

ZDENĚK MURDYCH

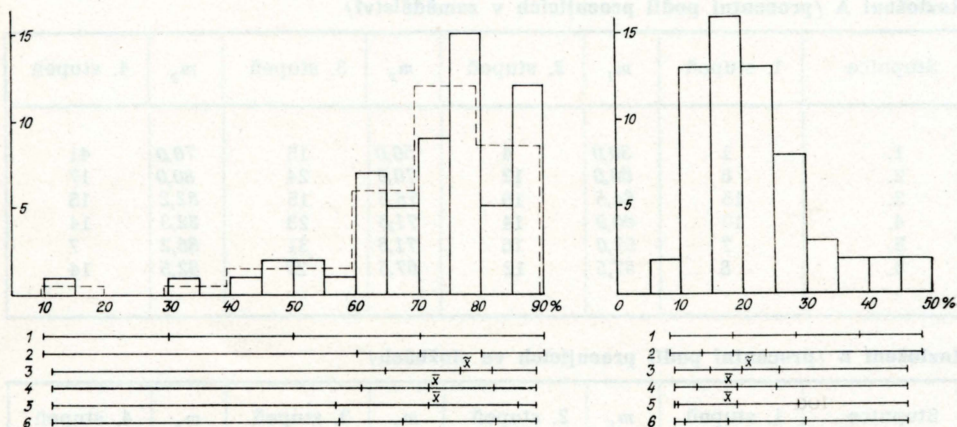
K OTÁZCE STANOVENÍ INTERVALŮ STUPNIC KARTOGRAMŮ

Stanovení intervalů stupnic patří k jednomu z prvních kroků při rozhodování o celkové koncepci mapy či kartogramu. Jde o problém, kolik zvolit stupňů stupnice a jak určit hranice mezi nimi. Tato otázka se někdy sleduje i v učebnicích kartografie i ve speciálních pojednáních, avšak nelze říci, že by byla již zcela rozpracována. Možností určení stupnic je celá řada a stále zůstává problémem, kterou metodu je třeba vybrat pro daný úkol jako nevhodnější. Přesto však se ať již při tvorbě mapových a atlasových děl nebo při vyhotovování kartografického doprovodu geografických prací používá ve velké většině stále jen nejjednodušších metod určení stupnic, někdy pro vyjadřování příslušných jevů i nevhodných. Stupnice se většinou určují tak, že se celá variační šíře souboru sledovaných jednotek pravidelně rozdělí určitými okrouhlými hodnotami. Při tom se obvykle nebere v úvahu, jaké je rozložení četností v jednotlivých intervalech stupnice, takže celková plocha kartogramu může být různými barvami či rastry nevhodně zaplněna. Předvedeme zde několik druhů stupnic stanovených jednak běžnými metodami empirickými, jednak jednoduchými metodami statistickými a provedeme jejich porovnání a posouzení jejich použití. Základem naší analýzy bude sledování rozložení četností jednotlivých hodnot souboru pomocí frekvenčního grafu. K získání potřebných výsledků lze používat i křivek sumáčnic, jimiž se například zabývá ve své práci, která je poněkud jinak zaměřena, americký autor J. R. MACKAY (1).

Zde je pro porovnání použito šesti druhů stupnic, které se aplikují pro dvojí ukazatele: podíl obyvatelstva (z celkového počtu aktivního obyvatelstva) zaměstnaného jednak v zemědělství (rozložení A), jednak ve službách (rozložení B), obojí na území Dolního Egypta, který je tvořen v podstatě deltou Nilu. Tyto dva ukazatele byly vybrány proto, že se vyznačují odlišným rozložením četností; jejich frekvenční graf je v prvním případě nepravidelný, ve druhém případě jde naopak o pravidelnější plynulé rozložení skupinových četností, blízké rozložení normálnímu. Kromě toho jsou tyto jevy na sobě — jak bude dále blíže ukázáno — statisticky nepřímou závislé. K uvedeným dvěma typům rozložení by bylo možno připojit ještě typy další, jejichž frekvenční čáry by měly podobu přímek buď rovnoběžných s vodorovnou osou (všechny velikostní skupiny by se vyznačovaly zhruba stejnou četností) nebo přímek či křivek plynule rostoucích či klesajících, a to podle různé posloupnosti. V případě, že by frekvenční křivka měla tvar paraboly, bylo by na místě užít stupnice představující geometrickou řadu, podíl sousedních mezi intervalů by tedy byl konstantní. Této problematice se věnuje ve své studii J. R. MACKAY (1). Zmíněná rozložení ale nejsou v geografii příliš běžná; jde naopak často o soubory jevů, které je možno vyjádřit frekvenční křivkou, mající blízko ke Gaussově křivce (normálního rozložení). Jedním z problémů při volbě intervalů stupnic je stanovení počtu tříd stupnice. Odhlédneme zde od této otázky, která není přímým předmětem naší studie a vy-

jděme z požadavku, aby stupnice měla 4 stupně, což pro méně četné soubory může být postačující a má naopak tu výhodu, že orientace v kartogramu je snazší.

Frekvenční grafy obou našich rozložení jsou vyjádřeny obrázky 1 a 2. V obou



1. Frekvenční graf rozložení A (podíl pracujících v zemědělství na území Dolního Egypta), znázorňující (podle údajů na svislé ose) počet okresů v jednotlivých velikostních třídách. Dole jsou znázorněny vždy 4 intervaly šesti použitých stupnic. (Bližší výklad všech obrázků v textu).

2. Frekvenční graf rozložení B (procentní podíl pracujících ve službách na stejném území).

případech byl pro grafické vyjádření celý soubor hodnot rozdělen do intervalů po pěti procentech, což se jeví pro uvažovaný soubor (61 okresů) jako přiměřeně velký interval. Stanovení menšího intervalu by vzhledem k nepříliš velkému počtu jednotek nebylo vhodné, použití většího intervalu by zase vedlo k setřetí významných nepravidelností frekvenčního grafu. To je znázorněno na obr. 1, kde je přerušovanou čarou vyjádřen tvar grafu při použití dvojnásobně velkého intervalu — deseti procent. Je zřejmé, že takto se pak dobře neprojeví členitost grafu právě v intervalu 60 % — 90 %, kde se jinak vyskytují dvě sedla, důležitá pro stanovení stupnice. Variační šíře druhého souboru je, jak zřejmo, vzhledem k prvnímu zhruba poloviční.

Budeme zde uvažovat a porovnávat následujících 6 druhů stupnic, jejichž intervaly jsou stanoveny takto:

1. rovnoměrným rozdělením celé variační šíře okrouhlými hodnotami,
2. rozdělením variační šíře souboru okrouhlými hodnotami tak, že pravidelně je rozdělen jen úsek velkých četností,
3. mediánem a dolním a horním kvartilem,
4. průměrem a průměrnou odchylkou od průměru,
5. průměrem a standardní odchylkou,
6. mezemi modálního intervalu a ostatními význačnými hranicemi vyjadřujícími minima (sedla) frekvenčního grafu.

V obr. 1 a 2 jsou všechna rozdělení vyznačena pod frekvenčními grafy. Zmíněná dvojitá rozložení lze charakterizovat následujícími tabulkami, které ukazují kolik případů je za použití 1.—6. stupnice obsaženo v 1.—4. stupni. Meze dělicí soubor jsou označeny m_1 , m_2 a m_3 .

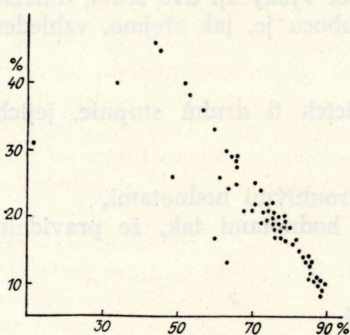
Rozložení A (procentní podíl pracujících v zemědělství)

Stupnice	1. stupeň	m_1	2. stupeň	m_2	3. stupeň	m_3	4. stupeň
1.	1	30,0	4	50,0	15	70,0	41
2.	8	60,0	12	70,0	24	80,0	17
3.	15	64,5	16	76,5	15	82,2	15
4.	10	60,9	14	71,6	23	82,3	14
5.	7	57,0	16	71,6	31	86,2	7
6.	8	57,5	12	67,5	27	82,5	14

Rozložení B (procentní podíl pracujících ve službách)

Stupnice	1. stupeň	m_1	2. stupeň	m_2	3. stupeň	m_3	4. stupeň
1.	22	18,0	26	28,0	8	38,0	5
2.	9	13,0	13	18,0	20	23,0	19
3.	15	14,5	16	19,5	15	25,5	15
4.	9	12,3	17	18,7	20	25,1	15
5.	1	9,0	25	18,7	24	28,4	11
6.	4	10,5	18	17,5	30	29,5	9

V prvním případě (procento pracujících v zemědělství) je nakupení hodnot na pravé straně variační řady (jde tedy o pravostranně asymetrické rozložení), v druhém případě je tomu obráceně (viz obr. 1 a 2). Služby jsou totiž — podobně jako průmysl — v podstatě komplementárním sektorem zemědělství: stoupá-li zaměstnanost v zemědělství, klesá v průmyslu a ve službách. Služby jsou do jisté míry jakýmsi doprovodným sektorem průmyslu (neplatí to ve všech jednotlivých případech); oba tyto sektory jsou totiž silně vázány na města, resp. silněji urbanizovaná a industrializovaná území.



Nepřímou úměrnost vztahu zaměstnanosti ve službách a v zemědělství názorně dokumentuje kolerační diagram (obr. 3). U většiny okresů (v Egyptě zvaných markazů) jde téměř o pravidelný pokles podílu pracujících ve službách, stoupá-li zaměstnanost v zemědělství. Pouze u několika nověji industrializovaných okresů, např. v blízkosti Káhiry, je tato pravidelnost porušena, neboť jde sice o místa silně průmyslová, ale ne s odpovídajícím podílem služeb: obyvatelé patrně využívají služeb blízké metropole.

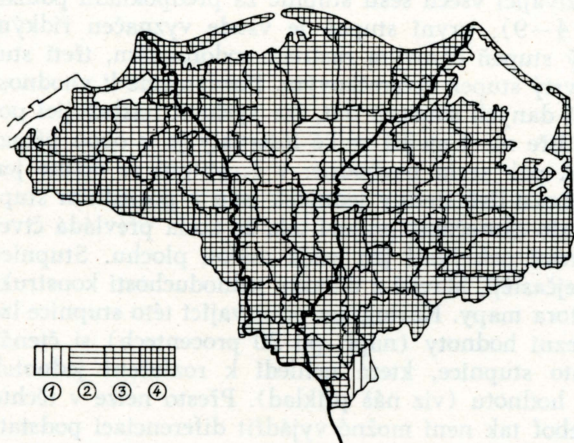
3. Korelační diagram závislosti zaměstnanosti ve službách (na svislé ose) na zemědělské zaměstnanosti (na ose vodorovné).

Pro porovnání byly pro soubor A (podíl obyvatelstva pracujícího v zemědělství) sestaveny kartogramy používající všech šesti stupnic za předpokladu použití stejných typů rastrů (viz obr. 4—9): první stupeň je všude vyznačen řídkým čárovým rastrem svislým, druhý stupeň čárovým rastrem vodorovným, třetí stupeň rastrem obdélníčkovým a čtvrtý stupeň čtverečkovým. Posuzujeme-li vhodnost uvedených stupnic pro vyjádření daných souborů dat, lze vyslovit následující poznatky. Rozdělení celé variační šíře na 4 stejně velké intervaly je v obou případech nepatřičné, pro rozložení A pak zcela nevhodné: v 1. stupni se ocitne jen 1 okres, ve druhém pouze 4, zatímco dvě třetiny jednotek jsou v posledním stupni. Na kartogramu (obr. 4) se tato skutečnost projeví tak, že zcela převládá čtverečkový rastr, kdežto rastry čárové pokrývají jen zcela malou plochu. Stupnice tohoto typu se obecně používá nejčastěji, hlavně z důvodu jednoduchosti konstrukce, někdy i neinformovanosti autora mapy. Kartogram používající této stupnice lze snadno sestavit i čist, neboť mezní hodnoty (např. po 20 procentech) si čtenář snadno může zapamatovat. Tato stupnice, která nehledí k rozložení jednotek v souboru, zdůrazňuje extrémní hodnotu (viz náš příklad). Přesto nelze v těchto případech stupnici doporučit, neboť tak není možné vyjádřit diferenciaci podstatného množství jednotek souboru a kartogram je málo informativní. Na extrémní hodnotu (v našem případě markaz Al-Kheima severně od Káhiry) lze upozornit v textu pojednání a není třeba na tuto jednu plošku vyplývat celý jeden stupeň stupnice.

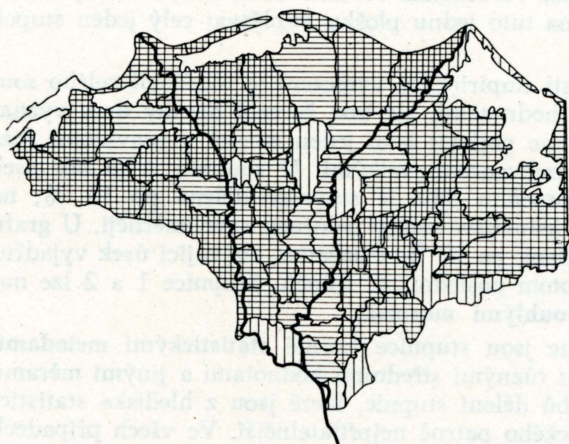
Jinou alternativou z této oblasti empirických vymezení je rozdělení celého souboru hodnot rovněž okrouhlými hodnotami, ale tak, že celý dlouhý úsek vyznačující se malými četnostmi budeme pojímat jako jeden stupeň a pravidelně rozdělíme jen úsek vyznačující se frekvencemi velkými. U souboru A je tak úsek 10—60 % stanoven jako 1. stupeň a dále je stupnice dělena po 10 %, na kartogramu (obr. 5) jsou potom jednotlivé rastry položeny rovnoměrněji. U grafu souboru B je naopak stupnice dělena po 10 % z počátku, zbývající úsek vyjadřující 25 % (23—48 %) tvoří potom poslední, 4. stupeň. Stupnice 1 a 2 lze nazvat souhrnně **stupnicemi s okrouhlými mezemi**.

Jinou velkou skupinou stupnic jsou stupnice určené **statistickými metodami**. Určení mezi intervalů lze provést různými středními hodnotami a jinými měrami. Zde se předkládá několik způsobů dělení stupnic, které jsou z hlediska statistického, geografického a kartografického patrně nejpříjemnější. Ve všech případech jde především o získání střední hodnoty souboru, kterou je postupně medián, průměr a modus. Pro konstrukci 3. stupnice používáme *mediánu* a dále dolního a horního kvartilu; tím se celý soubor rozdělí na 4 díly o stejném počtu okresů (je to třikrát 15 a jednou 16 okresů, protože zde, bohužel, celkový počet není dělitelný čtyřmi). Tímto postupem se dosáhne toho, že všechny druhy rastrů zabírají na kartogramu přibližně stejnou plochu (viz obr. 6). Podobnou metodu lze aplikovat, chceme-li získat jakýkoliv počet stupňů; potom je nutno celý soubor dat rozdělit příslušnými kvantily, např. decily, oktily, septily atd. Tato metoda získání mezi intervalů kvantily byla v širokém měřítku použita v Atlasu Londýna a Londýnské oblasti (9), kde pro řadu mapových listů znázorňujících rozmístění obyvatelstva a hospodářství je použito dělení stupnice na 8 stupňů, jejichž meze jsou určeny oktily. Tato koncepce je pro svou svéráznost a jednotné použití pozoruhodná; její užitečnost nesnižuje příliš ani to, že je často v atlase použita pro kartogramy vyjadřující současně různě velká území (malé, hustě obydlené městské čtvrti a velké, ale řídké osídlené vesnické obce), takže celkové plochy stejných barev jsou někdy dosti odlišné.

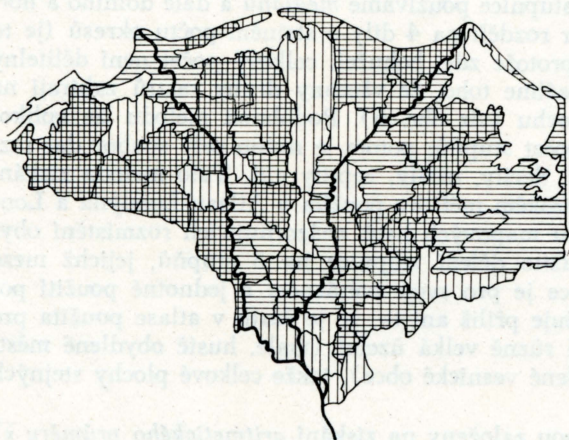
Další dvě metody (4. a 5.) jsou založeny na získání *aritmetického průměru* \bar{x} ;



4. Kartogram znázorňující podíl zemědělského obyvatelstva v okresech Dolního Egypta při použití 1. stupnice (vnitřní meze intervalů: 30—50—70 %). Rastry pro 4 stupně platí pro všechny stupnice.

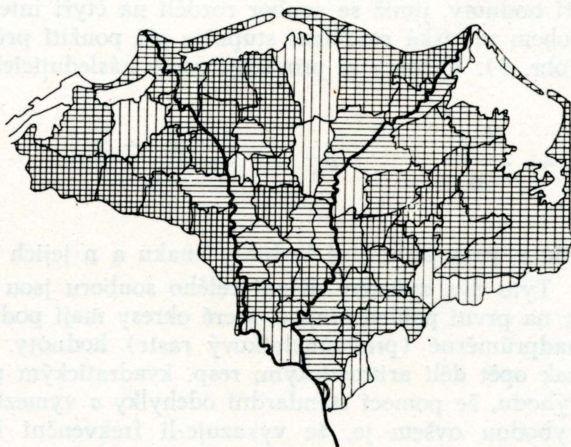


5. Kartogram používající stupnice 2 (meze intervalů: 60—70—80 %).

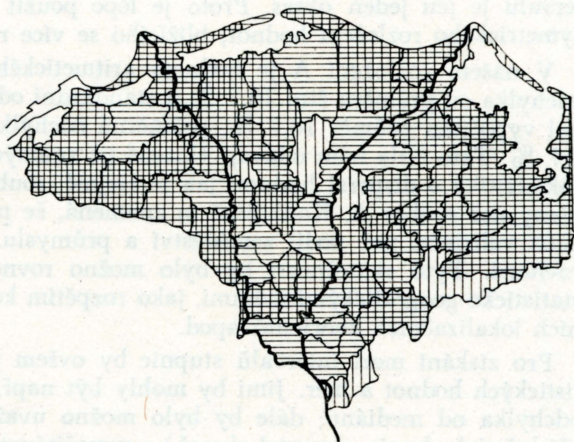


6. Kartogram používající stupnice 3 (meze stanoveny mediánem a kvartily: 64,5—76,5—82,2 %).

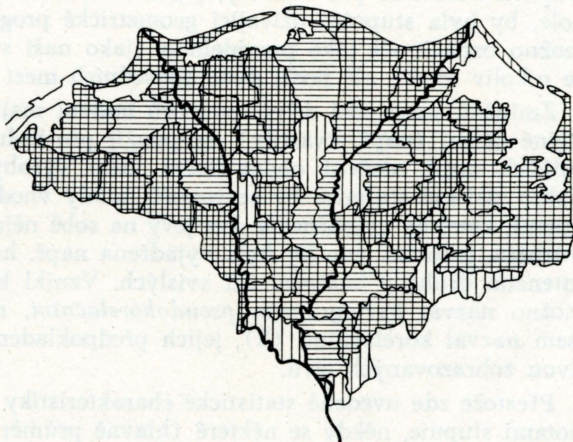
7. Kartogram používající stupnice 4 (meze stanoveny průměrem a průměrnou odchylkou od průměru: 60,9—71,6—82,3%).



8. Kartogram používající stupnice 5 (meze stanoveny průměrem a standardní odchylkou: 57,0—71,6—86,2 %).



9. Kartogram používající stupnice 6 (meze stanoveny individuálně — v minimech grafu: 57,5—67,5—82,5 %).



od této hodnoty se pak vypočítá průměrná odchylka \bar{d}_x , a tak se získají celkem tři hodnoty, jimiž se soubor rozdělí na čtyři intervaly (obr. 7). Obdobným způsobem se získá rozdělení stupnice při použití průměru a standardní odchylky σ (obr. 8). Výpočty se provádějí podle následujících vzorců:

$$\bar{x} = \frac{\sum x_i}{n} \quad \bar{d}_x = \frac{\sum |x_i - \bar{x}|}{n} \quad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

kde x_i jsou jednotlivé hodnoty znaku a n jejich celkový počet.

Tyto dva způsoby dělení celého souboru jsou vhodné proto, že z kartogramů je na první pohled zřejmé, které okresy mají podprůměrné (čárový rastr) a které nadprůměrné (pravouhelníkový rastr) hodnoty. Tyto dvě poloviny souboru se pak opět dělí aritmetickým, resp. kvadratickým průměrem; druhý způsob má tu výhodu, že pomocí standardní odchylky σ vymezíme normální obor souboru. Nevýhodou ovšem je, že vykazuje-li frekvenční křivka větší sešikmení doprava (doleva), pak na pravé (levé) straně stupnice se v krajním intervalu ocitne jen málo jednotek. To se projevuje zejména u našeho rozložení B, kde v prvním intervalu je jen jeden okres. Proto je lépe použít tohoto dělení v případech více symetrického rozložení hodnot, blížícího se více rozložení normálnímu.

V našem rozložení A je hodnota aritmetického průměru 71,6 %, průměrná odchylka od průměru činí 10,7 % a standardní odchylka 14,6 %. Pro rozložení B byl vypočítán průměr 18,7 %, průměrná odchylka 6,4 % a standardní odchylka 9,7 %. Tato čísla nám udávají (kromě již známých variačních rozpětí) další charakteristiky o disperzi hodnot: jak je zřejmé, soubor B je zhruba dvakrát více sevřený než soubor A. Konkrétně to znamená, že podíl služeb v okresech je stabilnější veličinou než podíl zemědělství a průmyslu, které jsou naopak dosti nevyrovnané. Tyto skutečnosti by bylo možno rovněž měřit a vyjadřovat i jinými statisticko-geografickými měrami, jako rozpětím koncentračních indexů nebo místních lokalizačních kvocientů apod.

Pro získání mezi intervalů stupnic by ovšem bylo možno použít i jiných statistických hodnot a měr. Jimi by mohly být např. geometrický průměr, průměrná odchylka od mediánu; dále by bylo možno uvažovat váhy (dané např. počtem příslušných skupin obyvatel okresů) a vypočítávat vážené průměry atd. Zajímavou stupnicí hodící se pro soubory, jejichž frekvenční křivka by byla podobná parabole, by byla stupnice užívající geometrické progrese. Takovou stupnici by bylo možno označit též jako pravidelnou (jako naši stupnici 1.), ale s tím rozdílem, že nikoliv rozdíl, ale podíl dvou sousedních mezi by byl konstantní.

Zmíněná statistická dělení souborů hodnot mají své výhody zejména v tom, že stejné dělení (např. kvartily) lze použít pro řadu různých ukazatelů charakterizujících např. složení obyvatelstva nebo výroby příslušných územních celků. Takto získané stupnice by se rovněž mohly vhodně použít při konstrukci kartogramů, které by vyjadřovaly dva jevy na sobě nějak stochasticky závislé, při čemž intenzita jednoho jevu by byla vyjádřena např. hustotou vodorovných čar rastru, intenzita druhého hustotou čar svislých. Vznikl by tak kartogram, který by bylo možno nazvat kartogramem *pseudokorelačním*, na rozdíl od kartogramů, které jsem nazval *korelačními* (8), jejich předpokladem je zjištění korelační závislosti dvou zobrazovaných jevů.

Přestože zde uvedené statistické charakteristiky nebývají obvykle mezními hodnotami stupnic, někdy se některé (hlavně průměr) na stupnici vyznačují. Příkla-

dem toho je náš Atlas ČSSR (5), kde jsou průměrné hodnoty vyznačeny na stupnici červenou čárkou a číslem.

Poslední zde probíranou metodou konstrukce stupnic kartogramů je metoda označená v grafech číslem 6, použita je na obr. 9. Lze ji označit za metodu geograficko-statistickou; její aplikace je individuální; sleduje se získání stupnice tak, aby mezní hodnoty sledovaly charakteristické body té které frekvenční křivky. Při tomto dělení se uplatní třetí druh střední hodnoty, a sice *modus*, v našich případech spíše modální interval. Je to interval vyznačující se nejvyšší četností (v grafech nejvyšším sloupkem). Je-li frekvenční graf více členitý, tj. má-li více vrcholů a sedel, je namíste použít jako mezních hodnot právě těchto sedel, tedy středů intervalů, které mají malé frekvence a jsou obklopeny frekvencemi většími. Jsou to u našeho souboru A hodnoty 57,5 %, 67,5 % a 82,5 %; úsek mezi dvěma posledními hodnotami je právě modálním intervalem. Druhé rozložení je vyjádřeno grafem, který není rozčleněn sedly. Zde je střední hodnota modálního intervalu (17,5 %) použita pro dělení souboru na dvě poloviny a další dvě mezní hodnoty jsou voleny přibližně v místech, která určují inflekční body frekvenční křivky sestrojené jako vyrovnávací křivky našeho frekvenčního histogramu. Toto členění, které v podstatě nejvíce přihlíží ke geografické realitě (zřejmě zejména u souboru A), se v geografii a kartografii bohužel rovněž málo používá.

Jak bylo zmíněno, metod vymezení intervalů stupnic je veliké množství. Ze zde uvedených metod je rovněž možné tvořit metody kombinované; je však třeba též hledět na to, aby užitý způsob byl metodicky čistý a nebyly násilně spojovány metody různých koncepcí. Celkově je však nutno prosazovat v naší geografii a kartografii větší pestrost v tvorbě a používání stupnic tak, aby užitá metoda vždy co nejlépe vyhovovala charakteru znázorňovaného jevu a autorově pojetí.

Literatura

1. MACKAY J. R.: An Analysis of Isopleth and Choropleth Class Intervals. *Economic Geography* 31: 71—81, Worcester 1955.
2. ROBINSON A. H.: *Elements of Cartography*. New York 1960. 344 str.
3. *The Population Census of Egypt 1960*. Cairo 1963.
4. KOVAŘÍK J., DVORÁK K.: *Kartografie*. SNTL, Praha 1964. 384 str.
5. *Atlas Československé socialistické republiky*. ÚSGK, Praha 1966. 53 map.
6. ARNBERGER E.: *Handbuch der thematischen Kartographie*. Wien 1966. 566 str.
7. WITT W.: *Thematische Kartographie*. Hannover 1967. 384 str.
8. MURDYCH Z.: Korelační plošné kartogramy. *Sborník Čs. společnosti zeměpisné* 72: 235—243, Praha 1967.
9. *Atlas of London and the London Region*. Oxford — London 1968. 70 map.
10. MURDYCH Z.: K otázce kartografických metod studia areálů dojížděky. *Demografie* 11: 245—253, Praha 1969.

ON THE PROBLEM OF CHOROPLETH CLASS INTERVALS

The paper deals with possible methods of determination of class intervals in choropleth maps. The application of six chosen methods is used for the territory of Lower Egypt (the Nile Delta), which is divided into 61 counties (in Egypt named „markazes“). The selection of class intervals is based on frequency graphs. Two features are expressed by these graphs: the ratio of workers employed in agriculture and fishing (distribution A) and the ratio of workers employed in services (distribution B), both in percentage out of the total number of working population. The first distribution is an example of varying distribution, the latter one is more uniform and it does not differ too much from the Gaussian distribution. There is a close negative relationship between these distributions as it is shown by Fig. 3.

Four class intervals are used for all six class interval scales and each interval is expressed in choropleth map by the same pattern. The limits of class intervals are selected by the following methods: 1. uniform arithmetic increase, 2. uniform division only in the section of great frequencies, 3. median and lower and upper quartile, 4. arithmetic mean and mean deviation from the mean, 5. arithmetic mean and standard deviation, 6. modal interval and other intervals of great frequencies. The number of counties in individual class intervals is shown by the tables; three limits of intervals are marked m_1 , m_2 , m_3 . One of the purposes of the article is to compare several methods and to show the advantages of various methods; namely the methods based on statistical calculations are important. In spite of this fact these methods are used in practical geography and cartography very rarely.

Explanations of the figures

1. The frequency graph of the distribution A. The class interval grouping is scaled off along the horizontal axis, the frequency of occurrence along the vertical axis.
2. The frequency graph of the distribution B.
3. The scatter diagram showing the dependance of the distribution B on the distribution A.
4. Choropleth map showing the regional differences in dislocation of agricultural population; the map uses the 1st method of selection of class intervals. The territory of Lower Egypt is divided into 61 counties. Line patterns (1, 2, 3, 4) are the same for all methods.
- 5.—9. Choropleth maps for the same feature and territory constructed according the class interval scales 2—6.