

Tárgy: Válasz Prof. Dr. Dobos Imre az MTA Doktora „Török Ádám: A személygépjármű közlekedés környezetterhelésének becslése városi környezetben” című MTA doktori értekezés bírálatra

Prof. Dr. Dobos Imre

MTA doktora

Tisztelt Prof. Dr. Dobos Imre Úr!

2021. március 31-én kelt bírálatát az MTA Doktori Tanácsának szabályzatának megfelelően megkaptam.

Köszönöm, hogy Bírálóm az értekezésemet értékesnek, a statisztikai elemzésem megbízhatónak, a leírt elemzéseket érthetőnek, követhetőnek, átláthatónak, és jól is olvashatónak tartja.

Kérdés: A 2. táblázat illeszkedésvizsgálatát hogyan értelmezzük? Miért kellett öt különböző módszert kipróbálni? Nem lett volna egy nagyon megbízható teszt, mint a Kolmogorov-Szmirnov-teszt elég? Továbbá hogyan dönthető el, hogy az egyes tesztek melyik a legjobb illeszkedés? Mi történik akkor, ha a módszerek más eredményt adnak, azaz nem egyöntetűek? Akkor mi alapján választhatunk egy megfelelő eloszlást?

Válasz: Köszönöm Bírálóm tisztázó kérdését. A 2. táblázatban az egyik azonosított forgalmi szituáció, mint részhalmaz, illetve példaként kiragadott 5 eloszlás típus került bemutatásra. A különböző statisztikai tesztek különböző minta és eloszlás jellemzők mellett alkalmazhatók hatékonyan. Az elérhető szakirodalom alaposan bemutatja az alkalmazhatóság korlátait (Razali and Wah, 2011). Egyes tesztek (mint pl. Anderson-Darling próba) érzékenyebbek az eloszlás szélére, míg más tesztek kis méretű, vagy asszimmetrikus minták esetében teljesítenek rosszabbul (Kolmogorov-Szmirnov). Fentiek tükrében, a tesztek kombinált alkalmazása hozzájárul a levont következtetés megbízhatóságának további növeléséhez. A Kolmogorov-Szmirnov próba a statisztikai következtetés olyan módszere, amely nem tartalmaz feltevést a vizsgált populáció eloszlásáról, azon nullhipotézis ellenőrzésére, hogy egy minta eloszlásfüggvénye megegyezik-e egy adott függvénnyel (Mondal, Gupta, 2021), (Abdulmawjoud, Jamel, Al-Taei, 2021), (Wang et al., 2021). A Kolmogorov-Szmirnov-próba lehetséges alternatívái a Cramér-von Mises próba, ami egy és két mintás esetre is alkalmas (Zhou et al, 2016), vagy az Anderson-Darling próba, ami csak az egymintás esetre (Mudigonda, Ozbay, 2015). Az egymintás Cramer-von-Mises teszt az elméleti eloszlás és az empirikusan megfigyelt eloszlás illeszkedését vizsgálja. Anderson-Darling próba, az elméleti eloszlás és az empirikusan megfigyelt eloszlás illeszkedését vizsgálja szintén (Stephens, 1974), (Jabari, Liu, 2013). Az Akaike információs kritérium (AIC) (Rabbouch, Saadaoui, Mraïhi, 2018) a statisztikai modellek közötti kiválasztást segíti. Az AIC lényegében az elérhető statisztikai modellek minőségének becsült mérőszáma, mivel ezek egy adott adatkészlethez kapcsolódnak, így ideális módszer a modell kiválasztására. Az AIC azt írja le, hogy a modell mennyire "illeszkedik" az adatokhoz vagy a megfigyelések halmazához. De az AIC-nek mint modellválasztó eszköznek is vannak korlátai. Például, ha a tesztelt statisztikai modellek mindegyike egyformán nem kielégítő vagy rosszul illeszkedik az adatokhoz, az AIC nem adna valós visszajelzést. A statisztikában a Bayes-i információs kritérium (BIC) segít a modellek kiválasztásában a véges modellkészlet között (Gutierrez-Osorio, Pedraza, 2020). A legalacsonyabb értékkel rendelkező modellt részesítik előnyben. A modellek illesztésekor lehetőség van paraméterek hozzáadásával növelni a valószínűséget, de ez túlillesztést eredményezhet. A BIC és az AIC is megpróbálja megoldani ezt a problémát a modellben szereplő paraméterek számára vonatkozó büntetés bevezetésével; a büntetés nagyobb a BIC-ben, mint az AIC-ben. A fentiek alapján látható, hogy a kiragadott példák alapján a tesztek az empirikus adatok és teoretikus függvények közötti eltérést mérik. Minél kisebb az érték annál jobb az illeszkedés. Az öt teszt együttes alkalmazásával konfliktus nélkül sikerült azonosítani a forgalmi állapotokhoz tartozó empirikus adatokra illesztendő teoretikus függvényt. A vonatkozó táblázatban az egyes próbák a teoretikus függvények és az empirikus adatok közötti távolságokat, az illesztés jóságát reprezentálják különböző metrikák szerint (Kolmogorov-Szmirnov próbánál a távolságok szuprénuma; a Cramer-von Mises próbánál a távolságok négyzete; az Anderson-Darling próbánál a távolságok súlyozott négyzete; AIC szintén a távolságokat adja meg; BIC szintén a távolságokat adja meg)

Kérdés: A 3. táblázat nehezen értelmezhető. Miért szerepel a „szabad áramlás” kétszer, és miért kellett az amerikai és német szabványok szerint azt még tovább bontani? Nem lehetett volna áttekinthetőbben megszerkeszteni a táblázatot? A 16. és 17. lábjegyzet is hiányzik.

Válasz: *Bírálómnak igaza van, a szabad áramlás valóban kétszer szerepel, abból adódóan, hogy eltérő eloszlás típust tudtam illeszteni a forgalmi helyzetet leíró empirikus adatokra (az egyik esetben normális eloszlást, a másik esetben log-normális eloszlást). Az amerikai és német szabványokban definiált városi közlekedési forgalmi helyzetek és útkapacitás-kihasználtság szolgáltatási szintjei, valamint a mért és illesztett eredményeket szerettem volna egy összefoglaló táblázatban megjeleníteni. Így a 3. táblázat 1-7 oszlopa az empirikus és illesztett eredményeket tartalmazza, 8-9. oszlop az ezeknek megfeleltethető amerikai szabványt, a 10-11. oszlop pedig az ezeknek megfeleltethető német szabványt kívánja bemutatni. A 16. és 17. lábjegyzet valóban hiányzik, elnézést, nem töröltem, tartalma felkerült a táblázat alá forrásmegjelölésként. 16. lábjegyzet: Tunver et al., 2010; 17. lábjegyzet: Schmietendorf 2010.*

Kérdés: Nehezen értelmezhető az is, hogy a forgalmi szituációk hogyan alakulnak időben, azaz milyen időintervallumokra vonatkoznak a nap folyamán. Egy-egy forgalmi helyzet hány alkalommal fordulhat elő egy nap folyamán. Laikusként azt gondolnám akár többször, de erről a leírtak nem győztek meg.

Válasz: *Bírálómnak észrevétele helytálló. Az értekezés terjedelmi korlátai nem tették lehetővé a forgalmi szituációk azonosításán túl, azok dinamikájának vizsgálatát. Az azonosított és vizsgált forgalmi helyzetek természetesen egy nap akár többször, több helyen is kialakulhatnak, kialakultak. További kutatási cél lehet az azonosított forgalmi helyzet dinamikájának vizsgálata.*

Megjegyzés: A 3. fejezetben arra történő utalás, hogy a modellek minél több paramétert és/vagy változót alkalmaznak annál pontosabbak, trivialisnak hat. Ugyanis minél bonyolultabb egy modellrendszer, annál nagyobb számításigénye is van.

Válasz: *Köszönöm Bírálóm megállapítását, hogy a probléma és annak modellezése korrekt, az eredmények jól értelmezhetők. A bíráló megállapításán túl célszerű még megemlíteni, hogy ha modellek minél több paramétert és/vagy változót alkalmaznak annál pontosabbak ugyan, de minél bonyolultabb egy modellrendszer, annál nagyobb számításigénye is van. A két tendencia egymás ellen hat a véges számítási kapacitás okán, így meghatározható egy optimális számítási kapacitás, és a hozzá tartozó optimális bonyolultságú rendszer.*

Megjegyzés: A bemutatott (4.3) és (4.4) modell, és az abból származtatható következtetések ugyanakkor helytállóak.

Válasz: *Köszönöm Bírálóm megállapítását.*

Megjegyzés: A bioetanol és biodízel anyagok csak kevésbé befolyásolják a zajszennyezést, amennyiben ezt is egyfajta emisszióknak tekintjük. A fejezetben képletezési hiba van: két darab (5.1) egyenlet szám van.

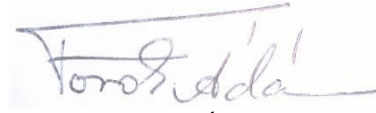
Válasz: *Köszönöm Bírálóm megállapítását, hogy elfogadja eredményeimet. Elgépelés történt. Helyesen a második (5.1)-es egyenletnek (5.3)-nak kellene lennie!*

Kérdés: A függelék I. és II. mellékletében lévő táblázatok áttekinthetetlenek. Szerencsés lett volna a táblázatokban a számokat három tizedesjegyben felülről korlátozni. Miért éppen ezt az öt függvényillesztés jóságának mértékét választotta a jelölt?

Válasz: *Köszönöm Bírálóm kritikai észrevételét, a függelékekkel kapcsolatban, valóban elegendő lett volna a három tizedesjegy alkalmazása, elnézést. A legnagyobb valószínűségi illesztés jóságának ellenőrzésére általánosan a Kolmogorov-Szmirnov próba, a Cramér-von Mises próba, az Anderson-Darling-próba, az Akaike információs kritérium (AIC) és a Bayes-i információs kritérium (BIC) az elfogadott ellenőrzési próbák. Segítségükkel akár több modell is összehasonlítható az empirikus adatokkal. Itt kell megemlíteni, hogy a szakirodalom más próbákat is ismer az illesztés jóságának ellenőrzésére (Khi-négyzet teszt, Schwarz-információs kritérium; Hannan – Quinn információs kritérium; Kullback – Leibler-divergencia; Jensen – Shannon divergencia), azonban ezek korlátozó feltételeik miatt általánosan nem elterjedtek.*

Budapest,

Tisztelettel: 2021. április 10.



Török Ádám
pályázó

Felhasznált Irodalom

- Abdulmawjoud, A. A., Jamel, M. G., & Al-Taei, A. A. (2021). Traffic flow parameters development modelling at traffic calming measures located on arterial roads. *Ain Shams Engineering Journal*, 12(1), 437-444.
- Gutierrez-Osorio, C., & Pedraza, C. (2020). Modern data sources and techniques for analysis and forecast of road accidents: a review. *Journal of traffic and transportation engineering (English edition)*, Volume 7, Issue 4, August 2020, Pages 432-446
- Jabari, S. E., & Liu, H. X. (2013). A stochastic model of traffic flow: Gaussian approximation and estimation. *Transportation Research Part B: Methodological*, 47, 15-41.
- Mondal, S., & Gupta, A. (2021). Speed distribution for interrupted flow facility under mixed traffic. *Physica A: Statistical Mechanics and its Applications*, 570, 125798.
- Mudigonda, S., & Ozbay, K. (2015). Robust calibration of macroscopic traffic simulation models using stochastic collocation. *Transportation Research Procedia*, 9, 1-20.
- Rabbouch, H., Saadaoui, F., & Mraïhi, R. (2018). A vision-based statistical methodology for automatically modeling continuous urban traffic flows. *Advanced Engineering Informatics*, 38, 392-403.
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of statistical modeling and analytics*, 2(1), 21-33.
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347), 730-737.
- Wang, J., Zhang, Z., Liu, F., & Lu, G. (2021). Investigating heterogeneous car-following behaviors of different vehicle types, traffic densities and road types. *Transportation Research Interdisciplinary Perspectives*, 9, 100315.
- Zhou, X. Y., Schmidt, F., Toutlemonde, F., & Jacob, B. (2016). A mixture peaks over threshold approach for predicting extreme bridge traffic load effects. *Probabilistic Engineering Mechanics*, 43, 121-131.