1.2 DEFINICE PRAVDĚPODOBNOSTI

1.2.1 Úvod

Historii matematické teorie pravděpodobnosti a obecněji také formování a vývoji pravděpodobnostního uvažování byla věnována již celá řada prací vydaných v edici Dějiny matematiky i v dalších českých sbornících a časopisech. Připomeňme zde především knihy a pojednání Karla Mačáka: *Počátky počtu pravděpodobnosti* [223], *Poznámky k formování kombinatoriky v 16. a 17. století* [224], *Poznámky k formování teorie pravděpodobnosti v 17. a 18. století* [222], *Bernard Bolzano a počet pravděpodobnosti* [220] a *Vývoj teorie pravděpodobnosti v českých zemích do roku 1938* [225], a dále pojednání Ivana Saxla: *Pravděpodobnost ve starověku a středověku* [319], *Pravděpodobnost před Pascalem a Fermatem* [323] či *Pravděpodobnost ve středověku* [324].

V návaznosti na předchozí kapitolu si nyní jen připomeňme základní přístupy k zavedení pojmu pravděpodobnosti a uveďme některé důležité mezníky ve vývoji této teorie.[30]

Život a náhodné jevy

Náhodné jevy a nutnost rozhodovat se v nejistých situacích a na základě nejistých informací člověka provázejí od pradávna, ať už hovoříme o výkyvech počasí, přírodních katastrofách, nemocech, epidemiích, nehodách či například o soudnictví. Prehistorie teorie pravděpodobnosti je tedy velmi dlouhá a rozhodně nezahrnuje pouze házení kostkami, popř. kůstkami různých tvarů, a jiné hry, v nichž hraje hlavní roli náhoda.

Upozorněme alespoň na židovský Talmud z první poloviny prvního tisíciletí, který obsahuje velké množství zajímavých úvah, jež bychom dnes nazvali pravděpodobnostními, anebo bychom je dokonce zařadili do teorie her (viz [6], [162]). K řešení mnohých nejednoznačných situací zde bylo využíváno losování z urny, jehož výsledek byl považován za projev Boží vůle. Na základě losování se rozhodovalo například o dělení majetku v nejasných případech, ale také o zvířecích obětech, službách v chrámu či dělení masa obětovaných zvířat mezi sloužící kněžstvo. Pravděpodobnostní úvahy lze rozeznat i v pozadí rozhodování o tom, jaký počet zemřelých svědčí o morové epidemii či jaký počet rituálních

[30] Zájemce o historii teorie pravděpodobnosti nalezne podrobné informace kromě výše citovaných prací také v klasickém spisu [372] Isaaca Todhuntera či v novější dvoudílné knize [133] a [134] Anderse Halda.

řemínků je třeba prověřit, aby byl celý soubor schválen – dnes bychom hovořili o „výběrovém šetření“.

Typickou oblastí velmi úzce spojenou s teorií pravděpodobnosti je dodnes pojišťovnictví. Smlouvy, jejichž předmětem bylo pojištění rizika v námořní dopravě, se začaly objevovat již ve 14. století. Pojistné bylo udáváno jako procentuální část hodnoty pojištěného majetku a lze jej považovat za numerický odhad pravděpodobnosti, že náklad bude zničen. Jeho výše závisela na aktivitách pirátů v příslušných vodách, místních válkách, navštívených přístavech a jejich vzájemných vzdálenostech, na kvalitě a zkušenostech kapitána i pojištěných obchodníků, na typu a stavu lodi, druhu přepravovaného zboží aj.

Dlouhé prehistorii, která zde byla jen stručně naznačena a která je mimořádně zajímavá také z didaktického hlediska, je věnována kniha *The Science of Conjecture: Evidence and Probability before Pascal* [111] Jamese Franklina; podrobně se jí zabývá také dvojice zmíněných pojednání [319] a [323] Ivana Saxla.

Počátky matematické teorie pravděpodobnosti

Za počátek matematické teorie pravděpodobnosti bývá obvykle považována korespondence, kterou spolu vedli Blaise Pascal (1623–1662) a Pierre de Fermat (1607–1665) v roce 1654, v níž byly kromě jiného úspěšně vyřešeny některé speciální případy úlohy o rozdělení sázky.[31]

V roce 1985 ovšem Laura Toti Rigatelli zveřejnila informace o italském rukopisu *Codice Magliabechiano CL.XI.120*, datovaném kolem roku 1420, který nalezla v Národní knihovně ve Florencii a který obsahuje správné řešení úlohy o rozdělení sázky pro dva hráče, z nichž jednomu chybí k vítězství jedna a druhému tři vyhrané hry – viz [374].

Ve dvojici článků [109] a [110] z roku 2002 pak Raffaella Franci upozornila na jiný středověký rukopis, který objevila ve Vatikánské apoštolské knihovně. Zde je kromě jiného úspěšně vyřešen problém rozdělení sázky mezi tři hráče v případech, kdy prvnímu hráči chybí jedna nebo dvě hry, druhému jedna, dvě nebo tři a třetímu dvě nebo tři hry. Zmíněné rukopisy jsou podrobně rozebrány v pojednání [237] Norberta Meusniera, který navíc dokládá, že úloha o rozdělení sázky vychází z problematiky dělení podílů při jednostranném porušení dohody, rozvíjené a předávané italskými učiteli „komerční aritmetiky“ ve 13. a 14. století.[32] Bohužel, později tyto spisy upadly v zapomnění dokonce i v Itálii. V druhé polovině 15. a v první polovině 16. století se tak řada matematiků pokoušela o řešení úlohy znovu a neúspěšně, např. Luca Pacioli (1445–1517), Niccolo Fontana zvaný Tartaglia (1499–1557), Giovanni Francesco Peverone (1509–1559) či Girolamo Cardano (1501–1576).

[31] Připomeňme, že v nejjednodušším případě se uvažuje hra dvou hráčů, kteří hrají sérii her o danou částku V s tím, že tuto částku získá hráč, který jako první vyhraje určitý počet her. Přitom se předpokládá, že v každé hře je pravděpodobnost výhry pro oba hráče stejná. Série je ukončena předčasně a hráčům v tu chvíli chybí k celkovému vítězství různé počty her. Otázka zní, jak má být mezi ně částka V spravedlivě rozdělena. O úlohách řešených Pascalem a Fermatem podrobně pojednává K. Mačák v knize [223].

[32] V češtině je tomuto tématu věnován článek [324] I. Saxla.

Skutečnost, že některé výsledky, k nimž dospěli Pascal, Fermat a později také Christian Huygens (1629–1695),[33] lze nalézt již o více než dvě století dříve, však nikterak nesnižuje význam těchto osobností pro formování teorie pravděpodobnosti. Huygensův spis *De ratiociniis in ludo aleae* (*O uvažování v hazardních hrách*) [160] z roku 1657 byl první tištěnou prací v této oblasti a základní publikací zůstal po dobu více než padesáti let. Potom se stal součástí dalšího slavného spisu: byl přetištěn v první části knihy *Ars conjectandi* (*Nauka o domněnce*) [29] Jacoba Bernoulliho (1654–1705), který jej opatřil podrobným komentářem, přibližně čtyřikrát rozsáhlejším než původní text, v němž na mnoha místech vylepšil, doplnil či zobecnil Huygensovy myšlenky a podal nová řešení různých problémů.

Christian Huygens: *De ratiociniis in ludo aleae*

Je zajímavé a z dnešního pohledu možná překvapivé, že Huygens ve svém spise [160] nepoužíval pojem „pravděpodobnost“, i když v úvodu můžeme nalézt zmínku, že *kdyby se například někdo sázel, že hodí šestku prvním hodem jednou kostkou, pak je sice nejisté, zda vyhraje, ale samotnou věcí je dáno, kolikrát je pravděpodobnější, že prohraje než že vyhraje, a lze to vypočítat.* Při řešení úloh však Huygens pracoval s pojmem *očekávaná výhra* či *očekávání* (*expectatio*) ve smyslu naší střední hodnoty (pro diskrétní náhodnou veličinu), a tedy také ve smyslu *matematické naděje*, s níž jsme se několikrát setkali v předcházející části. Spis začíná podrobným zdůvodněním následujících tvrzení:[34]

Tvrzení I. Jestliže bych očekával a nebo b, které bych obě mohl získat stejně snadno, pak je třeba říci, že mé očekávání má hodnotu $\frac{a+b}{2}$.

Potom Huygens uvažuje tři hodnoty, a, b, c, a nakonec odvozuje:

Tvrzení III. Jestliže by počet případů, v nichž mi připadne a, byl roven p, ale počet případů, v nichž mi připadne b, byl roven q, pak za předpokladu, že všechny případy jsou stejně možné, mé očekávání bude mít hodnotu

$$\frac{pa+qb}{p+q}\,. \tag{1.9}$$

[33] Christian Huygens pocházel z Haagu. V letech 1645 až 1649 studoval práva a matematiku na univerzitách v Leidenu a Bredě, potom se věnoval především vědeckému bádání za ekonomické podpory svého otce. Při návštěvě Paříže v roce 1655 se dozvěděl o Pascalově korespondenci s Fermatem, týkající se problémů rytíře de Méré, jejich konkrétní řešení však podle svých slov neznal. Po návratu do Holandska napsal spis [160], o němž se zmíníme v následujících odstavcích a který se stal důležitým mezníkem ve vývoji teorie pravděpodobnosti. V roce 1663 byl Huygens zvolen členem Královské společnosti v Londýně, o tři roky později se stal členem právě založené francouzské akademie věd a usadil se trvale v Paříži. V roce 1681 odcestoval na léčení do Haagu. Návrat do Francie nejprve zkomplikovala smrt zakladatele pařížské akademie a Huygensova osobního patrona, Jeana-Baptisty Colberta (1619–1683), v roce 1685 pak byl zrušen edikt nantský, který zaručoval protestantům náboženskou svobodu a stejná práva, jakých požívali katolíci, a Huygens se jako protestant do Francie již vrátit nemohl. Zemřel v Haagu 8. července 1695.

[34] Viz [160], str. 2–5; český překlad citován podle [223], str. 45–46. Zlomek (1.9) jsme zde pro názornost a přehlednější odkazování odsadili a očíslovali.

Uvědomme si, že my bychom dnes postupovali obráceně: řekli bychom, že $p/(p+q)$, resp. $q/(p+q)$, jsou pravděpodobnosti, že získáme částku a, resp. b (povšimněme si, že Huygens ve svých tvrzeních hovoří i o tom, že jsou všechny případy „stejně možné“, popř. že obě částky je možné získat „stejně snadno“), a spočítali bychom střední hodnotu výhry:

$$\frac{p}{p+q} \cdot a \; + \frac{q}{p+q} \cdot b = \frac{pa+qb}{p+q}\,.$$

Ovšem vzhledem k tomu, že Huygensův spis je zaměřen na hazardní hry, kde je důležitou otázkou hodnota výhry či prohry, je založení všech úvah na pojmu *očekávání* poměrně přirozené a názorné.

Bernoulli ve svém komentáři v knize [29] později přidal k Huygensovu sice jednoduchému, ale trochu delšímu odvození tvrzení III velice stručné a srozumitelné zdůvodnění, které by i dnes mohlo posloužit jako zajímavé objasnění pojmu střední hodnoty: Uvažujme loterii, v níž je p losů, z nichž každý vyhrává částku a, a dále q losů vyhrávajících b. Představme si, že si celkem $p+q$ hráčů koupí po jednom losu. Hodnota celkové výhry je $pa+qb$, a protože očekávání všech hráčů je stejné, musí jeho hodnota být $(pa+qb)/(p+q)$.

Následujících šest tvrzení se týká spravedlivého rozdělení sázky,[35] potom se Huygens obrací ke hře v kostky, s níž souvisí zbývajících pět tvrzení. Spis potom uzavírá pět neřešených úloh;[36] poslední z nich je v různých modifikacích velmi dobře známá dodnes a označuje se jako úloha o ruinování hráče.

Rozbor Huygensova spisu [160] i pozdějších řešení jeho úloh lze nalézt v Mačákově článku [221] a knize [223], kde je také přetištěn celý původní text s českým překladem. Pro ilustraci Huygensova přístupu zde proto již jen uveďme několik ukázek. V prvním z problémů týkajících se hry v kostky Huygens zkoumá šance na to, že padne aspoň jedna šestka při opakovaném hodu jednou kostkou, a hned na začátku se vrací k výše citované větě z úvodu spisu:

Jestliže by se někdo vsadil, že prvním hodem hodí šestku, pak zřejmě je jeden případ, kdy vyhrává a dostává to, co je vloženo jako základ; naproti tomu je pět případů, kdy prohrává a nedostává nic. Je tedy 5 hodů proti němu a jen jeden pro něho. Označme a to, co je vsazeno. Může tedy jednou očekávat a, ale pětkrát 0, což má podle třetího tvrzení hodnotu $\frac{1}{6}a$*. A zůstává* $\frac{5}{6}a$ *pro toho, kdo mu tuto příležitost nabídl. Může tedy vsadit jen 1 proti 5, kdo by chtěl uspět napoprvé.*[37]

Jinak řečeno, jestliže v případě, že nastane daný jev (např. padne šestka), vyhrajeme částku a a v opačném případě nevyhrajeme nic, získáme očekávanou výhru ze zlomku (1.9) jednoduše dosazením $b=0$. Povšimněme si, že kdybychom navíc dosadili $a=1$, udával by výraz (1.9) přímo pravděpodobnost daného jevu ve smyslu tzv. klasické definice – viz str. 40.

[35]Tento problém Huygens řeší nejprve pro dva hráče, z nichž jednomu zbývá jedna nebo dvě hry a druhému dvě, tři nebo čtyři hry, a potom přechází ke třem hráčům; popisuje obecnou metodu řešení a udává konkrétní řešení pro 17 různých případů.

[36]U úloh 1, 3 a 5 je uveden výsledek. Řešení všech pěti úloh podal Jacob Bernoulli ve spise [29].

[37]Viz [160], str. 16; český překlad citován podle [223], str. 59.

Následující problém zní takto: *Zjistit, kolika hody se lze odvážit dvěma kostkami hodit 12 bodů.*[38] Dnes bychom řekli, že Huygens hledal takový počet hodů dvěma kostkami, kdy pravděpodobnost, že padne aspoň jednou 12 bodů, není menší než 1/2. Huygens však nepoužíval pojem *pravděpodobnost*, ale *očekávání* a hledal takový počet hodů dvěma kostkami, při kterém by očekávaná výhra hráče sázejícího na to, že aspoň jednou padne dvojice šestek, byla větší než očekávaná výhra hráče, který by ve stejné hře vsadil na neúspěch. Huygens uvažoval postupně 1, 2, 4, 16 a 24 hodů a počítal očekávané výhry obou hráčů. Přitom opakovaně používal výraz (1.9) z tvrzení III: při jednom hodu existuje jeden příznivý případ a 35 případů nepříznivých, očekávání je proto rovno $[(1 \cdot a + 35 \cdot 0)/36] = (1/36)a$. Při dvou hodech Huygens uvažuje takto: v prvním z nich existuje jeden případ, kdy padnou dvě šestky a hráč získá výhru a, a 35 případů, kdy padne něco jiného – hráč pak může uspět v druhém hodu, v němž opět očekává $(1/36)a$. Před začátkem hry tedy hráč podle (1.9) očekává

$$\frac{1 \cdot a \ + 35 \cdot \frac{1}{36}\, a}{36} = \frac{71}{1296}\, a\,, \quad \text{zatímco protihráč očekává} \quad \frac{1225}{1296}\, a\,.$$

Dále stačí uvažovat přímo čtyři hody: v 71 případech z 1296 hráč uspěje v jednom z prvních dvou hodů a získá a, v ostatních případech nikoli a jeho očekávání před dalšími dvěma hody je $(71/1296)a$; celkem je tak hráčovo očekávání

$$\frac{71 \cdot a \ + 1225 \cdot \frac{71}{1296}\, a}{1296} = \frac{178\,991}{1\,679\,616}\, a\,, \quad \text{zatímco protihráč očekává} \quad \frac{1\,500\,625}{1\,679\,616}\, a\,.$$

Potom Huygens poznamenává, že stejným způsobem lze pokračovat s 8, 16 a 24 hody, a uzavírá, že první hráč začne být ve výhodě při 25 hodech.

Jako třetí ukázku zde uveďme problém, při jehož zobecnění Bernoulli později dospěl k binomickému rozdělení (viz část 1.2.2): *Zjistit, kolika kostkami by se někdo mohl odvážit hodit prvním hodem dvě šestky.*[39] Huygens říká, že se jedná o stejný problém, jako kdyby nám měla padnout alespoň dvakrát šestka při opakovaném hodu jednou kostkou. Při dvou hodech by bylo očekávání zřejmě $(1/36)a$. Nyní uvažujme tři hody. V jednom ze šesti případů se stane, že padne šestka hned napoprvé; pak stačí již jen jedna šestka ve dvou hodech, kdy je očekávaná výhra $(11/36)a$.[40] V pěti případech šestka napoprvé nepadne; pak je k výhře třeba, aby padla šestka v obou zbývajících hodech, a očekávání je $(1/36)a$. Před začátkem hry je tedy očekávaná výhra

$$\frac{1 \cdot \frac{11}{36}\, a + 5 \cdot \frac{1}{36}\, a}{6} = \frac{16}{216}\, a = \frac{2}{27}\, a.$$

Nakonec Huygens píše: *Přidává-li se tímto způsobem postupně jeden hod navíc, zjistí se, že buď deseti hody jednou kostkou nebo prvním hodem deseti kostkami je možno se pokusit, aby padly dvě šestky, a to se ziskem.*

[38] Viz [160], str. 18; český překlad citován podle [223], str. 59.

[39] Viz [160], str. 20; český překlad citován podle [223], str. 61.

[40] Jak Huygens ukázal již dříve, v jednom případě ze šesti padne šestka v prvním hodu a hráč vyhraje a, ve zbývajících 5 případech ještě může šestka padnout v druhém hodu a očekávaná výhra je $(1/6)a$; podle (1.9) je pak celkové očekávání $[1 \cdot a \ + 5 \cdot (1/6)a]/6 = (11/36)a$.

1.2.2 Jacob Bernoulli: pravděpodobnost jako míra jistoty

Definice pojmu *pravděpodobnost* (*probabilitas*) v souvislosti s náhodnými hrami se objevuje až v knize *Ars conjectandi* [29] Jacoba Bernoulliho,[41] vydané v roce 1713. Jak již bylo zmíněno, první část tohoto spisu obsahuje podrobně okomentované vydání Huygensovy práce [160]. Jedno z řady zajímavých zobecnění, která Bernoulli ve svých poznámkách přináší, lze nalézt v rozsáhlém komentáři k úloze, v níž se Huygens ptal, *kolika kostkami by se někdo mohl odvážit hodit prvním hodem dvě šestky* – viz 1.2.1, str. 37. Bernoulli zde kromě jiného zkoumá šance na výhru pro hráče, jemuž by se během n opakování téhož pokusu měl podařit určitý výsledek právě m-krát; přitom předpokládá, že počet příznivých, resp. nepříznivých případů se nemění – označme je r, resp. s, a dále označme $t = r + s$. Bernoulli pak ukazuje, že hráčovo očekávání je

$$\frac{n(n-1)(n-2)\dots(n-m+1)}{1\cdot 2\cdot 3\cdot 4\cdot\dots\cdot m}\cdot\frac{r^m s^{n-m}}{t^n}. \tag{1.10}$$

Řečeno dnešními slovy, Bernoulli dochází k *binomickému rozdělení* $\mathrm{Bi}(n,p)$, kde $p = r/t$ a pravděpodobnosti možných hodnot jsou

$$P(X=m) = \binom{n}{m} p^m (1-p)^{n-m}, \quad m = 0, 1, \dots, n. \tag{1.11}$$

Pojem pravděpodobnosti

V druhé části spisu [29] se Bernoulli zabývá permutacemi a kombinacemi, v třetí části je potom tato teorie aplikována na různé náhodné hry. Závěrečná čtvrtá část je věnována využití teorie pravděpodobnosti v občanských, morálních a ekonomických záležitostech a zde konečně nalezneme definici pravděpodobnosti:

Pravděpodobnost totiž je stupněm jistoty a liší se od ní jako část od celku. Jestliže totiž celá a absolutní jistota, kterou označíme písmenem a nebo jednotkou 1, se podle předpokladu skládá například z pěti pravděpodobností jako částí, z nichž tři jsou pro existenci nebo budoucí výskyt nějakého jevu, ostatní jsou proti: o onom jevu bude řečeno, že má $\frac{3}{5}a$ nebo $\frac{3}{5}$ jistoty.[42]

Bernoulli pak uvažoval následující speciální případy.

[41] Jacob Bernoulli absolvoval studia filosofie a teologie na univerzitě v Basileji; proti vůli svých rodičů přitom studoval také matematiku a astronomii. Potom působil jako vychovatel v Ženevě a ve Francii, později cestoval po Holandsku, Anglii a Německu, kde se seznámil s řadou význačných učenců a prohloubil si své vzdělání zejména v matematice a astronomii. Od roku 1683 konal přednášky o experimentální fyzice na basilejské univerzitě, kde byl o čtyři roky později jmenován profesorem matematiky; tuto funkci pak zastával až do konce života. V roce 1699 byl jmenován členem francouzské akademie věd, o dva roky později rovněž akademie pruské. Jacob Bernoulli zemřel 16. srpna 1705 v Basileji. Jeho spis *Ars conjectandi* [29] vydal o osm let později Nicolaus Bernoulli (1687–1759), syn Jacobova bratra a malíře Nicolause „staršího“ (1662–1716). O životě a díle Jacoba Bernoulliho, stejně jako o okolnostech vzniku a vydání jeho spisu [29], pojednává například E. D. Sylla v úvodu k překladu [34]. Zajímavé informace o vztahu Jacoba a jeho bratra Johanna lze nalézt v článku [274].

[42] [29], str. 211; český překlad citován podle [223], str. 22.

Možné [possibile] je to, co má alespoň malý stupeň jistoty, zatímco nemožné [impossibile] nemá žádnou nebo jen nekonečně malou jistotu. Možné je tedy například to, co má 1/20 *nebo* 1/30 *jistoty. Morálně jisté [moraliter certum] je to, co má pravděpodobnost tak blízkou úplné jistotě, že je rozdíl nezjistitelný. Naproti tomu morálně nemožné [moraliter impossibile] je to, co má jen tolik pravděpodobnosti, kolik chybí morální jistotě do jistoty úplné. Je-li například morálně jisté to, co má* 999/1000 *jistoty, pak to, co má jen* 1/1000 *jistoty, bude morálně nemožné. Nutné [necessarium] je to, co nemůže nebýt, nyní, v budoucnosti nebo v minulosti.* [...] *Náhodné [contingens] je to, co může nebýt, nyní, v budoucnosti nebo v minulosti.*[43]

Povšimněme si, že pravděpodobnost zde ještě není shora omezená hodnotou 1 a její vyčíslení připomíná Huygensovo *očekávání*. Tam, kde je to možné, Bernoulli vyjadřuje pravděpodobnost pomocí výrazu (1.9), kde klade $a = 1$, $b = 0$, tedy v duchu klasické definice pravděpodobnosti (viz část 1.2.3) jako podíl příznivých a všech možných případů, o nichž předpokládá, že *mohou nastat stejně snadno*.

Je však třeba zdůraznit, že Bernoulli nechápal teorii pravděpodobnosti jen jako nástroj pro rozbor hazardních her, ale mnohem obecněji. Již název spisu, který bychom mohli přeložit jako *Umění domněnky*, napovídá, že se jeho autorovi jednalo o usuzování v nejistých situacích nejrůznějšího druhu a že měl blízko k tzv. logické interpretaci pravděpodobnosti (viz část 1.3.2). Pravděpodobnost považoval za míru naší neúplné znalosti. V úvodu čtvrté části napsal:

Všechno pod Sluncem, co existuje nebo vzniká, všechno minulé, současné a budoucí má samo o sobě maximální jistotu. Co se týče současných a minulých věcí, je toto tvrzení jasné, protože každá věc tím, že existuje nebo existovala, vylučuje možnost, že by neexistovala. Ani v případě věcí budoucích však nelze pochybovat o tom, že nastanou, když ne s neodvratnou nutností, pak na základě Boží prozřetelnosti a předurčenosti. Protože kdyby se budoucí nemělo udát s jistotou, nebylo by jasné, proč by Nejvyššímu Stvořiteli měla příslušet neomezená sláva vševědoucnosti a všemohoucnosti.[44]

Podle Bernoulliho bychom například zítřejší počasí nebo výsledek hodu mincí byli schopni předpovědět stejně dobře jako zatmění planety, kdybychom jen uměli přesně změřit počáteční podmínky a znali všechny fyzikální zákony, jimiž se dané jevy řídí. Znalosti založené na zkušenosti jsou však obecně nejisté, a tak je teorie pravděpodobnosti použitelná ve všech empirických vědách i v běžném životě.

Výpočet pravděpodobnosti – váhy různých důvodů

Bernoulli tedy chápe pravděpodobnost jako stupeň jistoty, přisouzený určité domněnce na základě dostupné evidence. Numerické vyjádření pravděpodobnosti proto spočívá ve vyhodnocení známých důvodů svědčících pro a proti dané

[43] [29], str. 211–212; tento a následující citáty byly přeloženy s přihlédnutím k existujícím překladům do němčiny ([31]), ruštiny ([32]) a angličtiny ([33], [34]).

[44] Dalším filosofickým problémům se Bernoulli vyhnul poznámkou: *Avšak o tom, jak se tato jistota budoucího bytí slučuje s náhodností a nezávislostí působících příčin, nechť se dohadují jiní; my se tím nebudeme zabývat, protože je to příliš vzdálené od našeho cíle.* Viz [29], str. 210–211.

hypotéze, ve vyčíslení vah jednotlivých argumentů a jejich složení ve výslednou pravděpodobnost. Bernoulli podrobně rozebírá podmínky, které je přitom třeba dodržet, a potom odvozuje, jak hledanou pravděpodobnost ovlivňují různé typy důvodů, které rozlišuje podle toho, zda pro danou domněnku svědčí nutně nebo nikoli a zda je jejich samotná existence nutná nebo ne.

Toto rozlišení objasňuje na jednoduchém příkladu, v němž uvažuje tři možné důvody, proč mu bratr už dlouho nenapsal. Prvním důvodem by mohla být jeho *lenost*; tento důvod je podle něj přítomen nutně (Bernoulli podotýká, že považuje za známé, že bratr je líný), ale v psaní mu zabránit nemusel. Druhým důvodem by mohla být bratrova *smrt*; přítomnost tohoto důvodu je nejistá, bratr může být stále naživu. Pokud by však byl mrtev, jistě by nic nenapsal. Třetím důvodem by mohla být bratrova *zaneprázdněnost*; zde není jisté ani to, že je bratr skutečně zaneprázdněn, ani to, že by mu v kladném případě zaneprázdněnost zabránila v psaní.

Dále Bernoulli rozlišoval důvody *smíšené* a *ryzí* podle toho, zda v případech, kdy nesvědčí o pravdivosti dané domněnky, svědčí o pravdivosti hypotézy opačné, anebo zda z hlediska pravdivosti nesvědčí o ničem. Jako příklad uvádí situaci, kdy na ulici dojde k vraždě a očití svědkové uvedou, že pachatel měl tmavý plášť. Představme si, že z lidí, kteří mohli čin spáchat, měli takový plášť čtyři. Pro libovolného z nich je tmavý plášť smíšeným argumentem, který v jednom případě svědčí o jeho vině a ve třech o jeho nevině. Naproti tomu například zblednutí obviněného při výslechu je argumentem ryzím: je-li příčinou špatné svědomí, svědčí o jeho vině; je-li bledost způsobena něčím jiným, pak nikterak nepodporuje domněnku, že dotyčný je nevinen.

Při konkrétním vyčíslování pravděpodobnosti Bernoulli pro každý důvod uvažuje počty případů, kdy existuje či neexistuje, a počty případů, kdy danou domněnku dokazuje či nedokazuje, a využívá Huygensův výraz (1.9) pro očekávání. Přitom předpokládá, že jsou všechny případy *stejně možné*, neboli mohou nastat *stejně snadno*. Například pro důvod, který existuje nutně a o domněnce svědčí v p z celkových $n = p + q$ případů, Bernoulli ve výrazu (1.9) klade $a = 1$, $b = 0$ a uzavírá, že daný důvod *dokazuje p/n jistoty věci*.[45] Je-li důvod smíšený, pak navíc *dokazuje q/n jistoty opaku.*

„Sílu" důvodu, který existuje v σ z celkových $\nu = \sigma + \tau$ případů a o hypotéze svědčí s jistotou, Bernoulli podobně vyčísluje jako podíl $(\sigma \cdot 1 + \tau \cdot 0)/\nu = \sigma/\nu$. Při stejném značení uvažujme důvod, který o hypotéze svědčí v p z celkových $n = p + q$ případů; Bernoulli říká, že *hodnota tohoto důvodu pro důkaz věci je*[46] $(\sigma \cdot p/n + \tau \cdot 0)/\nu = (p\sigma)/(n\nu)$, *a je-li smíšený, jeho hodnota pro důkaz opaku je* $(\sigma \cdot q/n + \tau \cdot 0)/\nu = (q\sigma)/(n\nu)$.

Při současné existenci dvou [nezávislých] ryzích důvodů Bernoulli uvažuje takto: druhý důvod svědčí pro danou domněnku v p_2 z celkových n_2 případů; ve zbývajících q_2 případech nesvědčí o ničem, pro domněnku však může svědčit

[45] Viz [29], str. 219. Značení je upraveno tak, aby odpovídalo výrazu (1.9).

[46] Ve výrazu vpravo bychom dnes viděli součin podmíněné pravděpodobnosti p/n, že důvod svědčí o dané hypotéze, pokud je přítomen, a pravděpodobnosti σ/ν, že přítomen je.

ještě první důvod. Výsledná váha je proto rovna[47]

$$\frac{p_2 \cdot 1 + q_2 \cdot \dfrac{p_1}{n_1}}{n_2} = \frac{(n_2 - q_2) \cdot n_1 + q_2 \cdot (n_1 - q_1)}{n_1 n_2} = \frac{n_1 n_2 - q_1 q_2}{n_1 n_2} = 1 - \frac{q_1 q_2}{n_1 n_2} \,.$$

Podobně pro tři ryzí důvody dostává výslednou váhu $1 - (q_1 q_2 q_3)/(n_1 n_2 n_3)$. Pro libovolný počet ryzích důvodů pak uvádí: *všechny tyto argumenty vedou k pravděpodobnosti, která se liší od úplné jistoty nebo od jednotky o takovou část, která je získána jako podíl součinu případů, které nic nedokazují, a všech případů ve všech argumentech.* ([29], str. 220)

Toto formulaci můžeme vyjádřit vzorcem, který odpovídá Šimerkově *výsledné síle přesvědčení* – viz (4.4) na str. 149:

$$V = 1 - \frac{q_1 q_2 \dots q_k}{n_1 n_2 \dots n_k} \,. \tag{1.12}$$

Při současném výskytu tří smíšených argumentů Bernoulli vyjadřuje výslednou *pravděpodobnost* domněnky, resp. jejího opaku, ve tvaru[48]

$$\frac{p_1 p_2 p_3}{p_1 p_2 p_3 + q_1 q_2 q_3} \,, \qquad \text{resp.} \qquad \frac{q_1 q_2 q_3}{p_1 p_2 p_3 + q_1 q_2 q_3} \,,$$

který lze snadno zobecnit na libovolný počet takovýchto důvodů. Konečně uvažuje kombinaci důvodů obou typů a na příkladu tří ryzích a dvou smíšených důvodů ukazuje, jak se nalezne výsledná pravděpodobnost:

$$\frac{(n_1 n_2 n_3 - q_1 q_2 q_3) \cdot 1 + q_1 q_2 q_3 \cdot \dfrac{p_4 p_5}{p_4 p_5 + q_4 q_5}}{n_1 n_2 n_3} = 1 - \frac{q_1 q_2 q_3 q_4 q_5}{n_1 n_2 n_3 (p_4 p_5 + q_4 q_5)} \,.$$

Bernoulli zdůraznil, že je třeba vždy uvážit nejen všechny ryzí a smíšené argumenty svědčící o dané domněnce, ale i všechny ryzí argumenty svědčící o domněnce opačné, a podotkl, že obě pravděpodobnosti mohou *přesáhnout polovinu jistoty*. Jsou-li tyto hodnoty různé, lze jednu z alternativ prohlásit za pravděpodobnější, a to tolikrát, kolikrát vyšší je její pravděpodobnost. Vyjdeli například pravděpodobnost domněnky 2/3 a pravděpodobnost domněnky opačné 3/4, je daná domněnka méně pravděpodobná než její opak v poměru 2/3 ku 3/4, tj. 8 ku 9. Bernoulliho pravděpodobnosti tedy nemusejí splňovat obvyklou podmínku aditivity.

V závěru uvedené části Bernoulli také upozorňuje, že při využívání popsaných pravidel je třeba postupovat opatrně – nehovoří sice explicitně o „nezávislosti“ argumentů, na příkladech však takový požadavek popisuje.

---

[47] Povšimněme si, že předposlední zlomek vyjadřuje přímo podíl počtu případů $n_1 n_2 - q_1 q_2$, kdy o domněnce svědčí aspoň jeden z důvodů, a počtu všech případů $n_1 n_2$; poslední výraz pak udává doplněk pravděpodobnosti $(q_1 q_2)/(n_1 n_2)$, že pro domněnku nesvědčí ani jeden z důvodů, do jedné.

[48] Uvědomme si, že nemají-li si argumenty odporovat, musí všechny svědčit buď pro danou domněnku, nebo proti ní; celkový počet případů je proto roven $p_1 p_2 p_3 + q_1 q_2 q_3$.

Experimentální nalezení pravděpodobnosti

Bernoulli rozlišuje pravděpodobnost *apriorní* a *posteriorní*.[49] V prvním případě má na mysli pravděpodobnost, kterou lze určit deduktivně, na základě úvah o symetrii apod. Přitom podotýká, že tímto způsobem se pravděpodobnost podaří vypočítat u náhodných her, které byly již od počátku navrženy tak, aby bylo možné porovnat šance hráčů na výhru, ale jinak už téměř nikdy. Například při hodu kostkou existuje tolik případů, kolik má kostka stěn, a jsou-li všechny stěny shodné a hmotnost je rozložena rovnoměrně, *není důvod, aby jedna stěna padla snáze než jiné*. Podobně při losování z urny, v níž je známý počet bílých a černých lístků, mohou být všechny lístky vytaženy *stejně snadno*, protože *nelze uvést žádný důvod, proč by tento nebo onen lístek měl být vytažen snáze než jiný*. Jakmile však daný jev závisí na vlivu přírody nebo lidské vůli, jsou podobné úvahy obtížné:

Kdo ze smrtelníků je například schopen určit počet nemocí, [...] *které v každém věku mohou napadnout nesčetné části lidského těla a způsobit smrt; a kolikrát snáze může člověka zabít jedna nemoc (například mor) než jiná (například vzteklina; anebo vzteklina než horečka)? Kdo pak může na tomto základě něco usuzovat o budoucím stavu života a smrti?* ([29], str. 224)

Nyní se Bernoulli dostává asi k nejslavnějšímu výsledku své práce. Říká, že tam, kde se nám nepodaří pravděpodobnost vypočítat *a priori*, můžeme ji určit alespoň *a posteriori*, tedy induktivně, na základě pozorování mnoha podobných jevů. Přitom hovoří přímo o *empirické metodě určení počtu případů pomocí experimentu* a zdůrazňuje, že tento přístup není nikterak nový a každý jej běžně intuitivně používá:

Jestliže například bylo v minulosti pozorováno, že ze sledovaných 300 mužů stejného stáří a konstituce jako Titius jich do deseti let zemřelo 200, můžeme z toho usuzovat s dostatečnou důvěrou, že i pro Titia existuje dvakrát tolik případů, kdy zemře do deseti let, než případů, kdy bude žít déle. (tamtéž)

Bernoulli byl ovšem první, kdo tomuto přístupu dal matematický základ: s využitím několika pomocných tvrzení týkajících se binomického rozvoje dokázal větu, kterou později Poisson nazval *zákonem velkých čísel* (viz [285]). Je však třeba podotknout, že tato věta hovoří o tom, jaké výsledky pozorování můžeme za daných podmínek očekávat, nikoli o tom, jaké závěry o pravděpodobnosti můžeme učinit z provedeného pozorování.

Bernoulliův zákon velkých čísel

Základní myšlenku bychom mohli zjednodušeně popsat takto: pravděpodobnost, že při n [nezávislých] opakováních téhož pokusu nastane uvažovaný jev právě m-krát,[50] je rovna

$$\frac{n!}{m!(n-m)!}\, r^m s^{n-m} : t^n, \tag{1.13}$$

[49] Jak bude hned patrné, význam těchto pojmů se liší od moderní bayesovské terminologie a odpovídá tomu, s čím jsme se setkali v části 1.1.4 u učebnic z 19. století.

[50] Požadavek nezávislosti Bernoulli výslovně neuvádí, můžeme jej však odhalit v předpokladu, že počet příznivých a nepříznivých případů je stálý. V části 1.2.3 uvidíme, že explicitně o pojmu nezávislosti hovoří Abraham de Moivre – viz str. 46.

kde stejně jako v (1.10) značí r počet příznivých případů při každé realizaci pokusu, s počet případů nepříznivých a $t = r+s$. Po vynásobení číslem t^n, které nezávisí na m, představují výrazy (1.13) členy binomického rozvoje $(r+s)^n$. Bernoulli uvažoval speciálně $n = \nu(r+s) = \nu t$, kde ν je přirozené číslo, a ukázal, že největší z uvedených členů je ten, kde $m = \nu r$, tedy

$$\frac{n!}{(\nu r)!(\nu s)!}\, r^{\nu r} s^{\nu s}\,. \tag{1.14}$$

Potom dokázal, že součet $2\nu + 1$ členů „kolem" tohoto maxima,

$$\sum_{k=0}^{2\nu} \frac{n!}{(\nu(r+1)-k)!\,(\nu(s-1)+k)!} r^{\nu(r+1)-k}\, s^{\nu(s-1)+k}\,, \tag{1.15}$$

je alespoň c-krát větší než součet členů zbývajících, je-li

$$\begin{aligned} &\nu \geq \frac{\log c + \log(r-1)}{\log(s+1) - \log s}\left(1 + \frac{r}{s+1}\right) - \frac{r}{s+1} \\ \text{a zároveň}\quad &\nu \geq \frac{\log c + \log(s-1)}{\log(r+1) - \log r}\left(1 + \frac{s}{r+1}\right) - \frac{s}{r+1}\,. \end{aligned} \tag{1.16}$$

Součet (1.15) udává až na násobek t^n pravděpodobnost, že m leží v intervalu $[\nu(r-1),\ \nu(r+1)]$, tj. $m/(\nu t)$ leží v intervalu $[(r-1)/t,\ (r+1)/t]$. Pro všechna ν splňující podmínky (1.16) tedy platí:

$$P\left(\frac{r-1}{t} \leq \frac{m}{\nu t} \leq \frac{r+1}{t}\right) \geq c \cdot P\left(\frac{m}{\nu t} < \frac{r-1}{t} \vee \frac{r+1}{t} < \frac{m}{\nu t}\right), \tag{1.17}$$

což je předmětem Bernoulliovy věty:

Nechť je počet příznivých případů ku počtu případů nepříznivých přesně nebo přibližně v poměru $\frac{r}{s}$ a nechť je počet příznivých případů ku počtu všech případů v poměru $\frac{r}{r+s}$ nebo $\frac{r}{t}$, omezeném hodnotami $\frac{r+1}{t}$ a $\frac{r-1}{t}$. Má být dokázáno, že je možné učinit tolik pokusů, aby bylo libovolněkrát pravděpodobnější (řekněme c-krát), že počet příznivých pozorování padne mezi tyto hranice, než že padne mimo ně, tj. že poměr počtu příznivých ku počtu všech pozorování nebude větší než $\frac{r+1}{t}$ ani menší než $\frac{r-1}{t}$. ([29], str. 236)

Dodejme, že nerovnost (1.17) můžeme po úpravě vyjádřit ve tvaru:[51]

$$P\left(\left|\frac{m}{n} - \frac{r}{t}\right| \leq \frac{1}{t}\right) \geq 1 - \frac{1}{1+c}\,. \tag{1.18}$$

Zlomek m/n bychom dnes nazvali *relativní četností* jevu A v posloupnosti n nezávislých opakování pokusu; označme jej symbolem $h_A(n)$. Označíme-li dále

[51] Vztah (1.17) lze přepsat takto:

$$P\left(\left|\frac{m}{n} - \frac{r}{t}\right| \leq \frac{1}{t}\right) \geq c \cdot P\left(\left|\frac{m}{n} - \frac{r}{t}\right| > \frac{1}{t}\right) = c \cdot \left[1 - P\left(\left|\frac{m}{n} - \frac{r}{t}\right| \leq \frac{1}{t}\right)\right].$$

$p = r/t$, $\varepsilon = 1/t$, $\delta = 1/(1+c)$, obdržíme obvyklý tvar Bernoulliova zákona velkých čísel:

$$P\left(|h_A(\nu) - p\,| \le \varepsilon\right) \ge 1 - \delta\,, \tag{1.19}$$

kde se požaduje, aby nerovnost platila pro libovolná ε, $\delta > 0$ a dostatečně vysoký počet opakování n. Jak jsme viděli výše, Bernoulli uvažoval jen racionální hodnoty pravděpodobnosti p, vyjádřené zlomkem r/t, kde r, t jsou přirozená čísla, počet pokusů ν byl přirozeným násobkem čísla t a hodnota ε byla speciálně rovna $1/t$. Velikost ε Bernoulli libovolně zmenšoval tím, že zlomek r/t rozšiřoval vhodným násobkem 10.

Nalezením podmínek (1.16) učinil Bernoulli důležitý krok na cestě k číselnému vyjádření nejistoty. Z praktického hlediska byl však požadovaný počet pokusů příliš vysoký. Například pro $r/s = 3/2$ Bernoulli kladl $r = 30$, $s = 20$ a ukázal, že *při uskutečnění 25 550 pokusů bude více než 1000krát pravděpodobnější, že poměr počtu příznivých ku počtu všech pozorování padne mezi hodnoty* 31/50 *a* 29/50, *než že padne mimo ně* ([29], str. 238). Podle (1.16) pak vypočítal, že pro $c = 10\ 000$ by muselo být uskutečněno alespoň 31 258 pokusů a pro $c = 100\ 000$ dokonce 36 966.

Za tímto příkladem a krátkým komentářem celý spis poměrně nečekaně končí. Jak poznamenal S. M. Stigler v knize [355], jedním z důvodů mohl být právě vysoký počet pokusů. Hodnota 25 550 například převyšovala počet obyvatel Basileje; katalog hvězd, který sestavil anglický astronom John Flamsteed (1646–1719), obsahoval jen 3000 položek. V příkladu, který uvedl na začátku této části, Bernoulli přitom uvažoval vzorek 300 mužů. I kdybychom upustili od požadavku *morální jistoty* vyjádřené poměrem 1000/1001 a položili například jen $c = 1$, počet pokusů splňující podmínky (1.16) by stále přesahoval 8400.

Jak již bylo zmíněno, název čtvrté části spisu *Ars conjectandi* [29] avizoval *využití předchozí teorie v občanských, etických a ekonomických záležitostech.* Tento cíl se Bernoullimu bohužel nepodařilo splnit a jak o tom svědčí i jeho korespondence s Leibnizem, absence reálných numerických příkladů byla dalším z důvodů, proč spis nedokončil a nevydal již za svého života. Bernoulli Leibnize v dopisech opakovaně žádal o zaslání knížky [389] Johana de Witta (1625–1672) *o přednosti doživotních důchodů ve srovnání s důchody vypověditelnými* a o uvedení příkladů z oblasti práva, v níž mu chybělo systematické vzdělání. Leibniz Wittův spis nenašel, jen Bernoullimu napsal, že by pro něj stejně nebyl příliš zajímavý (podrobněji viz [361]). Je škoda, že jej alespoň neupozornil na pojednání [135], které v roce 1693 uveřejnil Edmond Halley (1656–1742) a které je všeobecně uznáváno jako základ správné teorie pojištění důchodu; Halley navíc na rozdíl od de Witta vycházel z reálných dat, totiž z přehledu narození a úmrtí ve Vratislavi, který získal právě díky Leibnizovi (viz [191] a [254]).

Inverzní pravděpodobnost

Bernoulli ve své větě i podmínkách (1.16) pracuje se známou hodnotou pravděpodobnosti p sledovaného jevu v jednotlivých pokusech. I když se zdá, že měl původně ambici odpovědět na otázku, jaké závěry o pravděpodobnosti daného jevu lze učinit na základě provedených experimentů, odpověděl v dokázaném tvrzení „jen“ na otázku, jaké výsledky lze za daných podmínek při pokusech očekávat. Zajímavou poznámku o řešení tzv. inverzního problému lze nalézt v knize [137], kterou v roce 1749 vydal anglický filosof a psycholog David Hartley (1705–1757):

Geniální přítel mne seznámil s řešením opačného problému, ve kterém ukázal, jaká je v případě, že určitý jev p-krát nastal a q-krát nenastal, naděje, že původní poměr příčin pro uskutečnění a neuskutečnění jevu se bude od poměru p ku q lišit o libovolnou danou hodnotu. A z tohoto řešení je patrné, že pro velký počet pokusů musí být odchylka zanedbatelná, což ukazuje, že máme šanci rozhodnout [...] *o neznámých příčinách na základě dostatečného množství pozorování jejich účinků.* ([137], str. 339)

Otázka, kdo je oním *geniálním přítelem*, jehož jméno není v textu uvedeno, je dodnes předmětem diskuse, která se rozhořela v 80. letech 20. století poté, co B. Singer upozornil v knize [347] na citovanou pasáž z Hartleyova spisu. S. M. Stigler v článku [353] vyjádřil a podrobně zdůvodnil názor, že zmiňovaným přítelem byl Nicholas Saunderson (1682–1739) nebo Thomas Bayes (1702(?)–1761); první jméno přitom pokládal za třikrát pravděpodobnější než druhé. A. W. F. Edwards v článku [85] oponoval tím, že Hartley nemusel chápat inverzní problém v našem smyslu, a přiklonil se k názoru, že zmíněným přítelem byl Abraham de Moivre (1667–1754). A. I. Dale pak v článku [68] předložil argumenty podporující domněnku, že se jednalo o Bayese. Dodejme, že podobné úvahy lze nalézt ve spise [129] Wilhelma Jacoba s'Gravesanda (1688–1742), který vyšel v roce 1736 a o kterém se zmíníme v části 3.2.2 (viz str. 131).

1.2.3 Klasická definice pravděpodobnosti

Abraham de Moivre

Obvyklé znění tzv. klasické definice pravděpodobnosti lze nalézt ve spisech Abrahama de Moivrea (1667–1754), francouzského matematika žijícího v Anglii.[52] V latinské práci *De Mensura Sortis ...* [244], která vyšla v roce 1711 a jejíž plný název v překladu zní: *O míře náhody nebo o pravděpodobnosti jevu v hrách na náhodě založených*, Moivre hned za předmluvou píše: [53]

Nechť r je počet případů, ve kterých nějaký jev může nastat, a s počet případů, ve kterých nastat nemůže; jak možnost, tak nemožnost jevu mají určitý stupeň pravděpodobnosti: mohou-li všechny tyto případy nastat stejně snadno,

[52] Abraham de Moivre pocházel z rodiny francouzského protestantského lékaře. Od roku 1681 studoval filosofii na protestantské akademii v Saumuru, kde se seznámil s Huygensovou prací [160], od roku 1684 studoval matematiku a fyziku v Paříži. Hned v následujícím roce však byl zrušen edikt nantský (viz pozn. 33 na str. 35) a Moivre byl po odmítnutí konverze vězněn v převorství Saint-Martin. Po propuštění v roce 1688 uprchl do Londýna, kde zůstal až do konce života. Jako cizinci se mu nikdy nepodařilo získat místo na univerzitě ani jiné stálé zaměstnání, přestože byl vynikajícím a uznávaným matematikem, členem předních vědeckých organizací a měl přátele a známé ve vědeckých kruzích i ve vysoké anglické společnosti. Svůj život tak prožil jako soukromý učitel šlechtických žáků; přivydělával si také poskytováním rad hazardním hráčům a soukromým pojišťovatelům a šachovou hrou. Podrobněji viz článek [322] I. Saxla a L. Ilucové nebo knihu [24] D. R. Bellhouse.

[53] Moivre značil počet příznivých, resp. nepříznivých případů písmeny p, q. Značení bylo upraveno, aby odpovídalo předchozí části a aby nedocházelo k záměně s pravděpodobností.

pak pravděpodobnost možnosti bude k pravděpodobnosti nemožnosti ve stejném poměru jako r ku s. ([245], str. 215)

Při výpočtech potom pracuje s pojmem *očekávání*, který zavádí takto: za výše uvedených podmínek uvažujme dva hráče, A a B, kteří hrají o předem složenou částku a. Jestliže daný jev nastane, vyhraje částku a hráč A, jestliže jev nenastane, vyhraje ji hráč B. Moivre říká, že očekávání hráče A, resp. B, je

$$\frac{ra}{r+s}, \quad \text{resp.} \quad \frac{sa}{r+s}\,. \tag{1.20}$$

My bychom dodali, že výrazy (1.20) odpovídají součinu hodnoty výhry a pravděpodobnosti, že ji daný hráč získá, tj.

$$\frac{r}{r+s}, \quad \text{resp.} \quad \frac{s}{r+s}\,, \tag{1.21}$$

která (až na násobek konstantou) přímo plyne z citované definice.

V tomto smyslu Moivre zavádí pojem očekávání v anglicky psané knize *The Doctrine of Chances* [245], která podstatně rozšiřuje výsledky pojednání [244]. První vydání tohoto spisu vyšlo v roce 1718 a jeho úvodní kapitola začíná slovy:

Pravděpodobnost jevu je větší nebo menší podle počtu možností, kdy nastane, ve srovnání s počtem všech možností, při nichž nastane nebo nenastane. Jsou-li tedy 3 možnosti, kdy jev nastane, a 2 možnosti, kdy nenastane, může být pravděpodobnost, že jev nastane, odhadnuta číslem $\frac{3}{5}$, *a pravděpodobnost, že nenastane,* $\frac{2}{5}$ *. Proto sečtou-li se pravděpodobnosti uskutečnění a neuskutečnění jevu, bude součet vždy roven jedné.* ([245], str. 1)

Potom Moivre zavádí pojmy *šance* a *očekávání*:

[...] *šance pro nebo proti jsou úměrné počtu možností, při nichž jev nastane nebo nenastane. Očekávání nějaké věci je odhadnuto hodnotou této věci, vynásobenou pravděpodobností jejího získání.* ([245], str. 2)

V dalších vydáních je za úvodní odstavec vloženo ještě upřesnění, které již představuje typickou formulaci klasické definice pravděpodobnosti:[54]

Sestavíme-li tedy zlomek, jehož čitatelem bude počet možností, kdy jev nastane, a jmenovatelem počet všech možností, kdy nastane nebo nenastane, bude tento zlomek vhodným stanovením pravděpodobnosti uskutečnění jevu.
([245], 2. a 3. vydání, str. 1)

Po zavedení výše uvedených pojmů Moivre již v prvním vydání hovoří o jejich využití v oblasti pojišťování. Cenu pojištění navrhuje stanovit tak, aby byl poměr pojistného a pojištěného jmění roven pravděpodobnosti, že bude jmění ztraceno. Dále se potom zabývá závislými a nezávislými jevy. Kromě jiného odvozuje a vysvětluje na řadě příkladů, že není-li mezi dvěma jevy *žádná závislost*, vypočítá se pravděpodobnost, že nastanou oba dva, jako součin jejich pravděpodobností, a nastiňuje možnost zobecnění na libovolný počet jevů. Pravděpodobnost opačného jevu pak počítá podle výše uvedeného tvrzení jako doplněk do jedné.

[54] Předpoklad, že všechny případy jsou „stejně možné", zde Moivre na rozdíl od původní latinské verze explicitně neuvádí, prováděné výpočty mu však odpovídají.

Moivreovy matematické výsledky jsou popsány v citovaném článku [322] a dále například v publikacích [24], [133] a [337]. Zde jen s využitím dnešní terminologie uveďme, že již v latinské práci [244], publikované ještě před vydáním Bernoulliova spisu *Ars conjectandi* [29], Moivre zavedl binomické rozdělení $\mathrm{Bi}(n, p)$ (viz (1.11)) a dospěl i k jeho limitnímu případu pro $p \to 0$, $n \to \infty$ a $np = \lambda$ (pevné), tedy k rozdělení Poissonovu $\mathrm{Po}(\lambda)$, kde pravděpodobnost, že náhodná veličina X nabude hodnoty m, je rovna $P_\lambda(X = m) = (\lambda^m/m!)e^{-\lambda}$.

Z hlediska Bernoulliova zákona velkých čísel je velmi zajímavé pojednání *Approximation Summam Terminorum Binomii* $(a+b)^n$ *in Seriem*, které Moivre v roce 1733 v několika exemplářích vytiskl a rozdal přátelům; později tuto práci přeložil do angličtiny a zařadil ji do druhého vydání spisu *The Doctrine of Chances* [245] z roku 1738. Pro speciální případ $(1 + 1)^n$, tedy pro $p = 1/2$, Moivre odvodil aproximaci, kterou bychom mohli přepsat ve tvaru

$$P\left(\left|\frac{x}{n} - \frac{1}{2}\right| \le \frac{t}{\sqrt{n}}\right) \approx \frac{4}{\sqrt{2\pi n}} \int_0^{t\sqrt{n}} e^{-\frac{2x^2}{n}}\, dx = \frac{4}{\sqrt{2\pi}} \int_0^t e^{-2y^2}\, dy\,. \tag{1.22}$$

Bez důkazu potom uvedl také vztah pro asymetrické binomické rozdělení, tedy pro $p \neq 1/2$, a jako limitní případ odvodil rozdělení normální. Používal rovněž *De Moivreovu-Laplaceovu větu*, kterou bychom dnes zapsali takto:

$$P\left(\left|\frac{x}{n} - p\right| \le 2t\sqrt{\frac{pq}{n}}\right) \approx \frac{4}{\sqrt{2\pi}} \int_0^t e^{-2y^2}\, dy\,. \tag{1.23}$$

Na závěr dodejme, že do druhého vydání spisu [245], které vyšlo v roce 1738, Moivre zařadil také rozsáhlé pojednání o životním pojištění.

Thomas Bayes

Na konci části 1.2.2 jsme se zmínili o diskusi na téma, kdo jako první přišel s myšlenkou inverzní pravděpodobnosti. Jak však zdůraznil Andrew I. Dale v knize [69], na metodě jejího nalezení má přece jen zcela zásadní zásluhu Thomas Bayes (1702(?)–1761).[55] Věta, která nese jeho jméno, je navíc důležitá pro epistemologické interpretace pravděpodobnosti. Věnujme proto této osobnosti několik odstavců.

Tři roky po Bayesově smrti vyšlo zásluhou Richarda Price (1723–1791) v časopise Philosophical Transactions jeho pojednání *Essay towards Solving a Problem in the Doctrine of Chances* [17], které začíná slovy:

Je dáno, kolikrát se neznámý jev uskutečnil a neuskutečnil: Požadována je šance, že pravděpodobnost jeho uskutečnění v jednotlivém pokusu leží mezi libovolnými dvěma stupni pravděpodobnosti.

Potom Bayes uvádí poněkud nezvyklou definici pravděpodobnosti:

Pravděpodobnost nějakého jevu je poměr mezi hodnotou očekávání, závisející na uskutečnění tohoto jevu, a hodnotou věci při uskutečnění očekávané.[56]
([17], str. 376)

[55] Thomas Bayes studoval teologii na univerzitě v Edinburku a později působil jako presbyteriánský duchovní v anglickém městečku Tunbridge Wells. Byl také uznávaným matematikem. Podrobné informace o Bayesově životě a díle lze nalézt v knize [68] A. I. Dalea.

[56] Jako bychom tedy hodnoty (1.20) vydělili číslem a.

Hlavním výsledkem Bayesova spisu je odpověď na otázku uvedenou v úvodu, kterou můžeme s využitím dnešního značení vyjádřit takto: představme si, že bylo provedeno n [nezávislých] pokusů, při kterých m-krát nastal určitý jev A a $(n-m)$-krát nastal jev opačný. Předpokládejme, že v každém z pokusů byla pravděpodobnost uskutečnění tohoto jevu rovna téže hodnotě x (apriorní pravděpodobnosti), která je neznámá, ale kterou považujeme za náhodnou veličinu s rovnoměrným rozdělením na intervalu $[0,1]$. Potom pro libovolná x_1, x_2, kde $0 \le x_1 < x < x_2 \le 1$, platí:

$$P(x_1 < x < x_2 \,|\, m, n) = \frac{\displaystyle\int_{x_1}^{x_2} x^m (1-x)^{n-m} dx}{\displaystyle\int_0^1 x^m (1-x)^{n-m} dx} . \tag{1.24}$$

Pierre Simon de Laplace

Důležitý mezník ve vývoji klasické teorie pravděpodobnosti představuje kniha *Théorie analytique des probabilités* [208] z roku 1812, v níž francouzský matematik a fyzik Pierre Simon de Laplace (1749–1827)[57] shrnul jednak své vlastní dřívější výsledky, jednak důležité výsledky svých předchůdců. Do druhého vydání, které vyšlo o dva roky později, Laplace navíc zařadil rozsáhlou úvodní stať *Essai philosophique sur les probabilités*, kterou téhož roku vydal i jako samostatnou publikaci [209]. Zde popsal základní pojmy, výsledky i filosofické problémy téměř bez použití matematických vzorců a své myšlenky tak zpřístupnil širšímu okruhu čtenářů. Uveďme nejprve Laplaceovu definici pravděpodobnosti, z níž se stala nejčastější formulace tzv. *klasické definice*:

Teorie pravděpodobnosti spočívá v redukci všech jevů, které mohou nastat v dané situaci, na jistý počet stejně možných případů, to je takových, že jejich uskutečnění je stejně nejisté, a ve stanovení těch případů, které jsou příznivé jevu, jehož pravděpodobnost hledáme. Poměr tohoto čísla k počtu všech možných případů je míra pravděpodobnosti, která je zlomkem, v jehož čitateli je počet příznivých případů a ve jmenovateli je počet všech případů. ([208], str. 178)

Z této definice pak Laplace odvozuje řadu tvrzení. První zní takto:

Pravděpodobnost jevu, složeného ze dvou jevů jednoduchých, je rovna součinu pravděpodobnosti jednoho z těchto jevů a pravděpodobnosti, že nastane-li tento jev, nastane i druhý. ([208], str. 181)

V dnešním značení:

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2|E_1) . \tag{1.25}$$

Potom Laplace pokračuje: *Pravděpodobnost budoucího jevu, vyvozená z určitého pozorovaného jevu, je podílem pravděpodobnosti jevu složeného z těchto*

[57] Pierre Simon de Laplace začal studovat teologii na univerzitě v Caen, před dokončením studia, v roce 1767, však odešel do Paříže za Jeanem d'Alembertem a začal se naplno věnoval matematice. Zanedlouho byl jmenován profesorem matematiky na École Militaire a začal produkovat velké množství pozoruhodných matematických prací. Ve 24 letech byl zvolen členem pařížské akademie; od roku 1795 vyučoval na nově založené École Normale.

dvou jevů, určené a priori, a pravděpodobnosti pozorovaného jevu, rovněž určené a priori.

Použijeme-li opět dnešní zápis, pak toto tvrzení můžeme vyjádřit vztahem:

$$P(H|E) = P(H \cap E)/P(E)\,. \tag{1.26}$$

Patrně nezávisle na Bayesově stati [17] Laplace dále dospěl k následující formulaci tzv. Bayesovy věty pro diskrétní náhodnou veličinu:

Může-li být pozorovaný jev výsledkem působení k různých příčin [...] *, potom je pravděpodobnost působení libovolné z nich rovna zlomku, v jehož čitateli je pravděpodobnost pozorovaného jevu, plynoucí z dané příčiny, a ve jmenovateli je součet všech obdobných pravděpodobností pro všechny příčiny.* ([208], str. 182)

V dnešním značení:

$$P(H_i|E) = \frac{P(E|H_i)}{P(E|H_1) + P(E|H_2) + \cdots + P(E|H_k)}\,, \tag{1.27}$$

kde H_i značí hypotézu, že působila příčina i, E je dostupná evidence, tedy pozorovaný výsledek. Přitom se předpokládá, že apriorní pravděpodobnosti $P(H_i)$ všech příčin jsou stejné. Laplaceův důkaz vychází ze vztahu (1.26) pro $H = H_i$, kde se na základě (1.25) položí $P(H_i \cap E) = P(H_i) \cdot P(E|H_i)$ a za $P(E)$ se dosadí součet těchto pravděpodoboстí. Jsou-li všechny příčiny *a priori stejně možné*, je $P(H_i) = 1/k$, a proto platí (1.27). Jestliže tento předpoklad vypustíme, dostaneme ihned obvyklý tvar Bayesovy věty (viz část 1.40, str. 60):

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{P(E|H_1)P(H_1) + P(E|H_2)P(H_2) + \cdots + P(E|H_k)P(H_k)}\,. \tag{1.28}$$

Ve stati [209] Laplace explicitně uvádí:

Jsou-li tyto různé příčiny, uvažované a priori, různě pravděpodobné, pak místo pravděpodobnosti výsledku, plynoucí z určité příčiny, je třeba uvažovat součin této pravděpodobnosti s možností příčiny samotné. ([209], str. 10)

O pravděpodobnosti určitého budoucího jevu B potom Laplace tvrdí:

$$P(B|E) = P(B|H_1)P(H_1|E) + \cdots + P(B|H_k)P(H_k|E)\,. \tag{1.29}$$

Připomeňme, že tvrzení o platnosti vztahu (1.28) se záhy začalo označovat jako *věta o pravděpodobnosti příčin.* Pro ilustraci uveďme následující příklad: předpokládejme, že určitý počet chorob $C_1, C_2, \ldots, C_k$ může být spojen s určitými symptomy $S_1,\ S_2,\ \ldots,\ S_n$. Pacient, u něhož byl pozorován symptom S_m, navštíví lékaře. Předpokládejme, že u každé nemoci je známa pravděpodobnost $P(S_j|C_i)$, že se u pacienta s nemocí C_i projevuje symptom S_j. Otázkou je, jaká je naopak „inverzní" pravděpodobnost, že pacient trpí určitou nemocí C_i, je-li u něj pozorován příznak S_m, tedy pravděpodobnost $P(C_i|S_m)$. Odpověď dává vzorec (1.28), kde E je dostupná evidence, tedy skutečnost, že se u pacienta projevuje symptom S_m, a H_i je hypotéza, že příčinou je nemoc C_i.

Dodejme, že dnes je tento vzorec základem tzv. teorie bayesovských metod, která má bohaté praktické aplikace a je všeobecně považována za užitečný statistický nástroj; na druhé straně má však také své odpůrce, kteří poukazují především na problém s apriorními pravděpodobnostmi $P(H_i)$, jež jsou často neznámé a jejich odhad bývá více či méně subjektivní.[58]

Pro spojitou náhodnou veličinu Laplace odvodil vzorec, jehož speciálním případem je vztah (1.24):[59]

$$P\,(x_1 < x < x_2 \,|\, E) = \int_{x_1}^{x_2} y dx \Big/ \int_0^1 y dx \ , \tag{1.30}$$

kde x je neznámá pravděpodobnost určitého *jednoduchého jevu* (například pravděpodobnost, že při libovolném z n opakování určitého pokusu nastane jev A – viz str. 48) a y je pravděpodobnost pozorovaného výsledku E pro danou hodnotu x (např. $y = \binom{n}{m} x^m (1-x)^{n-m}$). Laplace přitom opět vychází z předpokladu, že všechny hodnoty $x \in [0,1]$ jsou *stejně možné*. Vzápětí poznamenává, že kdyby tento předpoklad splněn nebyl, vynásobila by se funkce y v obou integrálech funkcí w proměnné x, *vyjadřující její pravděpodobnost*. Stejný výsledek bychom však dostali i v případě, že bychom vyšli z předpokladu rovnoměrného rozdělení a součin yw bychom považovali za pravděpodobnost, že nastanou současně dva nezávislé jevy, jejichž pravděpodobnosti jsou y a w. Proto Laplace rovnoměrnost rozdělení předpokládá i nadále.

Později odvozuje také vztah pro *celkovou pravděpodobnost budoucího výsledku* (např. pravděpodobnost, že v příštím pokusu nastane jev A), jehož pravděpodobnost pro danou hodnotu x je rovna z (viz [208], str. 394):

$$P\,(B|E) = \int_0^1 zy dx \Big/ \int_0^1 y dx \ . \tag{1.31}$$

Využití vzorce (1.31) Laplace ilustruje příkladem, v němž hledá pravděpodobnost, že při následujících m pozorováních nastane pokaždé určitý jev A, nastal-li zatím při každém z n pozorování:

$$P\,(B|E) = \frac{\displaystyle\int_0^1 x^m x^n dx}{\displaystyle\int_0^1 x^n dx} = \frac{\displaystyle\int_0^1 x^{m+n} dx}{\displaystyle\int_0^1 x^n dx} = \frac{n+1}{m+n+1} \,. \tag{1.32}$$

Ve stati [209] Laplace v souvislosti s tímto tématem uvádí známý příklad, k němuž se vrátíme ještě v části 3.2.2 (viz str. 133): Jestliže v uplynulých 5000 letech, tedy celkem 1 826 213krát, každý den vyšlo slunce, pak je pravděpodobnost, že vyjde i zítra, rovna 1 826 214/1 826 215.

[58] O Laplaceových výsledcích souvisejících s problémem inverzní pravděpodobnosti se lze dočíst v citované Daleově knize [69] nebo v článku [65] M. Čiháka.

[59] Laplace psal pouze podíl integrálů se slovním popisem mezí; levou stranu rovnice se zápisem podmíněných pravděpodobností jsme doplnili pro názornost. Viz [208], str. 363.

1.2.4 Axiomatická definice pravděpodobnosti

Hlavní potíž klasické definice pravděpodobnosti spočívá v tom, že předpokládá, že všechny výsledky jsou „stejně možné“, přitom se však náležitě nedefinuje, co to znamená. K rozhodnutí o tom, zda případy jsou stejně možné či nikoli, se často používá princip, označovaný jako *princip nedostatečného důvodu* (tento výraz použil J. von Kries v knize [199]) nebo též *princip indiference* (viz knihu [193] J. M. Keynese), podle něhož se jevy považují za „stejně možné“, není-li důvod předpokládat, že jeden nastane snáze než jiný. Tuto myšlenku lze objevit již u J. Bernoulliho (viz část 1.2.2) a potom explicitně u P. S. de Laplace. Bohužel, její použití vede k řadě paradoxů, kdy lze na základě principu indiference odvodit pro jeden jev více různých hodnot pravděpodobnosti. Jedním z nich je tzv. *Bertrandův paradox*, o němž je pojednáno v části 6.2.3.

Andrej Nikolajevič Kolmogorov

Všeobecně uznaným východiskem se nakonec stala axiomatická definice, kterou ve spise *Grundbegriffe der Wahrscheinlichkeitsrechnung* [196] z roku 1933 publikoval Andrej Nikolajevič Kolmogorov[60] a díky níž také byla teorie pravděpodobnosti přijata jako plnohodnotná matematická disciplína. Přitom je třeba připomenout, že tato kniha byla završením vývoje teorie míry a pravděpodobnosti první třetiny 20. století a lze v ní spatřovat určitou syntézu prací, které v uvedené době zveřejnili kromě jiných Émile Borel (1871–1956), Maurice Fréchet (1878–1973), Francesco Cantelli (1875–1966), Alexandr Alexandrovič Čuprov (1874–1926), Paul Lévy (1886–1971), Hugo Steinhaus (1887–1972), Stanisław Ulam (1909–1984) a Richard von Mises (1883–1953).[61] Uveďme zde původní znění Kolmogorovovy definice:

Definice (Kolmogorov). *Buď E množina prvků $\xi, \eta, \psi, \dots$, které se nazývají elementární jevy* [*elementare Ereignisse*], *a $\mathfrak{F}$ množina podmnožin množiny E; prvky množiny $\mathfrak{F}$ se v dalším budou nazývat náhodné jevy* [*zufällige Ereignisse*].

I. *$\mathfrak{F}$ je množinové těleso.*[62]
II. *$\mathfrak{F}$ obsahuje množinu E.*
III. *Každé množině A z $\mathfrak{F}$ je přiřazeno nezáporné reálné číslo $P(A)$. Toto číslo $P(A)$ se nazývá pravděpodobnost* (*Wahrscheinlichkeit*) *jevu A.*
IV. $P(E) = 1$.
V. *Jsou-li A a B disjunktní, pak platí $P(A + B) = P(A) + P(B)$.*

[60] Andrej Nikolajevič Kolmogorov (1903–1987) studoval na moskevské univerzitě matematiku, metalurgii a ruskou historii. Po absolutoriu v roce 1925 působil čtyři roky jako aspirant u Nikolaje Nikolajeviče Luzina (1883–1950), potom nastoupil jako vědecký pracovník do Ústavu matematiky a mechaniky moskevské univerzity. V roce 1931 podnikl studijní cestu do Německa a Francie a získal profesuru na moskevské univerzitě. V roce 1933 byl jmenován ředitelem zmíněného Ústavu matematiky a mechaniky, o dva roky později se stal vedoucím nově založené katedry teorie pravděpodobnosti. Od roku 1938 navíc vedl oddělení teorie pravděpodobnosti v Matematickém ústavu V. A. Stěklova Akademie věd SSSR.

[61] Podrobněji o této problematice pojednávají G. Shafer a V. Vovk v článku [333].

[62] Tj. $\mathfrak{F} \neq \emptyset$ a pro libovolné dvě množiny $A, B \in \mathfrak{F}$ je rovněž $A \cap B$, $A \cup B$, $A \setminus B \in \mathfrak{F}$; pro průnik, sjednocení a rozdíl množin Kolmogorov používal tehdy obvyklé značení AB, $A + B$ a $A - B$. Uvědomme si, že podmínky I a II říkají, že $\mathfrak{F}$ je algebra podmnožin množiny E.

VI. *Pro klesající posloupnost* $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ *jevů z* $\mathfrak{F}$, *kde* $\mathfrak{D}_n A_n = 0$,[63] *platí rovnost* $\lim P(A_n) = 0$.

Soustava množin $\mathfrak{F}$ *spolu s určitým přiřazením čísel* $P(A)$, *splňující axiomy I.–VI., se nazývá pravděpodobnostní pole* (*Wahrscheinlichkeitsfeld*).[64]

Po zavedení axiomu VI Kolmogorov dokazuje větu o spočetné aditivitě, kterou lze v dnešním značení přepsat takto: Jestliže $A_1, \ldots, A_n, \ldots \in \mathfrak{F}$, $\bigcup_n A_n \in \mathfrak{F}$ a $A_i \neq A_j$ pro všechna $i \neq j$, potom platí: $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

Pak poznamenává, že definici lze jednoduše vyjádřit slovy:

Buď E libovolná množina, $\mathfrak{F}$ *těleso podmnožin množiny* E, *které obsahuje množinu* E, *a* $P(A)$ *nezáporná úplná aditivní množinová funkce* [tj. σ-aditivní míra] *definovaná na* $\mathfrak{F}$; *těleso* $\mathfrak{F}$ *spolu s množinovou funkcí* $P(A)$ *pak tvoří pravděpodobnostní pole.* ([196], str. 15)

Nakonec Kolmogorov dokazuje, že každé pravděpodobnostní pole $(\mathfrak{F}, P)$ lze jednoznačně rozšířit na pravděpodobnostní pole $(B\mathfrak{F}, P)$, kde $B\mathfrak{F}$ je *Borelovo těleso*, tj. těleso uzavřené také na spočetná sjednocení; na tato tělesa se pak v dalším omezuje. Dnešními slovy bychom tedy mohli říci, že popsaným způsobem se Kolmogorov dostal k obvyklému zavedení pravděpodobnosti jako reálné funkce definované na σ-algebře $\mathfrak{F}$ podmnožin množiny E, která pro $A \in \mathfrak{F}$, $A_1, A_2, \ldots \in \mathfrak{F}$, kde $A_i \cap A_j = \emptyset$ pro všechna $i \neq j$, splňuje podmínky:

$$P(E) = 1; \quad P(A) \geq 0; \qquad P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)\,. \tag{1.33}$$

První ohlasy

V roce 1934 vyšly dvě pochvalné recenze Kolmogorovovy knihy [196]. První z nich byla otištěna v referativním časopise *Zentralblatt für Mathematik* a její autor, soukromý docent na univerzitě v Kielu Willy Feller (1906–1970), zde kromě jiného vyzdvihl, že Kolmogorov vystavěl teorii pravděpodobnosti axiomaticky, v plné obecnosti, bez mezer a jako součást teorie míry.[65] Druhou recenzi napsal Henry Lewis Rietz (1875–1943), profesor matematiky na univerzitě v Iowě, pro časopis *Bulletin of the American Mathematical Society*. Rietz je poněkud opatrnější, práci však hodnotí kladně a poznamenává, že podle jeho názoru přispěje k zajištění *logičtějšího vývoje teorie pravděpodobnosti*.[66] Další recenzenti, němečtí matematikové Gustav Doetsch[67] (1892–1977) a Karl Dörge[68] (1899–1977), byli spíše neutrální a poukazovali na právě vydaný spis *Grundlagen der Wahrscheinlichkeitsrechnung* [373] Erharda Torniera (1894–1982), který měl blíže k četnostnímu pojetí a nepožadoval axiom spočetné aditivity – podrobněji viz [142].[69]

---

[63] Tímto symbolem Kolmogorov značí množinový součin $A_1 A_2 \ldots A_n \ldots$, tj. průnik $\bigcap_n A_n$.

[64] [196], str. 2 (axiomy I–V) a str. 13 (axiom VI).

[65] Viz Zentralblatt für Mathematik und ihre Grenzgebiete 7 (1934), str. 216.

[66] Viz Bulletin of the American Mathematical Society 40 (1934), str. 522–523.

[67] Viz Jahresbericht der Deutschen Mathematiker-Vereinigung 45 (1933), str. 153.

[68] Viz Jahrbuch über die Fortschritte der Mathematik 59 (1933), str. 1152.

[69] Tornier v té době působil na univerzitě v Kielu, kam přišel v roce 1929 jako suplent za mimořádného profesora Roberta Schmidta (1898–1964). Od roku 1933 se jako člen NSDAP

Z prvních vědeckých článků, jež se explicitně hlásily ke Kolmogorovovým *Grundlagen* [196], zde uveďme pojednání, která v roce 1934 vydali Zbigniew Łomnicki and Stanisław Ulam ([217]), Eberhard Hopf ([143]) a Joseph L. Doob ([77], [78]). Konečně poznamenejme, že první učebnicí založenou na Kolmogorovově axiomatice byla kniha *Random Variables and Probability Distributions* [58], kterou v roce 1937 vydal Harald Cramér (1893–1985), profesor pojistné matematiky a matematické statistiky na univerzitě ve Stockholmu. Z Kolmogorovových axiomů vycházela také učebnice *Kurs teorii verojatnostej* [122] sovětského matematika a logika Valerije Ivanoviče Glivenka (1897–1940), vydaná v roce 1939, a v poválečném období pak další Cramérova kniha *Mathematical Methods of Statistics* [59], učebnice *Kurs teorii verojatnostej* [124] Kolmogorovova žáka Borise Vladimiroviče Gnedenka (1912–1995) či kniha *Probability Theory* [213], kterou v roce 1955 vydal matematik a statistik Michel Loève (1907–1979). V následujících letech se pak Kolmogorovova axiomatická teorie stala takřka nedílnou součástí vysokoškolské výuky teorie pravděpodobnosti.

Ohlasy v českých zemích: Karel Rychlík

V kontextu předchozího odstavce je velmi zajímavý pohled do českých zemí 30. let 20. století. Na uveřejnění Kolmogorovovy axiomatiky takřka okamžitě zareagoval Karel Rychlík (1885–1968),[70] který jako řádný profesor přednášel *Počet pravděpodobnosti* na pražské technice a k tomu jako soukromý docent vypisoval výběrové přednášky zejména z oblasti algebry a teorie čísel na pražské univerzitě.[71] Před zahájením zimního semestru školního roku 1933/34 zrušil původně vypsanou univerzitní výběrovou přednášku z lineární algebry a nahradil ji přednáškou *Úvod do počtu pravděpodobnosti (se stanoviska axiomatického)*. Přednášku s podobným názvem, *Počet pravděpodobnosti s hlediska axiomatického*, na univerzitě vypsal také v zimním semestru školního roku 1936/37.

V roce 1936, ještě před vydáním Cramérovy učebnice [58], byla do vydavatelského programu JČMF zařazena Rychlíkova skripta *Úvod do počtu pravděpodobnosti* [313],[72] založená na Kolmogorovově axiomatické teorii; tiskem tato kniha vyšla v roce 1938. Poznamenejme, že Rychlík zde nejprve zavádí základní

aktivně podílel na nacifikaci kielské univerzity a upozorněním na „neárijský“ původ přispěl i k nucenému odchodu W. Fellera, s nímž do té doby úzce spolupracoval. Feller potom působil v Kodani, ve Stockholmu (u H. Craméra) a Lundu a v roce 1939 emigroval do USA.

[70] Karel Rychlík studoval v letech 1904–1907 matematiku a fyziku na filosofické fakultě české univerzity v Praze, akademický rok 1907/08 pak strávil jako student na pařížské Sorbonně. Po návratu složil zkoušky učitelské způsobilosti pro učitelství na gymnáziích a reálkách. Od roku 1909 působil jako asistent na pražské univerzitě, kde ve stejném roce získal doktorát z matematiky a kde se o tři roky později habilitoval. Od roku 1913 působil jako placený asistent na české technice v Praze, kde byl po vypuknutí první světové války pověřen konáním základních kurzovních přednášek z matematiky a jednosemestrální přednášky o počtu pravděpodobnosti pro studenty pojistné matematiky. V roce 1920 byl na technice jmenován mimořádným profesorem, v roce 1923 se pak stal profesorem řádným; úvodní matematický kurz i počet pravděpodobnosti zde vyučoval až do uzavření vysokých škol na počátku druhé světové války. Podrobněji viz autorčinu monografii [161].

[71] V roce 1931/32 se jeho výběrová přednáška na univerzitě poprvé týkala pravděpodobnosti: na letní semestr vypsal přednášku *Počet pravděpodobnosti (Misesova teorie)*, která byla patrně inspirovaná nedávno vydanou Misesovou knihou [243].

[72] Viz ČPMF 66 (1937), str. D57.

pojmy (elementární jev, množina elementárních jevů aj.) pomocí teorie množin, potom se krátce zabývá klasickou definicí pravděpodobnosti. V dalším se obrací k axiomům pro rozložení pravděpodobnosti v množinovém tělese, pak se věnuje zobrazení a ekvivalenci pokusů, dále podmíněné pravděpodobnosti a odvození Bayesových vzorců, nezávislosti rozkladů a náhodných jevů, pojmům matematické naděje, rozptylu a vytvořující funkce. Potom se Rychlík zabývá spojitým rozložením pravděpodobnosti na přímce (neobjevuje se zde však přímo pojem spojité náhodné veličiny), spočetným rozložením pravděpodobnosti a zákonem velkých čísel. Zbývající část spisu je věnována posloupnostním modelům pro rozložení pravděpodobnosti a jejich aplikacím. V dodatku Rychlík uvádí přehled základních pojmů teorie množin; užitečný je i závěrečný, velmi podrobný seznam literatury. Oproti Kolmogorovově vědecké práci [196], kde bylo žádoucí podat teorii v plné obecnosti, se Rychlík omezuje v podstatě jen na konečné množiny, a tedy axiomy I–V; to je však dáno odlišným účelem práce.

Je třeba zdůraznit, že Rychlíkova kniha byla svou aktuálností v tehdejší české literatuře zcela ojedinělá a ojedinělou ještě poměrně dlouho zůstala. Kromě toho, že záhy po jejím uveřejnění zpřístupnila studentům Kolmogorovovu axiomatickou teorii pravděpodobnosti, spočívá její přínos i v tom, že vedle Kolmogorovovy abstraktní a skutečnosti poněkud vzdálené teorie věnuje poměrně velkou pozornost také o něco starší a v té době již převážně kritizované teorii Misesově a ukazuje její použitelnost v praktických aplikacích.

Záhy po vydání skript [313] vyšla velice podrobná a pochvalná recenze, kterou napsal Rychlíkův asistent na technice Otomar Pankraz (1903–1976), jenž zde ocenil jak Rychlíkovo zpracování problematiky, tak i Kolmogorovovo pojetí. Vyzdvihl zde například samotnou definici náhodného jevu jako množiny elementárních jevů, popsal Rychlíkovy axiomy a jeho posloupnostní model rozložení pravděpodobnosti a kromě jiného poznamenal: *Průkaznost nových metod jest nesporná a záleží nyní na didaktických výkladech, které by nejjednodušeji umožnily přesunouti myšlení ze starších forem k novým směrům. To jest již jen otázkou škol a jest sympatické, že České vysoké učení technické v Praze tyto nové metody podporuje. Spis prof. Rychlíka, ačkoli jest skromně označen jako Úvod, podává na četných místech původní autorovy úvahy a naznačuje, v jakém směru jest možno studium p. p. prohloubiti.*[73]

Otomar Pankraz

V letech 1939 a 1940 vydal Otomar Pankraz dvojici článků [P28] a [P30], v nichž Kolmogorovovu definici kritizoval a představil vlastní axiomatiku založenou na pojmu podmíněné pravděpodobnosti. Nezávisle na Pankrazovi podobnou axiomatickou definici o řadu let později představil Karl Popper (1902–1994) a Alfréd Rényi (1929–1970). Myšlenka, že pravděpodobnost by měla být definována jako dvouargumentová funkce, je živá i na počátku 21. století; výše zmíněné Pankrazovy práce však v širších kruzích zůstaly téměř neznámé. Podrobněji je o této problematice pojednáno v části 8.2.2.[74]

[73]ČPMF 67 (1938), D301–D303; citovaný odstavec na str. D303.

[74]Počátkům axiomatizace pravděpodobnosti v českých zemích je věnován autorčin článek [163] a stať [39] Š. Bilové, L. Mazliaka a P. Šišmy; o Pankrazově příspěvku píše rovněž F. Fabian (viz [91]).