

vzájomné pôsobenie subjektu a objektu? Veď zmyslová skúsenosť sa nedá oddeliť od vzájomného pôsobenia človeka a poznávaného predmetu. Slabiny v teórii „primárnych“ a „sekundárnych“ kvalít odhalil subjektívny idealizmus 18. storočia, ktorý poukázal (celkom oprávnené) na nemožnosť presne rozhraničiť subjektívne a objektívne kvality z hľadiska metafyzického empirizmu.

V antike teda takisto ako aj vo filozofii 17.—18. storočia výskum problémov vedomia a poznávania bezprostredne súvisí s filozofickou analýzou povahy reality, s odhalením prvotných základov skutočností, pričom v antickej filozofii sa teória poznania neoddeľuje jasne od ontologicko-kozmozologickej sústavy, zatiaľ čo v európskej filozofii 17.—18. storočia nadobúda teória poznania relatívnu samostatnosť. Avšak i v tom čase sa poznanie chápe ako nerozlučne späté s bytím. Podľa chápania charakteru reality prejavuje sa teória poznania alebo v súvislosti s ontologickou sústavou (kde sa realita chápe ako objektívne, od individuálneho vedomia nezávisle jestvujúce bytie — idealistický racionalizmus, metafyzický materializmus) alebo v súvislosti so systémom psychologickéj metafyziky (realita sa stotožňuje s empirickými „dátami“, so zmyslovými dojmami — Berkleyho a Humov subjektívny idealizmus). U Descarta je teoreticko-poznávací problém vzťahu poznania reality neoddeliteľný od ontologického problému vzťahu ideálnej a materiálnej substancie. Metafyzickí materialisti skúmajú problémy teórie poznania chápujú človeka ako biologické individuum, úplne závislé od prírody a ňou obdarené všetkými nevyhnutnými schopnosťami pre poznanie. V Humovej filozofii sa gnozeologické otázky redukujú na analýzu psychologických problémov vzájomného pôsobenia zmyslových vnemov, predstáv, pamäti, na výskum asociácií, zvykov atď.

(Pokračovanie)

METODOLÓGIA A ŠTATISTIKA

Cieľom tohto prehľadu je informovať o základných metodologických problémoch štatistického spracovania empirických dát.

1. Významným momentom v dejinách empirických vied je zavedenie meracích metód a kvantitatívnych pojmov. Týmto sa empirické vedy dostali od bežných pozorovaní a povrchných (často chybných) zovšeobecnení a analógií na vyššiu úroveň. Meranie a kvantitatívne pojmy umožnili podrobnejší opis, a tým aj hlbšie, rozsiahlejšie poznanie. Pripravili pôdu pre odhalenie nových zákonitostí. Stali sa dôležitou súčasťou experimentálnej metódy.

Dôsledné využitie kvantitatívnych pojmov a merania v empirických vedách musí byť sprevádzané aplikovaním matematiky. V tejto oblasti dosahovali úspechy spočiatku len fyzikálno-chemické vedy. Biologické a lekárske vedy stáli pred oveľa zložitejšími problémami, hlavne kvôli ohromnej zložitosti, komplexnosti, rozsiahlosti, premenlivosti, spletnosti a bohatstvu vzájomných vzťahov predmetov svojho skúmania.

Rozvoj štatistických metód otvoril cestu aplikáciám matematiky aj v biologických a lekárske vedách. Štatistické metódy vytvárajú predpoklady pre netriviálne používanie meraní a kvantitatívnych pojmov v biologických a lekárske vedách. Výborná uplatniteľnosť štatistiky pri vyhodnocovaní výskumných výsledkov v biologických, lekárske i spoločenských vedách sa zakladá na spomínaných zložitostiach predmetov týchto vied. Prirodzene, že rozsah oblasti aplikovateľnosti štatistických metód je širší.

2. Ak skúmame nejaký problém na živých organizmoch, máme veľmi ťažkú situáciu pri odhalovaní a izolovaní činiteľov, ktoré ovplyvňujú skúmaný problém. Napríklad zisťujeme účinnosť nejakej novej terapie. A tým, že ju porovnáваме s účinnosťou staršej terapie B. Predpokladajme, že ich účinnosť skúsime na dvoch pacientoch, každú na jednom z nich. Z toho, že A nebola úspešná a B bola úspešná, nemožno uzatvárať, že B je účinnejšia ako A.

Výsledok spomínaného pokusu mohol byť ovplyvnený mnohými inými činiteľmi. Napríklad stavom choroby pacienta, jeho vekom, pohlavím, určitými vlastnosťami jeho organizmu (aj takými vlastnosťami, od ktorých závisí reagovanie na skúšanú terapiu), psychickými dispozíciami atď. Tieto okolnosti mohli u skúmaných pacientov podstatným spôsobom a s odlišným výsledkom ovplyvniť náš pokus. Nepomohlo by nám ani to, keby sme účinnosť našich terapií — ak by to nevyučovala povaha choroby — skúšali na jedinom pacientovi v dvoch rôznych časových obdobiach. Stav, reagovanie a dispozície organizmu sú totiž v rôznych časových obdobiach rôzne, a práve ony môžu podstatne ovplyvniť výsledok pokusu.

V tom prípade, keď nevieme jednoznačne určiť, aké činitele a akým spôsobom ovplyvňujú výsledok pokusu, hovoríme o pôsobení náhodných činiteľov na pokus. To, čo sa nazýva biologickou variabilitou, môžeme považovať za jednu z foriem pôsobenia náhodných činiteľov na reálne deje. O náhodných činiteľoch vieme len toľko, že nejakým spôsobom ovplyvňujú výsledok pokusu a že ich pôsobenie sa vyrovnáva pri veľkom počte pokusov. Keď napríklad účinnosť našej terapie skúmame na väčšom počte pacientov, pravdepodobnosť toho, že výsledok pokusu je ovplyvnený náhodnými činiteľmi, sa znižuje. Inými slovami: väčší počet skúmaných prípadov umožňuje odhaliť v pôsobení náhodných činiteľov istú zákonitosť. Keď sme empirickým výskumom zistili výsledky pôsobenia náhodných činiteľov v nejakej skupine javov, je vhodné aplikovať štatistické metódy. Práve štatistika pomáha nachádzať zákonitosť vo výsledkoch pôsobenia náhodných činiteľov. Týmto úvahám však treba dať konkrétnejšiu podobu.

3. Výskumné výsledky (t. j. výsledky pozorovaní, meraní, experimentov) možno racionálne vyhodnocovať štatistickými metódami iba za určitých podmienok. Chaotický súbor výsledkov pozorovaní, meraní a experimentov, ktoré neprebíhali podľa dobre premysleného plánu, sa dá ťažko štatisticky spracovať. Prvoradým predpokladom uplatnenia štatistických metód je jasné formulovanie problému, ktorý chceme pozorovaním, meraním alebo experimentom vyriešiť. Teda základnou metodologickou zásadou štatistického vyhodnotenia výskumných výsledkov je to, že treba jasne sformulovať problém a na základe formulácie problému treba výskum premyslieť, rozplánovať, t. j. orientovať sa na pozorovania, merania a experimenty relevantné pre vyriešenie daného problému.

Najjednoduchším typom problémov sú problémy, ktorých vyriešením má byť nejaký opis. Je zrejmé, že opis jedného javu (alebo veľmi malého počtu javov), ktorý je výsledkom pôsobenia náhodných činiteľov, nemá veľkú hodnotu. Pre empirickú vedu je poučnejší až opis veľkého počtu takýchto javov (často sa hovorí o opise tzv. hromadného javu). Elementárnym problémom v tejto skupine problémov je zistiť, v kolkých prípadoch nastane v n pokusoch (slovo pokus sa tu chápe vo veľmi všeobecnom význame) niektorý z dvoch možných výsledkov. Napríklad: koľko chlapcov a dievčat prípadne na 1000 novonarodeniatok a podobne. Tieto opisy nazývame kvalitatívnymi opismi. Ak výsledok nejakého pokusu možno vyjadriť kvantitatívne, potom príslušný opis nazývame kvantitatívnym opisom.

Kvalitatívne i kvantitatívne opisy možno koncentrovane vyjadriť pomocou tzv. rozdelení pravdepodobnosti náhodnej premennej. Náhodnou premennou sa nazýva skú-

maný jav alebo pokus, ktorý vedie k rôznym možným výsledkom, závislým od náhodných činiteľov. Napríklad náhodnou premennou je pohlavie novonarodeniatka, výška maturanta v ČSSR v roku 1971 atď. Rozdelenie pravdepodobnosti poznáme vtedy, keď vieme, s akou pravdepodobnosťou nadobúda príslušná náhodná premenná každú svoju možnú hodnotu. Ak náhodná premenná môže nadobudnúť kontinuum hodnôt (sú to tzv. spojité náhodné premenné), musíme vedieť určiť, s akou pravdepodobnosťou padne hodnota náhodnej premennej do určitého intervalu.

Z uvedeného je zrejmé, že keď je dané rozdelenie pravdepodobnosti nejakej náhodnej premennej, je tým vytvorený aj nejaký opis: v toľkých a toľkých prípadoch zo všetkých pozorovaných (alebo keď chceme idealizovať: zo všetkých pozorovateľných) nadobúda náhodná premenná takú a takú hodnotu. Rozdelenie pravdepodobnosti nám to umožňuje povedať o každom pozorovanom prípade i o každej hodnote (alebo aspoň o každom intervale hodnôt).

Zložitejšie problémy sú tie, pri ktorých ide o zistenie, či určité vlastnosti, javy alebo veličiny sú alebo nie sú závislé od iných vlastností, javov alebo veličín. V takýchto prípadoch sú opisy iba východiskovým materiálom pre odhalenie hľadaných závislostí.

Iné štatistické metódy sú vhodné pre opisy a iné pre určenie závislosti či nezávislosti nejakých vlastností, javov, či veličín.

Aby sme sa vrátili k formulovaniu výskumného problému, ktorý chceme riešiť štatistickými metódami: Keď sa pýtame na nejaký opis, musíme predovšetkým vedieť, čo je skúmaná náhodná premenná a musíme stanoviť, aké hodnoty tejto premennej nás budú zaujímať. Až potom môžeme zisťovať, v koľkých prípadoch nadobúda skúmaná náhodná premenná aké hodnoty. Keď sa pýtame na nejakú závislosť, musí nám byť jasné, závislosť ktorých náhodných premenných budeme skúmať.

4. Opisy. Hromadné javy sú ovplyvňované náhodnými činiteľmi, preto nie je možné predpovedať, akým spôsobom budú determinovať náhodné činitele priebeh každého konkrétneho prípadu skúmaného javu. Štatistika sa aplikuje práve pri opisoch hromadných javov. Používa pri tom nasledovnú terminológiu:

Daný je základný súbor (populácia) predmetov, ktoré sa opisujú. Na základnom súbore sa zvolí náhodná premenná, ktorá sa podrobí výskumu. Určia sa možné hodnoty tejto premennej. Zo základného súboru sa zvolí náhodný výber a zisťuje sa, ktoré predmety zo zvoleného výberu nadobúdajú tie-ktoré hodnoty náhodnej premennej. Tieto zistenia získavame tzv. náhodnými pokusmi alebo náhodnými pozorovaniami. Sú to činnosti, ktorých výsledkom je registrovanie určitej hodnoty náhodnej premennej. Učebnicovým príkladom náhodného pokusu je hádzanie hracou kockou. Zvolenou náhodnou premennou môže byť počet šesťiek v sérii desiatich hodov kockou. Vidno, že náhodný pokus a náhodná premenná sú navzájom súvisiace, ale nie totožné pojmy. Pomocou vhodných štatistických metód sa z výsledkov získaných výskumom nejakého výberu (alebo výberov) odvodí opis základného súboru. Základný súbor sa opisuje štatistickými pojmami. Štatistický je každý taký pojem, ktorý sa vzťahuje na celý súbor predmetov, ale nemá zmysel pri opise izolovaného predmetu z tohto súboru.

Najjednoduchšie, čo nás pri opise môže zaujímať, je rozdelenie predmetov základného súboru na tie, čo majú nejaký hľadaný znak a na tie, čo ho nemajú. Štatistický opis hromadného javu v tomto prípade spočíva v tom, že sa vyjadrí, koľko percent predmetov základného súboru má (a koľko percent nemá) hľadaný znak. Skúmaná náhodná premenná zrejme nadobúda dve hodnoty. Takouto náhodnou premennou je už spomínané pohlavie novonarodeniatok (novonarodeniatko je alebo nie je chlapec). Pozorovanie, ktorým zisťujeme pohlavie novonarodeniatok nejakého výberu zo základného

súbore je náhodným pozorovaním preto, lebo vopred nemožno predpovedať s istotou výsledok jednotlivého pozorovania. Iba na základe štatistického opisu môžeme s určitou pravdepodobnostne vyčísľiteľnou spoľahlivosťou predpovedať o väčšom súbore novonarodeniatok, koľko v ňom bude chlapcov a koľko dievčat. Štatistické metódy umožňujú navyše vypočítať príпустné odchýlky od predpovedaných hodnôt.

Ak nás na základnom súbore zaujíma viac kvalitatívnych znakov, štatistický opis tohto hromadného javu spočíva opäť v tom, že sa vyčíslia, koľko percent zo základného súboru má ktorý znak (prípadne sa to vyjadrí v absolútnych číslach). V oboch uvedených prípadoch nie je dôležité, ktoré individuálne predmety zo základného súboru (alebo z výberu) majú ktorý znak, dôležitý je celkový pohľad na rozdelenie jednotlivých znakov medzi predmety základného súboru. Iba takýto celkový pohľad umožňuje abstrahovať od pôsobenia náhodných činiteľov na konkrétne prípady a ukazuje určité zákonitosti, platné pre súbor ako celok. Miesto terminológie používajúcej percenta možno používať terminológiu počtu pravdepodobnosti. Vtedy tvrdíme, že predmety zo základného súboru budú mať s pravdepodobnosťou p kvalitatívny znak A . Táto terminológia je výhodná najmä vtedy, keď závery o základnom súbore vynášame na základe skúmania určitého výberu (alebo výberov) zo základného súboru.

Každý kvalitatívny znak je hodnotou skúmanej náhodnej premennej. Prakticky nás zaujímajú len také náhodné premenné, ktoré majú konečný počet kvalitatívnych hodnôt. Predpokladá sa pritom, že náhodné pokusy alebo náhodné pozorovania možno pripraviť tak, aby výsledkom ich aplikovania na nejaký predmet z výberu bola ktoráási z hodnôt náhodnej premennej.

Od štatistických opisov kvalitatívnych znakov rozdelených medzi predmetmi nejakého základného súboru sa podstatne nelíšia opisy diskkrétnej číselnej náhodnej premennej (diskkrétnej náhodnej veličiny). Rozdelenie pravdepodobnosti náhodnej premennej, ktorej hodnotami sú kvalitatívne znaky, je špeciálnym prípadom rozdelenia pravdepodobnosti diskkrétnej náhodnej veličiny, pretože každý kvalitatívny znak možno reprezentovať určitým celým číslom.

V procese vedeckej idealizácie sa však vytvárajú spojité číselné náhodné premenné, t. j. náhodné premenné, ktorých hodnotou môže byť ktorékoľvek reálne číslo z určitého intervalu. Pri skúmaní výberu nikdy nenameriame ako hodnotu reálne číslo, vždy len racionálne číslo. Napriek tomu je výhodnejšie z výsledkov získaných o výbere odvodiť závery o základnom súbore tak, aby sme sa nevyhýbali iracionálnym číslam. Teória pravdepodobnosti pomáha dochádzať k veľmi plodným záverom, keď sa počíta so všetkými reálnymi číslami. Keď berieme do úvahy spojité náhodné veličiny, môžeme opísať mnohé dôležité typy rozdelenia pravdepodobnosti, z ktorých najznámejší je tzv. normálny rozklad. Pri normálnom rozklade je najpravdepodobnejšie, že predmety zo základného súboru nadobudnú priemernú číselnú hodnotu zo všetkých možných hodnôt spojitaj náhodnej veličiny a najmenej pravdepodobné, že budú mať niektorú extrémnu hodnotu. Niekedy je však postačujúce brať do úvahy iba diskkrétne premenné. Všetko závisí od typu výskumnej úlohy a od cieľov, ktoré sa sledujú.

Na základe skúmania výberu môžeme odvodiť opis základného súboru skoncentrovaný v nejakom rozdelení pravdepodobnosti. V prípade spojitaj premennej sa to všetko dá skoncentrovať v jedinej matematickej funkcii. Tento opis je samozrejme výsledkom určitej idealizácie a výsledky získané z výberov ho môžu potvrdzovať iba aproximatívne. Problém idealizácie i problém overovania štatistických hypotéz budú predmetom neskorších úvah.

Základný súbor možno približne opísať aj bez toho, že by sme skonštruovali úplné rozdelenie pravdepodobnosti. Určitý obraz o rozdelení pravdepodobnosti získame aj

tak, že určíme strednú hodnotu a rozptyl. Rozptyl určuje, ako sa kvantitatívne znaky predmetov základného súboru líšia od strednej hodnoty. Existuje viac druhov stredných hodnôt i rozptylov. Používajú sa v závislosti od typu úlohy. Jednou zo stredných hodnôt je i aritmetický priemer, častejšie sa používa vážený aritmetický priemer, modálna hodnota (hodnota, ktorú nadobúdajú prvky základného súboru s najväčšou pravdepodobnosťou), medián (hodnota, ktorú náhodná premenná nadobúda s pravdepodobnosťou $1/2$). Rozptyl sa najčastejšie určuje od váženého aritmetického priemeru. Najbežnejším rozptylom je tzv. variácia. Je to súčet štvorcov rozdielov hodnôt náhodnej premennej a strednej hodnoty násobených pravdepodobnosťou toho, že náhodná premenná nadobudne danú hodnotu. Odhliadnuc od matematických detailov možno konštatovať, že stredná hodnota a rozptyl dovoľujú pomocou dvoch veličín dostatočne prehľadným spôsobom opísať rozdelenie pravdepodobnosti náhodnej premennej v nejakom základnom súbore.

5. Keď už máme stručne naznačené, v čom spočíva štatistický opis určitého hromadného javu, môžeme sa začať zaoberať metodologickými problémami opisu.

Vo väčšine prípadov, a najmä vo výskumne dôležitých prípadoch, nie je možné opis založiť na výskume celého základného súboru. Preskúma sa určitý výber zo základného súboru a na základe výsledkov získaných pre výber sa opisuje celý základný súbor.

S týmto postupom sú späté dva vážne problémy:

a) ako zvoliť výber, aby výsledky výskumu boli reprezentatívne pre celý základný súbor,

b) v akej miere platia výsledky zistené pre výber (i reprezentatívny) pre celý základný súbor?

Prvý problém sa rieši rôznymi metodikami tzv. náhodných výberov, ktoré by mali zaistiť, že výber zo základného súboru nie je motivovaný subjektívnymi sklonmi alebo záľubami výskumníka, alebo že tento výber neobsahuje neúmerne veľa predmetov determinovaných iba určitými činiteľmi z celého komplexu náhodných činiteľov. Detailný rozbor týchto metodík nie je v tejto súvislosti na mieste.

Matematická štatistika má vypracovaný register metód aj na vyriešenie druhého problému. Tieto metódy patria do jadra matematickej štatistiky. Práve preto mnohí autori považujú štatistiku za metodologickú disciplínu. Jej úlohou je hľadať a zdôvodňovať (dokazovať) určité metódy overovania štatistických hypotéz skúsenosťou (opis nejakého základného súboru možno považovať za štatistickú hypotézu) a metódy usudzovania z výberu na základný súbor.

Prvá veľká skupina metód, umožňujúcich odpovedať na druhú otázku, je skupina metód odhadu štatistických parametrov. Metodami odhadu štatistických parametrov sa odvodzujú na základe výsledku zisteného o výbere závery o základnom súbore. Pritom je možné vypočítať aj spoľahlivosť tohto odhadu. Štatistickými parametrami sa nazývajú ľubovoľné štatistické charakteristiky základného súboru, ako sú napríklad rôzne druhy strednej hodnoty, rozptylu, rozdelenie pravdepodobnosti náhodnej premennej a podobne. Metódy odhadu štatistických parametrov dovoľujú na základe charakteristík vypočítaných z výberu (resp. výberov) odhadnúť príslušný parameter základného súboru. Pritom existujú postupy umožňujúce vypočítať, v akom okolí odhadovanej hodnoty parametra leží skutočná hodnota tohto parametra a s akou pravdepodobnosťou tam leží. Keď je daná pravdepodobnosť, s ktorou chceme poznať interval, do ktorého patrí skutočná hodnota hľadaného parametra, vieme vypočítať, aké okolie odhadnutej hodnoty hľadaného parametra bude tvoriť tento interval. Keď je daná veľkosť intervalu, do

ktorého by mala patriť skutočná hodnota hľadaného parametra, vieme vypočítať pravdepodobnosť, s akou tam táto hodnota bude patriť.

Druhú významnú skupinu metód, vhodných aj pre riešenie nášho problému, týkajúceho sa toho, do akej miery platia výsledky zistené na základe skúmania výberu pre celý základný súbor, tvoria metódy testovania štatistických hypotéz. Ak sa prijme hypotéza, že odhadnutá hodnota parametra je jeho skutočnou hodnotou alebo že skutočná hodnota je z určitého okolia odhadnutej hodnoty (takéto hypotézy sa nazývajú nulovými hypotézami), potom možno postupovať takto: 1. Zvolí sa malá pravdepodobnosť, zvaná hladina významnosti. 2. Podľa toho, aká hladina významnosti sa určila, dá sa odvodiť zodpovedajúca množina kritických bodov. Kritickými bodmi sa nazývajú také možné hodnoty náhodnej premennej, ktorých pravdepodobnosť je menšia ako zvolená hladina významnosti za predpokladu, že nulová hypotéza je pravdivá.

Ak sa podarí realizovať taký náhodný pokus, ktorého výsledkom bude niektorý kritický bod, nulová hypotéza o tom, že odhadnutá hodnota parametra je jeho skutočnou hodnotou, možno zamietnuť. To isté platí o akejkolvek inej testovanej hypotéze, ktorej kritický bod sa pozoroval.

To znamená, že metódy testovania štatistických hypotéz sa opierajú o voľbu nejakej hladiny významnosti. Podľa okolností to bývajú stupne pravdepodobnosti 0,01, 0,05 alebo podobne. Keď napríklad zvolíme hladinu významnosti 0,05 a predpokladáme, že nulová hypotéza je pravdivá, a pritom sa nám podarí pozorovať jav, ktorého pravdepodobnosť je za tohto predpokladu menšia ako 0,05, nulovú hypotézu odmietať. Ak sa nám to nepodarí, považujeme ju naďalej za platnú. Tento postup môže viesť k dvom chybám. Môžeme prijať nepravdivú hypotézu, alebo môžeme odmietnuť pravdivú hypotézu. Preto voľba hladiny významnosti je metodologickým problémom. Ak je hladina významnosti príliš vysoká, zvyšuje sa možnosť odmietnutia pravdivej nulovej hypotézy. Ak je hladina významnosti príliš nízka, znižujeme možnosti odmietnuť nepravdivú nulovú hypotézu. Obyčajne sa používa hladina významnosti 0,05. Prispôbujeme však túto voľbu rôznym okolnostiam. Niekedy je riskantnejšie prijať nepravdivú hypotézu, niekedy je riskantnejšie odmietnuť pravdivú hypotézu. Žiada sa ešte pripomenúť, že metódy testovania štatistických hypotéz sa používajú aj pri riešení iných otázok, ako sú štatistické opisy.

Metodologickým záverom z týchto štatistických techník by mohlo byť tvrdenie, že výsledky, ktoré sa sformulovali na základe nejakého výskumu, sú hodnotné iba vtedy, keď sa ukáže, že rozdiely z dvoch alebo viacerých výberov sú náhodné. Ak nájdeme výber, ktorého výskum vedie k výsledkom významne (signifikantne) odlišným od hypotézy sformulovanej na základe iného výberu či výberov, musíme túto hypotézu odmietnuť a uvažovaný základný súbor nepovažujeme vzhľadom na skúmanú náhodnú premennú a vzhľadom na prijatú hypotézu za štatisticky homogénny. Samozrejme, že kritériá významnej odlišnosti poskytujú štatistika.

6. Opisy sú väčšinou len základným stupňom k hlbšiemu poznaniu. Hlbšie poznanie spočíva v odhaľovaní závislostí medzi javmi, v odhaľovaní zákonitostí. Štatistika poskytuje užitočné metódy aj pre nachádzanie závislostí vo výsledkoch empirického výskumu. Keď máme sformulovaný problém o závislosti nejakého hromadného javu od iného hromadného javu a prijmeme aj nejakú hypotézu, môžeme použiť metódy testovania štatistických hypotéz. Na zisťovanie závislosti medzi dvoma hromadnými javmi alebo v presnejšej terminológii medzi dvoma náhodnými premennými môžeme používať aj iné metódy. Rôzne metódy sú vypracované pre zisťovanie závislostí medzi kvalitatívnymi náhodnými premennými a iné pre zisťovanie závislostí medzi kvantitatívnymi náhodnými premennými.

Závislosti medzi kvalitatívnymi znakmi sa zisťujú často pomocou tabuliek. Najjednoduchší prípad je, keď sledujeme závislosť (či nezávislosť) dvoch vlastností (kvalitatívnych znakov A a B) v nejakom základnom súbore. V skúmanom výbere môžeme nájsť k predmetov, na ktorých sme pozorovali znaky A i B, l predmetov, ktoré nemajú znak A a majú znak B, m predmetov, ktoré nemajú ani znak A ani znak B, n predmetov, ktoré majú znak A a nemajú znak B. Tieto výsledky sa ľahko dajú zhrnúť v štvorcovej tabuľke o štyroch políčkach. V štatistike sú vypracované rôzne testy, ktoré na základe takýchto pozorovaní umožňujú odvodiť, či rozdiely medzi jednotlivými číslami k až n sú alebo nie sú štatisticky významné, t. j. či existuje alebo neexistuje závislosť medzi znakmi A a B v uvažovanom základnom súbore. Podobným spôsobom sa zostavujú zložitejšie tabuľky, ktoré sa používajú ako podklad pre objavovanie závislostí medzi viacerými kvalitatívnymi znakmi.

Štatistika vypracovala i metódy hľadania závislostí medzi náhodnými veličinami (závislosť medzi kvantitatívnymi znakmi). Metódy odhadovania príslušného typu závislosti skúma tzv. regresná analýza a metódy zisťovania odchýlok od týchto odhadov skúma tzv. korelačná analýza. Tá nám umožňuje ukázať, či medzi dvoma kvantitatívnymi znakmi existuje alebo neexistuje závislosť, t. j. či korelujú alebo nie.

Nie je cieľom týchto úvah venovať sa štatistickej stránke veci, preto bude vhodnejšie prejsť k otázke kauzality. Za poznámku však stojí to, že metodologické otázky nemožno mechanicky oddeliť od štatistických otázok, pretože sama štatistika je vlastne metodologický nástroj a sama sa pokúša riešiť mnohé metodologické ťažkosti empirického výskumu i ťažkosti aplikácií vlastného aparátu pri spracovaní výskumných výsledkov.

7. Štatistické metódy skúmania závislostí medzi javmi sú podnetné pre moderné chápanie kauzality. Tieto metódy ukazujú, že mechanistické chápanie kauzality a indeterminizmus nie sú dve jediné a všetky ostatné možnosti vylučujúce alternatívy. Medzi hromadnými javmi či náhodnými premennými nie je prísna jednoznačná závislosť v zmysle mechanisticky chápanej kauzality. Keď v nejakom základnom súbore má určitý predmet vlastnosť A a má aj vlastnosť B, nie je zaručené, že i každý iný predmet s vlastnosťou A má vlastnosť B. Naopak, je mnoho takých predmetov, ktoré majú vlastnosť A a nemajú vlastnosť B. Táto situácia však nie je argumentom pre indeterministickú pozíciu. Videli sme, že štatistika umožňuje rozlišovať medzi situáciami, keď rozloženie vlastností v súbore ukazuje na ich závislosť a keď toto rozloženie neukazuje na ich závislosť. Ak sledujeme rozloženie dvoch kvalitatívnych znakov A a B, nezávislosť sme zistili vtedy, keď počty prvkov zo základného súboru, ktoré majú vlastnosť A i B, nemajú A ani B na jednej strane, a predmetov, ktoré majú jednu z vlastností a nemajú druhú na strane druhej, sú čisto náhodné. Keď sme zistili závislosť znakov A a B, tieto rozdiely nie sú náhodné, ale štatisticky významné — samozrejme, že na základe príslušných testov.

Zavrnutie mechanistického chápania kauzality ako striktného, jednoznačného vzťahu, kde po určitej príčine vždy bez výnimky nasleduje určitý účinok, teda nevedie k indeterminizmu, k stanovisku tvrdiacemu, že neexistujú vzťahy závislosti medzi javmi. Štatistika ukazuje, že vzťahy závislosti môžu byť rôzneho stupňa — veľmi striktné až veľmi voľné.

Ku koreláciám zisteným štatistickými metódami však treba pristupovať veľmi opatrne. Zistenie korelácie nemusí znamenať, že sme objavili nejakú závislosť. Buď je táto korelácia príliš voľná a závislá od výberu zo základného súboru, buď je metodologická ťažkosť menej zjavná: môže sa stať, že je veľmi tesná korelácia medzi znakmi A a B — skoro každý predmet s vlastnosťami A má vlastnosť B a napok — zistili sme

koreláciu aj medzi znakmi B a C, a napriek tomu sotva možno považovať C za závislý znak od B. Je to vtedy, keď C je znak závislý od A. Presvedčiť sa môžeme o tom vtedy, keď zistíme, že predmety, ktoré nemajú znak A a majú znak B, majú významne menej častejšie znak C ako predmety, ktoré majú znak A, a nemajú znak B. Takýto rozbor však vyžaduje v konkrétnych situáciách dosť značnú vynaliezavosť a opatrnosť vo vynášaní záverov.

Na druhej strane, keď sme neodhalili koreláciu medzi znakmi A a B, to ešte nemusí znamenať, že tieto znaky v danom základnom súbore nie sú od seba závislé.

S problémom kauzality úzko súvisí problematika plánovania experimentov. Kauzálna analýza nejakej situácie vyžaduje izolovanie faktorov, ktoré túto situáciu ovplyvňujú, prípadne môžu ovplyvňovať, a vedie k preskúmaniu vplyvu týchto faktorov na danú situáciu. Dôležitou vetvou štatistickej teórie je teória plánovania experimentov. Cieľom tejto práce nemôže byť rozbor základných pojmov a obsahu tejto teórie. Dôležité je to, že táto teória ukazuje, ako študovať procesy, na ktoré súčasne pôsobí viac faktorov a ako možno s dostatočnou presnosťou odkrývať stupeň pôsobenia každého uvažovaného faktora na daný proces alebo situáciu. Samozrejme, že metodologickou ťažkosťou, spätou s používaním týchto metód, je to, že skúmať sa môžu len tie faktory, ktoré sa vopred berú do úvahy ako možné príčiny skúmaného procesu, situácie či variability nejakého javu.

Druhou vážnou úlohou, ktorou sa zaoberá teória plánovania experimentov, je rozhodnutie o tom, koľko pozorovaní alebo experimentov treba na rozriešenie nejakého výskumného problému. Niekedy sa po rozbore výsledkov pozorovaní zistí, že by bolo treba rozšíriť počet pozorovaní, aby bolo možné prijať nejaké riešenie. Nie je však vylúčené, že po takomto dodatočnom rozbore z nejakých dôvodov už nemožno pokračovať v príslušnom experimente. Na druhej strane sa stáva, že sa zozbiera oveľa viac pozorovaní, ako treba k uspokojujúcemu rozriešeniu problému. Takýto experiment je zbytočne neekonomický, i z hľadiska časového i z hľadiska finančného. V medicíne tu pristupujú i problémy etického rázu.

Na odstránenie ťažkostí takéhoto druhu sú rozpracované metódy plánovania experimentov, rozvrhujúce experiment na niekoľko izolovaných na seba nadväzujúcich etáp. Po každej etape sa hodnotia výsledky a rozhoduje sa, či napríklad používať preparát A, či preparát B alebo, či treba pokračovať v experimente, pretože ešte niet dôvodov pre dôveryhodné rozhodnutie.

8. Dosiaľ uvedené fakty by mohli vyvolať dojem, že štatistika je všemocný nástroj umožňujúci riešiť všetky základné vedecké problémy. Rozhodne to však nie je tak. Rovnako ako každá iná veda, má svoje hranice aj štatistika. Hranice obmedzujúce možnosti štatistiky sú dosť silné a úzke. Potvrzuje to napokon aj názor, nie celkom ojedinelý, že štatistika je nástroj, pomocou ktorého možno dokázať čokoľvek. Tento názor sa azda opiera o mnohé nekvalifikované a nezdôvodnené použitia štatistických metód. Hranice použitia štatistických metód môžu vyplynúť z rozboru hrubého náčrtu základného cyklu vedeckej metódy. Prvá etapa vedeckej metódy je etapou zbierania empirického materiálu. V tejto etape sa hromadí množstvo skúsenostného materiálu. Materiál sa získava jednak bežnými, neplánovanými pozorovaniami, jednak pomocou systematických pozorovaní, experimentov, či meraní. Samozrejme, že pri zbieraní tohto materiálu svoju úlohu hrajú aj predstavy, koncepcie, či ciele výskumníkov, existujúce teórie, ba dokonca aj svetonázor v širokom zmysle slova. Tieto činitele vplyvajú aj na predbežné systematizovanie skúsenostného materiálu i na formuláciu vhodných hypotéz. Formulovaním takejto hypotézy, domyslením jej dôsledkov, jej zosúladením s existujúcimi názormi a načrtnutím zdedpovedajúcej teórie by mohla byť charakteri-

zovaná druhá etapa vedeckej metódy. V ďalšej etape sa príslušná hypotéza podrobuje premyslenej a systematickej experimentálnej previerke. Jej výsledky alebo potvrdia danú hypotézu, alebo nie. V druhom prípade možno sformulovanú hypotézu celkom odmietnuť alebo len modifikovať. Odmietať ju spravidla vtedy, keď to je izolovaná hypotéza, keď nie sú známe vzťahy tejto hypotézy k existujúcim systémom poznatkov, k existujúcim vedeckým teóriám. Ak sformulovanú hypotézu možno dobre zaradiť do existujúcich teórií, ak je určitým spôsobom týmito teóriami podopieraná viera v zásadnú správnosť danej hypotézy, je na mieste pokúšať sa o vhodnú modifikáciu tejto hypotézy. Tým sa dostávame opäť na začiatok cyklu, ktorým možno charakterizovať vedeckú metódu, samozrejme, že materiál, ktorý je tu už zozbieraný, je na vyššej úrovni ako materiál v predchádzajúcom „obehu“ cyklu a problémy, ktoré treba riešiť, sú zložitejšie. Takýto zjednodušený pohľad na vedeckú metódu dáva istý obraz aj o mechanizmoch vedeckého pokroku a vytvárania čoraz zložitejších problémov.

Je zrejme, že štatistika sa nemôže uplatniť v prvej spomínanej etape. Ohromný a chaotický komplex veľku náhodne pospájaných pozorovaní a poznatkov nemôže byť predmetom štatistického spracovania. Druhá etapa je kľúčová z hľadiska aplikability štatistických metód. Nie každá hypotéza môže byť overovaná prostriedkami štatistiky. Preto, keď sa zamýšľa aplikovanie štatistických metód, treba príslušným spôsobom aj formulovať hypotézu. Hypotéza sa musí týkať nejakej náhodnej premennej, môže sa preformulovať pomocou pojmov teórie pravdepodobnosti a musí byť vyvrátiteľná alebo overiteľná štatistickými metódami. To je teda prvá obmedzujúca podmienka, ktorá vystupuje v druhej etape vedeckej metódy. Druhé obmedzenie je dané tým, že formulovanie každej hypotézy je výsledkom určitej idealizácie. Ak je táto hypotéza formulovaná pomocou pojmov teórie pravdepodobnosti, hovorí sa, že hypotéza je matematickým modelom skutočnosti. Samozrejme, že každý matematický model skutočnosti je jej zjednodušením, výsledkom idealizácie, iba približným obrazom skutočnosti.

Z toho vyplýva, že štatisticky overované hypotézy platia len približne, nehovoriac už o tom, že len s určitou pravdepodobnosťou. Okrem toho takto sformulované hypotézy možno overovať alebo vyvracať len takým materiálom, ktorý spĺňa podmienky utvorené príslušnou idealizáciou. Problém idealizácie je metodologicky i filozoficky závažný, preto si zasluhuje väčšiu pozornosť. Každý štatistický opis základného súboru, utvorený na základe skúmania nejakých výberov, i každý štatistický výskum nejakej závislosti, vychádzajúci z výsledkov zistených o výberoch, je určitou idealizáciou.

Keď štatistické charakteristiky skúmaných výberov kolíšu okolo nejakej hodnoty, dovoľujeme si odhadnúť, aká je hodnota príslušného parametra, alebo v akom intervale leží skutočný parameter. Takéto tvrdenie je výsledkom určitej idealizácie. Je to obdobné ako pri dobre známom príklade ideálneho plynu. Namerané hodnoty tlaku, objemu a teploty skutočných plynov iba približne vyhovujú známej stavovej rovnici, ktorá vymedzuje správanie ideálneho plynu. Odhadnutie nejakého parametra alebo hypotéza o ideálnom plyne sú výsledkom idealizácie preto, lebo idealizujú situáciu v tom, že pripúšťajú akési ideálne podmienky dovoľujúce stanoviť požadované hodnoty (v prvom prípade hodnotu príslušného parametra, v druhom prípade hodnoty vyhovujúcej stavovej rovnici) a neberú do úvahy reálne okolnosti znemožňujúce tieto hodnoty získať.

Podľa dosť rozšíreného názoru biologické javy sú natoľko zložité a rozmanité, že idealizácie a používanie matematických modelov v tejto oblasti vedie k neprímerane veľkým zjednodušeniam, a tým aj k chybným záverom. I keď na tomto názore môže byť zrnko pravdy, najmä keď sa berú do úvahy niektoré príliš zjednodušujúce matematické modely, vo svojom jadre je konzervatívny.

Dôležitou idealizáciou sú rôzne rozdelenia pravdepodobnosti náhodných premen-

ných. Pozorovania výberov vždy len približne potvrdzujú, že určitý jav podlieha niektorému typu rozdelenia pravdepodobnosti. Idealizácia je prípustná alebo dobrá vtedy, keď pozorovania výberov sú v dobrej zhode s teoretickými predpokladmi, keď sa od nich podstatne nelíšia. Rozhodnutie o dobrej zhode možno urobiť na základe už spomínaných testovacích metód. Prijaté idealizované teoretické rozdelenia pravdepodobnosti náhodnej premennej v základnom súbore potom dovoľujú používať vhodné štatistické metódy pre overovanie rôznych hypotéz, týkajúcich sa štatistických vlastností základných súborov. Tým sme sa dostali k ďalšej hranici použiteľnosti štatistických metód. V tretej etape vedeckej metódy, pri overovaní hypotéz, nemožno napríklad používať také testy alebo metódy odhadu parametrov, ktoré sú vhodné pre súbory s inými rozdeleniami pravdepodobnosti, ako je skúmaný súbor. Štatistické metódy sa všeobecne viažu na isté typy rozdelení pravdepodobnosti. Niektoré metódy plánovania experimentov, ktorých cieľom je preskúmať pôsobenie jednotlivých faktorov, predpokladajú normálne rozklady uvažovaných veličín. Všeobecne by sa dalo povedať, že použitie určitej štatistickej metódy je závislé od príslušných štatistických vlastností skúmaného súboru.

Z konkrétnych typov chýb, ktoré sa môžu vyskytnúť pri aplikovaní štatistických metód, treba uviesť aspoň tieto: voľba nereprezentatívneho výberu, zovšeobecnenie na základe skúmania príliš malého výberu, nesplnenie predpokladov, o ktoré je opretá použitá štatistická metóda, pretože nie každá metóda je aplikovateľná na ľubovoľný výber alebo ľubovoľný základný súbor. Rôzne testy štatistickej významnosti a dobrej zhody majú rôzne hranice použiteľnosti, sú vhodné len pre určité typy dát atď.

Záver. I keď sa v tejto práci venovalo málo pozornosti štatistickým detailom, možno vzbudí dojem, že neúmerne málo miesta venovala metodologickým problémom. Ospravedlnením tohto prístupu alebo skôr vysvetlením tohto dojmu nech je to, že štatistika je skutočne vedou metodologického charakteru. Jej aplikovanie na výskumný materiál je vlastne používaním vhodných metód spracovania výskumných výsledkov. Existuje veľa styčných plôch filozofie a metodológie s aplikáciami štatistiky. Štatistika sa všade v práci chápa, prirodzene, ako určitá teória a metodologická veda, nie ako súhrn nejakých relatívnych alebo absolútnych čísel, charakterizujúcich výskyt určitých javov. Cieľom práce bolo opísať problémy, na riešenie ktorých možno použiť štatistiku pri spracúvaní výskumných výsledkov, ďalej poukázať na filozofické problémy kauzality a idealizácie s tým späť a upozorniť na hranice štatistických metód.

J. Š.