

## Statisztikai módszerek alkalmazása klinikai genetikai kutatásokban

Dr. Firneisz Gábor, egyetemi adjunktus  
Semmelweis Egyetem, II. sz. Belgyógyászati Klinika  
Telefon: +36-20-825-0469

E-mail: [firneisz.gabor@med.semmelweis-univ.hu](mailto:firneisz.gabor@med.semmelweis-univ.hu)

[firneisz@yahoo.com](mailto:firneisz@yahoo.com)

Téma leírása:

### I. Téma:

Háttér: Az elmúlt évtizedek jelentős áttörést hoztak a genetikai kutatásokban, különösen a nagy „genom széles” asszociációs vizsgálatok lehetővé tették több népbetegségben – így pl. cukorbetegségben, magasvérnyomás-betegségben, bizonyos daganatos betegségekben – a rendszerint több tucat hajlamosító vagy védő hatású génváltozat (un. egy pontos nukleotid polimorfizmus, vagy SNP) azonosítását. Egy génváltozat hordozása önmagában általában kismértékű kockázatnövekedést jelent (az Odds Ratio típusosan 1.5 alatti) egy betegség kialakulásában, ezért a gyakorlatban a betegségre való hajlamot egy személyben általában több génváltozat együttesen alakítja ki. Ezt az összjátékot és a kialakuló összhatást akárcsak egy nagy szimfonikus zenekarban fontosnak érezzük, ennek vizsgálata a kutatási téma terhe miatt cukorbetegségben, egy jelenleg is zajló nemzetközi kutatás keretében.

A kutatás jelenlegi állása és résztvevői: a kutatáshoz már évekkel ezelőtt csatlakozott Pongor Lőrinc és Kontsek Endre, BME Biomérnök Mesterképzés hallgatói, illetve sok segítséget kaptunk, Varga László matematikus Ph.D. hallgatótól és témavezetőjétől, Zemléni András Docens Úrtól az ELTE Matematika intézetéből. Az elmúlt időszakban a génváltozatok azonosítására és az SNPSNP interakciók vizsgálatára a Khi-négyzet próbán alapuló tesztet fejlesztettünk ki, amelyet R és C programnyelvekben is megírtunk.

Megoldandó feladatok: A program közzétételéhez és „teljesítőképeségének” felméréséhez szükséges az irodalomban már ismertett módszerekkel <sup>1,2,3</sup> való összevetés. A publikált módszerek két főbb megközelítésen alapulnak a logisztikus regresszió („logit modell”) <sup>3</sup>, valamint dimenziócsökkentési módszereken (MDR) <sup>1,2</sup>. A saját – már létező – programunkkal való összevetés mellett szóba jött esetlegesen további megközelítések fejlesztése is (logit modell + wild bootstrap), azonban ez utóbbi lehetőség vélhetően számottevő többlet-kapacitást igényelne.

Különösen érdekes, hogy vizsgálataink terhe miatt kórképben történnek, ezért szükséges az egy személyben megvalósuló génváltozat interakciók (SNP-SNP interakciók) matematikai modellezése mellett a várandós állapot modellezése is, hiszen ebben az esetben két (egymástól nem független) individuum (anya és baba) génváltozatai közti interakció is létrejön. Utóbbi kérdés megválaszolásával még adós az irodalom (azaz újdonságértéke nagy), ugyanakkor vannak saját előzetes eredményeink a megközelítés tekintetében.

## II. Téma

Az előző témához képest elméletibb jellegű felvetésből fakad, így ennek a témának „igazi” (azaz publikált) külső előzményei sem léteznek, ugyanakkor több, mint egy éve foglalkozunk a kérdéssel. Megközelítésünk (és személyes orvosi tapasztalatom szerint is) az egészség, mint állapot igen nagyszámú elem és sok rendszer nagyfokban összerendezett, harmonikus működését jelenti, amiből az is ered, hogy a betegség nem más, mint egy olyan állapot, amely rendezetlenebb, mint az egészség.

A fizikai tudományokban a rendezetlenség mértéke az entrópia - célunk egy új dimenzió, a genetikai entrópia létrehozása (amelyre tapasztalatunk szerint a Shannon-féle információs entrópia nem alkalmas). Az eltelt időben a kérdés megközelítésében egy olyan leíró képlethez jutottunk el, amely az általunk felállított 5 kritériumból 4-nek megfelel, de további magas szintű matematikai és programozási ismeretek szükségesek annak eldöntéséhez, hogy ennek a képletnek a további módosításával lehetséges-e az 5. kritérium teljesítése, vagy egy teljesen új megközelítést kellene alkalmaznunk. Ha az elgondolásunk jó, egyszer talán arra is sor kerülhet, hogy ezt a dimenziót összhangba hozzuk a termodinamikai entrópiával, amelyre egyelőre csak az élő és nem élő tekintetében (betegségek nélkül), de találtunk már nívós, nemzetközi lapban publikált kísérletet is.<sup>4</sup>

### *Irodalomjegyzék:*

1. Calle ML, et al.: mbmdr: an R package for exploring gene–gene interactions associated with binary or quantitative traits. *Bioinformatics* 2010; 26(17): 2198–2199. doi:10.1093/bioinformatics/btq352
2. Briollais M et al.: Methodological issues in detecting gene-gene interactions in breast cancer susceptibility: a population-based study in Ontario. *BMC Medicine* 2007; 5:22. <http://www.biomedcentral.com/1741-7015/5/22>
3. Cordell HJ: Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet.* 2009; 10(6): 392–404. doi: 10.1038/nrg2579
4. Sanchez IC: Entropy of Living versus Non-Living Systems. *Journal of Modern Physics.* 2011; 2:654-657.doi:10.4236/jmp.2011.27077

Kulcsszavak: Genetika, Terhesség, Cukorbetegség, Gén-gén interakció, Biológiai rendszerek, Logit, MDR, Bootstrap, Entrópia

Nyelvtudás igény: Angol (a téma várhatóan sok utánaolvasást igényel, döntően angol nyelven közölt cikkekben)

Programozás-tudás igény: R, Matlab, C (utóbbi programnyelv ismerete nem szükséges feltétel)