

**Balácsi Ildikó**

## **Az OECD PISA, IEA TIMSS és PIRLS adatbázisainak bemutatása**

Ebben az előadásban szeretném bemutatni az adatbázisok felépítését, valamint a méréseknek és az adatelemzési szabályoknak néhány alapvető jellemzőjét. Véleményem szerint azoknak, akik ezekből az adatokból másodelemzéseket vagy számításokat szeretnének végezni, szükségük van ezekre az ismeretekre.

### **A tartalmi keret, a füzetek összeállítása**

Minden méréshez készül tartalmi keret, amely meghatározza, hogy mit is mérünk, milyen típusú feladatok szerepelnek a felmérésben és milyen arányban. A tartalmi keret azt is meghatározza, hogy milyen egyéb változók szerepelnek a mérésben, valamint hogy milyen indexeket készítenek majd ezekből a változókból. A mérések (PISA, PIRLS, TIMSS) tartalmi keretei elérhetőek és nyilvánosak. A nemzetközi mérésekhez szigorúan követendő standardizált eljárásrend készül, amit minden országnak kötelező követnie, és az előírt eljárásrend betartását ellenőrzik is. Hasonlóan szabályozott a feladatokat tartalmazó füzetek összeállítása, a megoldások kódolása és a nemzetközi mérések minden részlete.

A mérésekben sokkal több feladatot mérünk, mint amit egyetlen gyerek meg tudna oldani. A PISA-ban például hat és fél órára való feladat van, amelyet 13 füzetbe rendezünk. Minden tanuló egy füzet feladatait oldja meg. A füzetekben négy feladatblokk van, amelyek megoldására fél-fél órát szánunk. Összesen 13-féle feladatblokk készül, egy blokk négy füzetben szerepel.

Az összes ilyen mérésre jellemző, hogy a tanulók nem pontosan ugyanazokat a feladatokat oldják meg, ugyanakkor a füzetek feladatállományában vannak átfedések. Előfordul az is, hogy az egyes füzetek nehézségi szintjei nem teljesen egyformák. A nemzetközi mérések fontos tulajdonsága, hogy trendeket igyekszik feltárni, nem egyetlen időpillanatban akarja mérni az oktatási rendszereket. Az egymást követő mérésekben nem feltétlenül ugyanazok a feladatok szerepelnek, ugyanakkor vannak átfedések, ezért is titkos a feladatok többsége.

## Statisztikai elemzések

### *A statisztikai modellek*

A statisztikai elemzéshez különböző modelleket használnak. Az úgynevezett részpontoszám-modelleket<sup>1</sup>, valamint a tippelési paraméter-modelleket<sup>2</sup> többnyire vegyesen szokták használni. A feladatok között vannak feleletválasztós feladatok, egy pontos nyílt végű feladatok és több pontos nyílt végű feladatok. Minden feladat eredményére a megfelelő görbét próbálja illeszteni a használt statisztikai program, és a megfelelő itemparamétereket számolja ki hozzá. A nemzetközi mérések esetén a különböző statisztikai modelleknek az egyes változatait használják. A PISA-ban a RASH-modellel számolnak, a PIRLS-ben és a TIMSS-ben két- és három-paraméteres modelleket használnak a feladat típusától függően. A részpontoszám-modellt mind a kettőben alkalmazzák. A RASH-modellek használatához szükséges, hogy a feladatok meredeksége közel azonos legyen, vagyis hogy minden feladat megoldási valószínűsége közel azonos ütemben emelkedjen a gyengébb képességű tanulóktól a jobb képességű tanulók felé haladva. A feladatok kiválogatásánál és a próbamérésnél is ezt figyelembe veszik a teszt összeállítói, vagyis a próbamérések eredményeit felhasználva ennek megfelelően válogatják a feladatokat a füzetekbe. A másik két mérésnél, mivel ott más statisztikai programot használnak, lehetnek a meredekségek különbözőek. Minden mérés esetén a feladatok nehézsége és a tanulók képessége közötti különbség az, ami számít, amit ki tudunk mutatni.

Az egyes mérések esetén a célpopuláció vagy egy teljes évfolyam, vagy egy teljes korosztály. Természetesen ezek közül mindig csak azok a tanulók, akik benne vannak az oktatási rendszerben, hiszen a mérések az iskolákban történnek, azokat a gyerekeket, akik nincsenek az oktatási rendszerben ezek a mérések el sem érik.

### *A mintaválasztás*

A nemzetközi mérésekben mindig egy kiválasztott mintát mérnek. A mintaválasztás két lépcsőben történik. Először iskolákat választanak. Vannak olyan iskolák, akiket eleve kizárnak a mérésből. Egyes országok esetén bizonyos földrajzi helyzetben lévő iskolákat zárnak ki, máshol esetleg a nagyon kis iskolákat vagy a speciális tanulókat oktató

---

<sup>1</sup> A részpontoszám-modellt a többpontos feladatoknál alkalmazzák. A modell azt az összefüggést írja le, hogy egy többpontos feladat különböző pontszámait mekkora valószínűséggel éri el a tanulók képességeik függvényében.

<sup>2</sup> A tippelési paraméter-modellt a feleletválasztós feladatok esetében alkalmazzák. A modell figyelembe veszi a feladat megoldási valószínűségének számítása során, hogy a tanulók tippeléssel is eltalálhatják a jó választ.

intézményeket. A mintaválasztásnál különböző szempontokat vesznek figyelembe, különböző strátumokat használnak. Ezek között vannak implicit és explicit strátumok. Implicit strátumnak nevezzük, ha külön-külön választanak mintát, vagyis ha valamelyik ország egyik részpopulációt mélyebben szeretné elemezni, és ezért nagyobb mintát választanak abból a strátumból. Ilyen lehet például, ha az általános iskolából több tanulót választanak, mint amennyi az országos populáció arányaiból következne, mivel róluk külön is érvényes állításokat szeretnének megfogalmazni az adott országban. Az explicit strátumok szerint igyekeznek a mintavétel során olyan arányban választani a tanulókat, ahogy azok az egész populációban is találhatóak. A mintaválasztás során előfordul, hogy kellő információ hiányában olyan típusú iskolák kerülnek a mintába, akiket eredetileg kizártak onnan. Ilyenkor ezeket utólag zárják ki. Mindig bekerülnek a mintába olyan iskolák is, akik nem hajlandók részt venni a felmérésben.

A második lépcsőben az iskolaválasztás után az osztályok kiválasztása történik a TIMSS- és a PIRLS felmérésben. A PISA felmérés esetén, ahol nem osztályokat, hanem a 15 éveseket mérjük, ott iskolánként 35 tanuló kerül kiválasztásra, egyszerű mintavétellel. Vannak a mintaválasztás utáni tanuló-kizárások, ahol az SNI-s tanulókat és különböző más tanulócsoportokat utólag zárunk ki. Ezek a kizárások azonban nem haladhatják meg a minta 2,5%-át.

A válaszadási arányokra és a részvételi arányokra is fontos szabályok vonatkoznak. A PISA-ban a kiválasztott iskolák 85%-ának és a kiválasztott tanulók 80%-ának részt kell vennie a felmérésben ahhoz, hogy az eredmények megjelenjenek a nemzetközi jelentésben és az adatbázisban. A PIRLS- és a TIMSS-jelentésekben megjelennek azok az országok is, ahol a mintaválasztás valamiért nem volt megfelelő, és ezeknek az országoknak az eredményei mellett jelzik, hogy a kiválasztás mely szempontjából nem érték el a megfelelő szintet. A megfelelő arányok eléréséhez pótlólag is lehet iskolákat választani (előre kiválasztott tartalék iskolákból), de a jelentésekben az is megjelenik, hogy csak pótiskolákkal sikerült elérni a kívánt arányokat. A PISA-mérés esetében, ha sok tanuló hiányzik a felmérés napján, lehetőség van egy pótmérés megtartására.

### *A súlyozás*

Minden mintaválasztási egységhez – iskola, osztály, tanuló – tartozik egy súly, egy érték, ami egyrészt függ attól, hogy mekkora valószínűséggel kerülhet be az adott egység a mintába a

mintaválasztás során, másrészt függ a válaszadási aránytól. Akik nem válaszolnak, azoknak az eredményét pótolni kell, erre szolgál a súlyozás. Ha a válaszadási arány 100%-os, vagyis mindenki válaszolt, akit kiválasztottunk (megírta a felmérést és van eredménye), akkor a súly az adott mintavételi egységre - mondjuk egy iskolára - az összes lehetséges minta száma osztva azoknak a mintáknak a számával, amelyekbe az iskola belekerülne. Persze a végén csak egyetlen mintánk van, de az nagyon fontos, hogy az adott populációra érvényes becslést kapjunk. A súlyok biztosítják, hogy a populációra vonatkozó becsléseink torzítatlanok legyenek. A teljes körű felméréssel kapott érték, a censusérték az, amire kíváncsiak vagyunk. Ha nem használjuk a súlyokat, akkor nem a tényleges populációra vonatkozó értéket becsüljük, hanem a mintaválasztás szabályaiból adódó lehetséges minták becsléseinek átlagértékét. Ha például valamelyik csoportot felülreprezentáltuk, akkor a súlyozás nélküli átlageredményben a felülreprezentált csoport nagyobb súllyal szerepel, mint amekkora arányban a teljes populáció átlageredményében megjelenne, így a súlyozatlan átlageredmény nem jó becslése a tényleges átlageredménynek.

### *Hibaszámítás*

A kétlépcsős mintaválasztás – iskola, tanuló – miatt az egyszerű hibaszámítási eljárások nem használhatók a méréseink hibájának kiszámítására. Ez azért van, mert a statisztikai programokba épített hibaszámítási eljárások akkor használhatók, ha egyszerű, véletlen mintát választunk. Ez esetünkben akkor valósulna meg, ha az összes tanulóból teljesen véletlenszerűen választottuk volna ki a mérésben részt vevő 5000 tanulót. Így ezek a hibaszámítási eljárások a mi mintáinkból téves eredményeket adnak, hiszen az egy iskolába, vagy az egy osztályba járó tanulók eredményei erősen összefüggnek, tehát nem független a mintánk. Ha egy osztály 30 tanulója mérjük, akkor azok eredményei jobban fognak hasonlítani egymásra, mintha véletlenszerűen választottunk volna ki 30 tanulót, egymástól függetlenül.

A másik probléma ezekkel a hibaszámítási eljárásokkal, hogy a mérőeszköz hibáját nem veszik figyelembe. Mi egy tesztet íratunk sok feladattal, ezeknek a mérőeszközöknek van hibája. Ugyanaz a tanuló ugyanazt a tesztfüzetet vagy egy másik tesztfüzetet ugyanabban a felmérésben vagy egy másik felmérésben más eredménnyel írta meg. Ezt a hibát nem veszi figyelembe az SPSS hibaszámítása. A mérési eljárásból adódó hibát úgy lehet pontosítani, hogy ismételt hibaszámítási eljárásokat alkalmazunk, amely megengedi, hogy összefüggés

legyen a tanulók között (egy iskola adott osztályából valók, vagy egy iskola 35 azonos korosztályú tanulója). A mérőeszközök hibájában pedig plauzibilis értékeket<sup>3</sup> alkalmazunk

\*

A mintaválasztásnál az iskolatípus és a településtípus szokott lenni az implicit strátum, attól függően, hogy melyik évfolyamot mérjük. Az explicit strátumoknál a régiókat is figyelembe vesszük. A súlyozás pedig mindig úgy történik, hogy a tanulóknak a teljes populációban adott jellemzőit visszaadják, a tanulói populáció legyen teljes.

### **A háttérkérdőívek**

A háttérkérdőívek kérdéseit is előre meghatározott és dokumentált szempontok szerint választják ki. Minden felméréshez vannak *tanulói kérdőívek*, amelyben különféle demográfiai, a tanulmányokkal kapcsolatos, szociális, kulturális helyzetet felmérő, vagy a teszt által felmért területekkel kapcsolatos kérdések szerepelnek. Minden háttérkérdőív nyilvános. Minden méréshez kapcsolódik egy *iskolai kérdőív*, hiszen az iskola az egyik kiválasztási egység. Ebben a kérdőívben az iskolavezetéssel, az iskola felszereltségével kapcsolatos kérdések szerepelnek. Természetesen ezek is előre kidolgozottak és dokumentáltak. Vannak olyan mérések, amelyekhez *szülői kérdőív* is tartozik. A 2009-es PISA-méréshez van, és a PIRLS-mérések mindegyikéhez tartozik. A *tanári kérdőív* nagyon fontos része az IEA-méréseknek.

A két mérés eltérő céljaiból adódóan nagyon sok a különbség az IEA által kezdeményezett és a PISA-mérések között. Az IEA-mérések főként a kutatók, az oktatáskutatók, a tanárok, az oktatással foglalkozók számára készülő mérések. Így ezekhez egy *tantervi kérdőív* is tartozik. Ezekben a mérésekben ugyanis nagy hangsúlyt fektetnek arra, hogy mit is oktatnak az iskolában. Így olyan tesztek állítanak össze, amelyek a részt vevő országoknak a tanterveiből építkeznek, az a kiindulási alap. A PISA-mérés összeállításánál nem foglalkoznak a tantervekkel, így azokban sem tanári, sem tantervi kérdőív nincs, és a mérési területek sem kapcsolódnak a tantárgyakhoz, illetve a tantervekhez. Sokkal inkább alkalmazásképes tudást mérnek. Ez a megrendelőnek az igényeihez illeszkedik. Az IEA egy oktatáskutató társaság, az OECD pedig egy gazdasági társaság. A háttérkérdőívek alapján nagyon sok indexet számítanak mind a kétfajta mérés esetében, ezek megtalálhatók az adatbázisokban.

\*

---

<sup>3</sup> A plauzibilis értékek a tanuló teszt eredménye alapján becsülünk lehetséges képességpontokat, egy-egy tanulóra több, rendszerint 5 ilyen valószínű képességpontot számítanak a nemzetközi mérésekben.

Az IEA honlapjain minden mérés eseményeit, eredményeit követni lehet. Ilyenek: a mérések menetrendjei, a résztvevők, a nemzetközi jelentések, a résztvevők oktatási rendszerének bemutatása, a mérésben alkalmazott módszerek. Megtalálhatóak a kérdőívek – tanulói, iskolai, szülői. Lehet az adatok alapján eredményeket kiszámítani, az elkészült elemzéseket elolvasni és tovább értelmezni. Lehetséges ez az egyes részterületek esetén, valamint országonként is. Az IDBAnalyzer programot biztosítja ehhez a PIRLS- és a TIMSS-mérést elkészítő bizottság. Ez a program alkalmas a korábban elmondott szabályok betartására. A PISA-adatbázis elemzése annyival egyszerűbb, hogy ahhoz az SPSS-t használják, makrók segítségével. Annyival bonyolultabb, hogy nem menüsorból irányítható, hanem az alkalmazásához parancssorban kell megadni az utasításokat.