

Diplomás kutatás 2010

Módszertani leírás

A kutatás alapsokasága

Vizsgálati populáció: államilag elismert felsőoktatási intézmények alapképzésben, kiegészítő alapképzésben és diplomás képzésben (minden tagozaton és finanszírozási formában) résztvevő 2007-ben végzett (diplomát szerzett) hallgatói.

A vizsgálat tíz képzési területre terjed ki:

- agrár
- bölcsészettudomány
- gazdaságtudomány
- informatikai
- jogi és igazgatási
- műszaki
- orvos- és egészségtudományi
- pedagógusképzés
- társadalomtudomány
- természettudomány

A vizsgálatba a négy legkisebb végzett létszámú képzési terület (sporttudományi, nemzetvédelmi és katonai, művészeti és művészetközvetítői) épp az alacsony létszámú populáció miatt nem került be (e területek végzettjeinek interjú vizsgálatát a kérdőíves kutatással párhuzamosan zajlik).

Az alapsokaság 49 382 fő. (Forrás: OKM statisztika)

Az adatok forrásai a felsőoktatási intézmények tanulmányi adatbázisából származó, részleges adatköröket tartalmazó címlisták, amelyeket az intézmények bocsátottak rendelkezésünkre. A minta végleges és 2010-es adatokra frissített címlistájának összeállítása a KEK-KH adatszolgáltatása révén valósult meg.

Az adatfelvétel módszere

Személyes lekérdezésen alapuló standard kérdőíves kutatás, laptopos adatrögzítéssel. Az adatfelvételt a megkérdezettnek átadott kártyafüzet segítette.

Az adatfelvételt a Medián Közvélemény- és Piackutató Intézet valamint a GFK Hungária Piackutató intézet kérdezőbiztosi hálózata végezte. A Budapesti Gazdasági Főiskola végzettjeinek esetében az adatfelvétel a Medián szakmai háttértámogatása mellett hallgatói részvétellel valósult meg 2010. május-júniusban.

A címlisták kezelése

A mintavétel alapjául szolgáló címlista 4879 főcímet tartalmazott, és egy főcímhöz ideálisan három, hierarchikusan, egy-egy értelműen hozzárendelt pótcím tartozott. A teljes minta 17 793 rekordból állt, mivel nem sikerült minden főcímhöz négy megfelelő (nem, képzési terület és intézmény tekintetében azonos) pótcímet rendelni.

Az adatfelvétel indulását követően az első néhány száz felkeresési kísérlet során az adatfelvételt végző cég javaslatára a címlista tervezett hierarchikus kezelése módosult. Ennek oka az volt, hogy a visszajelzések alapján a kutatás terepmunkája ebben a struktúrában, tehát a főcímek, illetve a megfelelő főcímek egymást követő felkeresése révén hatékonyan nem megvalósítható. (Ebben a területi tagoltság mellett a nehéz, többszöri megkeresést igénylő kapcsolatfelvételen túl az is szerepet játszott, hogy nagyon sokan másik településen éltek, mint az a mintában szerepelt, így – miután az új cím kiderült -, újabb megkeresések voltak szükségesek a főcímen, és addig a pótcímek használatára nem volt lehetőség. Ez olyan bonyolult logisztikát és részben ennek következtében akkora többlet időigényt jelentett, ami veszélyeztette a kutatás befejezését a tervezett időkereten belül.)

A problémát felismerve az adatfelvétel irányításával megbízott cég egy olyan szisztémát dolgozott ki, melyben vállalta, hogy a cím besorolásától függetlenül megkísérli elérni a teljes mintakeretet és ezeket utólag rendezni az eredeti rendszerbe. Az adatfelvétel torzulásának minimalizálása érdekében a főcímen történő elérést továbbra is előnyben részesítettük mind a többszöri megkeresések, mind a cím változásának követése tekintetében. A 17 793 cím közül összesen 14 284-ről (80 százalék), a 4879 főcím közül 4051 (83 százalék) sikerült valamit megtudni. Az elért 4051 főcím közül 1647-ben (41 százalék) készült el az interjú.

A címek elérése, a sikeres interjúk száma és aránya

cím	elért minta	tényleges minta	lekérdezési arány
főcím	4051	1647	41
első pótcím	3725	1066	29
második pótcím	3525	968	27
harmadik pótcím	2983	826	28
Total	14 284	4507	31

Azt, hogy a módosított mintavételi eljárás nem rontotta az adatfelvétel minőségét, alátámasztja, hogy a háromdimenziós súlyváltozó átlaga független attól, hogy főcímen vagy (hányadik) pótcímen készült az interjú.

A súlyváltozó jellemzői a főcím és pótcímek szerint a súlyozatlan mintában

cím	Mean	N	Std. Deviation
főcím	,9799	1647	,48675
első pótcím	,9974	1066	,53372
második pótcím	,9910	968	,48875
harmadik pótcím	1,0543	826	,54503
Total	1,0000	4508	,51009

Mintavétel

A mintavétel módszere egyszerű véletlen mintavétel. A mintát intézményi alminták alkotják, melyhez a végzetek listáját a kutatásban részt vevő intézmények bocsátották rendelkezésünkre. (A Budapesti Gazdasági Főiskola és a Szegedi Tudományegyetem esetében a címlisták átadása nem történt meg, e két intézmény részletes mintavételi instrukciók alapján előállított minta-arányos címlistáját közvetlenül az adatfelvevő cégek kezelték, az Educatio Nonprofit Kft. közbeiktatása nélkül.)

A kutatás mintájába az adott képzésterületen képző intézmények szinte teljes köre bekerült. Mivel rendelkezésre állt a végzetek közel teljes listája, az egyszerű véletlen módszerrel kialakított minta nem és a végzés helye (kar) szerint reprezentálja az egyes képzési területeken végzetek országos megoszlását.

A minta kialakításának elve egymást közt nem arányos intézményi almintákon alapult. Ennek indokoltsága a későbbi képzési területi elemzések számára megfelelő elemszám biztosítása volt a kisebb létszámú képzési területeken. Ennek következtében az egyes képzési területek nem számarányuknak megfelelően kerültek be a mintába. Ezt a szándékolt torzítást az adatbázis létrejötte után súlyozással korrigáltuk.

A tervezett minta elemszáma: 5500 fő.

A tervezett minta megoszlása képzési területek szerint

	Végzetek száma	Ebből nő	Mintanagyság
Agrár képzési terület	3879	2318	500
Bölcsettudományi képzési terület	7592	6092	700
Gazdaságtudományi képzési terület	11504	8595	800
Informatika képzési terület	3293	666	475
Jogi és igazgatási képzési terület	3714	2412	500
Műszaki képzési terület	6016	1787	700
Orvos- és egészségügyi képzési terület	3195	2527	475

Pedagógusképzés	3732	3552	500
Társadalomtudományi képzési terület	4499	3631	500
Természettudomány képzési terület	1958	1184	350
Végösszeg	49 382	32 764	5500

Súlyozás

A tervezett minta nagysága: 5500 fő

A létrejött adatbázis elemszáma: 4511 fő

A súlyozás alapja a teljes mintakeret nem, intézmény és képzési terület szerinti megoszlása volt. (A pótcímek számának különbözőségéből eredő torzulás mértéke pontosan nem számszerűsíthető, mindazonáltal csekély hatást kell tulajdonítanunk neki, tekintettel arra, hogy a címlistáknak intézményenként az egyes képzési területek nemek szerinti megoszlását kellett tükröznie, és ehhez elsősorban a számosságuknak van jelentősége, az pedig biztosított (a teljes alapsokaság 36 százaléka). Ahol nem volt elég pótcím, ebben éppen az játszott szerepet, hogy azon a képzési területen nem volt elégséges számú eset az alapsokaságban.) Mivel a mintakeretben nem szerepeltek a BGF és a Szegedi Tudományegyetem diplomásai, az e két intézményben végzetteket létszamarányosan, és a súlyozási szempontok szerint megoszlásának megfelelően vettük figyelembe.

Több szempontú cellasúlyozást alkalmazunk. A súlyváltozó háromdimenziós: a nem (2 kategória), a képzési terület (10 kategória) és az intézmény (24 kategória) szerint összesen elméletileg 480, a valóságban (mivel nincs oktatás minden intézményben minden képzési területen) 244 cellát tartalmaz. A cellasúlyozás lényegesen megbízhatóbb eljárás, mint a peremsúlyozás, mert az utóbbi révén a súlyozáshoz használt „peremváltozók” közel függetlenek lesznek egymástól, így minden intézményben minden képzési területen pontosan annyi férfi és nő szerepel a mintában, mint a valóságban.

A képzési területeken belüli reprezentativitást biztosító súlyváltozó értéke 0,2 és 7,1 között alakul, átlaga 1, szórása 0,509. A súlyváltozó átlaga minél közelebb van egyhez, az adott csoport annál inkább valós arányának megfelelő mértékben szerepel a teljes mintában, az érték egynél nagyobb azoknak a csoportoknak az esetében, amelyeket a valósnál kisebb arányban sikerült elérni (ezek a súlyozatlan tényleges mintában alulreprezentáltak), és egynél kisebb azokéban, amelyek a valósnál nagyobb arányban sikerült elérni. A férfiak és a nők elérése tekintetében alig van különbség, a Zsigmond Király Főiskolán, a Pázmány Péter Tudományegyetemen és a (más keretek között, az intézmény saját adatbázisát felhasználva, saját hallgatói által készített interjúkkal kérdezett) Budapesti Gazdasági Főiskolán végzett hallgatók számottevően alulreprezentáltak, a Kaposvári Egyetemen, a Nyugat-Magyarországi Egyetemen végzettek pedig felülreprezentáltak. A képzési területek elérése tekintetében lényegesen kisebbek a különbségek, mint az intézményekben: a pedagógiai és a társadalomtudományi szakokon végzettek elérése volt a legkönnyebb, és a gazdaságtudományi, valamint a jogi és igazgatási képesítésűeké volt a legnehezebb.

A képzési területi (suly1_1) súlyváltozó jellemzői a súlyozási szempontok szerint a súlyozatlan mintában

Nem	Mean	N	Std. Deviation
férfi	1,0922	1401	,61148
nő	,9585	3110	,45083
Intézmény	Mean	N	Std. Deviation
BCE	1,2022	176	,38403
BGF	1,9163	253	,06853
BME	1,4265	208	,36174
DE	,8666	482	,22245
DF	,6783	65	,11405
EKF	,7165	110	,24582
ELTE	1,2887	377	,31220
KE	,5796	151	,18459
KF	,6369	176	,18999
KJF	1,0125	43	,30617
KRE	,9868	55	,34646
KRF	,9865	144	,12217
ME	1,2092	181	,39478
NYF	,8142	143	,28866
NYME	,5839	389	,12914
PE	,7576	145	,26515
PPKE	2,1014	33	,57450
PTE	,7714	485	,20790
SE	1,2436	131	,05625
SZE	,6960	231	,19425
SZF	,8482	65	,32575
SZIE	,8223	213	,44250
ZSKF	2,1616	32	,53744
SZTE	1,4036	223	1,03652
Képzési terület	Mean	N	Std. Deviation
agrártudományi	,9600	383	,37465
bölcsészettudományi	1,0516	536	,46357
gazdaságtudományi	1,3179	733	,55822
informatikai	1,1464	278	,62166
jogi és igazgatási	1,2536	308	,40308
műszaki	,9303	569	,33096

orvos- és egészségtudományi	1,0327	355	,33107
pedagógiai	,6446	581	,31877
társadalomtudományi	,7627	516	,62369
természettudományi	,9709	249	,45378
Total	1,0000	4508	,51009

Mivel a mintakeretben annak érdekében, hogy a kevésbé népes képzési területek is megbízhatóan elemezhetőek legyenek, az egyes képzési területek nem tényleges arányuknak megfelelően szerepeltek, a cellasúlyozás révén előállított súlyváltozót peremsúlyozással korrigáltuk (a minta kialakításának alapjául szolgáló OKM statisztika alapján). A korrekció a másik két súlyozáshoz használt változó megoszlását alig érintette, egyedül a BGF aránya tekintetében van érdemi (1 százalékpontnyi) különbség.

Az adatfájl megoszlása a súlyozáshoz használt változók szerint (százalék)

	Cellasúly	Korrigált súly
Nem		
férfi	33,9	33,7
nő	66,1	66,3
Intézmény		
BCE	4,7	4,6
BGF	10,7	11,7
BME	6,6	6,7
DE	9,3	8,7
DF	1,0	1,0
EKF	1,7	1,7
ELTE	10,8	10,6
KE	1,9	1,8
KF	2,5	2,4
KJF	1,0	1,1
KRE	1,2	1,2
KRF	3,1	3,2
ME	4,9	4,8
NYF	2,6	2,6
NYME	5,0	5,1
PE	2,4	2,5
PPKE	1,5	1,5
PTE	8,3	8,4
SE	3,6	2,9

SZE	3,6	3,7
SZF	1,2	1,3
SZIE	3,9	3,9
ZSKF	1,5	1,7
SZTE	6,9	7,0
Képzési terület		
agrártudományi	8,2	7,9
bölcsészettudományi	12,5	15,4
gazdaságtudományi	21,4	23,3
informatikai	7,1	6,7
jogi és igazgatási	8,6	7,5
műszaki	11,7	12,2
orvos- és egészségtudományi	8,1	6,5
pedagógiai	8,3	7,6
társadalomtudományi	8,7	9,1
természettudományi	5,4	4,0

Az egyes képzési területek sajátosságainak feltárása érdekében és a képzési területek adatainak egymás közti összevetésére a cellasúlyt (suly1_1), a teljes, 10 képzési területet magába foglaló alapsokaság elemzéséhez a peremsúlyozással korrigált súlyváltozó (vegsulyminta) használatát javasoljuk.