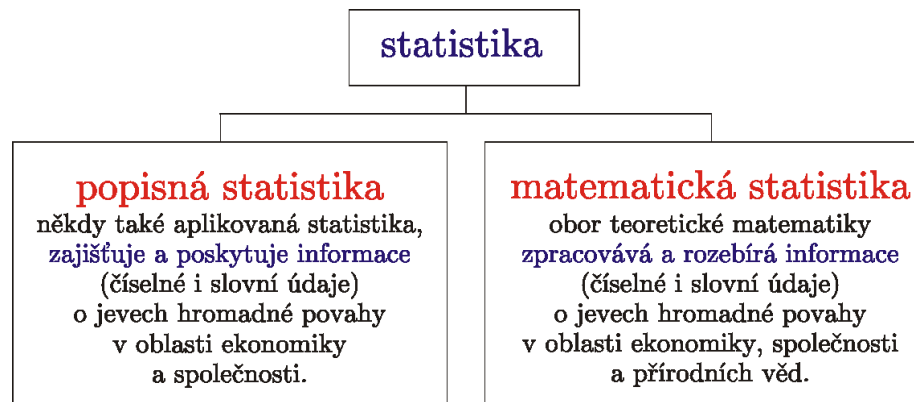


Statistika

Statistika je vědecká disciplína zabývající se studiem dat popisujících existující variabilitu (zajímají nás rozdíly i podobnosti hromadných jevů) a hodnocením hypotéz tato data vysvětlujících.

Slovo *statistika* pochází z latinského „*status*“, což znamená stav. Původně se jednalo pouze o stav nějaké země či státu a statistikou se rozuměla činnost spočívající ve zjišťování tohoto stavu. Později se pole působnosti statistiky značně rozšířilo, statistika navíc přestala být pouze praktickou činností a stala se vysoce propracovanou vědeckou naukou. Dnes tato nauka zahrnuje velmi širokou škálu kvantitativních metod umožňujících zjišťovat „stav“ věcí a poměrů v rozličných strukturách. Kromě přírodních, společenských a hospodářských poměrů v daném státě lze zjišťovat např. hospodářské poměry v nějaké firmě, stav zásob v obchodním domě, stav vody na českých tocích, stav lesů v České republice apod.

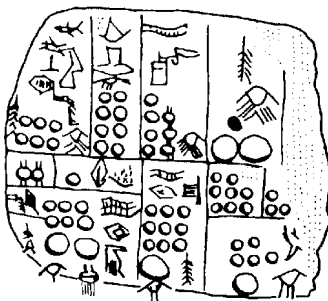
1. Poznámka Dnes se pod pojmem statistika skrývají dva vědecké obory, a to **popisná statistika** a **matematická statistika**. Viz Obrázek 1.



Obrázek 1: Rozdělení statistiky

Střípky z dějin popisné statistiky

Nejstarší písemné památky pocházející z oblasti Sumeru mají statistickou povahu – jednalo se o záznamy o časových intervalech, počtech osob a kusů domácího zvířectva a úrodě. Na Obrázku 2 vidíte záznamy, které jsou psány protosumerským písmem, které je vlastně přechodem mezi pouhými pikto-gramy a skutečným znakovým písmem.



Obrázek 2: Záznam psaný protosumerským písmem

Starověké říše, mezopotámskými městskými státy počínajíce, byly finančně zcela závislé na úspěšném výběru daní, ať již se jednalo o pracovní povinnost obyvatelstva, dávky naturální či peněžní. Pro tyto účely existovaly již propracované „statistické metodiky“ a z četných písemných památek jsme schopni si udělat představu o systému tehdejší státní správy a nakládání s informacemi. Například v Egyptě je od roku 2850 př.n.l. pravidelně jednou za dva roky prováděn soupis dobytka, od roku 2000 př.n.l. se vybírá

rovná daň z hlavy, jejíž zavedení si vyžádalo aktualizované sčítání obyvatelstva, také je zdaněna půda (výše daně vychází z přesně zaměřené rozlohy a vydatnosti záplav na tom kterém místě).

Zvlášť propracovaný systém evidence, census, měl starověký Řím. Census byl pravidelně prováděn v republikánském období; jednalo se původně o soupis nemovitého majetku, později i otroků a dobytka, na jehož základě byla vypočtena výše daně pro jednotlivé občany. Snad nejslavnějším se dnes jeví jeden z posledních pravidelných censů, provedený za císaře Augusta kolem přelomu letopočtu, který vešel do dějin jako časové určení narození Ježíše Krista.

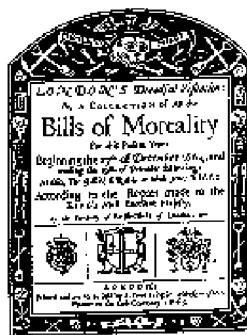
Období raného středověku přineslo do střední a západní Evropy všeobecný rozvrat, negramotnost byla běžným jevem i mezi panovníky. Určitá ohniska vzdělanosti udržovala církev, jež také byla schopna vést evidenci svého majetku a jeho změn a členové církevních řádů byli zaměstnáváni aristokracií jako „statistici“. Nejstarší zachovalou písemností z tohoto období na našem území je soupis majetku litoměřického kostela z roku 1058, který je součástí zakládací listiny vydané knížetem Spythněvem II.

Středověkými „statistickými ročenkami“ byly vrchnostenské urbáře, z nichž kromě evidence příjmů poddaných bylo možno vyčíst i rozsah pozemkového vlastnictví šlechty a církve, a městské berní knihy a berní rejstříky, obsahující výčet poplatníků spolu s jejich nemovitým majetkem a výši odvedených daní. Ve 14. století se v Evropě objevují první církevní matriky, cenný zdroj informací o přirozené změně obyvatelstva.

Šetření však nebyla vždy prováděna jen kvůli eráru, důvody pro statistické zjišťování byly mnohdy i zcela jiného – humánnějšího – rázu. Například za vlády císaře Rudolfa II. v roce 1583 vypukla v českých zemích epidemie moru. V jejím důsledku bylo zahájeno šetření o „zdraví populace“, které mělo zmapovat vznik a rozvoj zhoubných epidemií a umožnit přijímání včasných protiopatření.

16. století s sebou přineslo nový pohled na svět a člověka, rozvoj filozofie a počátky moderních vědních disciplín, je to období velkých společenských změn, kdy se společnost více diferencuje a z potřeby jejího efektivního řízení pramení i potřeba jejího zkoumání. V duchu popisné statistiky vyšlo v Benátkách roku 1562 jedno z prvních státovědných děl, *O vládě a správě v různých královstvích* od Francesca Sansoviny. V roce 1589 použil Ital Girolamo Ghilini jako první termín *statistika*, v původním smyslu se jednalo o stav státu. Roku 1662 vyšla kniha *Německý knížecí stát* autora Veita Ludwiga von Seckendorff, která se během 17. a 18. století dočkala řady vydání.

V Anglii v 17. století vznikl jiný okruh statistiky, takzvaná politická aritmetika (dnes mluvíme o demografii), která vycházela z údajů o narozeních a úmrtích a pokoušela se na jejich základě zkoumat vývoj stavu obyvatelstva v delších časových údobích, jednalo se vlastně o první výzkumy časových řad. Zakladatelem této disciplíny byl John Graunt (1620-1674), který jako první považoval demografické jevy za jevy hromadné, viz Obrázek 3. Odhalil poměr mezi počtem mužů a žen v populaci a stabilní poměr mezi počtem narozených chlapců a dívek a sestavil úmrtnostní tabulky na základě zkoumání vymírání jednotlivých věkových skupin. Dalším reprezentantem anglické školy byl Edmund Halley (1656-1742), který na konci 17. století při konstrukci úmrtnostních tabulek propojil statistiku s počtem pravděpodobnosti a použil geometrický náhled na pravděpodobnostní úlohu.



Obrázek 3: Ukázka z Anglie ze 17. století

Významnou osobností evropské statistiky byl belgický matematik, statistik a astronom Adolphe Lambert Quételet (1796-1874), který vypracoval zásady moderních sčítání lidu – poprvé byly uplatněny při belgickém sčítání lidu v roce 1846. Na základě velkého souboru dat vypočítal rozměry „průměrného člověka“ a odchylky jednotlivců od tohoto průměru – je tedy duchovním otcem pojmů jako průměr, střední

hodnota, rozptyl a rozdělení.

Statistika 19. a počátku 20. století, to je především vytváření rozsáhlých souborů dat, sběr mnoha informací od co nejširšího okruhu respondentů, se zjevným cílem: obsáhnout ve svém šetření celou populaci a tím získat maximálně přesný obraz stavu společnosti.

Časová i finanční náročnost takových šetření vedla k úvahám, zda je opravdu třeba zkoumat celou populaci, nebo postačí-li vybrat pouze její reprezentativní vzorek. Na základě této myšlenky se počátkem 20. století zrodila matematická statistika, disciplína, jejímž charakteristickým rysem je hledání metod, jež by umožnily vytvoření závěrů o celku na základě výběru. Matematická statistika jako samostatný obor vyšší matematiky si tak během prvních třiceti let 20. století vytvořila vlastní aparát s postupy jako je analýza rozptylu, korelační počet a ověřování hypotéz.

Zrod počtu pravděpodobnosti

Odhlédneme-li od prací indických a čínských matematiků, kteří se řešením kombinatorických úloh zabývali již ve starověkém období, můžeme položit základy této matematické disciplíny v Evropě do 16. století. Vysvětlení tak pozdního zájmu matematiků o náhodu může být následující.

Nebyla rozpoznána žádná souvislost mezi matematikou a náhodnými jevy: matematika je disciplína, v jejíž povaze se nic náhodného nevyskytuje, a náhoda je nahlížena buď jako projev vůle božstva, nebo je chápána jako neznalost všech příčin a kauzálních vazeb – tedy jev, který by mohl být zevrubným zkoumáním zcela vyloučen. Dalším důvodem mohl být prostý fakt, že matematické zkoumání náhodných jevů prostě dříve nikdo nepotřeboval (k jistě velkému potěšení falešných hráčů) a nutnost jejich přesného popisu se objevila až v souvislosti s rozvojem obchodu, demografie, astronomie a moderní fyziky.

Za „základní materiál“ teorie pravděpodobnosti lze považovat dva typy problémů z oblasti hazardních her a sázek:

První typ problémů byl v podstatě kombinatorický, jednalo se o otázky kolika způsoby může padnout jistý počet ok při házení určitým počtem hracích kostek – podobné úlohy se objevují v teorii pravděpodobnosti a jejich aplikacích i dnes.

Druhý typ problémů se týkal tzv. úlohy o rozdělení sázky, kterou lze stručně popsat následovně. Dva hráči hrají sérii her o určitou částku s tím, že tuto sumu získá ten, který jako první dosáhne předem domluveného počtu vítězství. Série je však předčasně ukončena a hráči si chtějí částku spravedlivě rozdělit v závislosti na tom, kolik her který z nich vyhrál.

Prvními (dnes) známými, ale neúspěšnými řešiteli úlohy o rozdělení sázky byli Luca Pacioli (asi 1445–1514) a Nicolo Tartaglia (1499–1557). Nejstarší práci věnovanou speciálně problémům spadajícím dnes pod teorii pravděpodobnosti je spis Hieronyma Cardana (1501–1576) *De ludo aleæ* datovaný rokem 1526, teorii pravděpodobnosti se zabýval také Galileo Galilei (1564–1642).

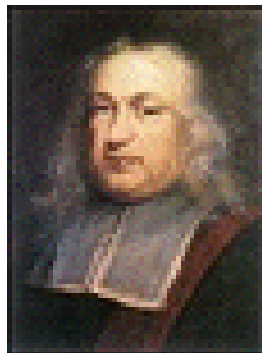


Obrázek 4: Blaise Pascal

Za opravdový počátek teorie pravděpodobnosti je pak považována korespondence, kterou v roce 1654 vedli Blaise Pascal (1623–1662) a Pierre de Fermat (1601–1665) o problémech, se kterými se na Pascala obrátil rytíř de Méré. Jednalo se o již zmíněnou úlohu o rozdělení sázky a úlohu o kostkách, tedy otázku, kolik hodů jednou či dvěma kostkami je třeba, aby šance, že padne aspoň jednou šestka, respektive dvě šestky, byla nadpoloviční. Tento elementární problém vyřešil Pascal okamžitě, podstatné však je, že našel i správné řešení úlohy o rozdělení sázky.

Blaise Pascal byl všestranně nadaný a svůj poměrně krátký život rozhodně nepromarnil – v 16ti letech publikoval práci o kuželosečkách, ve dvaceti se stal tvůrcem Pascaliny – počítacího stroje, v letech 1648–1653 se zabýval zkoumáním atmosférického tlaku, o rok později již pracoval s pravděpodobností a svými pracemi z let 1658–59 předjímal infinitezimální metody Newtona a Leibnitze.

Pierre de Fermat vystudoval práva, byl vynikajícím znalcem klasických i živých jazyků a matematice se věnoval jen ze záliby. O jeho práci parlamentního rady v Toulouse již mnoho nevíme, svému koníčku se však věnoval natolik poctivě, že je dnes považován za zakladatele teorie čísel a spoluzakladatele teorie pravděpodobnosti.



Obrázek 5: Pierre de Fermat

Dalším z mužů, kteří přispěli ke zrodu nové matematické disciplíny, byl Christiaan Huygens (1629–1695). Ač vzděláním (opět!) právník, zabýval se přírodními vědami: fyzikou, astronomií a také matematikou. Jeho přínosem k teorii pravděpodobnosti je spis *De ratiociniis in ludo aleæ* z roku 1656. V této době žil v Paříži a znal téma korespondence Pascala a Fermata. Huygens se v tomto spisu stal duchovním otcem střední hodnoty (i když tento pojem tak dosud nenazýval) a klasické definice pravděpodobnosti, přestože nemluvil o „pravděpodobnosti“, ale o „očekávané výhře“. Pascalovy a Fermatovy úvahy ve své práci posunul od řešení konkrétních úloh již k poloze obecných pojmů a postupů.



Obrázek 6: Christian Huguens

Počet pravděpodobnosti se zrodil ve Francii Ludvíka XIV. díky mužům, kteří matematiku nepovažovali za své životní poslání, ale do jejích hlubin pronikali s neobyčejným zaujetím a poznatky, ke kterým dospěli, si nenechávali pro sebe. A tak vděčíme šťastné náhodě, která svedla rytíře de Méré s Blaisem Pascalem, a Pascalovu a Fermatovu pochopení pro problémy vášnivého hráče, za zrod nové matematické disciplíny.