

В Параграф 2.2.1 се предполага $\gamma=1/\eta$, където $\eta>0$. Теорема 2.2.1 дава съвместното разпределение на статистиките a' и b' при неравни извадки. За съжаление тук пак възниква недоразумение, тъй като от една страна се посочва Дефиниция 1.10 от Параграф 1.4, а от друга страна приведената на страница 94 дефиниция е съвсем различна.

Разпределението на M-статистиката при неравни извадки е пресметнато в Теорема 2.2.1, която почива на 3 интересни лема.

В Параграф 2.3.1 е намерено разпределението на обобщените статистики A_r и B_r (Лема 2.3.1) и статистиката M_r (Теорема 2.3.1).

Получените резултати се използват в Параграф 2.4 за изследване мощта на критериите при тези Леманови алтернативи. Генерирани са различни извадки и са построени съответните криви. Използвани са различни критерии и разпределения. Резултатите са интересни и са представени в таблици и фигури, придружени със съответните изводи.

Глава 3 съдържа 46 страници и е посветена на две задачи: частично нареждане и непараметрично оценяване на плътности при неточни данни. Така, в Параграф 3.1 се изследва хипотезата за липса на тренд, т.е. за близост на извадка от частични ранжировки до фиксирана такава. Използува се статистиката на Fligner-Verducci (1988), породена от метриката footrule на Спирмън, изведена в (1.3.4). За така дефинираната статистика D^* е показано в Теорема 3.1.1, че е асимптотично нормална, като са пресметнати точно съответните математическо очакване и дисперсия. В доказателството на теоремата фигурират обаче случайни величини X_{nj} , които не са дефинирани и за които читателят трябва да гадае, какво означават.

В Параграф 3.2 се изследва функцията на риска при разпределяне в две групи и са намерени някои интересни свойства.

В Параграф 3.3 е предложен алгоритъм за йерахичен клъстер анализ, използващ частична ранжировка.

В Параграф 3.4 са разгледани някои методи за непараметрично оценяване на плътности при непълни и закръглени данни. Предложени са някои ядрени оценки. В условието на Теорема 3.4.5 обаче е допусната грешка, която прави твърдението неясно. Разгледани са и модификации на оценките на Надарая-Уотсън за цензурирани данни. Накрая са дадени примери с графични представяния и различни илюстративни изследвания.

В **Глава 4**, която съдържа 44 страници текст, са изложени три интересни статистически приложения. Най-напред в Параграф 4.1 са предложени статистически модели в механиката, описващи т.н. характеристична крива на взаимодействие вода-почва. Така в точка 4.1.1 е описан модел на двуфазна регресия с неизвестна точка на свързване γ . Подробно е изследван линейният случай, за който в 4.1.2 е намерена оценка на параметъра γ . Разгледана е модификация при известно ниво на сатурация и са намерени оценки на неизвестните параметри. Резултатите от стъпкова процедура за пресмятане са дадени в две таблици. В 4.1.3 са дадени и дискутирани 8 нелинейни модела, за

които са предложени доверителни интервали за водното съдържание θ . Накрая в 4.1.5 и 4.1.6 са са разгледани примери с реални данни. Получените резултати се илюстрират с интересни графики.

В Параграф 4.2 се дискутира накратко един модел, свързан с геофизиката и се търси оценка на т.н. политропен индекс при слънчев вятър. Използвани са данни, получени от спътник.

Накрая в последния Параграф 4.3 са изследвани някои статистически проблеми, свързани с тестови измервания. Въвежда се понятието надеждност на изпит и надеждност на скала и се изследат техни свойства и оценки. Дадени са интересни примери, илюстрирани с таблици и графики.

След основния текст на дисертацията следва **Заклучение** на 3 страници, където са сумирани основните постижения, групирани в 8 подточки. Считаю, че те ясно и точно отразяват най-важните постижения на дисертацията, които са на сериозно научно ниво и които се потвърждават в направения по-горе анализ.”

4. Научни приноси в работите [1-4] и [16-18], излезли от печат след дисертацията за дмн и не включени в нея.

В статия [1] е разгледан непаметричен критерий за сравнение на две извадки, основаващ се на **M**-статистика. В Теорема 2.1 е намерено точното разпределение при Леманова алтернатива в общия случай на неравни извадки. Доказателството се базира на 4 лема, доказани в работата, които имат представляват също самостоятелен интерес..

Основните резултати в статия [2] са представени в Теорема 4 и Теорема 6, където е намерено разпределението на съответната **M**-статистика, респективно при нулева хипотеза за еднородност и при Леманова алтернатива.

В работа [3] е намерено съвместното разпределение (Теорема 1) на статистики на предхождане и превъзхождане (съответно A_r и B_r). В Теорема 2 е представено съвместното разпределение при изпълнена нулева хипотеза (т.е. при еднородност на двете извадки).

Във всички тези статии освен това са направени интересни Монте-Карло симулации с графики, таблици и фигури, които добре илюстрират получените теоретични резултати. Изследвани са и мощностите на критериите. Да отбележим, че първата статия е самостоятелна, а последните две са с известния статистик Balakrishnan. Работите са публикувани в престижни статистически списания.

Статия [4] е пленарен доклад, публикуван в сборник от международна конференция в Минкс и отразяващ идеи и резултати, на които се основават работите [2] и [3].

Статиите [16-18] са определени в авторската справка като “приложно статистическо моделиране” и са резултат на ползотворно сътрудничество с три различни групи учени.

Работа [16] е свързана с механика и автоматика, [17] – с механика и устойчивост на материали, а [18] – с генетика на растенията.

Последната работа е в съавторство с наши и немски учени и е публикувана в престижно международно списание (издание на *Cambridge University Press*).

5. Анализ на предствената по конкурса документация.

Централно място във всеки конкурс заемат естествено публикациите и авторската справка.

За съжаление публикациите не са сложени на страницата на института (въпреки че ги има на PDF файлове) и поради това са трудно достъпни. Все пак има едно хартиено копие от тях и членовете на НС биха могли да се запознаят с работите при секретаря на НС.

Авторската справка по принцип е всъщност авторрецензия, а рецензентите би трябвало само да утвърдят нивото на достоверност. В нашия случай бихме могли да кажем, че авторската справка правилно и подробно отразява научните постижения в публикациите. За съжаление, четейки само справката читателят може да остане с впечатление, че в работите няма доказана нито една теорема (понеже не е цитирана нито една такава), нещо, което рецензентът се е опитал да опровергае, както се вижда от направения анализ на научните приноси.

Както се вижда от приложената справка за четени лекции и упражнения, кандидатката има много интензивна преподавателска дейност. Списъкът на научноизследователски проекти с ръководство или участие на кандидатката е също впечатляващ, както и списъкът на издадени учебници и учебни помагала – на хартия и/или електронни. За съжаление, тези важни и подробни материали не са сложени също на интернет-страницата на Института. Въобще, четейки страницата на Института, може да се помисли, че кандидатката няма не само цитируемост, но и автобиография. Да се надяваме, че поне членовете на НС ще се запознаят с тези материали от хартиения носител.

6. Заключение.

От направения анализ се вижда, че доцент дмн Евгения Стоименова има сериозни научни приноси, както в областта на теоретичната математическа статистика, така и в нейните приложения в други научни области. Тя е изявен специалист по непараметрична статистика и нейните най-важни научни и приложни резултати са публикувани в престижни международни списания с импакт-фактор. Доцент дмн Е. Стоименова има сериозна преподавателска дейност и е р-л на съвместна магистърска програма м/у ИМИ и НБУ. Била е н. консултант на 3 защитили докторанти и в момента е р-л на един аспирант.

Заслужава да се подчертае, че след защитата на дисертация за дмн, Е. Стоименова продължава възходящо своето научно развитие с нови сериозни публикации във важни международни списания.

Всичко изложено до тук ми дава основание да заключа, че доцент дмн Евгения Стоименова удовлетворява всички условия за получаване на обявената научна длъжност (звание) професор съгласно ЗРАС и правилникът на БАН и ИМИ. Ето защо предлагам научното жури да предложи на НС на ИМИ доцент дмн Евгения Стоименова да бъде избрана по обявения конкурс за **професор** в област на висше образование **4. Природни науки, математика и информатика**, научна специалност 01.01.10 „Теория на вероятностите и математическа статистика”, обявен в „Държавен вестник” бр. 58/29.07.2011 г.

03.01.2012 г.

Рецензент:

/професор дмн Николай М. Янев/