

刘 玮,李燕凌,胡扬名. 县域农业信息化发展水平评价[J]. 江苏农业科学,2014,42(5):399-403.

县域农业信息化发展水平评价

刘 玮,李燕凌,胡扬名

(湖南农业大学公共管理与法学学院,湖南长沙 410128)

摘要:以湖南省 78 个县(市、区)的 2012 年农业信息化服务等数据为基础,采用因子分析与聚类分析法对湖南省各县(市、区)的农业信息化发展水平进行比较研究。分析结果显示:(1)农业信息化发展水平的大部分信息能够通过“农村信息化建设主体”“农村信息化服务队伍”“农村信息化基础设施”等 7 个公因子较好地反映出来。(2)湖南省农业信息化发展水平可以划分为“偏高服务地区”“较高服务地区”“一般服务地区”和“需服务地区”4 种不同的类型。

关键词:农业信息化;因子分析;聚类分析;湖南省;县域;信息化水平

中图分类号: F323.3 **文献标志码:** A **文章编号:** 1002-1302(2014)05-0399-05

农村信息化建设是解决“三农”问题的一个重要途径,是当前和今后我国“三农”发展和新农村建设的一项重要内容,也是全球信息化形势下我国面临的一项紧迫任务。农村信息化建设是一项具有开放性、动态性和复杂性的巨型系统工程,目标具有阶段性、层次性、多样性,建设过程所处的环境极其复杂,在全国范围内要面临不同区域、不同生态环境、不同社会经济条件和不同产业特点等差异性大的建设环境,即使在一个省范围内其环境差异性也很显著,这也决定了我国农村信息化建设的长期性,建设过程所面临的问题的复杂多样性。要求我们要抓住农村信息化建设过程中所面临众多问题中的关键问题,要善于抓住信息化建设中的热点、难点和人民群众关心的主要问题。尽管农村信息化建设问题显得非常重要(这从近 11 年中央一号文件均强调农村信息化建设就可看出),但是与城市信息化建设相比,农村信息化建设更为复杂、发展也更缓慢、面临的问题更多。尤其是我国农村信息化建设还处于起步阶段,无论是人们的思想认识,还是物资、资金、技术方案、政策等方面的准备都不充分,农村信息化建设面临的困难重重,问题纷繁复杂,令人眼花缭乱。

相关学者专门研究了农村农业信息化发展的现状和影响因素。李玉霞等分析了当今农业信息化发展现状,以及农业信息网络发展的制约因素^[1]。肖惊对广东省湛江市进行了定性的半结构化调研和焦点式访谈,并在定量问卷调查的基础上,通过对调查数据的获取、整理、描述性统计、相关分析和二元 Logistic 回归,从实证角度验证了农村信息服务需求内容和需求渠道的影响因素模型^[2]。

相关学者用计量分析方法分析了影响信息化发展的因素。梁春阳等用波拉特方法对宁夏的信息化发展水平进行了客观分析评价,通过国民生产总值比重法和劳动力结构比重法得出评价宁夏信息化水平的 2 项重要指标,即信息部门所

创造的财富和收入在国民生产总值中的比重及信息劳动者占总就业人口的比例^[3]。田安意统计了 2000—2005 年世界各国信息化发展水平的面板数据,从宏观上确定了影响一个国家和地区信息化发展水平的有技术、社会经济发展水平、政府作为等因素^[4]。宋燕华等用因子分析法综合评价了农村信息化水平,认为农村信息化社会环境是影响我国农村信息化发展的关键因素^[5]。丁疆辉等选取农村信息化水平为被解释变量,利用相关分析与回归分析相结合的方法,测定各种因素对农村信息化空间作用的影响强度^[6]。在这些研究中,大多数学者使用了宏观数据对影响农业信息化发展的因素进行了分析。李德华采用波拉特法对农业信息化水平进行测度分析,指出我国农业信息化总体水平较低。他根据波拉特的“国民生产总值比重法”,得出农业信息部门的产值占农业产值的比重,从测算结果来看,农业信息部门产值占农业产值的比重呈上升趋势,但农业信息化水平总体较低^[7]。

为了深入揭示湖南省不同地区农业信息化发展水平的异同,本文以实地调研数据为基础,采用因子分析与聚类分析法进行了研究。与现有研究相比,本研究的主要特色与创新之处在于:(1)以 2013 年实地调研数据为基础,利用微观数据分析影响农业信息化发展的相关因素,比较各地信息化发展水平。(2)将因子分析与聚类分析结合起来对农业信息化发展水平进行研究。先以因子分析法从影响因素中提炼出若干公因子,再以提炼出的公因子为基础进行聚类分析。

1 农业信息化发展水平的评价模型

1.1 数据来源与指标选取

湖南省是我国中部省份,又是第二个国家农村信息化示范省建设试点省,湖南农村信息化建设在我国具有典型价值。本研究设计了问卷并于 2013 年 5 月通过湖南省科技厅农村处对全省 122 个县(市、区)中的 80 个县(市、区)科技部门就农村信息化建设情况进行了全面的摸底调查,调查 2012 年农村信息化建设的情况,回收有效问卷 78 份,有效率达 97.5%。

根据农村信息化的内涵,结合研究目的及有关专家的反馈意见和调查区域的实际情况,并遵循代表性、可比性、可操作性、科学性等原则,设置了 16 个变量,用于农村信息化发展水平的因子分析(表 1)。

收稿日期:2013-12-24

基金项目:湖南省社会科学基金(编号:2010YBB166);湖南省科技计划(编号:2012NK3068);湖南省省情决策项目(编号:2012BZZ185)。

作者简介:刘 玮(1976—),女,湖北襄阳人,博士研究生,主要从事农业科技服务与管理。Email:liuwei0901@126.com。

表 1 农村信息化建设公因子分析指标体系

控制层	指标层	标志
信息化服务资源	县域乡镇政府建立网站的比例(%)	X_1
	配备电脑的农民专业合作社的数量(个)	X_2
	进行网上交易的涉农企业的比例(%)	X_3
服务队伍	每万农村人口拥有的乡(村)级信息员数量(人)	X_4
	乡(镇)均拥有科技特派员数量(人)	X_5
	农村农业信息服务业从业人员比重(%)	X_6
	乡(镇)信息咨询专家的数量(人)	X_7
信息化服务现状	信息咨询专家人均接受农民访问(咨询)次数	X_8
	网上农产品交易额占农业总产值比重(%)	X_9
	县域涉农网站数量(个)	X_{10}
	基层信息服务站年均发布信息条数	X_{11}
基础设施	乡(镇)均有基层信息服务站数量(个)	X_{12}
	移动电话普及率(%)	X_{13}
	电脑普及率(%)	X_{14}
信息化发展社会环境	县距离中心城市(市政府所在城市)的距离(km)	X_{15}
	城镇化水平(%)	X_{16}
	年末总人口(万人)	X_{17}
	移动电话运营商信息平台服务用户数量(个)	X_{18}
	农林牧渔业总产值(万元)	X_{19}
	农村从业人员(万人)	X_{20}

农业信息化服务资源是指县域乡镇政府建立网站的比例、配备电脑的农民专业合作社的数量、进行网上交易的涉农企业的比例,这些指标可以表示出农业信息化的发展需要借助于一定的平台才可以有效进行。

农村信息化服务队伍建设包括村级信息员数量、乡(镇)均拥有科技特派员数量、农业信息服务业从业人员比重、乡(镇)均信息咨询专家的数量。科技特派员、信息咨询专家往来自于农业大专院校、农业科研院所、农技推广机构等事业单位。科技特派员、信息咨询专家、农村信息员、信息服务业从业人员直接关系到农村信息化服务状况,构成了农村信息化服务的主力军。这些指标主要反映的是数量问题,实际上农村信息化服务队伍还涉及质量、结构、发展等方面。

农村信息化服务现状方面,平均每个基层信息服务站年发布信息条数,反映的是信息发布情况。信息咨询专家人均接受农民访问(咨询)次数和网上农业产品交易额占农业总产值的比重实际上反映的是农村信息化建设效果。综合文献研究及实地座谈得知,农村信息化服务状况实际上是农村信息化建设的核心,直接关系到拿什么样的信息为农服务的问题。

农村信息基础设施是农村信息化的载体,也是农村信息化建设情况的直接表现,移动电话普及率、电脑普及率可以用来反映农村信息化基础设施建设情况。农村信息基础设施不仅包括这两方面,还有诸如广播电视网络设备建设、固定电话网络等多方面。

农村信息化发展社会环境包括移动电话运营商信息平台服务用户数量,这一指标实际上反映是移动电话运营商在农村信息化建设中的作用——提供信息服务平台。除此之外,还包括农村经济发展情况和农村人口状况,共同反映农村信息化建设的经济社会状况。也印证了某些学者所研究的信息

化与农村经济社会发展具有相关关系,信息化的推进能促进农村经济社会发展,而农村经济社会发展也能促进农村信息化建设向前发展。

1.2 农业信息化发展水平的模型

在实际研究中,为了客观、全面地分析问题,需要记录若干个反映这些问题的观测指标并考虑众多的影响因素,数据虽然可以提供丰富的信息,但同时也使得数据的分析工作非常繁杂。事实上,如果能够找到少数几个可以概括原始指标的主要信息的综合指标并对其进行多元统计分析就能使研究工作简化,又能抓住问题的关键。因子分析就为这种研究提供了可能。因子分析是多元统计分析技术的一个分支,主要目的是浓缩数据。它通过研究众多变量之间的内部依赖关系,探求观测数据中的基本结构,并用少数几个假想变量来表示基本的数据结构。这些假想变量能够反映原来众多的观测变量所代表的主要信息,并解释这些观测变量之间的相互依存关系,把这些假想变量称为基石变量,即因子。因子分析就是研究如何以最少的信息丢失把众多的观测变量浓缩为少数几个因子^[8]。因子分析通过建立相关系数矩阵,并计算出特征值,从中选择相符合的公因子,得到载荷矩阵,最后计算出公因子得分和综合得分。

1.2.1 因子分析模型 设有 m 个指标 X_1, X_2, \dots, X_m , 欲寻找可以概括这 m 个指标主要信息的综合指标 F_1, F_2, \dots, F_m , 从数学上讲,就是寻找一组常数 $a_{i1}, a_{i2}, \dots, a_{im} (i = 1, 2, \dots, m)$, 使这 m 个指标的线性组合

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1m}X_m \\ F_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2m}X_m \\ \vdots \\ F_m = a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mm}X_m \end{cases} \quad (1)$$

能概括 m 个原始指标 X_1, X_2, \dots, X_m 的主要信息,且各 $F_i (i = 1, 2, \dots, m)$ 互不相关。为叙述方便,引入如下的矩阵形式:

$$F = \begin{Bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{Bmatrix}, A = \begin{Bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{Bmatrix}, \Delta = \begin{bmatrix} a'_{11} \\ a'_{12} \\ \vdots \\ a'_{1m} \end{bmatrix}, X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix}$$

则公式(1)可表示为:

$$F = AX \quad (2)$$

或者

$$\begin{cases} F_1 = a'_{11}X \\ F_2 = a'_{12}X \\ \vdots \\ F_m = a'_{1m}X \end{cases} \quad (3)$$

如果 $F_1 = a'_{11}X$ 满足: $a'_{11}a_1 = 1$, 并且 $\text{var}(F_1) = \max_{a'a=1} \{\text{var}(a'X)\}$, 则称 F_1 是原始指标 X_1, X_2, \dots, X_m 的第一公因子。

一般来讲,如果 $F_i = a'_iX$ 满足:

(1) $a'_i a_i = 1$, 当 $i > 1$ 时, $a'_i a_j = 0 (j = 1, 2, \dots, i - 1)$;

(2) $\text{var}(F_i) = \max_{\substack{a'a=1 \\ a'_j a_j = 0 (j=1, 2, \dots, i-1)}} \{\text{var}(a'X)\}$ 。

则称 F_i 是原始指标的第 i 个公因子($i = 2, \dots, m$)。

1.2.2 因子分析法的计算步骤 (1) 对各原始数据进行标准化。为消除量纲不同带来的影响,需对原始数据进行标准

化处理。先按下式:

$$X'_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}, (j=1, 2, \dots, m)$$

将原始指标标准化,然后用标准化的数据 X'_{ij} 来计算公因子。为方便计算,仍用 X_{ij} 表示标准化后的指标数据, \mathbf{X} 为标准化后的数据矩阵,则

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix}。$$

(2) 求出 \mathbf{X} 的相关矩阵 \mathbf{R} 。 \mathbf{R} 是反映标准化后的数据之间相关关系密切程度的统计指标,值越大,说明越有必要对数据进行公因子分析。其中 $R_{ij} (i, j=1, 2, \dots, m)$ 为原始变量 X_i 和 X_j 的相关系数, \mathbf{R} 为实对称矩阵(即 $R_{ij} = R_{ji}$)。标准化后, \mathbf{X} 的相关矩阵即为协方差 $cov(\mathbf{X})$ 。

$$\mathbf{R} = cov(\mathbf{X}) = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ \vdots & \vdots & 1 & \vdots \\ r_{m1} & r_{m2} & \cdots & 1 \end{bmatrix}$$

(3) 求出相关矩阵的特征值和特征值所对应的特征向量。求公因子的问题,实际上就是要求出 \mathbf{X} 的协方差矩阵 $cov(\mathbf{X})$ 的特征值和特征向量。可由特征方程 $|\mathbf{R} - \lambda \mathbf{I}| = 0$ 求得 m 个非负特征值,将这些特征值从大到小排列为: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ 。再由

$$\begin{cases} (\mathbf{R} - \lambda \mathbf{I}) \mathbf{a}_i = \mathbf{0} \\ \mathbf{a}_i' \mathbf{a}_i = 1 \end{cases} \quad i=1, 2, \dots, m$$

解得每一特征值 λ_i 对应的单位特征向量 $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{im})'$, 列出公因子的表达式。

(4) 计算因子载荷,由此可以进一步计算公因子得分。

1.2.3 聚类分析 聚类分析是统计学所研究的物以类聚问题的一种方法,能够将一批样本数据(或变量)按照它们在性质上的亲疏程度在没有先验知识的情况下自动进行分类^[9]。根据分类对象的不同,聚类分析分为样本聚类分析(Q型)和变量聚类分析(R型)2种。Q型聚类分析使具有共同特征的样本聚集在一起,分类的依据是个案之间的“距离”。R型聚类分析使具有共同变化的特征聚集在一起,实现用较少变量刻画研究对象的目的。考察这2种方法的距离可以用系统聚类法,即将每个样品各自作为一类,并规定样品之间的距离和类与类之间的距离,然后将距离最近的2类合并成1个新类,计算新类与其他类的距离;重复进行2个最近类的合并,每次减少1类,直至所有样品合并为1类。常用方法有最短距离法、最长距离法、中间距离法、类平均法、离差平方和法等。本研究使用 Ward 提出的离差平方和聚类法(又称 Ward 法)。其基本思想是先将 n 个样本数据各自成一类,每次缩减1类,每缩减1类离差平方和就会增大;然后,再选择类内离差平方和的增量最小的2类进行合并,直到全部样本数据归类完毕^[10]。

本研究首先采用因子分析法对湖南省78个县域的农业信息化发展水平进行分析,可获得78个区域的农业信息化发展水平的综合得分与排名。因子分析只反映了78个地区农

业信息化发展水平排名的先后次序,未能对各地区进行层次划分,本研究通过Q型聚类分析获得不同县域农业信息化发展水平的层次分类。

2 数据处理与结果分析

2.1 数据的标准化和描述性分析

由于对农村农业信息化的测评研究所得数据不存在负值,所以使用无量纲化处理的方法进行数据标准化。本研究对原来的20个指标进行标准化,以消除变量在水平和量纲上的影响。

在进行正式分析之前,先对78个样本的20个变量进行统计,描述各变量的均值、标准差(表2)。

表2 各变量的描述性统计

变量	最小值	最大值	均值	标准差	离散系数
X_1	14.00	100.00	86.99	24.795 0	0.285 0
X_2	2.00	292.00	70.08	63.802 0	0.910 4
X_3	1.00	90.00	32.60	25.847 0	0.792 9
X_4	0.55	45.25	8.52	6.481 8	0.760 7
X_5	0.06	13.33	2.43	2.260 8	0.930 8
X_6	1.00	90.00	34.26	24.622 0	0.718 7
X_7	0.23	38.90	5.73	7.820 0	1.364 7
X_8	6.25	750.00	101.17	134.604 8	1.330 4
X_9	0.00	0.496	0.014	0.056 5	4.175 9
X_{10}	1.00	96.00	24.31	0.251 700	1.035 4
X_{11}	0.30	86.39	14.97	20.556 9	1.373 6
X_{12}	0.34	29.29	8.11	6.446 1	0.794 4
X_{13}	22.00	93.00	59.82	18.620 0	0.311 3
X_{14}	1.00	100.00	19.35	22.069 0	1.140 5
X_{15}	9.00	2 250	64.19	47.488 0	0.739 8
X_{16}	3.00	95.05	42.97	16.465 7	0.383 2
X_{17}	5.00	98.00	55.14	25.220 0	0.457 4
X_{18}	10 000	60 000	21 912.92	11 203.07	0.511 3
X_{19}	100 158	843 970	293 172.96	180 163.00	0.614 5
X_{20}	1.00	74.00	24.39	17.445 0	0.715 3

从表2中可以看出,各地区的发展存在较大差异,离散系数大于0.8,且离散程度较高的支出项目有配备电脑的农民专业合作社的数量、乡(镇)均拥有科技特派员数量、乡(镇)信息咨询专家的数量、信息咨询专家人均接受农民访问(咨询)次数、网上农产品交易额占农业总产值比重等项目,说明在这些因素上各地存在的差异较大;离散系数较小的有城镇化水平、年末总人口(万人)等项目,表明各地的这些因素在影响信息化发展水平方面差异相对较小。

2.2 因子分析及结果评价

由于不同县域的农业信息化发展水平的20个相关变量间可能存在严重的内部相互依赖性,难以从中清晰地看出不同地区财政支出结构的异同,而因子分析可以将具有错综复杂关系的变量综合为数量较少的几个因子^[11]。本研究先采用因子分析的方法对支出变量进行降维。

2.2.1 KMO 和 Bartlett 检验 因子分析的前提条件是变量之间存在着较强的相关性,经 SPSS 分析得出在 KMO 检验中 $KMO = 0.785$, Bartlett 的球形度检验显著性水平为 0.000,小于显著性水平 0.01,显著为单位阵,拒绝 Bartlett 检验的零假

设,代表总体的相关矩阵间有共同因素存在;并且由于 KMO 统计量在 0.7 以上,表明本研究选取的指标适合做公因子分析。

2.2.2 公因子确定及命名 通过总方差解释可以得出,前 7

个公因子的特征值均大于 1,可以解释原来 20 个指标变量 71.167%的信息量(表 3),基本保留了原始 78 个样本所反映的信息情况。选择前 7 个因子对湖南省农村信息化发展状况进行分析,做到简化多指标数据结构的目的。

表 3 特征根和累积贡献率

因子	初始特征值			提取的载荷平方和			旋转后的载荷平方和		
	总特征根	方差贡献率 (%)	累积贡献率 (%)	总特征根	方差贡献率 (%)	累积贡献率 (%)	总特征根	方差贡献率 (%)	累积贡献率 (%)
1	3.579	17.894	17.894	3.579	17.894	17.894	3.507	17.534	17.894
2	2.320	11.600	29.494	2.320	11.600	29.494	1.855	9.276	29.494
3	1.806	9.931	39.426	1.806	9.931	39.426	1.728	8.638	39.426
4	1.563	8.813	48.239	1.563	8.813	48.239	1.725	8.624	48.239
5	1.332	7.658	55.897	1.332	7.658	55.897	1.492	7.460	55.897
6	1.256	8.282	64.178	1.256	8.282	64.178	1.481	7.405	64.178
7	1.198	6.988	71.167	1.198	6.988	71.167	1.266	6.329	71.167
8	0.988	4.041	75.207						
9	0.943	4.015	79.222						
10	0.778	3.092	82.314						
11	0.751	2.753	85.067						
12	0.614	2.571	87.638						
13	0.559	2.496	90.134						
14	0.513	2.166	92.300						
15	0.445	1.725	94.024						
16	0.387	1.533	95.557						
17	0.341	1.407	96.964						
18	0.261	1.306	98.270						
19	0.218	0.990	99.260						
20	0.148	0.740	100.000						

接下来采用主成分分析法进行因子提取,并以正交旋转法进行因子旋转(表 4)。通过载荷矩阵可以得到每个因子的得分系数,当线性相关系数越大,说明该公因子对该变量的代表性就越大。

第 1 主成分所占信息量为 17.90%,主要包括县城乡镇政府建立网站的比例、配备电脑农民专业合作社的数量、进行网上交易的涉农企业的比例、移动电话运营商信息平台服务用户数量共 4 个变量。第 2 主成分包括乡(镇)均拥有科技特派员数量、乡(镇)均信息咨询专家的数量、每万农村人口拥有的乡(村)级信息员数量、农村农业信息服务业从业人员比重共 4 个变量。第 3 主成分包括移动电话普及率和电脑普及率共 2 个变量。第 4 个成分包括平均每个基层信息服务站年发布信息条数、乡(镇)均有基层信息服务站数量、县域涉农网站数量共 3 个变量。第 5 主成分包括信息咨询专家人均接受农民访问(咨询)次数、网上农产品交易额占农业总产值的比重共 2 个变量。第 6 主成分包括农林牧渔业总产值和城镇化水平共 2 个变量。第 7 主成分包括年末总人口和农村从业人员共 2 个变量。

将降维后的主成分重新命名,第 1 主成分命名为农村信息化建设主体,第 2 主成分命名为农村信息化服务队伍,第 3 主成分命名为农村信息化基础设施,第 4 主成分命名为农村信息资源开发利用,第 5 主成分命名为信息化建设运行效率,第 6 主成分命名为农村经济发展情况,第 7 主成分命名为农村人口状况。

根据公因子分析模型和因子载荷,可以得到 7 个公因子

表 4 旋转矩阵

变量	因子						
	1	2	3	4	5	6	7
X_{18}	0.666						
X_2	0.666						
X_3	0.631						
X_1	0.540						
X_7		0.602					
X_4		0.595					
X_5		0.562					
X_6		0.501					
X_{14}			0.749				
X_{13}			0.632				
X_{15}							
X_{12}				0.615			
X_{10}				0.544			
X_{11}				0.502			
X_8					0.670		
X_9					0.558		
X_{16}						0.618	
X_{19}						0.591	
X_{17}							0.669
X_{20}							0.554

注:表中因子载荷按大小进行排列,低于 0.5 以下的系数被抑制输出。

与原来 20 个变量之间的线性组合表达式如下(表达式中的 X_i 不是原始变量,而是标准化后的变量):

$$\begin{aligned}
Y_1 &= 0.540X_1 + 0.666X_2 + 0.631X_3 - 0.278X_4 - 0.202X_5 + \cdots + 0.475X_{19} + 0.345X_{20} \\
Y_2 &= 0.383X_1 - 0.122X_2 + 0.104X_3 + 0.595X_4 + 0.562X_5 + \cdots + 0.132X_{19} - 0.113X_{20} \\
Y_3 &= 0.361X_1 + 0.193X_2 + 0.345X_3 + 0.333X_4 + 0.145X_5 + \cdots - 0.055X_{19} - 0.048X_{20} \\
Y_4 &= -0.212X_1 - 0.034X_2 + 0.051X_3 - 0.140X_4 - 0.212X_5 + \cdots + 0.125X_{19} + 0.180X_{20} \\
Y_5 &= 0.385X_1 - 0.003X_2 + 0.312X_3 + 0.085X_4 + 0.005X_5 + \cdots - 0.127X_{19} - 0.179X_{20} \\
Y_6 &= -0.186X_1 - 0.108X_2 + 0.147X_3 + 0.127X_4 + 0.431X_5 + \cdots + 0.591X_{19} - 0.115X_{20} \\
Y_7 &= -0.446X_1 + 0.148X_2 + 0.233X_3 - 0.119X_4 + 0.043X_5 + \cdots - 0.133X_{19} + 0.554X_{20}
\end{aligned}$$

计算出 2012 年湖南省 78 个县(市、区)的 7 个公因子得分,7 个公因子分别从不同侧面反映了不同县域在信息化服务水平上的差异。

为了将湖南省农业信息化服务水平划分成几种类型,以

7 个公因子作为自变量进行 Ward 法聚类分析,距离测量采用欧氏距离平方法,利用系统聚类法进行聚类分析。根据聚类分析结果,可将 78 个县域划分为 4 个类型(表 5)。

表 5 湖南省信息化服务水平公因子得分及类别

地区	F ₁	F ₂	F ₃	F ₄	F ₅	F ₆	F ₇	综合得分	排名	类别
安乡	0.586 6	0.893 9	1.160 9	-0.704 3	-0.842 9	-0.429 6	-0.510 8	0.13	24	1
鼎城	0.292 0	-0.855 2	-1.098 7	0.016 3	0.546 8	-0.335 7	0.521 1	-0.10	45	2
汉寿	0.396 4	-0.568 2	-0.734 2	0.572 0	1.034 7	-0.118 9	-0.963 1	-0.01	35	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
津市	-1.342 9	-0.423 1	-0.535 3	-0.657 4	0.642 1	2.095 2	0.996 6	-0.11	49	2
桃源	1.44 12	1.040 5	-0.193 5	-0.289 9	1.377 7	1.768 4	-0.885 0	0.52	59	2
雨湖	-0.657 5	-0.642 4	0.265 60	-0.325 87	0.719 5	-1.280 21	0.076 6	-0.24	46	2
湘潭	0.893 3	-1.120 76	-0.191 0	-0.972 98	0.558 8	-0.700 7	-0.078 65	-0.10	44	1

3 讨论与结论

根据分析结果,类别 1 的农业信息化发展水平的基本特征是这些地区在第 1 公因子上得分远高于其他地区,该类型地区包括长沙、株洲、湘潭等。这些地区大多数是湖南省经济发达地区或者政府重视程度较高的地区,说明随着农业信息化工作的不断推进,政府对于农业信息化重视和投入所占比重会逐渐提升,这类地区可以称为“偏高服务”地区。类别 2 的农业信息化发展水平的基本特征是有 3 个公因子的得分较均衡,包括湖南省绝大部分地区,可称为“较高服务地区”。类别 3 的农业信息化发展水平正好处于中间状态,可称为“一般服务”地区。聚类分析中划分为第 4 类的县域,因公因子大部分是负值,表现出这类地区可能刚开始进行信息化建设或者出现了严重的建设不足情况,所以信息化发展水平受到的影响很大,可称为“需服务地区”。

农业信息化发展水平呈现出一定程度的趋同化。78 个县域的农业信息化发展水平集中在类型 2 和类型 3,说明湖南省在农业信息化建设方面,虽然政府很重视这项工作,也投入了一定的资金,但总体上而言还处于起步阶段,大部分县域在信息化建设方面都还需要进一步加强。

本研究利用因子分析和聚类分析方法,对湖南省的农业信息化发展状况做了定量分析。结果表明这些变量的绝大部分信息能够通过“农村信息化建设主体”“农村信息化服务队伍”等 7 个公因子较好反映出来,可作为评估农业信息化发展水平高低的依据。

系统聚类表明,2012 年湖南省农业信息化发展水平可以划分为“偏高服务地区”“较高服务地区”“一般服务地区”和“需服务地区”4 种不同的类型。衡量农业信息化发展水平须

兼顾多种因素,不同因素组合将直接影响着县域信息化发展状况,而因子分析和聚类分析可以为农业信息化的发展提供理论支持。

参考文献:

- [1]李玉霞,王 辉,周智男. 农业信息网络体系建设的制约因素及对策探讨[J]. 河北农业大学学报:农林教育版,2009,11(3): 376-379.
- [2]肖 惊. 农村信息服务需求影响因素研究——基于广东湛江市农村调研的实证分析[D]. 上海:复旦大学,2009.
- [3]梁春阳,史光明. 运用波拉特方法对宁夏信息化发展水平的测评[J]. 图书馆理论与实践,2000,23(1):53-57.
- [4]田安意. 信息化发展水平的宏观动力因素研究:理论和实证[J]. 经济经纬,2010(3):21-26.
- [5]宋燕华,施韶亭,由源清. 我国农村信息化发展影响因素分析[J]. 图书与情报,2012(4):138-140.
- [6]丁疆辉,刘卫东,吴建民. 农村信息化空间作用的影响因素与影响机理[J]. 地域研究与开发,2012,31(6):151-155.
- [7]李德华. 我国农业信息化水平的实证研究[J]. 情报理论与实践,2009,32(1):48-50,47.
- [8]郭志刚. 社会统计分析 SPSS 软件应用[M]. 北京:中国人民大学出版社,2009:87.
- [9]章文波,陈红艳. 实用数据统计分析及 SPSS 12.0 应用[M]. 北京:人民邮电出版社,2006.
- [10]何晓群. 多元统计分析[M]. 北京:中国人民大学出版社,2004:192.
- [11]张文彤. 统计分析高级教程[M]. 北京:高等教育出版社,2010:218.
- [12]匡小平,杨得前. 基于因子分析与聚类分析的中国地方财政支出结构的实证研究[J]. 中国行政管理,2013(1):105-110.