

A VALÓSZÍNŰSÉGELMÉLET BIZONYOS FEJEZETEIRŐL, MELYEK KÖZVETLEN KAPCSOLATBAN VANNAK A BIOLÓGIA ÉS AZ ORVOSTUDOMÁNY PROBLÉMÁIVAL*

Írta: B. V. GNYEGYENKO

1. Bevezető megjegyzések. A Nagy Szovjet Enciklopédiában (2. kiadás) a „Biológia” címszó után a szöveg a következő szavakkal kezdődik: „A biológia az életről szóló tan; az a tudomány, amely feltárja az élőlények életének és fejlődésének törvényszerűségeit (LISZENKO, T. D.: *Agrobiológia*. 5. kiadás, 1949. 609. o.)” Az élő természet törvényeinek megismerése az egyik legnehezebb és ugyanakkor legfontosabb probléma; ez a megismerés nemcsak mezőgazdasági problémák (például a mezőgazdasági termékek bőségének biztosítása) megoldásához szükséges, hanem felbecsülhetetlen szolgálatot tesz az orvostudománynak is a legkülönbözőbb betegségekkel folytatott harcában. Meg vagyok győződve arról is, hogy egyedi élő sejtekben, illetve egy egész, magas szervezettségű szervezetben lejátszódó folyamatok megértése óriási távlatokat nyit meg a jövő technikája előtt.

A biológusok érdeklődési köre a legnagyobb hozamú növényfajták és a legtermékenyebb állatfajták kitenyésztésének problémájától (amely már sok évtizeddel ezelőtt felmerült az emberiség előtt) a gondolkodás mechanizmusába való behatolás problémájáig terjed. Természetes, hogy ilyen sokféle kérdéssel foglalkozván, a biológia nem korlátozódhat vizsgálati módszerek valamilyen zárt csoportjára. A fizika, kémia, rádiótechnika behatolása a biológiába elkerülhetetlenül arra vezet, hogy idővel a matematika is állandóan bele fog tartozni a biológus kutatási eszközei közé. És ebben nincs semmi csodálatos, mert a biológiában már beszélhetünk „az anyag olyan homogén és egyszerű elemeinek megközelítéséről, amelyek mozgási törvényei megengedik a matematikai tárgyalásmódot” (V. I. LENIN: *Összegyűjtött művei*, 14. k. 294. o. (Oroszul)).

A matematika biológiai vizsgálatokba való bevonása nem mindig megy simán, nem mindig sikerül annyira megbízható eredményeket kapni, hogy azok nem szorulnak további élesítésre és átgondolásra. De vannak-e általában a tudományban olyan eredmények, amelyeket nem kell később pontosabban megfogalmazni, korrigálni és módosítani? Ismeretes, hogy minden új felfedezés csak egy állomás, csak egy lépcsőfok abban a folyamatban, ahogy a természetnek és jelenségeinek törvényszerűségeit megismerjük. Mármost mi hasznunk lehet ilyenformán matematikailag megfogalmazott elméleteknek? Mindenekelőtt az, hogy segítségükkel lehetőségessé válik bizonyos alapvető biológiai alapelvekből logikai úton egész sor következményt levezetni. Ezeket a következményeket kísérletekkel ellenőrizve, kiegészíthetjük a kiindulási alapelvek helyességét alátámasztó adatainkat; ez az ellenőrzés magában véve is érdekes a biológia szempontjából. De még ha ténylegesen a kísér-

* „Применение математических методов в биологии” (Matematikai módszerek alkalmazása a biológiában). Издательство Ленинградского Университета, 1960; 6–16.

letek által nyújtottakkal ellentétben álló eredményekre vezet is egy matematikailag megfogalmazott biológiai elmélet, érdekessége kétségtelenül megvan, mert arra vezet, hogy elkerülhetetlenül revideálni kell ama elsődleges biológiai megállapításokat, amelyeken az elmélet felépült. Az ilyen elméletek már ezzel is elősegítik ismereteink előrehaladását.

Nem ritkán halljuk biológusok között azt, hogy a matematikai analízis egyáltalán nem alkalmazható biológiai jelenségekre, azok különleges bonyolultsága és változékonysága miatt. Az ilyen okoskodások azonban nyilvánvaló félreértésen alapulnak, mert létezik egész sor rendkívül bonyolult fizikai jelenség is, és ez a körülmény mégsem zárja ki matematikai apparátus alkalmazását tanulmányozásukkor. Az pedig, hogy látszólag egyforma külső feltételek mellett egyes jelenségek lefolyása változékonyságot mutat, csak egy valamire utal: arra, hogy a legnagyobb mértékben segítségül kell hívni a valószínűségelméletet és a matematikai statisztikát.

Meg vagyok győződve arról, hogy bizonyos kölcsönös meg nem értés, amely biológusok és matematikusok között fennáll, jelentős mértékben abból ered, hogy elszigetelten dolgozunk. Bár napjainkban matematikusok és fizikusok ezrei választották életfeladatuknak fizikai jelenségek matematikai eszközökkel való tanulmányozását, mégsem ismerek egyetlen szovjet matematikust sem, aki teljesen vagy akár csak jelentős mértékben is biológiai jelenségek matematikai eszközökkel való tanulmányozásának szentelte volna idejét és energiáját. Nem arra gondolok most, hogy már szükségessé vált megalkotni a „matematikai biológia” sajátos tanát, — hasonlóan a „matematikai fizikához”. — Semmi kétségem azonban afelől, hogy elkövetkezett az az idő, amikor matematikusok és biológusok kollektíváinak el kell kezdeniök a rendszeres együttes munkát, hogy megoldják a sarkalatos biológiai problémákat; ez olyan munka lesz, amelynek során a matematikusnak bele kell mélyednie a biológiai jelenségek lényegébe, a biológusnak pedig a matematikai módszerek alapfogalmainak, s nem csak számolásokban való használhatósága által nyújtott lehetőségekbe. Ilyen hozzáállás mellett fokozatosan kikovácsolódnak majd a kutatás azon matematikai eszközei, amelyek legalkalmasabbak az előttünk egyáltalán ismeretes legbonyolultabb jelenségeknek — az élő természet jelenségeinek — tanulmányozására.

A tudomány története ismer olyan eseteket, amelyekben a matematikai elméletek haladása közvetlen kapcsolatban volt azzal, hogy feleletet adhatnak a biológusok kérdéseire —, azzal, hogy kísérletet tehetek a kvantitatív tényező tanulmányozására oly biológiai jelenségekben, amelyekben az döntő szerepet játszik. Így folyt le; nagymértékben a biológia, orvostudomány és a mezőgazdasági kísérletezés közvetlen és állandó ráhatása közben, a matematikai statisztika alapjainak kialakítása. A sztochasztikus folyamatok elmélete kialakulásának kezdeti szakaszára nagy befolyása volt egyes biológiai problémáknak. Biológiai kérdésekből ered az elágazó sztochasztikus folyamatok elmélete. Ma már számos és fontos alkalmazása van ennek a biológia területén kívül is. Mindamellett ebben az elméletben felettebb jelentős helyet foglalnak el azok a problémák, melyek a „születési és elhalási” folyamatok körébe tartoznak. Ezt a felsorolást tovább is lehetne folytatni.

2. Az egy- és többdimenziós sztochasztikus folyamatok elméletéről. Nem kétséges, hogy a biológia szempontjából legnagyobb jelentőségűnek az egy- és többdimenziós sztochasztikus folyamatok elméletének kell lennie, más szóval az olyan sztochasztikus változók elméletének, melyek egy vagy több folytonosan változó paramétertől

függnek. Ilyen mennyiségekre tipikus példaként szolgálhatnak olyan fontos biológiai jellemzők, mint bizonyos populációk adott időpontbeli elemei számának becslése, elektrokardiogramok, azon ingerlések mértékei, melyek szükségesek egyes idegsejtek ingerelt állapotba hozásához, — és a többi. A felhozott példák igen különfélék, nemcsak biológiai sajátosságaik szerint, hanem azért is, mert mindegyik tanulmányozásához sajátos matematikai apparátus kell.

Nem kétséges, hogy a biológiában nem szerencsés csupán a sztochasztikus folyamatok elmélete valamelyik részének felhasználására korlátozni tevékenységünket. A biológia erre a maga teljes sokoldalúságában támaszt igényt. És valóban: a Markov-féle sztochasztikus folyamatok elméletét (vagyis azon folyamatokét, amelyek lefolyásának véletlenszerű menete a $t > t_0$ időpontban — ha ismeretes állapotuk a t_0 időpontban — csupán ettől az állapottól függ és nem függ azoktól az előzményektől, melyek a folyamatot ebbe az állapotba eljuttatták) egész sor szerző már széles körben felhasználta. Az ez irányban végzett mélyenjáró kutatások most is állandóan folynak.

Viszonylag nemrég (1951) jelent meg W. FELLER szemleszerű cikke, amelyet a Markov-folyamatok és láncok bizonyos biológiai feladatokra való alkalmazásairól szóló munkáknak szentelt. LOTKA, VOLTERRA, A. N. KOLMOGOROV régebbi, a létért való küzdelem matematikai elméleteiről írott munkái, FISHER és RIGHT az evolúció matematikai elméletével kapcsolatos vizsgálatai ma is kétségtelenül érdeklődésre számíthatnak. Az utóbbi években nagy figyelmet szenteltek annak, hogy kidolgozzák járványok terjedésének matematikai modelljeit. Nemrég (1957) jelent meg erről a témáról BAILEY speciális monográfiája. Óriási elméleti és gyakorlati jelentősége van olyan elmélet megalkotásának, amely lehetővé teszi ilyen vagy olyan típusú populációk növekedésének tanulmányozását, sőt módszerek kidolgozását is közéleti szempontról fontos halfajták állományának számszerű megállapítására. Nagyfokú érdeklődés mutatkozik Markov-folyamatok iránt napjainkban a kibernetika és információelmélet teréről is. Nagy figyelmet szentel nekik többek közt W. R. ASHBY: „*Introduction to cybernetics*” című ismert könyve is.

A Markov-folyamatok speciális esetét képezik az úgynevezett elágazó folyamatok. Az ilyen folyamatok elméletéről jó áttekintést nyújt B. A. SZEVASZTYANOV cikke (1951); ez főleg fizikusok és kémikusok számára íródott ugyan, azonban biológusok is igazán haszonnal olvashatják.

A biológiai jellemzők egy részének időbeli változása véletlenszerű; ezek semmiféle periodikus folyamatot nem képviselnek. Példaként megemlítjük az elektrokardiogramot, amely a szív működés változásainak egyik objektív jelzője. Sok esetben az említett típusú jellemzők tanulmányozásához elegendően komoly alapot szolgáltat, ha azokat a stacionárius sztochasztikus folyamatok elméletének szemszögéből vizsgáljuk. Reméljük, hogy ezen az úton sikerül majd eléggé megbízható és meggyőző kritériumokat találni bizonyos szívbetegedések diagnosztizálásához. Speciálisan arra számítunk, hogy elektrokardiogramok spektrális függvényének analízise lehetővé fogja tenni, hogy jellemző kritériumokat találjunk a tiszta mitrális stenosis és a tiszta mitrális elégtelenség megkülönböztetéséhez. Nem kétséges, hogy a stacionárius sztochasztikus folyamatok elméletének a magasabb idegtevékenység bizonyos aspektusainak analízisére való felhasználása még jelentős szolgáltatásokat fog tenni az orvostudománynak.

A többdimenziós sztochasztikus folyamatok elmélete, — melynek kidolgozása különösen intenzíven folyt, minthogy meteorológiai jelenségek tanulmányozásával

volt kapcsolatban — bizonyára sikeresen alkalmazható lesz biológiai kérdések megoldásában is. Ha pl. a zsuzsok nevű kártevő bábjainak a répa földön való eloszlása érdekel minket, kérdés, számolhatunk-e azzal, hogy a báboknak telelés előtt a vizsgált területen való szétszóródása valamilyen határozott törvényszerűséget követ? Kétségtelen, hogy ennek a megvizsgálása bizonyos útmutatásokkal szolgálhat az említett kártevővel folytatott harc mértékének jellegét illetően. Vagy tekintsük a plankton eloszlását egy vízmedence meghatározott rétegében. Világos, hogy a plankton sűrűsége a tér minden egyes pontján más és más és változásai nagymértékben véletlen jellegűek. Milyenek ennek a térbeli (és időbeli) változásnak valószínűségi jellegű törvényszerűségei? Összehasonlíthatjuk-e őket porszemek Brown-féle mozgásával a levegőben, vagy pedig másfajta törvényszerűségről van itt szó? Ilyen típusú kérdések felvethetők baktériumoknak vagy spóráknak a levegőben való eloszlásával kapcsolatban is; lehetséges, hogy ugyanilyen módon kell jellemezni halak eloszlását is.

Az a sejtésem, hogy a sorbaállítás elmélete, amely napjainkban a technika és fizika új irányzataival való kapcsolata folytán a viharos fejlődés periódusában van, komoly alkalmazásra találhat egész sor biológiai probléma megoldásában, például a lakosság orvosi ellátása megszervezésének kérdéseivel kapcsolatban.

A sorbaállítás elméletének körébe tartozó problémák a következőképpen vetődnek fel: adva van bizonyos számú személy vagy készülék, amelyeknek az a feladata, hogy kiszolgáljanak bizonyos „fogyasztókat”, melyek időben véletlenszerűen jelentkeznek a kiszolgálási pontban. Ha a „fogyasztó” olyan időpontban jelentkezik, amelyben legalább egy kiszolgáló szabad, akkor rögtön elkezdik kiszolgálni. Ha az összes kiszolgálók foglaltak, a következő lehetőségek vannak: 1. a „fogyasztó” sorbaáll a kiszolgáláshoz és annyi ideig vár, ameddig szükséges; 2. a „fogyasztó” kilép a folyamatból (nem marad a sorban); 3. a „fogyasztó” sorbaáll, de csak korlátozott időtartamon át vár, amelynek hossza valamely τ -t nem halad meg; 4. a várakozási, ill. kiszolgálási idők hosszainak összege nem lépi túl a τ időtartamot. A felsorolt esetek közül az utolsóban a „fogyasztó” kimaradhat az elintézésből, akár azért, mert nem várta ki a kiszolgálást, akár pedig azért, mert sokáig várt és így a kiszolgálásra csak igen kis idő maradtott s az nem volt elegendő. Végül az is megtörténhetik, hogy a „fogyasztó” kivárta a kiszolgálást és az idő elegendő is volt a kiszolgálás lefolytatására. Szokásos a kiszolgálási idő hosszát sztochasztikus változónak felfogni. — A τ mennyiség, amelyről előbb szó volt, lehet állandó is, véletlenszerűen változó is. Az utóbbival van dolgunk az elsősegély állomások mindennapi munkája során. Egy személy, aki egy baleset eredményeként megsérült, ellátása elvégzésére valamilyen τ -nál hosszabb ideig nem várhat. Ez az időtartam egyénről egyénre változik és függ az elszenvedett sérülés jellegétől is. Ha a τ időtartam folyamán nem történik meg az ellátás, akkor a „fogyasztó” kilép a folyamatból (a balesetet szenvedett meghal). Nem kétséges, hogy ilyenfajta problémákkal kell foglalkozni sok más esetben is, amely érdekes a biológia és az orvostudomány számára, — például idegingerületek és más idegrendszeri jelek továbbítási elméletének megalkotása esetében is.

Sztochasztikus folyamatok biológiai problémák megoldásában való felhasználására példaként vegyük a járványok elterjedésének modelljeit, amelyeket M. BARTLETT angol tudós vizsgált könyvében (1956), valamint a Berkeleyben megtartott 3. valószínűségelméleti és matematikai statisztikai szimpóziumon elmondott előadásában (1956). Adva van valamilyen populáció, amely a t időpontban áll $1. s(t)$ egyedből,

amelyek fogékonyak valamilyen betegség iránt; 2. $i(t)$ már megbetegedett egyedből. A populáció kívülről csak s típusú egyedekkel egészülhet ki. A megbetegedni képes egyedek csoportja nem egészülhet ki a betegséget átvészelték közül (azt az esetet is vizsgálhatjuk, amidőn az immunitás csak időleges). A rendszer állapotát minden egyes időpontban két szám jellemzi, s és i — a megbetegedésre szóba jövők száma és a betegek száma. h hosszúsági időtartam alatt a következő átmenetek lehetségesek:

1. $s \rightarrow s-1, \quad i \rightarrow i+1, \quad is\lambda h$
2. $s \rightarrow s, \quad i \rightarrow i-1, \quad i\mu h$
3. $s \rightarrow s+1, \quad i \rightarrow i, \quad v h.$

Jobb oldalt a nyilakkal jelölt átmenetek valószínűségeit írtuk fel. Ha az (s, i) állapot valószínűségét a t időpontban $p_t(s, i)$ -nek jelöljük, $\pi_t(u, v)$ -vel pedig a megfelelő generátorfüggvényt, akkor

$$\pi_t(u, v) = \sum_{i, s=0}^{\infty} p_t(s, i) u^s v^i.$$

A $\pi_t(u, v)$ függvény kielégíti a

$$\frac{\partial \pi}{\partial t} = \lambda u(u-v) \frac{\partial^2 \pi}{\partial u \partial v} + \mu(1-u) \frac{\partial \pi}{\partial u} + v(v-1)\pi$$

differenciálegyenletet.

Az elméleti eredményeket BARTLETT összehasonlította kanyarós megbetegedésekre vonatkozó klinikai megfigyelések eredményével: A valóságos viszonyok elméleti megközelítése az említett megbetegedések esetében teljesen kielégítőnek mutatkozott.

Kétségtelen, hogy a BARTLETT vizsgálta modell primitív volt. Minden esetben számításba kellett volna venni az

$$s \rightarrow s-1, \quad i \rightarrow i$$

átmenet lehetőségét, minthogy valamely egyed más okból kifolyólag is elhalhat, vagy elszigetelődhet a megbetegedést létrehozó okoktól stb. Komoly kétségek merülhetnek fel azzal kapcsolatban is, hogy lehetséges állandó λ , μ és v -vel dolgozni. Mindazonáltal már az ilyen primitív modellkészítés is valamivel közelebb visz bennünket járványok terjedése mechanizmusának megértéséhez.

A LJAPUNOV-féle centrális határeloszlástételnek, melynek a valószínűségelmélet technikai és fizikai alkalmazásaiban alapvető szerepe van, számos alkalmazási területe nyílik a biológiában is. Mindazon esetekben, melyekben feltehető, hogy valamilyen jellemző kialakulása nagy számú független vagy gyengén függő mennyiség összegének ráhatása alatt folyik le, ennek a jellemzőnek az eloszlása közel lesz a normálhoz. Ha például tekintjük egyazon típusú, fajtájú és nemű, nagyszámú egyed összességét, melyek közelítőleg egyforma feltételek között fejlődtek ki, akkor meghatározott méretük, például valamelyik csontjuk hossza, közelítőleg normális eloszlású lesz. Ezt a körülményt széles körben felhasználják az antropológiában. Az antropológiai normáknak pedig, mint ismeretes, jelentős szerepe van az alkalmazások szempontjából: ezek alapján határozzák meg, milyen százalékban ésszerű készíteni bizonyos méreteket lábbelik, ruhák, kalapok gyártásakor. Ezeket a normákat, valamint ipari célokra való felhasználásuk alapelveit behatóan tanulmányozta M. V. IGNATYEV.

3. Valószínűségek becslése; statisztikai megfigyelések megszervezése. A valószínűségelmélet és a statisztika klasszikus feladata: ismeretlen valószínűségek becslése, napjainkban is az érdeklődés középpontjában áll. Az orvostudomány számára is különösen hasznos ilyenfajta feladattal találkoztunk nemrég, amikor Kijevben megkíséreltük matematikailag analizálni a szívbetegségek diagnosztikáját. Ennek során, hogy ne kelljen foglalkoznunk az összes szívbetegségekkel (melyekből kimeríthetetlenül sok van) csak két szívbetegséggel foglalkoztunk: a mitrális stenosiszal és a mitrális elégtelenséggel. Ismeretes, hogy mindkét betegség eléggé széles körben elterjedt és sokban egymásnak ellentétei: az ilyen esetek közül az egyikben a modern orvostudománynak az operatív beavatkozás jól kidolgozott módszerei állnak rendelkezésére; ezek a módszerek komoly sikerrel járnak, csak idejében állapítsák meg a diagnózist. A diagnózis pontos felállítása az orvostudomány jelen állása mellett azonban még nagyon fáradságos, minthogy igen ritkán találkozunk az említett megbetegedés világos megnyilvánulásával. A nehézségek akkor lépnek fel, amikor az alapmegbetegedésre komplikációk rétegződnek rá, amelyek anélkül is megzavarják a betegség mibenlétéről alkotott elképzelésünket. Ezenfelül ugyanazok a tünetek, melyek alapján a megbetegedéseket diagnosztizálják, másféle betegségeknél is előfordulnak. Végeredményben nem ritkán az a helyzet, hogy egy és ugyanazon beteggel kapcsolatban különböző orvosok lényegesen különböző megállapításokra jutnak. Világos, hogy az így előadódó helyzetből az lesz a kiút, hogy szisztematikusan tanulmányoznak lehetőleg nagyszámú különböző diagnosztikai adatot, ezek valószínűségeinek becsléseit különböző megbetegedések esetében és aztán mindezeket az adatokat egyidejűleg tekintik át. Lehetséges, hogy a beteg ilyen „körültekintő megvizsgálásához” adatok ezreit kell áttekinteni, melyek mindegyike a maga részéről csak kevésbé járul hozzá a diagnózis felállításának nagy ügyéhez, — mindenesetre azonban mégis csak elősegítik a helyes következtetések elérését. Lehetséges, hogy az összes adatoknak az áttekintése már túlhaladja az orvos erőit, minthogy a szokásos munkamódszerek mellett óriási időráfordítást követelnek. Mármost ezt a már csak technikai jellegű feladatot rábízhatjuk elektronikus számológépekre is.

Végeredményben a következő, tisztán statisztikai jellegű problémát vehetjük fel itt: válasszuk meg az A_1, A_2, \dots, A_n tünetek olyan komplexusát, hogy előre meghatározott fajta tünetek meglétének, illetve hiányának adott kombinációja 1-hez elég közeli valószínűséggel arra a következtetésre vezessen, hogy a beteg megadott B_1, B_2, \dots, B_m betegség-összesség közül épp a meghatározott B_k betegségben szenved. Bizonyos értelemben mondhatjuk, hogy a probléma ilyen kezelése teljesen megfelel a már létező diagnosztizálási módszerek alap gondolatának. Az, amiről előbb beszéltünk, csak kvantitatív becslésekkel egészíti ki az évezredek tapasztalatán alapuló kvalitatív becslést. Világos, hogy a szóban forgó feladat arra vezethető vissza, hogy becsléseket kell adnunk mind

$$P\{B_i|A_j\} \quad \text{és} \quad P\{A_j|B_i\}$$

típusú valószínűségekre, mind pedig a

$$P\{A_{i_1}, A_{i_2}, \dots, A_{i_s}; \bar{A}_{j_1}, \bar{A}_{j_2}, \dots, \bar{A}_{j_r} | B_i\}$$

és

$$P\{B_i | A_{i_1}, A_{i_2}, \dots, A_{i_s}; \bar{A}_{j_1}, \bar{A}_{j_2}, \dots, \bar{A}_{j_r}\}$$

valószínűségekre, az i_1, i_2, \dots, i_s és j_1, j_2, \dots, j_r indexek tetszőleges megválasztása mellett.

Az adatok közé természetesen besoroljuk mind az anamnézis adatait, melyeknek kiváltképpen kvalitatív jellegük van (szóba jöhet elkékülés, oedema, különböző fájások panaszolása és hasonlók), mind pedig más jellegű adatok, az arteriás és vénás vérnyomás értékei, elektrokardiográfiai vizsgálat, elektrokardiogram, fonokardiogram adatai és a többi. Továbbá figyelembe kell venni természetesen, hogy a beteg neme és kora, valamint foglalkozása is befolyásolja a végső megállapítást.

Meg kell mondanunk, hogy már az első lépések alkalmával számunkra nem várt mértékben nagy nehézségekbe ütköztünk. Kezdetben úgy látszott, hogy kimeríthetetlen statisztikai anyag van kezünkben, mely kórtörténetekből, számos profilaktikus jellegű orvosi vizsgálati eredményből, újoncok és a lakosság más kategóriái szemléjének eredményeiből gyűlt össze. Azonban már korán különböző okokból le kellett mondanunk ezen adatok tekintélyes részének feldolgozásáról. Mindenekelőtt azt tapasztaltuk, hogy az általános vizsgálódásokat különböző eszközökkel hajtják végre, e vizsgálódások részben nem tartalmazzák a szükséges adatokat, emellett kizárólag vegyes és teljes bizalmat nem mindig érdemlő anyagot képviselnek. Másodsor, az ilyenféle vizsgálatok eredményei nagymértékben az orvos szubjektív benyomásának jellegét viselik magukon és nem támasztják alá őket objektív adatok. Közelióleg ugyanilyen körülmények készítettek minket arra, hogy eltekintsünk nagyszámú kórtörténet tanulmányozásától és egyelőre csak ama betegek kórtörténeteire korlátozzuk vizsgálódásunkat, akik szívműtéten estek át és így esetükben bizonyos lehetőség volt objektív ellenőrzésre. Sajnos, gyűjtésünkben csupán kishatómú ilyen adat volt. Kijev, Moszkva és Leningrád más klinikai anyagának felhasználása bizonyos nehézségekbe ütközött, minthogy a diagnosztikai módszerek, az elektrokardiogramok kidolgozása ezeken a helyeken lényegesen más volt, mint N. M. AMOSZOV professzor klinikáján, amellyel együttműködve kezdtük el munkánkat. Az utóbbi időben elhatároztuk, hogy a statisztikai anyag teljesebbé tételére céljából felhasználjuk a hullaházakban lefolytatott boncolások eredményeit is.

Nem mondhatjuk, hogy a fentebb felvázolt feladatok megoldásában már messzire jutottunk volna. A munka megindulásának még csak a legkezdetén vagyunk. Egyelőre részeredményeket kaptunk, amelyek alapján kétségbe vonhattuk a megbetegedések bizonyos használatos, vagy kiemelni javasolt szimptomái használhatóságát — és olyan szimptomákét is, amelyek a betegség bizonyos stádiuma beálltának becslését teszik lehetővé. Ugyanakkor statisztikai alátámasztást is kaptunk bizonyos diagnosztikai kritériumokhoz is.

Vizsgálatainkban bizonyos figyelmet szenteltünk a statisztikai adatgyűjtés megszervezése bonyolultságának is. A megfigyelések megszervezésének, a következtetések mennyisége és jellegének kérdése egyik a legfontosabbak közül. Hogyan szervezzünk meg például kísérleteket adott körzetben elszaporodott rágszálók, vagy meghatározott fajtájú állatok adott területen élő mennyisége becslésére? Hogyan becsüljük fel a közélélmzési halkészleteket adott halastóban és milyen megfigyeléseket kell ehhez elvégezni? Ezek a kérdések biológusok és matematikusok tartós együttműködését kívánják. Mindenesetre rendkívül szükségesnek tartom matematikusok részvételét hosszabb expedíciókban, melyek folyamán a matematikus nemcsak a biológusokat érdeklő problémákat ismeri meg, hanem a jelenségeknek azokat a jellemző sajátosságait is, melyek kvantitatív elméletét neki kell majd kialakítania.

4. A statisztika nem-paraméteres módszereiről. Az utóbbi években gyors ütemben fejlődő statisztikai módszerek közül különösen figyelmet érdemelnek az úgy-

nevezett nem-paraméteres módszerek. A statisztika említett fejezete tárgyalásának van szentelve SIDNEY SIEGEL könyve (1956), melynek „receptkönyv” jellege van és különösen alkalmas azok számára, akik a statisztikai módszereknek csupán elvi oldalát kívánják megismerni. Egy másik nemrég megjelent könyv, FRASER munkája (1957), túlnyomórészt távolról sem foglalkozik nem-paraméteres módszerekkel és mindenesetre messzire esik azok érdeklődési körétől, akik arra törekednek, hogy gyorsabban létesüljön kapcsolat biológiai problémák és kísérleti adatok kidolgozásának statisztikai módszerei között.

A matematikai statisztika klasszikus problémafelvetési módjai változatlanul feltételezik, hogy a vizsgálat alá vont sztochasztikus változók olyan valószínűség-eloszlással rendelkeznek, amelyeknek analitikus alakja ismert. Felteszik, hogy ezek az eloszlások véges számú paramétertől függenek és ezek értékének becslése, valamint ezekkel kapcsolatos feltételekre vonatkozó különböző hipotézisek ellenőrzése a statisztikai vizsgálatok alapfeladatai. Ennek során — olykor elegendő alapról is — különös figyelmet szenteltek a normális eloszlásnak. Tudjuk, még ma is az a helyzet, hogy ha egy vizsgált sztochasztikus változó eloszlásáról semmi sem ismeretes, akkor azt normálisnak tekintik, aztán pedig az egész feladatot arra vezetik vissza, hogy jól kidolgozott szabályok szerint becsléseket keressenek annak paramétereire. Az ilyen eljárás nem mindig jogosult és nem is mindig hasznos. Nem ritkán az a probléma, amely a biológust érdekli, egyáltalán nincs is befolyással arra a feltételezésre, hogy a vizsgált eloszlás ebbe vagy abba a meghatározott eloszlásosztályba tartozik. Például csak annak a hipotézisnek a lehetőségét kell megvizsgálni, hogy a vizsgált megfigyelés-sorozat nem változó eloszlású sokaságból vett mintának tekinthető. Rendkívül fontos az orvostudomány, a szántóföldi kísérletezés vagy tetszőleges kísérleti munka szempontjából csupán annak a ténynek a megállapítása, hogy két független mintát egy és ugyanazon sokaságból valónak tekinthetünk-e.

Nem kevésbé lényeges olyan szabály kidolgozása, amely lehetővé tenné annak megállapítását, hogy egy bizonyos megfigyelés-csoport egyes eredményei oly sztochasztikus változóra vonatkoznak, amely „kisebb” egy másik sztochasztikus változónál, amelyre vonatkozólag egy másik megfigyelés-csoport áll rendelkezésünkre. Ilyen fajta feladat nevezetesen mindenkor felmerül, amikor azt kell eldönteni, nagyobb terméshozama van-e valamely új fajtának, mint a korábban termesztettnek, vagy pedig azt, hogy egy kitenyésztett tehénfajta jobban örökíti-e ezt vagy azt a jellegzetességet, mint mások és így tovább. Minden ilyen esetben a kutatót csak az említett kérdések érdeklik, és számára teljesen közömbös az a kérdés, hogy valamely szereplő eloszlás a normális eloszlás családba tartozik-e vagy egy másikba. Az összes ilyen típusú problémák, melyekben statisztikai szabályok kidolgozásáról van szó, a kiindulási eloszlások tág osztályával kapcsolatban felvethetők (felvethetők pl. az összes folytonos vagy abszolút folytonos eloszlásokkal kapcsolatban); az ilyeneket manapság nem-paraméteres problémáknak nevezzük.

Nem kívánok példákat felsorolni itt bármilyen nem-paraméteres statisztikai szabályra, minthogy ezt természetesebb máshelyütt megtenni. Ilyen másik helynek tekintem viszonylag nem terjedelmes könyvek oly sorozatát, amely bevezethetné az olvasót a modern statisztika módszereibe és amelyekben elég sok jó, biológiai jellegű illusztráló példa volna. Az ilyenféle nem nagyterjedelmű monográfiák formáját persze jól át kellene gondolni. Nem szabadna ezeket megterhelni tisztán numerikus

részletekkel; feladatuk főleg abból állna, hogy megvilágítsák a statisztikai módszerek alapjait, alkalmazási feltételeit és felhasználási szkémáit.

Ilyen, a biológus, az orvos, a mezőgazdasági szakember számára életbevágóan fontos könyvsorozat kiadásának megvalósulása lényeges segítséget nyújt majd az említetteknek. Amellett ez a lépés abban is segítséget hoz majd, hogy elhárítsuk a nem-matematikus koponyákban a statisztikai módszerek konkrét orvosi és biológiai vizsgálatokban való felhasználásával kapcsolatban fel-felmerülő ama erős emocionális reakciót, amelyről oly színesen írt A. BRADFORD HILL nemrég orosz nyelven is megjelent „*Az orvosi statisztika alapjai*” című könyvében. Engedjék meg, hogy idézzem itt idevágó szavait. „Bosszantó, ha aközben, hogy nagy munkával elsajátítható módszerekkel tanulmányozunk egy problémát, megtudjuk, hogy következtéseink kétkedésbe döntenek — sőt, elriasztanak tőlünk — bárkit, aki nem tudja önállóan reprodukálni megfigyeléseinket. Ahhoz, hogy elismerjük, hogy a hiba bennünk magunkban van, több hidegvér szükséges, mint amennyivel pillanatnyilag rendelkezünk.” Célunk nem annyira az, hogy ellenőrizzük mások megállapításait, hanem inkább az, hogy lehetőleg nagyszámú értékes, fontos és igaz megállapítást tehessünk a természet — speciálisan az élő természet — folyamatairól.

IRODALOM

- BARTLETT, M. S.: *An introduction to stochastic processes*. With special reference to methods and applications. Cambridge University Press, London, 1956. (Orosz fordítása: Бартлет, М. С.: Введение в теорию случайных процессов. Издательство Иностранной Литературы, Москва, 1958.)
- Севастьянов, Б. А.: Теория ветвящихся случайных наук, т. VI. (1951), вып. 6(46), 47—99.
- HILL, A. BRADFORD: *Principles of medical statistics*, 4th edition. Lancet, London, 1949. (Orosz fordítása: Хилл, Бредфорд А.: Основы медицинской статистики. Издательство Медицинской Литературы, Москва, 1958.)
- ASHBY, W. R.: *An introduction to cybernetics*. Chapman and Hall, London, 1956. (Orosz fordítása: Эшби, Росс У.: Введение в кибернетику. Издательство Иностранной Литературы, Москва, 1959.)
- BAILEY, N. T. J.: *The mathematical theory of epidemics*. Hafner Publishing Co., New York, 1957.
- BARTLETT, M. S.: Deterministic and stochastic models of recurrent epidemics. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954—1955, vol. IV, pp. 81—109. University of California Press, Berkeley and Los Angeles, 1956.
- FELLER, W. K.: Diffusion processes in genetics. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950, pp. 227—246. University of California Press, Berkeley and Los Angeles, 1951.
- FRASER, D. A. S.: *Nonparametric methods in statistics*. Wiley and Sons, New York, 1957.
- SIEGEL, S.: *Nonparametric statistics for the behavioral sciences*. McGraw—Hill, New York, 1956.

Fordította: dr. Medgyessy Pál,
a matematikai tudományok kandidátusa