

РЕЦЕНЗИЯ

на труд за академичната длъжност „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.

Автор : доц. дмн **Евгения Асенова Стоименова, ИМИ-БАН**

РЕЦЕНЗЕНТ: проф.д.т.н. Благовест Шишков,ИМИ-БАН

Евгения Асенова Стоименова е родена на 13.07.1959 г. в гр. Пелово, Ловешка област. През 1984г. завършва висше образование по математика, специализация „Вероятности и статистика” на СУ „Климент Охридски”. През 1994 г. получава образователната и научна степен „доктор по математика”. по научната специалност 01.01.10: „Теория на вероятностите и математическа статистика”. През 2003 г. става доцент (старши научен сътрудник II ст.), Институт по математика и информатика, БАН. През 2010 г. става доктор на математическите науки по научната специалност 01.01.10 „Теория на вероятностите и математическа статистика”.

Научните интереси на кандидата се отнасят към следните области:

- Непараметрични и вероятностни модели върху пермутации: вероятностни разпределения върху пермутации, метрики в пространства на частично наредени данни, мерки за асоциация на пълно и частично наредени данни. Статистически методи за избор и нареждане: инвариантни функции на загубите, оценка на риска. Проверка на хипотези: непараметрични критерии за пълно частично наредени данни.

- Приложна статистика: нелинейна регресия, статистическо моделиране, статистически методи в геомеханиката, статистически модели в психологията и образователните измервания.

- Статистика /математическо консултиране: Консултации на учени от различни научни области (биология, медицина, инженерни науки, физика, психология, образование и др.).

На тези области е посветена дисертацията на Евгения Стоименова за доктор на математическите науки, на която имах удоволствието да бъда един от рецензентите.

В **Увода** се прави преглед на ранговите критерии, които възникват в задачите за проверка на съгласуваност или сравняване на две извадки от разпределения с неизвестен параметър на положението. Те са приложими, когато в нулевата хипотеза се предполага инвариантност на разпределението на извадката спрямо група от пермутации, докато в алтернативата такава отсъства. Статистиките на критериите са бързи и лесни за пресмятане и могат да се използват дори когато наблюденията са представени само чрез тяхната наредба. За реални случайни величини замената на стойностите на извадката с тяхната наредба е загуба на информация, която може да се определи чрез Питменовият коефициент на относителна ефективност. Някои рангови статистики притежават естествено представяне чрез разстояния върху пермутации. Съществен принос в развитието на вероятностните модели, базиращи се на метрики имат Mallows (1957), Diaconis (1988), Critchlow (1985), Marden (1995), Deza (1998) и др. Метричният подход за конструиране на рангови статистики позволява да се обобщават по естествен начин статистики, дефинирани за някоя непараметрична задача за проверка на хипотези до статистики за други непараметрични хипотези.

1. Рангови критерии за проверка на непараметрични статистически хипотези [1,2,3,4,5,6,7,8,9,10,11,12].

В този раздел се разглеждат основните непараметрични статистически хипотези за разпределението на една, две или повече извадки. Предположенията за разпределенията са общи, например непрекъснатост или монотонност, и не се свързват с определени параметрични свойства. За проверка на такива непараметрични хипотези са развити редица критерии, основаващи се на наредбата на наблюденията в извадките. Задачите, които се решават в глава първа са: дефиниране на метрики в множество от пермутации, които да се използват за дефиниране на статистически критерии; построяване на статистически критерий за две извадки, породен от Чебишевата метрика в множеството от пермутации; обобщаването на критерия за две извадки при нулевата хипотеза; установяване на съответствието между новия критерий и други рангови критерии за същите задачи; построяване на семейство от критерии за две извадки чрез обобщаване на критерия, породен от Чебишевата метрика. В глава втора са изследвани разпределенията на рангови статистики при различни алтернативи. Задачите са свързани с изследване на свойствата на критериите от първата глава при конкретизирана алтернатива. Конкретно те са: намиране на разпределението на M – статистиката за две извадки, породена от Чебишевата метрика, при Леманови алтернативи; изследване на мощността на рангови критерии при Леманови алтернативи; симулационни изследвания на мощността на рангови критерии при алтернатива на отместване.

Трудността при непараметричните задачи е, че алтернативите са твърде общи и практически рядко разпределенията могат да се изразят в явен вид. В тази глава е изследвано разпределението на рангови критерии при някои подкласове на H_A . Първият такъв подклас е когато сравняването на две разпределения с плътности f_1 и f_2 е по отношение на тяхното положение (параметър на отместване).

Използвани са Монте Карло методи за изследване на рангови критерии при такива алтернативи.

Статиите [5,6,7,8,9,10,11,12] съдържат предхождащи резултати по тематиката за рангови критерии. Резултатите от тези статии са свързани с рангови статистики, породени от метрики в пермутационната група и са основа за първите две глави на дисертацията за дмн.

Новите резултати са публикувани в [1,2,3,4]. Дължа да отбележа че идеите на Евгения Стоименова от [1] са доразвити в [2,3,4] от Stoimenova and Balakrishnan – Journal of Statistical Planning and Inference.

Статиите от тази група са посветени на непараметрични критерии за сравняване на две или повече разпределения. В задачата за две разпределения нулевата хипотеза е че двете разпределения съвпадат срещу общата алтернатива, че разпределението на втората извадка е изместено. Критерият за две извадки, изследван в [1] е предложен в по-ранната статия [10]. В статията [1] е получено точното разпределение за неравни извадки при Леманова алтернатива. Лемановата алтернатива е частен случай на алтернатива на отместване, която за някои практически важни разпределения е аналогична на параметрична хипотеза с параметър на отместване (тя включва експоненциално разпределение, разпределение на Вейбул и по-общо - разпределение с пропорционална хазартна функция). Разпределението при алтернативата е използвано за определяне на мощността на критерия. Мощността на критерия е сравнена с мощността на някои други популярни рангови критерии.

При избор на подходящи разстояния се получат редица от известните критерии, в това число статистиката на Mann-Whitney, статистиката на Wilcoxon, медианният критерий на Mood, критерия на Колмогоров-Смирнов и др. Критерият за две извадки с наредена алтернатива, породен от Чебишевата метрика, е обобщен за сравняване на две разпределения срещу двустранна алтернатива, за сравняване на повече от две разпределения срещу наредена и ненаредена алтернатива и за проверка на тренд.

2. Непараметрично оценяване на плътността на разпределение [13,14,15].

Резултатите в тези статии са свързани с непараметрично оценяване на вероятностната плътност когато данните са интервално ограничени (цензурирани). Те са част от съдържанието на трета глава на дисертацията за дмн.

Задачите се отнасят за приложение на метричния подход в статистическите задачи за избор и нареждане и за непараметрично оценяване на вероятностната плътност. Конкретно те са: построяване на статистически критерий за проверка на тренд при частично нареждане; дефиниране на функция на загубите чрез матричен подход; оценяване на функцията на риска в задачата за разпределяне в две множества; дефиниране на непараметрични оценки на вероятностната плътност при непълни данни; приложение на метрики за частично наредени данни в клъстерния анализ.

В редица статистически задачи, обект на изследване са пълните и частичните ранжировки на n обекта.

Трите работи са самостоятелни. Интервално ограничените данни са резултат от наблюдаване на непрекъсната случайна величина посредством функция, която съпоставя на всяка наблюдавана стойност интервал, който я съдържа. Такива данни възникват естествено в практиката. В статията [13] е направено обобщение на оценката на Надарая-Уотсън за оценяване на вероятностната плътност по цензурирани наблюдения. Чрез нея се коригира максимално правдоподобната оценка (МПО) в интервалите, в които оценката на плътността е занижена поради цензурирането. Оценката е състоятелна при условие за асимптотики за редки данни (sparse asymptotics), докато МПО не е. Резултатите са представени в пленарен доклад на международната конференция CDAM в Минск. В статията [14] е предложено в модифицираната оценка на Надарая-Уотсън да се използват ядра с променливи дължини. Построен е алгоритъм за интерактивно графично изследване на различните оценки. Повечето непараметрични техники (все още) не са включени в стандартните статистически програмни системи. Ето защо активните изследвания в областта на непараметричното оценяване на функции е свързано и с практическа реализация на алгоритмите. Статията е по покана на редакторите на *Austrian Journal of Statistics*. В статията [15] е направена връзка между интервално цензурираните данни и размитите числа (fuzzy numbers). Реализиран е алгоритъм интерактивно изучаване на хистограмни оценки по цензурирани данни.

3. Приложно статистическо моделиране -- [16,17,18,19,20,21,22].

Резултатите в тези статии са свързани с приложно статистическо моделиране. Те са част от съдържанието на четвърта глава на дисертацията за дмн, в която са предложени и изследвани статистически модели. Приложими в някои области на механиката, геофизиката, инженерните науки и педагогическите измервания. Специално може да се отбележи възможността за ефективни приложения в статистическата теория на комуникациите, радиолокацията и автоматичното управление. Резултатите имат научно-приложен характер. Конкретно те са: построяване на модел на характеристичната крива почва-вода (SWCC) чрез двуфазен нелинеен регресионен модел; построяване на нови модели на (SWCC) и паралелен анализ със съществуващи такива модели; статистическо оценяване на политропния индекс при слънчев вятър; статистическо моделиране на тестови измервания и оценяване на качествата на моделите.

Статиите имат научно-приложен характер и са в съавторство с колеги от съответните области. Статиите [16,17,18] не са включвани в предхождащи конкурси и дисертации.

В статията [17] е изследвано влиянието на епоксидното покритие на стоманобетонни пръти от стоманобетонни структури върху издръжливостта на връзките между бетона и арматурата. В статията [16] е изследвано влиянието на геометричните параметри на капков

емитер, състоящ се от лабиринт от канали с триъгълен профил при освобождаването на водата.

Статиите [20], [21] и [22] са посветени на модели от областта на геотехниката. Характеристичната крива почва-вода (SWCC) описва зависимостта между водното съдържание в почвата и хидравличното налягане в почвата. Известните в литературата модели имат емпиричен произход и са подходящи за различни видове почви. Моделите се ограничават до подбор на параметрична функция, описваща характеристичната крива. Тъй като зависимостта е нелинейна, оценките на параметрите се получават чрез нелинеен регресионен анализ. В статията [22] се предлага двуфазен регресионен модел за описване на SWCC. Допълнително се предполага, че точката на свързване на двете регресионни криви е неизвестна. Предложен е алгоритъм за оценяване на този допълнителен параметър заедно с параметрите на двете регресионни уравнения. В статията [21] се разглежда фамилия от нелинейни модели, които допускат линеаризация и позволяват прилагането на линеен регресионен анализ за оценка на неизвестните параметри. Популярните модели от литературата се включват в тази фамилия. Линеаризацията позволява оценка на грешката на прогнозиране на първоначалните нелинейни модели, при условие че данните от всяка конкретна почва удовлетворяват регресионните условия. Предложени са нови модели, описващи SWCC. Моделите са сравнени с класическите модели на ван Генухтен и Фредлунд. В статията [20] е предложен модел, описващ SWCC, който включва хистерезиса при двата процеса на изсушаване и намокряне. Моделът включва две независими променливи, налягане и дълбочина на уреда за отчитане на налягането. Моделът позволява да се предвижда водното съдържание и сканиращите криви при различни дълбочини.

Статията [19] е от областта на геокосмичната хидродинамика. В нея се оценява политропния индекс при слънчев вятър. Предложена е оценка на базата на регресионен модел за логаритъма на плътността на плазмата и температурата. Количественото определяне на политропния подход се състои в дефиниране на газово-динамичните уравнения за запазване на енергията на флуида, което включва флуидното налягане, плътността и политропният индекс γ , който трябва да се определи. Обикновено коефициентът γ се определя по експериментални данни. Уравнението е приложимо при предположение, че разглеждаме състен флуиден елемент или достатъчно хомогенна флуидна подобласт. Затова политропният индекс γ се оценява за всеки хомогенен участък на относителната плътност. Методът е приложен за анализ на реални онлайн данни, получавани от спътник.

В статията [18] се изследва влиянието на генотипа, условията на съхранение и др. фактори върху кълняемостта и растежа на житни растения. Статистическите задачи са породени от изследванията на група от учени генетици. Приложен е дисперсионен анализ за няколко значими задачи на генетичното изследване. Установено е, че растежа на растенията зависи от генотипа, както се вижда от високата им статистическа значимост на два главни ефекта и при двете изследвани групи от растения. Размерът на основния генотипен ефект върху параметрите на растежа зависи от потенциала на водата, което се потвърждава статистически от анализа на факторните взаимодействия.

4. Модели на тестове за постижения - [23]

Книгата ``Измерителни качества на тестове'' [23] е в обем от 176 стр., състои се от 5 глави с графики и приложения. Предназначена е главно за специалисти от областта на образованието, ангажирани с конструирането на тестове за постижения. Ролята на тестовите като инструмент за измерване на постижения налага определени изисквания за тяхното качество. Това неизбежно включва оценка на точността на измерване, което от своя страна се основава на статистически анализ на данни. Традиционната измерителна характеристика на тест - неговата надеждност, е дефинирана като компонента на класическия модел на тест. Представени са основните статистически оценки на надеждността - алфа на Кронбах,

формула на Кудер-Ричардсън и др. Важно място е отделено на надеждността на тестове, съставени от като композиция от няколко теста. Изяснена е ролята на надеждността за постигане на крайната цел в изследването на тестове - осигуряване на желана точност на измерване. Друга основна тема, подробно застъпена в книгата, е изследването на измерителните качества на компонентите на теста, тестовите задачи. Освен популярните мерки за надеждност на задача, се дискутират съвременни мерки, основани на корелационни коефициенти. В книгата са включени и редица приложни теми като методика за анализ на учебни тестове, препоръки за тълкуване на резултатите от психометричния анализ, определяне на големината на извадката, осигуряваща желаната точност на оценките на параметрите на теста и др. Книгата е станала популярна сред създателите на тестове и специалисти от областта. Тя се използва като учебно пособие в няколко магистърски курса по психология и методика на обучението по математика, физика, история, български език и др.

5. Приноси в труда за академичната длъжност „професор” по научната специалност 01.01.10 „Теория на вероятностите и математическа статистика”

1. Дефинирана е обща структура на рангови статистики, която включва няколко основни непараметрични хипотези.
2. Построен е клас от рангови статистики за проверка на пет стандартни непараметрични хипотези. Приложен е общ подход при дефинирането на критичните функции, включващ разстояния между класове от пермутации.
3. Доказани са монотонни свойства на Чебишевата метрика върху пермутации. Свойствата са използвани за изразяването на ранговите статистики в явен вид.
4. Изследвани са свойствата на рангов критерий за две извадки, в т.ч. точното и асимптотичното разпределение на статистиката на критерия при нулевата хипотеза, точното разпределение при Леманови алтернативи. Направено е обобщение на ранговите критерии на предходане и са получени точните разпределения при нулевата хипотеза и Леманови алтернативи на обобщения критерий.
5. Предложен е и е изследван модел за оценяване на вероятностната плътност по интервално ограничени данни. Предложена е ядрената оценка от типа на Надарая-Уотсън, която позволява да се получат състоятелни оценки на вероятностната маса в интервалите на цензуриране. Направена е модификация чрез ядра с променливи дължини. Построен е алгоритъм и програма на R , която го реализира за интерактивно графично изследване на различни оценки.
6. Предложена е процедура за разпределяне на множество от k независими случайни величини във две подмножества с фиксиран обем. Предложена е и е изследвана фамилия от функции на загубите за тази задача, основаващи се на метрики за частично наредени данни. Рискът (средните загъби) е изследван като функция на параметъра Беховер. Горната граница за функцията на риска е получена в интегрален вид за случайни величини от еднопараметрично семейство от разпределения с параметър на отнемване.
7. Предложена е и е изследвана фамилия от нелинейни модели, описващи зависимостта между степента на насищане и водния потенциал ($SWCC$). Моделите допускат линеаризация и позволяват да се оцени големината на грешката в прогнозата. Предложени са двуфазни нелинейни модели на $SWCC$ и алгоритъм за оценяване на параметъра на свързване на двете фази, заедно с параметрите на двете регресионни уравнения. Предложен е модел на $SWCC$, включващ хистерезиса при двата процеса на изсушаване и намокряне. Моделът позволява да се предвижда водното съдържание на сканиращите криви при различни начални условия.
8. Получена е долната граница за надеждността на тест, състоящ се от няколко компоненти, която използва коефициентите на Кронбах на отделните скали и корелациите между скалите. Изследвано е влиянието на тегловни коефициенти на отделните задачи върху надеждността на един тест, а също и влиянието на тегловни коефициенти на няколко теста върху

надеждността на съставния тест. Получените тегловни коефициенти осигуряват максимална композиционна надеждност.

6. Публикации по конкурса за „професор” и цитирания, импакт фактор , преподавателска дейност, участия в международни проекти

В конкурса Евгения Стоименова участва с 22 научни публикации(10 от които самостоятелни) и една книга „Измерителни качесрва на тестове” ,издание на Нов български университет – всичките в областта на математическата статистика.

Преподавателска дейност – ще спомена само курсовете и програмите, които Евгения Стоименова е чела и е изготвила в:

- ИМИ – БАН
- Софийски университет
- Висше училище по финанси и застраховане
- Bauhaus-Universität Weimar, Germany
- Американски университет в България
- Нов български университет

Към работите, които са цитирани от чужди автори като **Lins, Schanz, Ziemmerer, Krzanowski** могат да се посочат [10],[11],[12],[13],[14],[18],[19],[21]. Заслужава да се отбележи Статия [18] която е ситирана 5 пъти, Статия [21] която е цитирана 9 пъти и др.

Към работите, които са цитирани от наши автори могат да се посочат [15] , [16], [17], [20] и особено [23]

Общият Импакт фактор на Евгения Стоименова за натоящия конкурс е **9.159**. Тук особено заслужават да бъдат отличени Статия [2] от **Stoimenova and Balakrishnan – Journal of Statistical Planning and Inference** с IF **0.691**, Статия [10] от **E. Stoimenova, Ann. Inst. Stat. Math.** IF **0.966**, Статията [12] от **E. Stoimenova в Statistics**, IF **0.721**, Статията [18] **Landjeva S.,V. Kozrun, E. Stoimenova, B. Truberg, G. Ganeva and A. Boerner - Journal of Agricultural Science** IF **1.418**, Статията [19] **M. Kartalev, M. Dryer, K. Grigorov, and E. Stoimenova – Journal of Geophysical Research – Space Physics** IF **3.303**.

Евгения Стоименова е участвала успешно в следните по-значими международни проекти:

- 1. Statistical Data Analysis and Modeling of processes in food Science Biotechnology (Technical University Berlin/Project (2004-2006)).**
- 2. New Techniques in Space Weather Forecasting (Project of the Institute of Mechanics BAS and EOARD, Contract No. 05324) 2006.**
- 3. Study of Smart Antenna Arrays for the Wireless Communication Networks (Project of the Bulgarian-French Program for joint scientific researches –PAI-“RILA” Project No. 7/2005) 2005.**

7. Критични бележки

За реални случайни величини, замяната на стойностите на извадката с тяхната наредба е загуба на информация. Добре би било, ако бе формулиран и използван Питменовият коефициент на относителна ефективност, особено в асимптотичен план, което ни дава възможност да пресметнем каква е загубата на информация при използването на ранговите статистики.

ЗАКЛЮЧЕНИЕ

В конкурса за академичната длъжност „професор” по научна специалност 01.01.10 „Теория на вероятностите и математическа статистика”, Евгения Стоименова е представила реализацията на една мащабна изследователска програма в областта на Математическата статистика и по конкретно: „Метрични методи в непараметричните статистически задачи” Непараметричните статистически задачи и алгоритми са един от основните подходи за преодоляване на априорната неопределеност в Математическата статистика.

Трудовете на Евгения Стоименова представляват творческо научно постижение в Математическата статистика. При това тя работи извън стандартните общи преположения за „линейност”, „стационарност” и „гаусовост”.

Те съдържат теоретични обобщения и решения на големи научни и научно-приложни проблеми, които съответстват на съвременните постижения в науката.

Изследвани са нови рангови статистики, които са инвариантни, към дадена група преобразувания и робстни към изменения на разпределението за нулевата хипотеза.

Преодолени са значителни математически трудности за намирането на разпределението при алтернативата.

Определям като голямо постижение асимптотическия анализ на ранговите критерии за нулевата и особено за алтернативната хипотеза.

Приносите от 1. до 6. са научни и представляват новост в Математическата статистика.

Приносите 7. и 8. са научно-приложни и значими за Математическото моделиране.

Във всичките трудове се чувства почерка на висок професионализъм

В заключение ще отбележа, че получените в трудовете на кандидата резултати надхвърлят основните изисквания на ЗРАСРБ по отношение на претендираното звание. Въз основа на всичко това си позволявам да предложа на уважаемите членове на НС на ИМИ да гласуват за присъждането на научното звание ”ПРОФЕСОР” на Евгения Асенова Стоименова

01.12.2011
София

Рецензент:
/Проф.д.т.н. Благвест Шишков/