

**Институт по Математика и Информатика,
Българска академия на науките**

РЕЦЕНЗИЯ

по конкурс за ДОЦЕНТ в област на висше образование
4. Природни науки, математика и информатика,
професионално направление **4.5. Математика, научна
специалност 01.01.10. Теория на вероятностите и
математическа статистика** – един, за нуждите на ИМИ-
БАН със срок 2 месеца от обнародването в „Държавен
вестник“, бр. 8 от 30.01. 2015 г.

Рецензент: професор д-р Николай Михайлов Янев
Секция “Изследване на операциите,
Вероятности и Статистика”
И-т по Математика и Информатика, БАН
Тел. 979-3861

По гореописания конкурс е постъпила само една кандидатура, тази на **д-р Деян Йорданов Палежев**, асистент в секция „Изследване на операциите, Вероятности и Статистика“ към ИМИ-БАН.

а) Кратки биографични бележки.

Деян Йорданов Палежев е роден на 25.08.1972 г. в гр. Русе. От приложеното CV се разбира, че Деян Й. Палежев е завършил специалност ”математика” във ФМИ на СУ през периода 1990-1995 със специализация “математически анализ”. Представена е диплома със среден успех “Отличен 5.81” и защита на дипломна работа с “Отличен 6.00”. През 1998-2001 е продължил образованието си (Graduated Program in Mathematics) в Департамента по Математика, Boston University. През периода 2001-2006 продължава образователния процес в Yale University, Департамента по Статистика, където получава М.А. in Statistics през 2003 г., а през 2006 г. защитава успешно дисертация (PhD in Statistics) на тема: «Тест за бимодалност в многомерни пространства». Приложени са копия от оригинала на съответната диплома със заверен превод на български, а също и приравняване към образователната и научна степен «Доктор», като последната диплома е подписана от Председателя на БАН, Гл.н.с-р,

Директора на ИМИ-БАН и Председателят на журито. За съжаление в цялата тази документация никъде не е отбелязано името на научния ръководител. Липсва и аврореферат.

Д. Палежев е заемал следните образователни и изследователски позиции през периода от 1999 до сега:

- Boston University, Департамент по Математика и Статистика – Teaching Assistant (1999-2001); Instructor, Summer Term (2000-2001).

- Yale University:

- (i) Департамент по Статистика – Teaching Fellow (2003-2005); Instructor, Summer Session (2005).

- (ii) Департамент по Генетика, School of Medicine – Postdoctoral Associate (2006-2008); Associate Research Scientist (2008-2010).

- (iii) Child Study Center, School of Medicine – Associate Research Scientist (2010-2011).

- Софийски Университет, ФМИ – хоноруван лектор (2012-2013).

- ИМИ –БАН – асистент (2012 до сега). Бил е и лектор в съвместната с НБУ магистърска програма (2014).

Съгласно приложения списък, Д. Палежев е участвал в 5 научно-изследователски проекта в САЩ. В момента е член на колектив от ИМИ-БАН, които работи по проект с ФНИ към МОН (2015-2016).

Д. Палежев има публикувани 13 статии в реномирани списания, сред които *Nature*, *Science*, *Proc.Natl.Acad.Sci.USA*, със сериозна цитируемост (772 цитата съгласно SCOPUS). Дадени са заглавията на доклади в 5 конференции (2 в чужбина и 3 у нас). Има сериозна преподавателска дейност като упражнения и четене на лекции в областта на ТВ и МС.

б) Научни приноси в представените публикации.

Представените за участие в конкурса 13 статии са отпечатани в периода 2007 – 2014 години в реномирани списания с висок IF. Те могат да бъдат представени, както следва.

Статия [8] е публикувана в *Nature*, ИФ-42.351, като съавторите са общо 18, а работа [1] – в *Science*, ИФ-31.48, с 23 съавтори. Трябва да отбележим, че по-големият брой съавтори в тези списания е нещо обичайно, тъй като там обикновено се публикуват комплексни изследвания, които изискват участието на специалисти от различни научни общности. Аналогично би могло да се отбележи и за *PNAS (Proc.Natl.Acad.Sci.) USA*, ИФ-9.809, където в [2, 3, 6] съавторите са съответно 5, 12 и 8. Статиите [5, 7, 10, 12] са

публикувани в психиатрично ориентирани списания: [5] – *Neuropharmacol.*, ИФ-4.819, 7 съавтори, [7] – *Dev.Psychopathol.* ИФ-4.40, 8 съавтори, [10] – *Appl. Psycholinguist.*, ИФ-1.58, 4 съавтори, [12] – *Biol.Psychiatry*, ИФ-9.472, 10 съавтори. Работа [4] е отпечатана в *Behav. Genet.* ИФ-2.839, 13 съавтори, а [11] – в *Plant Genetic Resources*, ИФ-1.057, 6 съавтори. Статия [13] е от двама съавтори в *Stat.Appl.Genet.Molec.Biol.*, ИФ-1.055, а статия [9] е самостоятелна и публикувана в *Serdica J. Computing*.

Посочените импакт фактори са от 2013 г. (последните известни). Публикациите [6-8] и [10-13] са излезли след получаване на степента PhD и дори след заемането на академичната длъжност “асистент” в ИМИ-БАН. Публикации [11-13] са излезли след признаването на научно-изследователската степен “доктор” от БАН. Работа [10] е излязла от печат след признаването от БАН на научно-изследователската степен “доктор”, но е приета и публикувана онлайн преди това. Публикация [9] е базирана на дисертацията, но е публикувана след получаването на научната степен PhD.

Всички представени работи съответстват на темата на конкурса и се приемат за рецензиране. Най-общо казано тези публикации могат да бъдат класифицирани като развитие и приложение на статистически и информатични методи в различни области на биологията. Прието е днес тези направления да се наричат съответно биостатистика и биоинформатика, като често границата между тях не е ясно очертана. Но безспорно е, че днес те бурно се развиват. Една от причините е, че сложните биологични проблеми привличат все повече квалифицирани математици и информатици. Този процес се наблюдава особено активно в областта на клетъчната и молекулярна биология, а също и в областта на генетиката. Изключително важни са изследванията, свързани със стволови клетки и ДНК/РНК анализи. Безспорно, че всеки математик и информатик, който иска да работи в тези области, трябва да навлезе в нова за себе си биологична проблематика, да усвои един нов език, за да може да комуникира и да работи със съответните специалисти от тази област. Изследванията на молекулярно ниво са много трудни и свързани със сложни технологии, а математическото моделиране и съответните анализи имат комплексен характер и изискват субординация на специалисти от различни области. При това често в съответните публикации математическите и информатични проблеми остават на втори план, като се представят най-вече

някои таблици и графики, а се акцентира на биологическите изследвания и съответните биологически изводи.

Такъв пример са публикации [2, 3, 4, 6], където се използват *microarrays*, т.е. анализ на данни от ДНК и РНК микрочипове. Да отбележим, че [2, 3, 6] са публикувани в *PNAS*, а [4] – в *Behav. Genet.* За съжаление на тези интересни работи е отделено недостатъчно внимание в авторската справка. Все пак е отбелязано, че в [4] е използван методът *elastic net* и са идентифицирани 55 гена, чиято пропорция е по-голяма от тази в една случайна извадка. За статия [2] в авторската справка добре е дадена същността на метода за амплификация на генетичен материал, но не са посочени методите за оценка, които се използват в работата. В работа [3] става въпрос за генна експресия на две клетъчни линии. Показано е, че в присъствието на антигена EBNA 1 статистическата значимост на средните нива е по-висока. Посочени са и други ефекти. Интересна е работа [6], където се използват РНК микрочипове, като се изследват т.н. iPS клетки (т.е. индуцирани плурипотентни стволови клетки) и връзката им с централната нервна система.

Стволовите клетки iPS са обект на изследване и в статиите [5, 8]. Интересна е работа [5], публикувана в *Neuropharmacol*, която има по-скоро обзорец характер и показва насоки за приложение в някои нервно-психични заболявания. В статия [8], отпечатана в *Nature*, се анализират данни от ДНК секвениране на iPS клетки, т.е. разглеждат се варианти с различен брой копия (структурни варианти) в сравнение с референтен геном. Установен е факт на т.н. репрограмиране, т.е. не се наблюдават нови варианти. Чрез допълнително РНК секвениране е показана близост до стандартната линия от iPS клетки (от гледна точка на генната експресия).

Всъщност статия [1], публикувана в *Science*, е една от пионерските работи, в която е развит метода на ДНК секвениране. Използват се различни биоинформатични и биостатистически методи за намиране координатите на секвенираните сегменти, като точните области за отхвърляне се определят непараметрично. Освен това се използват и някои параметрични методи.

В [10, 11] се изследват популации с различни фенотипни признаци от различни райони. Така в работа [10], публикувана в *Appl. Psycholinguistics*, използвайки някои статистически критерии, е показано, че честотата на известни речеви нарушения в изследвания географски район

е по-голяма отколкото в контролния район. В работа [11], публикувана в *Plant Genetic Resources*, се изследва генетичното разнообразие в два черноморски региона, използвайки различни статистически методи. Изследвани са 52 сорта пшеница и се използват генетични данни от 263 алели, избрани от 25 микросателитни локуса. Прилагат се някои метрики за измерване на генетичното разнообразие и се използва клъстерен анализ. Намерени са интересни сравнения.

В статиите [7, 12] се работи с данни от секвениране на РНК. Най-общо казано, при секвенирането се извършва разпознаване (употребява се и термина «четене») на основните нуклеогиди (т.е. на първичната структура) на ДНК или РНК. В работа [7], публикувана в *Dev.Psychopathol.*, освен това се изследват измененията в генната експресия, зависещи от възрастта. Резултатите от статистическия анализ, приведени в Таблица 1, показват съществени различия в изследваните три когнитивни процеса. Статия [12] е публикувана в *Biol.Psychiatry* и в нея се намират статистически значими разлики в генната експресия от носители на Синдром на Турет и контролна група от здрави индивиди.

Публикацията [9] е отпечатана в *Serdica J. Computing* и има чисто теоретичен статистически характер. Основната цел, реализирана с работата, е в конструирането на тест за многомерна нормална унимодалност срещу смес от две сферични нормални разпределения при определени условия на голяма размерност (Теорема 3.2). Доказателството се основава на няколко леми, които освен това представляват самостоятелен интерес. Изследва се асимптотичното поведение на „Хи-квадрат“ разпределения с d степени на свобода, когато d расте неограничено. При определени условия се намира и асимптотична апроксимация на максимумите. Накрая се предлага и по-опростен асимптотично еквивалентен вариант (Теорема 4.1) като приведените симулации показват добри резултати.

Статия [13] е публикувана в *Stat.Appl.Genet.Molec.Biol.* и в нея се разглеждат някои проблеми, свързани с така наречените „p-value“ изследвания, въведени в модерната статистика от Пирсън и Фишер. В работата са разгледани някои непараметрични и параметрични подходи при многократно тестиране, каквито ситуации често се срещат в биологията. Интересна е предложената трансформация (7) за корекция разпределението на p -values. Получените статистически резултати се прилагат при РНК анализи, като

е посочено, че по този начин може да се намали и големината на извадката.

Авторската справка правилно отразява основните научни приноси във всички представени работи, които, както вече беше отбелязано, са на високо научно ниво. За това говори и фактът, че сумарният импакт фактор на 13-те публикации е 128.48, а списъкът на цитиранията е внушителен – 772, като цитиранията са дадени поотделно за представените работи. От Авторската справка личи, че д-р Д.Палежев е навлезъл дълбоко в една сложна проблематика от Молекулярната биология и генетика, свързана с първичната структура на ДНК и РНК, което е естествено задължително условие за успешно математическо моделиране в тази област. За съжаление в авторската справка не е отделено достатъчно подобаващо внимание на биостатистическите и биоинформатичните методи, които се прилагат в статиите. Освен това към повечето статии [1,2,3,6,7,8,11,12,13] има т.н. Supporting Information, която представлява допълнителен интерес, но това също не е отразено в Авторската справка. Въобще Авторската справка трябва да бъде като една подробна авторска справка, при това специално насочена към специфичната тема на конкурса.

в) Учебно-преподавателска дейност.

От приложената справка се вижда, че д-р Д. Палежев е имал сериозна преподавателска дейност и е водил разнообразни курсове в следните университети:

Boston University - Статистика I; Анализ I.

Yale University – Въведение в Статистиката; ТВ; МС; Стохастични процеси.

ФМИ – СУ – Статистика.

ИМИ-БАН – НБУ – Статистически методи.

е) Лични впечатления.

Д-р Д. Палежев е изнасял доклади на Националния семинар по стохастика, а също и на Международната конференция по ТВ и МС (Поморие, 2014). Освен това е изнасял серия от доклади по Биостатистика на работен семинар в ИМИ. Докладите бяха на високо професионално ниво.

ж) Заключение.

От направения анализ се вижда, че д-р Деян Йорданов Палежев има сериозни научни приноси, които съответствуват на темата на конкурса. Той е защитил публично и успешно дисертация за придобиване на научната степен “Доктор” в областта на обявения конкурс в Yale

University през 2006 като тази степен е приравнена към о.н.с. “Доктор” в БАН. Освен това неговите най-важни научни и приложни резултати са публикувани в престижни международни списания с висок импакт-фактор и имат много сериозна цитируемост. Д-р Д. Палежев е имал и сериозна преподавателска дейност. Трябва изрично да се отбележи, че наукометричните данни на д-р Д. Палежев, както и тези за преподавателска дейност, надхвърлят изискванията, които се предлагат от Правилника на ИМИ – БАН.

Всичко изложено до тук ми дава основание да заключа, че д-р Деян Йорданов Палежев удовлетворява напълно всички условия за получаване на обявената научна длъжност (звание) доцент съгласно ЗРАС и съответните правилници.

01.06.2015 г.

Рецензент:

/професор дмн Николай М. Янев/