A POPULÁCIÓRA VONATKOZTATOTT HASONLÓSÁG FOGALMÁRÓL*

J. PERKAL *(Vrocław)*

Jóllehet a (geometriai) hasonlóság az ókor óta ismeretes fogalom (szobrok, festmények, térképek, stb.), mégsem alkalmas arra, legalábbis elemi matematikai értelmében, hogy kifejezze az élő világ egyedei közötti hasonlósági viszonyokat. Nem létezik két geometriailag hasonló lény, kivéve azt az esetet, midőn azonos méretűek. Valóban, például ahhoz, hogy egy fa képes legyen nedveit cirkuláltatni, az szükséges, hogy törzsének keresztmetszete körülbelül arányos legyen v térfogatával, tehát átmérője $\sqrt{v}$-vel legyen arányos. Ilyen meggondolások alapján egy $4v$ térfogatú fa átmérője $\sqrt{4} = 2$-ször nagyobb kellene, hogy legyen. Mármost ahhoz, hogy geometriai hasonlóság megvalósuljon, ez az átmérő, mint lineáris méret, nem nőhet csak $\sqrt[3]{4} < 2$-szeresére.

Mindazonáltal, mint az közismert, a természettudósok a hasonlóságnak egy, az azonosságnál sokkal általánosabb fogalmát használják, és kétségkívül, ilyen fogalom szükséges is. Ennek nyilvánvalóan több különböző értelmet tulajdoníthatunk. Ez az előadás azt tűzi ki céljául, hogy két ilyen kísérletet javasoljon.

Legyenek $x, y, z, \ldots$ egy P populáció egyedei. Ezeket az egyedeket a k számú $c_1, c_2, \ldots, c_k$ mennyiség határozza meg, és úgy kezelhetők, mint a k-dimenziós euklideszi tér pontjai. Az xSy szimbólum azt fogja jelenteni, hogy x és y *hasonlóak*. Tegyük fel, hogy

1. ha $x = y$, akkor xSy (a hasonlóság általánosabb, mint az azonosság),
2. ha xSy, akkor ySx (a hasonlóság szimmetrikus reláció),
3. ha xSy és ySz, akkor xSz (a hasonlóság tranzitív reláció).

A tranzitivitás törvénye leszűkíti a természettudósok által használt hasonlósági fogalmat. Valóban, ők kieszelik a két nem-hasonló egyed közötti láncok létezését, mely láncokat olyan egyed-láncszemek alkotják, melyek mindegyike hasonló közvetlen szomszédjaihoz. A 3. tulajdonság alapján ez a lehetőség számunkra ki van zárva.

Két egyedet *azonosnak* fogunk tekinteni, ha mért tulajdonságaik $c_1, c_2, \ldots, c_k$ nagyságai révén megkülönböztethetetlenek, vagyis ha a k-dimenziós

* A Biometriai Symposionon (Budapest), 1959. szeptember 8-án elhangzott előadás szövege.

térben nekik megfelelő egyedi pontok egybeesnek. Két egyedet az S *értelemben hasonlónak* fogjuk nevezni, ha a nekik megfelelő egyedi pontok valamely előre definiált S család ugyanazon halmazában vannak; erről az S családról azt fogjuk mondani, hogy az az xSy relációt *indukálja* (tekintettel arra, hogy ez az indukció egyértelmű). Például a koordináták kezdőpontján áthaladó összes egyenesek S családja a geometriai hasonlóságot indukálja.

Ha megadtunk egy S családot, és következésképp az általa indukált xSy hasonlóságot, akkor csak egyetlen, az S hasonlóságot *invariánsul hagyó transzformáció* létezik, vagyis egyetlen olyan transzformáció van, melynél az x pont és transzformáltja, az S család ugyanazon halmazához tartozik. Ezt a transzformációt $S(x)$-szel jelöljük. Például abban az esetben, ha S a geometriai hasonlóság, ilyen transzformáció az $S(x) = a \cdot x$ transzformáció, ahol a állandó.

Minden mennyiséget, mely az $S(x)$ transzformációnak invariánsa, az S hasonlóság *indexének* fogjuk nevezni. Például a $\frac{c_i}{c_j}$ alakú hányadosok a geometriai hasonlóságnak indexei. Ezen túlmenően a $\frac{c_1}{c_2}, \frac{c_2}{c_3}, \ldots, \frac{c_{k-1}}{c_k}$ indexek ezen hasonlóság *teljes* index-rendszerét alkotják, vagyis valamely egyed összes ilyen indexeinek egy másik egyed megfelelő indexeivel való egyenlősége szükséges és elegendő feltétele geometriai hasonlóságuknak.

Mármost a hasonlóság szokásos értelme viszonylagosnak mutatkozik. Két olyan kínai, akiket Európa lakosainak populációjában hasonlóknak minősítenek, a kínaiak valamely populációjában lehetnek nem-hasonlók. Természetes dolog tehát a hasonlóság jelentését a populációhoz kötni.

Itt a populációra vonatkoztatott hasonlóság két fajtáját fogom megkülönböztetni. Az elsőt *deviációs hasonlóságnak*, a másodikat pedig *természetes hasonlóságnak* fogom nevezni.

Ha valamely populáció homogén, tekintethetjük annak *centrumát*, azaz egy olyan virtuális egyedet, melynek tulajdonsági mennyiségei egyenlők az átlagosokkal. Ezt jelölje p. Két egyedet, x-t és y-t *deviációsan hasonlóknak* fogunk nevezni, ha tulajdonságaik nagyságai tekintetében a p centrumtól arányos módon térnek el. Így, ha x és y a p-től a c_i (ahol $i = 1, 2, \ldots, k$) mennyiségekben térnek el, továbbá ezen differenciák állandók és i-től függetlenek, akkor x és y ebben az értelemben hasonló. Ha például egy ember a populáció centrumától csak orrának hosszában tér el, akkor mindazon emberek, akik hozzá deviációsan hasonlók, a centrumtól csak ebben a tulajdonságban fognak eltérni.

Egy karikatura lényege abban áll, hogy hangsúlyozzuk a modell jellemző vonásait, amit azáltal lehet elérni, hogy a modell és a centrum tulajdonságai közötti különbségeket arányosan szorozzuk. Ebben az esetben a karikatúra deviációsan hasonlít a modellre.

A deviációs hasonlóságot indukáló S halmaz-család a populáció p centrumán áthaladó összes egyenesek családja (és nem az origón áthaladó egyenesek családja, mint a geometriai hasonlóság esetében). Tehát két egyed deviációsan hasonló egymáshoz, ha a k-dimenziós térbeli egyedi pontjaikat összekötő egyenes a p centrumon át halad. Jelöljük p koordinátáit $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_k$-val. Az ezt a hasonlóságot invariánsul hagyó transzformáció (vektori) alakja a következő: $S(x)=a\,(x-p)$, vagyis numerikus alakban $S(c_i)=a\,(c_i-\bar{c}_i)$, ahol az a együttható állandó.

Ennek a hasonlóságnak az indexei a $\frac{c_i-\bar{c}_i}{c_j-\bar{c}_j}$ hányadosok, és a $\frac{c_1-\bar{c}_1}{c_2-\bar{c}_2}$, $\frac{c_2-\bar{c}_2}{c_3-\bar{c}_3}, \ldots, \frac{c_{k-1}-\bar{c}_{k-1}}{c_k-\bar{c}_k}$ rendszer ezen indexek teljes rendszere.

Mégis, a deviációs hasonlóság fogalma nem mindig bír érdekes természetes értelemmel. Tekintsük például különböző korú és két tulajdonság: termet és súly által jellemzett gyermekek populációját. Itt a centrum függ az életkorok populáción belüli eloszlásától, és következésképpen nem bír biológiai érdekességgel, éppúgy, mint a deviációs hasonlóság.

Ilyenkor a természetes hasonlóság fogalma az, mely lehetővé teszi, hogy a természettudós intuicióját jobban megragadjuk. Ez esetben a centrum szerepét nem pont, hanem vonal játssza.

Ha felosztjuk a gyermekek ugyanezen populációját életkorok szerinti csoportokra és megkeressük minden egyes csoport centrumát, akkor látjuk, hogy ezen centrumok egy vonal (a fejlődési vonal) mentén helyezkednek el. Azonban a populáció centruma más esetekben is meg lehet fosztva biológiai értékétől. A térben valamely populációt egy többé vagy kevésbé kiterjedt pontfelhő (mage des points) ábrázol. Mennél inkább gömb alakú ez a felhő, annál nagyobb szerep jut a populáció centrumának. Ezzel szemben mennél elnyúltabb ez a felhő, annál nagyobb szerepet játszik egy vonal. Ha úgy tekintjük, hogy a populációk nem túlságosan térnek el a normális eloszlástól, fel lehet tételezni, hogy ez a vonal egyenes (HOTELLING tengelye [1]).

Ekkor két egyént természetes értelemben hasonlónak tekinthetünk, ha egyedi pontjaik a főtengellyel párhuzamos egyenesen helyezkednek el.

Minthogy ezen tengely numerikus meghatározása rendszerint meglehetősen unalmas dolog, én ezt egy másik egyenessel helyettesítem, mégpedig azzal, melyet *természetes tengely*nek neveztem el (lásd [2]), és amelyet már *Teissier* (lásd [3]) is vizsgált. Ezen azt a tengelyt értem, melynek komponensei rendre a $c_1, c_2, \ldots, c_k$ tulajdonságok $\sigma_1, \sigma_2, \ldots, \sigma_k$ *szórásai.* Következéskép két egyedet *természetes értelemben hasonlónak* nevezek, ha a c_i tulajdonságaik értékei közötti különbségek ezen σ_i szórásokkal arányosak, állandó (i-től független) arányossági együttható mellett.

Ebből következően a természetes hasonlóságot indukáló család a természetes tengellyel párhuzamos egyenesek családja. Tehát azt lehet mondani, hogy két egyed ebben az értemben hasonló valamely populációra nézve, ha az egyedi pontjaikat összekötő egyenes párhuzamos ezen populáció természetes tengelyével. A természetes hasonlóságot invariánsul hagyó transzformáció a térnek a természetes tengellyel párhuzamos transzlációja. Invariánsai és így természetes indexei, a természetes tengellyel párhuzamos egyenesek tetszés szerinti jellemzői.

Képezzük ezen indexek egy rendszerét. Normáljuk e célból a $c_1, c_2, \ldots$ $\ldots, c_k$ tulajdonságokat olyan módon, hogy várható értékük 0 és szórásnégyzetük 1 legyen:

$$\gamma_i = \frac{c_i - \bar{c}_i}{\sigma_i} \qquad (i = 1, 2, \ldots, k).$$

Normáljuk még egyszer a γ_i-ket a

$$\bar{\gamma} = \frac{1}{k} \qquad (\gamma_1 + \gamma_2 + \cdots + \gamma_k)$$

átlaguk segítségével olyképp, hogy minden egyes egyed bármely tulajdonságának várható értéke 0 legyen:

$$w_i = \gamma_i - \bar{\gamma} \qquad (i = 1, 2, \ldots, k).$$

Az így definiált w_i-k a természetes hasonlóság számára index-rendszert alkotnak. Ez a rendszer nem teljes, mert az indexek nem függetlenek, tudniillik $w_1 + w_2 + \cdots + w_k = c$. Elegendő azonban belőle bármelyiket elhagyni ahhoz, hogy teljes rendszerré váljék. Bebizonyítottam (lásd [2]), hogy a most definiált természetes indexek egyenlősége valóban szükséges és elegendő feltétele két egyed természetes hasonlóságának.

MILICEROWA (lásd [4]) érdekes eredményeket ért el a természetes indexek alkalmazásával boxolók és más sportolók szomatikus analízisére.

IDÉZETT MŰVEK

[1] HOTELLING, H.: Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.*, **24** (1933) 417—441 és 498—520.

[2] PERKAL, J.: O wskaźnikach antropologicznych (lengyelül). *Przegl. Antrop.*, **19** (1953) 209—221. Összefoglalás franciául ibid. **20** 679—680.

[3] TEISSIER, G.: Allométrie de taille et variabilité chez Maia squinado. *Zool. Exper.*, **42** (1955) 221—264.

[4] MILICEROWA, H.: Zastosowanie wskaźników Perkala de charakterystyki budowi ciała bokserów. *Materiały i Prace Antrop.*, **20** (1956) 1—85. (Angolul) ibid. 71—77. Összefoglalás oroszul ibid. 78—84.