KÜLÖNLENYOMAT

A MAGYAR TUDOMÁNYOS AKADÉMIA
III. (MATEMATIKAI ÉS FIZIKAI) OSZTÁLYÁNAK
KÖZLEMÉNYEIBŐL

XII. KÖTET 4. SZÁM

ARATÓ MÁTYÁS

NÉHÁNY MEGJEGYZÉS AZ I-DIVERGENCIA FOGALMÁVAL KAPCSOLATBAN

1962

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK OSZTÁLYÁNAK
KÖZLEMÉNYEI

XII. kötet 4. szám

Szerkesztőség: Budapest, V., Nádor utca 7.

Kiadóhivatal: Budapest, V., Alkotmány utca 21.

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának Közleményei változó terjedelmű füzetekben jelennek meg, és az Akadémia III. Osztályának felolvasóülésein bemutatott matematikai dolgozatokat, valamint egyéb dolgozatokat, referátumokat, továbbá az Osztály munkájára vonatkozó közleményeket, könyvismertetéseket stb. közölnek. Évenként egy kötet jelenik meg (négy szám alkot egy kötetet).

Kéziratok a következő címre küldendők:

A Magyar Tudományos Akadémia
III. Osztályának Közleményei
Budapest, V., Nádor u. 7.

Ugyanerre a címre küldendő minden szerkesztőségi levelezés.

Minden szerzőt 100 különlenyomat illet meg megjelent munkájáért.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de felelősséget a beküldött kéziratok megőrzéséért vagy továbbításáért nem vállal.

A Közlemények előfizetési ára kötetenként belföldi címre 40 forint, külföldi címre 60 forint. Belföldi megrendelések az Akadémiai Kiadó, Budapest, V., Alkotmány u. 21. (Magyar Nemzeti Bank egyszámlaszám: 05-915-111-46), külföldi megrendelések a „Kultúra" Könyv- és Hírlap Külkereskedelmi Vállalat, Budapest, I., Fő utca 32. (Magyar Nemzeti Bank egyszámlaszám: 43-790-057-181) útján eszközölhetők.

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica
2. Acta Physica Hungarica.

NÉHÁNY MEGJEGYZÉS AZ I-DIVERGENCIA FOGALMÁVAL KAPCSOLATBAN

Írta: ARATÓ MÁTYÁS

VINCZE I. két cikkében [1], [2] foglalkozik az információ fogalmának értelmezésével és javasolja, hogy az $f(x)$ sűrűségfüggvényű ξ valószínűségi változónak $\varphi(x)$ érdeklődési sűrűség esetén az információja legyen értelmezés szerint

$$I_\varphi(\xi) = \int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{\varphi(x)}\, dx. \tag{1}$$

Ezt a fogalmat absztrakt terekben KULLBACK vezette be, és általános valószínűség-eloszlások esetén — melynek (1) speciális esete — I-divergenciának nevezte el, mely azóta is általánosan használt. Ez a fogalom mind az információelméletben, mind a matematikai statisztikában igen hasznosnak bizonyult. Az [1], [2] cikkekben szereplő irodalmon kívül az olvasó erről meggyőződhet, pl. PINSKER [3] könyve, vagy ROZANOV [4] dolgozata alapján. Az I-divergencia fogalma azonban elsősorban kisegítő fogalom, mely az információelmélet és a statisztika különböző meggondolásaiból, tételeiből adódik és amelynek segítségével igen szép tételek bizonyíthatók sok más területen is. Példaként megemlítem a következőket: annak a tételnek a bizonyítása, hogy két normális mérték szinguláris, vagy abszolút folytonos egymásra nézve, az I-divergencia tulajdonságainak a felhasználásával történik. (HAJEK [5], ROZANOV [4].) A fenti (1) képlethez hasonlóan történik bizonyos számelméleti leképezések (pontos endomorfizmusok) entrópiájának kiszámítása is (ROHLIN [10]).

VINCZE [2] dolgozatában szerepel még alkalmazásként konfidenciaintervallumok szerkesztése is, bizonyos eloszlások ismeretlen paramétereire. Nem foglalkozik azonban eljárásának elvi kérdéseivel, a módszer érvényességi körének vizsgálatával. Kérdés például, hogy az általa javasolt konfidencia halmazok jobbak-e, mint az eddig ismertek, és hogy az általa javasolt eljárás esetén mennyiben válik egyszerűbbé a megfelelő valószínűségi szintek kiszámítása. Nem látható be, hogy az említett elv alapján kapható-e az ismeretlen paraméterre becslés, és hogy a kapott halmazok mely esetben lesznek konfidencia halmazok.

Az alábbiakban ismertetem WILKS, még 1938-ban végzett azokat a vizsgálatait [6], amelyekben az aszimptotikusan l[illegible]obb konfidencia halmaz szerkesztésének a problémájával foglalkozik. WILKStől eltérően a jelenleg szokásos jelöléseket használom, s az alábbi — WILKStől származó — állítás lényegesen egyszerűbb bizonyítását adom.

TÉTEL: *Legyen a $P_\theta(\cdot)$ mértékek egy $\theta \in \Theta$ paramétertől függő serege abszolút folytonos a P_0 mértékekre nézve és jelöl[illegible]*

$$L_\theta(\omega) = \log \frac{P_\theta(d\omega)}{P_0(d\omega)}$$

az ún. likelihood hányadost. Ismeretes, hogy $M_\theta \frac{\partial}{\partial\theta} L_\theta(\omega) = 0$, *másrészt* $\sigma(\theta) = D \frac{\partial}{\partial\theta} L_\theta(\omega) = -M_\theta \frac{\partial}{\partial\theta} \varphi_\theta(\omega)$, *ahol* $\varphi_\theta(\omega) = \frac{1}{\sigma(\theta)} \frac{\partial}{\partial\theta} L_\theta(\omega)$ (felhasználva, hogy $M_\theta \frac{\partial^2 L_\theta(\omega)}{\partial\theta^2} = -\sigma^2(\theta)$). Ezen összefüggések hasonlóan bizonyíthatók, mint CRAMER [7] könyvének 502. oldalán levő összefüggések, természetesen bizonyos megszorítások esetén, melyeket nem részletezünk. *Legyen* $\psi_\theta(\omega)$ *tetszőleges az* $M_\theta\psi_\theta(\omega)=0$ *és* $D\psi_\theta(\omega)=1$ *feltételeknek eleget tevő függvény, akkor igaz a következő egyenlőtlenség*

$$\left| M_\theta \frac{\partial\psi_\theta(\omega)}{\partial\theta} \right| \leqq -M_\theta \frac{\partial\varphi_\theta(\omega)}{\partial\theta} = \sigma(\theta).$$

BIZONYÍTÁS: A $p_\theta(\omega) = \frac{P_\theta(d\omega)}{P_0(d\omega)}$ jelöléssel és az $M_{P_0}\left(\frac{\partial\psi_\theta p_\theta}{\partial\theta}\right) = 0$ összefüggés alapján kapjuk, hogy

$$M_\theta \frac{\partial}{\partial\theta}\psi_\theta(\omega) = \int_\Omega \frac{\partial\psi_\theta(\omega)}{\partial\theta} p_\theta(\omega) P_0(d\omega) = -\int_\Omega \psi_\theta(\omega) \frac{\partial p_\theta(\omega)}{\partial\theta} P_0(d\omega) =$$

$$= -\int_\Omega \psi \frac{\partial L_\theta(\omega)}{\partial\theta} p_\theta(\omega) P_0(d\omega) = -M_\theta\left(\psi_\theta(\omega) \frac{\partial L_\theta(\omega)}{\partial\theta}\right) = -\sigma(\theta)\, M_\theta(\psi_\theta(\omega)\varphi_\theta(\omega)).$$

Innen a *Cauchy–Bunyakovszkij egyenlőtlenség* alapján kapjuk állításunkat. Másrészt bizonyos általános feltételek esetén a maximum likelihood egyenlet megoldásának eloszlása aszimptotikusan megegyezik a likelihood hányados deriváltjának eloszlásával (lásd pl. WALD [8]), s az általuk szolgáltatott konfidencia intervallumok hossza azonos (lásd LUVSZANCEREN [11]). Ily módon, mivel a $\psi_\theta(\omega)$ függvény „átlagban" laposabban metszi a θ tengelyt, a megfelelő $|\psi_\theta(\omega)| \leqq x$ intervallum hossza „átlagosan" hosszabb lesz, mint a $\varphi_\theta(\omega)$ függvény esetén nyert intervallum. Éppen ebben az értelemben nevezi WILKS a maximum likelihood egyenlet megoldásából adódó konfidencia halmazokat optimálisnak. Ha φ függ a megfigyelési intervallum hosszától (pl. sztochasztikus folyamatok esetén), — legyen ez $(0, T)$ — akkor

$$-\frac{1}{\sigma(\theta)} \frac{\partial\varphi_\theta^T(\omega)}{\partial\theta} = -\frac{1}{\sigma^2(\theta)} \frac{\partial^2 L_\theta^T(\omega)}{\partial\theta^2} + \frac{1}{\sigma^3(\theta)} \frac{\partial L_\theta^T(\omega)}{\partial\theta} \cdot \frac{\partial\sigma(\theta)}{\partial\theta},$$

ahol az első tag -1-hez való tartása $T \to \infty$ esetén, az információs stabilitást jelenti (lásd pl. PINSKER [3], vagy DOBRUSIN [9]), miközben a második tag nullához tart. Az elmondottakból látható a konfidencia intervallum és az információ fogalmának a kapcsolata.

IRODALOM

[1] VINCZE I.: *MTA III. Osztályának Közleményei* (1962) 1, 7–14.
[2] VINCZE I.: *Mat. Lapok* (1960) 1.
[3] М. С. Пинскер: Информация и информационная устойчивость случайных величин и ироцессов, Изд. А. Н. СССР, Москва, 1960.
[4] Ю. А. Розанов: Теория вер. и её ирим. **7** (1962) 1, 84—89.
[5] J. HAJEK: Чехослов. Мат. журнал **6** (81) (1956) 94—114.
[6] S. S. WILKS: A. M. Stat. **9** (1938) 166–175.
[7] H. CRAMER: *Mathematical Methods of Statistics*, London 1946.
[8] A. WALD: *A. M. Stat* **13** (1942) 127–137.
[9] R. L. DOBRUSIN: *MTA III. Osztályának Közleményei* **11** (1961) 427, **12** (1962) 51.
[10] В. А. Рохлин: Известия А. Н. СССР. сер. Мат. (1961) Н°4, 499—530.
[11] Ш. Лувсанцерен: Д. А. Н. **98** (1954) 723—726.

(Beérkezett: 1962. VII. 19)