

Jegyzetek

1. rész. Kísérletek megtervezése

1. FEJEZET. KONTROLLÁLT KÍSÉRLETEK

1. Az összehasonlításos módszert először feltehetően a XIX. század elején használták, annak kimutatására, hogy az érvágás mégsem annyira hatékony ellenszere a tüdőgyulladásnak. L. Pierre Louis; *Recherches sur Les Effets de la Saignée dans quelques Maladies inflammatoires: et sur l'Action de l'Emétique et des Vésicatoires dans la Pneumonie* (Párizs: J.B. Baillière, 1835; angolul, 1836; újranyomva: The Classics of Medicine Library, 1986, Birmingham, Alabama.) Részletesebben tárgyalja R.H. Shryock, *The Development of Modern Medicine* (University of Pennsylvania Press, 1936, 163. o.) Továbbá említenünk kell Lindnek a C-vitamin skorbutellenes hatásával kapcsolatos vizsgálatát; L. K.J. Carpenter, *The History of Scurvy and Vitamin C* (Cambridge University Press, 1986.)
2. Thomas Francis, Jr. et al., „An evaluation of the 1954 poliomyelitis vaccine trials – summary report,” *American Journal of Public Health* vol. 45 (1955), 1-63. o. Lásd még P.Meier cikkét, „The biggest public health experiment ever: the 1954 field trial of the Salk poliomyelitis vaccine”, in J.M.Tanur et al., *Statistics: A Guide to the Unknown*, 3rd. ed. (Wadsworth, 1989). Olvasmányosabb összefoglalást találnak Jane S. Smith, *Patenting the Sun* (Anchor, 1990) c. könyvében.
3. Újabb keletű példa: sok halálesetet okoztak szív-aritmia elleni gyógyszerek. Thomas J. Moore, *Deadly Medicine* (Simon & Schuster, 1995.)
4. A statisztikusok szerint „Amit lehet, tarts kézben – minden mást randomizálj”. Az illesztés és a blokkosítás azon az áron csökkenti a varianciát, hogy bonyolítja az elemzést. Lásd a 19. fejezet 12-es jegyzetét és a 27. fejezet 17-es jegyzetét is.
5. H. K. Beecher, *Measurement of Subjective Responses* (Oxford University Press, 1959., 66-67. old.). L. még Berton Roueché, *The Medical Detectives* (New York: Washington Square Press, 1984., II. kötet, 9. fejezet). Újabb keletű hivatkozások pl. K. B. Thomas, „General practice consultations: is there any point in being positive?” *British Medical Journal* vol. 294 (1987), 1200-1202. o. és J. A. Turner és mtsai., „The importance of placebo effects in pain treatment and research”, *Journal of the American Medical Association* vol. 271 (1994) 1609-1614. o.
6. N. D. Grace, H. Muench and T. C. Chalmers, „The present status of shunts for portal hypertension in cirrhosis”, *Gastroenterology*, vol.50 (1966) 684-691. o. Mi J.P. Gilbert, R.J. Light és F. Mosteller „Assessing social innovations: an empirical guide for policy” c. közleményében (*Benefit Cost and Policy Analysis Annual*, 1974) találtuk ezt a példát.
7. H. Sacks, T. C. Chalmers, and H. Smith, „Randomized versus historical controls for clinical trials”, *American Journal of Medicine*, vol. 72 (1982), 233-240. o. A „sorsolt kontrollú vizsgálat” kifejezést nem szigorú értelemben használják. Az eredeti táblázatban véralvadást okozó szívinfarktust követő alkalmazására vonatkozó adatok is szerepeltek. Az alvadást okozó klinikai vizsgálatoknak értelmezéséről még a 80-as években is viták zajlottak. Azóta jelentős változáson mentek át a thrombolitikus terápiák, sok az új kísérlet. Áttekintést adnak:

- J. D. Talley, „Review of thrombolytic intervention for acute myocardial infarction – is it valuable?“ *Journal of the Arkansas Medical Society*, 91 (1994), 70-79. o.
- C. H. Hennekens, „Thrombolytic therapy: pre- and post-GISSI-2, ISIS-3, and GUSTO-1“, *Clinical Cardiology* vol. 17 suppl. I (1994), 115-117. o.
- R. Collins, R. Peto, S. Parish and P. Sleight, „ISIS-3 and GISSI-2: no survival advantage with tissue plasminogen activator over streptokinase, but a significant excess of strokes with tissue plasminogen activator in both trials“, *American Journal of Cardiology* vol. 71 (1993), 1127-1130. o.
- M. J. Stampfer et al., „Effect of intravenous streptokinase on acute myocardial infarction: pooled results from randomized trials“, *New England Journal of Medicine* vol. 307 (1982), 1180-1182. o. *Coronary Artery Disease* vol. 5 no. 4 (1994).
8. T. C. Chalmers, „The impact of controlled trials on the practice of medicine“, *Mount Sinai Journal of Medicine* vol. 41 (1974), 753-759. o.

2. FEJEZET. MEGFIGYELÉSES VIZSGÁLATOK

1. Hivatkozások:

- J. Berkson, „The statistical study of association between smoking and lung cancer“, *Proceedings of the Mayo Clinic* vol. 30 (1955), 319-348. o.
- R. A. Fisher, *Smoking: The Cancer Controversy*. (Oliver and Boyd, 1959)
- J. Cornfield, W. Haenszel, E. C. Hammond, A. M. Lilienfeld, M. B. Shimkin and E. L. Wynder, „Smoking and lung cancer: recent evidence and a discussion of some questions“, *Journal of the National Cancer Institute* vol. 22 (1959), 173-203. old.
- U.S. Public Health Service, *Smoking and Health: Report of the Advisory Committee to the Surgeon General*. (Washington, D.C., 1964).
- International Agency for Research on Cancer, *Tobacco Smoking*. Monograph 38 (Lyon, Franciaország: 1986)
- U.S. Public Health Service, *The Health Benefits of Smoking Cessation. A Report of the Surgeon General*. (Washington, D.C., 1990)

Újabb keletű adatok a hatásmechanizmusról:

- M. F. Denissenko et al., „Preferential formation of Benzo[a]pyrene adducts at lung cancer mutational hotspots in p53“, *Science* vol. 274 (1996), 430-432. o.

2. The Coronary Drug Project Research Group, „Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project“, *New England Journal of Medicine* vol. 303 (1980), 1038-1041. o. A kipróbált többi gyógyszer: ösztrogének (kétfajta dózisban), dextrothyroxin és nikotinsav. A kezelések 1966. márciusától 1969. októberéig folytak. Az ösztrogének és a dextrothyroxin adagolását, ártalmas mellékhatások miatt, félbeszakították. A clofibrate és a nikotinsav lejjebb vitte a vér koleszterinszintjét, de nem csökkentette a mortalitást. Lásd „Clofibrate and niacin in coronary heart disease“, *Journal of the American Medical Association* vol. 231 (1975) 360-381. o. Lásd továbbá *Lancet* (1989. márc. 4.) 473-474. o., mely a nikotinsavnak a nyomonkövetés kései szakaszában mutatózó pozitív hatására utal. További hivatkozások a 29. fejezet 7-es jegyzetében.

Ugyanerről, másik megvilágításban: 40 kockázati tényezőt mértek induláskor. Ezek alapján úgy tűnik, hogy mint csoport, a szedést elhanyagolók rosszabb bőrben voltak a vizsgálat kezdetekor. De a rendszeren szedők és a szedést elhanyagolók közötti különbséget nem lehetett a mért kockázati tényezők alapján számított regresszióval megfogni; az orvosi előírások követése és a kisegítő változók között nem volt elég erős az összefüggés. Ez arra mutat, hogy óvatosnak kell lennünk, amikor megfigyeléses vizsgálatoknál regressziós modellekkel próbáljuk az összemosódást kézben tartani.

3. Az idézet forrása: D. A. Roe, *A Plague of Corn* (Cornell University Press, 1973). Egy másik nagyszerű forrás, sok eredeti közleményt idéz, magyarázatokkal: K. J. Carpenter, *Pellagra* (Academic Press, 1981). Lásd még M. Terris (szerk.) *Goldberger on Pellagra* (Louisiana State University Press, 1964).

A kukorica Amerikából került Európába. Az indiánok főzés előtt lúggal kezelték – ez felszabadította a nikotinsavat; ők nem ismerték a pellagrát. Úgy tűnik, az Egyesült Államokban a pellagra-járvány akkor kezdődött, amikor a molnárok elkezdtek a kukoricából kivonni a csírt; a kukoricában a

nikotinsav, illetve triptofán nagy részét a csíra tartalmazza. (A triptofán egy aminosav, melyet a szervezet nikotinsavvá tud alakítani.) Az Egyesült Államokban a pellagraeredetű halálozások a harmincas években tetőztek; számuk azóta elég egyenletesen csökken, feltehetően a javuló általános gazdasági helyzetnek és táplálkozásnak köszönhetően; a liszt adalékolása túl késői fejlemény ahhoz, hogy jelentős hatása legyen. Afrika és India egyes részein a pellagra ma is népbetegség. A nyolcvanas évek vége felé egy Dél-Afrikában végzett vizsgálat arra utalt, hogy a betegség kialakulásában mycotoxinoknak (gombamérgeknek) is szerepe lehet; eszerint egy kis igazság mégis lehet a fertőzéses elméletekben. A példához nyújtott segítségéért szeretnénk köszönetet mondani K. J. Carpenternek (University of California, Berkeley).

4. Hivatkozások:

E.L. Wynder, J. Cornfield, P.D. Schroff and K.R. Doraiswami, „A study of environmental factors in carcinoma of the cervix“, *American Journal of Obstetrics and Gynecology* vol. 68 (1954), 1016-1052. o.

Újabb keletű bizonyítékok:

J. Cairns, *Cancer: Science and Society* (Salt Lake City: W. F. Freeman & Co., 1978, 59. old.).

R.Peto and H. zur Hausen (szerk.), *Viral Etiology of Cervical Cancer*, Banbury Report no. 21 (1986)

N. Muñoz, F. X. Bosch, K. V. Shah and A. Meheus, (szerk.), *The Epidemiology of Cervical Cancer and Human Papillomavirus*. International Agency for Research on Cancer, Scientific Publication no. 119 (1992).

S. H. Swan and W. L. Brown, „Oral contraceptive use, sexual activity, and cervical carcinoma“, *American Journal of Obstetrics and Gynecology* vol. 139 (1981) 52-57. old.

S. H. Swan and D. B. Petitti, „A review of problems of bias and confounding in epidemiologic studies of cervical neoplasia and oral contraceptive use“, *American Journal of Epidemiology* vol. 115 (1982), 10-18. o.

S. A. Cannistra and J. M. Niloff, „Cancer of the uterine cervix“, *New England Journal of Medicine* vol. 334 (1996), 1030-1038. o.

Egy ideje csökken a méhnyakrák-okozta halálozások aránya; az okok nem ismeretesek. A betegség rizikófaktorai között a dohányzás is szerepel. A példát Michael Kramer epidemiológus professzor (McGill University) javasolta.

5. R. M. Moore et al., „The relationship of birthweight and intrauterine diagnostic ultrasound exposure“, *Journal of Obstetrics and Gynecology* vol. 71 (1988), 513-517. o. A számításba vett összemosó változók: bórszín, magánbeteg-e, dohányzik-e, szülési státusz (rendes, vagy koraszülés), spontán vetéléses előzmények, szeszesital-fogyasztással kapcsolatos előzmények, amniocentesis vizsgálat eredménye, magzatvizsgálat eredménye, szülés módja, iskolai végzettség, hány hetes terhes, amikor jelentkezik, születés előtti jelentkezések száma, az anya testsúlya és súlygyarapodása, hány hetes a magzat a szüléskor.

A klinikai kísérlet: U. Waldenstrom et al., „Effects of routine one-stage ultrasound screening in pregnancy: a randomized controlled trial“, *Lancet* (1988. szept. 10.), 585-88. old. Az ultrahanggal vizsgált újszülötteknek, átlagosan, nagyobb volt a születéskori súlya. A kezelt csoportban az anyák megnézhatték magzatuk ultrahangos képét. Ennek hatására többen felhagytak a dohányzással, márpedig a dohányzás alacsony születéskori súlyt okoz. A magzatvédő hatást feltehetően a dohányzási szokások megváltozása magyarázza.

6. „Suicide and the Samaritans“, *Lancet* (1978. okt. 7) 772-773. o. (szerkesztőségi cikk). Az első kutató C. Bagley (*Social Science and Medicine*, 1968). A gáz fajtája nem szerepelt a városokat illesztő változók között; úgy tűnik, visszamenőleg ezek az adatok nem elérhetőek. A megismételt vizsgálat: B. Barraclough et al. (*Lancet*, 1977; *Psychological Medicine*, 1978). A példát D. C. Hoaglin, R. J. Light, B. McPeck, F. Mosteller és M.A. Stoto, *Data for Decisions* (University Press of America, 1982, 133. o.) könyvében találtuk.
7. A Berkeley adatok paradoxonát Eugene Hammel – a posztgraduális részleg akkori dékánja vette észre. Megoldásában két kollégája, P. Bickel és J. W. O'Connell volt segítségére. Az ő beszámolójukat követjük („Is there a sex bias in graduate admissions?“, *Science* vol. 187 (1975) 398-404. old.). A felvételi adatok 1973 ősziek.

8. Áttekintést ad: Myra Samuels, „Simpson's Paradox and related phenomena”, *Journal of the American Statistical Association* vol.88 (1993), 81-88. old.

9. Néhány jellegzetes példa:

(i) Az összemosó változó okozza a külső hatást is, a betegséget is.

(ii) Az összemosó változó okozza a betegséget és összefügg a külső hatással.

(iii) Az összemosó változó okozza a külső hatást, és összefügg a betegséggel.

Egy olyan dolog, mely közös következménye a külső hatásnak és a betegségnek, általában nem magyarázat az együttjárásra. És paradox módon negatív összefüggés keletkezhet, ha egy ilyen együttes hatás negatív szelekciót okoz: lásd az 1. jegyzetben hivatkozott Berkson-közleményt.

10. *Statistical Abstract*, 1988., 117. táblázat; 1993., 118. táblázat.

11. Lásd az 1. fejezet 2. jegyzetét.

12. Hivatkozások:

L. M. Friedman, C. D. Furberg and D. L. DeMets, *Fundamentals of Clinical Trials*, 2nd ed. (Littleton, Mass., PSG Publishers, 1985).

T. L. Lewis, T. R. Karlowski, A. Z. Kapikian, J. M. Lynch, G. W. Shaffer, D. A. George and T. C. Chalmers, „A controlled clinical trial of ascorbic acid for the common cold”, *Annals of the New York Academy of Science* vol. 258 (1975), 505-512. o.

T.R. Karlowski, T. C. Chalmers, L. D. Frenkel, A. Z. Kapikian, T. L. Lewis and J. M. Lynch, „Ascorbic acid for the common cold”, *Journal of the American Medical Association* vol. 231 (1975), 1038-1042. o.

K. J. Carpenter, *The History of Scurvy and Vitamin C* (Cambridge University Press, 1986.)

13. Nikotinsav - ez a niacinnak, azaz a pellagra megelőzésére alkalmas anyagnak a tudományos elnevezése. A niacin elnevezést feltehetően azért vezették be, mert a „nikotinsav” felírat nagyon rosszul mutatott volna a lisztes zacskókon. A nikotinsavat is kipróbálták a Coronary Drug Project során, hatásatlannak találták. A feladatbeli táblázat fiktív.

14. Az életeket mentő hatás hosszú éveken át tart, és ezt más vizsgálatok is megerősítik. A szűrés néhány hónappal korábbra hozza a felismerést, és ez, úgy látszik, már számít. Publikálatlan adatok Sam Shapiro (az epidemiológia emeritus professzora, Johns Hopkins University) szíves közlése alapján. A HIP vizsgálat egy kezdeti szűrésből, majd három éven át évi egy szűrésből állt (egy-egy szűrés pedig egy orvos által végzett emlővizsgálatból és egy mammográfiából).

Az emlőrák kockázatát befolyásolja a hormonegyensúly - a terhességnek védő hatása van; a korai első terhességnek kifejezett védő hatása van. Valószínűleg ez magyarázza a jövedelmi szint szerinti eltéréseket.

Hivatkozások:

J. Cairns, „The treatment of diseases and the war against cancer”, *Scientific American* vol. 253 (1985) 51-59. o.

R. Doll és R. Peto, *The Causes of Cancer* (Oxford University Press, 1981).

S. Shapiro, „Evidence on screening for breast cancer from a randomized trial”, *Cancer* vol. 39 (1977) 2772-2782. o.

S. Shapiro et al., „Current results of the breast cancer screening randomized trial”, in N. E. Day and A. B. Miller (ed.), *Screening for Breast Cancer* (H. Huber, 1988).

L. Tabar et al., „Reduction in mortality from breast cancer after mass screening with mammography”, *Lancet* (1985. ápr. 13.) 829-832. o.

Az 1990-es években szakirodalmi áttekintések alátámasztották a mammográfia hasznát 50 évesnél idősebb nőknél, kérdésesnek találták azonban fiatalabbaknál. Hivatkozások:

H. Kattlove et al., „Benefits and costs of screening and treatment for early breast cancer”, *Journal of the American Medical Association* vol.273 (1995) 142-148. o.

K. Kerlikowske et al., „Efficacy of screening mammography”, *Journal of the American Medical Association* vol. 273 (1995) 149-154. o.

Más szakértők szerint a mammográfia éppúgy hasznos fiatal nők számára, mint idősebbek számára:

D. B. Kopans et al., „Statistical power in breast cancer screening trials and mortality reduction among women 40-49 years of age with particular emphasis on the National Breast Cancer Screening Study of Canada”, *Cancer* vol. 74 (1994), 1196-1216. o., bővebben kifejtvé.

Néhány szakértő továbbra sincsen meggyőződve a mammográfia hasznáról, még idősebb nők esetében sem:

- C. J. Wright és C. B. Mueller, „Screening mammography and public health policy – the need for perspective“, *Lancet* vol. 346 (1995. július 1) 29-32. old; levél, vol. 346 (1995. szeptember 23), 852. o.

A HIP-vizsgálat érdekes bírálata:

- D. Plotkin, „Good news and bad news about breast cancer“, *Atlantic Monthly* vol. 277 (1996. június), 53. skk.

A betegség kockázata különböző társadalmi csoportokban – további hivatkozásokkal:

- J. N. Morris et al., „Levels of mortality, education, and social conditions in the 107 local education authority areas of England“, *Journal of Epidemiology and Community Health* vol. 50 (1996), 15-17. o.

- J. Pekkanen et al., „Social class, health behavior, and mortality among men and women in eastern Finland“, *British Medical Journal* vol. 311 (1995), 589-593. o.

15. Hivatkozások: lásd a 4. jegyzetet.

16. A példát Dr. Shanna Swan (Department of Health Services, State of California) javasolta egy olyan megfigyeléses vizsgálat adataira alapozva, melyet a kaliforniai Walnut Creek-ban, a Kaiser Permanente-nél végeztek.

17. *Statistical Abstract*, 1993: 1056., 1063. és 1066. táblázat. Az esetszámok kicsik, de a tendencia hosszabb ideje meglehetősen következetes.

18. *Federal Register*, vol.59, no.151, 1994. aug.8., 40409-40414. o.

19. *Statistical Abstract*, 1971, 118. táblázat. A vizsgálat 1964-es; ugyanez a hatás sok más vizsgálatban is felbukkan. Aki leszokik a dohányzásról és ezt túléli néhány évvel, annak egészségügyi kockázatai ezután a továbbdohányzókra vonatkozó kockázatoknál kisebbek lesznek. Lásd U. S. Public Health Service, *The Health Benefits of Smoking Cessation. A Report of the Surgeon General* (Washington, D.C., 1990).

20. Hivatkozások:

- L. M. Friedman, C. D. Furberg and D. L. DeMets, *Fundamentals of Clinical Trials*, 2. kiadás (PSG Publishers, 1985.)

- P. J. Schechter, W. T. Friedewald, D. A. Bronzert, M. S. Raff and R. I. Henkin, „Idiopathic hypoguesia: a description of the syndrome and a single-blind study with zink sulfate“, *International Review of Neurobiology* (1972), Supplement 1., 125-139. o.

- R. I. Henkin, P. J. Schechter, W. T. Friedewald, D. L. DeMets and M. S. Raff, „A double blind study of the effects of zink sulfate on taste and smell dysfunction“, *American Journal of the Medical Sciences* vol. 272 (1976), 285-299. o.

21. A példát Dr. Shanna Swan javasolta. Lásd: E. Peritz et al., „The incidence of cervical cancer and duration of oral contraceptive use“, *American Journal of Epidemiology* vol. 106 (1977), 462-69. old. További kiküszöbölt összemosó változók voltak: vallás, dohányzás (a méhnyakrák egyik rizikófaktora), Papanicolau-kenetek száma a vizsgálatba kerülés előtt, és „bizonyos fertőzések“. További hivatkozások: lásd 4. jegyzet.

22. Idézi Herb Caen, *San Fransisco Chronicle* 1995. augusztus 9., szerda.

23. Hivatkozások:

- E. R. Greenberg et al, „A clinical trial of antioxidant vitamins to prevent colorectal adenoma“, *New England Journal of Medicine* vol. 331 (1994) 141-147. o.

- O. P. Heinonen et al., „Effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers“, *New England Journal of Medicine* vol. 330 (1994), 1029-1035. o.

Kommentárok találhatóak: *New England Journal of Medicine* vol. 331 (1994), 611-614. o. Két további kísérlet is nemleges volt:

- C. H. Hennekens et al., „Lack of effect of long-term supplementation with beta-carotene on the incidence of malignant neoplasms and cardiovascular disease“, *New England Journal of Medicine* vol. 334 (1996), 1145-1149. o.

- G. S. Omenn et al., „Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease“, *New England Journal of Medicine* vol. 334 (1996), 1150-1155. o.

24. S. L. Johnson and L. L. Birch, „Parents’ and children’s adiposity and eating style“, *Pediatrics* vol. 94 (1994), 653-61. old., „A gyermekük táplálékfelvételét inkább korlátozó anyák gyermekei kevésbé voltak képesek energiabevitelük szabályozására ($r = -0,76$; $P < 0,0001$).“
25. A feladat alapjául a *San Francisco Chronicle* 1993. január 19.-i száma szolgált. Az idézetet módosítottuk: egyszerűsítettünk a vizsgálati terven. Az elítélteknek, az önkéntes jelentkezést ösztönzendő, általában felajánlják a feltételes szabadlábra helyezés korábbra hozatalát.

2. rész. Leíró statisztika

3. FEJEZET. A HISZTOGRAM

1. Antoine de Saint-Exupéry, Rónay György fordítása. A kötetben a Harcourt Brace Jovanovich, Inc. kiadó engedélyével szerepel.
2. *Money Income in 1973 of Families and Persons in the United States*, Current Population Reports, Series P-60, No. 97 (January, 1975). U.S. Department of Commerce.
3. Az 1973-as adatok forrását lásd a 2. jegyzetben. Az 1992-es adatok a Bureau of the Census CD-ROM-járól származnak, melyet a U.C. Survey Research Center bocsátott a rendelkezésünkre. Lásd még: *Money Income of Households, Families and Persons in the United States: 1992*, Current Population Reports, Series P-60, No. 184, U.S. Department of Commerce (September, 1993). Sok esetben a kérdőívben szereplő kérdés megfogalmazása miatt kell eltérő hosszúságú osztásközökkel dolgoznunk. A megkérdezettektől például csak azt kérdezték meg, hogy a felsoroltak közül melyik sávba esik a jövedelmük; ezek a sávok - osztásközök - jellemzően különböző hosszúságúak.
4. Lásd a 2. jegyzetet
5. A 3. fejezet adathisztogramjai - többek közt - az elméleti hisztogramok tárgyalását készítik elő a 18. fejezet számára. Utóbbiakat sokszor sűrűségfüggvényekkel közelítjük; például gyakori, hogy egy hisztogram a „normálgörbét követi” abban az értelemben, hogy a görbe jól közelíti a hisztogramot. Ha az f görbe egy sűrűségfüggvényt ábrázol, az f alatti terület mutatja a valószínűséget; f -nek az x pontban felvett értéke az egységnyi hosszúságra eső valószínűséget jelenti, nem pedig x valószínűségét: x valószínűsége valójában 0, nem pedig $f(x)$. Általában a matematikusok nem törődnek a fizikai egységekkel, centiméterekkel, hüvelykekkel, kilogrammokkal, dollárokkal. A statisztikusoknak azonban számítanak. Ezért adjuk meg a jövedelemhisztogram sűrűségskáláját „százalék per ezer dollárban” vagy a súlyhisztogramot „százalék per kilogrammban”.
6. Ez az osztásközökre pontos érték, más intervallumokra közelítés.
7. *Statistical Abstract*, 1971, 118. táblázat
8. Sok változó mindkettőbe besorolható, attól függően, hogyan nézzük. A jövedelmek például legfeljebb egy pennyvel (vagy forinttal) térhetnek el egymástól. Mindazonáltal célszerű a jövedelmet folytonos változóként kezelni, mivel terjedelme a minimális különbségnél jóval nagyobb.
9. Túl keskeny osztásközökkel dolgozva a hisztogram annyira csipkézett lehet, hogy lehetetlen felismerni az alakzatot. Szélesebb osztásközöknél könnyebb átlátni a hisztogram alakját, még ha bizonyos információk el is vesznek. Ennek részletes tárgyalását lásd P. Diaconis és D. Freedman: „On the histogram as a density estimator: L_2 theory.” *Z. Wahrscheinlichkeitstheorie* vol. 57 (1981) 453-476. o.
10. Lásd I. R. Fisch, S. H. Freedman, A. V. Myatt: „Oral contraceptives, pregnancy, and blood pressure”, *Journal of the American Medical Association* vol. 222 (1972) 1507-1510. o. Gondolatmenetünk a cikkben foglaltakat követi. Köszönettel tartozunk Dr. Shanna Swan-nek (California Department of Health Services) és Dr. Michael Grossman-nek gyakorlati tanácsaikért.

Vérnyomásméréskor két értéket, *szisztolés* és *diasztolés* vérnyomást mérnek. Mi a szisztolés értékeket vizsgáljuk. A diasztolés értékre vonatkozó eredmények nagyon hasonlóak ehhez. A Contraceptive Drug Study keretében géppel történt a vérnyomásmérés. A vizsgálatból kizártak 3500 nőt, akik terhesek voltak, nemrég szültek vagy más hormonális készítményt szedtek; ezek a tényezők ugyanis befolyásolják a vérnyomást. A vizsgálatban elegendőnek mutatkozott négy életkori csoporttal dolgozni: 17-24, 25-34, 35-44 és 45-58. Az egyes korcsoportokon belül hasonlóan alakult a tablettaszedők és a nem szedők életkori megoszlása.

11. R. C. Tryon: „Genetic differences in maze-learning techniques in rats”, 39th yearbook, *National Society for the Study of Education* part I (1940) 111-119. o. A cikk megjelent egy kiváló szöveggyűjteményben is: Anne Anastasi: *Individual Differences* (John Wiley & Sons, 1965). Tryon nemlineáris skálát alkalmazott a hisztogramoknál, így azok a mi vázlatainktól eltérően néznek ki.
12. *1970 Census of Population*. Lásd vol. 1, part 1, section 2, appendix, 14. o. U. S. Department of Commerce. Az 1880-ra vonatkozó oszlopban a 23-99 éves, az 1970-re vonatkozóban a 23-82 éves személyek szerepelnek.
13. K. Bemesderfer - J. May: *Social and Political Inquiry* (Belmont, California: Duxbury Press, 1972, 6. o.).
14. Hivatkozások:
 - R. A. Baron - V. M. Ransberger: „Ambient temperature and the occurrence of collective violence: the 'long, hot summer' revisited” *Journal of Personality and Social Psychology* vol. 36 (1978) 351-360. o. A cikkben foglaltakat csak felszínesen érintjük.
 - J. M. Carlsmith - C. A. Anderson: „Ambient temperature and the occurrence of collective violence: a new analysis” *Journal of Personality and Social Psychology* vol. 37 (1979) 337-344. o.
 Az ábrát Baron és Ransberger eredeti cikke alapján készítettük, melynek szerzői jogtulajdonosa az American Psychological Association. Közléséhez a jogtulajdonos és a szerző hozzájárultak.

4. FEJEZET. AZ ÁTLAG ÉS A SZÓRÁS

1. *Natural Inheritance* (London: MacMillan, 1889; újabb kiadásban megjelentette az American Mathematical Society Press, 1973).
2. Az adatsor közepének megragadására használatos olykor a *módusz* is: az a pont, ahol a legmagasabb a hisztogram. Ezt azonban nem javasoljuk, mivel a móduszt az adatok kis változása is jelentősen befolyásolhatja.
3. Tom Alexander: „A revolution called plate tectonics”, *Smithsonian Magazine* vol. 5, no. 10 (1975). A. Hallam: „Alfred Wegener and the hypothesis of continental drift”, *Scientific American* vol. 232, no. 2 (1975). Ursula Marvin: *Continental Drift* (Smithsonian Press, 1973).
4. Az egyik lehetséges forrás: *Blood Lead Levels for Persons Ages 6 Months-74 Years: United States, 1976-1980*. Data from the National Health Survey, series 11, no. 233, U.S. Department of Health and Human Services, Washington, D.C. Vannak bizonyos viták az alacsony ólomszintnek a gyermekek fejlődésére gyakorolt hatásáról: lásd a 12. fejezet 14. lábjegyzetét.

A Public Health Service (Közegészségügyi Hivatal) és a National Center for Health Statistics (Országos Egészségstatisztikai Központ) az Egészségügyi és Népjóléti Minisztérium (Department of Health and Human Services) intézményei. Az adatok a „Vital and Health Statistics” 11. sorozatából származnak, valamint az Egészségstatisztikai Központ és az Inter-University Consortium for Political and Social Research által rendelkezésünkre bocsátott adatszalogokról. Az adatok értelmezéséért minden felelősség bennünket terhel, legyenek azok jók vagy rosszak. Az első kiadáshoz nyújtott segítségéért köszönettel tartozunk Mr. Arthur J. McDowellnek, aki a Health Examination Statistics részleg vezetője volt 1976-ban. A második kiadáshoz nyújtott segítségükért Dorothy Rice professzort (UCSF) és Dale Hitchcockot (NCHS) illeti köszönet. A harmadik kiadásnál Bob Murphy (NCHS) segített bennünket. A HANES III az 1988 és 1994 közötti időszakban készült; az adatokat 1997-ben bocsátották közhasználatra.

A 4., 8., 9. ábrákon szereplő hisztogramok a minta súlyozatlan esetszámain alapulnak; ugyanígy az összegző statisztikák is, melyeket kerekítettünk. A súlyozás csak kis különbséget okoz. A pontos értékek (átlag ± szórással) a következők:

	18-74 éves férfiak súlyozatlan adatok	18-74 éves férfiak súlyozott adatok	18-74 éves nők súlyozatlan adatok	18-74 éves nők súlyozott adatok
Testmagasság (hüvelykben)	68,78 ± 2,83	69,11 ± 2,82	63,46 ± 2,62	63,71 ± 2,60
Testsúly (fontban)	170,92 ± 30,13	172,19 ± 29,75	145,71 ± 32,65	144,18 ± 32,27
Szisztolés vérnyomás	131,55 ± 19,29	128,91 ± 17,70	127,37 ± 23,42	122,81 ± 21,22
Diasztolés vérnyomás	81,18 ± 11,53	80,81 ± 11,31	77,66 ± 11,91	76,58 ± 11,60

5. A 3. ábránál használt életkori csoportok: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74. Lásd: *Anthropometric Reference Data and Prevalence of Overweight: United States, 1976-1980*. Data from the National Health Survey, series 11, no. 238, U.S. Department of Health and Human Services, Washington, D.C. Az akcelerációs tendenciát évtizedenként 1 cm-re (0,4 hüvelyk) becsülik; az 1960-80 közötti húsz éves időszakban az amerikai népesség átlagosan 2 cm-rel (0,8 hüvelykkel) lett magasabb. Azonkívül úgy tűnik, hogy 50 és 75 éves kor között 0,5-1,5 hüvelyknyit veszít az ember a magasságából. (Az egyik lehetséges magyarázat: a testmagasságból körülbelül 2 hüvelyket a testben lévő csontok közötti légrések tesznek ki; az életkorral a test kissé „leülepszik”, azaz a légrések egyre kisebbek lesznek.) Az akcelerációs tendencia és az „összemenés” együtt összesen 2,3-3,5 hüvelyk csökkenést jelentenek 20 és 70 év között. A megfigyelt különbség 2,3 hüvelyk a férfiak, 2,1 hüvelyk a nők esetében, tehát további tényezőnek is szerepet kell játszania. Szeretnénk köszönetet mondani Dr. Reubin Andresnek (NIH) és Dr. Stanley Garn-nek (University of Michigan) a segítségükért. A testmagasságtrendeknek a társadalmi változások indikátoraként való tárgyalását lásd: R. Floud, K. Wachter, A. Gregory: *Height, Health, and History* (Cambridge University Press, 1991)
6. Ez pontosan igaz, ha egész számokkal dolgozunk és az intervallumok közepét ezekre tesszük. Egyéb esetben csak közelítés.
7. Lásd a 3. fejezet 3. jegyzetét.
8. Az alapvető érv az, amit a statisztikusok *ortogonalitás*nak neveznek. Ha több, egymástól független forrásból származik a hiba, a teljes hiba négyzetes közepét egzakt és egyszerű képlettel kaphatjuk meg: a hibák négyzetes közepe úgy összegződik, mint a derékszögű háromszög oldalai. Két ortogonális hibaforrás esetén:

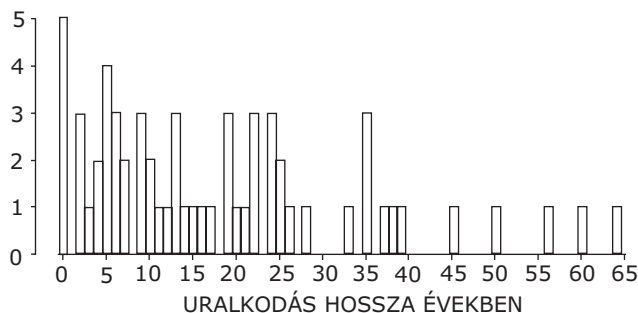
$$c = \sqrt{a^2 + b^2}$$

ahol a az egyik forrásból, b a másik forrásból származó hiba négyzetes közepe, c pedig a teljes hibáé. Ezt a tényt többször is felhasználjuk majd a könyvben: a regressziónál (III. rész), összeg standard hibájának kiszámításánál (V. rész), és a különbség standard hibájának meghatározásánál (VIII. rész). Az abszolút értékek átlagára nem létezik ilyen képlet.

9. A számokat kerekítettük. Az összegző statisztikák pontos értékei a 4. jegyzetben szerepelnek. A sávokon belüli esetek pontos aránya a magasságadatoknál (hüvelykben):

	átlag ± 1 szórás	átlag ± 2 szórás	átlag ± 3 szórás
férfiak	68,1%	95,1%	99,6%
nők	68,3%	95,2%	99,6%

A 68%-95% szabály meglehetősen jól működik sok olyan adatsor esetében is, amelyek nem követik a normálgörbét. Vegyük például 61 angol uralkodó uralkodásának hosszát VI. Györggyel bezárólag. Az átlag 18,1 év, a szórás 15,5 év. Alább láthatjuk a hisztogramot, mely cseppet sem hasonlít a normálgörbére. Mégis 69% (61 közül 42) egy szóráson belül van az átlagtól. És a 61-ből 57, azaz 93% esett a szórás kétszeresén belül. (Az uralkodás hosszát az 1988-as *Information Please Almanac* 274-275. oldalán szereplő első és utolsó év közti különbségként definiáltuk; a példa David Lane-től származik (Modenai Egyetem Politikai Gazdaságtan Tanszék, Olaszország).



10. A szórás négyzetét *varianciának* nevezzük. Sokszor használjuk a szóródás mérésére, de leíró statisztikaként nem ajánljuk. Például: az amerikai férfiak testsúlyának szórása 30 font (13,5 kg) körül van – az egyes emberek testsúlya durván 30 fonttal tér el az átlagostól. A testsúly varianciája viszont

$$(30 \text{ font})^2 = 900 \text{ négyzetfont (182,25 négyzetkilogramm)}$$

11. Ez a képlet azonban érzékeny a kerekítési hibákra.
 12. Az adatok a National Center for Health Statistics és az Inter-University Consortium for Political and Social Research által rendelkezésünkre bocsátott mágnesszalagról származnak.
 13. Patricia Ruggles, *Drawing the Line* (Urban Institute Press, Washington, D. C., 1990). Az alsó osztály leírását a könyv 105. oldaláról vettük át. A könyv tárgyalja az alkalmazott definíciók hatását is; lásd erről még az 5. fejezetben az idődimenzióról szóló részt. Újabb adatok a Survey of Income and Program Participation (SIPP) alapján állnak rendelkezésre. Lásd: *Poverty, 1990 to 1992. Current Population Reports P70-42*, Bureau of the Census, Washington, D. C. (1995).

5. FEJEZET. ADATOK NORMÁLIS KÖZELÍTÉSE

- Standardizált értéknek, z-score-nak, szigma score-nak is nevezik.
- A „becslés”, „megbecsülni” szavakat itt köznapi értelemben használjuk: „határozzuk meg közelítőleg”. A „becslés” emellett szakkifejezés is a statisztikában, melyet a VI. részben tárgyalunk majd.
- Lásd a 3. fejezet 3. jegyzetét. A családi jövedelmek átlagát az 1993-as felmérés 14. táblája közli. A hisztogramot, a szórást és a percentiliseket a CD-ROM-ról számoltuk ki. A közvetlen családokra (családegységekre) vonatkozó súlyozott adatokkal dolgoztunk; az átlag 43 400\$, a szórás 32 500\$ volt.
- A SAT pontszámok csökkenésének tárgyalását lásd: Willard Wirtz et al.: *On Further Examination. Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline* (New York: College Entrance Examination Board, 1977). Az adatok forrása a College Boardnak a sajtó számára készült tájékoztatója, 1988. szeptember 20; valamint az *1994 Profile of SAT and Achievement Test Takers* című kiadvány (Educational Testing Service, Princeton, N.J.). Lásd még: *Statistical Abstract*, 1993, 265. táblázat. A táblázatban azon végzős középiskolások szerepelnek, akik középiskolai tanulmányaik során bármikor is kitöltötték a SAT-ot. A szórás valójában nőtt valamelyest; 1975-ben 109, 1994-ben 113 volt a nyelvi SAT szórása. 1994-ben újristandardizálták a tesztet úgy, hogy az átlag 500, a szórás 100 legyen; a változtatás 1995-ben lépett érvénybe. Köszönettel tartozunk Susan Bryce-nak (ETS), Paul Holland-nek (Berkeley) és Howard Wainernek (ETS) a harmadik kiadáshoz nyújtott segítségükért.
 A matematikai SAT (4. feladat) és sok más hasonló teszt esetében is nagyobb a szórás a férfiak körében – 120 a 110-zel szemben; ez az egyik ok, amiért több férfi található a megoszlás szélein. Lásd L. V. Hedges és A. Nowell: „Sex differences in mental test scores, variability, and numbers of high-scoring individuals,” *Science* vol. 269 (1995) 41-45. o.
- A rendszeres népszámlálás 1993 márciusi adatai; az adatok a Bureau of the Census által, a U.C. Survey Research Center közvetítésével rendelkezésünkre bocsátott CD-ROM-ról származnak. Egy személy összes jövedelme különböző tételek összegeként áll elő; minden tételhez meg van határozva a lehetséges maximum, legfeljebb 100 000\$ (a társadalombiztosítástól kapott összeg például legfeljebb 30 000\$ lehet); a súlyozott adatokkal dolgoztunk.

6. FEJEZET. A MÉRÉSI HIBA

- A National Bureau of Standards (NBS) elnevezése NIST-re változott: The National Institute of Standards and Technology. Dr. H. H. Kunak szeretnénk köszönetet mondani a Hivatal részéről nyújtott segítségért.
- A köznapi súly szót használjuk a *tömeg* szakkifejezés helyett. A nemzetközi mérésügyi konferencia 1983-ban megváltoztatta a méter definícióját, de a súlyt továbbra is a referencia-tárggyal definiáljuk. A „kilogramm”-ról szóló legfrissebb közlést lásd: *Science* (1995. május 12.) 804. o.

3. A Hivatalban végzett precíziós méréseknél a véletlen hiba két főbb feltételezett forrása:
 - a mérleg-szerkezet – főként az ék élének – kismérvű holtjátéka,
 - árnyalatnyi eltérések a súlyok helyzetében a mérlegserpenyőkben.
4. P. E. Pontius, „Measurement phylosophy of the pilot program for mass calibration,” *NBS Technical Note No. 288* (1966). A Mérésügyi Hivatal magányos eseteket csak „valamilyen okból, például ajtócsapódás vagy a berendezés hibás működése miatt” vet el. Lásd még H. H. Ku (szerk.) *Precision Measurement and Calibration*. NBS Special Publication no. 300, vol. 1 (Washington, D. C., 1969).
5. Az adatok a National Center for Health Statistics és az Inter-University Consortium for Political and Social Research által rendelkezésünkre bocsátott mágnesszalagról származnak.
6. Az adatok a Bureau of the Census által, a U.C. Survey Research Center közvetítésével rendelkezésünkre bocsátott CD-ROM-ról származnak. Az elemi családokra (családjegységekre) vonatkozó adatok.
7. J. N. Morris-J. A. Hardy, „Physique of London busmen,” *Lancet* (1956) 569-570. o. Eric Peritz, a jeruzsálemi Héber Egyetem Sziszatika Tanszékének professzora hivatkozott a vizsgálatra.
8. *Chance News*, Dartmouth (February, 1995). Lásd Kopans et al., a 2. fejezet 14. lábjegyzetében hivatkozott cikkét; lásd továbbá: *Science* vol. 275 (1997) 1056-1059. o. A kutatás statisztikai bizonyító erejének növelése érdekében láthatólag véletlenszerűen soroltak be olyan nőket is, akiknek orvosiilag felismerhető rákjuk volt – ami másfajta kérdéseket vet fel a kutatási elrendezéssel kapcsolatban. Nem találtak jótékony hatást az 50 éven aluli nők körében, amit Kopans erősen vitat.
9. A megfigyeléseket tárgyalja: H. Zeisel, H. Kalven, Jr., B. Buchholz, *Delay in the Court*, 4th ed. (Little, Brown & Co., 1959). A kísérletről beszámol: Maurice Rosenberg, *The Pretrial Conference and Effective Justice* (Columbia University Press, 1964). A definíciók a 19. oldalon, a számok a 20. oldalon, a 8. és 9. Táblázat adatai a 48., illetve 52. oldalon szerepelnek. A kérdésről lásd még: H. Zeisel, *Say It with Figures*, 6th ed. (Harper & Row, 1985, 141. o.).

3. rész. Korreláció- és regressziószámítás

8. FEJEZET. A KORRELÁCIÓ

1. Kettőnél több változó kezelésére is vannak módszerek, csak kissé bonyolultabbak. Szükség van némi mátrixalgebrára tárgyalásuk követéséhez. Alapvető könyvek e tárgyban:
 - M. L. Eaton, *Multivariate Statistics: A Vector Space Approach* (John Wiley & Sons, 1983).
 - C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. (John Wiley & Sons, 1973).
 - J. A. Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. (Duxbury Press, 1955).
 - H. Scheffé, *The Analysis of Variance* (John Wiley & Sons, 1961).
 - G. A. F. Seber, *Linear Regression Analysis* (John Wiley & Sons, 1977).
 A könyv 12. fejezetének 3. szakaszában röviden szót ejtünk majd a többváltozós regressziószámításról is.
2. K. Pearson, A. Lee, „On the laws of inheritance in man,” *Biometrika* vol. II. (1903) 357-462. o. Hüvelykre kerekítve közzétették a testmagasságok együttes megoszlását. Hogy folytonos adatokat kapjunk, tizedes értékeket rendeltünk véletlenszerűen az adatokhoz. Az itt szereplő ábrák kis mértékben eltérnek a könyv első kiadásában megjelent változattól, minthogy a most használt adatok az előzőtől független randomizálással készültek.
3. Az „átlagpont” nem bevett szakkifejezés.
4. H. N. Newman, F. N. Freeman, K. J. Holzinger, *Twins: A Study of Heredity and Environment* (University of Chicago Press, 1937). Ikervizsgálatoknál az ikerpárokat kétszer szokás feltüntetni; egyszer (x,y) -ként, egyszer pedig (y,x) -ként.
5. A rendszeres népszámlálás 1993 márciusi adatai; a korrelációkat a Bureau of the Census által, a U. C. Survey Research Center közvetítésével rendelkezésünkre bocsátott CD-ROM alapján számítottuk ki, a súlyozatlan adatok felhasználásával. Az adatok az 1992. évi jövedelemre vonatkoznak. A befejezett iskoláévek számát az iskolai végzettség csoportosított adataiból kaptuk. 1992-től kezdődően kategóriánként közlik az iskolai végzettséget, mint például „1-4 osztály”, „befejezetlen főiskola”. Lásd *Monthly Labor Review*, 1993. szeptember, 34-38. o.

8. Lásd az 5. jegyzetet. A „gyerekszám” a saját, 18 éven aluli nőtlen, illetve hajadon gyerekek számát jelenti. Az anya életkorával 0,05 körüli a korreláció. Az idősebb nők kevesebb osztályt végeztek; viszont a gyerekeik is idősebbek, tehát kevesebb 18 éven aluli gyermekük van. Ez pozitív összefüggést eredményez, mely ellensúlyozza a negatív kapcsolatot. (Az adatok a 25 éven felüli anyákra vonatkoznak; a korreláció bizonyos mértékig függ a választott életkori határoktól.)
9. Ha a korreláció 0, bármelyik meredekség használható. A „szórás egyenes” nem bevett szakkifejezés.
8. Ez a képlet azonban érzékeny a kerekítési hibákra.
9. *Consumption Patterns of Household Vehicles 1985*. Residential Transportation Energy Consumption Survey. Energy Information Administration, Washington, D.C.
10. Dr. Marjorie Honzik (Institute of Human Development, Berkeley) volt kedves rendelkezésünkre bocsátani az adatokat.

9. FEJEZET. KICSIT BŐVEBBEN A KORRELÁCIÓRÓL

1. Az adatokat Sam Cohen és Doug Hale (Energy Information Administration, Department of Energy) voltak szívesek rendelkezésünkre bocsátani. New Yorkban a Central Parkban, Bostonban a Logan repülőtéren működő mérőállomáson mérték ezeket az adatokat.
2. R. Doll, „Etiology of lung cancer,” *Advances in Cancer Research* vol. 3 (1955) 1-50. o. Report of the U.S. Surgeon General, *Smoking and Health* (Washington, D.C., 1964.)
3. Az ötlet forrása: W.S. Robinson, „Ecological correlations and the behavior of individuals,” *American Sociological Review* vol. 15 (1950) 351-357. o. Robinsonnál az írni - olvasni tudás és a borszín szerepel az 1930-as népszámlálás adatai alapján. Példánk ennek megismétlése; az adatok forrását lásd a 8. fejezet 5. lábjegyzetében.

Amennyiben minden egyes csoport kétváltozós normális eloszlást követ és közös a regressziós egyenes, akkor az átlagokból becsülni lehet a meredekséget és a tengelymetszetet. Lásd még L. Goodman, „Ecological regression and the behavior of individuals,” *American Sociological Review* vol. 18 (1953) 663-664. o. A kérdés további tárgyalását lásd: S. Klein and D. Freedman, „Ecological regression in voting rights cases,” *Chance* vol. 6 (1993) 38-43. o.

4. Magyar kiadása: E. Durkheim, *Az öngyilkosság* (Közgazdasági és Jogi Könyvkiadó, 1967, 153. o.). Ezt a korrelációt mi számoltuk ki. Durkheim tartományok csoportjaira nézte meg az átlagokat, melyekre 0,9 volt a korreláció. Arra a következtetésre jutott, hogy „a népoktatás és az öngyilkosság számai pontosan azonos módon oszlanak el”.
5. Valamelyes segítséget kínál a többváltozós regressziószámítás, ám sokszor több kérdést vet fel, mint amennyire választ ad (lásd a 12. fejezet 3. szakaszát).
6. A kérdés bővebb tárgyalását lásd: H. Zeisel, *Say It With Figures*, 6th ed. (Harper & Row, 1985, 152ff. o.)
7. M. Russel bocsátotta rendelkezésünkre a tanulmány 1. táblázatának adatait: D. Jablonski, „Larval ecology and macroevolution in marine invertebrates,” *Bulletin of Marine Science* vol. 39 part 2 (1986) 565-587. o. Lásd még: *Science* vol. 240 (1988) 969. o.
8. Hivatkozások:

R. Doll and R. Peto, *The Causes of Cancer* (Oxford University Press, 1981).

B. E. Henderson, R. K. Ross and M. C. Pike, „Toward the primary prevention of cancer,” *Science* vol. 254 (1991) 1131-1138. o.

B. N. Ames, L. S. Gold and W. C. Willett, „The causes and prevention of cancer,” *Proceedings of the National Academy of Science USA* vol. 92 (1995) 5258-5265. o.

B. S. Hulka and A. T. Stark, „Breast cancer: cause and prevention,” *Lancet* vol. 346 (1995. szeptember 30.) 883-887. o.

Komoly epidemiológiai – valamint állatkísérletek alapján nyert – bizonyítékok szólnak amellett, hogy a túlzott táplálékbevitel karcinogén hatású. A zsiradékok hatása (azonos kalóriatartalmú étkezés esetén) kevésbé egyértelmű. Lásd A. Schatzkin et al., „Serum cholesterol and cancer in the NHANES I epidemiologic followup study,” *Lancet* ii (1987) 298-301. o.; ez a vizsgálat védőhatást tulajdonít a zsiradékoknak. Egy másik vizsgálat az ökológiai elemzést támasztja alá: W. C. Willett et al., „Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among

- women,” *New England Journal of Medicine*, 1990. december 13., 1664-1671. o. Ezzel szemben lásd D. Hunter et al., „Cohort studies of fat intake and the risk of breast cancer – a pooled analysis,” *New England Journal of Medicine*, vol. 334 (1996) 356-361. o. A 8. ábra életkor szerint kontrollálva készült, ám a gyermekek száma is befolyásoló tényező lehet (lásd a 2. fejezet 14. jegyzetét). Az '50-es és '60-as évek táplálkozási szokásai is kérdéseket vetnek fel az ábrával kapcsolatban.
9. National Assessment of Educational Progress, *The Reading Report Card* (Princeton: ETS/NAEP, 1985, 53. o.). Negatív a korreláció a standard tudásteszték pontszámaival is. Lásd Lee R. Jones et al., *The 1990 Science Report Card: NAEP's Assessment of 4th, 8th, and 12th Graders*. (Washington, D.C., U.S. Department of Education, Office of Educational Research and Improvement, 1992.)
 10. T. R. Dawber et al., „Coffee and cardiovascular disease: observations from the Framingham study,” *New England Journal of Medicine* vol. 291 (1974) 871-874. o.
 11. M. P. Rogin and J. L. Shover, *Political Change in California* (Westport, Connecticut: Greenwood Press, 1970, xvii. o.).
 12. Lásd a 8. fejezet 5. jegyzetét.
 13. M. és B. Rodin vizsgálatát idéztük, lásd: „Student evaluation of teachers,” *Science* vol. 177 (1972) 1164-1166. o. Az egyének szintjén gyengébb lenne a korreláció; az összefüggés mindazonáltal jelzésértékű.
 14. Az adatok a College Board sajtó számára készített, 1993. augusztus 19-iki tájékoztatójából származnak. Az egyes államok átlagos SAT pontszámainak átlaga 501, a szórás 37 volt. Az emeltszintű érettségis írók államonkénti arányának átlaga 37, a szórás 28 volt. New York államban 74%, Wyomingban 13% írt emeltszintű érettségit.

10. FEJEZET. REGRESSZIÓSZÁMÍTÁS

1. Kissé kerekítettünk. A pontos számok (a súlyozatlan adatokra):

magasságtálg = 69,6 hüvelyk	szórás = 2,79 hüvelyk	
testsúlyátlag = 162 font	szórás = 29,1 font	$r = 0,448$

Az adatok az Inter-University Consortium for Political and Social Research által rendelkezésünkre bocsátott mágnesszalagról származnak.
2. Lásd a 8. fejezet 5. jegyzetét.
3. Az „átlagdiagram” nem bevett szakkifejezés. Elméletileg az ábra függ az x tengely beosztásának finomságától is.
4. A medián görbéje meglepő módon sokkal kevésbé szabályos.
5. Az apák magasságtálg 67,7 hüvelyk, a szórás 2,74 hüvelyk; a fiúk 68,7 hüvelyk, 2,81 hüvelyk szórással; $r = 0,501$. Az eredeti táblázatnál a szórás 2,72, illetve 2,75, $r \approx 0,514$; az átlagokat csak minimálisan befolyásolta a kerekítés. Lásd a 8. fejezet 2. jegyzetét.
6. D. Kahnemann and A. Tversky, „On the psychology of prediction,” *Psychological Review* vol. 80 (1973). 237-251. o.
7. Dr. Marjorie Honzik bocsátotta rendelkezésünkre az adatokat. Összehasonlításképpen: a rendszeres népességfelmérés 1993 márciusi adatai szerint a házastársak iskolázottsági szintje közötti korreláció 0,66.

11. FEJEZET. A REGRESSZIÓS EGYENES NÉGYZETES KÖZÉPHIBÁJA

1. *Statistical Methods for Research Workers* (Oliver and Boyd, 1958, 182. o.)
2. Többváltozós regresszióanalízis maradékokat a függő változóval, az egyes független változókkal, a becslésekkel és be nem vont változókkal összevetve is ábrázolhatjuk.
3. 60 családban volt 64 hüvelyk magas az apa (kerekítve); a fiúk magasságtálg 66,7 hüvelyk, 2,29 hüvelyk szórással. 50 családban volt 72 hüvelyk magas az apa (kerekítve); a fiúk magasságtálg 70,7 hüvelyk, a szórás 2,30 hüvelyk.
4. Egy „homoscedasztikus” pontdiagram bármely keskeny függőleges sávjában többé-kevésbé ugyanakkora a szórás, de nincs kizárva, hogy a tendencia nem lineáris. Az általunk „rögbilabda alakúnak”

nevezett pontdiagram homoscedasztikus és lineáris tendenciát mutat; az ilyen pontdiagramok úgy néznek ki, mint egy kétváltozós normális eloszlásból vett minta adatai.

5. A könyvnek ebben a részében alapvetően a leírásra koncentrálunk. Adott egy véges adatsor, minden ponthoz egyforma súlyt rendelünk. A regressziós egyenest úgy tekintjük, mint amely kisimítja $E\{Y | X\}$ -et ebben a véges sokaságban. Hasonlóan, az adott sokaság véletlenszerűen kiválasztott elemére előre jelezzük Y -t az X -ből. A $\sqrt{1 - r^2}$ képlet pontos az előrejelzési hibára ebben az összefüggésben.

Amikor egy minta adataira illesztünk regressziós egyenest, melyet azután előrejelzésre akarunk felhasználni, fellép a fentebb tárgyalt regressziós egyenes körüli varianciakomponens. Emellett a mintából kapott regressziós egyenes is tartalmaz a populációra érvényes egyenes körüli varianciakomponenset. Ez utóbbi természetesen függ X -től, de ennek tárgyalásával könyvünkben már nem foglalkozunk. Nagy minta esetén az utóbbi komponens elég kicsi lesz. Képzeljük el az 5. szakasz 1. példáját 100 fős mintával. Az első (a regressziós egyenes körüli) komponens 64 körül van. Egy olyan hallgatóra, aki felvételin z szórásnival tér el az átlagtól, a második—a regressziós egyenes mintavételi hibája miatt fellépő—komponens $0,64(1 + z^2)$ körül van. Igen extrém esetnek számít $z = 3$. Ekkor a teljes négyzetes középhiba 8,4 körül lesz; a második tényező figyelmen kívül hagyásával pedig 8,0.

Persze korántsem szabad a vásár! Elég távol a pontdiagram középpontjától már meglehetősen megbízhatatlannak válnak a regressziós becslések. Egyrészt megtörhet a linearitás, másrészt a mintavételi hibának is nagyobb ára van. A következőképpen képzelhetjük el magunknak ez utóbbit. A regressziós egyenesnek át kell mennie az átlagpontra, amely viszonylag stabilan rögzített, legalábbis elég nagy mintánál. Az egyenes meredekségében azonban van valamelyes bizonytalanság. Minél távolabb kerülünk az átlagponttól, annál erősebben számít a meredekség: bizonytalansága felnagyítódik.

6. Lásd a 8. fejezet 5. jegyzetét. Az összegző statisztikákat súlyozatlanul és leegyszerűsítve közöljük. A mintában a munkaerőforrásba sorolt nők szerepelnek. Egy-egy ponthoz több eset is tartozhat.
7. A Lake Mead-i 3-as állomáson mért adatok; a rendelkezésre álló adatok mértani közepe 53 hónapra; az 1976-86-os mintavételi időszakból. Az adatokat az azóta elhunyt Jerome Horowitz bocsátotta közre egy, a vízminőséggel kapcsolatos meghallgatás alkalmából.
8. Az adatokat Dr. Marjorie Honzik bocsátotta rendelkezésünkre.
9. A dobók nem szerepelnek az adatok közt; a „másodéves elslamposodás” azonban a dobójátékosoknál is megfigyelhető az elfutásokban. Van olyan év, hogy két „Év Újoncát” is választanak ugyanabban a ligában. A problémát eredetileg David Lane vetette fel. Szeretnénk köszönetet mondani Sam Buttreynek és Oren Tverskynek a baseball statisztikákkal kapcsolatos szakértő tanácsaikért. Az adatokat a STATS (Skokie, Illinois) bocsátotta rendelkezésünkre.

Az összegző statisztikák függenek attól, hogy hányszor szerepelt ütőként valaki, és ezért attól is, hogy súlyozzák-e a játékosok átlagát az ütések számával. (A jó ütőket gyakrabban állítják be ütéshez, de ez a hatás az ütések számának növekedésével csökken.) Akadnak eltérések az egyes évek között, és a két liga között is. A szövegben szereplő 260-as ütőátlag tipikus értéknek mondható akkor, ha súlyoznak az ütések számával, vagy csak azokat a játékosokat veszik figyelembe, akik a szezonban legalább 100-szor ütöttek. Az „Év Újoncai” ütőjátékosok, és jellemzően sokszor kerültek ütőhelyzetbe; sima átlagolással jött ki a 285 és a 257.

A most következő adatok az 1992-es és az 1993-as szezonra vonatkoznak; egyesítettük a két liga adatait; és egyszerű átlagokat számoltunk. 588 játékos szerepelt mindkét szezonban; 438 ütött legalább 25-ször mindkét szezonban; a 438 ütőátlag-párra vonatkozó összegző statisztikák:

1992	átlag = 241	szórás = 55
1993	átlag = 250	szórás = 55

az évek közötti korreláció = 0,52.

298 játékos ütött mindkét szezonban legalább 100-szor. Az összegző statisztikák:

1992	átlag = 260	szórás = 30
1993	átlag = 269	szórás = 35

az évek közötti korreláció = 0,26.

A korreláció a leszűkítés miatt gyengült le: a 25-100 ütéshez beállított játékosok közül soknak 200 alatt volt az ütőátlaga, a 100-nál többször ütőknél alig fordul elő ilyen gyenge teljesítmény. A mérési

hiba is szerepet játszik. 186 játékos volt, aki legalább 250-szer ütött mindkét szezonban. Az összegző statisztikák náluk:

1992	átlag = 268	szórás = 27
1993	átlag = 276	szórás = 31

az évek közötti korreláció = 0,40.

12. FEJEZET. A REGRESSZIÓS EGYENES

1. *Abhandlungen zur Methode der kleinsten Quadrate* (Berlin, 1887, 6. o.). A könyvben angolul L. Le Cam és J. Neyman, *Bayes-Bernoulli-Laplace Seminar* (Springer-Verlag, 1965, viii. o.) című könyvéből idézik.
2. Lásd a 8. fejezet 5. jegyzetét. A munkaerőforrásba sorolt férfiak mintája. Az egyes pontok több személyt is jelenthetnek.
3. Lásd a 11. fejezet 6. jegyzetét.
4. Az iskolarendszerbe való visszatérésről a következő művekből tájékozódhatunk:
 - J. D. Angrist és A. B. Krueger, „Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery,” *Princeton University Industrial Relations Section Working Paper #290*, August 1991.
 - O. Ashenfelter és A. Krueger, „Estimates of the economic return to schooling from a new sample of twins,” *American Economic Review* vol. 84 (1994), 1157-1173. o.
 - Z. Griliches, *Technology, Education, and Productivity: Early Papers with Notes to Subsequent Literature* (Basil Blackwell, Oxford, 1988).
5. Tegyük fel például, hogy $y = a + bx + cx^2$. Az alanyokat véletlenszerűen hozzárendeljük a különböző x értékekhez és regressziós egyenest illesztünk az adatokhoz. A meredekség ilyenkor nem ad előrejelzést az x megváltozásának hatására létrejövő, az alanyokra és x értékeikre kiátlagolt y -változásról, vagy legfeljebb csak valamilyen általános vagy átlagos értelemben. Természetesen a pontos képzet szerint is adhatunk becslést az adatok alapján.
6. Az IRRI által rendelkezésre bocsátott összesítő statisztikák kerekített értékeiből kaptuk az egyenletet.
7. Franklin Fisher (MIT) tanúvallomását parafráztuk, melyet a Cuomo kontra Baldrige ügyben tett a regressziós modellekről a népszámlálással kapcsolatban. (80 civ. 4550 SDNY 1987), a jegyzőkönyv 2149ff oldala.
8. William Fretter professzor (University of California, Physics Department) végezte a kísérletet demonstrációképpen bevezető fizika előadásának keretében; a részleteket leegyszerűsítettük.
9. A Berkson-féle hibát tartalmazó változó modell alkalmazandó itt. A súlyok névleges értékeit rögzíti a kutató, a tényleges érték hibának van kitéve; a regressziószámításba a névleges érték kerül. Ha a mérési hibának kitett súly értékét mérnénk, és a mért értékekkel dolgoznánk, akkor a szokásos regressziós becslés torzítana. Lásd G. W. Snedecor és W. G. Cochran, *Statistical Methods*, 6th ed. (Iowa State University Press, 1973).

Az 1. ábrában szereplő adatokat a két átlaggal, a két szórással és az r -rel összegeztük. Valójában itt ésszerűbb összegzés lenne megadni

- (i) a regressziós egyenes meredekségét, tengelymetszetét és négyzetes középhibáját, és
- (ii) a súlyok átlagát és szórását

Az (i) pont alatti statisztikák többé-kevésbé invariánsak (a mintavételei hibára) a különböző súlyokkal végzett kísérletek során. Talán ezért tartja sok statisztikus természetesebbnek a regressziós egyenessel kezdeni, és azután térni rá a korrelációra. Az általunk vizsgált témakörökben azonban nem túl sok példát találunk olyanfajta invarianciára, amelyet a Hooke-törvény mutat. Más típusú példánál – mondjuk a férj és a feleség iskolázottsága esetében – a korrelációs együttható a kézenfekvőbb összegzés. Továbbá az egyenesből kiindulva nehéz elmagyarázni r -et (legalábbis saját tapasztalataink szerint). Ezért kezdtük a tárgyalást a korrelációs együtthatóval – és reméljük, hogy a másik megközelítés hívei elnézik ezt nekünk.

10. Az Inter-University Consortium for Political and Social Research által rendelkezésünkre bocsátott adatszagról számítottuk ki ezeket a statisztikákat. A minta elemszáma 1036, a meredekség tehát valós, nem a mintavételei véletlen műve. (Kihagytuk azokat az eseteket, ahol hiányzott a jövedelem vagy az iskolázottság adata; az összesítő statisztikákat kerekítettük.) Lásd még T. W. Teasdale et al., „Fall

- in association of height with intelligence and educational level,” *British Medical Journal* vol. 298 (1989) 1292-1293. o. Az NHANES III (1988-91) előzetes adatai szerint a testmagasság - iskolázottság regressziós egyenes meredeksége a 25-34 éves férfiakra 0,20 hüvelyk körül volt befejezett iskolai osztályonként; a minta 763 fős volt. (Az adatokat az NCHS, Division of Health Examination Statistics biztosította számunkra; a számításokat magunk végeztük.)
13. A regressziós modellek társadalomtudományi alkalmazásáról lásd a *Journal of Educational Statistics* 1987 nyári számát, továbbá *Sociological Methodology* 1991, *Foundations of Science* vol.1, no. 1 (1995), és *Advances in Applied Mathematics* vol. 18 (1997) 59-110. o.
 14. Lásd a 10. jegyzetet. A jövedelmeket 2-vel szoroztuk az 1995-ös vásárlóértékre konvertáláshoz. Az NHANES III előzetes adatai szerint a magasság és a jövedelem közötti korreláció a 25-34 éves férfiak körében 0,13 volt.
 15. Lásd a 10. jegyzetet. Az NHANES III felvételben részletesen rákérdeztek az étkezési szokásokra, beleértve a pizza- és a sörfogyasztást is. A feladatban szereplő statisztikák nagyjából hűen tükrözik a 18-24 éves amerikai populáció fogyasztási szokásait.
 16. A szakirodalom áttekintéséhez lásd S. J. Pocock, M. Smith és P. Baghurst, „Environmental lead and children’s intelligence: a systematic review of the epidemiological evidence,” *British Medical Journal* vol. 309 (1994), 1189-1197. o.
 17. „Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion,” *British Medical Journal* vol. 297 (1988) 319-328. o. A szerzők szerint kölcsönhatás van a sófogyasztás és az életkor között. A kommentárokat lásd: *British Medical Journal* vol. 312 (1996) 1283ff o. Pozitív kísérleti eredmények találhatók (az alacsony sóbevitelről): „Fall in blood pressure with modest reduction in dietary salt intake in mild hypertension,” *Lancet* (1989. február 25.) 309-402. o.

4. rész. Valószínűség

13. FEJEZET. MIK AZ ESÉLYEK?

1. A valószínűség más szemléletű tárgyalását lásd pl.:
 - R. A. Fisher, *Statistical Methods and Scientific Inference*, 13th ed., reprinted by Oxford University Press, 1993., in: J. H. Bennett (ed.) *Statistical Methods, Experimental Design and Scientific Inference*.
 - L. J. Savage, *Foundations of Statistics*, 2nd ed. (Dover, 1972.)
 Áttekintést ad: *Foundations of Science*, vol. 1. (1995) no.1.
2. A 3. kiadás 1756-ban, de Moivre halála után jelent meg. Reprint: New York, Chelsea Publishing, 1967.
3. *Statistical Abstract*, 1994., 20. táblázat.
4. W. Fairley and F. Mosteller, „A conversation about Collins“, *University of Chicago Law Review* (1974).
5. A vádló két „eseményre” vonatkozóan, két „valószínűséggel” végzett számításokat, oda-vissza ugrálva köztük. Az első esemény az, hogy a vádlottak bűnösök. A második esemény az, hogy Los Angelesben nincs másik, a leírásnak megfelelő pár. Annak, aki a valószínűségszámítással a klasszikus elmélet alapján foglalkozik, a valószínűség fogalma nem igazán tűnik ide illőnek. Bayes követőjének is nehézségei volnának, mert nem találna alkalmas valószínűségi modellt, amelynek révén az adatokat a bűnösség, illetve ártatlanság hipotézisével összekösse. (Lásd még a 6. jegyzetet is.)

Voltak-e Los Angelesben a leírásnak megfelelő más párok? Elvileg statisztikai kérdésnek látszik, amire mintavétellel választ lehetne adni. De ha utánaszámolunk, kiderül, hiába vennénk mintát a Los Angeles-i párok közül, mintavétel a vitát nem döntené el elfogadható megbízhatósági szinten. Teljeskörű cenzusra volna szükség.
6. Az összehasonlításnál használt DNS-„jellemzők”: az ismétlődések száma bizonyos, fehérjét nem kódoló kódszegmensek sorozatos előfordulásainál (Variable Number of Tandem Repeats, VNTR). Hivatkozások:

Jurimetrics, vol.34, no.1 (1993)

D. H. Kaye, „DNA evidence: probability, population genetics and the courts“, *Harvard Journal of Law and Technology* vol. 7(1993), 101-172. o.

J. Cohen, „DNA fingerprinting for forensic identification: potential effects on data interpretation of subpopulation heterogeneity and band number variability“, *American Journal of Human Genetics* vol. 46 (1990), 358-368. o.

J. Slimowitz and J. Cohen, „Violations of the ceiling principle: exact conditions and statistical evidence“, *American Journal of Human Genetics* vol. 53 (1993), 314-323. o.

National Academy of Sciences/National Research Council, *DNA Technology in Forensic Science* (Washington, D.C., 1992)

National Academy of Sciences/National Research Council, *DNA Technology in Forensic Science: An Update* (Washington, D.C., 1996)

„Az ügyész logikai hibája“ abban áll, hogy összekever két dolgot: milyen gyakorisággal fordul elő a lakosság körében a vádlott DNS-e (akár pontos az adat, akár pontatlan), illetve hogy milyen valószínűséggel ártatlan a vádlott; általánosabban – legalábbis a Bayes-féle nézőpontból – abban, hogy összekeveri a $P(\text{bizonyíték} \mid \text{kártatlan})$ valószínűséget a $P(\text{ártatlan} \mid \text{bizonyíték})$ valószínűséggel. Lásd W. C. Thompson and E. L. Schumann, „Interpretation of statistical evidence in criminal trials: the prosecutor’s fallacy and the defense attorney’s fallacy“, *Law and Human Behavior* vol. 11 (1987), 167-187. o.

7. A feladat ötletét D. Kahnemann and A. Tversky („Judgment under uncertainty: heuristics and bias“, *Science* vol. 185 (1974) 1124-31. old.) cikke sugallta. Lásd még D. Kahnemann, P. Slovic and A. Tversky (eds.), *Judgment under Uncertainty: Heuristics and Biases* (Cambridge University Press, 1982)

14. FEJEZET. MÉG MINDIG A VALÓSZÍNŰSÉGRŐL

1. A *Doctrine of Chances* ajánlásából.
2. Természetesen $P(A \text{ vagy } B) = P(A) + P(B)$ ha $P(A \text{ és } B) = 0$. Általánosabban,

$$P(A \text{ vagy } B) = P(A) + P(B \text{ de nem } A),$$

ami analóg az összefüggő eseményekre vonatkozó szorzási szabállyal.

3. A Pascal - Fermat levelezés történetebb ismertetése található F. N. David *Games, Gods and Gambling* (Buckinghamshire, England: Charles Griffin & Co., 1962) 88-89. oldalán. (Magyar nyelven ugyanezről olvashatnak Rényi Alfréd: *Levelek a valószínűségről*, Budapest, Typotex kiadó, 1994. A szerk.)

15. FEJEZET. A BINOMIÁLIS FORMULA

1. Úgy látszik, hogy a binomiális formulát Jia Xian már egy kicsit korábban – a tizenegyedik században felfedezte. Lásd Li Yan and Du Shiran, *Chinese Mathematics* (Oxford University Press, 1987., 121. skk.)
2. Ez csak közelítő modell; az újszülöttek valamivel nagyobb az esélye arra, hogy fiú legyen, és van némi összefüggés egy adott családban az egymást követő születések között.
3. A feladat ötletét D. Kahneman and A. Tversky, „Judgment under uncertainty: heuristics and bias“ c. cikkéből vettük át (*Science*, vol. 185 (1974), 1124-1131. o.).
4. A finn ikervizsgálatról bővebben: J. Kaprio and M. Koskenvuo, „Twins, smoking and mortality: a 12-year prospective study of smoking-discordant twin pairs“, *Social Science and Medicine* vol. 29 (1989) 1083-1089. o. A dohányzás oki szerepére vonatkozó hivatkozások a 2. fejezet 1. jegyzetében található. A dohányzás ártalmosságára vonatkozó állítások közelmúltbeli bírálója P. R. Burch („The Surgeon General’s ‘Epidemiologic Criteria for Causality.’ A Critique.“, *Journal of Chronic Diseases* vol.36 (1983), 821-836. old.).

- A táblázat adatai az aktuális (a vizsgálat idején folytatott) dohányzásra vonatkoznak; férfiak és nők együtt szerepelnek. Kaprio és Koskenvuo elemzett néhány egybemosó változót is. Dohányosok és nemdohányosok között jelentéktelen volt az eltérés a következőkben: szeszesitalfogyasztás, vérnyomás, koleszterinszint, cukorbetegség, kávéfogyasztás, iskolázottság, foglalkozás, családi állapot, Eysenck-féle személyiségletár. A dohányosok kevesebbet mozogtak, ez növelheti a szívinfarktus kockázatát; viszont soványabbak voltak, és ez csökkenti a kockázatot. Amerikai adatok: D. Carmelli and W. F. Page, „Twenty-four year mortality in World War II US male veteran twins discordant for cigarette smoking”, *International Journal of Epidemiology* vol. 25 (1996) 554-559. old.
5. *Statistical Abstract*, 1988, 268. táblázat; 1994, 305. táblázat.
 6. *San Francisco Chronicle*, 1975. december 9. A kutatásról megjelent tanulmány sokkal józanabb: D. J. Ullyot et al., „Improved survival after coronary artery surgery in patients with extensive coronary artery disease”, *Journal of Thoracic and Cardiovascular Surgery* vol. 70 (1975) 404-413. o.
 7. *Bouman v. Block*, 940 F.2d 1211 (9th Cir. 1991)
 8. E. L. Dey, A. W. Astin and W. S. Korn, *The American Freshman: Twenty-Five Year Trends* (University of California, Los Angeles: Higher Education Research Institute, Graduate School of Education, 1991.)
 9. 65 év fölött a grafikon megint nőni kezd. Feltehető, hogy a magas fizetésűek egy része tovább dolgozik, s az ő jövedelmük hat ilyen mutatósan az átlagra. Lásd még *Economic Report of the President 1974*, 147. skk.
 10. T. W. Teasdale et al., „Degree of myopia in relation to intelligence and educational level”, *Lancet* (1988. dec. 10.) 1351-1354. o.

5. rész. Véletlen ingadozás

16. FEJEZET. A NAGY SZÁMOK TÖRVÉNYE

1. *An Experimental Introduction to the Theory of Probability* (University of Witwatersrand Press, 1964). A második világháború után Kerrich Dél-Afrikába ment tanítani.
2. Itt különbséget teszünk a darabszámban („abszolút számokban“) és a százalékban kifejezett eltérés között. Technikaibb szinten az abszolút érték is előkerül. Legyen X_n az n dobás utáni véletlen hiba, azaz a fejek száma, mínusz a dobások számának fele. Ekkor X_n martingál, tehát $E\{X_{n+m} | X_n\} = X_n$; de $E\{|X_{n+m}| | X_n\} > |X_n|$.
3. A „véletlen folyamat“ kifejezést nem szakkifejezésként használjuk. „Egy véletlen folyamatból kapott szám“ annyi, mint egy valószínűségi változó megfigyelt értéke.
4. A számítógépes programok determinisztikusak, igazán véletlenszerű számokat ezért nem tudnak létrehozni. Elő tudnak azonban állítani olyan számokat, melyek egészen véletlenszerűeknek látszanak. Az egyik módszer szerint van egy nagyon nagy szorzó, M ; aztán a programozó választ egy „mag“ (seed) értéket – jelölje ezt x ; x valahol 0 és 1 között lehet. A számítógép kiszámítja, mennyi M -szer x ; az eredménynek van egész része és tizedestört része:

... aaaaaaaaaaaaaaaaa,bbbbbbbbbbb ...

A tizedesvesszőtől balra eső jegyekből lesz az első véletlen szám, s a törtrész szolgál a soron következő véletlen szám magjaként. Bővebb kifejtését lásd

- P. L'Ecuyer, „Efficient and portable combined random number generators”, *Communications of the ACM* vol.31 (1988) 742-74. o.
- D. Knuth, *The Art of Computer Programming*, vol. II, 2nd ed. (Addison-Wesley, 1981). Magyarul: Donald E. Knuth, *A számítógép-programozás művészete 2*, (Műszaki Könyvkiadó, Budapest, 1987)
- P. A. W. Lewis and E. J. Orav, *Simulation Methodology for Statisticians, Operations Analysts, and Engineers* (Wadsworth & Brooks/Cole, 1989, chapter 5).
- G. Marsaglia, „Random numbers fall mainly in the planes”, *Proceedings of the National Academy of Sciences* vol. 60 (1968), 25-28. o.

_____, „A current view of random number generators“, *Proceedings of the Sixteenth Symposium on the Interface between Computer Science and Statistics* (1985), 3-10. o.

- A „húzáások összege“ nem bevett szak kifejezés, de egyszerűbb, mint a „független, azonos eloszlású valószínűségi változók összege“. A „dobozmodell“ sem bevett – bár mintha terjedőben lenne.
- Legyen S_N egy N -edrendű, p paraméterű binomiális eloszlású változó. Azt állítjuk, hogy $P\{S_{2n+2} > n+1\} > P\{S_{2n} > n\}$ minden olyan p -re, melyre $1/2 \leq p < 1$. Valóban, legyen $f_{Nm}(p) = P\{S_N > m\}$. Ekkor
(1) $f'(p) = NP\{S_{N-1} = m\}$.

Legyen $g_n(p) = P\{S_{2n+2} > n+1\} - P\{S_{2n} > n\}$. (1) szerint

$$(2) \quad \frac{d}{dp} g_n(p) = \frac{(2n)!}{n!(n-1)!} \left[\frac{4n+2}{n} p(1-p) - 1 \right] p^n (1-p)^{n-1}.$$

Mármost $(4n+2)/n > 4$. Létezik olyan p_0 , melyre $1/2 < p < p_0$, hogy $g'_n(p)$ pozitív ha $1/2 \leq p < p_0$, negatív ha $p_0 < p < 1$, és 0 ha $p = p_0$. Azaz g_n növekszik $1/2$ és p_0 között, majd csökken p_0 és 1 között. Tegyük fel, hogy $n > 1$. Ekkor $g_n(1) = 0$, s ez elég is, hogy az állítást a $p = 1/2$ esetre bizonyítsa. Vizsgat

$$f_{2n,n}\left(\frac{1}{2}\right) = \frac{1}{2} - P\left\{S_{2n} = n \mid p = \frac{1}{2}\right\}$$

n szerint szigorúan monoton növekedik. Ha $n = 1$, az állítás arra egyszerűsödik, hogy megmutassuk, $3p^2 - 4p + 1 < 0$ ha $1/2 \leq p < 1$, ami könnyen látható.

17. FEJEZET. A VÁRHATÓ ÉRTÉK ÉS A STANDARD HIBA.

- A kenő a bingó Las Vegas-i megfelelője. Nyolcvan, 1-től 80-ig megszámozott golyóval játszá. Minden fordulóban kihúznak 20-at véletlenszerűen, visszatevés nélkül. Ha egy száma, például a 17-esre teszünk, ezzel arra fogadunk, hogy a 17-es ott lesz a kihúzott 20 golyó között. Nyerési esélyünk $20/80=1/4$.
- Ha X_i független, azonos eloszlású, μ várható értékű és σ^2 varianciájú (szórásnégyzetű) változók, akkor

$$E(X_1 + \dots + X_n) = n\mu$$

és

$$\text{var}(X_1 + \dots + X_n) = \text{var} X_1 + \dots + X_n = n\sigma^2.$$

Így a standard hiba (azaz a variancia négyzetgyöke) $\sqrt{n} \sigma$. Ez a négyzetgyökszabály.

- E könyvben szórásnak nevezzük az adatok szórását, míg standard hibának, azaz SH-nak a véletlen mennyiségek (valószínűségi változók) szórását. E különbségtétel nem bevett – mindkét helyzetre a szórás elnevezés a használatos.
- A II. és III. részben adatokra vonatkozóan használtunk standard egységeket – az átlagtól mért eltéréseket fejeztünk ki, egységül a szórás szolgált; most véletlen változókra térünk át – a várható értéktől mért eltéréseket fejezzük ki, egységül a standard hiba szolgált.
- A „becslés“ szót itt köznapi „közelítés“ értelmében használjuk. A statisztikusok egy másik, szakmai értelemben is használják a „becslés“ szót – ez a VI. részben fog előkerülni.
- Tekintsünk tizenegy, -5-től 5-ig megszámozott rekeszt, és 10 golyót. Összesen

$$\binom{20}{10} = 184\,756$$

féleképpen lehet a 10 golyót a 11 rekeszbe kiosztani; minden kiosztás egy 10-lapos „dobozt“ definiál: ha például 4 golyót teszünk a „-1“-es rekeszbe és másik 6-ot a „3“-as rekeszbe, így a

$\boxed{4} \boxed{-1}$ -es $\boxed{6} \boxed{3}$ -as dobozt kapjuk. Mind a 184 756 dobozt megvizsgáltuk; közülük 5448-nak 0 az átlaga, de ez utóbbiak egyikében $10 \boxed{0}$ van. Most a fennmaradó 5447 lehetőséget ellenőriztük, hogy az intuitíve nyilvánvalónak tűnő eredményt bebizonyítsuk. A MATLAB-ban írt program 90 MHz-es

Pentiumon két és fél óra hosszan futott. A kombinatorikához lásd W. Feller, *An Introduction to Probability Theory and its Applications* (3rd ed., John Wiley and Sons, 1968.), I. kötet, II.4. szakasz. Magyarul megjelent: W. Feller: *Bevezetés a valószínűségszámításba és alkalmazásaiba* (Műszaki Könyvkiadó, Budapest, 1978.) Lásd még a 16. fejezetét. Nemmonoton viselkedésre példa lehet, ha 100 illetve 200 húzás összegét vesszük az

$$\boxed{5 \quad -16} \text{-os} \quad 5 \quad \boxed{16} \text{-os}$$

dobozból. Legyen S_n az n húzás összege; ez szükségképpen 32 többszöröse lesz. Így, ha $-15 \leq S_{100} \leq 15$, akkor $S_{100} = 0$; ha $-30 \leq S_{200} \leq 30$, akkor $S_{200} = 0$. És $P\{S_{100}=0\} > P\{S_{200}=0\}$.

7. A Southern Methodist University statisztika csoportjától származik az az ötlet, hogy a dobozban a lapokat „Kicsi”-nek és „Nagy”-nak nevezzük.
8. E. O. Thorp, *Beat the Dealer* (Random House, 1966). Pozitív várható értékűek továbbá a bakkarátban bizonyos kiegészítő fogadások. És a lóversenyen is ez a helyzet, sajátságosan éppen az esélyesekre kötött fogadások egy részénél; lásd R. T. Thaler and W. T. Ziemba, „Parimutuel betting markets: race-tracks and lotteries”, *Journal of Economic Perspectives* vol.2 (1988) 161-174. o. [A lóversenyen a fogadás ún. *parimutuel* szisztéma szerint zajlik: miután bizonyos hányadot levonnak a személyzet számára, a vesztesek összes tétjét felosztják a nyertesek között. – *A ford.*] Hogy visszaálljon az esélyek normális rendje, a kaszinók igyekeznek megátolni a kártyák számolgatását.

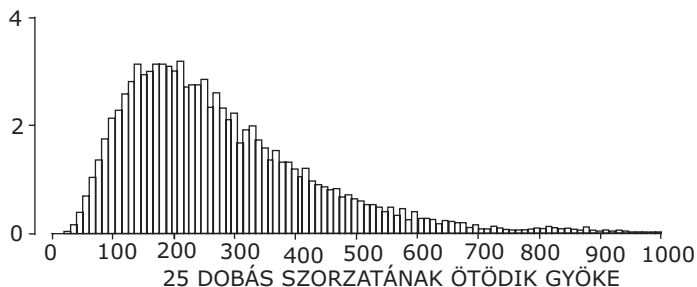
18. FEJEZET. ÉLMÉLETI HISZTOGRAMOK NORMÁLIS KÖZELÍTÉSE

1. Számítások Hewlett Packard HP 15C géppel.
2. Matematikai tárgyalása W. Feller, *An Introduction to Probability Theory and its Applications*, I. kötet, 7. fejezetében (3rd ed., John Wiley & Sons, 1968.) Magyarul megjelent: W. Feller: *Bevezetés a valószínűségszámításba és alkalmazásaiba* (Műszaki Könyvkiadó, Budapest, 1978.)
3. Ha nagyon nagy a húzások száma, célszerű lehet az előforduló értékek csoportosítása, ahogy ez (szorzatokra) a 10. ábrán látható. Hasonlóképpen, ha a lapokon nem egész számok vannak. De lehet értelme a szélesebb osztásközök használatának úgy is, hogy nem nagy a húzások száma, s egész számok vannak a lapokon: ilyen a helyzet, amikor a lapokon lévő számok közötti különbségeknek van 1-nél nagyobb közös osztójuk. Például tegyük fel, hogy 1 dollárral fogadunk érmedobásra, 100-szor. Tiszta nyereségünk olyan, mint 100 véletlenszerű, visszatevéses húzás összege egy $\boxed{-1} \quad \boxed{+1}$ dobozból. A tiszta nyereség lehetséges értékei párosak: 0, ± 2 , ± 4 ,... Rajzolhatjuk a hisztogramot ezen értékek körüli, 1 szélességű téglalapokkal. Viszont jobb eredményre vezet, ha 2 szélességű téglalapokat használunk, a 4. szakasz módszerével.
4. Számítások a HP 15C-vel.
5. A ferdeség (skewness) matematikai elemzését adja az Edgeworth-sorfejtés. Lásd W. Feller, *An Introduction to Probability Theory and its Applications*, II. kötet, 16. fejezetében (2nd ed., John Wiley & Sons, 1970.) Magyarul megjelent: W. Feller: *Bevezetés a valószínűségszámításba és alkalmazásaiba* (Műszaki Könyvkiadó, Budapest, 1978.)
6. A hullámok magyarázata a következő. Ha $\boxed{1} \quad \boxed{1} \quad \boxed{9}$ lenne a doboz összeállítása, az összeg lehetséges értékei: 25, 33, 41,... közöttük 8-as ugrások lennének. Ha $\boxed{2} \quad \boxed{2} \quad \boxed{9}$ -es dobozt néznénk, 50, 57, 64,... lennének az összeg által felvehető értékek, 7-es ugrásokkal. A 9. ábrán látható doboz a kettő közötti átmenet – a csúcsok közötti távolság 7 és 8 között váltakozik. Vagy, másként: a csúcsok megfelelnek annak, hány 9-es kerül a 25 húzás közé.
7. A 10. ábra hisztogramjainak alakja talán kicsit meglepő. Azonban, ha X_1, X_2, \dots egy dobókocka egy-más utáni dobásai, akkor közelítőleg

$$(X_1 \cdot X_2 \cdots X_n)^{1/\sqrt{n}}$$

lesz centrálás után lognormális. Alább látható a 25 dobás szorzatának ötödik gyökére vonatkozó elméleti hisztogram; megfelelő az alakja. A valószínűségeket kombinatorikus algoritmussal számított-

tuk; a kiugrások valódiak. (A 25 dobás szorzata $2^a 3^b 5^c$ alakú, ahol a , b és c nemnegatív egészek – ez magyarázza a kiugrásokat és a hézagokat.)



25 dobás szorzatának 5. gyöke

A 25 dobás szorzatának (10-es alapú) logaritmusa: 25 logaritmus összege. Ezek mind 0,4762 átlagúak és 0,2627 szórásúak, így a 25 logaritmus összegének várható értéke $25 \cdot 0,4762 \approx 11,91$, standard hibája $\sqrt{25} \cdot 0,2627 \approx 1,31$. A 25 logaritmus összege már majdnem normális eloszlású. Nézzük a 10. ábra alsó rajzát, 25 dobás szorzatáról. A tengely 10^{13} -nál van elvágtva, ez a logaritmikus skálán 13, azaz 0,83 standard egység. A valószínűségnek körülbelül 20%-a esik ettől az értéktől jobbra. A hisztogramon a téglalapok 10^{11} szélesek. Az első téglalap a logaritmikus skála szerint a $-\infty$ és 11 közötti intervallumot fedi le – ez standard egységekben $(-\infty; -0,69)$. Ez az intervallum a valószínűség 25%-át tartalmazza!

8. Hallgatólagos feltevések: nullától különböző szórás, és a dobozban véges számú lap, rajtuk egész számok. Mi mondható a standard egységekre való átváltásról? Az egyszerűség kedvéért tegyük fel, hogy a dobozban a számok nem periodikusak; átlaguk legyen μ , szórásuk pedig σ . Legyen $h_n(x)$ az n húzás összegére vonatkozó hisztogram, a nálunk szokásos konvenciók szerint rajzolva: minden téglalap 1 széles, s egy lehetséges érték van a felezőpontjában. Legyen $\phi(z)$ a standard normális eloszlásfüggvény. Ekkor

$$\sigma \sqrt{n} h_n(n\mu + \sigma \sqrt{n}z) \rightarrow \phi(z)$$

Ezt a skála-transzformációt hangsúlyozzuk, amikor standard egységekben rajzolt hisztogramról beszélünk. Lásd W. Feller, *An Introduction to Probability Theory and its Applications*, II. kötet (2nd ed., John Wiley & Sons, 1971., 517. és 540. o.) Magyarul megjelent: W. Feller: *Bevezetés a valószínűség-számításba és alkalmazásaiba* (Műszaki Könyvkiadó, Budapest, 1978.)

9. Tegyük fel, van két doboz, melyekben a lapok átlaga, és az átlagtól való átlagos abszolút eltérések is megegyeznek. Ha egyezik a szórásuk is, akkor egyforma lesz az összegek aszimptotikus viselkedése – ha nem, nem. Egy példa:

$$\text{A) } \begin{array}{|c|c|} \hline -1 & 1 \\ \hline \end{array} \quad \text{B) } \begin{array}{|c|c|c|c|} \hline -2 & 0 & 0 & 2 \\ \hline \end{array}$$

Mindkét dobozban 0 a lapok átlaga, 1 az átlagos abszolút eltérés az átlagtól. De az A doboznál a szórás =1, B-nél a szórás $\approx 1,4$; következésképp a B-ből végzett 100 húzás összegének a szóródása körülbelül 1,4-szer akkora lesz – a szóródást bármilyen ésszerű módon mérve –, mint az A-ből végzett 100 húzásénak. Az összeg aszimptotikus viselkedését a dobozbeli számok átlaga és szórása határozza meg – a középérték és a szóródás többi mértékszámai nem képesek erre.

10. Jelölje a húzások számát n , az ismétlések számát k . Implicit feltevés, hogy $k/\sqrt{n} \log n \rightarrow \infty$. Lásd D. A. Freedman, „A central limit theorem for empirical histograms“, *Z. Wahrscheinlichkeitstheorie* vol.41 (1977) 1-11. o.

6. rész. Mintavétel

19. FEJEZET. NAGY MINTÁN VÉGZETT FELMÉRÉSEK

1. A fejezet nyitóidézete a *The Adventure of the Copper Beeches* című kötetből származik. Don McNeill *Interactive Data Analysis* című könyvében (John Wiley & Sons) bukkantunk rá.
2. Néhány alampmú a mintavételről:

MATEMATIKAI STATISZTIKAI ELŐKÉPZETTSÉGET KEVÉSSÉ IGÉNYLŐK:

N. M. Bradburn-S. Sudman, *Polls and Surveys* (Jossey-Bass Inc., 1988).

A. Campbell-G. Gurin-W. Miller, *The Voter Decides* (Evanston: Row, Peterson, 1954).

Jean M. Converse, *Survey Research in the United States: Roots and Emergence, 1890-1960* (University of California Press, 1987).

George Gallup, *The Sophisticated Poll Watcher's Guide* (Princeton Opinion Press, 1972).

Herbert Hyman et al., *Interviewing in Social Research* (University of Chicago Press, 1954).

Frederick Mosteller et al., *The Pre-Election Polls of 1948* (New York: Social Science Research Council, 1949).

Mildred Parten, *Surveys, Polls and Samples* (Harper & Row, 1950).

F. F. Stephan-P. J. McCarthy, *Sampling Opinions* (John Wiley & Sons, 1989).

MATEMATIKAI STATISZTIKAI ELŐKÉPZETTSÉGET IGÉNYLŐK:

W. G. Cochran, *Sampling Techniques*, 3. kiadás (John Wiley & Sons, 1977).

Robert M. Groves, *Survey Errors and Survey Costs* (John Wiley & Sons, 1989).

M. H. Hansen-W. N. Hurwitz-W. G. Madow, *Sample Survey Methods and Theory* (John Wiley & Sons, 1953).

Leslie Kish, *Statistical Design for Research* (John Wiley & Sons, 1987).

Seymour Sudman, *Applied Sampling* (Academic Press, 1976).

3. Az idézetek a *New York Times*ből származnak (1936. október 1-15.)
4. Köszönetet szeretnénk mondani a Gallup Poll következő munkatársainak: Diane Colasanto, Laura Kalb, Jack Ludwig, Coleen McMurray, Paul Perry, akik minden információ iránti kérésünket teljesítették. A 3. ábra a Gallup Intézet által rendelkezésünkre bocsátott példány alapján került kinyomtatásra, szíves engedélyükkel közöljük. A harmadik kiadáshoz nyújtott segítségükért a következőket illeti köszönet: Lydia Saad, David Moore, Kim Neighbor. Az 1992-es felvételt a velük folytatott beszélgetések alapján írtuk le.
5. Lásd Parten, 393. o., Stephan - McCarthy, 241-270. o. (lásd 2. jegyzet).
6. M. C. Bryson „The *Literary Digest* poll: making of a statistical myth” című cikkében (*American Statistician* vol. 30 (1976) pp. 184-185.) másképp látja azt. Bryson egyetért azzal, hogy a válaszhíányok miatti torzítás tönkretette a *Digest* közvéleménykutatását, a mintavételi torzítást azonban sokkal kisebbnek tartja és megkérdőjelezi, hogy valóban telefonkönyvekből merítették-e a mintát címlistájukhoz (azaz a megkérdezendők listájához). A mi legfontosabb információforrásunk a *Digest* közvéleménykutatásával kapcsolatban George Gallup volt—éles elméjű, a kérdésben érdekelt, közvetlen megfigyelő. Gallup határozottan állította, hogy a *Digest* telefonkönyveket, automobiltulajdonosok listáit és saját előfizetőinek listáját használta a címlista összeállításához. Ezt a leírást lényegében megerősítik mások is, mint Parten vagy Stephan és McCarthy (lásd 2. jegyzet).

Nem tudunk fellelni olyan publikációt, melyben a *Digest* részletesen számot adott volna az alkalmazott eljárásról. A lap 1936 augusztus 22 és november 14 közötti számaiból azonban kiderül valami a mintavételről. Például:

Közvéleménykutatásunk harminc éves fejlődés és tökéletesítés eredménye. Olyan „kereskedelmi célú” mintavételi módszereken alapul, melyeket több mint száz éve alkalmaznak már a könyvkiadók forgalmuk növelésére. A mostani címlista összeállításához az *Egyesült Államok összes telefonkönyvét*, klubok és egyesületek tagnyilvántartásait, városi címlistákat, a regisztrált szavazók listáit, csomagküldő szolgálatok listáit és foglalkozási adatokat használtunk fel. [augusztus 22. 3. oldal, kiemelés tőlünk.]

A cikk a továbbiakban kifejti, hogy a címlista eredetileg az 1924-es választásokra készült, de azt „képzett szakértők” folyamatosan átdolgozták. A listán szereplő nevek többségét évről évre követték, és két választás között is folytattak közvéleménykutatásokat körükben (augusztus 22. 3. oldal). A re-

gisztrált választók névjegyzékeinek felhasználása az 1936-as év újtása volt, és úgy tűnik, hogy csak egyes „nagyvárosokban” alkalmazták (október 17. 7. oldal).

Modern mércével mérve a *Digest* címlistáját eléggé önkényesen állították össze, és torzított: a lakosság egy jelentős, jól meghatározható része kimaradt belőle. Bryson szerint, ha valamiképpen sikerült volna elérni, hogy mind a 10 millió listán szereplő személy válaszoljon, akkor meg tudták volna jósolni a választás kimenetelét. Ám ez valószínűtlennek tűnik. A *Digest* nyilvánvalóan bízott abban, hogy a helyes arányokat kapja meg (augusztus 22. 3. oldal):

Hangsúlyozzuk, hogy A DIGEST több mint 10 millió embert kérdezett meg – minden negyedik választópolgárt az ország minden tájáról –, hogy már októberben megállapítsa a novemberi választási eredményeket. Jövő héten megérkeznek az első kitöltött szavazócédulák, melyeket további milliók özöne követ majd, hogy *háromszor is ellenőrizzék, igazolják, öt szempont szerint osztályozzák és összesítsék* őket. Amikor az utolsó szám is összeadásra és ellenőrzésre került, ha hagyatkozhatunk a múltbeli tapasztalatokra, *1 százalékos pontossággal* megismerheti majd ország-világ a negyvenmillió szavazat jelenlegi állását. [Az újság kiemelései.]

A *Digest* 19 százalékpontot tévedett. Mint az kifejtettük, két fő oka volt ennek: a mintavételi torzítása és a nem-válaszolásból fakadó torzítás.

7. Ez a 65% a '80-as évek végén jellemző a négy megkereséses valószínűségi mintákra. A válaszolási arány erősen lecsökkent: 1975-ben még 75%, 1960-ban 85% volt. Ez a csökkenés a közvéleménykutató szervezetek egyik legnagyobb gondja.
8. Ez a szakasz Mosteller et al. könyvéből (lásd 2. jegyzet) merít jelentős mértékben.
9. Stephan and McCarthy, 286. o. (2. jegyzet).
10. Nagy a kísértés a kvótás mintavétel és a rétegzett mintavétel összekeverésére, pedig a kettő nagyon is különbözik egymástól. Képzeljük el példaképpen, hogy egy városban 200 fős mintát kell vennünk úgy, hogy a nem szerinti megoszlást megtartsuk; nevezetesen a férfiak és a nők számának egyenlőnek kell lennie. A kvótás mintavétellel dolgozó lényegében felfogadhatna két kérdezőbiztost, az egyiket arra, hogy megkérdezzen 100 férfit, a másikat arra, hogy keressen 100 nőt. Egyéb tekintetben azt választhatnák, akit akarnak. Ez bizony nem igazán jó megoldás. Egy rétegzett mintát ezzel szemben a következő módon választanánk:

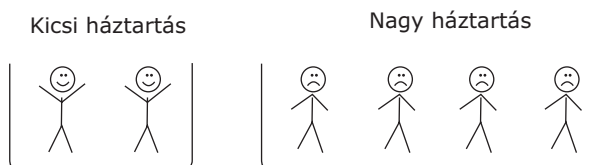
- egyszerű véletlen mintavétellel kiválasztanánk 100 férfit.
- ettől függetlenül, egyszerű véletlen mintavétellel kiválasztanánk 100 nőt.

Ez jobb megoldás, hiszen kizárjuk az egyéni torzítást.

11. „Mintavételi keretnek” nevezzük azt a listát, amelyből a mintát vesszük; a valószínűségi mintavétel első lépése tehát a mintavételi keret összeállítása. Ez meglehetősen bonyodalmas is lehet; és sokszor valamekkora eltérés is van a populáció és a mintavételi keret között. Területi mintáknál a földrajzi egységek listája a mintavételi keret.
12. A 22. fejezetben tárgyaljuk majd részletesen az ilyen mintavételi megoldásokat. Azt állítjuk, hogy szükséges a rétegzés ahhoz, hogy a költségeket ésszerű keretek között tartva vehessünk mintát, de sok közvéleménykutatásnál a rétegzés valójában nem csökkenti különösebben a varianciát. Hipotetikus példaként tegyük fel, hogy egy ország két régióból – Keletből és Nyugatból – áll. Keleten a szavazók 60%-a demokrata; Nyugaton csak 40%. A két régió nagysága egyforma, tehát összességében 50% a demokraták aránya. Két közvéleménykutató cég vesz most mintát a demokraták arányának becsléséhez. Az első n elemű egyszerű véletlen mintát alkalmaz. A standard hiba ekkor $50\% / \sqrt{n}$. A másik rétegzést alkalmaz, és $n/2$ elemű egyszerű véletlen mintát vesz Keleten, és ettől független, ugyanilyen mintát Nyugaton is. A standard hiba $\sqrt{0,4 \cdot 0,6 \cdot 100\% / \sqrt{n}}$. Mínt hogy $\sqrt{0,4 \cdot 0,6} \approx 0,49$, minimálisan csökkent csak a variancia. Ebben a kitalált példában ráadásul sokkal nagyobb a régiók közötti különbség az igazi választásokkor tapasztaltnál. Még kisebb tehát a rétegzés előnye az igazi előrejelzéseknél. (Viszont amikor gazdasági egységek, például cégek vagy intézmények közül kell mintát vennünk, a rétegzés valóban segíthet csökkenteni a varianciát; lásd még a 20. fejezet 5. jegyzetét is.)
13. A Gallup a véletlen kezdőpontú listás mintavétel különböző változatait alkalmazza. Az első három lépcsőben a kiválasztási valószínűség az egység nagyságával arányos: az egyes egységek a méretükkel arányos darabszámban szerepelnek a listán. A négy földrajzi régió mindegyikében külön rétegre sorolják a rurális területeket, melyeket a városias területektől némileg eltérően kezelnek.

14. A Gallup kiadványát idézve: „A háztartáson belüli kiválasztás ezen eljárását empirikusan dolgoztuk ki oly módon, hogy a férfiak, illetve a nők külön életkori megoszlása közel hasonló lesz a populáció megoszlásához.”
15. Szigorúan véve csak a háztartásokra lehet kiszámolni a kiválasztási valószínűséget, az egyénekre nem – a háztartáson belül az egyének kiválasztására vonatkozó szabály következtében. A nem válaszolók szintén bonyodalmat jelentenek. Köszönet illeti Ben Kinget (Florida) ezen kérdés kimerítő elemzéséért. Többnyire úgy tervezik meg a valószínűségi mintavételt, hogy az alapsokaság minden tagja egyforma eséllyel kerüljön be a mintába – ekkor a minta „önsúlyozó”. A Gallup Intézet azonban egyetlen személyt kérdez meg a kiválasztott háztartásokból. Ez hátrányos megkülönböztetést jelent a nagy háztartásban élőkkel szemben, túl kevesen kerülnek közülük a mintába. A torzítás kiigazítására korrekciót alkalmaznak: nagyobb súlyt rendelnek a nagyobb háztartásban élő mintába került személyekhez. A háztartás mérete a kérdőív 18. kérdéséből tudható (4. ábra).

A háztartásméret torzítása. Képzeld el, hogy az alábbi két háztartás közül véletlenszerűen kiválasztunk egyet. Azután véletlenszerűen kiválasztunk egy személyt a kiválasztott háztartásból. Így egyelemű mintához jutunk. A kis háztartásban élő személyeknek nagyobb esélyük van a mintába kerülésre.



16. Paul Perry, „A Comparison of the voting preferences of likely voters and likely nonvoters,” *Public Opinion Quarterly* vol. 37 (1973) 99-109. o. Nyilvános adat, hogy ki szavazott; hogy kire szavazott, az természetesen nem.
17. A Gallup-féle „titkos szavazás” nem titkos, össze lehet kapcsolni a szavazócédulát a kérdőívvel.
18. A válaszmegtagadási arány nevezőjében a kérdezők által elért személyek száma áll. A felhívott számok mintegy 30%-ában a kérdezőknek még egy válaszmegtagadást sem sikerül kieszközölniük: soha senki nem veszi fel a telefont, vagy legfeljebb csak az üzenetrögzítő – a huszadik század végi telefónia nagy átka.
19. L. Belmont-F. Marolla, „Birth-order, family-size, and intelligence,” *Science* vol. 182 (1973) 1096-1101. o. Az intelligencia átlagosan csökken a születési sorrend és a családméret növekedésével még akkor is, ha kontrollváltozóként bevonjuk a családi hátteret. Lásd még: R. B. Zajonc, „Family configuration and intelligence,” *Science* vol. 192 (1976) 227-236. o. A kutatásról a 29. fejezet nagy ismétlő feladat-sorának 40. feladatában ismét szó lesz.
20. Kenneth Stampp, a történettudomány emeritusz professzora, University of California, Berkeley. A WPA (az öregedéssel foglalkozó nemzetközi program) keretében az alanyoknak a hetvenes éveikben kellett járniuk!
21. R. W. Fogel-S. L. Engerman, *Time on the Cross* (New York: W. W. Norton & Company, 1989, pp. 39); *Evidence and Methods* (Little, Brown & Company, 1974, 37. o.). Alapos kritikát nyújt Richard Sutch, „The treatment received by American slaves,” *Explorations in Economic History* vol. 12 (1975) 335-438. o.
22. L. L. Bairds, *The Graduates* (Princeton: ETS, 1973).
23. A. L. Cochran fejtegetése az Orvosi Kutatások Tanácsában (The Medical Research Council): *The Application of Scientific Methods to Industrial and Service Medicine* (London: HMSO, 1951, 36-39 o.).
24. A. C. Nielsen, 1987 *Annual Report on Television*; *New York Times*, 1997. március 10. p. C1.
25. Lásd Raymond A. Eve-Dana Dunn, „Psychic powers, astrology and creationism in the classroom,” *American Biology Teacher* vol. 52 (1990) 10-21. o. A kutatók a 20.000 névből álló listából kiválasztott 387 fős mintától 190 választ kaptak vissza. A lista pedig a középiskolai biológia és élettudományi tanárok országos nyilvántartásából választott szisztematikus minta volt. Ez egy figyelemre érdemes, jó kutatás. Sajnálatos módon mi a könyv második kiadásának első nyomásaiban az újság leírására ha-

- gyatkoztunk, mely döntő fontosságú részletekről feledkezett meg, és így helytelen következtetést vontunk le a nem-válaszolásból eredő torzításról.
26. Parten könyvében (2. jegyzet) szereplő feladat alapján.
 27. A *Time on the Cross* című könyvből (2. jegyzet). Anne Aroundel Cecil Calvertnek, Maryland sorrendben második urának volt a felesége. Az időszak két legfontosabb rabszolgapiaca Annapolisban (Aroundel megye) és Charlestonban (Dél-Karolina) volt. Sharon Tuckernek köszönjük ismereteinket Maryland történelméről.
 28. E. K. Strong, *Japanese in California* (Stanford University Press, 1933).
 29. *San Francisco Chronicle*, 1987. december 10.; Stephen Peroutka levele a *New England Journal of Medicine* egyik cikkéhez (vol. 317 (1987) 1542-1543 o.).
 30. A példa ötlete a következő cikkből származik: D. Kahnemann–A. Tversky, „Judgement under uncertainty: heuristics and bias,” *Science* vol. 185 (1974) 1124-1131. o.

20. FEJEZET. VÉLETLEN HIBÁK MINTAVÉTELNÉL

1. A feladat ötletét a Health Examination Survey első ütemében végzett követéses vizsgálatok adták. Az alapfelmérés a 18-79 éves amerikaiak valószínűségi mintáján készült 1960-61-ben. A mintanagyság 6672 fő volt, ebből 3091 fő volt a férfi.
2. Sir Arthur Conan Doyle, *The Sign of the Four* (J. B. Lippincott, 1899; Ballantine Books, 1974, 91. o.). Holmes Winwood Reade-nek tulajdonítja a gondolatot.
3. A 3. ábrán látható hisztogram, akárcsak a 2. példa számításai, visszatevéses mintavételre vonatkozik. Ennél a példánál – 400 fős minta egy 100 000-es alapsokaságból – alig van különbség visszatevéses és visszatevés nélküli mintavétel között. Részletesebben lásd a következő szakaszban. A függőleges tengelyen százalék per standard egység szerepel.
4. Az országos adatok megtalálhatók: *Statistical Abstract*, 1994; a családi állapotra vonatkozó adatok a 60. táblázatban; a 25 éven felüliek iskolai végzettsége a 232. táblázatban; a családi jövedelem a 720. táblázatban, az adókedvezmények az 526. táblázatban szerepelnek.
5. Más összefüggésben többféleképpen is felmerülhet a probléma. Tegyük fel, például, hogy két különböző rétegből veszünk mintát, és rögzített számú kérdőívet akarunk felosztani a két réteg között. Ha az a cél, hogy egyformán pontos legyen a százalékarányok becslése, akkor első megközelítésként ésszerű azonos nagyságú mintákat választanunk. Ha a darabszámok becslésénél törekszünk egyforma pontosságra, vagy összesített (a két rétegre együtt vonatkozó) százalékarányra vagyunk kíváncsiak, akkor általában a nagyobb rétegből nagyobb mintát érdemes venni. Ne becsljük túl, hogy réteggéssel mennyi nyerhető pontosságban az egyszerű véletlen mintavételhez képest (19. fejezet 12. jegyzete).
6. A szavazókorú népesség száma államonként az 1993. évi *Statistical Abstract* 35. táblázatában, az államonkénti választási eredmények a 431. táblázatban találhatók.

21. FEJEZET. A SZÁZALÉKARÁNYOK PONTOSSÁGA

1. Sir Arthur Conan Doyle, *A Study in Scarlet* (J.B. Lippincott, 1983; Ballantine Books, 1975, 136. o.).
2. A szövegben leírt módszer speciális esete annak, amit a statisztikusok manapság a standard hiba becslésére szolgáló „bootstrap” eljárásnak neveznek. Lásd Brad Efron és Rob Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall, 1993).
3. A felsőoktatási intézménybe járók számáról az USA területén lásd: *Statistical Abstract*, 1994, 269ff táblázatok.
4. A családi jövedelemre vonatkozó adatok a háztartásfő iskolai végzettsége szerint megtalálhatók: *Statistical Abstract*, 1994, 708 táblázat.
5. *Census of Manufactures*, 1982, *Subject Series, General Summary, Part 2: Industry Statistics by Employment Size of Establishment*. A 4. és 5. feladatok számai nagyjából megfelelnek a valós adatoknak.

6. Tegyük fel, hogy visszatevéssel húzunk véletlenszerűen. Ha az n mintanagyságra $n \rightarrow \infty$, a $P\{\hat{p} - k \text{ SH} < p < \hat{p} + k \text{ SH}\}$ a $-k$ és k közötti normálgörbe alatti területhez tart; ez pedig kisebb 1-nél.
7. Könyvünkben szigorúan a klasszikus gyakorisági megközelítéshez tartjuk magunkat. Másfajta felfogásokat a 13. fejezet 1. l. ábjegyzetében idéztünk. Sok kollégánk fogja úgy érezni, hogy ebben a szakaszban becsukott szemmel járkáltunk egy intellektuális aknamezőn, de reméljük, hogy könyörületesen fogják megítélni a végeredményt.
8. Az ábra ötlete Juan Ludlow-tól származik (CIMASS/UNAM, Mexico).
9. Az egyszerű véletlen mintákra vonatkozó standard hibát sokszor a „kézreeseő” mintákra is kiszámolják. Ezek a számítások bizonyos összefüggésben hasznosak is lehetnek. Például, ha azt kívánják bemutatni, hogy a kérdéses mintavételi eljárás mennyire különbözik az egyszerű véletlen mintavételtől (amint az a 23. fejezet 6. szakasz ismétlődő feladatsorának 26. feladatában szerepelt).
10. *Statistical Abstract*, 1994, 1223. táblázat. Az autótulajdonosok országos adatait lásd az 1026. táblázatban; a 3. feladat adatai nem tekinthetők reprezentatívnak az egész országra.
11. D. Ravitch és C. E. Finn, Jr., *What Do Our 17-Year-Olds Know?* (Harper & Row, 1987). A mintavételt gondosan megtervezték.
12. Az országos adatokat lásd: *Statistical Abstract*, 1994, 883. táblázat.
13. A Kenő szabályait a 17. fejezet 1. jegyzetében magyaráztuk el. Szimpla számnál 20/80 az esély, mivel 20 számot húznak. A dupla szám esélye $(20 \cdot 19) / (80 \cdot 79)$.

22. FEJEZET. A FOGLALKOZTATOTTSÁG ÉS A MUNKANÉLKÜLISÉG MÉRÉSE

1. Köszönettel tartozunk a Népszámlálási Hivatal munkatársainak: Paul Bettin, Charles Jones, Donna Kostanich és Jay Waite volt segítségünkre. Meg kell említenünk Earl Gerson (nyugdíjba vonult) és Daniel Levine (CNSTAT), valamint Morton Boisen és Julius Shiskin (azóta elhunyt) nevét is. A harmadik kiadáshoz nyújtott segítségükért a következőknek szeretnénk köszönetet mondani: Marty Riche, Sherry Courtland, Greg Weyland. A Népszámlálási Hivatal felelősségi körébe tartozik a mintavétel, az adatgyűjtés, az adatok előállítása, valamint a becslések és a megfelelő standard hibák kiszámítása. A Munkaügyi Statisztikai Hivatal végzi el a szezonális kiigazításokat, és a kiadványokért valamint az eredmények közgazdasági értelmezéséért felelős. Néhány hasznos kiadvány a Rendszeres Népeségfelmérésről:
 - Bureau of the Census, *The Current Population Survey—A Report on Methodology*, Technical Paper No. 7 (1963).
 - Bureau of the Census, *The Current Population Survey: Design and Methodology*, Technical Paper No. 40 (1978).
 - Bureau of Labor Statistics, *Handbook of Methods*, Bulletin No. 1171 (1971).
 - Employment and Earnings* vol. 41, no. 4 (April, 1994).
 - President's Committee to Appraise Employment and Unemployment Statistics, *Measuring Employment and Unemployment* (Washington, D.C., 1962).
 - M. Thompson and G. Shapiro, „The Current Population Survey: an overview,” *Annals of Economics and Social Measurement* vol. 2 (1973).
2. Pár kivétel van az ország északkeleti részén; Hawaii szintén kivételt jelent.
3. Néhány nagy végső mintavételi egységet másképpen kezelnek. Az elsődleges mintavételi egységen belül úgy sorolják fel a végső mintavételi egységeket, hogy csökkentsék a különbségeket, azaz rétegzést végeznek. A mintavételi arány nagyjából ekkora a '90-es években, de idővel változik – különösen a 2000-es népszámlálást követő átdolgozás után.
4. A minta úgy készült, hogy a 11 legnagyobb államban egyforma legyen a havonkénti becslések precizitása, a többi 40 államban pedig az éves átlagok pontossága egyezzen meg. (Ezért számít külön államnak Washington D.C.) „Precizitás” alatt a munkanélküli személyek számára vonatkozó becslés variációs együtthatóját értjük.
5. Foglalkoztatottnak számítanak a saját vállalkozásukban vagy saját földjükön dolgozók; valamint azok is, akik legalább 15 órát dolgoztak valamilyen családi vállalkozásban (akár pénzbeli fizetés nélkül).

- Akiket elbocsátottak, de remélik, hogy visszahívják őket a munkába, akkor is munkanélkülinek számítanak, ha nem kerestek más munkát a megadott időszakban. Külön kategóriába („elcsüggedtek”) kerülnek, akik szeretnének munkát vállalni, de már nem keresnek aktívan, mivel nem hisznek abban, hogy munkát találnának. A hivatalos definíciók az első, 1940-es felmérés óta lényegében változatlanok. Végrehajtottak valamelyes módosításokat 1994-ben, a felmérés átdolgozásának részeként; a kérdőív viszont érzékelhetően változott. Lásd a 13. jegyzetet.
- Nem tartoznak bele a büntetésvégrehajtási és az elmeegógyintézetek lakói, valamint a katonaság. Az 1-3 táblázatok adatai nem tartalmazzák szezonális kiigazítást. Az adatok az 1990-es évek elején alkalmazott mintavételi eljárással álltak elő, mely kissé eltér a szövegben leírttól; utóbbit 1995 júliusában vezették be. 1994 márciusára az átdolgozott adatok is rendelkezésre állnak: lásd az *Employment and Earnings* 1995. januári számát. Az eltérések elhanyagolhatóak.
 - A teljes munkaerőforrás a civil munkaerőforrásból plusz a katonaságból áll.
 - A súlyok meghatározására szolgáló eljárást hányadosbecslésnek (ratio estimation) is szokás nevezni. A Népszámlálási Hivatal által alkalmazott módszer valójában a leírtnál valamivel bonyolultabb, ugyanis más demográfiai változók mentén is készítenek összetett csoportosításokat. Ezen felül korrekciót végeznek a mintába került végső mintavételi egységek és az országos adatok közötti eltérések korrigálására a népszámlálási adatok alapján; és korrigálják az adott havi becslést az előző havi minta adatai alapján is. Végezetül arra törekedve is igazítanak a súlyokon, hogy kompenzálják az eltérő mértékű részvételt a népszámlálásban (lásd a 12. jegyzetet).
 - A '70-es években alkalmazott eljárásnál először linearizálták a becsléseket, és mintafelezéssel kiszámították néhány tömbök közötti varianciát. Vázlatos leírását lásd: R. S. Woodruff, „A simple method for approximating the variance of a complicated estimate,” *Journal of the American Statistical Association* vol. 66 (1971) 411-414. o. A teljeskörű leírás megtalálható: Technical Paper No. 40. A 80-as években egy részlegesen kiegyensúlyozott replikációs módszert alkalmaztak. Lásd Janice Lent, „Variance estimation for Current Population Survey small area estimates,” *Proceedings of the Section on Survey Research Methods* (American Statistical Association, 1991. augusztus).
 - A rétegzés csökkenti a standard hibát, ugyanígy a hányadosbecslés is. A csoportos mintavétel viszont megnöveli.
 - Statistical Abstract*, 1994, 448. és 449. táblázatok.
 - Az 1980-as népszámláláskor számba nem vettek tárgyalja R. E. Fay-J. S. Passel-J. G. Robinson-C. D. Cowan, *The Coverage of the Population in the 1980 Census*, Bureau of the Census, 1988; lásd még *Survey Methodology* vol. 18, no. 1, 1992. június. Az 1990-es hiányok tárgyalását és korrekciós javaslatokat lásd *Jurimetrics* vol. 34, no. 1 (1993. ősz), *Statistical Science* (1994 november), *Evaluation Review* (1996. augusztus). A népszámlálásban és Rendszeres Népeségfelmérésben (CPS) való részvétel közötti különbségekről lásd *Money Income of Households, Families, and Persons in the United States 1992, Current Population Reports*, P60-184 (Bureau of the Census, U.S. Department of Commerce, Washington, D.C., D-4 o.).
 - A tapasztalat szerint meglehetősen jól sorolják be a teljes munkaidőben dolgozókat és a munkaerőforrásba nem tartozókat. Probléma a harmadik csoporttal van: a határeseteket hol részmunkaidős foglalkoztatottnak, hol munkával rendelkező, de nem dolgozónak, hol munkanélkülinek sorolják be. A következőképpen néz ki például az 1987. utolsó negyedévi megismételt kérdézésről készült keresztábra:

Gazdasági aktivitás a megismételt kérdézésnél	Gazdasági aktivitás az eredeti kérdézésnél						
	Mezőgazdasági	Nem mezőgazdasági			Munkanélküli	Munkaerőforráson kívüli	Összesen
		Teljes munkaidős	Részmunkaidős	Nem dolgozik			
Mezőgazdasági	117	1	2	0	0	2	122
Nem mezőgazdasági							
Teljes munkaidős	0	2967	22	2	3	3	2997
Részmunkaidős	1	45	1187	5	4	28	1270
Nem dolgozik	0	2	0	137	2	4	145
Munkanélküli	0	0	2	2	226	21	251
Munkaerőforráson kívüli	0	0	2	1	7	2716	2726
Összesen	118	3015	1215	147	242	2774	7511

MEGJEGYZÉS: Egyeztetés után, súlyozás előtt; 75%-os minta
 FORRÁS: Bureau of the Census, Statistical Methods Division

Tehát 7511 embert kérdeztek meg ismételten. Az eredeti kérdezőskor 3015 főt soroltak be teljes munkaidőben dolgozóként, de csak 2997 főt minősítettek így – feltehetőleg helyesen – az ismételt kérdezőskor. A csökkenés 0,6%. Másfelől 4,5%-kal nőtt a részmunkaidőben dolgozók, 3,7%-kal a munkanélküliek száma. A munkanélküliek teljes számát az eredeti kérdőívek alapján mintegy 7 millióra becsülték. Ennek 3,7%-a 250.000, tehát a becslés torzítása pár százezer embert jelentett. A munkanélküliek száma ebben az adatsorban kicsi, a számolás tehát csak illusztrálni kívánja a lényegét. Lásd még K. W. Clarkson - R. F. Meiners, „Institutional changes, reported unemployment, and induced institutional changes,” a *Journal of Monetary Economics* (1979) melléklete.

1994-ben jelentősebben átdolgozták a felmérés kérdőívét. Új „ellenőrző” kérdéseket tettek bele a ledolgozott órák számáról és a munkanélküliség időtartamáról; megváltoztatták az „elcsüggedtek” és a kényszerből részmunkaidősek definícióját. Lásd *Monthly Labor Review*, 1993. szeptember és *Employment and Earnings*, 1994. február. A kérdések átdolgozása észrevehetően megváltoztatta a számadatokat, megerősítve, hogy az adatok torzítása (noha kicsi,) valószínűleg nagyobb a mintavételi hibánál. Lásd még T. J. Plewes, „Federal agencies introduce redesigned Current Population Survey,” *Chance* vol. 7, no. 1 (1994) 3541. o.

Elméletben a hányadosbecslés okozhat kisebb torzításokat. A gyakorlatban azonban, kellően nagy minta esetén, az ilyenfajta torzítás elhanyagolható. A felmérés során nem merül fel az egyik probléma: a háztartásméret torzítása (lásd a 19. fejezet 25. jegyzetét). Ennek oka, hogy a kiválasztott háztartások összes 16 éves és idősebb tagja bekerül a mintába, nem csupán egy fő, akit a kérdező otthon talál.

14. Hyman könyvében (19. fejezet 2. jegyzet) szereplő példa alapján.

23. FEJEZET. AZ ÁTLAGOK PONTOSSÁGA

1. A húzások visszatevéssel és visszatevés nélkül is történhetnek. Utóbbi esetben nagyak kell lennie a húzások számának – és a dobozban maradó cédulák számának is; a korrekciós faktorra is szükség lehet a standard hiba kiszámításához (20. fejezet 3. szakasz). Lásd: T. Höglund, „Sampling from a finite population: a remainder term estimate,” *Scandinavian Journal of Statistics* vol. 5 (1978) pp. 69-71. Kis-számú húzás esetén az összeg eloszlása erősen függ a doboz tartalmától, és a normális eloszlástól meglehetősen távol eshet: lásd a 18. fejezetet vagy a 26. fejezet 6. szakaszában a *t-próbáról* mondottakat.
2. *Statistical Abstract*, 1994, a 269-es és a rákövetkező táblázatokban szerepelnek a beiratkozottakra vonatkozó, életkor és nem szerint bontott országos adatok.
3. *Statistical Abstract*, 1994, az 1217-1221. táblázatok szolgálnak információval a lakásállományról az 1990-es népszámlálás és az „American Housing Survey” (amerikai lakásfelmérés) alapján. A definíciók az 1219. táblázat lábjegezetében található. Segítségéért köszönettel tartozunk dr. Harry Scarnnak (Bureau of the Census).
4. *Statistical Abstract*, 1994, a 882. táblázatban szerepelnek az egész országra vonatkozó adatok. A tévé nézésrel töltött időről lásd George Gallup, Jr., *The Gallup Poll: Public Opinion 1991* (Wilmington, Delaware: Scholarly Resources, Inc., 49. o.). Lásd még *Statistical Abstract*, 1994, 884. táblázat.
5. *Statistical Abstract*, 1994, a 1026. táblázatban szerepelnek az egész országra vonatkozó adatok. További információval szolgálhat a lakossági energiafogyasztásról készült felmérés (Residential Energy Consumption Survey) közlekedés blokkja és a Federal Highway Administration közlekedési szokásokra vonatkozó felmérése (Personal Transportation Survey).
6. Lásd Ina V. Mullis et al., *NAEP 1992 Trends in Academic Progress* (Washington, D.C., U.S. Department of Education, Office of Educational Research and Improvement, 1992). Bonyolultan megtervezett minta volt, de 1000 fős egyszerű véletlen mintával számolva nagyjából a tényleges standard hiba jön ki. Lásd az 1. fejezet 1. szakaszát és a 26. fejezet feladatainak megoldását.
7. A felmérés 1976-ban készült. A Carnegie-vizsgálat előző fázisáról lásd: Martin Trow (szerk.), *Teachers and Students* (McGraw-Hill, 1975). Valójában rétegzett mintával dolgoztak. 1992-ben körülbelül 3600 intézmény működött; a beiratkozottak száma átlagosan 4000 körül volt: *Statistical Abstract*, 1994, 275. táblázat. Lásd még: *Digest of Educational Statistics*. A különböző források kategóriái azonban némileg eltérnek egymástól.
8. Lásd a 7. jegyzetet és Trow könyvének 6-7. oldalát. Valójában rétegzett mintavétel történt.

9. E. S. Pearson–J. Wishart, eds., *Student's Collected Papers* (Cambridge University Press, 1942).
10. A. Jensen, „Environment, heredity and intelligence,” *Harvard Educational Review* (1969, 20. o.). Az eredeti szöveget kissé megváltoztattuk.
11. 380 US 202 (1965). A későbbiekben a bíróságok azon az állásponton voltak, hogy az esküdtzéknek reprezentálnia kell a lakosságot; az egyszerű véletlen mintavétel adja meg, mekkora a még megengedhető eltérés a lakossági arányoktól. A precedenst alkotó döntés a *Castaneda* ügy (430 US 482, 1977), különösen a 17. megjegyzés (496); lásd még az *Avery* ügyet (345 US 559, 1953), melyben Frankfurter bíró azt rögzítette, hogy „az igazságszolgáltatásnak nem csupán vaknak, de tudatlannak is kellene lennie ahhoz, hogy egy ilyen esetet a pusztán véletlennek tulajdonítson”. A *Batson* ügyben (476 US 79, 1986) a Legfelsőbb Bíróság leszűkítette az esküdtekkel szembeni megfellebbezhetetlen kifogás jogát.
A „hány szórásnyi az eltérés” módszert mostanra nem csak az esküdtek kiválasztásánál, de a foglalkoztatási diszkriminációval kapcsolatban, sőt trösztellenes ügyekben is alkalmazzák. *McCleskey* szerint (107 S Ct 1756, 1987) a főbenjáró bűnökben ítélező bíróságok – bárhogy is minősítsük ezt – kevésbé hajlandók elfogadni a diszkrimináció statisztikai bizonyítékát. A statisztikai bizonyítékokról különböző megközelítésekben olvashatunk:
B. Black, „Evolving legal standards for the admissibility of scientific evidence,” *Science* vol. 239 (1988) 1508-1512 o.; vol. 241 (1988) 1413-1414. o.
S. Fienberg (ed.), *The Evolving Role of Statistical Assessments as Evidence in the Courts* (Springer-Verlag, 1989).
M. Finkelstein, „The application of statistical decision theory to the jury discrimination cases,” *Harvard Law Review* vol. 338 (1966) 353-356. o.
M. Finkelstein-B. Levin, *Statistics for Lawyers* (Springer-Verlag, 1990).
D. Kaye - D. Freedman, *Reference Guide on Statistics* (Washington, D.C., Federal Judicial Center, 1994).
P. Meier - J. Sacks - S. L. Zabell, „What happened in Hazelwood: statistics, employment discrimination, and the 80% rule,” *American Bar Foundation Research Journal* (1984) 139-186. o.
D. W. Peterson (ed.), „Statistical inference in litigation,” *Law and Contemporary Problems* vol. 46, no. 4 (1983).
D. L. Rubinfeld, „Econometrics in the courtroom,” *Columbia Law Review* vol. 85 (1985) 1048-1097. o.
12. Lee R. Jones et al., *The 1990 Science Report Card* (Washington, D.C., U.S. Department of Education, Office of Educational Research and Improvement, 1992). Lásd 135 és 165. o.

7. rész. Valószínűségi modellek

24. FEJEZET. MODELL A MÉRÉSI HIBÁRA

1. W. J. Youden, *Experimentation and Measurement* (Washington, D.C., National Science Teachers Association, 1962.)
2. A Toledo gyárt ilyen berendezést; ebben négy súlycella van, s a vagonok 9 km/órával mehetnek át a hídmérlegen.
3. Köszönet az adatokért Sam Cohennek és Doug Hale-nek (Energy Information Administration).
4. A hibadoz megkeletkezett cífra volt: a lapok 95%-a a 0 átlagú és 4 mikrogramm szórású normálgörbét követte; a fennmaradó 5% a 0 átlagú és 25 mikrogramm szórású normálgörbét. Két normálgörbére volt szükség: egyre a közepes, és egyre a szélsőséges (magányos) értékek (ún. *outliers*) számára.
5. Talán még jobb lenne a négyzetes közép: a doboz átlagáról feltettük, hogy 0.
6. Az NB10-adatok tapasztalati eloszlása ferde s hosszan elnyúlik az egyik irányba (6. fejezet 2. ábra). Mindazonáltal ennek az eloszlásnak önmagával vett 100-szoros konvolúciója a normálhoz nagyon közeli; a normalitástól vett kisebb eltéréseit nagyon jól leírja egy $1/n$ rendű Edgeworth-sorfejtés.
7. W. J. Youden, „Enduring values”, *Technometrics* vol.14 (1972), 1-11. oldal. Lásd továbbá M. Henion és B. Fischhoff, „Assessing uncertainty in physical constants”, *American Journal of Physics* vol.54 (1986), 791-97. oldal.

8. Az ismételt mérések közötti összefüggést sokszor „megfigyelői torzítás” okozza: a méréseket végző személy, tudattalanul, azt szeretné, hogy a második mérés legyen közel az elsőhöz. A Mérésügyi Hivatal kifinomult ellenintézkedésekkel igyekszik ennek a fajta torzításnak elejét venni. Például úgy, hogy többféle súlyösszeállítás ösztömeget összehasonlítva határozza meg az NB 10 értékét. Az összeállítások a Hivatal által megszabott rendben követik egymást. Az, aki a tényleges méréseket végzi, nem ismeri, milyen kapcsolat van az összeállítások között, így nem alakulhat ki véleménye arról, mikor milyen eredményt „kellene” kapnia.
9. A méréseket Michelson, Pease és Pearson végezte az Irvine Ranch-en, 1929 és 1933 között. Az adatokat a feladatban némileg kerekítettük. Az ő átlaguk, mérföld per másodperce átszámítva, körülbelül 186 270. A mérésekre több szakaszban került sor; bizonyos jegyek arra utalnak, hogy különböző időszakokban a szórás más és más volt.
- A fény sebessége ma lényegében meghatározás: „A Súlyok és Mértékegységek Általános Konferenciája 1983-ban kiadott új hivatalos meghatározása szerint egy méter az a távolság, melyet a másodperc 1/299 792 458 része alatt a fény légüres térben megtesz.” Lásd E. M. Purcell, *Electricity and Magnetism*, 2nd ed. (McGraw-Hill, 1985, Appendix E.).
10. Az idézet forrása R. D. Tuddenham és M. M. Snyder, *Physical Growth of California Boys and Girls from Birth to Eighteen Years* (University of California Press, 1954., 191. old.). Némileg módosítottunk rajta. Mint a szerzők folytatják,
- Az utólagos bölcsesség, néhány éves vizsgálati tapasztalat után, azt mondatja velünk, hogy az első feljegyzett mérési eredménynél és a „leginkább reprezentatív” mérésnél is pontosabb becslése lett volna az elméleti „valódi értéknek” pusztán a mérések [átlaga]“.

25. FEJEZET. VALÓSZÍNŰSÉGI MODELLEK A GENETIKÁBAN

1. Hálával tartozunk Everett Dempster és Michael Freeling (genetika tanszék, University of California, Berkeley) szakértő tanácsaiért (melyek közül néhányat megfogadtunk). Szintén rendkívül sokat segített G. A. Marx és D. K. Ourecky (New York Állam Mezőgazdasági Kísérleti Állomása). Köszönetet mondunk továbbá Ann Lane-nek (University of Minnesota). Két alapvető hivatkozás:
- I. M. Lerner, *Heredity, Evolution, and Society* (W. H. Freeman, 1968).
- E. Rosenberg, *Cell and Molecular Biology* (New York, Rinehart and Winston, 1971).
- Két közvetlenül idekapcsolódó munka:
- J. L. Gould, W. T. Keeton and C. G. Gould, *Biological Science*, 6th ed. (W. W. Norton & Company, 1996).
- M. W. Strickberger, *Genetics*, 3rd ed. (Macmillan, 1985).
- Néhány molekuláris genetikai tankönyv:
- J. D. Watson, N. H. Hopkins, J. W. Roberts, J. A. Steitz and A. M. Weiner, *Molecular Biology of the Gene*, 4th ed. (Benjamin Cummings, 1987).
- B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts and J. D. Watson, *Molecular Biology of the Cell*, 3th ed. (New York: Garland Publishing, 1994).
- Lenyűgöző dolgokat mutat a genetikából a rák biológiája. Két összefoglaló munka:
- J. Cairns, *Cancer: Science and Society* (W. H. Freeman, 1978).
- L. M. Franks and N. Teich (eds.), *Introduction to the Cellular and Molecular Biology of Cancer* (Oxford University Press, 1986).
2. Szigorúan véve a mag egy bizonyos részéé, a *cotyledoné* azaz sziklevéle.
3. A „gén” kifejezés nem egészen egyértelmű, vonatkozhat a DNS-nek egy fehérjét vagy egy fehérje fő láncát kódoló szakaszára, vagy egy ilyen régióban egy konkrét DNS-variánsra. (Talán az „allél” pontosabb megnevezés, ha adott kódpozíción lehetséges DNS-szegmensek közül az egyikről akarunk beszélni.) Feltehetőleg egynél több fehérje határozza meg a mag színét. Ha így van, akkor a tiszta sárga törzsből és a tiszta zöld törzsből egy hely kivételével megegyeznek ezek az allélek - ezen a helyen vannak a szövegben s/s-sel, illetve z/z-vel jelölt párok. Mendel a fenotípusokat meghatározó „entitások”-ról beszél.

4. A hímvarsejtet – a spermiumot – a virágpor tartalmazza, a petesejtet a termő. Szigorú értelemben véve ezek nem sejtmagok és nem is sejtek.
5. Ma úgy tartják, hogy néha – „imprinting” alapján – meg lehet különböztetni, hogy egy gén az anyai vagy az apai szervezetből származik-e. Lásd C. Sapienza, „Parental imprinting of genes”, *Scientific American* vol. 263 (1990. okt.), 52. skk. Lásd még K. Peterson and C. Sapienza, „Imprinting the genome – imprinting, genes, and a hypothesis for their interaction”, *Annual Review of Genetics* vol. 27 (1993), 7–31. o.
6. A gének kromoszómális elhelyezkedését Lamprecht (*Agric. Hortique Genetica*, 1961.) tárta fel. Angol nyelven, nemrégéről, ugyanerről: S. Blixt (ugyanebben a folyóiratban, 1972.). Lásd még: S. Blixt, „The Pea”, *Handbook of Genetics (Plants, Viruses and Protista)*, 2. kötet, 9. fejezet (ed.: R. C. King; Plenum Press, N.Y., 1974.); S. Blixt, „Why didn't Gregor Mendel find linkage?” *Nature* vol. 256 (1975), 206. o.; E. Novitski and S. Blixt, „Mendel, linkage, and synteny”, *Biosciences* vol. 28 (1978) 34–35. o.
7. *Experiments in Plant Hybridisation* (Oliver & Boyd, 1965, 53. o.). A könyv újraközli Mendel eredeti dolgozatát, Fisher néhány megjegyzésével, az *Annals of Science* 1. kötetében (1936.; 115–117. o.) megjelent cikk alapján. Fisher érvelését néhány genetikus erősen támadja; lásd például F. Weiling, „What about R. A. Fisher's statement of the 'too good' data of J. G. Mendel's *Pisum* paper?“, *Journal of Heredity* vol. 77 (1986), 281–283. o. Mindent összevéve, Fisher érvelése meggyőzőnek látszik.
8. A vizsgálatban öt tulajdonság szerepelt, nemcsak ez az egy. Egy kísérlet Mendel újra elvégzett, mert gyengének találta az egyezést. Minden kísérletben 100 növényt használt, így jön ki a szövegbeli 600.
9. „On the correlation between relatives on the assumption of Mendelian inheritance”, *Transactions of the Royal Society of Edinburgh* vol. 52, 399–433. o.
10. *Biometrika* (1903). Az 1,08-as szorzó nagyjából kiegyenlíti a két nem közötti magasságkülönbséget. A dolgozatban közölt egyenletből, kerekítve.
11. A vizsgálatban 1078 család szerepelt, így rendkívül valószínűtlen, hogy a véletlen ingadozás lett volna ekkora.
12. Hogy a (3) egyenletből megkapjuk az (5)-öt, képezzünk feltételes várható értéket az apa magassága alapján; válogatás nélküli párválasztást feltételezve, az anya magassága helyett vehetjük az anyai testmagasságok átlagát. Valójában a szülői testmagasságok között 0,25-ös korreláció volt.
13. A legújabb kutatások arra mutatnak, hogy közönséges sejtosztódáskor a kromoszómák nem mindig másolódnak pontosan. Lásd például J. Marx, „Chromosome ends catch fire”, *Science* vol. 265 (1994), 1656–58. o. Úgy tűnik, hogy a „telomér”-ek – a kromoszóma-végek – megrövidülnek, ha a sejtben nem termelődik a telomeráz nevű enzim. Lásd még C. W. Greider and E. H. Blackburn, „Telomeres, telomerase, and cancer”, *Scientific American* (1996. febr.) 92–97. o., M. Barinaga, „Cells count proteins to keep their telomeres in line”, *Science* vol. 275 (1997), 928. o.
14. E vázlatos ismertetésben figyelmen kívül hagyunk bizonyos bonyodalmakat, így a mutációt, és a genetikusok által átkeresztződésnek (crossover) nevezett jelenséget.
15. A feladatot, módosítva, Strickbergertől vettük; itt a *hüvely színe* a kérdés; ott a *sziklevél színe* (458.skk. o.); a kettő nem ugyanaz.
16. Rasmusson, *Hereditas* vol. 20 (1935). Ez a feladat is Strickbergertől való.

8. rész. Szignifikanciapróbák

26. FEJEZET. SZIGNIFIKANCIAPRÓBÁK

1. „A tapasztalásról”. Idézi Jerome Frank, *Courts on Trial* (Princeton University Press, 1949.) Magyarul: Európa Könyvkiadó, Bp., 1992 (42. és 45. old.)
2. A mikroszimulációs modellek többé-kevésbé az ismertetett módon működnek, adótörvények módosításainak hatásvizsgálatakor ilyeneket használnak. Az adóalanyok viselkedésében bekövetkező változásokat általában nem veszik figyelembe. A párbeszéd természetesen kitalált.
3. Az egyének által fizetendő adóösszegeket a *Statistical Abstract* 1993-as kiadásának 526. táblázata összegzi. Lásd még *Statistical Abstract*, 1988., 489. táblázat; *Statistical Abstract*, 1994., 524. táblázat; *Statistics of Income (SOI) Bulletin* vol. 11., no.4 (1992.tavaszi), Selected Historical and Other Data

- (green pages), 3. táblázat, 131. o. Az 1991. évi egyéni adók átlaga kb. 5000 dollár volt, a szórásuk kb. 20 000 dollár; a különbségeken alapuló becsléseket is közölnek. Amikor az eloszlás valamelyik irányba ennyire elnyúlik, megfontolandó a robusztus eljárások alkalmazása: lásd P. Huber, *Robust Statistics* (John Wiley & Sons, 1981.)
4. Kiegészítő olvasmányok (nehézségek esetére):
 - M. J. Moroney, *Facts from Figures*, 3rd rev. ed. (Penguin Books, 1968.), magyarul: *Számoktól a tényekig* (Gondolat, Budapest, 1970.)
 - J. L. Hodges, Jr. and E. Lehmann, *Basic Concepts of Probability and Statistics*, 2nd ed. (Holden-Day, 1970.)
 - L. Breiman, *Statistics with a View towards Applications* (Houghton Mifflin, 1973.)
 - J. Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. (Belmont, California: Duxbury, 1994.)
 - P. Bickel and K. Doksum, *Mathematical Statistics* (Holden-Day, 1977.)
 - E. Lehmann, *Theory of Point Estimation*, 2nd ed. (Pacific Grove, California: Wadsworth & Brooks/Cole, 1983; újryomás: Chapman and Hall, London)
 - _____, *Testing Statistical Hypotheses*, 2nd ed. (Pacific Grove, California: Wadsworth & Brooks/Cole, 1983; újryomás: Chapman and Hall, London)
 5. Bayesiánusok szempontjából a gyakorisági próba P -értéke jelentősen különbözik a nullhipotézis *a posteriori* valószínűségétől; utóbbinak ugyanis függenie kell (i) a próba erejétől és (ii) a nullhipotézis *a priori* valószínűségétől is. Bővebbet erről: J. Berger and T. Selke, „Testing a point null hypothesis: the irreconcilability of P -values and evidence“, *Journal of the American Statistical Association* vol.82 (1987) 112-39. old.
 6. 1985 és 1991 között az USA-ban 12%-ról 15%-ra nőtt a rugalmas munkaidőben foglalkoztatott dolgozók részaránya. *Statistical Abstract*, 1994, 634. táblázat.
 7. A parajelenségek közé soroljuk az ESP-t, a PK-t, és a *clairvoyance*-t is [ESP: Extrasensory Perception: észlelés érzékelés nélkül; PK: parakinesis: tárgyak mozgatása érintésük nélkül; *clairvoyance*: valaki térben vagy időben távollévő tárgyakat lát – *A ford.*]. A kísérlet leírása megtalálható: C. Tart, *Learning to Use Extrasensory Perception* (University of Chicago Press, 1976). Az alanyok egyike valójában nem csinálta végig az összes fordulót. A „Ten Choice Trainer“-ről is vannak eredmények, lásd az „E” feladatsor 6. feladatát.
 8. A szövegben adott okok miatt nem tekintjük megfelelőnek az ellenhipotézis olyan 0-1 dobozból végzett húzásokkal való modellezését, ahol az 1-esek aránya magasabb 1/4-nél. Ésszerűbb azt mondani, hogy a találatok száma sztochasztikusan nagyobb, mint 7500 húzás összege az

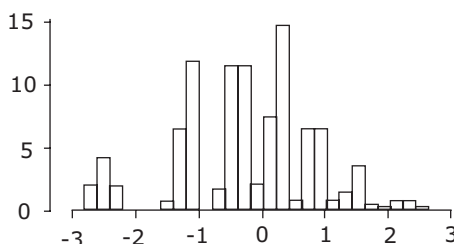
1	0	0	0
---	---	---	---

 dobozból.
 9. Valóban végzett ilyen kísérletet W. Meredith professzor – University of California, Berkeley, Lélektan Tanszék.
 10. A. N. Doob et al., „Effect of initial selling price on subsequent sales“, *Journal of Personality and Social Psychology*, vol.11 (1969), 345-50. old.



11. A Student-anekdotát W. J. Youden, *Experimentation and Measurement* (Washington, 1963) közli.
12. A *t*-próba a legnépszerűbb statisztikai eljárások egyike, igazán sajnáljuk, hogy olyan példán kell bemutatnunk, mely száraz és részben kitalált. (A szövegben előadott történet egészen odáig igaz, ahol *t*-próbát végeznek: valójában nem végeznek.) De egyetlen példára sem akadunk, mely valós, érdekes, és ugyanakkor elfogadható is lett volna. Problémánk a következő volt. A *t*-próbával szignifikanciaszinteket számítottunk. Kis mintáknál azonban jelentősen eltérítheti a számításokat, ha nem teljesül a normalitás.

Szemléltetésképpen bemutatjuk az alábbi ábrán, milyen lenne annak a *t*-statisztikának az elméleti hisztogramja, melyet a $\boxed{-3} \quad \boxed{-2} \quad \boxed{5}$ dobozból végzett 10 - véletlenszerű, visszatevéses - húzás alapján számítanánk.



A hisztogramot pontosan kiszámítottuk úgy, hogy figyelembe vettük mind a

$$\binom{12}{2} = 66$$

lehetőséget, ahogyan 10 elemet 3 csoportba lehet sorolni; minden ilyen csoportosításnak kiszámítottuk a valószínűségét, és hogy milyen *t*-értéket adna. A *t*-statisztika értéke az adódó csoportosítások közül 3-ban nincs értelmezve - de ezek együttes súlya sem éri el a 10^{-5} -t; további 2%-nál az érték a $[-3; 3]$ intervallumon kívülre esik. A példa könnyen módosítható úgy, hogy a sűrűségfüggvény sima legyen; az eltolás paraméterként kezelhető. A szükséges kombinatorika megtalálható: W. Feller, *An Introduction to Probability Theory and its Applications*, 1. kötet, 3-dik kiadás (John Wiley & Sons, 1968., II.4. szakasz). Magyarul: *Bevezetés a valószínűségszámításba és alkalmazásaiba*. Műszaki Könyvkiadó, Budapest, 1978.

Ez azt jelenti, hogy a *t*-próba alkalmazásához a hibaeloszlásról elég sokat kellene tudnunk, ugyanakkor, amikor a szóródás mértékéről viszont nincs világos elképzelésünk - mert ha volna, akkor nem *t*-próbát kellene alkalmaznunk. Nagy mintáknál nem ilyen fontos a normalitástól való eltérés - olyankor a *t*-statisztika a normálgörbét követi (lásd 18. fejezet, 9.ábra). Részben erre gondolnak a statisztikusok, amikor „a *t*-próba robusztusságáról” beszélnek. De ilyenkor a Student-eloszlások már a normálhhoz tartanak. A mi fogalmaink szerint itt a *z*-próba robusztusságáról van szó, nem a *t*-próbáérol. Két hivatkozás:

H. D. Posten, „The robustness of the one-sample *t*-test over the Pearson system“, *Journal of Statistical Computation and Simulation* vol.9. (1979), 133-49. old.

E. Lehmann and W.-Y.Loh, „Pointwise vs. uniform robustness of some large sample tests and confidence intervals“, *Scandinavian Journal of Statistics* vol.17 (1990), 177-87. old.

A *z*- és a *t*-próbák eredményeit jelentősen befolyásolhatják a függetlenségtől való - akár kis mérvű - eltérések. Lásd alább, 14-es jegyzet.

13. Mostani céljaink szempontjából ez csak konvenció: behelyettesíthetnénk a $\sqrt{n/(n-1)}$ -es tényezőt a Student-görbéből nyert szorzóba. Bizonyos kontextusokban azonban előnyben részesítik a populáció szórásának becsléseként a korrigált szórást a korrigálatlan szórással szemben: a korrigált szórás négyzete torzítatlan becslése a populációs szórásnégyzetnek, ami nem mellékes, ha sok kis elemszámú mintából számított szórásbecslést kell összevonni (pooling).

14. A görbe egyenlete

$$y = \text{konstans} \left(1 + \frac{t^2}{d} \right)^{-(d+1)/2}$$

$$\text{konstans} = 100\% \frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi d} \Gamma\left(\frac{d}{2}\right)}$$

d = szabadságfok

Γ = az Euler-féle gamma-függvény

A *t*-próba szigorú matematikai megalapozását R. A. Fisher végezte el, aki azt is megmutatta, hogy az eljárás jó közelítést ad akkor is, ha a hibák nem pontosan követik a normálgörbét: egy kis eltérés a normalitástól még nem baj. Ezt a kismintás tulajdonságot szintén „robustusság”-nak nevezik. (De lásd a 12. jegyzetet.)

15. Ha a dobozbeli számkártyák a normálgörbét követik, akkor a húzások összegére vonatkozó hisztogram is ezt teszi - már egészen kis számú húzásnál. Szaknyelven: normálgörbe önmagával való konvolúciója újra normálgörbét ad. Olyankor, amikor a dobozbeli számok eloszlása ismert de nem normális, konvolúciók segítségével ki lehet számítani a húzások összegére vonatkozó elméleti hisztogramot.
16. A. W. Astin, K. C. Green, W. S. Korn and M. Schalit, *The American Freshman: National Norms for Fall 1986* (American Council on Education, UCLA, 1987.)
17. Az 1969-es Zeisel-cikk megjelenése után Ford bíró által összeállított első esküdtszoportban 24% nő volt. Hivatkozások:

Hans Zeisel, „Dr. Spock and the case of the vanishing woman jurors“, *University of Chicago Law Review* vol.37 (1969), 1-18. o.

_____, „Race bias in the administration of the death penalty: the Florida experience“, *Harvard Law Review* vol.95 (1981), 456-468. o.
18. S. C. Truelove, „Therapeutic trials“ in L. J. Witts (eds.): *Medical Surveys and Clinical Trials* (Oxford University Press, 1959). A randomizáció „vakká tételéről“ l.: T. C. Chalmers, P. Celano, H. S. Sacks and H. Smith, Jr., „Bias in treatment assignment in controlled clinical trials“, *New England Journal of Medicine* vol.309 (1983), 1358-1361. o.
19. Az „irodalom olvasásáról“ *Statistical Abstract*, 1994., 412. táblázata közöl országos adatokat. Az iskolázottsági adatok a Rendszeres Népszámlálás 1993. márciusi felvételéből származnak. A piacuttatás kitalált.
20. Az adatok New York önkormányzatának egészségügyi osztályától származnak. Mi Sandy Zabelltól (a statisztika professzora, Northwestern University) kaptuk őket. Hivatkozás: A. J. Izenman and S. L. Zabell, „Babies and the blackout: the genesis of a misconception“, *Social Science Research* vol.10 (1981), 282-299. o. Valószínűleg az történt, hogy a *New York Times* elküldte egy riporterét néhány kórházba augusztus 8-án, hétfőn és augusztus 9-én, kedden, kilenc hónappal az áramszünet után. A kórházak beszámoltak arról, hogy szülészeti osztályaiknak a szokásosnál több a dolguk - talán amiatt, ami általános, hogy a hétfőgék békések, a hétfő és a kedd forgalmasak szoktak lenni. Ezek a „felfedezések“ kerültek azután címlapra szerdán, 1966. augusztus 10-én, „Megugrott születésszám kilenc hónappal a Nagy Áramszünet után“ szalagcímmel. Úgy tűnik, ez állhat a népességreobbanás mítosza mögött.
21. S. M. Stigler, „Eight centuries of sampling inspection: the Trial of the Pyx“, *Journal of the American Statistical Association* vol.72 (1977), 493-500. old. A mi változatunk kissé stilizált; lehet, hogy nem az egyszerűen véletlen mintavétel a pontos modell.
22. M. Rosenzweig, E. L. Bennett and M. C. Diamond, „Brain changes in response to experience“, *Scientific American* vol 210 (1964), 22-29. old. A kísérlet valójában nem párokkal, hanem hármassokkal dolgozott, ezek tagjait sorolták véletlenszerűen az ingerben gazdag, a standard és az ingersejény csoportba.

27. FEJEZET. TOVÁBBI PRÓBÁK AZ ÁTLAGRA

1. Tegyük fel, hogy X és Y véletlen változók, szórásnégyzetük rendre σ^2 , illetve τ^2 , korrelációjuk ρ . Ekkor $\text{var}(X - Y) = \sigma^2 + \tau^2 - 2\rho\sigma\tau$. Ha $\rho = 1$, akkor a „véletlen hibák“ (azaz az eltérések a várható értéktől) szükségszerűen mind azonos előjelűek, így a különbségben kiejtik egymást – ekkor a különbség standard hibája $|\sigma - \tau|$. Ha $\rho = -1$, akkor a hibák erősítik egymást, a standard hiba $\sigma + \tau$. A függetlenség megfelel annak, amikor $\rho = 0$, ilyenkor a két véletel között középtájon van a standard hiba: $\text{SH} = \sqrt{\sigma^2 + \tau^2}$.
2. Ina V. Mullis et al., *NAEP 1992 Trends in Academic Progress* (Washington D.C., U.S. Department of Education, Office of Educational Research and Improvement, 1992). Az átlagok az eredetiek; az eredeti minta lényegesen nagyobb volt, és erősen struktúrált; a szórásértékek kerekítettek; a standard hibák a valódiakhoz közeleiek. Az ezekben a standardizált tesztekben elért teljesítménypontszámok, úgy tűnik, 1975 körül jöttek ki az addigi zuhanórepülésből. Nem használhattuk most a próbastatisztika nevezőjében az összevont mintákból számított (pooled) szórás, az egyes populációk szórásai ugyanis – a nullhipotézis teljesülése esetén is – lehetnek különbözők.
3. Ilyen jellegű feladatokra három további próba is széles körben használatos, a szövegben ismertetetten kívül:

(i) A két dobozban az 1-esek részaránya a nullhipotézis szerint egyforma, és, összevonva a két mintát, a következőképpen becsülhető:

$$\frac{107 + 132}{200 + 300} = \frac{239}{500} \approx 48\%$$

Ebből kiindulva becsülhető a dobozok egyforma szórása:

$$\sqrt{0,48 \cdot 0,52} \approx 0,50.$$

Ennek az összevont szórásnak a segítségével számíthatjuk ki a próbastatisztika nevezőjében szereplő standard hibát. (Nem alkalmas azonban az összevont szórás más célokra – pl. a különbségre vonatkozó konfidenciaintervallumok számítására.)

(ii) A Fisher-féle exakt próba abból indul ki, hogy hány 1-es volt a két mintában összesen. Próbastatisztikája az 1-esek száma – mondjuk – az első mintában; a mintaelemszámokat előre adottnak tekintti. A nulleloszlás hipergeometrikus.

(iii) Számíthatunk χ^2 statisztikát a 2×2 -es táblából, és összevethetjük a χ^2 eloszlással.

Ha rögzítjük, hogy hány 1-es van a két mintában összesen, akkor az (i)-es próbastatisztika monoton függvénye az első mintában lévő 1-esek számának; tehát az (i)-es és a (ii)-es próba ekvivalens. Ez azt jelenti, hogy a hipergeometrikust közelítettük normálissal, ami – eléggé nagy minták mellett – elfogadható. Hasonlóképpen (iii) is ekvivalens (i)-gyel és (ii)-vel; sőt, $\chi^2 \equiv Z_1^2$, ahol Z_1 az összevont szórásból számított z-statisztika. Ha – mint a szövegben – Z a külön szórásokból számított z-statisztika, akkor $Z^2 > Z_1^2$. Azonban, mint alább megmutatjuk, ha rögzítjük a két mintában összesen előforduló 1-esek számát, akkor Z monoton függvénye az első mintában lévő 1-esek számának. Ez azt jelenti, hogy az általunk ismertetett próba ekvivalens az (i)-(ii)-(iii) próbákkal.

Vezessünk be néhány jelölést. Jelölje ξ a fejek számát n dobásból egy p -érmével, $\hat{p} = \xi/n$. Jelölje ζ a fejek számát m dobásból egy q -érmével, $\hat{q} = \zeta/m$. A két érme független. A nullhipotézis szerint $p = q$; az ellenhipotézis szerint $p \neq q$. Rögzítjük $s = \xi + \zeta$ -t. Legyen $r = s/(m + n)$. Ezt az r -et is rögzítjük.

Az (i)-es esetben a variancia

$$v = r(1-r) \left(\frac{1}{n} + \frac{1}{m} \right)$$

a próbastatisztika pedig

$$Z_i = (\hat{p} - \hat{q}) / \sqrt{v}$$

A mi próbastatisztikánkhoz a variancia a

$$w = \frac{\hat{p}(1-\hat{p})}{n} + \frac{\hat{q}(1-\hat{q})}{m}$$

képlettel adódik, a próbastatisztika pedig

$$Z = (\hat{p} - \hat{q}) / \sqrt{w}$$

m -et, n -et és $r = s/(m+n)$ -et rögzítettnek tekintjük; hallgatólagosan feltételezzük, hogy $0 < \hat{p}, \hat{q}, r < 1$. Továbbá vezessünk be egy új változót, legyen $x = \hat{p} - r$, s így $\hat{q} = r - (n/m)x$. Magától értetődik, hogy x -nek csak véges sok fajta értéke lehet. (Később azonban kényelmesebb lesz úgy elgondolni x -et, mintha a minimumától a maximumáig terjedő teljes intervallumon végigfutna.) Mármost

$$\hat{p}(1-\hat{p}) = r(1-r) + (1-2r)x - x^2$$

$$\hat{q}(1-\hat{q}) = r(1-r) - \frac{n}{m}(1-2r)x - \frac{n^2}{m^2}x^2$$

A két variancia között ez a kapcsolat:

$$w = v + bx - cx^2$$

ahol

$$b = \left(\frac{1}{n} - \frac{n}{m^2} \right) (1-2r), c = \frac{1}{n} + \frac{n}{m^2}.$$

Végül a próbastatisztikánk:

$$Z = \frac{m+n}{m} x / \sqrt{v+bx-cx^2}$$

Magától értetődik, hogy w , v és c mind pozitívak; b pozitív és negatív is lehet. Z -nek mint x függvényének a monotonítása az alábbi lemma következménye.

Lemma. Legyen $v, c > 0$, b valós. Tekintsük x -et csak azon az intervallumon, amelyen $v+bx-cx^2 > 0$. Legyen

$$f(x) = x / \sqrt{v+bx-cx^2}$$

Ekkor az $f(x)$ függvény x szerint monoton nő.

Bizonyítás. Ha $b \leq 0$, az állítás triviális; legyen tehát $b > 0$. Ekkor

$$\frac{df}{dx} = \frac{2v+bx}{2(v+bx-cx^2)^{3/2}} > 0$$

4. Országos adatokat közöl:

A. W. Astin, K. C. Green, W. S. Korn and M. Schalit, *The American Freshman: National Norms for Fall 1986* (American Council on Education, UCLA, 1987.),

E. L. Dey, A. W. Astin, and W. S. Korn, *The American Freshman: Twenty-Five Year Trends, 1966-1990* (Higher Education Research Institute, UCLA, 1991).

5. „Literacy among 12-17 years“, *Vital and Health Statistics*, series 11., no. 131. (Washington, D.C., 1973). A mintavételi terv a Rendszeres Népszécfelméréséhez hasonlított, a standard hibák becslésére a kutatók a mintafelezéses eljárást használták. A példában szereplő nagyságú egyszerű véletlen mintákkal körülbelül akkorák a hibák, amekkorák a valódi kutatásban voltak.

6. „Intellectual development of children by demographic and socioeconomic factors“, *Vital and Health Statistics*, series 11., no. 110. (Washington, D.C., 1971). A standard hibákat illetően l. az 5. jegyzetet. A gyermekek tesztpontszámái és a szülők iskolázottsága között 0,5-ös volt a korreláció – a szülők jövedelmét állandó szinten tartva ez 0,3-ra csökkent. „Nagyváros“ 3 milliós vagy nagyobb lakosságot jelent. Legjobban az 1–3 milliós városokból való gyermekek teljesítettek: átlag 28 pontot értek el.
7. *Statistical Abstract*, 1994., 210-es táblázat. U.S. National Institute on Drug Abuse, National Household Survey on Drug Abuse. A százalékokat kissé megváltoztattuk. A mintavételi terv erősen struktúrált volt; a példabeli mintaelemszámokkal a ténylegeshez közeli standard hibák adódnak. Lásd még *Morbidity and Mortality Weekly Report*, 1995. július 21.; lehet, hogy a kábítószerhasználat megint növekszik.
8. A forrást lásd a 4-es jegyzetben.
9. *Statistical Abstract*, 1994, 287. táblázat.
10. C-vitaminnal folytatott valódi kísérletekre, s velük kapcsolatban a csoportbasorolás vakká tételével kapcsolatos érdekes tudnivalókra található hivatkozások a 2. fejezet 12. jegyzetében.
11. Néha a nullhipotézis gyengébb változatát használják: a válaszok átlaga egyforma a két kezelés esetében. A szigorúbb nullhipotézisnél tulajdonképpen csak egy-egy szám van a lapokon (mindkét mezőbe ezt az egyet írjuk). Sokszor alkalmazható a szigorúbb verzió, melynek mindig sajátos bájt ad az egyszerűsége. Valószínűbb viszont az enyhébb változat pl. olyankor, mikor az új kezelés egyes pácienseknek árt, másoknak használ. Egyes ellenhipotézisek pedig megkínávják, hogy minden alanyt két számmal jellemezzünk – hogyan reagál a kezelésre és hogyan a placebóra. Továbbiak erről a 12. jegyzetben.
12. Tekintsünk egy klinikai tesztet, mely az A és B kezeléseket hasonlítja össze. A gyengébbik fajta nullhipotézissel dolgozunk (lásd a 11. jegyzetet); és olyan ellenhipotézissel, mely a reakciókra nézve semmit sem köt ki. Tegyük fel, hogy N alany van, $i = 1, \dots, N$ sorszámokkal. Legyen x_i az i -edik alany reakciója az A kezelésre; y_i pedig a B kezelésre. Mindegyik i -re vagy x_i -t, vagy y_i -t figyelhetjük meg, mindkettejüket nem. Legyen

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i & \bar{y} &= \frac{1}{N} \sum_{i=1}^N y_i \\ \sigma^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 & \tau^2 &= \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 \\ \text{cov}(x, y) &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})\end{aligned}$$

Ez a modell kellően hajlékony ahhoz, hogy a nullhipotézis gyenge alakját (11. jegyzet) is, és az ellenhipotézis alanyról-alanyra fellépő heterogenitását is kezelni tudja. Így például az A és B kezelés közötti átlagos különbség – a kísérletben résztvevő összes alanyra elvégezve az átlagolást – $\bar{x} - \bar{y}$. Ez az „átlagos oksági hatás“ azt méri, mekkora a különbség aközött, ha mindenkit az A, és ha mindenkit a B kezelésnek vetnének alá. Sokszor az átlagos oksági hatás a legfontosabb paraméter. Becsülni tudjuk, bár egyetlen alanyról sincs rá mód, hogy mindkét választ megfigyeljük. Sőt, \bar{x} , \bar{y} , σ^2 , és τ^2 is becsülhető; $\text{cov}(x, y)$ -t viszont nem tudjuk a mintabeli kovarianciával becsülni.

A kezelt és a kontrollcsoportbeli válaszokat gyakran modellezik két eltérő p paraméterű és független binomiállissal, vagy két eltérő μ paraméterű és független normálissal. E modellek kevésbé látszanak valószínűnek. A két mintaátlag függetlensége általában nem teljesül, s nincs ok feltételezni, hogy az egyes csoportokon belül az alanyok felcserélhetők. A randomizáció ezeket a feltevéseket *nem* támasztja alá, csak *mint csoportokat* teszi őket összehasonlíthatóvá. Emiatt nem teljesülnek pl. a *t*-próba alkalmazásának elméleti előfeltételei. Meglepő – és biztató –, hogy a hagyományos próbatastiztikák permutációs eloszlásai többé-kevésbé egybeesnek a modellünkre alapozott eloszlásokkal, legalábbis a szóban forgó összefüggésben.

Szemléltetésül nézzük, mekkorának adódik az ellenhipotézis szerint a permutációs elrendezésben $\bar{X} - \bar{Y}$ varianciája. Legyen S az $\{1, \dots, N\}$ halmaz n elemű véletlen részhalmaza; ez a csoport kapja az A kezelést, tehát $i \in S$ esetén x_i -t figyeljük meg. Legyen T az $\{1, \dots, N\}$ halmaz egy m elemű vé-

letlen, S -től diszjunkt részhalmaza. Ez a csoport részesül a B kezelésben, tehát $j \in T$ esetén y_j -t figyeljük meg. Az \bar{x} és \bar{y} populációs átlagokat az

$$\bar{X} = \frac{1}{n} \sum_{i \in S} x_i \quad \text{és} \quad \bar{Y} = \frac{1}{m} \sum_{j \in T} y_j$$

mintaátlagokkal becsljük. Kombinatorikai számításokból

$$\text{var } \bar{X} = \frac{N-n}{N-1} \frac{\sigma^2}{n} \quad \text{var } \bar{Y} = \frac{N-m}{N-1} \frac{\tau^2}{m}$$

$$\text{cov}(\bar{X}, \bar{Y}) = \frac{1}{N-1} \text{cov}(x, y)$$

Innen

$$\begin{aligned} \text{var}(\bar{X} - \bar{Y}) &= \frac{N-n}{N-1} \frac{\sigma^2}{n} + \frac{N-m}{N-1} \frac{\tau^2}{m} + \frac{2}{N-1} \text{cov}(x, y) = \\ &= \frac{N}{N-1} \left(\frac{\sigma^2}{n} + \frac{\tau^2}{m} \right) + \frac{1}{N-1} [2\text{cov}(x, y) - \sigma^2 - \tau^2] \leq \\ &\leq \frac{N}{N-1} \left(\frac{\sigma^2}{n} + \frac{\tau^2}{m} \right) \end{aligned}$$

mivel $\text{cov}(x, y) \leq \sigma\tau \leq 2\sigma\tau - \sigma^2 - \tau^2 \leq 0$. A szövegben jelzett „konzervatív becslés“ $\sigma^2/n + \tau^2/m$. σ^2 -et és τ^2 -et a gyakorlatban a mintabeli varianciákkal becsljük.

Az előjelek talán zavarbaejtőek. Általában, az összes alanyra nézve, azt váránk, hogy x és y pozitívan korreláljon. De ha nagyon sok magas x értékű alany kerül az A kezelésbe, akkor csak kevés magas y értékű marad a B kezelésre. Emiatt az \bar{X} és \bar{Y} mintaátlagok negatívan korrelálnak. Elvben $\text{cov}(x, y)$ -nak közel kellene lennie a saját felső korlátjához, $\sigma\tau$ -hoz, legalábbis, amikor x és y az összes alanyra nézve magasan korrelál. Ilyenkor, kellően nagy mintákra, a „konzervatív becslés“-nek elég pontosnak kell lennie. A szövegbeli szigorú nullhipotézis szerint $x \equiv y$. Ekkor $\sigma = \tau$, a nullhipotézis fennállása esetén a számítás egzaktt. Lásd az alábbi 15. jegyzetet is. Természetesen ha N nagy n -hez és m -hez képest, akkor \bar{X} és \bar{Y} jószertint függetlenek; a „konzervatív becslés“ ekkor is majdnem pontos.

A többi változók hatását az alábbi módon vehetjük számításba. Jelölje a kezelésfajta η . Jelölje a válaszreakcióra hatással bíró egyéb változók értékét ω . Tegyük fel, létezik olyan f függvény, hogy $f(i, \eta, \omega)$ megadja az i -edik alany válaszreakcióját a kezelésre. Legyen ρ a besorolás szerinti változó: $\rho(i) = A$, ha az i -edik alanyt az A kezeléshez sorolták, és ugyanígy B-re. Feltesszük, hogy ρ és ω független: adott ω mellett egyformán valószínű az alanyok minden lehetséges ρ -beosztása egy n számosságú, az A kezelésbe kerülő S , és egy m számosságú, a B kezelésbe kerülő T csoportra. A randomizálásnak, a vak kísérleti elrendezésnek stb. éppen az a céljuk, hogy ezt a feltételt biztosítsák. Ekkor érvelésünk minden ω -ra külön is elvégezhető úgy, ha

$$x_i = f(i, A, \omega) \quad \text{minden } i \in S\text{-re}$$

$$y_j = f(j, B, \omega) \quad \text{minden } j \in T\text{-re}$$

A modell valamelyest visszautal Neyman korai, mezőgazdasági kísérletekkel kapcsolatos munkáira.

Néhány forrás:

- J. Neyman, „Sur les applications de la théorie des probabilités aux expériences agricoles: Essai des principes“, *Roczniki Nauk Rolniczki* vol.10. (1923) 1–51. old, lengyelül; angol fordítás (D. Dabrowska and T. Speed): *Statistical Science*, vol.5 (1991) 463–480. o.
- H. Scheffé, „Models in the analysis of variance“, *Annals of Mathematical Statistics*, vol.27 (1936).
- J. L. Hodges, Jr. and E. Lehmann, *Basic Concepts of Probability and Statistics*, 2nd ed. (Holden-Day, 1970, 9.4)
- D. Rubin, „Estimating causal effects of treatments in randomized and nonrandomized studies“, *Journal of Educational Psychology*, vol.66 (1974), 688–701. o.
- J. Robins, „Confidence interval for causal parameters“, *Statistics in Medicine*, vol.7 (1988), 773–785. o.
- P. Holland, „Causal inference, path analysis, and recursive structural equations models“, *Sociological Methodology 1988*, ed.: C. Clogg (Washington, D.C., American Sociological Association; 13. fejezet).
- L. Dümbgen, „Combinatorial stochastic processes“, *Stochastic Processes and their Applications*, vol.52 (1994), 75–92. o.

Néhány apró technikai részlet: (i) a releváns centrális határeloszlástétel a visszatevés nélküli mintavételre vonatkozó (23. fejezet 1. jegyzet). (ii) A *t*-eloszlás kis mintáknál sem feltétlenül ad a normálisnál jobb közelítést: nem teljesülnek a *t*-próba kiinduló feltételei.

13. A. Tversky and D. Kahnemann, „Rational choice and the framing of decisions“, *Journal of Business* vol.59, no.4., 2. rész (1986), S251–78. oldalak. Lásd még D. Kahneman and A. Tversky, „On the reality of cognitive illusions“, *Psychological Review* vol.103 (1996), 582–596. o. (részletes kifejtéssel).
14. B. J. McNeil, S. G. Pauker, H. C. Sox, Jr. and A. Tversky, „On the elicitation of preferences for alternative therapies“, *New England Journal of Medicine* vol.306 (1982), 1259–1262. o.
15. A 3. és 12. jegyzetekben megkezdett téma tárgyalását folytatjuk. Összesen $80 + 87 = 167$ kísérleti személy volt (1. táblázat). Közülük $40 + 73 = 113$ -an álltak a műtét pártján; a maradék 54 sugárkezelést javasolta. A szigorú nullhipotézis szerint $x \equiv y$, így $\sigma \equiv \tau$, s mindkettő számítható az adatokból. Valóban, a nullhipotézis szerint a műtét pártján álló orvosok százalékaránya $113/167 \cdot 100\% \approx 68\%$. Így

$$\sigma = \tau \approx \sqrt{0,68 \cdot 0,32} \approx 0,47$$

Ehhez hasonlóan az \bar{X} és \bar{Y} közötti kovariancia is pontosan kiszámítható. A kifejezés eléri felső korlátját, a $\sigma\tau = \sigma^2$ értéket, mivel x és y között – az összes kísérleti személyekre nézve – 1 a korreláció. Mármost

$$\text{var}(\bar{X} - \bar{Y}) = \frac{N}{N-1} \left(\frac{1}{n} + \frac{1}{m} \right) \sigma^2$$

A próbastatisztika két alakja (összevont, illetve külön szórások – lásd a 3. jegyzetet) gyakorlatilag azonos. Ha például a nullhipotézis által meghatározott modellt nézzük, a két statisztika értéke közötti eltérés négyzetes közepe mindössze 0,013. Továbbá, a normális közelítés is meglehetősen pontos: mindkét próbastatisztika körülbelül 4,8% valószínűséggel haladja meg abszolút értékben a 2-t (normáleloszlás esetén 4,6% volna ez a valószínűség).

16. D. Kahneman and A. Tversky, „Choices, values and frames“, *American Psychologist*, vol.39 (1984), 341–50. old.
17. A valóságban ennél kicsit bonyolultabban zajlott a randomizáció. Az oltás 3 egymás utáni injekcióból állt, ennek megfelelően a placebo-csoport is 3 injekciót kapott (placebót). Az injekciós fiolákból 6-ot csomagoltak egy dobozba; e 6 közül 3-ban, közös kódszám alatt, oltóanyag volt, másik 3-ban, szintén közös kódszám alatt, placebo. Mindegyik fiola 10 injekcióra való folyadékot tartalmazott.

Amikor eljött az első injekció beadásának ideje, kivettek a dobozból egy fiolát, s a benne lévő anyaggal beoltottak 10 gyermeket; a kutató feljegyezte a gyermekek neve mellé a fiola kódját; ez a 10 gyermek a doboz ugyanilyen kódú másik 2 fiolájából kapta a 2. és a 3. injekciót. A következő 10 gyer-

mek a dobozban lévő másik 3 fiola egyikéből kapta a maga első injekcióját (ezen másik kódszám volt, mint a korábban kivett első fiolán); az ő nevük mellé ennek a fiolának a kódszámát jegyezték fel; a 2 ugyanilyen kódú fiolából kapták a hátralévő 2 injekciót.

Az eljárást tekinthetjük úgy, hogy a gyermekeket 10-es csoportok párhozba sorolták; s aztán minden párról érmefeldobás döntött; egyik csoport kapta a kezelést, másik lett a kontroll, 50-50 százalékos eséllyel. Ha elfogadjuk azt a - plauzibilis - feltevést, hogy a gyermekbénulásos esetek között nem volt olyan kettő, aki ugyanabból a dobozból kapta volna az injekciót, akkor a szövegben ismertetett számítás pontos. Másként módosítani kell a számításon. Ezt a konkrét példát általában kétmintás z-próbával szokták elemezni, a 10-es csoportba sorolás figyelembevétele nélkül (1. fejezet, 2. jegyzet). Ezt tesszük mi is.

18. D. N. Kershaw and F. Skidmore, *The New Jersey Graduated Work Incentive Experiment* (Princeton: Mathematica, 1974). A tőle kapott segítségért szeretnénk köszönetet mondani Rob Hollisternek (Swarthmore). A tényleges kísérleti terv bonyolultabb volt; és jelentős mértékű volt a lemorzsolódás. Lásd még: L. Neuberger, „Distorted transmission”, *Theory and Society* vol.17 (1988), 487-525. old.
19. D. Ravitch and C. E. Finn, Jr., *What Do Our 17-Year-Olds Know?* (Harper & Row, 1987., 52. old.). A felismerési hányad a Szovjetunió esetében volt a legmagasabb.
20. George Gallup, Jr., *The Gallup Poll* (Wilmington: Scholarly Resources, 1987.)
21. Hivatkozások:

K. Gray-Donald, M. S. Kramer, S. Munday et al., „Effect of formula supplementation in the hospital on duration of breast-feeding: a controlled clinical trial”, *Pediatrics*, vol.75 (1985), 514-518. o.

K. Gray-Donald and M. S. Kramer, „Causality inference in observational vs. experimental studies: an empirical comparison”, *American Journal of Epidemiology*, vol.127 (1988), 885-92. old.

A kutatók a kontrollós kísérletet megelőzően egy megfigyeléses vizsgálatot is elvégeztek; mindkét részt vevő szülészeti a szokásos táplálékkiegészítési gyakorlatát folytatta. Ahogy a korábbi vizsgálatokban, itt is erős negatív kapcsolat mutatkozott a szülészeten adott táplálékkiegészítés, és a továbbszoptatás között. Technikai értelemben véve nem volt random, hogy ki melyik szülészetre került: a jelentkező anyukát arra az osztályra küldték, ahol volt szabad ágy; ezt a beosztást a kórháznak a kísérletbe be nem vont személyzete végezte. A publikálatlan adatokhoz a kutatók szívességéből jutottunk.

22. Legyenek (X_i, Y_i) független, azonos együttes eloszlású valószínűségi változó párok, $E(X_i) = \alpha$, $\text{var } X_i = \sigma^2$, $E(Y_i) = \beta$, és $\text{var } Y_i = \tau^2$; legyen az X_i és Y_i korrelációja ρ , tehát $\text{cov}(X_i, Y_i) = \rho\sigma\tau$. Legyen $\bar{X} = (X_1 + \dots + X_n)/n$ és $\bar{Y} = (Y_1 + \dots + Y_n)/n$. A mintaátlagok korrelálnak, $\text{var}(\bar{X} - \bar{Y}) = v/n$, ahol

$$v = \sigma^2 + \tau^2 - 2\rho\sigma\tau,$$

A v variancia becslése a mintaadatokból

$$\hat{v} = \hat{\sigma}^2 + \hat{\tau}^2 - 2r\hat{\sigma}\hat{\tau},$$

ahol

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \hat{\tau}^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

és r a mintabeli korrelációs együttható,

$$r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) / \hat{\sigma}\hat{\tau}.$$

A z-próba az $(\bar{X} - \bar{Y}) / \sqrt{\hat{v}/n}$ statisztikával dolgozik.

Vessük ezt össze az $X_i - Y_i$ különbségeken alapuló z-próbával. Könnyen látható, hogy

$\bar{X} - \bar{Y} = \overline{X - Y}$, mivel az utóbbi $\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)$. Az $X_i - Y_i$ különbségek függetlenek és egyforma eloszlá-

súák, $E(X_i - Y_i) = \alpha - \beta$ várható értékkel és $\text{var}(X_i - Y_i) = \sigma^2 + \tau^2 - 2\rho\sigma\tau = v$ varianciával; természetesen $\text{var}(\overline{X} - \overline{Y}) = v/n = \text{var}(\overline{X} - \overline{Y})$, a fenti v értékre. v -nek a különbségeken alapuló természetes becslése

$$\frac{1}{n} \sum_{i=1}^n [(X_i - Y_i) - (\overline{X} - \overline{Y})]^2 = \hat{v}$$

ami egyezik a varianciának az adatpárokon alapuló becslésével. (Némi számolás kell az egyenlőség belátásához.) Következésképpen a párok alapján számolt z -statisztikának és a különbségek alapján számolt z -statisztikának meg kell egyeznie.

23. George Gallup, Jr., *The Gallup Poll: Public Opinion 1992*. (Wilmington: Scholarly Resources). Lásd a 118. oldalt.
24. Forrás: lásd 23. jegyzet. Az eredeti kérdés: „Milyennek értékeli e különböző foglalkozási ágakba tartozó emberek becsületességét és erkölcsi nivóját - nagyon magasnak, magasnak, átlagosnak, alacsonynak, nagyon alacsonynak?” A „nagyon magas vagy magas” osztályzatok százalékarányai:

Gyógyszerészek	66%
Orvosok	52%
Egyetemi oktatók	50%
Újságírók	27%
Építési vállalkozók	19%
Jogászok	18%
Kongresszusi képviselők	11%
Autókereskedők	5%

25. A. Tversky and D. Kahneman, „The framing of decisions and the psychology of choice”, *Science* vol.211 (1981), 453–458. old.
26. *The Third National Mathematics Assessment: Results, Trends and Issues* (Princeton: ETS/NAEP, 1983.) A kérdés a felmérésből való, az eredmények körülbelül megfelelnek az eredetieknek; a zseb-számológépes csoport lényegesen rosszabbul teljesített. A beszámolóból azonban nem derül ki világosan, kísérletről vagy megfigyeléses vizsgálatról volt-e szó.
27. P. H. Rossi, R. A. Berk and K.J. Lenihan, *Money, Work and Crime: Experimental Evidence* (San Diego: Academic Press, 1980.; főként az 5.1-es táblázat). A vizsgálatot 1976-ban végezték. A kísérleti tervet egyszerűsítettük, de semmilyen lényegi szempontból nem változtattunk rajta; hasonlóképpen, a százalékokat is módosítottuk valamelyest, hogy a szignifikancia-kérdést élesebbé tegyük. Rossi és munkatársai úgy vélik, hogy az anyagi támogatás igenis csökkentette a visszaesést, csak ezt a hatást elfedi a munkában töltött hetek hatása. Elemzésük bírálata, részletesen kifejtve: H. Zeisel, „Disagreement over the evaluation of a controlled experiment”, *American Journal of Sociology* vol. 88 (1982), 378–396. o.
28. S. J. Sherman, „On the self-erasing nature of errors of prediction”, *Journal of Personality and Social Psychology*, vol.19 (1980.), 211–221. o.
29. Dr. William Epstein; beszámol róla a *New York Times* 1988. szeptember 27.

28. FEJEZET. A χ^2 -PRÓBA

1. K. Pearson, „On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling”, *Phil. Mag.*, series V., vol.1 (1900), 157–175. o.
2. Ha valószínűségi modellünk helyes, mindegyik összeadandónak valamivel 1 alatti a várható értéke; az összeg várható értéke pedig $n-1$, ahol n az összeadandók száma.

3. A görbe egyenlete

$$y = \frac{100\%}{\Gamma(d/2)} \left(\frac{1}{2}\right)^{d/2} x^{(d/2)-1} e^{-x/2}$$

ahol

d = a szabadságfokok száma

Γ = az Euler-féle gamma-függvény

4. A pontos eloszlást a második kiadáshoz számítottuk ki, IBM PC-XT 286-oson egy 5 perc futásidejű True-BASIC programmal. (90 MHz-es Pentiumon ugyanez a program 10 másodperc alatt hajtódik végre.) A program ábécérendben végigment az összes 60-összegű számhatoson. Mindegyik számhatosra kiszámította, hogy mekkora érték adódik a χ^2 -statisztikára, továbbá (a polinomiális eloszlás képletével) kiszámította az adott számhatos valószínűségét. E valószínűségek összegzésével adódott a megoldás – és a 2. ábrán látható elméleti hisztogram. Úgy tűnt, hogy a számítás 15 tizedesjegyig pontos, mivel az összes valószínűség összege $1 - 10^{-15}$ -nek adódott. A 2. ábrán látható fogazottság valós. 600 kockadobásnál a szakadékok és a csúcsok kisebbek, de a hisztogram továbbra sem sima.
- Sok könyv javasolja a Yates-korrektciót (négyzetreemelés előtt az eltérés abszolút értékéből vonjunk ki 0,5-öt, ha ez – az eltérés abszolút értéke – nagyobb 0,5-nél). Egy szabadságfoknál ez a folytonossági korrekcióval (l. a 18. fejezet 4. szakaszát) egyenértékű, és valóban jó, ha így teszünk. Egynél nagyobb szabadságfok esetén azonban a számítások azt mutatják, hogy sokszor nem jó, ha így teszünk. A hisztogram túl sokkal eltolódik bal felé. Más számításokból kiderül, hogy ha cellánként 5 megfigyelés várható és a szabadságfok sem magas, akkor a χ^2 -görbében nagyjából az 5%-os pontig bízhatunk. Ha cellánként 10 megfigyelés várható, akkor az 1%-os ponton jóval túl is bízhatunk a görbében.
5. Amikor a dobozban csak kétfajta lap van, a χ^2 -statisztika megegyezik a z-statisztika négyzetével. Mivel egy normális eloszlású változó négyzete: egy szabadságfokú χ^2 , így a χ^2 -próba ebben az esetben pontosan azt az eredményt adja, mint a (kétoldali) z-próba. Lásd még 27. fejezet 3. jegyzet.
6. E példához és a 28. fejezet 2. szakasz 9-es feladathoz az adatokat a Kaliforniai Államsorsjáték (California State Lottery) szívességéből, statisztikus szakértőjük, Don Ylvisaker (UCLA) révén kaptuk. A P-értéket numerikusan – többszöri parciális integrálással – számítottuk. A szokásos normális közelítések nem működnek igazán jól, van azonban egy megdöbbentően pontos eredményeket adó közelítés: D. B. Peizer and J. W. Pratt, „A normal approximation for binomial, F, beta, and other common, related tail probabilities”, *Journal of the American Statistical Association*, vol. 63 (1968), 1416–1483. o.
7. Bizonyos esetekben – például ha a megfigyelések cellánkénti száma nagyon alacsony – érdemes az adatokat csoportosítani.
8. *UCLA Law Review*, vol. 20 (1973), 615. o.
9. Lásd a 25. fejezet 7. jegyzetét.
10. A. R. Luria, *The Working Brain* (New York: Basic Books, 1973).
11. A HANES-vizsgálat csoportos mintavétellel dolgozott, van tehát az adatokban némi összefüggés, amit a χ^2 -próba nem vesz figyelembe. Használhatjuk a mintafelezéses módszert, hogy nulleloszláshoz jussunk. Következtesen, minden korszakportban több a nők között a jobbkezes, mint a férfiak között. (Lásd: *Anthropometric Reference Data and Prevalence of Overweight: United States, 1976-1980*. Data from the National Health Survey, series 11, no.238., Washington, D.C., U.S. Department of Health and Human Services.) Az 5. táblázatbeli adatok közel vannak a valóságosakhoz, ugyanakkor könnyebben követhetővé teszik a számításokat.
12. Természetesen, ha a próbát úgy végezzük, hogy adottnak vesszük a marginálisokat, akkor tekinthetjük rögzítettnek a várható értékeket. Lásd a 27. fejezet 3. jegyzetét is.
13. Súlyozatlan adatok a Népszámlálási Hivatal által az 1988. márciusi Rendszeres Népszéfélmérés számára átadott szalagról. Igaz, a χ^2 -próba nem veszi figyelembe a mintavételi tervet, az eltérés azonban valóságos.
14. *UCLA Law Review*, vol. 20 (1973), 616. old.
15. A Népszámlálási Hivaltól származó súlyozatlan adatok, egy, a *U.C. Survey Research Center* által közreadott CD-ROM-ról. A hadseregénél szolgálatot teljesítőket alkalmazásban állóknak tekintettük.

A χ^2 -próba nem veszi tekintetbe a mintavételi tervet. Sok hasonló felmérés mutat arra, hogy a nőtlen férfiak kevésbé sikeresek a munkában, s hogy ez így van az összes korcsoportban. Ugyanakkor nőknél a hajadonok és a házások csoportjában nagyjából egyforma a munkanélküliek aránya. Az alábbi táblázatokban a Rendszeres Népeségfelmérés 1988. márciusi felvételéből adjuk a (nemcsak kaliforniai) 35–44 évesekre vonatkozó adatokat.

	FÉRFIAK		
	Özvegy, elvált,		
	Házás	külön él	Nőtlen
Alkalmazásban	7660	1190	856
Munkanélküli	308	88	67
Inaktív	278	148	175

	NŐK		
	Özvegy, elvált,		
	Házás	külön él	Hajadon
Alkalmazásban	5879	1798	699
Munkanélküli	218	120	26
Inaktív	2323	422	163

16. Az IRRI adatai, módosítva.

17. A James Smyth kiadási ügyében, a Szövetségi területi törvényszéken tartott meghallgatáson (N.D.Calif., 1993.) bemutatott bizonyítékok, átfogalmazva. Lásd a védelem 1993. dec. 10-i írásos előterjesztését (7–8. old.), a 72.15 sz. felperesi bizonyítékot, és Robert Koyak írásos tanúvallomását. A Területi törvényszéknek a kiadást a valószínűsíthető diszkrimináció alapján elutasító döntése a fellebbezés nyomán ellenkezőjére változott.

29. FEJEZET. SZIGNIFIKANCIAPRÓBÁK, KÖZELEBBRŐL

- 225 US 391; idézi: Jerome Frank, *Courts on Trial* (Princeton University Press, 1949).
- Ezekkel a szavakkal védelmezte e gyakorlatot a *Journal of Experimental Psychology* szerkesztőjeként Arthur Melton:

A cikk értékelésének következő lépéseként arról kellett ítéletet mondanunk, hogy mennyire lehetünk bizalommal az eredmények iránt – mennyire bízhatunk abban, hogy a kísérletet a megadott körülmények között megismételve ugyanezek az eredmények adódnának. A *Journal* szerkesztése során jelentős volt a vonatkozás a kutatás fő kérdéséhez kötődő eredmények elfogadása s közlése iránt, ha ezek – akár 1-, akár 2-oldalú próbával – [csak] a 0,05-ös szinten mutatkoztak szignifikánsnak! Ez, egyes bírálók esetleges ilyen véleménye ellenére sem jelenti sem a 0,01, sem valamely más szignifikanciaszint szolgai imádatát. Véleményünk szerint azonban a tudományban a kutató felelőssége, hogy úgy tevékenykedjék, hogy eredményeit senki se vonhassa kétségbe azt mondván, hogy ezeket egy golyó ide-oda pattogása idézte elő.

Hogy meggyőződjünk az eredmények megismételhetőségéről, annak jobb módszere is van: ragaszkodni kell a fontos kísérletek megismétléséhez. Az idézet a *Journal* – a 64. kötet (1962.), 553-557. oldalán található – vezércikkéből való. Mi David Bakan egy cikkében találkoztunk vele; újranyomva: J. Steger (ed.), *Readings in Statistics* (Holt, Rinehart and Winston, 1971.). Lásd az alábbi 4. jegyzetet is.

- A történet ismertetésénél G. A. Barnardnak, az Imperial College of Science and Technology volt statisztika professzorának tanúságtételére támaszkodtunk.
- Elég szomorú, de már egy szerény mértékű szignifikanciaavadászat jelentősen torzítja a P -értékeket. Természetesen nem megoldás, ha a kutatók – hogy a P -értékeknek baja ne essék – nem nézik meg az adataikat. Ajánlható módszer például a *cross-validation*: dolgozzuk ki az adatok egyik felén a modellt, majd nézzük meg, milyen az illeszkedés, ha egyenleteinket az adatok másik felére alkalmazzuk. Ennél is jobb a vizsgálat tényleges megismétlése. A vizsgálatok megismétlése alapvető fontosságú.

gú gondolat, mellyel nem foglalkoztunk érdeme szerint. Hivatkozások (szignifikanciavadászat és vizsgálatok megismétlése):

- R. Abelson, *Statistics as Principled Argument* (Hillsdale, N.J., Lawrence Erlbaum Associates, 1995).
- T. K. Dijkstra (ed.), *On Model Uncertainty and its Statistical Implications*. Springer Lecture Notes No. 307, in *Economics and Mathematical Systems* (1988).
- A. S. C. Ehrenberg and J. A. Bound, „Predictability and prediction“, *Journal of the Royal Statistical Society*, series A, vol.156, part 2 (1993), 167–206. old.
- D. A. Freedman, „A note on screening regression equations“, *The American Statistician* vol. 37 (1983), 152–155. o.
- M. Oakes, *Statistical Inference* (Chestnut Hill: ERI, 1986).
5. Példánk stilizált, de a probléma valós. Mi évenként, 100 000 főre számítva, 1 esettel számoltunk, a Poisson-modellrel használva. A környezetszennyezéssel kapcsolatos aggodalmak dacára a harmincas évektől kezdve folyamatosan csökken a májrákos esetek előfordulási aránya az Egyesült Államokban. Részletesebb tárgyalás és további hivatkozások találhatóak: D. Freedman and H. Zeisel: „From mouse to man: the quantitative assessment of cancer risks“, *Statistical Science*, vol. 3 (1988), 3–56. o., bővebben kifejtve. Lásd még B. N. Ames, L. S. Gold and W. C. Willett, „The causes and prevention of cancer“, *Proceedings of the National Academy of Science, U.S.A.*, vol. 92 (1995), 5258–5265. o. Egy vitatott halmozott-előfordulás esetről ír S. W. Lagakos, B. S. Wessen and M. Zelen, „An analysis of contaminated well water and health effects in Woburn, Massachusetts“, *Journal of the American Statistical Association*, vol. 81 (1986), 583–614. o., hozzászólásokkal. Izgalmas beszámoló a woburni perről: Jonathan Harr, *A Civil Action* (Random House, 1995).
6. A szignifikanciavadászat más eseteiben nem ilyen egyszerű a P -érték helyesbítése. Lásd a 4. jegyzetben hivatkozott Dijkstra-könyvet. A jelenségnek a tudományos közleményekre gyakorolt hatásáról lásd:
- L. J. Chase and R. B. Chase, „A statistical power analysis of applied psychological research“, *Journal of Applied Psychology*, vol. 61 (1976), 234–237. o.
- K. Dickersin, S. Chan, T. C. Chalmers, H. S. Sacks and H. R. Smith, Jr., „Publication bias and clinical trials“, *Journal of Controlled Clinical Trials* vol. 8 (1987), 343–53. old.
- A. Tversky and D. Kahneman, „Belief in the law of small numbers“, *Psychological Bulletin*, vol. 2 (1971), 105–10. old.
- C. B. Begg and J. A. Berlin, „Publication bias and dissemination of clinical research“, *Journal of the National Cancer Institute*, vol. 81 (1989), 107–115. o.
7. „The Lipid Research Clinics Primary Prevention Trial Results“, *Journal of the American Medical Association*, vol. 251 (1984), 351–64. old. A kutatók élettartam-elemzés és blokkos elrendezés alapján kapott, kb. -1,92-es z -értékről számolnak be. Az előzetes kísérleti protokoll nem utal rá, egy- vagy kétoldali próbával szándékoztak-e dolgozni; beszámol arról, hogy „szignifikáns morbiditás és mortalitás mutatkozott a koleszterinszintet csökkentő szerekkel összefüggésben“; s közli, hogy „mércéül, a kísérleti csoportok közötti eltérések meggyőző bemutatása érdekében“ az 1%-os szignifikanciaszintet választották. A szövegből határozottan úgy tűnik, hogy szándékuk szerint külön elemezték volna a végzetes, illetve a nem végzetes infarktuszokat – úgy az eltérések nem szignifikánsak. Lásd *Journal of Chronic Diseases*, vol. 32 (1979), 609–631. o. Úgy tűnik, hogy a kutatók eltértek az előzetes kísérleti protokolltól. Kevésbé formális beszámoló: T. J. Moore, *Heart Failure* (Random House, 1989) és *Lifespan* (Simon & Schuster, 1993).
- Egy másik kísérletről Buchwald et al. „Effect of partial ileal bypass surgery on mortality and morbidity from coronary heart disease in patients with hypercholesterolemia“, *New England Journal of Medicine*, vol. 323 (1990), 946–955. o. számol be. De lásd G. D. Smith and J. Pekkanen, „Should there be a moratorium on the use of cholesterol lowering drugs?“, *British Medical Journal*, vol. 304 (1992), 431–434. o.: több kísérlet adatai mutatnak, hogy a koleszterinszintet csökkentő gyógyszerek éppenhogy növelik a halálozási arányszámot. Másfelől egy, a *Simvastatin*-nal kapcsolatban végzett nagy skandináv vizsgálatban a mortalitás 30%-os csökkenését tapasztalták olyan páciensek körében, akiknek kórtörténetében szerepelt szívbetegség. Lásd: „Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study“, *Lancet*, vol. 344 (1994. nov. 19.), 1383–89. old. Van egy skót vizsgálat is, a *pravastatin*-ról, lásd *New*

- England Journal of Medicine* (1995. nov. 16). Összefoglalót ad A. M. Garber, W. S. Browner and S.B. Hulley, „Cholesterol screening in asymptomatic adults, revisited“, *Annals of Internal Medicine*, vol. 124 (1996), 518–531. o.
8. K. R. Rao, ed., „The Ganzfeld debate“, *Journal of Parapsychology*, vol. 49. no. 1 (1985) és vol. 50, no. 4 (1986). Lényeges az eloszlás diszkrét volta; a szignifikancia-valószínűségeket konvolúcióval lehet kiszámítani.
 9. A bioassay-eredmények kiértékelése komplikált dolog, de a sokvégpontú vizsgálatokkal kapcsolatos nehézség valós. És sok vegyület, ami májrákot okoz, ugyanakkor – egerekben – megelőzi a fehérvérűséget. Lásd az 5. jegyzetben hivatkozott Freedman–Zeisel dolgozatot. Lásd továbbá T. S. Davies and A. Monro: „The rodent carcinogenicity bioassay produces a similar frequency of tumor increases and decreases: implications for risk assessment“, *Regulatory Toxicology and Pharmacology*, vol. 20 (1994), 281–301. o. És lásd T. H. Lin et al., „Carcinogenicity tests and inter-species concordance“, *Statistical Inference*, vol. 10 (1995), 337–353. o.
 10. T. C. Chalmers, R. S. Koff and G. F. Frady, „A note on fatality in serum hepatitis“, *Journal of Gastroenterology and Hepatology*, vol. 69 (1965), 22–26. o.
 11. Még rosszabb a helyzet – még inkább keveredik a statisztikai értelemben vett szignifikancia a gyakorlati fontossággal – a korrelációs együtthatóknál: ahelyett, hogy r nagyságát vennék figyelembe, az $r = 0$ hipotézist tesztelik, és P -t használják a kapcsolaterősség mérőszámaként. A regressziós együtthatóknál is gyakran ez a helyzet. De legelősebben a baj a szórásелеmzésnél mutatkozik: vannak, akik mindent leírnak – P -értékeket, F -statisztikákat –, viszont nem közlik, hogy mekkorák a hatások. A téma kifejtését lásd P. E. Meehl, „Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology“, *Journal of Consulting and Clinical Psychology*, vol. 46 (1978), 806–834. o.
 12. Ugyanakkor nagyvárosi és falusi gyermekek olvasási készsége között jelentős különbségek lehetnek, különösen későbbi életkorokban. Lásd I. S. Kirsh and A. Jungeblut, *Literacy: Profiles of America's Young Adults* (Princeton: ETS/NAEP, 1986).
 13. A 6-pontos becslés egy hevenyészett regresszióelemzésből származik, mintavételi hatások is jelentkeznek benne. Lásd Mullis et al., 5. o. (idézjük a 27. fejezet 2-es jegyzetében).
 14. Egy (összefüggéséből kiemelt) megjegyzés pontos átfogalmazása; a megjegyzés forrása D. T. Campbell, „Reforms as experiments“, *American Psychologist*, vol. 24 (1969), 409–29. old. A néhai Merrill Carlsmith, a Stanford University volt pszichológia professzora hívta fel rá a figyelmünket.
 15. M. J. Mahoney, „Publication prejudices: an experimental study of confirmatory bias in the peer review system“, *Journal of Cognitive Therapy and Research*, vol. 1 (1977), 161–175. o. Valamelyest egyszerűsítettük a vizsgálati tervet; az idézeteket is.
 16. Daniel McFadden, „The revealed preferences of a government bureaucracy: empirical evidence“, *Bell Journal of Economics*, vol. 7 (1971), 55–72. old. A vizsgált időszak: 1958–1962. Egy-egy változó „hatása“: a modell egy együtthatója; természetesen maga a modell is vita tárgya lehet. Chris Achen, a politikatudományok professzora (University of Michigan) hívta fel figyelmünket a forrásra.
 17. W. Hogan és J. Kalt (Harvard) tanúvallomásának átfogalmazása; a tanúvallomásra egy 1987-es, az olajárszabályozás többszöri megsértésének tárgyában zajló kormányzati meghallgatáson került sor. Az árrugalmasság az együtthatók egyike egy regressziós modellben.
 18. Keynes után szabadon azt mondhatnánk, aki úgy végez szignifikanciapróbát, hogy azt hiszi, nincs szüksége dobozmodellre, annak is valószínűleg van, csak nem tud róla. J. M. Keynes, *The General Theory of Employment, Interest, and Money* (Harcourt Brace and Jovanovich, 1935, 383–384. old.)
A magukat mindennemű elméleti befolyástól mentesnek tudó gyakorlati emberek többnyire valamely rég halott közgazdász rabjai.
 19. *Statistical Abstract*, 1988., 20. táblázat; 1994., 20. táblázat.
 20. Ez a vizsgálat már szerepelt a 2. fejezet 4-es szakaszában; hivatkozásokért lásd az ottani 7-es jegyzetet. Ebben a példában $z \approx 5$, így P meglehetősen kicsiny. Értelmezhetjük leíró statisztikaként P -t. Összesen 933 jelentkező volt, 825 férfi és 108 nő. Ha úgy gondoljuk, hogy nem és felvétel között nincsen összefüggés, akkor a férfiakra illetve a nőkre vonatkozó felvételi arányszámok összehasonlítása olyan, mint a valamelyik 825 emberre vonatkozó felvételi arányszámnak az összehasonlítása a fenn-

maradó 108-ra vonatkozóval. (Végül is a jelentkezőket rengeteg irreleváns módon fel lehet osztani, pl. az ujjlenyomatuk stb. alapján.) Összesen

$$\binom{933}{825} \approx 7 \cdot 10^{43}$$

- féleképpen lehet a 933 jelentkezőt egy 825 és egy 108 létszámú csoportra kettébontani. Számítsuk ki mindegyik felbontásra z -t. A z -értékek ezen sokasága közel normális eloszlású lesz, így a megfigyelt 5-ös z -érték elég rendkívüli. Lásd D. Freedman and D. Lane, „A nonstochastic interpretation of reported significance levels“, *Journal of Business and Economic Statistics*, vol. 1 (1983), 292-298. o.
21. Kísérlet ritkán végeznek véletlen mintavétellel kapott alanyokon. Általában van valamiféle kezdeti szűrő fázis. Csak a szűrőn átjutott személyeket randomizálják, ők pedig leginkább *ad hoc* mintának tekinthetők. Ennélfogva bizonyos körülményekkel kell eljárni ilyenkor a következtetés feltételeinek megfogalmazásánál. Olyan modellre gondolunk (27. fejezet 11-es és 12-es jegyzet), melyben minden alanyon van két lehetséges válasza, az egyik a kezelésre, a másik a kontroll eljárásra. E válasz párok alkotják a „populációt“. Egyetlen alanyon sincsen mód mindkét válasz megfigyelésére. A kísérlet nem szolgáltat adatokat a teljes populációra, csak egy részére nézve. A kezelt csoporthoz sorolt alanyoknál a kezelésre adott választ, a kontrollhoz sorolt alanyoknál a kontroll eljárásra adott választ figyeljük meg. A statisztikai következtetésnél ezen megfigyelt adatokból következtetünk a randomizált alanyok válasz-pár-populációját jellemző paraméterekre. Nem következtetünk alanyok valamely bővebb populációjára – ilyen általánosításhoz a statisztikánál többre lenne szükség; s egy ilyen általánosítás meglehetősen problematikus lehet. Alapvetően olyan kísérletekre gondolunk, melyekben az alanyokat random módon két csoportba osztják. Mindazonáltal az iméntihez hasonló megjegyzések vonatkoznak például akkorra is, amikor valamilyen eseti eljárással párosítjuk az alanyokat, majd minden párnál feldobunk egy érmét, így sorsoljuk ki, melyikük kerül a kezelt, s melyikük a kontroll csoportba. Következtetni itt is a lehetséges válaszok összességét jellemző paraméterekre tudhatunk; a következtetés feltételes, az adott alany párokra és párosításra mint feltételre nézve.
 22. *Project Follow Through Classroom Evaluation*, kiadja az SRI (Menlo Park, California). Jane Stallings volt a kutatás vezetője. Az idézeteken árnyalatos változtatásokat hajtottunk végre. A vizsgálat 1972-73-ban zajlott.
 23. Ehhez azt kellene feltételeznünk, hogy a kontroll 60-as átlagát pontosan, hiba nélkül ismerjük. Az SRI valójában kétmintás *t*-próbat használ. Az SRI által alkalmazott pontozási eljárás viszont óhatatlanul összefüggést hozott létre a kezelt és a kontroll pontszámai között, mert összevont (pooled) rangsorokon alapult.
 24. A számok valódiak, 1976-ból. A tanársegédnek körülbelül a fele részt vett a vizsgadolgozat javításában; és sokan javítottak hasonló dolgozatokat más negyedekben.
 25. F. Mosteller and R. Rourke, *Sturdy Statistics* (Addison-Wesley, 1971., 54. old.)
 26. T. A. Ryan, B.L. Joiner and B. F. Ryan, *Minitab Student Handbook* (Boston: Duxbury Press, 1976., 228. old.)
 27. „Intellectual development of children by demographic and socioeconomic factors“, *Vital and Health Statistics*, series 11, no 110. (Washington, D.C., 1971).
 28. R. S. Erikson, J. P. McIver and G. C. Wright, Jr., „State political culture and public opinion“, *American Political Science Review*, vol. 81 (1987), 797-813. old. Többtenyezős regressziót végeztek, először a demográfiai kategóriákat (pl. alacsony jövedelmű, stb.) mutató „dummy“-változókon (0/1 változókon); majd bővítették a régiót és az államot jelölő dummykkal. Az állam bevonása 0,0898-ról 0,0953-ra növelte a korrigált R^2 -et – viszont 8,35 volt a beléptetési kritérium F -értéke, a számlálóban 40-es – a nevezőben pedig 55 072-es szabadságfokkal. A szerzők szerint az eltérések nemcsak statisztikai értelemben, hanem gyakorlatilag is jelentősek; az R^2 -ből nem ez látszik. A szerzők szerint is lehetséges, hogy az „állam-dummy“-k más, kimaradt változók helyetteseiként működjenek, de érveik szerint itt nem ez a helyzet. Az ebben és a 29-es jegyzetben említett dolgozatokról bővebb elemzést ad: D. Freedman, „Statistical models and shoe leather“, in: P. Marsden (ed.): *Sociological Methodology 1991* (Washington, D.C., American Sociological Association; 10. fejezet).

29. J. L. Gibson, „Political intolerance and political repression during the McCarthy era“, *American Political Science Review*, vol. 82 (1988), 511–39. old. „Hatások“: az út-modell egyítható. A becslésekben mutatkozó véletlenszerűséget a szerző feltehetőleg a modelltől eredőnek mondaná. A modell megfelelő volta azonban nem nyilvánvaló.
30. Bővebben tárgyalja a kísérletet C. E. M. Hansel, *ESP: A Scientific Evaluation* (Charles Scribner's Sons, 1966, 11. fejezet). Módosítottuk a számokat, hogy egyszerűbb legyen a számolás. A kísérlet célja egy, a szövegben tárgyalt következtetési hiba szemléltetése volt. A forrásra Dr. Charles Yarbrough (Santa Rosa, Kalifornia) hívta fel a figyelmünket.
31. Úgy látszik, hogy az *Aquarius*-ban használt véletlenszám-generátort nem ellenőrizték, ellenőriztek azonban hasonló típusúakat. A parajelenségekkel kapcsolatos kutatásban semmi sem egyszerű; annak a nagy részével, amit írunk, Tart nem értene egyet: C. Tart et al., „Effects of immediate feedback on ESP performance: a second study“, *Journal of the American Society for Psychological Research*, vol. 73. (1979), 151–165. old. Fordulatos összefoglalást ad a szóban forgó kérdésekről Martin Gardner, *Science: Good, Bad, and Bogus* (Avon Books, 1981.; 18. és 31. fejezet).
32. A kiadó (Harcourt, Brace & Jovanovich, Inc.) engedélyével.
33. A kérdés eredetijét A. Tversky and D. Kahneman használták. Lásd továbbá Steger – a fenti 2. jegyzetben hivatkozott – könyvének 298. oldalát.
34. Lásd a 25. jegyzetbeli hivatkozás 68. oldalán.
35. F. Arcelus and A. H. Meltzer, „The effect of aggregate economic variables on congressional elections“, *American Political Science Review*, vol. 69 (1965), 1232–69. old., hozzászólásokkal. A forrásra Chris Achen hívta fel a figyelmünket. Az érvelés regressziós modellt használ, tehát finomabb, mint a feladat mutatja. (Természetesen maga a modell érvényessége is vita tárgya lehet.) A kutatók hipotézisvizsgálattal kapcsolatos álláspontja azonban mondhatni brutális. Lásd Arcelus és Meltzer válaszát Goodman és Kramer hozzászólására.
36. *Statistical Abstract*, 1988., 21. táblázat; *Statistical Abstract*, 1994., 26. táblázat.
37. *Statistical Abstract*, 1994., 616. és 621. táblázat.
38. *Statistical Abstract*, 1994., 287. táblázat.
39. R. E. Just and W. S. Chern, „Tomatoes, technology and oligopsony“, *Bell Journal of Economics*, vol. 11 (1980), 584–602. old. A kérdést taglalja R. Daggett and D. Freedman, „Econometrics and the law: A case study in the proof of antitrust damages“, in: L. M. LeCam and R. A. Olshen (eds.), *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, vol 1, 123–72. old. (Belmont, Calif.: Wadsworth, 1985.) Just és Chern lineáris és loglineáris keresleti függvényeket is megbecsültek; a feladatban említett *t*-próbát egy lineáris keresletfüggvényben az ár együtthatójára alkalmazták.
40. Országos adatokat közöl George Gallup, Jr., *The Gallup Poll: Public Opinion* (Wilmington, Delaware: Scholarly Resources, Inc., 1991., 49. old.) Alább kivonatossan közöljük az adatokat. (Különböző tevékenységekre fordított idő „tegnap“, önbeszámoló alapján.)

	0	< 1 óra	1-3 óra	>3 óra
TV-t néz	12	8	42	38
Rádiót hallgat	22	20	33	25
Újságot olvas	29	39	30	2
Képeslapot olvas	64	22	14	0
Könyvet olvas (munkához)	62	11	18	9
Könyvet olvas (szórakozás)	67	14	16	3

41. Az idézet forrása: D. L. Hartl levele, *Nature* vol. 372 (1994.), 398. oldal; David Kaye-nek (Arizona State University) tartozunk köszönettel azért, hogy felhívta rá a figyelmünket. Áttekintés a DNS-alapú azonosításból származó bizonyítékokról: *Jurimetrics* vol. 34., no.1 (1993.)
42. A példa alapja egy olyan bizonyíték volt (módosítottunk rajta), melyet a szövetségi kerületi bíróságnak (N. D. Cal., 1993.) terjesztettek be a James Smyth kiadási ügyében tartott meghallgatáson. Defense Exhibit 31, *Secondary Analysis of the School Leavers Survey*, (1989), Standing Advisory Commission on Human Rights (Cormack et al.).

43. Az adatok forrása: „Civil jury cases and verdicts in large counties”, *Bureau of Justice Statistics Special Report* (1995. július). Összesen 12 026 esetről határoztak az esküdtszék.
44. Lásd R. C. Lewontin, „Sex, lies, and social science” (*New York Times Review of Books*, 1995. április 20, május 25 és augusztus 10). Lewontin az R. T. Michael és mtsai. által jegyzett *Sex in America: A Definitive Survey* (Little Brown) c. könyvet szemléli, mely az E. O. Laumann et al.: *The Social Organization of Sexuality: Sexual Practices in the United States* (University of Chicago Press) népszerűbb kiadása.
45. John A. Dossey et al., *Can Students Do Mathematical Problem Solving?* (Washington, D.C., U.S. Department of Education, Office of Educational Research and Improvement, 1992.; 141. és 172. o.)
46. A vonatkozó peranyagok: *Brock vs. Merrell Dow Pharmaceuticals, Inc.*, 874 F. 2d 307, 311-12 (5th Cir.), módosítás: 884 F. 2d 166 (5th Cir., 1989), felülvizsgálati kérelem megtagadása: 494 U.S. 1046 (1990); továbbá: D. H. Kaye and D. A. Freedman, *Reference Guide on Statistics*, Federal Judicial Center, Washington, D.C. (1994., 377. o.)
47. Lásd W. T. Keeton, J. L. Gould, and C. G. Gould, *Biological Science*, 5th ed. (W. W. Norton & Company, 1993., 445. o.)
48. Hivatkozások:
Statistical Abstract, 1994: a 70. táblázat szerint 94 401 millió a háztartások száma; a 309. táblázat szerint az FBI 2,980 millió betörésről számolt be; a Bűnügyi Felmérés 4,9%-os adata a 310. táblázatban látható.
Criminal Victimization in the United States (U.S. Bureau of Justice Statistics, évkönyv).
Crime in the United States (FBI, évkönyv).
- Az Országos Bűnügyi Felmérés erősen strukturált mintavételi tervvel dolgozik; mintájuk pontossága közelítőleg megegyezik az általunk megadott méretű egyszerű véletlen mintáéval. A számokat kerekítettük, de az eredményen nem változtattunk.
49. A randomizációnál blokkos elrendezést is használtak, ezt itt figyelmen kívül hagytuk. Az átlagokat közőlték; a szórások ismertetésére J. D. Neaton (University of Minnesota, a biostatistika professzora) szíveségéből nyílik módunk. Egy érdekes adalék: a Framingham-adatok alapján számított logisztikus regressziók – a rizikófaktorok amúgy szerénynek tűnő csökkenéséből (vérnyomás: 3 Hgmm; vérkoleszterin: 5 mg/dl; dohányzás: 13%) – a halálozási arányszámok igen jelentős esését jósolták. Felmerült, hogy a kezelt csoport esetleg a valóságosnál kisebb mérvű dohányzásról számol be; ennek nyomán – a vér kémiai vizsgálata alapján – az adatokat megfelelően korrigálták. Hivatkozások:
 „Multiple Risk Factor Intervention Trial”, *Journal of the American Medical Association*, vol. 248 (1982), 1465-1477. o.
 „Statistical design considerations in the NHLI Multiple Risk Factor Intervention Trial (MRFIT)”, *Journal of Chronic Diseases*, vol. 30 (1972), 261-275. o.
 „Mortality rates after 10,5 years for participants in the Multiple Risk Factor Intervention Trial”, *Journal of the American Medical Association*, vol. 263 (1990), 1795-1901. o.
50. George Gallup, Jr., *The Gallup Poll: Public Opinion 1992*. (Wilmington: Scholarly Resources). Lásd 118. oldal. A számok kerekítettek.
51. *Walsome vs. Port Authority*, 948 F.2d 1370, 1376 (2nd Cir. 1991); D. H. Kaye and D. A. Freedman, *Reference Guide on Statistics*, Federal Judicial Center, Washington, D.C. (1994., 381. old.). Az idézetten kissé módosítottunk.
52. M. S. Kanarek et al., „Asbestos in drinking water and cancer incidence in the San Francisco Bay Area”, *American Journal of Epidemiology*, vol. 112 (1980), 54-72. old. Bővebb taglalás végett lásd a fenti 28. jegyzetben hivatkozott cipő-bőr (shoe-leather) tárgyú cikket. Sem a feketék, sem a nők körében nem volt kapcsolat a vízben lévő azbeszt és a tüdőrák között. A publikált adatok erősen arra mutatnak, hogy a dohányzás egybemosó változó lehetett.
53. Lásd a 19. fejezet 19. jegyzetét. Az egyik kivételt képeznek: eredményeik a kétgyermekes családok elsőszülöttjeinek eredményeinél valamelyest gyengébbek.