

JAN HENDL A KOLEKTIV

ZÁKLADY MATEMATIKY, LOGIKY A STATISTIKY PRO SOCIOLOGII A OSTATNÍ SPOLEČENSKÉ VĚDY V PŘÍKLADECH



KAROLINUM

Jan Hendl a kolektiv

Základy matematiky, logiky a statistiky pro sociologii a ostatní společenské vědy

Recenzovali:

RNDr. Jindra Reissigová, Ph.D.

RNDr. Marek Malý, CSc.

Autoři:

Jan Hendl, Petr Hanzlík, Martin Moldan, Tereza Ranošová, Jakub Siegl,
Martin Štrobl a Jan Žáček



**Národní
plán
obnovy**



**Financováno
Evropskou unií**
NextGenerationEU

MSMT
MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Publikace byla vydána za podpory Ministerstva školství, mládeže
a tělovýchovy a Národního plánu obnovy v rámci projektu
Transformace VŠ na UK (reg.č. NPO_UK_MSMT-16602/2022).

Vydala Univerzita Karlova
Nakladatelství Karolinum
Praha 2022
Sazba Jan Hendl
Třetí, doplněné vydání

© Univerzita Karlova, 2022
© Jan Hendl a kolektiv, 2022

ISBN 978-80-246-5400-3
ISBN 978-80-246-5404-1 (online : pdf)



Univerzita Karlova
Nakladatelství Karolinum

www.karolinum.cz
ebooks@karolinum.cz

Obsah

I Matematika a logika

1	Úvod do matematiky	13
2	Lineární rovnice	32
3	Vektory	45
4	Matice	56
5	Logika	96
6	Množiny, relace	129
7	Booleova algebra	150
8	Millova pravidla, QCA	164
9	Analýza sociálních sítí	179
10	Kombinatorika	206
11	Pravděpodobnost	214
12	Markovovy řetězce	235
13	Racionální volba	250
14	Hry se ztrátou nebo ziskem	257
15	Derivace, integrál, modelování	270

II Statistika

1	Úvod do statistiky	301
2	Experimentální vs. observační design	309
3	Popisná statistika	318
4	Náhodné veličiny	335
5	Diskrétní pravděpodobnostní rozdělení	352
6	Normální rozdělení	359
7	Zákon velkých čísel, centrální limitní věta	369
8	Spojité rozdělení a rozdělení statistik	378
9	Teorie statistického usuzování	386
10	Statistická významnost, velikost účinku	398
11	Jednovýběrový z -test pro průměr	408

12	Jednovýběrový t -test pro průměr	413
13	Dvouvýběrový t -test (2 nezávislé výběry)	420
14	Párový t -test (2 závislé výběry)	430
15	Test dobré shody a kategoriální proměnná	436
16	Závislost mezi kategoriálními proměnnými	444
17	Jednovýběrový z -test pro relativní četnost	452
18	Dvouvýběrový z -test pro relativní četnosti	459
19	Neparametrické metody	466
20	Korelační koeficient	480
21	Lineární regresní model	489
22	Doporučení k využívání statistiky	506

III Přílohy

1	Model argumentace podle Toulmina	519
2	Základní vzorce statistiky	523
3	Matematické vzorce, grafy funkcí a tabulky	526
4	Řešení příkladů	539
5	Použitá a doporučená literatura	568

Předmluva

Porozumění základům matematiky a statistiky tvoří často předpoklad pro to, aby mohli studující navázat ve svých oborech na práci velkých předchůdců, aby se udrželi „na ramenou obrů“ a někdy i viděli dál než oni.

Znalostní požadavky v prvních semestrech studia sociálních věd, pokud jde o matematiku a statistiku, se liší od látky, jejíž zvládnutí se předpokládá při studiu přírodních a technických věd. Většinou jde o matematické pojmy a techniky, které lze využít při osvojování probíraných statistických metod (Beneš, Drulák 2020; Gill 2006; Moore, Siegel 2014; Novotná, Špaček, Šťovíčková Jantulová 2019).

Studující se matematice spíše vyhýbají. Je to způsobeno mimo jiné okolností, že střední úroveň vzdělávání je dostatečně nepřipravila, v horším případě se v nich utvrdil pocit jisté úzkosti z matematiky (Jones, Goldring 2014). Negativní pohled studujících na matematiku a vlastní matematické schopnosti dokládají empirické studie (např. Murton, Lehtinen 2003). Zdůvodnění pro tento stav lze přijmout, upřednostňujeme však jiný postoj. Perspektivu má poučený a tvořivý přístup k matematice, logice i statistice.

Matematika je potřeba při zvládání kvantitativních výzkumných metod. Především objasňování vícerozměrných statistických metod se neobejde bez znalostí základů vektorového a maticového počtu. Jak v kvalitativním, tak v kvantitativním výzkumu využíváme logické principy. S jejich pomocí jsou dokonce vyvinuty metody kvalitativní komparativní analýzy (QCA), která dnes přitahuje velkou pozornost. Nejjednodušší modely časových řad jsou dobrým příkladem mnoha aplikačních možností matematického modelování a využití principů vektorového a maticového počtu. Základy grafů a sociálních sítí je nutné pochopit, protože se pomocí nich řeší problémy v souvislosti s otázkami, které vyvstávají například při využívání internetu (Bonacich, Lu 2014).

Z těchto skutečností autorská skupina vycházela, když se rozhodovala, jaké partie z matematiky do textu zařadit a jak je podat. Také zohlednila, že se jedná o doplněk k přednáškám o základech matematiky a statistiky pro studující bakalářského studia sociologie a ostatních sociálních věd. Přednášející, případně jeho/její/jejich pomocníci a pomocnice, rozhodují, o čem se bude mluvit a co se bude procvičovat.

První část knihy je věnována matematickým základům. Nejprve jsou jednoduše představeny vybrané matematické koncepty, které jsou poté rozšířeny v dalších kapitolách o speciálnější témata, které matematicky zaměřené sociální vědy potřebují. Text se zabývá např. QCA metodou, aplikací jednoduchých náhodných procesů, modelů epidemií, analýzou sociálních sítí, teorií rozhodování a teorií her.

Druhá část knihy se zaměřuje na představení užitečných konceptů statistiky nezbytných pro elementární práci s daty a jejich použití pro vyvozování zobecnujících závěrů. Je probrána popisná statistika, náhodná proměnná, pravděpodobnostní rozdělení a pojmy pojící se s těmito tématy. Větší prostor je věnován porozumění a demonstraci testování statistických hypotéz. Text seznamuje nejprve s jednoduchými metodami (mezi které patří např. z -testy nebo t -testy) a poté s metodami složitějšími, jako jsou např. test dobré shody, analýza kontingenčních tabulek nebo neparametrické metody. Také se osvětlují principy a konstrukce intervalů spolehlivosti. Druhá část knihy je uzavřena úvodem do regresní analýzy.

Každá kapitola obsahuje základní a zjednodušeně pojatou teorii a několik řešených modelových příkladů typických pro dané téma. Modelové příklady jsou základním prvkem této publikace. Na konci každé kapitoly jsou navíc uvedeny další podklady k procvičení společně s výsledky. Text uzavíráme statistickými tabulkami a užitečnými vzorečky. Diagnostický test předpokládaných vstupních znalostí naleznou čtenáři v příkladech k procvičení v první kapitole.

Při aplikaci matematických a statistických metod hrají důležitou roli počítače a příslušný software. V tomto směru doporučujeme tabulkové programy *Excel* od Microsoftu, *Calc* z OpenOffice, komerční statistický systém *SPSS*, volné statistické systémy *PSPP*, *JASP* (oba tyto programy se snaží nahradit systém *SPSS* a internetový statistický program *SISA* a volný programovací jazyk *R* nebo volný matematický systém *Octave*. Také je cenná knížka Mareše, Rabušice a Soukupa *Analýza sociálně vědních dat [nejen] v SPSS* (vyd. Masarykova univerzita).

Tato kniha by nevznikla bez pomoci mnoha lidí. Dík patří všem, kteří přispěli k její přípravě: studentkám a studentům z FSV UK, zvláště T. Balouškové (podklady pro kapitoly 2, 3, 10 a 11 z matematiky), L. Rozumkové (podklady pro kapitoly 2, 3, 10 a 11 z matematiky) nebo K. Hanzlíkovi (podklady pro kapitoly 10, 11 a 12 ze statistiky), statistikovi a sociologovi P. Soukupovi z FSV UK, logikům M. Pelišovi a P. Arázimovi a oběma recenzentům J. Reissigové a M. Malému z AV ČR.

Koordinátor autorského kolektivu upřímně děkuje všem spoluautorům, zejména mgr. J. Žáčkovi, za jejich příspěvek k této knize. Také je vděčný za přátelskou asistenci PhDr. J. Koukalovi, Ing. K. Novotnému, doc. PhDr. Richardu Papíkovi a prof. MUDr. Š. Svačinovi.

Za chyby v textu přejímá zodpovědnost koordinátor projektu. Za případná pochybení se omlouváme. Budeme každému čtenáři vděční za jakékoliv připomínky a upozornění na nalezené nedostatky a možná vylepšení.

jan.hendl@ruk.cuni.cz

Značení:

$\mathbb{N}$... množina přirozených čísel

$\mathbb{Z}$... množina celých čísel

$\mathbb{Q}$... množina racionálních čísel

$\mathbb{R}$... množina reálných čísel

Část I

Matematika a logika

Kapitola 1

Úvod do matematiky

Tradičně byly teoretické vztahy v sociálních vědách formulovány především pomocí kvalitativních pojmů (nemáme na mysli ekonomii). Matematické metody se nepovažovaly za dostatečně vhodné pro rozvíjení teoretického porozumění. Matematika hraje velkou roli v přírodních vědách a biologii, od kterých se sociální vědy liší v mnoha ohledech. Z příslušných úvah proto plyne, že nemůže jít o vulgární přenos přírodovědeckého myšlení přímo do sociálních věd (Lazarsfeld¹ 1954). V posledních desetiletích se matematika a statistika staly důležitými prostředky modelování i v sociálních vědách. V literatuře se objevují speciální matematické metody a postupy. Těm by měli studující porozumět, aby mohli kriticky posoudit popisované skutečnosti, případně zvolený přístup přetvořit.

Příkladem rozšířené aplikace matematiky v sociálních vědách je síťová analýza. Ta se stala domovem výzkumníků a výzkumníků rozličných zájmů, které sahají od konstrukce osobní identity k epidemiologii, od etnografie po analýzu struktur v internetu.

¹Lazarsfeld, Paul, Felix (1901-1976), rakousko-americký sociolog. Věnoval se zejména statistickým metodám a výzkumu veřejného mínění.

Demokracie a matematika

„Pokud jsem viděl dále než ostatní, pak to bylo proto, že jsem stál na ramenou obrů“, řekl fyzik Isaac Newton (1643-1727) v souvislosti s hodnocením významu matematické analýzy pro jeho objevy. Podobně se vyjadřoval Albert Einstein (1879-1955), největší fyzik 20. století. Mezi Newtonem a Einsteinem byla časová mezera stovky let, ale pohled na matematiku měli podobný.

Uvedená myšlenka byla vyslovována mnohokrát autory reformy výuky matematiky na amerických základních a středních školách v 90. letech minulého století. Eseje v knížce „Na ramenou obrů“ (Steen 1990) poukazovaly k novým pohledům na didaktiku matematiky pro mládež přicházejícího století. „Na ramenou obrů“ pro pedagogy symbolizovalo snahu pomáhat studentům při hledání konceptů matematiky tak, aby docházelo k lepšímu pochopení světa a řešení jeho problémů. Tamějším pedagogům záleželo také na propojení matematiky a demokracie. Matematika hraje určitě významnou roli nejenom ve vědě, ale také při rozvíjení konceptu svobodného občana v demokracii.

Například politici a jejich příznivci se při volbách ohánějí čísly bez omezení a zodpovědnosti. Novináři informují o rozporných číslech s nedostatkem reflexe a odborného vhledu o jejich přesnosti a konzistenci. Pro fungování demokracie je tedy žádoucí jistá míra znalosti „matematiky voleb“ a statistiky, matematická gramotnost občana. Matematická gramotnost však není v tomto pojetí pouze soubor určitých matematických dovedností, nýbrž také zahrnuje uvažování o povaze a kvalitě interakce mezi jedincem s těmito dovednostmi a sociální situací nebo prostředím. To implikuje, že taková gramotnost nemá jednou provždy stejný význam, protože sociální situace a prostředí se neustále proměňují v čase.

Pro pedagogy jsou v této souvislosti aktuální myšlenky filosofa a pedagoga Johna Deweye¹. Ten byl velkým příznivcem moderní demo-

¹Dewey, John (1859-1952), americký filozof, pedagog, psycholog a reformátor vzdělávání. Je považován za jednoho ze zakladatelů filozofie pragmatismu.

kracie. Ve svém díle upomínal, že změna převažuje kontinuity. Proto je podle něho nutné, aby došlo k přeformulování gramotnosti, která musí držet krok s nikdy nekončícím procesem změn. Je to zapotřebí, protože také lidská přirozenost a společnost se neustále mění. Pro Deweyho se věc stává složitější, když se zeptáme, co gramotnost vlastně znamená pro demokratickou společnost a její ideály. Rozlišuje „inertní“ a „osvobozující“ gramotnost. První význam znamená úroveň verbálních a numerických dovedností, aby bylo možné porozumět instrukcím, provádět rutinní procedury a vykonávat úkoly standardním způsobem. Ze sociální perspektivy jde o gramotnost, kde převládají tradice a zvyky, kde je vše na svém místě a inovace jsou omezovány. „Osvobozující“ gramotnost předpokládá standardy, které podporují rozvoj dovedností při hledání informací a sílu kritiky, reflexi a rozhodnutí ke změně. Dewey má na mysli „osvícení mas“, které vede ke skutečně „vitální“ demokracii.

Matematika se mění. Mění se také možnosti aplikací (nástupem počítačů). Matematické vědy dnes zahrnují kromě čisté matematiky také statistiku, finanční matematiku, části počítačové informatiky, operačního výzkumu a bioinformatiky (například aplikace statistiky v genetice). Tyto oblasti matematiky třebaže mají mnoho společného se základy matematiky, mají zvláštní charakter, metodologii, standardy a cíle.

Z aplikované matematiky je nutné vyzdvihnout zvláště statistiku, nebo spíše vědu o datech. Bohužel běžná školní matematika jí věnuje málo pozornosti. Na školách se nevytváří most mezi školní matematikou a interpretacemi dat, mezi abstraktními výpočty a světem statistiky v medicíně, zemědělství, psychologii, ekonomii, sociologii a politice.

V demokracii vede numerická a matematická negramotnost zaručeně ke špatným rozhodnutím. Numerická a matematická gramotnost je však pouze jedním z předpokladů změny. Numerická a matematická gramotnost a sociální vědy mohou pomoci občanům dospět k rozhodnutím, která jsou zakotvena v aktuálních datech, matematických modelech, logice a ne v setrvačnosti a v nepřehledném uvažování.