

A STATISZTIKUS MECHANIKA NÉHÁNY ELOSZLÁSTÉTELÉRŐL*

Írta: VINCZE ISTVÁN

Bevezetés

BOLTZMANN, majd PLANCK munkáiban egy általános eljárás került kidolgozásra, amely megadja, hogy egy rendszer összes energiája hogyan oszlik meg a rendszert alkotó nagyszámú — azonos típusúnak és függetlennek feltételezett — részecskéje között. A módszer tisztán valószínűségszámítási, amely lényegében a maximum likelihood elv alkalmazása. JAYNES [5] 1957-ben a módszernek információelméleti megfogalmazását adta. A diszkrét eloszlás esetével foglalkozik és a folytonos esetre a kérdést alapvetően bonyolultabbnak ítélte meg [lásd i. h. 622. o. 6) lábjegyzet]. Ugyancsak JAYNES egy 1968. évben megjelent könyvismertetéséből kitűnik, hogy a kérdést ma sem tekinti a statisztikus fizika elmélete szempontjából lezártnak.

PLANCK „A normálspektrum energiaeloszlási törvényének elmélete” című előadásában maga is arra hivatkozik, hogy ha a teljes energia „minden határon túl részeire osztható, akkor végtelen sokféle eloszlás valósulhat meg”. A következő mondatában — mintegy e körülményt tekintve a nehézség okának — az energiát diszkrét értékeket felvevő mennyiségnek tekinti és bevezeti a h mennyiséget. PLANCK itt nyilván az eloszlás folytonos voltából eredő kombinatorikus nehézségekkel találta magát szemben, azonban nem állítható, hogy ez lett volna valóban és kizárólag az oka a kvantumhipotézis felállításának. Mind a kvantumelmélet sikerei, ahogyan a fizika számos problémáját megoldotta, mind pedig az a körülmény, hogy az energiaeloszlásra kapott diszkrét eloszlásfüggvényt akadály nélkül helyettesíthették volna a megfelelő folytonos eloszlással, arra utalnak, hogy több irányú megfontolás vezette a fizikusokat az elméletnek a kvantumhipotézis alapján való kifejesztésére.

Nem érdektelen azonban a felmerült valószínűségszámítási feladat vizsgálata és a fizikusok által e téren alkalmazott módszer egzaktabb matematikai megalapozása mind folytonos, mind diszkrét változó esetében. Hasonló nehézség merül fel a SHANNON által megadott entrópia-fogalomnak folytonos változóra való kiterjesztése esetén, mely kiterjesztést a szerző [11, 12] dolgozatában vitte keresztül, majd ettől függetlenül néhány évvel később JAYNES [5a] is erre az eredményre jut. A Boltzmann—Planck-féle alapproblémának egzakt megoldásmódja is hasonló ehhez. A Boltzmann—Planck- módszer a maximum likelihood módszerrel teljes analógiát mutat. Ez természetesen csupán analógia: a maximum likelihood módszer jogosságát (konzisztencia, efficiencia) bizonyos eléggé általános feltételek mellett bizonyítani lehet. A Boltzmann—Planck-eloszlás jogosságát semmiféle matematikai módszerrel, „bizonyítani” nem lehet, azt csupán a tapasztalattal való egyezés támaszthatja

* Szerző e vizsgálatairól beszámolt 1968. szeptember 3-án Amszterdamban a “European Meeting 1968 on Statistics, Econometrics and Management Science” konferencián.

alá. Mindamellett a kérdésnek érdekes valószínűségelméleti vonatkozásai vannak, amiről külön szólnunk.

Az alábbi tárgyalásban előjövő egyes matematikai megfontolások már ismertek; így JAYNES említett munkáján kívül SANOV [10], KULBACK [6a] munkájában megtalálhatók. Az alábbiakban a kialakult módszer matematikai szempontból való kifogástalan tárgyalását és olyan értelmezését adjuk, amely plauzibilis, semmi-féle információelméleti, vagy a valószínűségszámítástól, ill. fizikai megfontolásoktól idegen elv alkalmazását nem teszi szükségessé. A módszer mind folytonos, mind diszkrét eloszlású változó esetére alkalmazható. Az, hogy adott esetben diszkrét vagy folytonos modellt válasszunk-e, a jelenség természetétől függ, és a feltevés a tapasztalattal való egyezésében nyerheti igazolását. Ugyanakkor megmutatjuk, hogy ha a *Bose—Einstein*- vagy *Fermi—Dirac*-elvekre alkalmazunk folytonos modellt, akkor az ismerttől eltérő eredményre jutunk, amelynek a klasszikus eloszlások csak első közelítései. Minthogy azonban a folytonos modell a diszkrét modell (numerikus) közelítésének tekinthető, a valóságos helyzet jobb leírását adhatja. Ezek az eloszlások sajnos explicite nem adhatók meg és csupán numerikus módszerek vezethetnek eredményre.

Tárgyalásunk, amely a *Boltzmann—Planck*-féle eljárás kifejtése, még azzal az előnnyel is bír, hogy nincs szükség a változó értéktartományának egyenlő „apriori” valószínűségű részekre való felosztására; ugyanakkor természetes módon adódik a termodinamikai valószínűség logaritmusára, tehát a termodinamikai entrópiára az

$$S = -k \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)}$$

alakú kifejezés, amely mind diszkrét, mind folytonos változó esetén érvényes. Itt $F(x)$ a szóban forgó véletlen ingadozást mutató fizikai mennyiség eloszlásfüggvénye, míg $\Phi(x)$ az „apriori” eloszlás, melynek értelmezésére többféle lehetőség van; az apriori eloszlás lényegében már BOLTZMANN munkájában fellép.

Szerző ezúttal mond köszönetet RÉNYI ALFRÉD akadémikusnak és FÉNYES IMRÉNEK, a fizikai tudományok doktorának értékes megjegyzéseikért.

1. §. Az eloszlás kiválasztására vonatkozó elv

Tekintsünk egy $\mathcal{E}$ fizikai rendszert (pl. ideális gáz egy részecskéje) és legyen az X valószínűségi változó a rendszerrel kapcsolatos valamely mennyiség (pl. energia, sebesség stb.). Kiindulunk abból, hogy a rendszernek valamely állapotában az X változó eloszlását a $\Phi(x)$ eloszlásfüggvény írja le, amely kifejezi a rendszer magatartását az adott helyzetben. Kérdés, hogy a rendszert érő valamely hatás eredményeként mi lesz X eloszlása.

A következő két feltételezéssel élünk:

a) Az X változó a tartós hatásra is ugyanavval az értékkel rendelkezik, mint a rendszer kiinduló állapotában. Ezt a követelményt alább pontosabban megfogalmazzuk. (Más eset tárgyalására később kívánunk rátérni.)

b) Az X változó *alapvető magatartását* a feltételezett hatás létrejötte után is *lényegében* a $\Phi(x)$ eloszlástörvény szabja meg, bár attól a tényleges eloszlás a kényszer folytán el fog térni.

A hatást abban a legegyszerűbb formában vesszük tekintetbe, hogy az X változóra az $\int_{-\infty}^{\infty} x d\Phi(x) = m_0$ -tól különböző m első momentumot kényszerítünk.

Ha az X eloszlására tekintetbe jövő eloszlások összessége véges sok paraméterrel volna megadható, akkor a maximum likelihood módszerhez analóg eljárásnak nem volna akadálya. Így a következő eljárást követjük: a tekintetbe jövő eloszlások összességében egy pszeudosűrűséget vezetünk be; szemléletesen szólva: meghatározzuk annak valószínűségét, hogy a $\Phi(x)$ eloszlású X változóra vonatkozó igen nagy számú megfigyelés eredményéből konstruált tapasztalati eloszlásfüggvény nem a $\Phi(x)$ -hez, hanem valamely tekintetbe jövő $F(x)$ eloszlásfüggvényhez esik közel. Más szóval: annak valószínűségét határozzuk meg, hogy a $\Phi(x)$ eloszlású változóra vett nagy minta olyan képet mutat, mintha eloszlásfüggvénye $F(x)$ volna. Ennek az eseménynek a valószínűsége növekvő megfigyelésszámmal természetesen zéróhoz tart, miként annak a valószínűsége is, hogy egy folytonos eloszlású változó értéke egy előre megadott értéket vesz fel; ugyanakkor azonban létezik a folytonos valószínűségi változó *sűrűsége* az adott pontban. Esetünkben az $F(x)$ -beli *pszeudosűrűségről* beszélünk, amely a Boltzmann—Planck-eljárás következetes keresztülvitele esetén a következő kifejezés:

$$e^{-\int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)}}.$$

A kitevőben negatív előjellel az információelméletben használt ún. relatív információ vagy I -divergencia áll. E kifejezés veszi át szerző [10] és JAYNES [5a] munkájában a folytonos változó esetében a Shannon-féle entrópiakifejezés szerepét, illetve abból határártmenettel nyerhető.

Ily módon a szóba jöhető eloszlások összességében pszeudosűrűséget vezetünk be és a maximum likelihood módszer analogonját alkalmazhatjuk. Előbb azonban a mondottakat pontosabban megfogalmazzuk.

2. § Az I -divergencia mint pszeudo-likelihood függvény.

Sanov formulája

1. Legyen X valószínűségi változó az $(\Omega, \mathcal{A}, \mathcal{P})$ valószínűségi mezőn $\Phi(x)$ eloszlásfüggvénnyel:

$$P(X < x) = \Phi(x), \quad -\infty < x < \infty.$$

Jelölje μ_F azt a mértéket, amelyet az $F(x)$ eloszlásfüggvény a $(-\infty, \infty)$ számegyenes Borel-halmazain értelmez. Tekintsük az eloszlásfüggvények ama $\mathcal{F}_\Phi$ halmazát, amelynek $F(x)$ elemeire a μ_F mérték ekvivalens a μ_Φ -vel, azaz egymásra nézve kölcsönösen folytonosak:

$$\mathcal{F}_\Phi = \{F(x) : \mu_\Phi \ll \mu_F \text{ és } \mu_F \ll \mu_\Phi\}.$$

Tekintsük továbbá a valós egyenes egy felosztásrendszerét, azaz minden $n > 1$ egész számra a

$$\{-\infty = x_0^{(n)} < x_1^{(n)} < \dots < x_{n-1}^{(n)} < x_n^{(n)} = +\infty\}$$

osztópontrendszert, amely bírjon a következő tulajdonsággal: minden véges (a, b) $(a < b)$ intervallumban a $\max_{(i)} (x_i^{(n)} - x_{i-1}^{(n)})$, $a \leq x_{i-1}^{(n)} < x_i^{(n)} \leq b$ tartson zéróhoz, ha $n \rightarrow \infty$.

Jelöljék a nagyszámú részecskékből álló rendszerben az egyes részekre nézve az X mennyiséget az $X_1, X_2, \dots, X_N, \dots$ független valószínűségi változók (amelyek száma tehát megszámlálhatóan végtelen). Legyen e valószínűségi változók egy realizációja az $(x_1, x_2, \dots)$ végtelen dimenziós vektor. Jelentse $N_{iN}^{(n)}$ az $\{x_{i-1}^{(n)} \leq X < x_i^{(n)}\}$ esemény gyakoriságát a megfigyelt $(x_1, x_2, \dots, x_N)$ mintában; vagyis az illető intervallumba eső elemek számát az első N megfigyelésből.

Jelölje végül $F_N(x)$ az $(x_1, x_2, \dots, x_N)$ mintához tartozó empirikus eloszlásfüggvényt:

$$F_N(x) = \frac{1}{N} \sum_{x_i \leq x} 1, \quad F_N(-\infty) = 0, \quad F_N(\infty) = 1.$$

Tekintsük most az $\mathcal{F}_\Phi$ eloszlásösszesség valamely $F(x)$ elemét és olyan $(x_1, x_2, \dots)$ realizációját az $(X_1, X_2, \dots)$ valószínűségi vektor változónak, amelyre az egész $(-\infty, \infty)$ számegegyenesen $\lim_{N \rightarrow \infty} F_N(x) = F(x)$ egyenletesen.

E realizációkra nézve az is teljesülni fog, hogy minden rögzített n -re

$$(*) \quad \lim_{N \rightarrow \infty} \frac{N_{iN}^{(n)}}{N} = F(x_i^{(n)}) - F(x_{i-1}^{(n)}), \quad i = 1, 2, \dots, n.$$

A továbbiakban rögzített n mellett fogjuk az N megfigyelésszámot végtelenbe tartatni, majd a felosztások n számát növeljük minden határon túl.

2. Az $(X_1, X_2, \dots)$ végtelen dimenziós vektorváltozó elemei tehát függetlenek, $\Phi(x)$ eloszlású valószínűségi változók. Vegyük az első N számú $(X_1, X_2, \dots, X_N)$ megfigyelést, és jelölje a $v_{iN}^{(n)}$ valószínűségi változó az $(x_{i-1}^{(n)}, x_i^{(n)})$ intervallumba eső elemek számát: $\sum_{i=1}^N v_{iN}^{(n)} = N$. Ha a $q_{in} = \Phi(x_i^{(n)}) - \Phi(x_{i-1}^{(n)})$ jelölést használjuk, akkor a

$$v_{iN}^{(n)} = N_{iN}^{(n)}, \quad i = 1, 2, \dots, n$$

események együttes bekövetkezésének a valószínűsége

$$P_{n,N} = \frac{N!}{N_{1N}^{(n)}! N_{2N}^{(n)}! \dots N_{nN}^{(n)}!} q_{1n}^{N_{1N}^{(n)}} \dots q_{2n}^{N_{nN}^{(n)}}.$$

Nagy N -re alkalmazzuk a *Stirling*-formula $k! \sim \left(\frac{k}{e}\right)^k$ változatát, ami a következő eredményre vezet:

$$P_{nN} \sim \frac{\prod_{i=1}^n q_{iN}^{N_{iN}^{(n)}}}{\prod_{i=1}^n \left(\frac{N_{iN}^{(n)}}{N}\right)^{N_{iN}^{(n)}}}.$$

Válasszuk most az $\{N_i^{(n)}\}$ rendszert úgy, hogy az a $(*)$ feltételt kielégítse. Ha $F(x) \equiv \Phi(x)$, akkor megmutatható, hogy

$$\lim_{N \rightarrow \infty} P_{nN} = 1.$$

Ha azonban $F(x)$ eltér a $\Phi(x)$ -től, akkor a $P_{n,N}$ valószínűség N -nel zéró felé tart, ugyanakkor N -edik gyökére a következő relációt nyerjük:

$$\lim_{N \rightarrow \infty} \sqrt[n]{P_{n,N}} = e^{-\sum_{i=1}^n p_{in} \log \frac{p_{in}}{q_{in}}},$$

ahol a

$$p_{in} = F(x_i^{(n)}) - F(x_{i-1}^{(n)}), \quad i = 1, 2, \dots, n$$

jelölést alkalmaztuk. Ha most a felosztások számát is végtelen felé tartatjuk, akkor a következő formulához jutunk:

$$\lim_{n \rightarrow \infty} \lim_{N \rightarrow \infty} \sqrt[n]{P_{n,N}} = e^{-\int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)}}.$$

A fenti aszimptotikus relációk pontosabb alakjai SANOV [10] munkájában találhatók.

A $\Phi(x)$ eloszlásról nem tettük fel sem hogy folytonos, sem hogy diszkrét. Folytonos esetben az F és Φ sűrűségfüggvényét f -fel, ill. φ -vel jelölve a kitevő alakja

$$-\int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{\varphi(x)} dx,$$

míg diszkrét esetben

$$-\sum_{(i)} p_i \log \frac{p_i}{q_i},$$

ahol p_i és q_i jelöli az $F(x)$, ill. Φ megfelelő ugráspontjaihoz tartozó valószínűségeket.

3. A kapott pszeudosűrűség logaritmus — amit pseudo-likelihood függvénynek nevezhetünk — negatív előjellel adja az I -divergenciát:

$$I_{\Phi}(F) = \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)}.$$

Ennek a következő nevezetes tulajdonságai jól ismertek:

α) Nemnegatív; akkor és csakis akkor zéró, ha $F(x) \equiv \Phi(x)$.

β) Az x tengely monoton transzformációival szemben invariáns.

Szemben az entrópiával, ami a bizonytalanság mértéke, az I -divergencia az információ mértékének tekinthető. Ez megfelel annak, hogy erre a kifejezésre nem a Shannon-féle entrópia formulájából jutunk, hanem annak kiegészítő kifejezéséből, a

$$\log m - \sum_{i=1}^m p_i \log \frac{1}{p_i} = \sum_{i=1}^n p_i \log \frac{p_i}{1/m}$$

kifejezésből. Ez az érték akkor 0, ha egyenletes az eloszlás, míg a $\log m$ értéket veszi fel, ha egyetlen eseménynek 1 a valószínűsége, a többié zéró, tehát amikor teljes információval rendelkezünk egy jövőbeni kísérlet kimenetelét illetően.

4. Ha tehát a $\Phi(x)$ eloszlásfüggvénnyel bíró X valószínűségi változóval kapcsolatban olyan feladatot tűzünk ki, hogy adott kényszerfeltételek mellett milyen eloszlást fog követni, akkor közelfekvő az $\mathcal{F}_\Phi$ összességéből annak az eloszlásfüggvénynek kiválasztása, amely — az adott mellékfeltételek mellett — a legnagyobb (pseudo-) valószínűségű. Ez a Boltzmann—Planck-eljárás matematikai szempontból következetes keresztülvitele és — heurisztikusan — megfelel az előző pontban tett a) és b) követelményeknek.

3. §. A Boltzmann—Planck-féle eloszlástörvényről

1. Mint az 1. §-ban említettük, a rendszerre ható kényszert (hatást) avval adjuk meg, hogy a valószínűségi változóra egy az eredetitől különböző első momentumot írunk elő. Ily módon azt az $F(x) \in \mathcal{F}_\Phi$ eloszlásfüggvényt keressük, amelyre a pseudo-likelihood függvény maximum:

$$- \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)} = \text{maximum}$$

a következő mellékfeltételek mellett

$$\int_{-\infty}^{\infty} dF(x) = 1$$

$$\int_{-\infty}^{\infty} x dF(x) = m,$$

ahol $m \neq \int_{-\infty}^{\infty} x d\Phi(x)$.

E feladatot a diszkrét eset analógiájára — a Lagrange-multiplikátor módszerrel szokás megoldani, amelynek eredményeként a következő adódik:

$$dF(x) = c(\lambda, \mu) e^{-\lambda - \mu x} d\Phi(x),$$

ahol

$$c(\lambda, \mu) = \frac{1}{\int e^{-\lambda - \mu x} d\Phi(x)};$$

itt μ értékét a második feltétel határozza meg; λ értéke $F(x)$ kifejezéséből eliminálható, ill. értéke fizikai megfontolásoknak megfelelően választható meg.

A szélsőérték feladat alább adott megoldása RÉNYI Alfréd-től származik.

Legyen a minimalizálandó kifejezés

$$A(F) = \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)},$$

legyen továbbá

$$dF_c(x) = \frac{e^{cx} d\Phi(x)}{\Psi(c)},$$

ahol

$$\Psi(c) = \int_{-\infty}^{\infty} e^{cx} d\Phi(x).$$

Határozzuk meg a c állandó értékét az

$$\int_{-\infty}^{\infty} x dF_c(x) = m$$

feltételből:

$$m = \int_{-\infty}^{\infty} x dF_c(x) = \frac{1}{\Psi(c)} \int_{-\infty}^{\infty} x e^{cx} d\Phi(x) = \frac{\Psi'(c)}{\Psi(c)}.$$

Vagyis a $\frac{\Psi'(c)}{\Psi(c)} = m$ egyenletet megoldjuk c -re; a továbbiakban jelentse c ezt az értéket. Ekkor

$$\begin{aligned} A(F) &= \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{\Psi(c) e^{-cx} d\Phi(x)} = \\ &= \log \frac{1}{\Psi(c)} - c \int_{-\infty}^{\infty} x dF(x) + \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{dF_c(x)}. \end{aligned}$$

Tehát minden $F(x) \in \mathcal{F}_\Phi$ -re

$$A(F) = \log \frac{1}{\Psi(c)} - cm + \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{dF_c(x)} \geq \log \frac{1}{\Psi(c)} - mc$$

és egyenlőség akkor és csakis akkor áll fenn, ha $F(x) = F_c(x)$, ami $c = -\mu$ jelöléssel a fentebb adott eredménnyel egyezik.

2. Az $F(x)$ -re kapott megoldásfüggvény ismeretéhez $\Phi(x)$ ismerete szükséges. Ezt legtöbbször nem ismerjük, hanem feltevással élünk erre vonatkozólag. A feltevést — az elméletet — a tapasztalat igazolhatja, vagy nem támasztja alá, amikor is újabb feltevessel kísérletezünk. Ha $F(x)$ -re vonatkozóan tapasztalati adataink vannak, akkor a

$$d\Phi = \frac{1}{c} e^{\mu x} dF(x)$$

reláció tájékozódást nyújt $\Phi(x)$ -re nézve.

PLANCK feltevése az energiaeloszlást illetően a $d\Phi(x_i) = 1$, $i = 1, 2, \dots$ volt, ami nem valószínűségeloszlás, sőt nem is normálható mérték egy diszkrét állapothalmazon. Ez önmagában nem volna baj; a tárgyalt módszer kiterjeszthető olyan

monoton nem csökkenő $\Phi(x)$ függvényekre, amelyekre $\Phi(0)=0$, $\Phi(\infty)=\infty$ és $\int e^{-\mu x} d\Phi(x) < \infty$. Kérdés azonban, szükséges-e ez a végtelen halmaz egyedein vett egyenletes eloszlás? Valamely kezdeti $\Phi(x)$ valószínűségeloszlással nem közelíthető-e jobban az $F(x)$ eloszlás tapasztalati alakja.

3. Ami a nyert eredmény „matematikai bizonyítását” illeti, a következőt jegyezzük meg. A felállított variációs feladatot a nyert eredmény kielégíti, sőt lényegében ez az egyetlen megoldása a feltételes szélsőérték-feladatnak. Azonban a következő rokon probléma vethető fel, amelyre nézve kézenfekvő az általunk felállított szélsőérték-feladathoz fordulni, legalábbis a feladat formális megoldása céljából:

Legyenek $X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók. Kérdés: mi X_1 eloszlása ama feltétel mellett, hogy az első N tag számtani közepe adott:

$$P(X_1 < x | X_1 + X_2 + \dots + X_N = Nm) = ?,$$

ahol $m \neq \int x d\Phi(x)$. Ha $N \rightarrow \infty$, akkor várható, hogy X_1 feltételes eloszlása a maximum likelihood elv szerinti lesz az általunk adott pseudo-értelemben. Ezt az állítást az exponenciális eloszláscsalád némely összességére (normális, exponenciális) nem nehéz bebizonyítani. A $\Phi(x)$ -re vonatkozó bizonyos feltételek mellett BARTFAI PÁL adott igenlő választ, amely kérdésre vissza kíván térni. Lásd még HINCSEN [4] munkájában 19–31. old.

4. Érdekes végül megjegyezni, hogy az

$$\int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{d\Phi(x)} = \text{minimum}$$

$$\int_{-\infty}^{\infty} T_i(x) dF(x) = m_i, \quad i = 1, 2, \dots, k$$

feltételes szélsőérték-probléma az exponenciális eloszláscsalád általános alakjához vezet:

$$dF(x) = c(\mu_1, \mu_2, \dots, \mu_k) e^{\sum_{i=1}^k \mu_i T_i(x)} d\Phi(x).$$

Ez a tény hasonló módon lehet magyarázata annak, hogy ez az eloszláscsalád alkalmazásokban gyakran fellép, mint ahogyan a normális eloszlásnak mint határeloszlásnak fellépése motiválható.

Ugyanitt érdekes megjegyezni, hogy LINNIK az I -divergencia tulajdonságai segítségével bizonyítja a normális eloszlásnak mint határeloszlásnak fellépését.

4. §. Megjegyzés a termodinamikai entrópia fogalmához

A termodinamikában szokásos a termodinamikai valószínűség bevezetése, amely megadja, hogy egy makroállapot hány mikroállapottal valószínűsíthető meg. Ehhez tudnunk kell az ún. apriori eloszlást, mert az ún. betöltési számok egyenlő apriori valószínűségű cellákhoz tartoznak. Fenti megfontolásaink azt mutatják,

hogya ha nem csupán a lehetséges állapotok számát, hanem a teljes valószínűséget, azaz az adott esetben az

$$\frac{N!}{N_1! N_2! \dots N_n!} q_1^{N_1} q_2^{N_2} \dots q_n^{N_n}$$

kifejezést tekintjük, ahol most az egyszerűsített $N_i = N_{iN}^{(n)}$ jelölést alkalmaztuk, nincs szükség arra, hogy egyenlő valószínűségű cellákra osszunk és a matematikai tárgyalás teljesebb. Fenti kifejezésből N -edik gyököt vonva és limesre térve diszkrét esetben az

$$e^{-\sum p_i \log \frac{p_i}{q_i}},$$

folytonos esetben az

$$e^{-\int f(x) \log \frac{f(x)}{\varphi(x)} dx}$$

alakú kifejezéshez jutunk.

Ily módon a termodinamikában használatos entrópia kifejezésnél egzaktabbnak, és így alkalmasabbnak mutatkozik a

$$H = -K \int dF(x) \log \frac{dF(x)}{d\Phi(x)}$$

formula. Ez ugyancsak bír az additivitás tulajdonságával. Legyen ugyanis két független rendszer a

$$H_1 = -K \int dF_1(x) \log \frac{dF_1(x)}{d\Phi_1(x)},$$

ill.

$$H_2 = -K \int dF_2(x) \log \frac{dF_2(x)}{d\Phi_2(x)}$$

entrópiával. Egyszerűen belátható, hogy az egyesített rendszer entrópiájára a

$$H = H_1 + H_2 = -K \iint dF_1(x) dF_2(y) \log \frac{dF_1(x) dF_2(y)}{d\Phi_1(x) d\Phi_2(y)}$$

érvényes.

Ugyancsak tekinthető az entrópia két nem független valószínűségi változóra. Legyen ezeknek együttes eloszlásfüggvénye $F(x, y)$, az apriori eloszlás $\Phi(x, y)$. Ekkor

$$H = -k \iint dF(x, y) \log \frac{dF(x, y)}{d\Phi(x, y)}.$$

Független esetben ez ismét összeadandókra bomlik.

Tekintetbe kell azonban venni, hogy a valószínűségek kiszámítása különböző elvek alapján történhet. Így az entrópia definícióját a következő módon adhatjuk meg:

Jelölje $P_N(\Phi(x); F(x)|\mathfrak{R})$ annak valószínűségét, hogy a rendszerben, amelynek valamely X fizikai mennyisége a rendszer kiinduló állapotában a $\Phi(x)$ eloszlás-

sal bír, az N nagyszámú részecske mégis az $F(x)$ eloszlás körül helyezkedik el; itt $\mathfrak{R}$ jelzi azt az elvet, ahogyan a valószínűség kiszámítandó. Ekkor az entrópia

$$H = -k \log \lim_{N \rightarrow \infty} \sqrt[N]{P_N(\Phi(x); F(x)\mathfrak{R})}.$$

Fentiekben az $\mathfrak{R}$ a polinomiális valószínűség alkalmazását jelentette. Most rátérünk a termodinamikai entrópia kiszámítására, ha a *Bose—Einstein*- ill. *Fermi—Dirac*-féle elveket alkalmazzuk.

5. A termodinamikai valószínűség a Bose—Einstein-eloszlás folytonos változata esetén

A számítás egyszerűbbé tétele érdekében az energiaskálát *egyenlő* $\frac{1}{n}$ (apriori) valószínűségű részekre osztjuk és mindegyik részben z számú cellát jelölünk ki. A cellanagyság tehát $h = \frac{1}{nz}$, amit zéró felé fogunk tartatni. Nagyszámú részecske és kis cellanagyság mellett ez az eredeti modell jó approximációjának tekinthető. $\Phi(x)$ -szel jelölve ismét az apriori eloszlásfüggvényt az i -edik felosztásban a szokásos

$$z_i = \frac{1}{nh} = \frac{\Phi(x_i^{(n)}) - \Phi(x_i^{(n-1)})}{h} = \frac{d\Phi(x_i^{(n)})}{h}$$

összefüggést kapjuk.

A *Bose—Einstein*elv szerint egy cellába akárhány részecske juthat, azonban a részecskék meg nem különböztethetők. Ily módon, ha a részecskék száma N , akkor az $A_{n,N} = \left\{ N_1, N_2, \dots, N_n; \sum_{i=1}^n N_i = N \right\}$ makroállapothoz tartozó valószínűség

$$P(A_{n,N}) = \frac{\prod_{i=1}^n \binom{N_i + z}{N_i}}{\binom{N + (z+1)n}{N}}.$$

További egyszerűsítés végett válasszuk a részecskék számát a következőképpen: $N = zn$. Ez nem jelent megszorítást, mert egy arányossági tényezőt véve, az a később választandó állandókba úgyis beolvasztható. Alkalmazva a *Stirling*-formulát az $N! \sim \left(\frac{N}{e}\right)^N$ alakban, a következőre jutunk:

$$P(A_{n,N}) \sim \frac{N^N [n(z+1)]^{n(z+1)}}{[N + (z+1)n]^{N + (z+1)n}} \prod_{i=1}^n \frac{\left(\frac{N_i}{N} + \frac{1}{n}\right)^{N_i + n}}{\left(\frac{N_i}{N}\right)^{N_i} \left(\frac{1}{n}\right)^z}.$$

A szorzatjel előtti rész a következőképp írható — mindjárt az N -edik gyökét véve

$$\frac{\left(1 + \frac{n}{N}\right)^{1 + \frac{n}{N}}}{\left(2 + \frac{n}{N}\right)^{2 + \frac{n}{N}}} \sim \frac{1}{4}.$$

Ez az állandó a szélsőérték-feladat szempontjából nem játszik szerepet; így el fogjuk majd hagyni. A Π jel utáni kifejezésben arra ügyelünk, hogy az $F(x)$ és $\Phi(x)$ eloszlásfüggvények egymásra nézve való abszolút folytonosságát biztosítsuk. Ezért a $\Psi(x) = \frac{dF(x)}{d\Phi(x)}$ bevezetésére törekszünk. Ezt figyelembe véve:

$$P(A_{u,n}) \sim \frac{1}{4^N} \prod_1^n \left(1 + \frac{1}{\frac{n}{N_i}}\right)^{N_i + n} \left(\frac{N_i}{\frac{1}{n}}\right)^z.$$

Mint hogy $\frac{N_i}{N} \sim dF(x_i)$ és $\frac{1}{n} = d\Phi(x_i)$, $i = 1, 2, \dots, n$, a következőt kapjuk

$$\frac{1}{N} \log P(A_{n,n}) \sim \log \frac{1}{4} + \sum_{i=1}^n \left[(dF(x_i) + d\Phi(x_i)) \log \left(1 + \frac{d\Phi(x_i)}{dF(x_i)}\right) + d\Phi(x_i) \log \frac{dF(x_i)}{d\Phi(x_i)} \right].$$

Az $N \rightarrow \infty$, majd az $n \rightarrow \infty$ határátmenetet elvégezve eredményül a termodinamikai valószínűsége, illetve a termodinamikai entrópiára a következő kifejezést kapjuk — az $\frac{1}{4}$ additív tag elhagyásával:

$$H = -k \int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{(1 + \Psi(x))^{1 + \frac{1}{\Psi(x)}} d\Phi(x)}.$$

Ha ezt maximalizálni akarjuk, az

$$\int_{-\infty}^{\infty} dF(x) = 1$$

$$\int_{-\infty}^{\infty} x dF(x) = m$$

feltételek mellett a következő szélsőérték-feladatra jutunk:

$$\int_{-\infty}^{\infty} dF(x) \log \frac{dF(x)}{c(\lambda, \mu) e^{-\lambda \mu x} (1 + \Psi(x))^{1 + \frac{1}{\Psi(x)}} d\Phi(x)} + \log c(\lambda, \mu) = \text{minimum},$$

ahol

$$c(\lambda, \mu) = \frac{1}{\int_{-\infty}^{\infty} e^{-\lambda\mu x} (1 + \Psi(x))^{1 + \frac{1}{\Psi(x)}} d\Phi(x)}.$$

A második tag most függ az ismeretlen $F(x)$ -től ami sokkal bonyolultabbá teszi a problémát. Az első tag minimális — azaz zéró — értékét a

$$dF(x) = \frac{e^{-\lambda-\mu x} (1 + \Psi(x))^{1 + \frac{1}{\Psi(x)}} d\Phi(x)}{\int_{-\infty}^{\infty} e^{-\lambda-\mu x} (1 + \Psi(x))^{1 + \frac{1}{\Psi(x)}} d\Phi(x)}$$

reláció mellett kapjuk. Ha a kitevőben levő $\frac{1}{\Psi(x)}$ tagot elhanyagoljuk — ami nem tehető minden további nélkül —, akkor a

$$dF(x) = ce^{-\lambda-\mu x} \left(1 + \frac{dF(x)}{d\Phi(x)} \right) d\Phi(x),$$

vagyis a

$$dF(x) = \frac{cd\Phi(x)}{e^{\lambda+\mu x} - 1},$$

klasszikussal analóg eredményre jutunk.

Hasonló módon határozható meg a *Fermi—Dirac*-esetben a termodinamikai entrópia. Erre a következő kifejezés adódik — egy additív állandótól eltekintve:

$$H = -k \int dF \log \frac{dF}{\left(1 - \frac{1}{\alpha} \Psi(x) \right)^{\frac{\alpha}{\Psi(x)} - 1} d\Phi(x)}, \quad \alpha > 1.$$

Az előzőhöz hasonló módon képzett közelítésként a következő eloszlás adódik:

$$dF(x) = \frac{cd\Phi(x)}{e^{\lambda+\mu x} + 1}.$$

6. További megjegyzések az I -divergenciára vonatkozóan

Az I -divergencia, mint említettük, széles körű alkalmazást talált eddig is egészen különböző irányú vizsgálatokban. Ez az alkalmazhatóság legtöbb esetben a 2. §-ban adott és elsőként SANOV által nyert formulában, ill. annak értelmezésében leli magyarázatát.

KULLBACK és LEIBLER [6], majd KULLBACK [6a] terjedelmes munkájában matematikai statisztikai alkalmazások egész sorát adja és hozza összefüggésbe az I -divergenciát különböző ismert módszerekben szereplő statisztikákkal. Szerző [12] konfidenciaintervallumok konstrukciójára alkalmazta, ami már KULLBACKnál is szerepel.

PEREZ [7] információelméleti alkalmazásait adta. CSISZÁR [1] általánosította az I -divergencia fogalmát és az általánosított fogalom számos tulajdonságát és alkalmazását adta meg.

RÉNYI [8, 9] munkáiban a matematikai statisztika információelméleti megalapozását adja. Dolgozataiban a statisztikai eljárás hibáját az entrópia mennyiséggel becsülte. A parameter itt diszkrét; szerző [13] RÉNYI bizonyos eredményeit folytonos paraméterterre terjesztette ki az I -divergenciának adva a diszkrét entrópia szerepét.

HÁJEK [3] Gauss-folyamatok vizsgálatára használta az I -divergenciát, annak bizonyos alapvető tulajdonságait is tisztázta.

FRITZ [2] információelméleti módszerrel határozta meg gravitációs térben a gáz nyomáseloszlását.

IRODALOM

- [1] CSISZÁR, I.: On topological properties of f -divergences, *Studia Sci. Mat. Hung.* **2** (1967), 329—339.
- [2] FRITZ, J.: Information Theory and Thermodynamics of Gas Systems, *Proc. Coll. on Information Theory*, Debrecen (Hungary) 1967
- [3] HÁJEK, J.: A property of I -divergence of marginal probability distributions, *Czechoslovak Journal of Mathematics* **8** (83) 1958. pp. 460—463.
- [4] HINCHIN, A. J.: *A statisztikai mechanika analitikus módszerei*, Akadémiai Kiadó. 1951.
- [5] JAYNES, E. T.: Information Theory and Statistical Mechanics, *Physical Reviews* **106** (1957) pp. 620—630.
- [5a] JAYNES, E. T.: see K. W. Ford, Ed.; *Statistical Physics*, W. A. Benjamin Inc. 1963. 4. fejt.
- [6] KULLBACK, S.—LEIBLER, R. A.: On information and sufficiency. *Annals of Mathematical Statistics*, **22** (1951) pp. 79—86.
- [6a] KULLBACK, S.: *Information and statistics*, John Wiley & Sons, Inc. New York, 1959.
- [7] PÉREZ, A.: Notions généralisées d'incertitude, d'entropie et d'information du point de vue de la théorie de martingales, *Transactions of the first Prague Conference on information theory, statistical decision functions, random processes*. Prague 1957 pp. 183—208.
- [8] RÉNYI, A.: Statistics and information theory, *Studia Scientiarum Mathematicarum Acad. Sci. Hung.* **2** (1967) 249—256.
- [9] RÉNYI, A.: On the dimension and entropy of probability distributions, *Acta Mathematica Acad. Sci. Hung.* **10** (1959) 193—215.
- [10] SANOV, I. N.: On the probability of large deviations of random variables. *IMS and AMS Selected Translations in Mathematical Statistics and Probability*, Vol. 1 (1961) 213—244.
- [11] VINCZE, I.: An interpretation of the I -divergence of information theory, *Transactions of the second Prague Conference on information theory, statistical decision functions, random processes*. Prague, 1959. pp. 681—684.
- [12] VINCZE, I.: Some questions concerning the probabilistic concept of information. *AMS and IMS Selected Translations in Mathematical Statistics and Probability*, Vol. V. (1965).
- [13] VINCZE, I.: On the information-theoretical foundation of mathematical statistics. *Proc. of the Coll. on Information Theory, Debrecen (Hungary)* 1967. 503—509. old.

(Beérkezett: 1968. XII. 22.)

ON SOME DISTRIBUTION LAWS OF STATISTICAL PHYSICS

by

ISTVÁN VINCZE

Summary

The aim of the paper is to establish an exact and unified formulation and treatment of the method given by BOLTZMANN and PLANCK for the derivation of the distribution laws of statistical physics valid for continuous as well as discrete random quantities. Indeed — as mentioned by PLANCK — in obtaining the distribution of the total energy of a gas-system among the large number of their particles, certain difficulties arise if the energy may take all possible values; these combinatorial difficulties are similar to those which occur when *Shannon's* entropy is to be extended to the continuous case. The latter problem was treated by JAYNES [5a] and earlier by the author [11,13]; the present consideration has a similar feature to this extension. — A pseudo-likelihood function or pseudo-density function is introduced in the set of the distribution functions which can be taken into consideration; this procedure has an obvious interpretation. With the aid of this pseudo-likelihood function a *principle* — analogous to the maximum likelihood *method* — can be introduced for the selection of the appropriate distribution function. What is discussed here is that this is only an analogy: the two procedures have different nature. — As a natural consequence we may obtain for the (thermodynamical) entropy the following expression:

$$S = -k \int dF(x) \log \frac{dF(x)}{d\Phi(x)};$$

$F(x)$ denotes here the distribution function of the considered random variable (discrete or continuous) and $\Phi(x)$ is the prior distribution, which may have different interpretations. The prior distribution occurs already in the work of BOLTZMANN. — PLANCK uses a prior distribution which leads to a distribution uniform on a countable infinite set of energy-values; the question whether this assumption is necessary or not is discussed. — Finally the entropy and the continuous version in the case of the *Bose—Einstein* and *Fermi—Dirac* statistics is considered; to obtain the exact distribution laws for these is considerably difficult. The classical forms can be considered as very first approximations. — In the different parts of the paper some known mathematical results due to KULLBACK [6a], SANOV [10] are used, one proof was given by A. RÉNYI.