

Bírálat

Horváth István MTA doktora cím elnyeréséhez benyújtott értekezéséről

Horváth István MTA doktora címért benyújtott értekezését a gammakitörések tárgykörében, a megfigyelési vonatkozások statisztikai értelmezéséről írta. A tudományterület az utóbbi négy-öt évtizedben alakult ki, de elegendően nagy számban detektált, kellően informatív adatok csak ezen időszak későbbi felétől-harmadától állnak rendelkezésre. A dolgozat témájának időszerűsége vitán felül áll, a bemutatott eredmények újszerűek és fontosak.

Ennek megfelelően javaslom a nagydoktori értekezés elfogadását és nyilvános vitára bocsátását.

A részletes értékelés és a kérdéseim a következők.

A kutatási háttér bemutatása

A dolgozat értékéhez nagyban hozzájárul, hogy a magyar nyelvű irodalom relatív szűkössége miatt átfogó diszciplináris alapozás is szükséges. Ezt a feladatot Horváth István sikeresen megoldotta, a bevezető rész különösen olvasmányos, jól összefogott szerkezetűre sikerült, érzésem szerint az oktatási és didaktikai szempontokat sem figyelmen kívül hagyva.

Ezért zavaró, hogy a kozmológiai modelltől függő átszámítások a dolgozatban úgy szerepelnek, hogy a háttérben álló kozmológiai modellre nem történik hivatkozás. Általában az átszámítások modellfüggése sem kerül megemlítésre. Erre példa, hogy a 13. oldalon a vöröseltolódás alapján fényút-távolság származtatására gondolhatunk, de nem tudjuk, hogy milyen modellben. A 20. oldalon is pontatlanra sikeredett az abszolút fényességek eloszlásának bevezetése, hiszen a mért fényesség függése a kozmológiai modelltől itt sem kerül megemlítésre. Eukleidészi geometriában érvényes diszkussziót kapunk, ami lehet egy jól kezelhető határeset, de nem tudjuk meg, hogy mennyire jó. Később, az 5.1.10. ábrán nem derül ki, hogy a fényességek kozmológiai z-korrekciójára is sor került-e, ezt is szerencsés lett volna kifejtteni.

A válaszmátrixok invertálásakor nem kerül említésre a zaj és a mintavételezettség hatása a detektálási formulában (5.1.1.), és nem olvasunk ennek lehetséges kezeléséről a fejezet további részében. Az invertálásra több eljárást ismertet referencia szintjén, ami nagyon szerencsés. Azonban hiányosságnak tartom, hogy csak a szintézises módszer alapjait mutatja be. A szintézis a mintavételezettséget és a zajt eléggé természetes módon kezeli, ez azonban nem ennyire automatikus a többi módszer esetében. Az alkalmazhatóság ebből következő korlátaira fel is hívja a figyelmet a dolgozat.

Ezzel együtt a bevezetés jól fókuszált és értékes anyag. A fentebb említett hiányosságok nagyjából didaktikai jellegűek, és különösen a didaktikai szempontok felmerülése esetén ajánlom ezeket megfontolásra.

Alkalmazott módszerek:

A dolgozatban haladó statisztikai módszerekkel végzett analízisekről olvashatunk. Horváth István kiemelkedően széles eszköztárról tesz tanúbizonyságot, a módszereket jól ismeri és nagy rutinnal, helyesen alkalmazza.

Inkább didaktikai jelentősége van a szignifikancia, valószínűség, likelihood hányados fogalmak olykori keveredésének, ezekre tételesen nem térek ki, de a dolgozat számos helyén felbukkan a keveredésük.

Az 50-53. oldalon log likelihood próba leírását olvassuk, amelynek szignifikanciáját – a mai kor követelményeinek megfelelően – parametrikus bootstrap segítségével erősíti meg. Felvethető azonban, hogy mivel parametrizálunk, vizsgálni kellene a bootstrap stabilitását a paramétermeghatározástól. Az adott hiányosság a konklúziót valószínűleg nem változtatja meg, mert a csoportok jól szeparálódnak, ez esetben viszont erre lehetett volna felhívni a figyelmet.

A 89. oldalon a kvadрупól tagok analízisének Student-féle t-tesztet találunk. Ezzel a teszttel mintaátlagokat lehet összevetni. Nem kerül kifejtésre, hogy a kvadрупól amplitúdók vizsgálatára ez a teszt hogyan alkalmazható. A t-teszt viszont feltételezi a normalitást, ha a feltételezés az alkalmazott implementációban nem áll meg, a normalitástól független alternatíva az u-teszt lenne. Elemszámok analízisére viszont inkább nem mintaátlagokat, hanem kategóriák populánságát szoktak összehasonlítani, és erre a Fisher-féle egzakt teszt valamilyen módosítása lehetne a legcélszerűbb eszköz. A bemutatott Bernoulli-teszt minden szempontból egzakt és konklúzív.

A 9.2.1. fejezetben a Peacock-féle többdimenziós teszt implementációját olvashatjuk. A leírás alapján a teszt helyesen lett végrehajtva. Azonban a Peacock-teszt a KS-teszt nem teljesen problémamentes kiterjesztési kísérlete, az eredmény függ a koordináta-rendszer megválasztásától és a tengelyek forgatásától. Sokkal jobb tulajdonságai vannak más megközelítéseknek. Ilyen például az információs távolság tesztje (Székely-Rizzo-teszt), amely tetszőleges dimenziójú mérhető térben adja kedvező erejű próba lehetőségét az eloszlásfüggvény tesztelésére, így például a homogenitás vizsgálatára is. Érdekes lenne ezzel is megismételni a vizsgálatokat.

A 9.2.3. fejezetben leírt teszt, ami a lefedett legnagyobb elemszámokat hasonlítja össze, olyan értelemben meggyőző, hogy bootstrap alapú kiértékelés alapján szignifikáns. Itt azonban a vizsgált statisztikai változó nem robusztus, ami numerikus instabilitások forrása lehet, és ezek tesztelése nélkül a teszt konkluzivitása vitatható. Itt valóban a k-szomszédok vizsgálata jelenthet robusztus alternatívát, amit a dolgozat helyesen mutat be.

4. Az eredmények értékelése

A disszertáció legfontosabb eredményének a közepes időtartamú kitörések azonosítását és ezek alapvető tulajdonságának meghatározását tartom. Szintén fontos eredmény az $z=1,6-2,1$ intervallumba eső kitörések anizotrópiájának kimutatása, amelyet validál egyfajta „teszt minta” is: a publikálás óta detektált GRB-k adott csoportja kvalitatíve és kvantitatíve is követi a talált anizotrópiát.

A tézispontokat négy nagyobb egységbe csoportosíthatónak érzem: a közepes kitörések azonosítását időtartam, időtartam-keményiség terekben és vöröseltolódás szerint, valamint a $z=1,6-2,1$ csoport anizotrópiájának kimutatását. A dolgozat ennél több – hét – tézispontot tartalmaz, annak megfelelően, hogy más-más adatsorokon is megismétlődtek hasonló vizsgálatok az elmúlt másfél-két évtized alatt. A tézispontokat tartalmilag elfogadom.

6. Kérdéseim a következők.

1. Az 5.1.11. ábrával (Tagliaferri és mtsai, 2015 munkájának bemutatásával) kapcsolatban hiányolom az I1 és I2 jelölések feloldását. Az utófényre a V és R tartományban limiteket közölnek az idézett szerzők, 2 magnitúdó körüli R-I színindexszel. Ez vajon mennyire informatív, tehát például I1=22,2 magnitúdó esetén az R fényességet milyen tartományban várnánk, 1 napos utófény esetén?

2. A 6. fejezetben olvassuk a ML módszer első alkalmazását a dolgozatban. Az ML módszerek számos előnye mellett két legfontosabb hátránya, hogy csak aszimptotikusan torzításmentesek, és relatíve lassan konvergálnak. Milyen értelemben „legjobb becslés” (az 50. oldal terminusának

megfelelően) az alkalmazott megközelítés?

3. A 6.1.2. ábrán impresszióm szerint túl sok háttér van levonva, a kitörés előtti meredekség szignifikánsan negatív. A 6.1.3. ábrán a kitörés előtti szakasz túl rövid, a háttér levonás nem ítélné meg szemre. Ez az impresszió azonban fölveti az általános kérdést: a lehetséges hibaforrásokat figyelembe véve, milyen módszerrel lehet T90 hibáját meghatározni? Hogyan függ ez a detektált fluxustól? Vannak-e T90 mérésének szisztematikussai?

4. A 7.1. fejezetben a dolgozat a három csoport létét azzal a hallgatólagos feltételezéssel vizsgálja, hogy mindhárom csoport két dimenziós lognorm eloszlást követ. Az alkalmazott próba relevanciája azon múlik, hogy teljesül-e a normalitás, ami viszont nem nyilvánvaló. A 7.5.1. ábra mintha éppen azt sugallná, hogy a harmadik csoportként az algoritmus nem a közepes időtartamú kitöréseket azonosítja, hanem a rövid és hosszú kitörések átfedésénél, a közepes csoport szárnyain létrejövő többletet. A kijelölt csoport szelektivitása igazából meglehetősen gyengének tűnik a közepes kitörésekre. Ezzel mintegy egybecsengően, a 86. oldalon olvashatjuk Zitouni és mtsai (2015) eredményeit, akik egy bővebb adatsoron is azt találták, hogy az azonosítható csoportok száma függ a normalitás feltételezésétől. Kérem a feltételeknek és a 7.5.1. ábra tartalmának részletesebb kifejtését.

5. A dolgozat záró fejezete beszámol arról, hogy a közepes időtartamú kitörések értelmezése 1998 óta vitatott kéréskör. Kérem, hogy a mostani állás szerint ismertesse a különböző ezzel kapcsolatos irodalmi álláspontokat, és a különböző álláspontok elterjedtségét.

Szombathely, 2017. március 13.

Dr. Szabó M. Gyula