Az érettségi vizsga prediktív validitása

*Az egyetemi előrehaladás sikerességének összefüggése
a matematika érettségi eredményeivel*

# Bevezetés

Magyarországon 2005 óta zajlik a középiskolai tanulmányok lezárásaként és a felsőoktatási rendszer felvételi vizsgájaként egyaránt funkcionáló kétszintű érettségi. A matematika érettségi vizsgák elemzésével Csapodi Csaba (2017.) foglalkozott részletesen PhD-disszertációjában. Jelen dolgozat a Csapodi tanár úr által vezetett kurzus keretében készült, és elsődleges célja annak bemutatása egy pilot-elemzés keretében, hogy a matematika érettségi eredményei milyen összefüggésben állnak az ELTE TTK matematika alapszakos hallgatóinak egyetemi sikerességével, előrehaladási, eredményeivel. A kurzus, illetve TDK-munka keretében a matematika érettségi, illetve az egyetemi kritériumteszt közötti összefüggést Jantner Anna vizsgálja.

Mivel a kétszintű érettségi bevezetésével a mérés automatikusan döntő szerephez jutott a felsőoktatási felvételi eljárásban számított pontszámok meghatározásában, mindenképpen szükségszerűnek tűnik annak vizsgálata, mennyiben képes az érettségi vizsga a felvételi funkciójának ellátására. Alapvető olvasata volna a kétszintű érettséginek, hogy a középszintű vizsga a középiskolai tanulmányok lezárására szolgál, és a megfelelő szakterület felsőoktatási intézményeiben való továbbtanulásra az emelt szintű vizsga sikeres letétele jogosítja fel a diákot. Ezzel szemben említésre méltó körülmény, hogy a matematika alapszak esetében a 2016. évi általános felvételi eljárás volt az első, amely *egyáltalán* emelt szintű érettségihez kötötte a felvétel lehetőségét, azonban még a jelenlegi, 2018-as tájékoztató szerint sem előírás éppen matematika tárgyból rendelkezni az emelt szintű vizsgával a szakra való felvételhez.

A matematika tantárgyra vonatkozóan nem ismert olyan kutatás, amely az érettségi vizsga prediktív validitását korábban átfogóan vizsgálta volna, továbbá az érettségi ezen jelenség meghatározásakor az egyetemi sikeresség, „beválás” skálájának definiálása önálló módszertani problémát vethet fel, mivel a felsőoktatásban tanulmányiátlag-jellegű mutatószámból is többet használnak párhuzamosan.

Csapodi tanár úr javaslatára és tanácsai, segítsége mentén ezért elvégeztem az ELTE TTK matematika alapszakára a 2013/14. tanév őszi félévében felvett hallgatók eredményeinek vizsgálatát. Mivel 2013-ban semmilyen szigorú megkötés nem volt érvényben a felvételi eljárásban rendelkezők számára a matematika érettségi meglétén kívül, így a kutatás célja annak vizsgálata, összességében a matematika érettségi eredmények milyen kapcsolatban állnak az egyetemi előrehaladás sikerességével, illetve ezen belül a közép- és emelt szintű eredménnyel rendelkező hallgatók körében kimutatható-e szignifikáns eltérés.

A kutatás kezdeti hipotézisei a következők voltak:

1a. Az emelt szinten érettségizők sikeresebben képesek teljesíteni a kritériumtesztet

1b. Az emelt szinten érettségizők eredményesebbek az egyetemi tanulmányaikban

2. Az érettségi szintjétől függetlenül gyenge kapcsolat áll fenn a százalékos eredmény és az egyetemi sikeresség mutatója között

# Anyag és módszer

## Elméleti, szakirodalmi háttér

Prediktív validitásnak valamely teszt azon mutatóját nevezik, amely megmutatja, a teszten elért eredmény milyen mértékben képes előre jelezni a tesztet kitöltők későbbi eredményességét valamely szituációban. Természetszerűen ez a mutatószám kiemelten fontos felvételi tesztek esetében. A prediktív validitás meghatározásának lehetséges módja a vizsgált teszt ismételt kitöltésének eredménye és az egyidejű eredményesség közötti összefüggés vizsgálata. Amennyiben a teszt prediktív validitása jó, úgy elvárható, hogy a két mutató kapcsolat között szignifikáns kapcsolat jelentkezzen. Összetettebb elemzés valósítható meg, amennyiben az eredeti tesztkitöltéstől kezdve folyamatosan nyomon követhető a kitöltők teljesítménye. Ebben az esetben a teszten felvett eredményeket ténylegesen a későbbi eredményességgel lehet összehasonlítani. A longitudinális vizsgálat elvégzése összetettebb, időigényesebb, azonban az alkalmazásával nyerhető eredmények a korábbi kutatások szerint megbízhatóbbak. (Badó & Józsa, 2017.)

A diákok középiskolai pályafutásával ellentétben az egyetemi hallgatók képzésébe nincsenek olyan egységes mérőeszközök beépítve, mint például az országos kompetenciamérés, továbbá az egyetemeken jellemző egyéni tantervek alkalmazása jelentős kihívást jelent az együtt felvett csoport tagjai esetében a későbbi eredmények összevetésében (eltérő félévben, eltérő oktatónál teljesítik esetleg az azonos nevű tárgyat is). Ezen túlmenően a képzés során jelentős a lemorzsolódás aránya, a szituációra jellemző, hogy az ELTE TTK hallgatói közül alapszakon és a mintatanterv által előírt időben diplomát szerző hallgatók részaránya a 25-30%-ot jellemzően nem haladja meg.

További különbség lehet, hogy míg az általános- és középiskolákban végzett kutatások (Badó & Józsa, 2017.) közepesen erős korrelációt mutatnak a például a nyolc évfolyamos gimnáziumok számára szervezett felvételi vizsgák eredményei, illetve a diákok előrehaladása között még az érettségi előtti teszten is, addig az egyetemeken jellemzően kritériumteszteket és felzárkóztató tárgyakat szerveznek. Az egyetem által elvárt tudásszintet a felvételkor nem teljesítő elsőéves hallgatók aránya igen jelentős az ELTE TTK matematika alapszakán is, innen származik a kutatás második hipotézise, mely szerint a matematika érettségi százalékos eredményei nem jelzik jól előre az egyetemi elvárásoknak való megfelelést.

Az előbbieket figyelembe véve tehát önálló elemzés témáját képezi az érettségi eredmények, illetve a matematika kritériumteszt közötti összefüggés vizsgálata. Jelen dolgozat keretében a felvett hallgatók egyetemi indexkivonatai alapján képzett mutatószámok alapján definiált kumulált eredményességi index szerinti rangsorban elfoglalt pozíció került összehasonlításra az érettségi eredményekkel.

## Az elemzésbe vont adatok

Az elemzés során a Csapodi tanár úr által rendelkezésemre bocsátott, anonimizált adatokkal dolgoztam, melyek a 2013/14. tanév őszi félévében az ELTE TTK matematika alapszakára felvételt nyert hallgatók érettségi eredményeit, illetve a hallgatók által az egyetemi képzés során az egyes tárgyakból szerzett érdemjegyeket tartalmazták.

A megkapott adatok táblázatkezelő programban kerültek eltárolásra, az érettségi eredmények előzetes feldolgozása pedig már megtörtént, így közvetlenül rendelkezésemre állt minden hallgató esetében a matematika érettségi százalékos eredménye, szintje / típusa (voltak még a korábbi rendszerben érettségizett jelentkezők is), az érettségi letételének éve, illetve a jelentkezők érettségi-átlageredményei.

Az előzetes feldolgozás után 138 egyedi hallgató adatait volt lehetséges vizsgálni, közülük 32 fő középszintű, 97 fő emelt szintű matematika érettségivel rendelkezett; régi rendszerű vizsgát 5 fő tett, illetve az adathiány 4 fő volt. (1. ábra)

1. ábra – A minta megoszlása az érettségizők által letett vizsga típusa szerint

A minta a hallgatók által letett érettségi vizsga időbeli megoszlása tekintetében a 2013. év hasonlóan domináns (64,5%), mint az előző esetben az emelt szintű érettségi. (2. ábra) A 2005 előtt érettségizett 6 fő közül egy fő matematikából és informatikából rendelkezik kétszintű érettségi eredményekkel. Megállapítható továbbá, hogy a 2013-ban érettségizettek esetében az emelt szintű matematika érettségivel felvettek aránya 84,3%, a 2012-ben érettségizettek között pedig 58,8%.

A kétszintű rendszerben vizsgát tett 129 fő közül 90% feletti érettségi átlageredménnyel rendelkezik 22 fő, további 5 százalékonként csökkenő sávokban közelítőleg egyenlő számú vizsgázó helyezkedik el, a leggyengébbek létszáma, átlagosan 70% alatti teljesítménnyel 13 fő.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *70% alatt* | *70-75%* | *75-80%* | *80-85%* | *85-90%* | *90% felett* |
| 13 fő | **17 fő** | **27 fő** | **22 fő** | **28 fő** | **22 fő** |

2. ábra – A minta megoszlása a felvettek által tett vizsga időpontja szerint

## A COCO-Y0 módszer[[1]](#footnote-1)

A vizsgálatba bevont hallgatók egyetemi előrehaladásának / sikerességének mérésére az antidiszkriminatív hasonlóságelemzés (*component-based object comparison for objectivity*, COCO) módszerét alkalmaztam.

A COCO-módszer egy speciális többváltozós regressziós modellt alkalmaz, s tetszőleges számú független ($X\_{i}\in N^{n}$) és a függő változó ($Y\in R^{n}$) közötti összefüggés feltárására alkalmas olyan módon, hogy az egyes változók nem egyetlen regressziós együtthatót kapnak, hanem az adott változó értékein értelmezett lépcsős függvény ($S\_{i}:N\rightarrow R$) adja meg a változó lehetséges hatásának értékeit. A modellalkotás lineáris programozási eljáráson alapul, a választott módszer esetében a változók additív becslését adják az $Y$ értékének.

$$\left(\begin{matrix}Y\_{1}\\\vdots \\Y\_{n}\end{matrix}\right)=\sum\_{i=1}^{k}\left(\begin{matrix}S\_{i}\left(X\_{1i}\right)\\\vdots \\S\_{i}\left(X\_{ni}\right)\end{matrix}\right)+\left(\begin{matrix}Δ\_{1}\\\vdots \\Δ\_{n}\end{matrix}\right)$$

A fenti egyenlet teljesül az $n$ objektumot tartalmazó vektor esetén, ahol a Delta hibatagok minimalizálása a becslést létrehozó modellalkotás célja.

A COCO-modellalkotás alapfeltétele az objektum–attribútum-mátrix (OAM) megléte, melynek sorai az objektumok, oszlopai pedig a független és függő változók. A nyers adatokat tartalmazó OAM-ot méretfüggetlenítés céljából sorszámozni szükséges, így minden objektum minden attribútuma az adott attribútum értékeinek sorában elfoglalt helyét kapja ebben a lépésben. A sorszámozás csökkenő vagy növekvő iránya attól függ, hogy az $X\_{i}$ tényleges és az $Y$ (szakirodalmi adatok vagy az elemző saját mérlegelése alapján) elvárt értékei közötti korreláció pozitív vagy negatív.

A lépcsők száma a modellekben alapértelmezés szerint az objektumok száma, így egy adott attribútum minden lehetséges rangsorszámához van mód önálló lépcsőértéket társítani, azonban speciális esetekben, amikor az objektumok között sok, egy-egy attribútumon belül azonos sorszámmal rendelkező fordul elő (például már az attribútumok maguk osztályzatok, amelyek különböző értékei eredendően is korlátozott számban fordulnak csak elő), mesterségesen lehetséges csökkenteni a lépcsők számát. Ezzel (a fajlagos számítási igény csökkenése miatt) a gyakorlatban futtatható mátrixok mérete növelhető.

A felhasználás céljától függően különböző COCO-módszerek érhetők el, melyek között alapvetően szükséges megkülönböztetni a COCO-STD (standard hasonlóságelemzés) és a COCO-Y0 (antidiszkriminatív hasonlóságelemzés) típusokat. A STD-modellek esetében a függő változó értéke valamely eleve adott mennyiség, például egy termék ára, és ennek a mutatószámnak a levezethetőségét vizsgálja a modellalkotás. Ezzel a modellel tipikusan olyan kérdésekre lehet választ keresni, hogy található-e a termékek között alul-, esetleg túlárazott.

Az Y0-modell elnevezésének alapját az a meglátás adja, hogy a függő változó értéke itt lényegében nulla (az LP-módszer technikai megszorításaiból adódóan valójában egy nagy konstans érték, az ún. norma-pontszám), amelytől a pozitív, illetve negatív irányú eltérés ténye és mértéke szolgáltat információt. A modell alkalmazásával összefüggő kérdés a következő: „lehet‑e minden objektum másként egyforma?”, az antidiszkriminatív elnevezést pedig az indokolja, hogy ebben a modellben hibátlan becslés akkor keletkezik, ha létezik a lépcsősfüggvényeknek egy olyan halmaza, amelyek alkalmazásával minden objektum becsült pontszáma egyformán a norma-pontszám.

A hasonlóságelemzési modell speciális jóságmutatója az inverz tesztelés. Ebben az esetben az eredeti attribútum-értékeket fordított irányítás mellett sorszámozva kerül sor az (azonos modell) futtatására, s azon objektumok esetében tekinthető validnak az eredeti hasonlóságelemzés, ahol a normál és az inverz Delta szorzata negatív (előjele ellentétes[[2]](#footnote-2)).

# Az elemzés végrehajtása

Amint az adatok bemutatásánál látható volt, az elemzés során az érettségi eredmények további feldolgozására nem volt szükség, az adatokkal való manipuláció az egyetemi indexkivonatokat érintette.

## Az adatok tisztítása

Az egyetemi eredmények egyetlen, 5396 sorból álló táblázatban voltak tárolva. Ez hallgatónként átlagosan kb. 39 bejegyzést jelent, azonban az erős lemorzsolódás miatt az egyes hallgatókra jutó teljesítések és teljesítési kísérletek száma szélsőséges ingadozást mutat. Az egy hallgatóra jutó bejegyzések maximális száma 67, és 20-nál több bejegyzéssel 98 hallgató (71%) rendelkezik – nagyságrendileg ennyi tehát azok aránya, akik egy évet meghaladó időt töltöttek a szakon. Az első év abból a szempontból is jelentős mérföldkő, hogy a kritérium teszt teljesítésére a második félév végéig van lehetősége a hallgatóknak, s ha ezalatt nem sikerült megfelelt minősítést szerezniük, elbocsátásra kerülnek a szakról.

Az indexekben nem teljesített, illetve elhagyott tanegységként (amelynek a teljesítését meg sem kísérelte a hallgató, tehát elégtelen / nem felelt meg minősítés bejegyzésére sem került sor) 21 hallgatónál (15%) szerepel a kritériumtárgy az első két félévben, ami azt jelenti, hogy körülbelül még egyszer ennyi hallgató a kritériumtárgy teljesítése ellenére kiszóródott a képzésről. Ezek a hallgatók egy kivételével szintén igen gyengén teljesítettek (átlagosan 8,5 bejegyzés, teljesített tárgyaik átlaga 2,5), és a teljes mintában található öt olyan hallgató közül, aki régi rendszerű érettségivel felvételizett, négy közöttük van, az ötödik hallgató pedig egyáltalán nem teljesített tárgyat.

A hallgatók, illetve a bejegyzések magas összesített száma mellett az egyetlen táblázatban tárolt adatok a tárgyak kezelése kapcsán is problémákat vetnek fel: 348 db egyedi tárgynév, illetve 411 db egyedi tárgykód található a listában, melyek között sok az olyan, szabadon választható tárgyat, melyet csak igen kevesen teljesítettek. Az egyértelműen a Matematika Intézethez tartozó tárgykódok száma 236 db, azonban az azonos tárgyhoz tartozó eltérő szintek / változatok[[3]](#footnote-3) egyesítésével, összevonásával ez még csökkenthető volt.

Az adatokat áttekintve úgy döntöttem, hogy a pilot-elemzés során kizárólag a matematika alapszak első félévében mintatanterv szerint kötelező tárgyak esetében vizsgálom a teljesítést. Az elemzésbe vont tárgyak a következők voltak:

* matematika kritériumtárgy
* elemi matematika 1 gyakorlat
* analízis 1 gyakorlat és előadás VAGY kalkulus 1 gyakorlat és előadás
* algebra 1 gyakorlat és előadás
* számelmélet 1 gyakorlat és előadás
* véges matematika 1 gyakorlat és előadás

Ezzel párhuzamosan megkezdtem a további alapvető tárgyak változatainak egységes azonosítóval való ellátását is, de további részletes elemzést megelőzően valószínűleg egy alaposan felépített adatbázis létrehozása volna szükséges, ahonnan SQL-lekérdezésekkel könnyebben testre szabható lekérdezések állíthatók elő, mint az Excel beépített szűrőlehetőségeivel vagy függvényeivel.

## COCO-attribútumok képzése

Annak érdekében, hogy a 138 hallgató eredményeinek összehasonlítása minél alaposabb lehessen, szándékomban állt minél nagyobb számú attribútum bevonása az elemzésbe, ugyanis a COCO-módszerrel a sorszámozás (méretfüggetlenítés) révén igen változatos adatok együttes elemzésére nyílik lehetőség olyan módon, hogy a mutatószámok kiválasztása után a műveleteket az elemző szubjektív és *ad hoc* súlyozással már nem befolyásolja.

Az elemzés[[4]](#footnote-4) során a mutatószámoknak három karakteres csoportját alkalmaztam:

* érdemjegyek; tárgyteljesítési mutatók
* a tárgy sikeres teljesítéséhez tartozó félév-adatok
* átlag- vagy darab-jellegű, aggregált mutatók

Minden hallgató esetében tíz „érdemjegy” került elemzésre azzal a döntési szabállyal, hogy minden hallgató esetén minden tárgyhoz az utolsóként bejegyzett értékelés kerül automatikusan figyelembe vételre. A tárgyak közül a matematika kritériumtárgy háromfokozatú skálán: elhagyott / nem felelt meg / megfelelt; a fejenként öt kötelező tárgy pedig hatfokozatú skálán szerepelt, ahol a tárgy elhagyása nulla érdemjegyként szerepelt. Az alternatív teljesítésű analízis / kalkulus tárgyakat összefésültem, azon hallgatók esetében pedig, akik analízisről sikertelenség miatt váltottak kalkulusra, csak az utóbbi tárgy teljesítésének eredményeit vettem figyelembe.

Ebben a blokkban minden mutatószám csökkenő irány mentén került sorszámozásra, mivel a végső eredményességben természetesen a jobb jegyek kell, hogy a jobb eredményhez tartozzanak.

Az tárgyteljesítéshez tartozó aggregált mutatószámok között az első a teljesítési kísérlettel rendelkező tárgyak eredményeinek hagyományos átlaga volt hallgatónként (az elhagyott tanegységek nem szerepeltek a számolásban, azonban a sikertelenség után analízisről kalkulusra történő váltás plusz egy jeggyel (elégtelen) itt beszámításra került. A második mutató az elhagyott tanegységek szankcionálásával számított, kreditindex-szerű mutatószám, ahol a kötelező tárgyakból megszerzett érdemjegyek (nullától ötig) összege került elosztásra kilenccel. A matematika kritériumtárgyat, mivel az egyetem szabályzatai szerint a kétfokozatú értékeléssel rendelkező tárgy nem szerepel az átlagolásban, egyik alkalommal sem vettem figyelembe.

További mutatószámok voltak a jegyszerzési kísérlettel zárult tárgyak összesített száma, a teljesített tárgyak száma, az elégtelen érdemjeggyel zárt tárgyak száma, illetve az elhagyott tanegységek száma. Az átlagok, illetve az érdemjeggyel zárt tárgyak száma csökkenő, a bukott és elhagyott tanegységek száma növekvő irány mentén került sorszámozásra.

Az előzőekben bemutatott tíz tárgyhoz egyenként feljegyzésre került az is, a hallgató mely félévben szerezte az előző blokkban figyelembe vett utolsó érdemjegyet – ez a félévszám ilyen módon a sikeres teljesítés félévét jelenti, kivéve azon tárgyak esetében, amelyekből a hallgató utolsó teljesítési kísérlete is elégtelen volt, vagy amelyeket elhagyott. Az elhagyott tanegységekhez a teljesítés féléveként a hatodik lett megjelölve, mivel egyébként olyan hallgató és olyan tárgy nem volt, amelyet természetes módon a hatodik félévben még megkíséreltek volna teljesíteni.

A teljesítési félévek értelemszerűen növekvő irány mentén kerültek sorba rendezésre, mivel az eredményesség fogalma ezúttal a minél korábbi teljesítéssel függ össze.

A félévszámokra vonatkozóan két aggregált mutató került az elemzésbe, egyrészt az előzőekben bemutatott félévszámok átlagértéke, másrészt pedig azon tárgyak darabszáma, amelyeket az első félévben érdemjeggyel (esetleg ki nem javított elégtelennel) zárt a hallgató. A félévszámok átlaga növekvő, az első félévben lezárt tárgyak száma csökkenő irány szerint került sorszámozásra.

Végül 28 attribútum mentén került sor a hallgatók értékelésére, így a végleges OAM (a függő változó, konstans 1000 érték, felvételével) 139x29 méretű. Ez a méret a jelenleg elérhető online LP-modul kapacitásának felső korlátjánál található, így a lépcsők számát a sorszámok „összesöprésével” tízre[[5]](#footnote-5) állítottam be. Mivel az elemzésben 10 tárgy szerepel, így a sorszámozás felbontását ez, a lépcsőkre vonatkozó módosítás számottevő arányban kizárólag a tanulmányi, illetve félév-átlag mutatószámok esetében csökkenti, a többi mutatószám esetében az egyedi értékek száma kevés.

A fenti attribútumokkal a COCO-Y0 modell futtatása sikeres volt, az inverz irányok mentén végzett ellenőrzés a 138 hallgatóból kettő kivételével minden esetben validálta az elsődleges hasonlóságelemzés eredményét.

## A hasonlóságelemzés eredménye

Az 1000 norma-pontszám körül 850 és 1100 pont körül szóródnak az eredményességi mutató értékei az egyes hallgatók esetében, vagyis a hallgatók eredményeit tekintve (természetesen) nem teljesül a „mindneki-másként-egyforma”-elv.

A 28 attribútum esetében a lépcsősfüggvények lefutása 26 esetben teljesen egységes, és jelentős eltérést a maradék esetek sem mutatnak, így kijelenthető, hogy az eredményességi mutató a bemenő attribútumok mindegyikével legalább közepesen erős korrelációt mutat (ld. 3. ábra). Nem jelentkezett az attribútumoknak olyan kisebb csoportja, amely önállóan kiemelkedően erős magyarázóerővel rendelkezne, így ezúttal a hasonlóságelemzési, illetve az egyszerű rangsor-összegzéses értékelés közel azonos eredményre vezet.

3. ábra – A hasonlóságelemzéssel maghatározott egyetemi eredményesség és az elhagyott tanegységek figyelembe vételével számított tanulmányi átlag közötti összefüggés
($R^{2}=0,934$)

Az eredményességi mutató eloszlása a következő ábrán látható. Megállapítható, hogy az eloszlás nem mutatja a normális eloszlásra jellemző tulajdonságokat, a mintában két csoport különül el markánsan, a norma felett, illetve alatt teljesítők.

4. ábra – A hasonlóságelemzéssel meghatározott egyetemi eredményesség eloszlása

# Következtetések

## A kutatás hipotéziseinek ellenőrzése

Az 5. ábrán látható az emelt, illetve középszintű érettségivel felvett hallgatók vizsgaeredménye, illetve az első féléves teljesítmények alapján kalkulált egyetemi eredményesség közötti kapcsolat.

5. ábra – Az egyetemi eredményesség és az érettségi eredmények kapcsolata
„Emelt” $R^{2}=0,070$ (p<0,01) „Közép”$ R^{2}=0,122$ (p=0,05)
*bekarikázott jelölő:* sikertelen kritériumteszt

Mind az emelt, mind a középszintű érettségi vizsgával felvételizők esetén kijelenthető, hogy az egyetemi eredményességet mérő hasonlóságelemzési indexhez csak gyengén korrelál az érettségi eredményük, mintegy 7%, illetve 12% magyarázó erővel bírva csupán. Ennek ellenére megállapítható, hogy mindkét esetben – 55 pont körüli abszolút hiba mellett – szignifikáns az összefüggés a két változó között.

6. ábra – A COCO-Y0 pontszám megoszlása az érettségi típusa szerint

A fenti 6. ábrán látható, hogy az egyetemi eredményesség a részmintákban sem közelít egyértelműen a normális eloszláshoz, azonban az emelt szintű érettségivel felvételizők kiugróan magasabb pontszámai világosan látszanak (az emelt szinten érettségizők eredményességének átlaga 1020 pont, a középszinten érettségizőké 960 pont).

A 18 fős csoportnak, akik nem tudták az első két félév alatt teljesíteni a kritériumtárgyat, a kétharmada érkezett középszintű érettségivel. Ezek az adatok azt is jelentik, hogy a középszinten érettségizők körében a sikertelen kritériumtesztet írók aránya 37,5%, míg ugyanez az emelt szintű érettségivel felvételizők esetében kevesebb, mint 6,5 %.

## További kutatási lehetőségek

Amint az adatok feldolgozásánál megállapítható volt, a rendelkezésre álló egyetlen táblázatnál rugalmasabb feldolgozást tenne lehetővé egy precízen felépített és kódolt adatbázis használata. Az adatbázis kialakításával együtt a mintatantervek / NEPTUN-rendszer alapján az egyes tárgyakhoz rendelt kreditértékek lekérdezését követően további attribútumokkal volna bővíthető az elemzés.

A COCO-Y0 hasonlóságelemzés alkalmazása esetén az elemzési mátrix mérete jelentős mértékben már csak akkor növelhető, ha az LP-szolgáltatásokat az online eszköz helyett megfelelő teljesítményű célgépen van lehetőség implementálni, ebben az esetben azonban mindenképpen érdemesnek látom az elemzés folytatását a további féléves teljesítmények vizsgálatba vonásával, végső soron az azóta esetleg csúszással is diplomát szerzettek adatainak a teljes alapképzésen végigfutó vizsgálatával.

# Összefoglalás

Szabadon választható egyetemi kurzus keretében valósult meg a matematika érettségi prediktív validitásának vizsgálatát előkészítő pilot-elemzés. A Magyarországon a felvételi funkcióját több mint tíz éve betöltő kétszintű érettségivel kapcsolatban ilyen jellegű kutatás matematika tantárgyra vonatkozóan nem ismert, a módszertan azonban a közoktatásban szélesebb körben előforduló felmérések alapján adaptálható az egyetemi mintára is.

Mivel az egyetemi képzés relatív rövidsége, a nagy lemorzsolódás és az egyéni tantervek nem teszik lehetővé a képzés során természetesen felmerülő közös tesztek használatát az egyetemi eredményesség mérésére, így az elérhető indexkivonatok adataiból származtatott különböző mutatószámok felhasználása mellett antidiszkriminatív hasonlóságelemzési modellben tettem kísérletet az eredményesség modellezésére.

A COCO-Y0 módszerrel meghatározott eredményességi pontszámok magas korrelációt mutatnak a felsőoktatásban elterjedten használt tanulmányi átlag, illetve kreditindex-jellegű mutatószámokkal, azonban érdemben képesek további szempontok hatását is megjeleníteni, akár az egyes tárgyak egyedi eredményeire, akár a képzési idő folyamán a teljesítés időpontjára vonatkozóan.

A kutatás kezdeti hipotézisei beigazolódtak: az emelt szintű, illetve a középszintű érettségivel felvételiző hallgatók mind a kritériumteszten való sikeresség, mind az egyetemi eredményességi pontszám tekintetében számottevően eltérő részmintát alkotnak. Az érettségi százalékos eredményei és az egyetemi eredményesség között gyenge pozitív, szignifikáns korreláció volt megfigyelhető.

# Felhasznált irodalom

Badó, Zs. & Józsa, K. (2017.) A gimnáziumi felvételi vizsga eredményének és a gimnáziumi tanulás sikerességének kapcsolata. *Új Ped. Szemle.* 2017/1-2
([online](http://folyoiratok.ofi.hu/uj-pedagogiai-szemle/a-gimnaziumi-felveteli-vizsga-eredmenyenek-es-a-gimnaziumi-tanulas); hozzáférés dátuma: 2018.01.06.)

Csapodi, Cs. (2017.) *„A MATEMATIKA ÉRETTSÉGI VIZSGA ELEMZÉSE 2005-2015” PhD-értekezés*. Debrecen.
([online](https://dea.lib.unideb.hu/dea/bitstream/handle/2437/236563/Csapodi_Csaba_doktori_ertekezes_vegleges.pdf); hozzáférés dátuma: 2018.01.05.)

**Ábrák jegyzéke**

[1. ábra – A minta megoszlása az érettségizők által letett vizsga típusa szerint 4](#_Toc503091139)

[2. ábra – A minta megoszlása a felvettek által tett vizsga időpontja szerint 5](#_Toc503091140)

[3. ábra – A hasonlóságelemzéssel maghatározott egyetemi eredményesség és az elhagyott tanegységek figyelembe vételével számított tanulmányi átlag közötti összefüggés 11](#_Toc503091141)

[4. ábra – A hasonlóságelemzéssel meghatározott egyetemi eredményesség eloszlása 12](#_Toc503091142)

[5. ábra – Az egyetemi eredményesség és az érettségi eredmények kapcsolata 12](#_Toc503091143)

[6. ábra – A COCO-Y0 pontszám megoszlása az érettségi típusa szerint 13](#_Toc503091144)

Tartalom

[Bevezetés 1](#_Toc503091293)

[Anyag és módszer 2](#_Toc503091294)

[Elméleti, szakirodalmi háttér 2](#_Toc503091295)

[Az elemzésbe vont adatok 3](#_Toc503091296)

[A COCO-Y0 módszer 5](#_Toc503091297)

[Az elemzés végrehajtása 7](#_Toc503091298)

[Az adatok tisztítása 7](#_Toc503091299)

[COCO-attribútumok képzése 8](#_Toc503091300)

[A hasonlóságelemzés eredménye 11](#_Toc503091301)

[Következtetések 12](#_Toc503091302)

[A kutatás hipotéziseinek ellenőrzése 12](#_Toc503091303)

[További kutatási lehetőségek 13](#_Toc503091304)

[Összefoglalás 14](#_Toc503091305)

[Felhasznált irodalom 15](#_Toc503091306)

1. Az elemző eszköz online hozzáférhető a következő linken: <http://miau.gau.hu/myx-free/coco/>

További információ a COCO-módszertanról [ebben a dokumentumban](http://miau.gau.hu/miau/196/My-X%20Team_A5%20fuzet_EN_jav.pdf) olvasható [↑](#footnote-ref-1)
2. Amennyiben ugyanis valamely objektum az első rangsort, a „szépségversenyt” megnyerte, úgy magától értetődő elvárás, hogy a fordított szempontok szerinti rangsorolást, a „csúnyaságversenyt” el kell veszítenie egy konzisztens modellben. [↑](#footnote-ref-2)
3. normál, haladó, intenzív [↑](#footnote-ref-3)
4. Az eredeti OAM a kapcsolódó munkafüzet azonos nevű munkalapján érhető el. [↑](#footnote-ref-4)
5. A munkafüzetben a „rank” munkalap 144. sorától kezdődően. [↑](#footnote-ref-5)