© Typotex Kiadó

ELŐSZÓ

A könyv alapjául azok a matematikai statisztikai előadások szolgálnak, amelyeket a szerző sok éven át tartott a novoszibirszki egyetem matematikai fakultásán. A lehetőleg minél jobban felépített, érthető, ugyanakkor a tudományág szintjének megfelelő variánsok keresése közben az idő folyamán az előadás anyaga nem egyszer megváltozott. Különböző variánsokat próbáltunk ki, kezdve az alapfeladatok (becslések, próbák és tulajdonságaik) főként receptúraszerű kifejtésével és végezve az általános játékelméleti jellegű felépítéssel, amelyben a becsléselmélet és a hipotézis vizsgálat mint egy egységes megközelítés speciális esetei jelennek meg. Az időbeli korlátok (egyetlen szemeszter) nem tették lehetővé, hogy egyesítsük ezt a két, egymást kiegészítő változatot, amelyek mindegyike különkülön nyilvánvaló hiányosságokkal rendelkezik. Az első esetben a sok konkrét adat zavarta az általános felépítésmódot. A második változatból viszont hiányoztak az egyszerű, konkrét eredmények, és úgy tűnt, hogy túlságosan megterhelő a sok új, bonyolult, nehezen elsajátítható fogalom. Szemlátomást a leginkább alkalmazhatónak az a variáns látszott, amelyben a becsléselmélet és a hipotézis vizsgálat ismertetése után egyenes út vezet az optimális eljárások megkereséséhez

A könyv alapanyaga a különböző időpontokban tartott kurzusok anyagából áll össze – kibővítve azokat olyan részekkel, amelyek jelenlétét maga a felépítés logikája diktálta. A fő cél a téma modern eredményeinek tárgyalása volt, párosítva ezt a maximálisan lehetséges érthetőséggel és a logikus matematikai felépítéssel.

A könyv tartalma 3 fejezetre és a Függelékre oszlik.

Az 1. fejezetben a matematikai statisztika alapjait képező empirikus eloszlások tulajdonságait (főként aszimptotikus tulajdonságait) vizsgáljuk.

www.interkonyv.hu © Borovkov, Michaletzky György

© Typotex Kiadó

A 2., illetve 3. fejezetben a becsléselmélet és a hipotézis vizsgálat elmélete szerepel. Ezen fejezetek mindegyikének első része a kitűzött feladatok megoldásaihoz és az optimális eljárások megtalálásához vezető lehetséges utak leírását tartalmazza. A második részek az aszimptotikusan optimális eljárások felépítését tartalmazzák.

A könyv ezen felül 7 Függeléket tartalmaz. Ezek az alapszöveg azon állításaival kapcsolatosak, amelyek bizonyításai vagy a jellegüknél fogva, vagy a nehézségük miatt kívül esnek a felépítés keretein.

Ezenkívül bibliografikus megjegyzések is szerepelnek a könyvben – egyáltalán nem törekedve a teljességre –; ezek lehetővé teszik, hogy követni lehessen a matematikai statisztika alapvető irányvonalainak keletkezését és fejlődését. Ennek során mindenütt, ahol ez lehetséges volt, inkább a monográfiákra (mint könnyebben elérhető irodalomra) hivatkoztunk, és nem az eredeti cikkekre.

Jelenleg elég sok matematikai statisztika könyv létezik. A következő négyet választjuk ki közülük – ezekben nagy mennyiségű, a tudomány jelenlegi állását tükröző, anyag szerepel – ezek a könyvek G. Cramer 37, E. Lehman 40, S. Zaks 30 és I. A. Ibragimov és R. Z. Hasminszkij 31. A jelen könyv felépítésére közülük legnagyobb hatással a 31 monográfia (e könyv némely ötletét felhasználtuk a 2. fejezet 23–35, 27–29 pontjaiban) és a 40 monográfia volt (a 3. fejezet 5–8 pontjainak felépítése tartalmilag közel áll a 40 monográfia megfelelő részeihez). A többi rész felépítésének szerkezete nincs szoros összefüggésben az ismert könyvekkel.

Jelen könyv az ismert eredményekkel és eljárásokkal együtt olyan új témaköröket is tartalmaz, amelyek egyszerűsítik a tárgyalásmódot; ezenfelül egy sor metodológiai tökéletesítés és néhány új eredmény, illetve a monografikus irodalomban először publikált eredmény is szerepel benne.

Alább megadjuk a könyv metodológiai struktúrájának rövid leírását (lásd ugyancsak a tartalomjegyzéket és az egyes fejezetek rövid előszavát).

Az 1. fejezet 1. és 2. pontjában a minta, az empirikus eloszlás fogalmát vezetjük be, itt mondjuk ki a Glivenko–Cantelli tételt. Ez utóbbit lehet a statisztikai következtetéseket megalapozó állításnak tekinteni.

A 3. pontban kétfajta típusú statisztikát definiálunk (I és II típusú statisztikát), ezek magukban foglalják a gyakorlatilag érdekes statisztikák túlnyomó többségét. Ezeket a statisztikákat mint egy (bizonyos feltételeket kielégítő) G funkcionálnak a $\mathbf{P}_n^*$ tapasztalati eloszlástól függő $G(\mathbf{P}_n^*)$ értékeként definiáljuk. Később, a 7. és 8. pontban ezen statisztikák határeloszlásáról szóló tételek vannak. Ez megkönnyíti a téma további kifejtését és szükségtelenné teszi, hogy minden egyes konkrét statisztika esetében végigkövessük lényegében ugyanazt a gondolatmenetet, amely akkor nem tartozik a dolog lényegéhez.

Az 5. pontban az eloszlások és a momentumaik konvergenciájáról szóló segédtételeket gyűjtöttük össze. (Ezeket a könyvben „folytonossági” tételeknek nevezzük.) Ez szintén megkönnyíti a későbbi tárgyalásmódot.

www.interkonyv.hu © Borovkov, Michaletzky György

© Typotex Kiadó

A 6. pontban (első olvasáskor nem feltétlenül szükséges átnézni) kimondjuk, hogy az $F_n^*(t)$ tapasztalati eloszlásfüggvény nem más, mint egy feltételes Poisson-folyamat, és megfogalmazzuk (az I. Függelékben szerepel a bizonyítás) a $\sqrt{n}(F_n^*(t) - F(t))$ folyamat Brown-hídhoz való konvergenciájáról szóló tételt.

A 10. pontban vezetjük be a simított tapasztalati eloszlások fogalmát, amelyek lehetővé teszik, hogy ne csak magukat az eloszlásokat, hanem a sűrűségfüggvényüket is közelíteni tudjuk.

A 2. Fejezet 3. pontjában, amely az ismeretlen paraméterek becsléséről szól, a becslések gyártásának egyfajta egységes eljárását mutatjuk be, ezt behelyettesítéses módszernek nevezzük. Ez abból áll, hogy a minta $\mathbf{P}$ eloszlásától függő $\theta = G(\mathbf{P})$ paraméter becslését $\theta^* = G(\mathbf{P}_n^*)$ alakban keressük, ahol $\mathbf{P}_n^*$ az empirikus eloszlás. Mindegyik „ésszerű" becslés, amelyet a gyakorlatban használunk, behelyettesítéses becslés. A G funkcionál alkalmas megválasztásával érhetjük el a becslés optimalitását. Ha a $\theta^* = G(\mathbf{P}_n^*)$ statisztika I vagy II típusú statisztika, akkor az 1. Fejezet eredményei alapján azonnal állíthatjuk a becslés konzisztenciáját és aszimptotikus normalitását. A 4. és 5. pontban ezt az eljárást olyan becslésekkel illusztráljuk, amelyeket a momentum módszer és a legkisebb távolság módszerének segítségével kapunk. Ugyanebből a nézőpontból lehetne szemlélni a maximum likelihood becslést is (6. pont), azonban ennek közvetlen vizsgálatával lehetővé válik, hogy – későbbiekben szükséges – mélyebb eredményeket kapjunk.

A 2. Fejezetben lényegében két fajta módon hasonlítjuk össze a becsléseket: négyzetes középben (az $\mathsf{E}_\theta(\theta^* - \theta)^2$ összehasonlításával) és aszimptotikusan (az aszimptotikusan normális becslések osztályán belül összehasonlítjuk a $\sqrt{n}(\theta^* - \theta)$ mennyiség határeloszlásának szórását). Parametrikus esetben ennek alapján lehetővé válik, hogy kiválasszuk az optimális becslések 3 féle osztályát: a b rögzített torzítással rendelkező becslések K_b osztályán belül hatásos becsléseket, a Bayes-féle és a minimax becsléseket. Ugyanezen alapelvek szerint lehet meghatározni az aszimptotikusan optimális becslések osztályait. A hatásos becslések elkészítésekor a követhető tradicionális módszereket lehet használni: közülük az elsőnek minőségi jellege van és az elégségesség elvével van összefüggésben (12–14. pont) a második a Cramer–Rao egyenlőtlenségből adódó mennyiségi viszonyokon alapul (16. pont), a harmadik az invarianciák szemügyrevételével kapcsolatos (17., 19. pont), amely lehetővé teszi, hogy leszűkítsük a figyelembe vett becslések osztályát. Az aszimptotikusan optimális becslések megkeresésének, és a likelihood függvény aszimptotikus tulajdonságai vizsgálatának szenteltük a 20–30. pontokat. A 20. pont tartalmazza az integrál típusú Cramer–Rao-egyenlőtlenséget, amely lehetővé teszi, hogy például egyszerű feltételeket kapjunk arra, hogy mikor aszimptotikusan Bayes-féle vagy minimax egy becslés, és hogy megalapozza a becslések alkalmas $\widetilde{K_0}$ részhalmazának kiválasztását; elegendő erre korlátozódni, ha aszimptotikusan hatásos becsléseket keresünk. Ez lehetővé teszi, hogy a maximum likelihood becslés aszimptotikus tulajdonságainak tanulmányozása útján (25. pont) a szóban forgó becslésekről

www.interkonyv.hu © Borovkov, Michaletzky György

© Typotex Kiadó

azonnal állíthassuk, hogy aszimptotikusan Bayes-félék és minimax becslések, és hogy aszimptotikusan hatásosak a $\widetilde{K}_0$ osztályon belül. A 21–24. pontok a segédállításokat tartalmazzák. A paraméterek intervallumbecslését a 31. és 32. pontban, illetve a 3. Fejezet 8. pontjában vizsgáljuk.

A 3. Fejezetet a hipotézisvizsgálatnak szenteltük. Az 1. és 2. pontokban a véges sok egyszerű hipotézis esetét tekintjük. Kiválasztjuk (a becsléselmélethez hasonlóan) az optimális próbák három típusát – részosztályokon belül legerősebb, Bayes-féle, minimax. Összefüggéseket fedezünk fel ezek között a próbák között, és megtaláljuk az általános alakjukat. Eközben a vizsgálat alapjául a Bayes-féle elv szolgál (és nem a Neyman–Pearson-féle lemma), amely – nézetünk szerint – leegyszerűsíti és világosabbá teszi a tárgyalásmódot. A 3. pontban két egyszerű hipotézis közötti döntés alapjául szolgáló próbák aszimptotikus megközelítése, illetve ezek összehasonlítása szerepel. A 4. pontban a két összetett hipotézis közötti döntés általános feladatát vizsgáljuk, és definiáljuk az optimális próbák különféle osztályait (egyenletesen legerősebb, Bayes-féle, minimax). Az 5. pont foglalkozik az egyenletesen legjobb próbák megkeresésével azokban az esetekben, amikor ez egyáltalán lehetséges. A 6., 7. pontban ugyanezt a feladatot oldjuk meg, azonban próbáknak a torzítatlanság és az invariancia alapján leszűkített osztályán belül. Eközben a vizsgálat alapjául, ugyanúgy, mint az 1. és 2. pontban, a Bayes-féle elv szolgál. A kapott eredmények segítségével a 8. pontban megkonstruáljuk a legpontosabb konfidenciahalmazokat. A 9. pontban a Bayes-féle és a minimax próbákat vizsgáljuk. A 10. és 13. pontok likelihood hányados próbával foglalkoznak. Ez sok esetben egyenletesen legjobb és teljesen általános feltételek mellett aszimptotikusan Bayes-féle. A 15–17. pontokban folytatjuk a likelihood hányados próba aszimptotikus optimalizálásának vizsgálatát. A 11. pontban a szekvenciális analízis feladatkörében vett optimalizálását mondjuk ki. A 14., 15. pont foglalkozik azzal, hogy közeli hipotézisek közötti döntés feladatára keressen aszimptotikusan optimális próbákat, és megadja ezek egyszerű, világos alakját az alapvető statisztikai feladatok esetében.

Az igazi különlegessége ennek a könyvnek abban rejlik, hogy ebben csak olyan statisztikai feladatok szerepelnek, amelyek egyetlen minta felhasználásával kapcsolatosak; a két vagy több mintával kapcsolatos feladatok, valamint a statisztikai feladatok általános játékelméleti megközelítése egy külön könyvben szerepelnek, amely egyenes folytatása és kiegészítése a jelen könyvnek.

A könyvnek szerteágazó a célkitűzése, Természetesen teljes terjedelmében közelebb áll a matematikai statisztikával foglalkozó specialisták kandidátusi minimumához, mint az egyetemi hallgatók tankönyvéhez. Ugyanakkor igyekeztünk különféle jelzések segítségével megkönnyíteni az első átolvasást is, ezek a hallgató számára is hozzáférhetővé teszik a könyvet. A kiemelkedően nehéz és a tartalmában nagyot ugró pontokat csillaggal jelöltük, ezeket első olvasáskor át kell ugrani, mint azokat a szövegeket is, amelyek kisbetűvel vannak szedve. Ezen felül a technikailag jóval bonyolultabb – a többdimenziós paraméterhez kapcso-

www.interkonyv.hu © Borovkov, Michaletzky György

© Typotex Kiadó

lódó – esetek tárgyalását majdnem mindig külön részben és pontban választottunk szét; ezeket szintén el lehet hagyni.

A különböző egyetemek oktatói, akik már legalábbis részben ismerősek a témakörrel, kiválaszthatnak a könyvből olyan pontokat (igen sok variáns lehetséges), amelyek felhasználásával (nem szükségképpen teljes felhasználásával) összeállíthatják egy matematikai statisztikai kurzus anyagát. Például egy változat: az 1. Fejezet 1.,3.,5. pontja; a 2. Fejezet 2–4., 6–12., 14., 16. (21. 23–25.), 31., 32. pontja, a 3. Fejezet 1., 2., 4., 5., 12. (13., 16.) pontja. A zárójelbe tett pontok az aszimptotikusan legjobb eljárásokkal foglalkoznak. Ezeket a hallgatók felkészültségétől függően vagy maximálisan könnyebbé kell tenni, vagy esetleg teljesen elhagyni.

A könyv tanulmányozása feltételezi a valószínűségszámítás elméletének ismeretét olyan mélységig, ahogy A. A. Borovkov tankönyve [11] tartalmazza. Az erre a könyvre való hivatkozások, ellentétben a többi hivatkozással, olyan helyeken jelennek meg, amelyekről feltesszük, hogy az olvasó ismeri őket, és ily módon inkább csak emlékeztetőül szolgál.

Az egyes pontok számozása fejezeteken belül önálló, ugyancsak az egyes tételek (lemmák és példák) számozása az egyes pontokon belül. A kényelmesebb olvasás kedvéért különböző módon hivatkozunk a tételekre, lemmákra, példákra, képletekre stb. attól függően, hogy milyen messze találhatók az olvasott helytől. Ha hivatkozni kell az 1. Tételre vagy a 12 képletre az olvasott ponton belül, arra a hivatkozás a következő alakot ölti: 1. Tétel, 12 képlet. Ha az 1. Tételre vagy a 12 képletre a Fejezet valamely korábbi pontjából kell hivatkozni, akkor ilyen alakú a hivatkozás: 13.1. Tétel, 13.12 képlet. Végezetül, ha a hivatkozás egy másik fejezetre vonatkozik, akkor a fejezet számának mutatója is megjelenik (az első szám). Például, a 2.13.1. Tétel a 2. fejezet 13. pontjának 1. tételét jelzi, a 2.13.12. képlet a 2. fejezet 13. pontjának 12 képletét. Ugyanez vonatkozik az egyes pontok jelzésére. A 13. pontra való hivatkozás a szóban forgó fejezet 13. pontjára vonatkozik, a 2.13. pontra való hivatkozás pedig a 2. fejezet 13. pontját jelzi.

A □ jel a bizonyítások végét jelzi.

A könnyebbség kedvéért a könyv végén tárgymutató és jelölésjegyzék van.

Ezen könyv megírása igen nehéz, sok lépésből álló munka volt.

Az eredeti előadásjegyzet nyomdára való előkészítésében és a hiányosságok megszüntetésében jelentős segítséget nyújtott nekem I. Sz. Boriszov. A kézirat második változatát kérésemre K. A. Borovkov olvasta át, aki a hasznos tanácsok és az általa észrevett hibák hosszú listáját adta át nekem. Újabb kritika után kutatva A. I. Szahanyenkohoz fordultam azzal a kéréssel, hogy ismerkedjék meg a kézirattal. Ő szintén hosszú listáját adta a felépítés megjobbítását célzó megjegyzéseinek és javaslatainak, amelyek közül sokat fel is használtam. A leglényegesebb változtatás a 2. Fejezet 16., 21., 23., 29. pontjaiban, a 3. Fejezet 13–15.

www.interkonyv.hu © Borovkov, Michaletzky György

© Typotex Kiadó

pontjaiban, a II. és V. Függelékben lévő bizonyításokat érte (lásd a bibliográfiai megjegyzéseket).

Igen sok értékes, a könyv megjobbítására szolgáló megjegyzést kaptam D. M. Csibiszovtól. A kéziratot V. V. Jurinszkij és A. A. Novikov nézte át, és egy sor hasznos megjegyzést tettek. Mindegyik megnevezett kollégámnak és azoknak is, akik így vagy úgy segítettek nekem a könyvvel kapcsolatos munkámban, itt szeretném kifejezni mély és szívből jövő hálámat és köszönetemet a munkájukért és a könyv megírásában való együttműködésükért.

1982. szeptember

A. A. Borovkov

www.interkonyv.hu © Borovkov, Michaletzky György