

METODOLOGICKÁ VÝCHODISKA ANALÝZ REGIONÁLNÍHO ROZVOJE

METHODOLOGICAL STARTING POINTS OF REGIONAL DEVELOPMENT ANALYSES

Libuše Svatošová

Anotace:

Při hodnocení regionálního rozvoje je nutné sledovat a analyzovat řadu ukazatelů, které charakterizují regiony z hlediska přírodních a geografických podmínek, ekonomické situace, lidského potenciálu, sociální úrovně, životního prostředí a dalších. Analyzovaná data tak představují vícerozměrný statistický soubor, ve kterém sledujeme velké množství proměnných, mezi nimiž existují, respektive mohou existovat vzájemné vztahy. Příspěvek diskutuje možnosti využití metod vícerozměrné analýzy dat.

Klíčová slova:

Vícekritériální hodnocení, disponibilní potenciál regionu, vícerozměrné statistické metody, analýza hlavních komponent, shluková analýza, regresní analýza.

Abstract:

When assessing regional development it is necessary to assess and analyse many indicators describing the regions as to the natural and geographical conditions, economic situation, human potential, social levels, environment and others. The data analysed represent a multidimensional statistical population, where many variables are observed and relationships among these variables exist or may exist. The paper deals with the possibilities of the multivariate statistical data analysis application.

Keywords:

Multicriterial assessment, available potential of the region, multivariate statistical methods, principal component analysis, cluster analysis, regression analysis.

ÚVOD

Otázky sledování a hodnocení regionálního rozvoje patří v souvislosti se vstupem České republiky do Evropské unie ke klíčovým ekonomickým otázkám. Smyslem a cílem regionálního přístupu je zajištění stejných šancí a možností rozvoje regionů tak, aby jejich demografický, přírodní a hospodářský potenciál byl plnohodnotně využíván. Pouze tržní prostředí není schopno zajistit vyvážený rozvoj celého území státu. Má-li být zajištěn rovnoměrný rozvoj všech regionů, je nutné přijímat určitá opatření jak na úrovni regionální, státní i nadnárodní (EU), která povedou ke zmírnění či odstranění regionálních rozdílů. K tomu, aby prostředky, které stát či EU vloží na rozvoj regionů, byly účelně využity a splnily svůj účel, je třeba znát důkladně situaci v regionech, zjistit zdroje a rezervy, prokázat životaschopnost a perspektivy daného regionu. Je tedy třeba provést důkladné hodnocení dosavadního vývoje jednotlivých ukazatelů s cílem vytvořit takovou soustavu indikátorů, které by kvantifikovaly jednotlivé složky rozvojového potenciálu regionů. Tento úkol se

neobejde bez kvalifikovaných statistických analýz.

METODIKA

Jednou z možností je posuzovat vybrané ukazatele regionálního rozvoje odděleně – samostatně popsat jejich vývoj a stanovit prognózy budoucího vývoje. Toto je dobrým východiskem, ale pro rychlé a zejména jednoduché posouzení rozvojového potenciálu není příliš vhodné, neboť orientace v řadě ukazatelů je velmi pracná a málo přehledná. Navíc z analýz jednotlivých ukazatelů lze jen velmi těžko usuzovat na vzájemné vztahy a souvislosti. Výhodnější formou se zdají být indikátory souhrnné (např. ve formě koeficientu), které v sobě zahrnou podstatné a charakteristické rysy jednotlivých složek regionálního rozvoje. Takové ukazatele sice nebudou na první pohled vyhovovat požadavku srozumitelnosti a jednoduchosti, to však lze kompenzovat důkladným a srozumitelným metodickým popisem jejich konstrukce. Budou však uvedenou problematiku pojímat v celém komplexu a souvislostech. Indikátory by pak měly sloužit jako informační systém vypovídající o míře rozvoje regionu a o schopnosti využívat prostředky do regionu plynoucí či v regionu vytvořené. Schopnost rozvoje je nutné průběžně vyhodnocovat a získané informace zpětně zabudovat do rozhodovacího procesu.

Analýza rozvoje regionů by měla zahrnovat všechny ukazatele, jejichž sledování a hodnocení má při stanovení rozvojového potenciálu velký význam. Jedná se zejména o ukazatele charakterizující:

- Souhrnný popis regionu
- Ekonomickou situaci v regionu
- Lidský potenciál v regionu
- Sociální úroveň
- Infrastrukturu
- Životní prostředí
- Průmysl
- Služby
- Venkov a zemědělství.

Každý z výše uvedených ukazatelů zahrnuje celou řadu proměnných. Statistické analýzy vycházejí tedy z poměrně široké datové základny. Zpracováváme-li takto rozsáhlý soubor ukazatelů, kdy výchozí počet znaků je značný a pro interpretaci nepřehledný, můžeme s výhodou užít některou z metod vícerozměrných statistických analýz. Částečně redukováný soubor proměnných může být posléze dále analyzován podstatně jednodušším způsobem. Metody vícerozměrné statistické analýzy slouží k řešení několika úkolů, jimiž jsou :

- 1) Redukce nadměrného počtu proměnných, respektive zhuštění informací do menšího počtu neměřitelných hypotetických veličin s co nejmenší ztrátou informace, která je obsažena ve sledovaných znacích. Do této skupiny metod řadíme analýzu hlavních komponent, faktorovou analýzu, kanonickou analýzu.
- 2) Vícerozměrná klasifikace - stanovení pravidel, podle nichž se zařazují objekty do jedné z několika skupin na základě měření určitého počtu znaků a vytvoření popisu těchto tříd. Zde se uplatňuje metoda diskriminační analýzy.
- 3) Typologie objektů - uspořádání, respektive hierarchické třídění do relativně stejných skupin, přičemž nemusí být znám ani jejich počet, ani nejsou přesně definovány. Současně lze určit pořadí těchto skupin podle zvolených kritérií. Pro tento úkol je velmi vhodné užití shlukové analýzy.

VÝSLEDKY A DISKUSE

Všechny výše uvedené metody lze při hodnocení rozvojového potenciálu využít. Každá z nich podá o analyzovaném souboru důležité informace. Jejich použití a zejména

následná interpretace jsou však poměrně náročné a zejména pro uživatele - nestatistiky se mohou zdát příliš komplikované. Z tohoto pohledu je vhodné, aby byla použita pouze jedna z těchto metod, která poskytne většinu žádoucích informací. Výstupy pak mohou být předány v odpovídající jednoduché formě tak, aby mohly být uživateli využity při rozhodování v otázkách regionální politiky.

Jako příklad je uvedena analýza potenciálu lidských zdrojů v regionech. Tato byla provedena pomocí analýzy hlavních komponent, shlukové analýzy a vícerozměrné regresní a korelační analýzy.

Pro analýzu potenciálu lidských zdrojů v regionech byly z dostupných dat vybrány proměnné :

- Hustota obyvatelstva (počet obyvatel na ha)
- Podíl městského obyvatelstva z celkového počtu obyvatel regionu
- Podíl obyvatel v obcích do 2000 osob z celkového počtu obyvatel regionu
- Podíl žen na celkovém počtu obyvatel regionu
- Migrace na tisíc obyvatel regionu (přistěhovalí- vystěhovalí)
- Index stáří (podíl obyvatel nad 65 let k počtu obyvatel do 15 let)
- Podíl vysokoškoláků na celkovém počtu obyvatel
- Podíl středoškoláků na celkovém počtu obyvatel
- Podíl příjemců starobních důchodů na celkovém počtu obyvatel
- Počet obyvatel na katastrální území regionu
- Míra nezaměstnanosti
- Podíl nezaměstnaných do 25 let a celkovém počtu obyvatel
- Relativní přírůstek obyvatel (narození - zemřelí + migrace k celkovému počtu obyvatel)

Výsledky regresní a korelační analýzy

Pomocí metod regresní a korelační analýzy byly zejména zkoumány vztahy mezi proměnnými a zjišťována multikolinearita. Ta byla zjištěna mezi proměnnými:

Hustota obyvatelstva $\Leftrightarrow$ počet obyvatel na katastrální území

Migrace na tisíc obyvatel $\Leftrightarrow$ relativní přírůstek obyvatel

Míra nezaměstnanosti $\Leftrightarrow$ podíl nezaměstnaných do 25 let z celkového počtu obyvatel

Podíl městského obyvatelstva $\Leftrightarrow$ podíl obyvatel v obcích do 2000 obyvatel

Druhý úkol regresní analýzy - vytvoření regresního modelu předpokládá stanovení závisle proměnné. Neměří se tedy vzájemné vztahy mezi proměnnými, ale vztahy mezi závisle proměnnou a nezávisle proměnnými. Nejde tedy o hodnocení stavu lidského potenciálu v regionu, ale o vliv lidského faktoru na zvolenou výslednou proměnnou. Jako jedna z možností byl zkoumán model postihující vliv lidského faktoru na velikost hrubého domácího produktu na obyvatele regionu. Analýza zahrnula všechny výše uvedené proměnné a použití metody Forward Selection byl vytvořen poměrně kvalitní model ($R^2 = 95,8\%$) zahrnující 5 proměnných : podíl městského obyvatelstva, podíl žen, index stáří, podíl vysokoškoláků a podíl nezaměstnaných do 25 let.

Výsledky shlukové analýzy

Metoda shlukové analýzy byla použita dvojím způsobem. V první fázi byla využita jako metoda redukční - pro výběr relevantních proměnných, ve druhé fázi pak jako shlukovací metoda, která poskytla informace o zařazení jednotlivých regionů do skupin podle úrovně stavu potenciálu lidských zdrojů v nich.

V první fázi bylo tedy přistoupeno ke shlukování proměnných. Celý soubor lze popsat pomocí tří shluků. V prvním shluku byly soustředěny proměnné mající vliv na zaměstnanost v regionu - míra nezaměstnanosti, podíl nezaměstnaných do 25 let, počet obyvatel na katastrální území, hustota obyvatelstva, migrace obyvatelstva. Druhý shluk

pak převážně charakterizuje věkovou strukturu obyvatelstva (index stáří, podíl příjemců starobních důchodů), třetí shluk pak složení obyvatelstva (podíl obyvatel v obcích do 2000 obyvatel, podíl vysokoškoláků, středoškoláků, žen, podíl městského obyvatelstva).

Druhá fáze - zařazení jednotlivých regionů do skupin s podobným chováním z hlediska lidského faktoru poskytla příliš obecné výsledky. Regiony byly rozlišeny pouze do dvou shluků, což neumožnilo žádoucí detailnější porovnání.

Výsledky analýzy hlavních komponent

Analýza hlavních komponent byla provedena pro zjednodušený model. Z původního souboru byly na základě výsledků regresní a korelační analýzy vyřazeny proměnné, u nichž byla zjištěna multikolinearita, a to :

Podíl městského obyvatelstva z celkového počtu obyvatel regionu

Migrace na tisíc obyvatel regionu (přistěhovalí- vystěhovalí)

Počet obyvatel na katastrální území regionu

Podíl nezaměstnaných do 25 let a celkovém počtu obyvatel.

Při zahrnutí zbylých 9 proměnných do analýzy hlavních komponent, byl model vysvětlen z 95,25% pomocí pěti komponent a je tedy možné pracovat pouze s těmito komponentami, které lze charakterizovat následujícím způsobem:

První komponenta vyčerpává 45,88 % celkové variability a koreluje nejsilněji s proměnnými :

Index stáří: $r = 0,443$

Podíl obyvatel v obcích do 2000 obyvatel : $r = 0,397$

Podíl středoškoláků : $r = 0,412$

Míra nezaměstnanosti: $r = -0,431$

Na základě těchto korelací lze tuto komponentu interpretovat jako ukazatel pohotové zaměstnatelnosti v regionu.

Druhá komponenta vyčerpává 29% celkové variability a koreluje nejsilněji s proměnnými :

Hustota obyvatel: $r = 0,407$

Podíl žen na celkovém počtu obyvatel : $r = 0,470$

Podíl vysokoškoláků : $r = 0,464$

Relativní přírůstek obyvatel: $r = -0,471$

Tuto komponentu lze tedy interpretovat jako ukazatel potenciálních zdrojů pracovních sil.

Třetí komponenta vyčerpává 10,85% celkové variability a s ní nejsilněji korelují proměnné :

Index stáří : $r = 0,483$

Podíl příjemců starobních důchodů : $r = -0,472$

Relativní přírůstek obyvatelstva : $r = 0,573$

Třetí komponentu lze tedy interpretovat jako ukazatel věkového složení (stáří) obyvatelstva..

Čtvrtá komponenta vyčerpává 5,57% celkového rozptylu a nejsilněji s ní korelují proměnné:

Podíl obyvatel v obcích do 2000 obyvatel : $r = 0,491$

Podíl žen: $r = 0,409$

Podíl středoškoláků: $r = -0,425$

Tuto komponentu lze tedy interpretovat jako ukazatel struktury obyvatelstva..

Pátá komponenta vyčerpává 3,92 % celkové variability a nejsilněji s ní korelují proměnné :

Podíl žen : $r = -0,055$

Podíl vysokoškoláků : $r = 0,509$

Tuto komponentu lze tedy interpretovat podobně jako komponentu čtvrtou - jako doplňující

ukazatel struktury obyvatelstva.

Vzhledem k tomu, že pátá komponenta podává doplňující informace ke čtvrté komponentě a navíc vyčerpává již jen malou část celkové variability modelu (3,97%), nebude dále uvažována.

Celkový model je tedy z 91,33% vysvětlen čtyřmi komponentami, které představují nejdůležitější faktory, které mají vliv na rozvoj lidského potenciálu v regionech :

Tab.č.1

Faktory ovlivňující lidský potenciál v regionech

Komponenta(faktor)	Váha
Zaměstnatelnost obyvatelstva	0,4588
Potenciální zdroje pracovních sil	0,2903
Věkové složení obyvatelstva	0,1085
Struktura obyvatelstva	0,0557

Pomocí této analýzy lze tedy definovat nejdůležitější faktory, které mají vliv na lidský potenciál regionu a dále je možné provést porovnání jednotlivých regionů z hlediska stavu tohoto ukazatele. Na základě vah komponent a korelačních koeficientů jednotlivých proměnných je možné vytvořit koeficient rozvojového potenciálu :

$$\text{Konstrukce ukazatele : } I_{i}^{LP} = \frac{\sum_{j=1}^n x_{ij} w_i}{\sum_{j=1}^m \sum_{i=1}^n x_{ij} w_i},$$

kde : i - počet proměnných v modelu

j - počet regionů

x_{ij} - hodnota i-té proměnné v j-tém regionu

w_i - váha i-té proměnné vypočtená z modelu analýz hlavních komponent

Koeficient umožňuje srovnání regionů, vytypování tzv. problémových regionů, kterým je třeba v této oblasti věnovat pozornost a dále pak představuje východisko k dalším rozborům situace v jednotlivých regionech.

ZÁVĚR

Všechny prezentované statistické vícerozměrné metody lze odpovídajícím způsobem pro zmíněné analýzy užít. Každá z nich představuje určitou kvantifikaci daného procesu a poskytuje relevantní informace. S ohledem na stanovený cíl a zejména uživatelskou základnu se jeví jako nejvhodnější použití jedné metody a to analýzy hlavních komponent. Tato metoda poskytuje informaci nejuniverzálnější, která kvantifikuje významnost jednotlivých proměnných - faktorů regionálního rozvoje a dále může posloužit, jak bylo naznačeno k výpočtu koeficientu, s jehož pomocí je možné provést srovnání regionů dle vytčených kritérií. Z tohoto pohledu přináší model cennou informaci pro rozhodování v regionální správě. Je však třeba si uvědomit, že samotný koeficient podává jen jednu informaci a je třeba, aby dále byl doplněn o detailní ekonomické zhodnocení. Jeho nízká výše totiž nemusí vždy informovat o nedostatečném rozvojovém potenciálu tzv. problémových regionů, ale může rovněž jít o rozvinuté regiony, které již nemají tak silnou potřebu rozvoje. Toto vše pomohou odhalit následné analýzy relevantních ukazatelů.

Uvedená metodika je flexibilní, do modelu analýzy hlavních komponent, z níž

konstrukce koeficientu vychází, mohou být zařazeny další proměnné podle toho, na co bude ve sledovaném období kladen největší důraz. Koeficient pak podá informaci o stavu využití vynaložených prostředků.

Literatura:

Boháčková, I., Hrabánková, M., Svatošová, L.: Role of agriculture in development of agricultural regions, Agricultural economics, 49, 2003(5), str. 229-232, ISSN 0139-570

Hebák, P., Hustopecký, J.: Vícerozměrné statistické metody, SNTL/ALFA 1987

Hindls, R., Hronová, S., Seger, J.: Statistika pro ekonomy, Professional Publishing, 2002

Parr Rud, O.: Data Mining, Computer Press, Praha, 2001

Svatošová, L.: Zdroje informací a možnosti analýz rozvojového potenciálu regionů, Sborník příspěvků z mezinárodní vědecké konference Agrární perspektivy XII, s. 859-865, ISBN 80-213-1056, 2003

Svatošová, L.: Analýza klíčových aspektů regionálního rozvoje, Sborník příspěvků z mezinárodní konference kateder statistiky a operačního výzkumu „Kvantitativne metody v ekonomii, metodologické a praktické aspekty výskumu v období vstupu do EÚ, Račkova dolina, září 2003, ISBN 80-8069-299-8

Kontaktní adresa :

Doc. Ing. Libuše Svatošová, CSc., Katedra statistiky, PEF ČZU Praha, e-mail: svatošova@pef.czu.cz