

Кратко описание на най-важните постижения на трудовете на проф. дмн Евгения Стоименова

Научните ми изследвания са основно в областта на математическата статистика и по-конкретно в развитието на нови модели за непараметрични статистически задачи. В тази област е дисертацията ми за доктор на науките и последващите изследвания с водещи чуждестранни учени и докторанти. За конкурса са представени 28 статии от тази област. Съществена част от другите ми научни интереси са от областта на приложното статистическо моделиране. Резултатите са свързани със статистически модели, приложими в някои области на механиката и инженерните науки. Статиите имат научно-приложен характер и са в съавторство с колеги от съответните области. За конкурса са представени 12 статии от тази област. Статистическите модели в областта на психометрията също са обект на моите научни интереси. Представени са две статии върху математически модели и една монография, станала популярна за сред създателите на тестове за постижения и специалисти от областта на психометрията.

Рангови критерии за сравняване на две разпределения. Голяма част от изследванията ми са посветени на непараметрични критерии за сравняване на две или повече разпределения. В задачата за сравняване на две разпределения нулевата хипотеза е че двете разпределения съвпадат срещу общата алтернатива, че разпределението на втората извадка е изместено, например наляво, по отношение на разпределението на първата извадка. Без предположения за конкретен вид на разпределенията, статистиките на критериите за проверка на тези хипотези могат да се основават само на ранговете на обединената извадка.

Построили сме клас от рангови статистики, породени от Чебишевата метрика върху пермутации. Получено е разпределението на ранговата статистика за сравняване на две разпределения, която е породена от Чебишевата метрика. Изследвани са асимптотичните свойства на ранговата статистика при нулевата хипотеза. Резултатите са приложени за доказване на аналогични свойства на E -статистиката на Шидак. Точното разпределение на статистиката на критерия е получено за 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 за разпределения с пропорционална хазартна функция. Разпределението при алтернативата е използвано за определяне на мощността на критерия.

Рангови критерии, основани на статистики на предхождане и превъзхождане, за сравняване на две разпределения. Рангови статистики, които се базират на броя на наблюденията от една извадка, които предхождат (превъзхождат) някоя фиксирана стойност или случайна величина наричаме рангови статистики на предхождане (превъзхождане). При подходящо реформулиране някои от ранговите критерии за две извадки, дефинирани от нас, са еквивалентни на други рангови критерии от статистическата литература. В обичайните изследвания в литературата нивата и статистиките са независими. Разглежданите от нас статистики са зависими, и за да могат да се прилагат в статистически модели, се изисква явен вид на тяхното съвместно разпределение. Изследвано е съвместното разпределение на статистики на предхождане и превъзхождане при две извадки. Изведено е разпределението на статистиката с нива определени от фиксирани порядкови статистики при еднакви разпределения на двете извадки.

Предложени са и са изследвани класове от рангови статистики, които са обобщение на статистиките на предхождане (Precedence tests). Изследвали сме нов клас от рангови критерии за

проверка на хомогенност на две извадки от две разпределения срещу едностранна алтернатива, че втората случайна величина е стохастично по-голяма от първата. Изведено е разпределението на статистиката на критерия при нулевата хипотеза и при Леманова алтернатива. Изследвана е мощността на критериите, основан на Чебишевската метрика. Друг клас от рангови критерии, който сме изследвали, са обобщение на критерий, изследван от Тюки, Хайек и Шидак. Направено е сравнително изследване на мощността на класове от критерии основани на статистики на предхождане и превъзхождане.

Статистически модели върху пермутации. Моделите, които сме изследвали се базират на разстояния върху пермутации и използват съществено статистическите свойства на тези разстояния. Изследвали сме статистическите свойства на Чебишевската метрика и метриката на Ли. Изследван е модел от тип на Малоус на разпределение върху пермутации основан на метриката на Ли. Построен итеративен алгоритъм за оценяване на параметрите в модела на Малоус. Метрика с фиксиран вектор е използвана за дефиниране на статистика на непараметричен критерий за проверка на тренд при частично нареждане. Доказано е асимптотичното разпределение на статистиката на критерия.

Изследвани модели на близостта на частични нареждания чрез метризиране на пространства от десни съседни класове на пермутационната група. Въведена е метрика за частични нареждания, породена от Чебишевската метрика върху пермутации, която е алтернатива на стандартната метрика за пълни нареждания (пермутации). Изследвани са статистическите свойства на метриката при равномерно разпределение на частичните ранжировки.

Статистически модели с използване порядкови статистики. В теорията на надеждността е популярна система k -out-of- n , състояща се от n компоненти и функционираща, тогава и само тогава, когато работят поне k от нейните компоненти. Изследвали сме надеждност на такава система с единичен резервен компонент в режим на готовност. Получили сме резултати, описващи остатъчния живот и неговата средна стойност в случаите, когато времената на живот на компонентите се представят чрез независими случайни величини с дискретно разпределение на времената на живот като активните компоненти са хомогенни. Резултатите са валидни и при по-общи условия на произволно съвместно разпределение на времената на живот.

Непараметрично оценяване на плътността на разпределение. Резултатите са свързани с непараметрично оценяване на вероятностната плътност когато данните са интервално ограничени (цензурирани). Интервално ограничените данни са резултат от наблюдаване на непрекъснатата случайна величина посредством функция, която съпоставя на всяка наблюдавана стойност интервал, който я съдържа. Направено е обобщение на оценката на Надарая-Уотсън за оценяване на вероятностната плътност по цензурирани наблюдения. Предложената модифицирана оценка на Надарая-Уотсън използва ядра с променливи дължини. 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*). Построен е алгоритъм за интерактивно графично изследване на различните оценки.

Статистически модели за изследване на зависимости. Изследвани регресионни модели с предполагаема зависимост между предикторните променливи. Предложен е нов подход за намаляване броя на променливите в многомерна линейна регресия с зависимост между предикторните променливи. Идеята е да се конструират последователни регресионни модели, които се подбират по критерии с крос-валидация. Избраните променливи в модела имат по-слаба корелация и са по-устойчиви към входните данни по отношение на грешки извън извадката (*out-of-sample errors*).

Изведена е формула за пресмятане на средноквадратичната грешка след крос-валидация.

Моделите от областта на геотехниката. Изследвана е поредицата от тестове за последователно изсушаване и намокряне в експериментално устройство с пясъчен стълб. Измерванията на съдържанието на вода, както и налягането в порите са пряко свързани и се описват с кривата SWRC (soil-water retention curve) на задържане на вода в почвата. Въз основа на експерименталните резултати е предложен нов модел на SWRC за задържане на вода в почвата. Сравнението между измерените данни и резултатите от числената симулация на преноса въздух-вода в ненаситени среди показват, че предложеният модел може да подобри точността при симулиране на водния пренос в ненаситени среди.

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.

Моделите на тестове за постижения. Построен вероятностен модел на изпит, включващ вероятностни разпределения в три множества – на учениците, изпитите и постиженията. Показана е връзката между надеждността и корелацията между два еквивалентни изпита. Доказано е, че коефициентът на Кронбах е долна граница за надеждността. Изследвана е нова долна граница за надеждността на изпит, състоящ се от няколко компоненти. Оценката използва коефициентите на Кронбах на отделните скали и корелациите между скалите. Изследвано е влиянието на теглови коефициенти на отделните задачи върху надеждността на един тест, а също и влиянието на теглови коефициенти на няколко теста върху надеждността на съставния тест. Получените теглови коефициенти осигуряват максимална композиционна надеждност. В определени случаи тази оценка (композиционна надеждност) е по-висока от коефициента на Кронбах на съставния изпит.