Nem-paraméteres statisztikai módszerek alkalmazási lehetőségei a pedagógiai kutatásban

A társadalomtudományok, így a pedagógia is, igen széles körben használnak matematikai statisztikai módszereket. A pedagógia számtalan területén – a tudásszintmérő tesztek elemzésétől az attitűdvizsgálatokig – sokféle elemzéshez nélkülözhetetlenek a korábban csak a természettudományok által használt matematikai statisztikai eljárások.

Magyarországon a pedagógia csak mintegy negyven éve – elsősorban *Kiss Árpád* tevékenységének köszönhetően – kezdte „felfedezni" a pedagógiamérést, és ezzel együtt az eredmények feldolgozásához szükséges statisztikai módszereket. Azóta több monográfia született a pedagógiában alkalmazott matematikai statisztikai eszközökről. *Ágoston György, Nagy József* és *Orosz Sándor* ma már klasszikusnak számító könyve elsősorban a tudásszintmérők adatainak feldolgozásához szükséges alapvető eljárásokkal, alapfogalmakkal foglalkozott (átlag, szórás, korreláció stb.). *(1)* A legújabb áttekintés sem helyez nagy súlyt arra, hogy ma az esetek túlnyomó részében számítógéppel történik az adatfeldolgozás, és így a statisztikai próbák kiszámítása is. *(2) Nahalka István* ismertet néhány nem-paraméteres próbát; a nem-paraméteres eljárások azonban korántsem nevezhetők elterjedtnek a pedagógiai kutatók körében. Jelen tanulmány célja az, hogy még több módszer ismertetésével és a gyakorlati alkalmazási lehetőségek felvázolásával újabb lépést tegyünk ennek megváltoztatására, hiszen a pedagógia számára releváns pszichológiai szakirodalomban gyakran találkozhatunk az ordinális és nominális változók elemzésére alkalmas módszerekkel (*Wilcoxon*-próba, *Kruskall–Wallis*-, vagy *Kolmogorov–Szmirnov*-próba stb.)

Paraméteres és nem-paraméteres statisztikák

Mit jelent a „nem-paraméteres" szóösszetétel? *Vargha András* és *Varbanova Mária* definíciója szerint akkor beszélünk paraméteres eloszlásösszességről, ha annak bármely elemét véges sok mennyiség teljesen meghatározza. Ilyen eloszlásösszesség például a normális eloszlások halmaza, amelynek elemeit egyértelműen meghatározza egy tetszőleges valós szám (a várható érték) és egy tetszőleges pozitív szám (a szórás). Vagyis véges sok (két) mennyiség teljesen meghatározza a normális eloszlások összességének bármely elemét.

Ezzel szemben a nem-paraméteres eloszlásösszesség esetében nem elegendő véges sok paraméter (az eloszlást jellemző mennyiség) megadása az eloszlás meghatározásához. A „nem-paraméteres" kifejezés nem paraméterek hiányára utal, hanem arra, hogy azokból nem elegendő véges sokat kiválasztani az eloszlás megadásához.

A statisztikai módszereket aszerint nevezzük paramétereseknek, illetve nem-paramétereseknek, hogy a statisztikai eljárás során milyen típusú eloszlásösszességgel számolunk.

A paraméteres statisztikai eljárások a pedagógiai kutatásokban igen széles körben elterjedtnek mondhatók. Ide tartozik például a korreláció-számítás, a t-próbák, a variancia-analízis. A nem-paraméteres statisztikai módszerek kevésbé ismertek, aminek három fő okát említjük: 1. a tudással kapcsolatban mért adataink gyakran jól közelítik a normális eloszlást; 2. a nem-paraméteres módszerek (statisztikai értelemben vett) érzékenysége kisebb; 3. a szakmai közösség számára könnyebben tálalhatók az eredmények az ismert statisztikai eljárások nyelvén.

Általánosságban azt mondhatjuk, hogy a nem-paraméteres módszerek minden olyan esetben is használhatók, amikor paraméteres módszert használhatunk. Fordítva azonban ez nem áll: hibát követhetünk el, ha paraméteres módszereket használunk ott, ahol nem ismerjük, nem ismerhetjük a változónk eloszlását. Fontos feladatunk ezért, hogy tisztázzuk, meddig terjed a paraméteres statisztikák hatóköre a pedagógiai kutatásokban, és amennyiben azok használatára nincs mód, úgy megtaláljuk a megfelelő nem-paraméteres eljárást.

Skálák és normalitás

A mérések során használt skálák típusairól bőségesen találunk leírást a szakirodalomban. *(3)* A négy legfontosabb skálatípus a nominális (névleges), az ordinális (rang-), az intervallum- és az arányskála. Az arányskála pedagógiai mérésekben csak elvétve szerepel. A másik három skálatípus esetében nem a mérés pontossága szerinti hierarchiáról van szó: egy névleges adat (pl. a tanuló neme) is lehet precízen meghatározott, csakúgy, mint egy intervallumskálán elhelyezhető adat (pl. a tanuló mennyi idő alatt futja le a száz métert). A skálák elkülönítése azon alapul, hogy milyen műveleteket végezhetünk a számokkal anélkül, hogy elszakadnánk a mért tulajdonságok közötti viszonyoktól. A nominális skála esetében a megegyező számértékek azonos tulajdonságot jelölnek, de a számok nagyságrendi viszonyainak nincs valódi jelentéstartalma. Példa erre egy pedagógiai vizsgálatokban sokszor szereplő változó, az iskolák sorszáma. Rangskálák esetében a számok nagyság szerinti sorrendje a mért tulajdonságok értékítéleten alapuló sorrendjét tükrözi, de a számok közti különbségek nem fejezik ki, hogy a mért tulajdonságok „értékessége" milyen mértékben különbözik egymástól. Intervallumskála esetén a számok közötti különbségeknek, arányskála esetén a számok arányainak is jelentéstartalma van.

Ebből a rövid leírásból is kitűnik, hogy a skálatípus meghatározása leginkább akkor okozhat gondot, ha rang- és intervallumskála között kell döntenünk. Sok esetben tisztán elméleti probléma, hogy a számok közötti különbségek nagysága tükrözi-e a mért objektumok tulajdonságai közötti különbségek volumenét. Például: Milyen mérési skálán helyezhetők el az osztályzatok? Gyakran találkozunk olyan változókkal, amelyekről nehéz eldönteni, hogy rang- vagy intervallumskálán helyezhetők-e el a számadatok. Ilyen lehet még a tantárgyak kedveltségének skálája és a szülők iskolai végzettségét vagy a továbbtanulási szándékot számszerűsítő mutató.

Bizonyos esetekben – példa rá az attitűd skálázásának problémaköre *(4)* – az a cél, hogy az elméleti modell alapján rangskálán levő adatokat intervallumskálán levőkké értékeljük föl. Egy lehetséges eljárás az egyenlőnek tűnő intervallumok módszere. *Kindler József* és *Papp Ottó* egyéb technikákat ismertet: a *Churchman–Akoff*-eljárást és *Guilford* módszerét.*(5)*

Azok a módszerek, amelyekkel rangskálából intervallumskála készíthető, sem az osztályzatokkal, sem a tantárgyak kedveltségével kapcsolatban nem használhatók. A Churchman–Akoff-eljárás a választási lehetőségek additivitására épül. Abszurdum lenne az osztályzatokkal kapcsolatban azt feltételezni, hogy elegendő sok gyenge osztályzat többet ér, mint egy jeles. Ez az eljárás inkább ételek vagy szabadidős programok rangsorolásakor működik. („Jobban szeretem a sarokházat a somlói galuskánál és a krémesnél, de a két utóbbi együtt már lehet, hogy többet ér egy darab sarokháznál.") A Guilford-eljárás páronkénti összehasonlításokból indul ki az egyes fokozatok közötti távolságok megállapításánál.

Ez szintén nem tűnik járható útnak, mert feltehetőleg akárhány embert is kérnénk föl a rangsorolásra, nagy valószínűséggel mindenki azt mondaná, hogy az ötös jobb, mint a négyes, a négyes pedig jobb, mint a hármas. A szélsőséges vélemények döntenék el tehát, hogy mekkora „távolság" van az egyes osztályzatok között. Az egyenlőnek tűnő intervallumok módszere a leginkább használható (bár munkaigényes) eljárás, erről azonban bebizonyosodott, hogy nem ad valódi intervallumskálát. *(6)*

Dönthetünk úgy, hogy az osztályzatokat intervallumskálán levőknek tekintjük. Két jó okunk is van erre. Egyrészt hivatkozhatunk a centrális határelosztás tételére (l. később), hiszen minden eddigi kutatási eredmény szerint az osztályzat rendkívül sok (talán túl sok) tényező eredőjének tekinthető. Másrészt, ha az osztályzatot teszten elért eredmény alapján adjuk, akkor – mivel a tesztpontszámokról számos esetben bizonyítható, hogy közel normális eloszlást adnak – elegendő a kvázi-normalitás biztosításához, hogy az osztályzatok megállapításának alapjául szolgáló ponthatárokat a megfelelő helyen húzzuk meg; ott, ahol a normalitás megőrzése azt megkívánja.

A másik megoldási mód szerint az osztályzatokat mint a társadalmi közeg számára értékítéletet hordozó kulturális-interakciós terméket rangskálán elhelyezkedőnek tekintjük. Az 1948 előtti osztályozási gyakorlat (amikor az alacsonyabb számérték jobb osztályzatot jelentett) sokkal inkább tükrözte ezt a felfogást, mint a ma megszokott módszer.

Miért fontos mindez? – kérdezhetjük, hiszen az előző pontban még a paraméteres és nem-paraméteres statisztikai módszerekről beszéltünk. Azért fontos ismernünk a pedagógiai mérésekben alkalmazott skálákat, mert a normális eloszlás két paramétere csak intervallum- és arányskálán levő adatok esetében számítható ki korrekt módon. Az átlag és a szórás kiszámítása nominális vagy rangskála esetén nem megengedett művelet.

A pedagógiai kutatásokban alkalmazott legismertebb statisztikai próbák esetében a próba alkalmazásának feltétele, hogy az eloszlás normális legyen. A normalitás bizonyítása sokszor nehézségekbe ütközik, de a t-próba nem túlságosan érzékeny a normalitási feltételre. *(7)* Emellett sokszor támaszkodhatunk a centrális határeloszlás tételére, *(8)* amely szerint több független (legfeljebb kismértékben összefüggő) azonos eloszlású változó (egyik sem lehet domináns) összegének eloszlása normális. A pedagógiai kutatásokban vizsgált tényezők nagyon gyakran tekinthetők sok-sok egymástól független, viszonylag azonos erősségű hatás eredőjének, ezért sokszor feltételezhetjük az eloszlás normalitását.

Mely pedagógiai vizsgálatok körében alkalmazhatók a nem-paraméteres statisztikai módszerek?

Mindenekelőtt ismételten hangsúlyozzuk, hogy a paraméteres eljárások helyett mindig alkalmazhatunk nem-paraméteres módszereket. A nem-paraméteres becslések és próbák ereje azonban 5–15%-kal kisebb, mint a megfelelő paraméteres próbáké, *(9)* ami azt jelenti, hogy mintegy 5–18%-kal nagyobb mintára van szükség ugyanolyan szignifikanciaszintű megállapításokhoz. Gazdaságossági megfontolások tehát a paraméteres próbák mellett szólnak, vagyis, ha csak lehet, azokat érdemes használnunk. A következőkben összegyűjtjük azokat az egymást részben átfedő eseteket, amelyekben nem használhatók a megszokott paraméteres eljárások.

Elsősorban azokat az eseteket kell számításba vennünk, amikor a változónk nem-paraméteres eloszlású.

Másodsorban azok az elemzések igényelhetik a nem-paraméteres módszerek alkalmazását, amelyekben nincs átlag, hanem a medián tölti be a várható érték becslésére szolgáló középérték esetét. Ebben az esetben a számadataink rangskálán helyezkednek el.

Harmadsorban az olyan változók jönnek számításba, amelyek két lehetséges számérték valamelyikét veszik föl. A matematikai leírás szerint ezek az úgynevezett dichotóm változók ugyanis egyszerre tekinthetők nominális, ordinális és intervallum-arányváltozónak is. Ezért a statisztikai próbák egy külön csoportja vonatkozik rájuk.

Negyedsorban a nominális nem-dichotóm változók esetében is nem-paraméteres próbák jöhetnek számításba, éspedig a módusz megállapításán túl elsősorban a kereszttáblás összefüggésvizsgálatok.

Ötödsorban pedig az olyan összefüggésvizsgálatokra kell figyelnünk, amelyekben különböző skálán helyezkednek el a mért adatok. Ilyen esetekben (a Liebig-féle minimumelv mintájára) a legalacsonyabb rendű skálához kell alkalmazkodni.

Fontosabb nem-paraméteres statisztikai módszerek

Három fő szempont alapján választottuk ki azokat a módszereket, amelyek alkalmazási lehetőségeit a továbbiakban ismertetjük:

1. igyekszünk a legismertebb paraméteres próbák és becslések nem-paraméteres megfelelőit bemutatni;
2. az 1. pont alapján kiválasztott módszereken túl azokat mutatjuk be, amelyeknek fontos szerepük lehet bizonyos pedagógiai problémák statisztikai elemzésében;
3. figyelembe vesszük a talán legismertebb statisztikai adatfeldolgozó program (az SPSS szoftver) által nyújtott lehetőségeket, és egyúttal a program korlátait.

A nem-paraméteres módszerek ismertetése e tanulmányban sem többet, sem kevesebbet nem jelent, mint azt, hogy megmondjuk, a becslés vagy próba milyen esetekben alkalmazható, és konkrét eseteket sorolunk fel, amikor a pedagógiai kutatás használhatja azokat. Nem vállalkozunk a képletek közlésére, a számítások részletezésére, és táblázatokat sem közlünk. *(10)*

Ordinális változókkal kapcsolatos eljárások

A rangskálán elhelyezkedő adataink elemzésére alkalmas módszereknek a legtöbb esetben megvan a megfelelő paraméteres analogonjuk. A rangskálán elhelyezhető adatok elemzésére szolgáló eljárásokat a nem normális eloszlású intervallumváltozók esetében is használhatjuk.

A következőkben Az iskolai tudás-vizsgálat *(11)* adatai segítségével mutatjuk be egyes nem-paraméteres módszerek lehetséges felhasználási területét.

A Wilcoxon-próbát a páros t-próba helyett használjuk ordinális adatok esetén. A 7. osztályosok esetében arra voltunk kíváncsiak, hogy van-e jelentős különbség az apa és az anya iskolai végzettsége között. Az SPSS szolgai módon elvégezte a páros t-próbát, és így megtudtuk, hogy az apák iskolai végzettsége átlagosan 2,96, az anyáké 3,07 (az iskolai végzettséget ötfokozatú skálán kellett bejelölni, a 3-as az érettségit jelöli). Az ezen adatok közti különbség nem értelmezhető, annak ellenére, hogy a t-próba szerint p=0,009 szinten szignifikáns az eltérés. A Wilcoxon-próba, nagyjából ugyanilyen valószínűségi szinten (p=0,008), annak bizonyítására alkalmas, hogy a két változó eloszlásának elhelyezkedése nem azonos. A mediánok egyenlők, mindkét esetben 3., de a Wilcoxon-próba megmutatta, hogy az eloszlások közti különbség alapján melyik változó esetében fordulnak elő általában a nagyobb rangszámok. A Wilcoxon-próba részeredményeinek, a rangszámok különbségeinek, összegeinek is valóságos jelentéstartalmuk van, ami nem volt elmondható a t-próbánál.

Az előjel-próba szintén a páros t-próba nem-paraméteres megfelelőjének tekinthető. A Wilcoxon-próbától abban különbözik, hogy a rangszámok eltéréseinek előjelét teszteli. Ez azt jelenti, hogy megvizsgálja, egy adott mintaelem esetén milyen a rangok különbségének előjele. A pozitív és negatív előjeles esetek számának különbségéből következtet az eloszlások egyezőségére, avagy különbözőségére. A fenti példa esetében p=0,004 szinten mutatható ki előjel-próbával, hogy az apák és anyák iskolai végzettségének eloszlása nem azonos.

Ha két független mintánk van, akkor a kétmintás t-próba (vagy a Welch-próba) a leggyakrabban használt paraméteres próba a középértékek összehasonlítására. Ebben az esetben a *Mann–Whitney*-próba lehet a megfelelő nem-paraméteres eljárás. Egy klasszikus kérdés, hogy vajon a fiúk vagy a lányok szeretik-e jobban a matematikát. A hetedik osztályosok körében a Mann–Whitney-próba alapján p=0,111 szinten az eloszlások megegyezőek. Egy másik példa: vajon azonos eloszlású-e az iskolai eredményekkel való elégedettség mutatója a matematikából négyessel és ötössel rendelkezők körében? A Mann–Whitney-próba szerint p=0,001 szinten elvethető az eloszlások elhelyezkedésének egyezését kimondó nullhipotézis, a rangszámok összege alapján pedig azt mondhatjuk, a matematikából jelessel rendelkezők inkább elégedettek iskolai eredményeikkel, mint akiknek négyesük van ebből a tárgyból.

Sokszor több részmintát hasonlítunk össze egyszerre. Normális eloszlású intervallumváltozók esetén ezt a feladatot varianciaanalízissel oldjuk meg. Azok az okok, amelyek ellene szólnak annak, hogy ilyen problémát sok-sok Mann–Whitney-próbával oldjunk meg ordinális változók esetén, ugyanazok, mint amelyek a paraméteres esetben a sok-sok kétmintás t-próba ellen szólnak. *(12)*

Ordinális változók esetében több független mintán a *Kruskall–Wallis*-próba alkalmas az eloszlások egyezésének vizsgálatára. Vizsgálhatjuk például, hogy valamely tantárgy osztályzatai szerint különböznek-e az iskolai eredményekkel való elégedettség mutatójának eloszlásai. Láttuk, hogy a matematika esetében már a négyes és ötös tanulók részmintáin megdőlne a nullhipotézis. Ha elvetjük a nullhipotézist, gyakran feltételezhető valamilyen növekvő sorrend a részminták elkülönítésének alapjául szolgáló változó növekvő értékei szerint. Az ilyen – csoportok sorrendjét feltételező – hipotézisek vizsgálatára a *Jonckheere–Terpstra*-próba alkalmas. A nullhipotézis itt is az, hogy minden részmintán megegyezik az eloszlások elhelyezkedése, az ellenhipotézis ugyanakkor a Kruskall–Wallis-próbával szemben nem az, hogy valamely két csoport különbözik egymástól, hanem az, hogy az imént említett növekvő (pontosabban: nem csökkenő) sorrendbe rakhatók a részcsoportok.

Hetedikeseink történelemosztályzata szerint öt csoportot képezve, a Kruskall–Wallis-próbát, majd a Joncheere–Terpstra-próbát alkalmazva körükben, mindkét esetben azt kapjuk, hogy p=0,000 szinten el kell vetnünk a nullhipotézist. Ez annyit jelent, hogy általában minél jobb jegye van valakinek történelemből, annál inkább elégedett az iskolai teljesítményével. (A többi tantárgy esetében természetesen ugyanez a következtetés adódik.) Ha viszont az apa iskolai végzettsége szerinti öt csoportot hasonlítjuk össze abban a kérdésben, hogy „Véleményed szerint kinek van több természetes képessége a matematikához?”, akkor a Kruskall–Wallis-próba alapján p=0,069 szinten megtarthatjuk a nullhipotézist, vagyis az apa iskolai végzettségétől független az, hogy a tanulók hogyan vélekednek a fiúk és a lányok természetes matematikai képességeiről. A Jonckheere–Trepstra-próba ezek után arra alkalmas, hogy a p=0,709 szintet látva még azt is megállapítsuk, hogy a Kruskall–Wallis-próba által mutatott elhanyagolhatón kicsi különbségekben sincs tendencia.

Azok a módszerek, amelyekkel rangskálából intervallumskála készíthető, sem az osztályzatokkal, sem a tantárgyak kedveltségével kapcsolatban nem használhatók. A Churchman-Akoff-eljárás a választási lehetőségek additivitására épül. Abszurdum lenne az osztályzatokkal kapcsolatban azt feltételezni, hogy elegendő sok gyenge osztályzat többet ér, mint egy jeles.

A Kruskall–Wallis- és a Jonckheere–Terpstra-próbák egymástól függetlenül adnak információt a változókról. Előfordulhat, hogy a Kruskall–Wallis-próbánál megtarthatjuk a nullhipotézist, de a sorrendben mégis határozott tendencia rajzolódik ki. Ez történt például, amikor hetedikeseinket a matematika jegy szerint soroltuk be öt részmintába, és ismét csak a természetes matematikai képességgel kapcsolatos kérdést elemeztük. A Kruskall–Wallis-próba alapján megtartható volt a nullhipotézis, a Jonckheere–Terpstra-próba ugyanakkor megmutatta a sorrendi tendenciát, ami azt jelenti, hogy minél jobb jegye van egy hetedikesnek matematikából, annál inkább hajlamos arra, hogy a lányoknak több „természetes matematikai képességet” tulajdonítson.

Szintén több minta összehasonlítására alkalmas a medián-próba, amely az előjel-próba többváltozós általánosításának tekinthető.

Azonos mintán több változó összehasonlítása Wilcoxon-próbák sorozatával is elvégezhető, de léteznek olyan eljárások, amelyek egyszerre vizsgálják a változók eloszlásának elhelyezkedését. Ezek közül a legismertebb a *Friedman*-próba. Egy lehetséges felhasználási terület: adott tanulócsoportban több tantárgy kedveltségének összehasonlítása. Ha elvethető a nullhipotézis (azaz van legalább két tantárgy, amelynek kedveltsége jelentősen különbözik egymástól), akkor Wilcoxon-próbák sorozatával lehetséges az elméleti szempontból érdekes párosításokat tovább elemezni.

Dichotóm változók problémái

A legegyszerűbb ezek közül a χ^2-próba speciális esete, a 2x2-es χ^2-próba. Szintén a χ^2-eloszlást használja a *McNemar*-próba, amely két dichotóm változó eloszlását hasonlítja össze. Követelmény, hogy a két változó pontosan ugyanazokat a számértékeket vegye föl (leggyakrabban 0-t és 1-et). Az egyik fontos felhasználási terület a kritériumorientált értékelés lehet, mivel ott „teljesítette–nem teljesítette" dichotóm változót használunk. Az elő- és utóteszten elért eredmények összehasonlítására a McNemar-próba a legalkalmasabb.

A McNemar-próba általánosítása több változó esetére a *Cochran*-próba. Használhatjuk abban az esetben, amikor egy adott tanulócsoportban több időpontban mértünk. nem elsősorban tudással kapcsolatos dichotóm változókra kell gondolnunk itt, hanem például kétértékű attitűdskálára. Érdekes kérdés lehet például (ezzel kapcsolatos kutatásról nem tudok), hogy a hét különböző napjaink jelentősen eltér-e egymástól a tanulással szembeni attitűd. Amennyiben elvethető a nullhipotézis, a McNemar-próbával folytathatjuk az elemzést.

Dichotóm változókra alkalmazható a *Wald–Wolfowitz*-sorozatpróba (az SPSS-ben a „Runs" alpont). Azt a hipotézist tesztelhetjük vele, hogy a változó két értéke véletlenszerűen követi-e egymást. Olyan esetben használhatjuk, amikor a mintaelemek sorrendjének jelentősége van. Mintaelemeknek adott időpontokat tekintve vizsgálhatjuk például azt, hogy egy konkrét személy dichotóm változóval jellemezhető teljesítménye, véleménye véletlenszerűen változik-e az idő múlásával.

Szintén dichotóm változókra alkalmazható a binomiális próba. Azt tesztelhetjük vele, hogy a mintánk adatai alapján kijelenthető-e (adott valószínűségi szinten), hogy a változó valamely értéke bizonyos valószínűséggel fordul elő a populációban. Lehet például az a nullhipotézisünk, hogy kritériumorientált tesztelésnél a tanulók 15%-a kap „nem megfelelt" minősítést. Ha például a reprezentatív 100 fős mintán 21-en nem feleltek meg, akkor megtartható a 15%-ra vonatkozó nullhipotézis, ha 22-en, akkor már nem.

Nem-paraméteres összefüggésvizsgálatok

Nominális változók összefüggéseit a kereszttábla-elemzésekkel vizsgálhatjuk. Az egyik változó értékeit a sorok, a másikét az oszlopok szerint feltüntetve számadatokkal jellemezhető, hogy van-e „sűrűsödés" a téglalap egyes részein, vagy minden cellába nagyjából ugyanannyi elem jut. Ha egyes cellák üresen maradnak, vagy sok cellába túl kevés elem tartozik, akkor össze kell vonnunk a változók bizonyos értékeit. Ha például a lányok és a fiúk továbbtanulási terveit szeretnénk összehasonlítani, akkor lehetséges, hogy össze kell vonnunk a főiskolai és egyetemi végzettségre törekvőket.

A kereszttábla-elemzések minden típusú skála esetében használhatóak. A χ^2-próba az összefüggés szorosságának vizsgálatára sok esetben a legmegfelelőbb módszer. A kontingencia-koefficiens a χ^2 értékéből számítható, és értéke a korrelációs együtthatókhoz hasonlóan -1 és +1 közötti. Két dichotóm változó esetében a 2x2-es χ^2-próbát alkalmazzuk, ami csak annyiban más, mint az általános változat, hogy ha nem teljesülnek a χ^2-próba alkalmazásainak feltételei, akkor a *Fisher*-féle egzakt próbát használhatjuk annak eldöntésére, hogy van-e szoros összefüggés a változók között.

Ordinális változóink között az összefüggés szorosságát a *Spearman*-féle rangkorrelációs együttható mutatja. Ennek értékei általában kicsit magasabbak (abszolútértékben), mint a legtöbbször használt és ismertebb Pearson-féle együtthatóé. Egy másik elterjedt mérőszám a *Kendall*-féle rangkorrelációs együttható. Rangskálán levő adatsorok összefüggésének jellemzésére alkalmas a Kendall-féle konkordancia-mutató is. Használhatjuk például akkor, ha a tesztfejlesztés során szakértőkkel rangsoroltatunk feladatokat aszerint, hogy melyik maradjon ki az új változatból első-, másod-, harmadsorban stb. A konkordancia-mutató 0 értéke a vélemények teljes mértékű különbözőségét jelzi (ez csak két bíráló esetében fordulhat elő), az 1 érték maximális véleményegyezést mutat. Két bíráló véleményének egészét jellemezhetjük a *Cohen*-féle κ (kappá)-val is. Ez a mutató azért jelentős, mert kritériumorientált tesztek esetében reliabilitás-mutatóként szerepelhet. *(13)*

Amennyiben többféle skálán elhelyezkedő változók összefüggéseit vizsgáljuk, általános alapelv, hogy a kevesebb megengedhető matematikai művelettel rendelkezőhöz kell alkalmazkodni. Vannak esetek azonban, amikor két adott skálán levő adatsor vizsgálatára speciális módszer létezik. Például a nominális és intervallumváltozó közötti összefüggés szorosságát jellemzik az úgynevezett ε (éta)-mutatók.

Több hasonló, az összefüggés szorosságát jellemző mutató kiszámítására képes az SPSS. A legtöbb esetben egy olyan statisztikai próbát is rögtön elvégez a program, amely az adott mutató által jelzett összefüggés szignifikanciáját teszteli. Komoly hiányérzetünk lehet ugyanakkor, hogy pontbiszeriális és biszeriális együtthatókat nem számol. Ezek a dichotóm és intervallumváltozók kapcsolatának szorosságát jellemzik, és kiszámításuk a *Vargha András* által javasolt módon történhet *(14)*.

Jegyzet

(1) ÁGOSTON GYÖRGY–NAGY JÓZSEF–OROSZ SÁNDOR: *Méréses módszerek a pedagógiában.* Tankönyvkiadó, Bp. 1974.

(2) NAHALKA ISTVÁN: *A változók rendszerének struktúrája. = Bevezetés a pedagógiai kutatás módszereibe.* Szerkesztette: FALUS IVÁN. Keraban Kiadó, Bp. 1993.

(3) Pl.: ÁGOSTON GYÖRGY–NAGY JÓZSEF–OROSZ SÁNDOR: *Méréses módszerek...,* i. m.; KINDLER JÓZSEF–PAPP OTTÓ: *Komplex rendszerek vizsgálata.* Műszaki Kiadó, Bp. 1978; REUCHLIN, M.: *Mérés a pszichológiában. = A kísérleti pszichológia módszerei.* Szerkesztette: PIAGET, J.–FRAISSE, P.–REUCHLIN, M. Akadémiai Kiadó, Bp. 1985; NAHALKA ISTVÁN: *A változók rendszerezésének struktúrája,* i. m.; CSÍKOS CSABA: *Tudásszintmérő tesztekkel kapcsolatos alapkérdések.* Megjelenés előtt.

(4) SELLTIZ, C.–JAHODA, M.–DEUTSCH, M.–COOK, S. W.: *Az attitűd skálázása. = Az attitűd pszichológiai kutatásának kérdései.* Szerkesztette: HALÁSZ LÁSZLÓ, HUNYADY GYÖRGY és MÁRTON L. MAGDA. Akadémiai Kiadó, Bp. 1979.

(5) KINDLER JÓZSEF–PAPP OTTÓ: *Komplex rendszerek vizsgálata*, i. m.

(6) SELLITZ, C.–JAHODA, M.–DEUTSCH, M.–COOK, S. W.: *Az attitűd skálázása,* i. m.

(7) HAJTMAN BÉLA: *Bevezetés a matematikai statisztikába.* Akadémiai Kiadó, Bp. 1968.

(8) Lásd pl.: MÉRÔ LÁSZLÓ: *A pszichológiai skálázás matematikai alapjai.* Tankönyvkiadó, Bp. 1992.

(9) Lásd: VARGHA ANDRÁS: *Pszichológiai statisztika gyakorlat II.* Tankönyvkiadó, Bp. 1989.

(10) Ezek megtalálhatók: VINCEZ ISTVÁN–VARBANOVA MÁRIA: *Nem-paraméteres matematikai statisztika.* Akadémiai Kiadó, Bp. 1994; VARGHA ANDRÁS: *Pszichológiai statisztika gyakorlat...,* i. m.

(11) Az iskolai tudás. Szerkesztette: CSAPÓ BENÔ. Osiris Kiadó, Bp. 1998.

(12) Lásd: HAJTMAN BÉLA: *Bevezetés a matematikai statisztikába,* i. m.

(13) CSAPÓ BENÔ: *A kritériumorientált értékelés.* Magyar Pedagógia, 1987. 87. sz., 247–266. old.

(14) VARGHA ANDRÁS: *Pszichológiai statisztika...,* i. m.

Csíkos Csaba