

تجزیه عاملی^۱ روی داده‌های حاصل از درصد فعالیتهای اقتصادی مردم در استانهای ایران^۲

راحله اصغری ابواسحق^۳

چکیده

تجزیه عاملی، یکی از مهمترین روشهای آماری چند متغیره می‌باشد. روشهای آماری بسیاری برای مطالعه بین متغیرهای وابسته و مستقل به کار برده می‌شود، اما تجزیه عاملی متفاوت از آنهاست. در واقع، تجزیه عاملی با هدف کشف متغیرهای مستقلی که عامل نامیده می‌شوند، برای یافتن مدل ارتباطی بین مجموعه‌ای از متغیرها که به ظاهر بی ارتباط هستند، به کار برده می‌شود. به عبارت دیگر هدف تجزیه عاملی پی بردن به الگوهای ساده از طریق ارتباطات بین متغیرهاست. در این مقاله قصد داریم تجزیه عاملی را بر روی داده‌های حاصل از درصد فعالیتهای اقتصادی مردم انجام داده و نتایج جالبی از آن را تفسیر نماییم. نتایج حاصل از این تجزیه عاملی، عموماً نمی‌تواند به تمامی سالها اختصاص داده شود و تفسیر تجزیه عاملی،

^۱ – Factor Analysis

^۲ – داده‌ها از نتایج طرح اشتغال و بیکاری بانک مرکزی جمهوری اسلامی ایران در سال ۱۳۸۴ اخذ شده‌اند، که فقط مناطق شهری را در بر می‌گیرد.

^۳ – محقق اداره آمار اقتصادی بانک مرکزی ج.ا.ا.

منوط به همان سال خواهد بود. برای رسیدن به این نتایج از نرم افزار آماری SPSS ۱۱.۰ استفاده شده است. جداول نیز به غیر از ۲ جدول کوچک بقیه در پیوست آورده شده‌اند.

واژگان کلیدی: تجزیه عاملی، مقادیر ویژه، عامل مشترک، واریانس اختصاصی، واریانس مشترک، چرخش، مقدار عامل

مقدمه

تجزیه عاملی، برای اولین بار، نزدیک به ۱۰۰ سال پیش توسط روان‌شناسی به نام "چارلز اسپیرمن"^۱ ابداع شد. او متوجه شد که تواناییهای فکری اعم از مهارتهای ریاضی، ادبی، شفاهی، هنری، منطقی و غیره می‌توانند تحت یک عامل هوش عمومی که آن را *g* نامید، توضیح داده شوند. او آزمون کرد که اگر بتوان *g* را اندازه‌گیری نمود و یک مجموعه از مردم را با همان *g* انتخاب کرد، بین این عامل با هیچ آزمون اندازه‌گیری توانایی فکری دیگری همبستگی وجود ندارد و به این نتیجه رسید که *g* تنها عامل مشترک بر روی اندازه‌های تواناییهای فکری است.

امروزه حداقل سه عامل مهم را در اندازه‌گیری تواناییهای فکری در نظر می‌گیرند. مهارتهای منطقی، شفاهی و ریاضی. بیشتر روان‌شناسان بر این عقیده‌اند که عوامل بسیار دیگری را در تشخیص تواناییهای فکری می‌توان شناسایی کرد.

در روش تجزیه عاملی، متغیرها به طور مستقیم در ساختار مدل ارتباطی ظاهر نمی‌شوند و همچنین تعداد عاملها که ترکیب خطی از متغیرهای اصلی هستند و ویژگی خاصی از ارتباط را نشان می‌دهند، به مراتب کمتر از تعداد متغیرهای اصلی هستند. بنابراین یکی از اهداف اصلی روش تجزیه عاملی، کاهش ابعاد داده‌هاست. فرض اساسی در به کارگیری این روش، وجود الگویی زیربنایی یا مدلی خاص در تعیین مفاهیم پیچیده ارتباطی بین

^۱ - Charls Spearman (۱۹۰۴)

متغیرهاست. این ارتباط در قالب یک عامل در این مدل فرضی ظاهر می‌شود. برای انجام تجزیه عاملی نیازی به داشتن کلیه داده‌ها نیست و تنها داشتن ماتریس همبستگی یا کوواریانس کافی است. اما در مورد مطلب ما که کاربردی خاص از تجزیه عاملی را مطرح می‌کنیم، داشتن داده‌های اصلی ضرورت دارد.

پرسشهای مطرح در روش تجزیه عاملی

عمدتاً تجزیه عاملی به چهار پرسش زیر پاسخ می‌دهد:

۱- چند عامل متفاوت برای توضیح الگوهای ارتباطی بین متغیرها مورد نیاز است؟

۲- ساختار این عاملها چیست؟

۳- چه میزان، عاملهای مفروض مشاهدات را توضیح می‌دهند؟

۴- هر متغیر مشاهده شده چه میزان از واریانس اختصاصی را در بر می‌گیرد؟

به منظور پاسخ به پرسشهای فوق، تجزیه عاملی می‌تواند تحت مدل‌های قطعی^۱ یا مدل‌های ابتکاری^۲ به کار برده شود. تمایز این دو مدل به نحوه تفسیر خروجی‌ها توسط کاربران بستگی دارد. به کاربردن مفهوم ابتکار در تجزیه عاملی می‌تواند در درک خصوصیتی ویژه از تجزیه عاملی که باعث سردرگمی مردم می‌شود، مفید واقع شود. مثلاً، چندین محقق ممکن است مجموعه‌های مشابه یا عیناً مثل هم را در تجزیه عاملی به کار ببرند و هر یک تعداد عاملهای متفاوت از دیگری را ارائه دهد: یکی ۳ تا، یکی ۶ تا، یکی ۱۰ تا. این عدم توافق درباره عاملها در اندیشه مردم بر اعتبار تجزیه عاملی می‌تواند تاثیرگذار باشد. اما این معضل به این صورت قابل حل است که وقتی تجزیه عاملی که توسط چند نفر ارائه شده اند، ظاهری متفاوت از یکدیگر در مدل‌بندی داده‌ها داشته باشند، در صورتی نتایج آنها از تجزیه عاملی

^۱ - Absolute

^۲ - Heuristic

یکدیگر را نقض می‌کنند که محققان ادعا کنند که نظریات قطعی را ارائه کرده‌اند و نه یک نظریه ابتکاری. عاملهای کمتر تنها نظریه را راحت‌تر می‌کنند و عاملهای بیشتر بیان نظریه را مشکل‌تر می‌کنند، اما داده‌ها را بهتر توضیح می‌دهند و مدل بهتری در جهت برازش داده‌ها ارائه می‌کنند. در موازنه بین راحتی و برازش بهتر هر کاربری ممکن است بسته به نوع کار و اهمیت کار نظر متفاوتی را ارائه بدهد. عموماً تفسیر عاملهای بیشتر، تفسیر عاملهای کمتر را تایید می‌کنند و به عبارتی در توضیح داده‌ها ریزبین‌تر می‌شوند.

تاریخچه

اگر روش آماری باشد که تاریخچه آشفته‌ای داشته باشد، قطعاً تجزیه عاملی همان روش است. حدود سال ۱۹۵۰ اعتبار تجزیه عاملی دستخوش مشتاقان افراطی این روش شد که به غلط از آن بهره می‌جستند. با نگاهی به گذشته، سه مطلب اشتباه در تفکرات مردم آن زمان نسبت به تجزیه عاملی دیده می‌شود که در اینجا آنها را بیان می‌کنیم. نخست آنکه، گروهی از مردم تجزیه عاملی را یک روش آماری ویژه و مشخص می‌پنداشتند تا یک خانواده از روشهای آماری. دوم آنکه مسائلی که باید به‌گونه‌ای ابتکاری تفسیر می‌شد، کاملاً قطعی تفسیر می‌شد. سوم آنکه، آنها می‌خواستند به وسیله تجزیه عاملی، کلیه متغیرهای موثر بر روی عاملها را در مدل خود پوشش دهند، در حالی که چنین چیزی همیشه عملی نیست. لذا آنها به سه طریق یاد شده، به استفاده از تجزیه عاملی بیش از حد قابلیت و تواناییهایش مبادرت می‌کردند. به نظر می‌آید در دهه‌های اخیر، تجزیه عاملی جایگاه واقعی خود را به عنوان خانواده‌ای از روشها که برای اهداف محدود و معینی به کار برده می‌شوند پیدا کرده باشد.

اجرای تئوری تجزیه عاملی^۱

در تجزیه عاملی ۴ مرحله جهت تعیین الگوهای ارتباطی بین متغیرها اجرا می‌شود:

^۱ - تئوری تجزیه عاملی در پیوست توضیح داده شده است.

- ۱- ماتریس ضرایب همبستگی بین متغیرها را محاسبه نموده، متغیرهایی را که به نظر می‌رسند با سایر متغیرها وابستگی ضعیفی دارند، معین می‌کنیم.
- ۲- دومین مرحله ساختن عاملها می‌باشد. این مرحله شامل تعیین تعداد و روش محاسبه عاملهاست. میزان برازش مدل به داده‌ها نیز در این مرحله محاسبه می‌شود.
- ۳- دوران و اعمال تبدیلاتی خاص روی عاملها برای آن که روابط میان داده‌ها را بهتر تفسیر کنند، در این مرحله صورت می‌گیرد.
- ۴- امتیاز هر عامل برای هر مشاهده محاسبه می‌شود. این امتیازات می‌توانند مبنای تحلیل‌های متنوعی قرار گیرند.

تحلیل داده‌ها

در اینجا قصد داریم روش تحلیل عاملی را به ترتیب جزئیاتی که در صفحات پیشین ارائه شد، بر روی داده‌های مربوط به درصد نیروی کار ۱۴ بخش مختلف فعالیت‌های اقتصادی در ۳۰ استان کشور ایران پیاده کنیم و نتایج جالبی به دست آوریم. این داده‌ها در جدول ۱ آمده‌اند. و ۱۴ بخش عبارتند از:

- ۱- کشاورزی ۲- معدن ۳- صنعت ۴- انرژی ۵- ساختمان ۶- بازرگانی ۷- هتل
- ۸- ارتباطات ۹- واسطه‌گریهای مالی ۱۰- املاک و مستغلات ۱۱- امور عمومی ۱۲- آموزش
- ۱۳- بهداشت ۱۴- سایر

جدول ۲، میانگین درصد اشتغال در هر فعالیت را در کلیه استانها و میزان پراکندگی آنها در کلیه استانها را نشان می‌دهد. بر اساس این داده‌ها ملاحظه می‌شود که میانگین درصد اشتغال در زمینه بازرگانی و صنعت از میانگین درصد اشتغال در سایر زمینه‌های فعالیت‌های اقتصادی مطرح شده بیشتر می‌باشد.

از طرفی بر اساس جدول ۲، میانگین درصد اشتغال در استانها در بخش صنعت دارای

پراکندگی بیشتری نسبت به سایر فعالیت‌های اقتصادی است.

بعد از اشاره‌ای اجمالی به این آمار توصیفی، برای اجرای تجزیه عاملی ابتدا ماتریس ضرایب همبستگی بین متغیرها را محاسبه می‌کنیم، تا متغیرهایی که با سایر متغیرها، همبستگی ضعیف‌تری دارند مشخص شوند. هرچه ضرایب همبستگی به یک نزدیک‌تر باشند مطلوبیت انجام تجزیه عاملی نیز بیشتر خواهد بود.

برای مطالعه نحوه محاسبه ماتریس همبستگی به پیوست مراجعه شود.

مقدار دترمینان این ماتریس همبستگی معادل $10^{-6} \times 1/244$ به دست می‌آید که بسیار نزدیک به صفر است، که نشان‌دهنده ارتباطات قوی بین لاقل برخی از متغیرها است که انجام تجزیه عاملی را مطلوب‌تر می‌سازد.

در تایید این گفته نتیجه آزمون بارتلت که در جدول ۴ آمده است، فرض صفر "برابری ماتریس واحد با ماتریس ضرایب همبستگی" را آزمون می‌کند، این فرض صفر را رد کرده و در هر سطح معنی‌داری فرض استقلال متغیرها از یکدیگر را مردود اعلام می‌کند و بیان می‌کند لاقل یک همبستگی معنی‌دار بین این متغیرها یافت می‌شود.

جدول ۴- آزمون بارتلت برابری ماتریس ضرایب همبستگی با ماتریس واحد

تقریب آماره χ^2	۳۱۹/۵۴۳
درجه آزادی	۹۱
سطح معنی‌داری	۰/۰۰۰

در جدول ۳، همبستگی‌های معنی‌دار با مربعات توپر نشان داده شده‌اند.

از این جدول ملاحظه می‌شود که ضعیف‌ترین همبستگی‌ها، به همبستگی‌های بین متغیر کشاورزی با سایر متغیرها مربوط می‌شود که تقریباً با هیچ یک از متغیرها در سطح معنی‌داری ۰/۰۵ همبستگی معنی‌دار ندارد. از جمله متغیرهایی که همبستگی‌های بیشتر و

بهتری با سایر متغیرها دارند می‌توان از صنعت، انرژی، ساختمان و ارتباطات نام برد. در اینجا برای تعیین متغیرهایی که خروج آنها از مدل تجزیه عامل به روند بهتر شدن کیفیت تجزیه عاملی کمک می‌کند، از ماتریس Anti-image استفاده می‌شود. این ماتریس، ماتریسی است که عناصر آن ضرایب همبستگی جزئی^۱ با علامت مخالف می‌باشند که عناصر روی قطر این ماتریس، مقادیر اندازه دقت نمونه‌گیری (Measure Of Sampling Adequacy) با نماد MSA نامیده می‌شوند. این معیار که برای هر متغیر جداگانه تحلیل می‌شود، بیان می‌کند که آیا متغیر خاصی از جهت وارد شدن به مدل تجزیه عاملی صلاحیت دارد یا نه؟ مقادیر بزرگ MSA در جهت تایید ورود متغیر به مدل حرکت می‌کنند. برای ملاحظه فرمول معیار MSA، به پیوست مراجعه کنید.

در اینجا با استفاده از ماتریس Anti-image متغیرها را به ترتیب اولویت از مدل خارج کرده و نتایج را در جدول ۵ ذخیره نموده‌ایم، که به تفسیر این جدول به تفصیل خواهیم پرداخت.

در مرحله اول، چنانکه از جدول پیداست هیچ تغییری از مدل حذف نشده است. از ماتریس ضرایب همبستگی در سطح معنی‌داری ۰/۰۵ ملاحظه شد که فعالیت کشاورزی با هیچ یک از سایر فعالیتها همبستگی ندارد. طبق جدول ۵، وقتی این متغیر در مدل حضور داشته باشد،

آماره MSA مربوط به آن ۰/۰۲۳ برآورد می‌شود که کمترین MSA را شامل می‌شود به همین دو دلیل، مطلوبیت چندانی برای حضور در مدل تجزیه عاملی برای این متغیر وجود ندارد. این نتیجه را آماره KMO، که در حالتی که همه متغیرها در مدل باشند (یعنی مرحله اول)، ۰/۰۵۲ برآورد شده است نیز تایید می‌کند. برای ملاحظه فرمول معیار KMO، به

^۱ - برای ملاحظه جزئیات ضریب همبستگی جزئی به بخش اول جلد دوم منبع [۲] مراجعه شود.

پیوست مراجعه کنید.

با توضیحات فوق، متغیر کشاورزی را از مدل خارج می‌کنیم و طبق مرحله ۲، جدول ۵، ملاحظه می‌شود که $KMO = 0/181$ افزایش قابل توجهی پیدا می‌کند. همین طور تمامی MSAها، مربوط به سایر متغیرهای باقیمانده در مدل نیز افزایش قابل توجهی دارند. در اینجا، برای به دست آوردن ضرایب عاملهای مشترک از روش مولفه‌های اصلی استفاده شده است که در پیوست توضیح داده شده است. طبق این روش تعداد عاملهای مشترک، تعداد مقادیر ویژه بزرگتر از یک، در نظر گرفته شده‌اند. د. بر اساس جدول ۵، با حذف متغیر کشاورزی، مدل با ۵ عامل تشکیل می‌شود که کل واریانس توضیح داده شده توسط این ۵ عامل مشترک ۷۳/۹۳٪ خواهد بود. با این وجود هنوز آماره KMO به مطلوبیت لازم نرسیده است. لذا به نظر می‌آید که با حذف متغیر دیگری به مطلوبیت بیشتری جهت انجام تجزیه عاملی برسیم.

با مراجعه به MSAها در مدل اخیر که متغیر کشاورزی حذف شده است (مرحله ۲)، ملاحظه می‌شود که متغیر معدن کمترین MSA را داراست، از طرفی از لحاظ سطح همبستگی نیز فقط با متغیر بازرگانی همبستگی معنی دار دارد. از طرفی افزایش KMO، پس از حذف این متغیر به میزان قابل توجه، حذف این متغیر را از مدل تایید می‌کند که نتایج در مرحله ۳، جدول ۵ آمده است و نشان می‌دهد که MSAها بازهم افزایش قابل ملاحظه‌ای یافتند. در این مرحله مدل با ۴ عامل مشترک ادامه می‌یابد که کل واریانس توضیح داده شده توسط این عاملها ۶۸/۹۳٪ خواهد بود. اما آماره KMO که در این مرحله ۰/۳۳۶ برآورد شده است، هنوز به مطلوبیت ویژه نرسیده است. لذا سعی می‌کنیم با حذف متغیر دیگری از مدل، KMO را افزایش دهیم.

بر اساس مرحله ۳، جدول ۵، متغیر آموزش با $MSA = 0/215$ که کمترین MSA نسبت به سایر متغیرهای باقیمانده در مدل را داراست، طی مرحله ۴، حذف می‌شود. با حذف

این متغیر، طبق مرحله ۴، باز هم $KMO = 0/486$ افزایش یافته، همین‌طور MSA های مربوط به متغیرها نیز افزایش می‌یابند و مدل با ۴ عامل ادامه یافته و کل واریانس تبیین شده توسط عاملهای منتخب به ۶۹/۷۱ افزایش می‌یابد. به همین ترتیب متغیرها را بر اساس آماره MSA از مدل حذف کرده و نتایج را در جدول ۵ آورده‌ایم.

این جدول پایه‌ای خواهد بود برای اینکه در چه مرحله، از حذف متغیرها و خروج آنها از مدل متوقف شویم و ادامه آن تا مرحله ۱۱، بدان معنی نیست که کلیه این ۱۱ متغیر از مدل ما حذف شده باشند، در واقع نکته جالب و در خور توجه این است که ما اکنون قصد داریم خانواده‌ای از مدل‌ها را برای رسیدن به نتایج و تفسیر آنها دنبال کنیم. خواهیم گفت که مدل را می‌توانیم براساس مرحله مثلاً ۸ یا مثلاً ۹، تشکیل دهیم و آنچه در این میان متفاوت خواهد بود، مطلوبیت انجام تجزیه عاملی در این مدل‌ها، روایی تفسیر و اعتبار نتایج بر اساس میزان مطلوبیت، واریانس تبیین شده توسط عاملها و ... خواهد بود. در هر مرحله که قرار بگیریم و مدل را تشکیل دهیم، عموماً بسته به میزان مطلوبیت و برخی فاکتورهای دیگر نتایج مرحله بعدی را تایید خواهد کرد. در این مرحله معمولاً تلاش بر این است که ویژگی عاملها مشخص گردند. این امر معمولاً براساس ضرایب عاملهای چرخش یافته صورت می‌گیرد، لذا لازم است توضیح مختصری در مورد چرخش عامل ارائه دهیم.

چرخش عاملها:

چون در بسیاری از موارد تعدادی از متغیرها به یک عامل ویژه یا حتی به تعدادی از عاملها بستگی دارند، تعبیر عوامل مشکل خواهد بود. لذا روش‌هایی ابداع شده‌اند که بدون تغییر میزانهای اشتراک، باعث تعبیر ساده‌تر عوامل می‌شوند. این روش‌ها، همان دوران عاملها هستند و بر دو نوع دوران عمود و دوران مایل تقسیم می‌شوند. دورانهای عمود استقلال میان عاملها را حفظ می‌کنند، اما دورانهای مایل استقلال عاملها را از بین برده و آنها را به هم وابسته

می‌کنند. این چرخش‌ها در زیر نام برده شده‌اند:

- واریماکس (Varimax)

- ایکواماکس (Equamax)

- کوارتیماکس (Quartimax) و ...

چرخش واریماکس به عنوان روش استاندارد توصیه می‌شود. ما نیز در این تحقیق از این روش بهره گرفته‌ایم. جزئیات این روش در پیوست آمده است. در مورد عملکرد سایر روش‌ها به منبع [۲] مراجعه شود.

جدول ۶، گویای ویژگی‌های عامل‌های دوران یافته و میزان واریانس تبیین شده توسط هر یک از آنها، طی مراحل است که شرح این مراحل در جدول ۵ داده شده است. در جدول ۶ اثری از مرحله ۱۱ که در جدول ۵ آورده شده بود به چشم نمی‌خورد. علت آن است که با حذف متغیر انرژی هر چند آماده KMO افزایش می‌یابد، اما تعداد عاملها به یک کاهش یافته، از طرفی واریانسی از مدل که توسط این عامل توضیح داده می‌شود. نسبت به مرحله ۱۰، شدیداً کاهش می‌یابد. از طرفی این متغیر با تعداد زیادی متغیر همبستگی دارد. به این دلایل، از حذف متغیر انرژی صرف‌نظر کرده و در این مرحله از حذف برای بررسی نتایج مراحل یادشده متوقف شدیم. توضیحات جدول ۶ را براساس مرحله ۹، که براساس آماره KMO برای انجام تجزیه عاملی، مطلوبیت متعادلی دارد، شروع می‌کنیم. نتایج ارائه شده در جدول ۷، ضرایب عامل مدلی است که براساس تجزیه به مولفه‌های اصلی و روش واریماکس مطابق آنچه قبلاً توضیح داده شد، به دست آمده است. چنانکه ملاحظه می‌شود و طبق آنچه در جداول ۵ و ۶ آورده‌ایم، این مدل شامل ۲ عامل مشترک و ۶ متغیر است. ضرایب عامل‌های به دست آمده برای هر متغیر، در واقع میزان توضیح هر متغیر به وسیله آن عامل است. زیر ضرایب عامل بزرگتر از ۰/۵ (صرف‌نظر از علامت) خط کشیده شده است.

چنانکه ملاحظه می‌شود عامل یک، ضرایب مثبت بزرگ برای متغیرهای ارتباطات، واسطه‌گریهای مالی و امور عمومی و ضرایب منفی بزرگ برای صنعت را داراست. پس عامل اول، میزان اشتغال در بخشهای ارتباطات، واسطه‌گریهای مالی و امور عمومی را در مقایسه با صنعت اندازه‌گیری می‌کند. در نتیجه این عامل می‌تواند به وسیله تاکید بر بخشهای ارتباطات، واسطه‌گریهای مالی و امور عمومی و فقدان صنعت مشخص شود. عامل دوم، ضرایب مثبت بزرگ برای ساختمان، انرژی و ضرایب منفی بزرگ برای صنعت را دارا است. بنابراین تاکید عامل دوم، بر بخشهای ساختمان و انرژی بیشتر از صنعت می‌باشد.

اینها همان نتایجی هستند که در جدول ۶ خلاصه شده اند و سایر مراحل نیز به شیوه مرحله ۹ تهیه گردیده اند.

جدول ۷- ضرایب عاملها برای مرحله نهم از جدول ۶

متغیر	ضریب عاملها	
	اول	دوم
صنعت	-۰/۶۹۷	-۰/۵۱۷
انرژی	-۰/۱۴۳	-۰/۷۸۴
ساختمان	-۰/۰۰۸	-۰/۸۳۵
ارتباطات	۰/۶۵۴	-۰/۳۴۱
واسطه‌گری مالی	۰/۷۶۸	-۰/۲۲۱
امور عمومی	۰/۶۹۸	۰/۰۸۸

براساس جدول ۶، حدود ۳۳/۵۴٪ واریانس توسط عامل اول و ۲۹/۲۰٪ واریانس توسط عامل دوم تبیین می‌شود، که مجموعاً حدود ۶۲/۷۴٪ واریانس کل توسط این دو عامل مشترک تبیین می‌شود. طبق جدول ۶، مطلوبیت تحلیل عاملی تا مرحله ۳ نامناسب، مرحله ۴

ناچیز، مرحله ۵ تا ۸ متوسط و مراحل ۹ و ۱۰، متعادل می‌باشد. اما همان‌طور که گفتیم، معمولاً نتایج هر مرحله، نتایج مراحل بعد را در دل خود گنجانده است. تنها چیزی که در اینجا باید در نظر گرفته شود، این است که میزان اطمینان به نتایجی که در مراحل دیگر گنجانده شده‌اند با تکرار در هر مرحله بالاتر می‌رود. مثلاً با مروری بر مراحل ۹ و ۱۰ ملاحظه می‌شود که ویژگی عاملهای مرحله ۱۰ در دل ویژگی عاملهای مرحله ۹ گنجانده شده‌اند و با دقتی در این جدول طی مراحل مختلف به نتایج جالبی از ویژگی عاملها می‌رسیم، اما همان‌طور که از علم آمار می‌دانیم، همیشه احتمال در بیان نتایج و تفاسیر درصدی را به خود اختصاص می‌دهد و در کنار هر بیان یا عبارتی، معیارهایی را جهت هرچه بیشتر اطمینان کردن به آن مطلب در کنار آن عبارت یا فرض قرار می‌دهیم. در اینجا نیز ما، برای بیان مطلب معیارهایی مانند مطلوبیت، میزان واریانس تبیین‌شده توسط هر عامل، میزان واریانس اختصاصی هر متغیر و ... را در کنار نتایجی که اعلام می‌کنیم، قرار می‌دهیم.

روشهای محاسبه مقدار عامل برای هر نمونه

بر اساس امتیازات عامل هر مرحله، مقدار عامل برای هر نمونه که در اینجا نمونه‌های ما، استانهای کشور ایران بوده‌اند، به چند طریق قابل محاسبه می‌باشند که خود آن مقادیر، پایه‌ای برای نتایج جالب دیگری خواهند بود. که ذیلاً به آن می‌پردازیم:

این روشها عبارتند از:

- رگرسیونی (Regression)

- بارتلت (Bartlett)

- اندرسن - رایبین (Anderson- Rubin)

که ما در این تحقیق از روش معمول رگرسیونی استفاده نموده‌ایم. جزئیات این روش در

پیوست آمده است. در مورد عملکرد سایر روشها به منبع شماره [۲] مراجعه شود.

مقادیر عاملها برای مرحله نهم

از آنجا که در صفحات پیشین ضرایب عاملها را برای مرحله نهم شرح دادیم، در ادامه نیز مقادیر عاملها را برای مرحله نهم، در قالب نتیجه‌گیری شرح می‌دهیم. و چون همان‌طور که از جدول ۵ به خاطر می‌آوریم، مدل ارائه شده در مرحله نهم دارای این خصوصیات بود که:

- ۱- از لحاظ مطلوبیت در سطح متعادل قرار داشت.
- ۲- حدوداً ۶۳٪ از واریانس کل، توسط دو عامل مرحله نهم که شرح آن رفت، توضیح داده می‌شد.

لذا نتایج اخذ شده برای مرحله نهم در سطح متعادلی از اعتبار قرار دارد.

نتیجه‌گیری

هم‌چنان‌که قبلاً توضیح دادیم تحلیل عاملی روشی آماری است که بین مجموعه‌ای از متغیرهای به ظاهر بی‌ارتباط، رابطه خاصی برقرار می‌کند. که یکی از اهداف آن کاهش ابعاد داده‌هاست. که در این مقاله همان‌طور که دیدیم تعداد متغیرها می‌تواند به ۲ عامل کاهش یابد و البته تعداد عاملهایی که ما در نظر می‌گیریم به معیارهایی مثل مطلوبیت و درصد واریانس توضیح داده شده ارتباط دارد. با نگاهی به مراحل طی شده برای انجام این روش آماری و دقیق شدن در جداول به‌دست آمده به نتایج جالبی می‌رسیم. از آن جمله نتایجی است که در جدول ۹ آمده‌اند. یعنی مقایسه استان‌ها از لحاظ میزان اشتغال افراد در بخش‌های مختلف اقتصادی. بررسی تک‌تک سلولهای این جدول با توجه به میزان اعتبار هر مرحله صورت می‌گیرد که خود می‌تواند بسیار مفصل باشد. لذا برای روشن شدن نحوه نتیجه‌گیری به تفسیر یکی از مراحل آن، یعنی مرحله ۹ که همان‌طور که در بند ۸ گفته شد از لحاظ مطلوبیت در سطح متعادلی قرار دارد می‌پردازیم. در مرحله نهم عامل اول بر بخشهای ارتباطات، واسطه‌گریهای مالی و امور عمومی در مقابل صنعت تأکید دارد و عامل دوم، بر بخشهای ساختمان و انرژی در مقابل

صنعت تأکید دارد.

جدول ۸، مقادیر عاملها برای مرحله نهم را نشان می‌دهند که به وسیله روش رگرسیونی که شرح آن رفت به دست آمده‌اند. هرچه مقادیر عاملها بزرگتر باشند (با علامت+) تاکید آن عامل در آن استان بیشتر خواهد بود و هرچه مقادیر عامل از لحاظ قدرمطلق بزرگتر باشند (با علامت-)، تاکید عکس آن عامل در آن استان بیشتر خواهد بود. به عنوان مثال برای عامل اول، در استانهای بوشهر، کهگیلویه و بویراحمد و هرمزگان مقادیر عاملهای بزرگتری به دست آمده‌اند. این نشان می‌دهد که در این استانها میزان اشتغال افراد در بخشهای ارتباطات، واسطه‌گریهای مالی و امور عمومی بیشتر از بخش صنعت می‌باشد، که این خود نشان دهنده اهمیت بخشهای فوق در قبال صنعت است. این استانها در جدول ۹ در هر مرحله تحت عنوان گروه اول معرفی شده‌اند. در استانهای آذربایجان شرقی، اصفهان، قم، مرکزی و یزد، مقادیر عامل اول اعداد منفی بزرگی را از لحاظ قدرمطلق شامل می‌شود. این نشان می‌دهد که در این استانها بخش صنعت نسبت به بخشهای ارتباطات، واسطه‌گریهای مالی و امور عمومی از اهمیت بیشتری برخوردار است. این استانها در جدول ۹، در هر مرحله تحت عنوان گروه دوم معرفی شده‌اند و اما به همین شیوه، از مطالعه مقادیر عامل دوم، در استانها ملاحظه می‌شود که میزان اشتغال افراد در بخشهای ساختمان و انرژی در استانهای اردبیل، ایلام، خراسان شمالی، خوزستان، سیستان و بلوچستان، فارس و هرمزگان بیشتر از بخش صنعت می‌باشد و این خود تاکید است بر بخشهای ساختمان و انرژی در مقابل صنعت در این استانها. همچنین در استانهای بوشهر، تهران، خراسان جنوبی، سمنان، قزوین و قم بخش صنعت نسبت به بخشهای ساختمان و انرژی از اهمیت بیشتری برخوردار است.

در مورد سایر مراحل جدول ۹ تفسیر نتایج به شیوه فوق‌الذکر شده است. بدون توجه به مناسب بودن یا نبودن تحلیل عاملی در هر مرحله، تمامی نتایج در جدول ۹ آورده شده‌اند.

جدول ۹ در واقع، خلاصه‌شده کلیه نتایج است. مشخصات عاملها در هر مرحله از این جدول، در جدول ۶ خلاصه شده است. لذا توجه شود که مشخصات عامل اول در هر مرحله لزوماً مشخصات عامل اول مراحل بعدی نیست. این جدول پایه خوبی است برای تحلیل نتایج هر مرحله، آن چه مهم است، این است که با مطالعه این جدول می‌توان به نگرش خوبی نسبت به عملکرد عاملها طی مراحل مختلف در استانهای کشور رسید و می‌تواند پایه‌ای برای برنامه‌ریزی در بخشهای مختلف اقتصادی در استانهای کشور باشد، چرا که نقاط قوت و ضعف استانها در زمینه‌های مختلف فعالیت‌های اقتصادی با توجه به این تحلیل آماری در مقایسه با یکدیگر مشخص شده است. البته توجه می‌کنیم که نتایج اخذ شده برای هر سال در همان مقطع زمانی مطرح می‌شود و ممکن است بسته به برنامه‌ریزی‌های مختلف شغلی در استانها طی سالهای آتی تغییراتی ایجاد گردد.

منابع و ماخذ

الف - منابع فارسی

۱- بی. اف. جی. مانلی، ترجمه: دکتر محمد مقدم، مهندس سید ابوالقاسم محمدی شوطی، مهندس مصطفی آقائی سربرزه، آشنایی با روشهای آماری چندمتغیره، ۱۳۷۳، تبریز، انتشارات پیشتاز علم

۲- شرکت آمارپردازان، اکبر گلدسته، راحله اصغری، محمود ترابی، SPSS/Win راهنمای کاربران، ۱۳۷۷- تهران ناشر: مرکز فرهنگی انتشاراتی حامی

ب- منابع انگلیسی

۱- Darlington, Richard. "Factor Analysis".

<http://compu⁹.psych.cornell.edu/Dalington/factor.htm> Accessed
۳۰ January ۲۰۰۲

۲. Gorsuch, Richard L. (۱۹۸۳) Factor Analysis. Hillsdale, NJ: Erlbaum

پیوست شماره (۱)

تئوری تجزیه عاملی

اگر مدل تجزیه عاملی به صورت زیر فرض شود :

$$X_i = \mu_i + \sum_{j=1}^p \lambda_{ij} f_j + e_i \quad i = 1, \dots, m \quad (1)$$

که در آن X_i متغیر i ام، μ_i میانگین متغیر i ام بر روی کلیه مشاهدات، f_j عامل j ام، e_i خطای مدل در رابطه با متغیر i ام و λ_{ij} ضریب عامل j ام در ارتباط با متغیر i ام یا در حقیقت کوواریانس بین متغیر i ام و عامل j ام است، اما اگر λ_{ij} ها با استفاده از ماتریس ضرایب همبستگی بین متغیرها یا ماتریس واریانس - کوواریانس متغیرهای استاندارد شده به دست آیند، آنگاه λ_{ij} ها ضریب همبستگی بین متغیر i ام و عامل j ام خواهند بود. یعنی :

$$\lambda_{ij} = \text{corr} \left(\frac{X_i - \hat{\mu}_i}{\hat{\sigma}_{ii}}, f_j \right) \quad i=1, \dots, m \quad j=1, \dots, p \quad (2)$$

که در آن σ_{ii} ، جذر عنصر i ام قطری ماتریس واریانس - کوواریانس متغیرهاست و f_j عامل مشترک i ام می‌باشد که دارای میانگین صفر و واریانس یک است. آن گاه :

$$\text{Var}(X_i) = \sum_{j=1}^p \lambda_{ij}^2 + \text{Var}(e_i) = h_i^2 + \tau_i^2 \quad i=1, \dots, m \quad (3)$$

h_i^2 میزان اشتراک متغیر X_i یا بخشی از واریانس X_i است که به عاملهای مشترک مربوط می‌شود. ولی τ_i^2 میزان عدم اشتراک متغیر X_i ، یا بخشی از واریانس است که به عاملهای مشترک ارتباط ندارد. m تعداد متغیرها و p تعداد عاملها را شامل می‌شود و $p < m$.

برای برآورد اولیه میزان اشتراک، یعنی، \hat{h}_i^r ها روشهای متفاوتی وجود دارند که در زیر به برخی از آنها اشاره می‌شود:

- مولفه های اصلی (Principle Component)

- حداکثر درستنمایی (Maximum Likelihood)

- کمترین مربعات موزون (Unweighted Least Squares)

در این تحقیق از روش مؤلفه‌های اصلی برای به دست آوردن مقادیر ویژه و میزانهای اشتراک استفاده شده است. لذا مختصراً آن را توضیح می‌دهیم. در روش تجزیه به مؤلفه‌های اصلی، برآورد ماتریس ضرایب همبستگی (R) را به دست می‌آوریم. با محاسبه مقادیر ویژه این ماتریس، p مقدار بزرگتر از یک را در نظر می‌گیریم:

در حالت کلی تعیین تعداد عاملها (p)، به فرد تجزیه‌کننده آن عاملها بستگی دارد، هرچند که گاهی می‌توان آنها را به وسیله ماهیت داده‌ها نیز مشخص کرد. زمانی که تجزیه به مؤلفه‌های اصلی استفاده می‌شود، یک روش تقریبی این خواهد بود که تعداد عاملها مساوی با تعداد مقادیر ویژه بزرگتر از یک در نظر گرفته شود. همان‌طور که گفتیم در این روش، مقادیر ویژه از ماتریس همبستگی متغیرها به دست می‌آیند و عاملی که دارای مقدار ویژه کمتر از یک باشد، تغییرات کمتری از کل داده‌ها را در مقایسه با یکی از متغیرها توضیح می‌دهد. به طور کلی افزایش تعداد عاملهای مشترک، میزان اشتراک متغیرها را افزایش خواهد داد.

برای این p مقدار، بردارهای ویژه را از رابطه:

$$(R - \lambda_i I) \mathcal{V}_i = 0 \quad i=1, \dots, p \quad (4)$$

محاسبه می‌کنیم که در آن λ_i مقدار ویژه i ام و \mathcal{V}_i بردارهای ویژه متناظر با آنها هستند. و R برآورد ماتریس ضرایب همبستگی می‌باشد. در این صورت برآورد ماتریس ضرایب

عامل برای p عامل اول، یعنی L برابر است با :

$$L = \left\{ \sqrt{\hat{\lambda}_1} \hat{\gamma}_1, \dots, \sqrt{\hat{\lambda}_i} \hat{\gamma}_i, \dots, \sqrt{\hat{\lambda}_p} \hat{\gamma}_p \right\} \quad (5)$$

برای ملاحظه عملکرد سایر روشها می‌توانید به مرجع [۲] مراجعه نمایید.

همبستگی بین X_i و X_j در این مدل برابر است با:

$$r_{ij} = \sum_{k=1}^p \lambda_{ik} \lambda_{jk} \quad (6)$$

که در آن λ_{ij} ضریب عامل j ام در متغیر i ام می‌باشد.

لذا دو متغیر X_i و X_j زمانی شدیداً همبسته‌اند که دارای ضرایب عامل بزرگ باشند.

چون h_i^2 ، یعنی میزان اشتراک حداکثر یک است. بنابراین $|\lambda_{ij}| \leq 1$ خواهد بود.

ضرایب همبستگی

نحوه محاسبه ضرایب همبستگی به شرح زیر است :

اگر X_1, \dots, X_n نمونه‌ای از توزیع چند متغیره $N_p(\underline{\mu}, \Omega)$ باشد، در این صورت

ماتریس واریانس - کوواریانس نمونه‌ای برابر خواهد بود با :

$$S = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \quad (7)$$

عنصر i ام قطری این ماتریس S_{ii} نامیده شده، لذا همبستگی بین مولفه i ام و j ام

برابر است با :

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}} \sqrt{S_{jj}}} \quad \forall i, j = 1, \dots, m \quad (8)$$

معیار MSA

معیار **MSA**، به کمک ضرایب همبستگی و ضرایب همبستگی جزئی و با رابطه ذیل

محاسبه می‌شود:

$$MSA_i = \frac{\sum_{j \neq i} r_{ij}^r}{\sum_{j \neq i} r_{ij}^r + \sum_{j \neq i} a_{ij}^r} \quad i=1,2,\dots,m \quad (9)$$

که در آن ضریب همبستگی ساده بین متغیرهای i و j و a_{ij} ضریب همبستگی (جزئی) متغیرهای i و j به شرط ثابت بودن سایر متغیرهاست. مقادیر بالاتر MSA_i مستلزم کوچک بودن ضرایب همبستگی جزئی هستند. زیرا این ضرایب برآورد ضرایب همبستگی جملات خطا هستند که در هر مدل تجزیه عاملی فرض می‌کنیم ناهمبسته هستند.

معیار KMO

این آماره که شاخصی برای مقایسه مقادیر ضرایب همبستگی ساده و جزئی بر روی کلیه متغیرهاست، به صورت زیر محاسبه می‌شود:

$$KMO = \frac{\sum_i \sum_{j \neq i} r_{ij}^r}{\sum_i \sum_{j \neq i} r_{ij}^r + \sum_i \sum_{j \neq i} a_{ij}^r} \quad (10)$$

که در آن ضریب همبستگی ساده بین متغیرهای i و j و a_{ij} ضریب همبستگی جزئی متغیرهای i و j به شرط ثابت بودن سایر متغیرهاست. مقادیر بزرگ KMO دلالت بر تایید تجزیه عاملی دارد. این معیار در واقع همان MSA کل می‌باشد به طور کلی مقادیر حدود $0/9$ این کمیت تجزیه عاملی بسیار مناسب، حول $0/8$ مناسب، $0/7$ متعادل، $0/6$ متوسط، $0/5$ ناچیز و کمتر از $0/5$ نامناسب را نشان می‌دهند.

درصد واریانس کل تبیین شده

درصد واریانس کل تبیین شده به وسیله عامل مشترک λ_m به صورت $\frac{\lambda_i}{\sum_{j=1}^m \lambda_j} \times 100$ و

درصد تجمعی آن که بیانگر کل واریانس تبیین شده توسط عاملهای منتخب (c عامل) می‌باشد به صورت پیرو محاسبه می‌شود:

$$\sum_{i=1}^c \frac{\lambda_i}{\sum_{j=1}^m \lambda_j} \times 100 \quad C=1, \dots, m \quad (11)$$

که به ازای $C=m$ کل واریانس تبیین شده توسط عاملهای مشترک را نشان می‌دهد. (m تعداد کل عاملها)

چرخش وریماکس

چرخش وریماکس به عنوان روش استاندارد توصیه می‌شود. مانیز در این تحقیق از این روش بهره گرفته‌ایم. لذا به شرح مختصری از آن می‌پردازیم.

"اچ.اف. کایزر"^۱ برای اولین بار این روش را پیشنهاد نمود که بعداً آن را اندکی تغییر داد. با به کار بردن این چرخش دورانی متعامد بر روی ضرایب عاملها صورت می‌گیرد. با این دوران تغییرات مربعات عناصر ستونی، برآورد ضرایب عاملها را ماکسیمم می‌کند. یعنی با فرض

$$d_j = \sum_{i=1}^m \lambda_{ij}^2 \quad j = 1, \dots, p \quad (12)$$

عبارت ذیل ماکسیمم می‌شود:

$$\sum_{i=1}^m \sum_{j=1}^p (\lambda_{ij}^2 - p^{-1} d_j)^2 \quad (13)$$

که در آن ضریب عامل j ام در متغیر i ام، m تعداد عاملها و p تعداد متغیرها می‌باشد و d_j همان مجموع توان دوم λ_{ij} هاست. این روش مقادیر نسبتاً بزرگ (از نظر قدر مطلق) یا صفر به ستونهای ماتریس ضرایب عاملها اختصاص می‌دهد. در نتیجه عواملی ایجاد می‌شود که

یا شدیداً با متغیرها وابسته‌اند یا مستقل از آنها هستند. این امر سبب ساده‌تر شدن تعبیر عاملها خواهد شد.

روش رگرسیونی برای محاسبه برآورد مقادیر عاملها

با فرض آنکه $\underline{X} \sim N_p(\underline{\mu}, \Omega)$ ، مقادیر برآورد عاملها از

$$\hat{\underline{f}}_j = \hat{\underline{L}}'_z R^{-1} \underline{Z}_j \quad j=1, \dots, n \quad (14)$$

به دست می‌آید که در آن $\hat{\underline{f}}_j$ برآورد مقادیر عاملها، R برآورد ماتریس ضرایب

همبستگی بین متغیرها، \underline{Z}_j بردار استاندارد شده مشاهده زام و $\hat{\underline{L}}'_z$ ، برآورد برگردان ماتریس ضرایب عاملهای متغیرهای استاندارد شده است.

پیوست شماره (۱)

جدول ۱ - داده‌های مربوط به درصد نیروی کار ۱۴ بخش مختلف فعالیت‌های اقتصادی در ۳۰ استان کشور^(۱)

ردیف	نام استان	کشاورزی	معادن	صنعت	انرژی	ساختمان	بازرگانی	هتل	ارتباطات	واسطه‌گری‌های مالی	مستغلات	امور عمومی	آموزش	بهداشت	سایر
۱	آذربایجان شرقی	۳/۹	۰/۹	۲۹/۹	۱/۵	۱۱/۳	۱۶/۷	۰/۶	۹/۰	۱/۰	۱/۳	۸/۵	۹/۵	۲/۷	۳/۳
۲	آذربایجان غربی	۶/۲	۰/۴	۱۲/۵	۱/۲	۱۴/۲	۲۱/۸	۱/۱	۱۰/۸	۲/۴	۲/۸	۱۲/۰	۸/۴	۳/۳	۲/۸
۳	اردبیل	۳/۱	۰/۰	۱۴/۴	۱/۵	۱۹/۱	۱۸/۲	۰/۷	۱۲/۰	۲/۱	۱/۹	۱۴/۹	۸/۱	۲/۱	۱/۸
۴	اصفهان	۳/۵	۰/۱	۲۹/۲	۱/۳	۱۱/۲	۱۹/۶	۰/۶	۸/۸	۲/۰	۱/۲	۷/۶	۹/۱	۲/۹	۲/۹
۵	ایلام	۶/۰	۰/۴	۸/۰	۳/۰	۱۹/۰	۱۷/۳	۰/۸	۷/۶	۲/۲	۱/۶	۱۵/۳	۱۱/۴	۴/۷	۲/۷
۶	بوشهر	۵/۱	۰/۲	۹/۷	۱/۵	۹/۰	۲۳/۸	۰/۱	۱۱/۵	۴/۴	۳/۹	۱۷/۰	۷/۷	۲/۴	۳/۸
۷	تهران	۰/۷	۰/۴	۲۲/۴	۱/۲	۸/۷	۲۱/۴	۰/۷	۱۲/۳	۲/۷	۳/۷	۱۰/۶	۷/۱	۳/۵	۴/۷
۸	چهارمحال و بختیاری	۲/۸	۰/۳	۱۶/۹	۰/۶	۹/۸	۱۴/۶	۰/۹	۱۱/۶	۱/۵	۱/۷	۱۶/۷	۱۵/۸	۴/۳	۲/۷
۹	خراسان رضوی	۴/۴	۰/۳	۱۸/۴	۱/۱	۱۲/۰	۲۳/۲	۱/۷	۱۰/۶	۱/۸	۱/۷	۹/۶	۹/۰	۲/۴	۳/۷
۱۰	خراسان شمالی	۶/۷	۰/۲	۱۰/۳	۱/۱	۱۹/۹	۱۶/۴	۰/۸	۱۱/۳	۱/۸	۲/۰	۱۳/۷	۱۱/۴	۲/۶	۱/۷
۱۱	خراسان جنوبی	۴/۷	۰/۱	۱۷/۵	۰/۲	۹/۲	۱۴/۹	۰/۱	۹/۶	۱/۳	۱/۲	۱۸/۷	۱۶/۹	۳/۳	۲/۲
۱۲	خوزستان	۵/۷	۱/۸	۱۴/۸	۲/۸	۱۲/۴	۱۹/۴	۰/۸	۱۱/۶	۱/۲	۲/۶	۱۱/۳	۸/۷	۳/۸	۳/۰
۱۳	زنجان	۴/۸	۱/۱	۲۴/۹	۱/۷	۱۳/۴	۱۲/۷	۱/۰	۹/۹	۲/۵	۲/۶	۱۲/۷	۶/۹	۲/۵	۳/۳
۱۴	سمنان	۵/۰	۰/۵	۲۱/۷	۰/۸	۱۱/۳	۱۱/۰	۰/۵	۹/۸	۲/۸	۴/۷	۱۵/۸	۱۲/۶	۱/۶	۱/۸
۱۵	سیستان و بلوچستان	۱/۴	۰/۰	۹/۷	۲/۸	۱۴/۱	۲۱/۸	۰/۳	۱۵/۰	۱/۷	۲/۱	۱۴/۰	۱۲/۶	۱/۷	۲/۸
۱۶	فارس	۷/۲	۰/۶	۱۲/۰	۲/۴	۱۵/۲	۲۰/۶	۰/۸	۱۱/۹	۱/۷	۱/۹	۸/۴	۱۱/۰	۳/۷	۲/۵

۱- داده‌ها بر اساس طرح اشتغال و بیکاری تهیه شده توسط دایره شاخص‌های اقتصادی، اداره آمار اقتصادی مربوط به سال ۱۳۸۴ به دست آمده‌اند.

۳/۰	۲/۹	۱۳/۰	۱۰/۴	۲/۰	۲/۰	۸/۱	۰/۲	۱۷/۵	۶/۸	۱/۴	۲۳/۸	۰/۴	۸/۴	۱۷	قزوین
۵/۱	۴/۲	۹/۲	۸/۴	۲/۷	۱/۰	۵/۳	۰/۶	۱۷/۵	۱۱/۲	-/۳	۳۲/۴	۰/۳	۱/۸	۱۸	قم
۴/۰	۲/۷	۱۱/۰	۱۱/۴	۲/۱	۲/۲	۱۰/۳	۰/۶	۲۱/۳	۱۶/۵	۱/۱	۱۲/۹	۰/۱	۳/۷	۱۹	کردستان
۲/۷	۱/۹	۱۱/۴	۱۰/۹	۳/۳	۲/۴	۹/۷	۲/۱	۱۷/۷	۱۴/۰	۱/۵	۱۳/۸	۱/۴	۷/۲	۲۰	کرمان
۵/۰	۳/۶	۹/۶	۱۴/۱	۱/۹	۱/۲	۹/۹	۱/۵	۲۳/۵	۱۳/۱	۱/۷	۱۰/۱	۰/۰	۵/۰	۲۱	کرمانشاه
۳/۲	۳/۸	۱۲/۳	۱۴/۴	۲/۲	۲/۸	۱۲/۲	۰/۴	۱۰/۵	۱۳/۴	۱/۸	۷/۱	۱۰/۸	۵/۱	۲۲	کهگیلویه و بویراحمد
۴/۶	۳/۷	۱۰/۴	۱۱/۴	۳/۱	۲/۷	۱۱/۵	۰/۸	۲۵/۰	۱۰/۷	۱/۲	۸/۲	۰/۰	۶/۷	۲۳	گلستان
۴/۳	۴/۳	۸/۸	۱۰/۵	۳/۳	۲/۴	۱۳/۱	۱/۰	۲۰/۱	۱۱/۱	۱/۵	۱۴/۳	۰/۲	۵/۳	۲۴	گیلان
۳/۴	۵/۴	۱۱/۶	۱۲/۹	۱/۴	۰/۹	۹/۱	۰/۸	۱۸/۴	۱۵/۸	۱/۵	۱۴/۰	۰/۹	۴/۰	۲۵	لرستان
۵/۱	۶/۲	۱۱/۶	۱۰/۲	۲/۳	۲/۳	۱۰/۶	۱/۱	۱۹/۶	۱۲/۴	۱/۴	۱۲/۹	۰/۲	۴/۰	۲۶	مازندران
۳/۲	۲/۴	۸/۶	۸/۶	۱/۶	۱/۴	۱۰/۳	۰/۴	۲۱/۲	۸/۶	۱/۱	۲۷/۲	۰/۲	۵/۱	۲۷	مرکزی
۳/۷	۵/۲	۴/۳	۱۵/۱	۱/۲	۲/۷	۱۸/۴	۱/۶	۱۸/۳	۱۴/۵	۲/۵	۹/۴	۰/۰	۳/۱	۲۸	هرمزگان
۴/۴	۲/۹	۱۲/۴	۹/۶	۱/۸	۲/۵	۱۰/۳	۰/۴	۲۵/۷	۱۱/۵	۰/۶	۱۱/۸	۰/۲	۶/۰	۲۹	همدان
۲/۹	۲/۲	۸/۴	۸/۱	۱/۲	۲/۰	۱۰/۲	۰/۶	۱۳/۸	۱۳/۴	۱/۶	۳۲/۰	۰/۲	۳/۴	۳۰	یزد

جدول ۲ - میانگین و انحراف معیار متغیرها در ۳۰ استان کشور

ردیف	نام متغیر	میانگین	انحراف معیار
۱	کشاورزی	۴/۷	۱/۸
۲	معدن	۰/۷	۱/۹
۳	صنعت	۱۶/۷	۷/۷
۴	انرژی	۱/۵	۰/۷
۵	ساختمان	۱۲/۸	۳/۲
۶	بازرگانی	۱۸/۸	۳/۹
۷	هتل	۰/۸	۰/۵
۸	ارتباطات	۱۰/۷	۲/۳
۹	واسطه‌گریهای مالی	۲/۱	۰/۷
۱۰	املاک و مستغلات	۲/۲	۰/۹
۱۱	امور عمومی	۱۲/۱	۳/۰
۱۲	آموزش	۱۰/۱	۲/۶
۱۳	بهداشت	۳/۳	۱/۱
۱۴	سایر	۳/۳	۱/۰

جدول ۵ - مقایسه مطلوبیت تجزیه عاملی بر اساس آماره KMO به ترتیب حذف متغیرها بر اساس آماره MSA

کل واریانس تبیین شده توسط عاملهای منتخب	تعداد عاملها	متغیرهای باقیمانده در مدل به ترتیب MSA													آماره MSA	آماره KMO	متغیر حذف شده	مرحله				
		صنعت	سایر	ارتباطات	واسطه‌گری مالی	امور عمومی	بازرگانی	آموزش	انرژی	ساختمان	مستغلات	هتل	بهداشت	معدن					کشاورزی			
۷۸/۲۲	۶																	متغیر کشاورزی	۰/۰۵۲	—	۱	
																			MSA			
۷۳/۹۳	۵																		متغیر معدن	۰/۱۸۱	کشاورزی	۲
																			MSA			
۶۸/۹۳	۴																		متغیر آموزش	۰/۳۳۶	معدن	۳
																			MSA			
۶۹/۷۱	۴																		متغیر بازرگانی	۰/۴۸۶	آموزش	۴
																			MSA			
۷۲/۷۰	۴																		متغیر سایر	۰/۵۴۲	بازرگانی	۵
																			MSA			
۶۲/۱۲	۳																		متغیر بهداشت	۰/۵۶۷	سایر	۶
																			MSA			
۶۸/۰۹	۳																		متغیر هتل	۰/۵۹۴	بهداشت	۷
																			MSA			

۶۰/۱۶۶	۲								صنعت	امور عمومی	انرژی	ارتباطات	واسطه‌گری مالی	ساختمان	مستغلات	متغیر	۰/۶۲۶	هتل	۸
									۰/۶۷۱	۰/۶۶۱	۰/۶۴۹	۰/۶۴۶	۰/۶۱۰	۰/۵۸۹	۰/۵۲۷	MSA			
۶۲/۷۴	۲								ارتباطات	واسطه‌گری مالی	صنعت	امور عمومی	انرژی	ساختمان	متغیر	۰/۶۴۵	مستغلات	۹	
									۰/۷۰۰	۰/۶۶۴	۰/۶۵۸	۰/۶۴۷	۰/۶۲۷	۰/۵۵۵	MSA				
۶۷/۴۰	۲								واسطه‌گری مالی	ارتباطات	صنعت	امور عمومی	انرژی	متغیر	۰/۶۵۴	ساختمان	۱۰		
									۰/۷۳۷	۰/۷۳۶	۰/۶۵۱	۰/۶۰۰	۰/۵۵۰	MSA					
۵۲/۴۶	۱									واسطه‌گری مالی	ارتباطات	امور عمومی	صنعت	متغیر	۰/۶۸۵	انرژی	۱۱		
										۰/۸۰۵	۰/۷۰۱	۰/۶۷۴	۰/۶۳۸	MSA					

جدول ۶- مشخصات عامل ها و میزان واریانس تعیین شده توسط آن‌ها طی مراحل مختلف^۱

مرحله	مطلوبیت تجزیه عاملی براساس آماره KMO	عامل ۱	عامل ۲	عامل ۳	عامل ۴	عامل ۵	عامل ۶
توسط عامل	توسط عامل	توسط عامل	توسط عامل	توسط عامل	توسط عامل	توسط عامل	توسط عامل
۱	نامناسب	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات دربرابر آموزش	اهمیت بخش امور عمومی و آموزش دربرابر صنعت	اهمیت واسطه‌گریهای مالی و مستغلات	اهمیت بهداشت و سایر	اهمیت معدن دربرابر بازرگانی	اهمیت کشاورزی
۲	نامناسب	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات دربرابر آموزش	اهمیت بخش امور عمومی و آموزش دربرابر صنعت	اهمیت واسطه‌گریهای مالی و مستغلات	اهمیت بهداشت و سایر	اهمیت معدن دربرابر بازرگانی	
۳	نامناسب	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات دربرابر آموزش	اهمیت بازرگانی، بهداشت و سایر	اهمیت واسطه‌گریهای مالی و مستغلات	اهمیت امور عمومی و آموزش دربرابر صنعت		
۴	ناچیز	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات دربرابر صنعت	اهمیت واسطه‌گری و مستغلات	اهمیت بازرگانی و سایر دربرابر امور عمومی	اهمیت سایر و بهداشت		
۵	متوسط	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات دربرابر صنعت	اهمیت واسطه‌گری و مستغلات	اهمیت امور عمومی دربرابر صنعت	اهمیت بهداشت و سایر		

۱- این عوامل، عوامل دوران یافته براساس روش وریماکس و به روش تجزیه به مولفه های اصلی، بادر نظر گرفتن مقادیر ویژه بالاتر از یک به دست آمده‌اند.

						۱۹/۲۹	اهمیت واسطه‌گری‌های مالی و مستغلات	۲۰/۰۷	اهمیت امور عمومی در برابر صنعت	۲۲/۷۶	اهمیت بخشهای انرژی، ساختمان، هتل و ارتباطات در برابر صنعت	متوسط	۶
						۱۵/۳۱	اهمیت هتل در برابر امور عمومی	۲۲/۲۶	اهمیت واسطه‌گری مالی و مستغلات	۲۰/۵۲	اهمیت بخشهای انرژی، ساختمان و ارتباطات در برابر صنعت	متوسط	۷
								۲۷/۰۳	اهمیت واسطه‌گری مالی و مستغلات	۲۲/۶۳	اهمیت انرژی، ساختمان، و ارتباطات در برابر صنعت	متوسط	۸
								۲۹/۲۰	اهمیت انرژی و ساختمان در برابر صنعت	۲۳/۵۴	اهمیت ارتباطات، واسطه‌گری مالی و امور عمومی در برابر صنعت	متعادل	۹
								۲۵/۰۱	اهمیت انرژی و ارتباطات در برابر صنعت	۲۵/۳۹	اهمیت امور عمومی و واسطه‌گری‌های مالی در برابر صنعت	متعادل	۱۰

جدول ۸- مقادیر عاملهای استانها برای مرحله ۹ ام^(۱)

ردیف	نام استان	مقادیر عامل اول	مقادیر عامل دوم
۱	آذربایجان شرقی	-۱/۷۶۱	-۰/۱۴۴
۲	آذربایجان غربی	۰/۳۳۰	۰/۰۴۳
۳	اردبیل	۰/۳۰۸	۱/۱۰۸
۴	اصفهان	-۱/۲۱۳	-۰/۶۴۸
۵	ایلام	-۰/۰۴۱	۲/۰۴۷
۶	بوشهر	۲/۶۳۳	-۱/۴۰۲
۷	تهران	۰/۴۳۹	-۱/۱۴۹
۸	چهارمحال و بختیاری	۰/۵۴۴	-۰/۹۴۶
۹	خراسان رضوی	-۰/۴۷۵	-۰/۲۶۶
۱۰	خراسان شمالی	۰/۰۳۴	۱/۱۹۰
۱۱	خراسان جنوبی	۰/۴۴۸	-۱/۳۸۷
۱۲	خوزستان	-۰/۵۶۲	۱/۲۸۱
۱۳	زنجان	-۰/۱۱۳	-۰/۱۶۶
۱۴	سمنان	۰/۷۵۲	-۱/۲۲۷
۱۵	سیستان و بلوچستان	۰/۶۳۳	۱/۵۸۲
۱۶	فارس	۰/۵۵۵	۱/۴۳۸
۱۷	قزوین	-۰/۵۴۹	-۱/۲۶۷
۱۸	قم	-۲/۲۲۸	-۱/۱۹۳
۱۹	کردستان	-۰/۰۵۴	۰/۴۱۷
۲۰	کرمان	-۰/۰۱۴	۰/۱۶۰
۲۱	کرمانشاه	-۰/۱۹۲	۰/۶۳۴
۲۲	کهگیلویه و بویراحمد	۱/۲۴۵	۰/۳۰۶
۲۳	گلستان	۰/۸۷۱	۰/۴۹۹
۲۴	گیلان	۰/۴۴۹	۰/۱۸۴
۲۵	لرستان	-۰/۸۹۴	۰/۹۵۶
۲۶	مازندران	۰/۰۷۵	-۰/۰۵۸
۲۷	مرکزی	-۱/۰۵۵	-۰/۸۹۶
۲۸	هرمزگان	۱/۸۶۰	۱/۱۸۴
۲۹	همدان	۰/۲۵۰	-۰/۷۹۸
۳۰	یزد	-۱/۲۱۳	-۰/۱۰۵

۱- اعداد تا ۳ رقم اعشار رند شده‌اند.

جدول ۹- گروه‌بندی شهرها براساس تاثیرگذاری عاملها طی مراحل مختلف

مرحله	گروه‌های استانی	عامل اول	عامل دوم	عامل سوم	عامل چهارم	عامل پنجم	عامل ششم
۱	گروه اول	هرمزگان- ایلام- فارس- خوزستان	بوشهر- چهارمحال و بختیاری- خراسان جنوبی- سیستان و بلوچستان	تهران- بوشهر- گلستان	قم- کرمانشاه- کهگیلویه و بویراحمد- گلستان- لرستان- مازندران	زنجان- کهگیلویه و بویراحمد	ایلام- خراسان شمالی- فارس- قزوین- کرمان- گلستان
	گروه دوم	خراسان جنوبی- قم- چهارمحال و بختیاری- قزوین- سمنان	آذربایجان شرقی- اصفهانی- زنجان- قم- مرکزی- یزد	آذربایجان شرقی- چهارمحال و بختیاری- خراسان جنوبی- لرستان	اردبیل- خراسان شمالی- سمنان- سیستان و بلوچستان- یزد	کرمانشاه- گلستان- همدان	تهران- سیستان و بلوچستان- هرمزگان- یزد
۲	گروه اول	اردبیل- ایلام- هرمزگان	ایلام- چهارمحال و بختیاری- خراسان جنوبی- سیستان و بلوچستان	بوشهر- تهران- سمنان- گلستان	قم- کرمانشاه- کهگیلویه و بویراحمد- گلستان- گیلان- مازندران	زنجان- کهگیلویه و بویراحمد	
	گروه دوم	خراسان جنوبی- چهارمحال و بختیاری- سمنان- قزوین- قم	آذربایجان شرقی- اصفهان- زنجان- قم- مرکزی- یزد	آذربایجان شرقی- ایلام- خراسان جنوبی- لرستان	اردبیل- خراسان شمالی- سمنان- سیستان و بلوچستان- یزد	کرمانشاه- گلستان- همدان	
۳	گروه اول	اردبیل- ایلام- فارس- هرمزگان	تهران- کرمانشاه- گلستان- گیلان- مازندران	بوشهر- تهران- سمنان	ایلام- چهارمحال و بختیاری- خراسان جنوبی- کهگیلویه		

		آذربایجان شرقی- اصفهان- خراسان رضوی- زنجان- قم- مرکزی- یزد	آذربایجان شرقی- ایلام- خراسان جنوبی- لرستان	اردبیل- خراسان شمالی- سمنان- یزد	خراسان جنوبی- چهارمحال و بختیاری- سمنان- قزوین- قم	گروه دوم	
		چهارمحال و بختیاری- خراسان جنوبی- کرمانشاه- لرستان-مازندران- هرمزگان	تهران- خراسان رضوی- کرمانشاه- گلستان- همدان	بوشهر- تهران- سمنان- گلستان	اردبیل- ایلام- خراسان شمالی- سیستان و بلوچستان- فارس- هرمزگان	گروه اول	۴
		اردبیل- اصفهان- سمنان- کرمان- یزد	ایلام- چهارمحال و بختیاری- خراسان شمالی- خراسان جنوبی- سمنان- کهگیلویه	آذربایجان شرقی- اصفهان- قم- لرستان- یزد	خراسان جنوبی- سمنان- قزوین- قم	گروه دوم	
		قم- کرمانشاه- گلستان- گیلان- مازندران- هرمزگان	ایلام- بوشهر- چهارمحال و بختیاری- خراسان جنوبی- کهگیلویه و بویراحمد- هرمزگان	بوشهر- تهران- سمنان- کرمان- گلستان- گیلان	ایلام- فارس- کرمان- هرمزگان	گروه اول	۵
		اردبیل- خراسان شمالی- سمنان- سیستان و بلوچستان- یزد	آذربایجان شرقی- اصفهان- خراسان رضوی- قم- کرمان- یزد	ایلام- آذربایجان شرقی- خراسان جنوبی- لرستان	خراسان جنوبی- سمنان- قزوین- قم- مرکزی	گروه دوم	

۶	گروه اول	فارس- کرمان- هرمزگان ایلام- بوشهر- خراسان جنوبی- سیستان و گلستان بلوچستان- کهگیلویه و بویراحمد- هرمزگان بوشهر- تهران- خراسان جنوبی- سمنان- کرمان- گلستان			
	گروه دوم	خراسان جنوبی- سمنان- قزوین- قم- مرکزی آذربایجان شرقی- اصفهان- خراسان رضوی- کرمان- مرکزی- یزد ایلام- خراسان جنوبی- لرستان			
۷	گروه اول	ایلام- سیستان و بلوچستان- هرمزگان بوشهر- تهران- سمنان- کرمان- گلستان- گیلان خراسان رضوی- کرمان			
	گروه دوم	آذربایجان شرقی- اصفهان- خراسان جنوبی- قزوین- قم- مرکزی آذربایجان شرقی- ایلام- خراسان جنوبی- لرستان بوشهر- چهارمحال و بختیاری- خراسان جنوبی- کهگیلویه و بویراحمد			
۸	گروه اول	اردبیل- ایلام- خراسان شمالی- سیستان و بلوچستان- هرمزگان بوشهر- تهران- سمنان- گلستان			
	گروه دوم	آذربایجان شرقی- اصفهان- قزوین- قم- مرکزی آذربایجان شرقی- لرستان- یزد			

				ارdebیل - ایلام - خراسان شمالی - زنجان - سیستان و بلوچستان - فارس - هرمزگان	بوشهر - کهگیلویه و بویراحمد - هرمزگان	گروه اول	۹
				بوشهر - تهران - خراسان جنوبی - سمنان - قزوین - قم	آذربایجان شرقی - اصفهان - قم - مرکزی - یزد	گروه دوم	
				ایلام - خوزستان - سیستان و بلوچستان - فارس - هرمزگان	بوشهر - خراسان جنوبی - سمنان - کهگیلویه و بویراحمد - هرمزگان	گروه اول	۱۰
				خراسان جنوبی - سمنان - قم	آذربایجان شرقی - اصفهان - خوزستان - فارس - قم - مرکزی - یزد	گروه دوم	

