تحلیل عاملی؛ تکنیک خلاصه‌سازی متغیرها

غلامحسین کرمی
دانشجوی دکترای رشتهٔ آموزش و ترویج کشاورزی دانشگاه شیراز

چکیده

تحلیل عاملی تکنیکی است که کاهش تعداد زیادی از متغیرهای وابسته به‌هم را به تعداد کوچک‌تری از ابعاد پنهان یا مکنون امکان‌پذیر می‌سازد. هدف اصلی آن، کاربرد کوچک‌ترین مفاهیم تبیین‌کننده به‌منظور تبیین بیشینهٔ مقدار واریانس مشترک در ماتریس هم‌بستگی است. مفروضهٔ اساسی تحلیل عاملی این است که عامل‌های زیربنایی متغیرها را می‌توان برای تبیین پدیده‌های پیچیده به‌کار برد. هم‌بستگی‌های مشاهده شده بین متغیرها، حاصل اشتراک آن‌ها در این عامل‌هاست. هدف تحلیل عاملی تشخیص این عامل‌های مشاهده‌ناپذیر بر پایهٔ مجموعه‌ای از متغیرهای مشاهده‌پذیر است.

تحلیل عاملی، متغیرهای وابستهٔ از قبل تعیین‌شده‌ای ندارد. موارد استفادهٔ تحلیل عاملی را به دو دستهٔ کلی می‌توان تقسیم کرد: مقاصد اکتشافی و مقاصد تأییدی. اگر هیچ حدسی از ساختار روابط میان متغیرها وجود نداشته باشد، از «تحلیل عاملی اکتشافی» استفاده می‌شود. اما اگر متغیرها براساس ابعاد شناخته‌شده‌ای بارگذاری شده باشند، باید از «تحلیل عاملی تأییدی» استفاده کرد.

کلیدواژه‌ها: مقدار ویژه، عامل مشترک، بار عاملی، تحلیل عاملی اکتشافی، تحلیل عاملی تأییدی، هم‌بستگی درونی

مقدمه

در روش تحقیق، تکنیک‌ها از اهمیت ویژه‌ای برخوردارند. تحلیل عاملی از جمله تکنیک‌های قدرتمندی است که در علوم گوناگون خصوصاً علوم اجتماعی، کاربردهای مختلفی دارد. با استفاده از این تکنیک محقق می‌تواند دست به شاخص‌سازی بزند و مجموعهٔ پیچیده‌ای از داده‌ها را به چند عامل تقلیل دهد. این روش دارای پیش‌فرض‌های متفاوتی همچون نرمال بودن توزیع داده‌هاست که با رعایت آن‌ها می‌توان نتایج دقیق‌تری به‌دست آورد. برای انجام آن باید ابتدا مقاصد را مشخص کرد و سپس با یکی از روش‌های موجود در آن، به تحلیل پرداخت. (نظری و مختاری، ۱۳۸۸: ۲۰) پیچیدگی و دشواری درک، اجرا و تفسیر نتایج تحلیل عاملی موجب می‌شود که بسیاری از کاربران بدون آشنایی با منطق زیربنایی، محدودیت‌ها و نیز دامنهٔ کاربرد آن، به اجرا و به‌کارگیری یافته‌های حاصل از آن بپردازند و بدین ترتیب زمینه‌ساز تفسیرهای نامعتبر و ناروا و نیز تعمیم‌های نادرست از پدیده‌ها و متغیرهای مورد مطالعه شوند.

تحلیل عاملی چیست؟

تحلیل عاملی نامی عمومی است برای برخی روش‌های آماری چندمتغیره که هدف اصلی آن‌ها، خلاصه کردن داده‌هاست. این روش به بررسی «هم‌بستگی درونی» تعداد زیادی از متغیرها می‌پردازد و در نهایت آن‌ها را در قالب عامل‌های کلی محدودی دسته‌بندی و تبیین می‌کند. تحلیل عاملی روشی هم وابسته است که در آن، تمامی متغیرها به‌طور هم‌زمان مد نظر قرار می‌گیرند، به‌عبارت دیگر، در این تکنیک هر یک از متغیرها به‌عنوان یک متغیر وابسته لحاظ می‌شود. (کلانتری، ۱۳۹۱: ۲۸٤)

تحلیل عاملی تکنیکی است برای شناسایی عواملی که به‌طور آماری اختلاف مشترک میان اندازه‌گیری‌ها را توضیح می‌دهند. به‌طور کلی، تعداد عوامل به‌صورت قابل ملاحظه‌ای کمتر از تعداد اندازه‌گیری‌هاست. در نتیجه، عوامل به‌صورت مختصر، مجموعه‌ای از اندازه‌ها را ارائه می‌کنند. از این دیدگاه، تحلیل عاملی را می‌توان به‌عنوان یک تکنیک کاهش‌دهندهٔ داده‌ها تلقی کرد. زیرا تعداد زیادی از متغیرهای اندازه‌گیری شدهٔ «همپوش»[1] را به مجموعهٔ کوچک‌تری از عوامل کاهش می‌دهد. اگر یک مطالعه به‌خوبی طراحی شود، به‌طوری که مجموعه‌های متفاوت مقادیر، ابعاد متفاوت یک سیستم مفهومی وسیع‌تر را منعکس کنند، تحلیل عاملی می‌تواند عواملی را ارائه دهد که نشان‌دهندهٔ این ابعاد باشند. به‌طور خاص، عوامل را می‌توان مطابق متغیرهای پنهان غیرمشهود یک نظریه تلقی کرد که به درک رفتار کمک می‌کنند. (زرگر، ۱۳۸۳: ٤۱۱)

بنیان‌های اولیهٔ تحلیل عاملی

گالتون اولین کسی بود که پایه‌های اولیهٔ تحلیل عاملی را بنا نهاد. پس از او، **کارل پیرسون** در اوایل قرن بیستم روشی برای تحلیل عاملی براساس یک فضای چندبعدی هندسی ارائه داد و در پی آن، **مک دونال** این روش را در زمینهٔ شناسایی جرائم و ارتباط آن با ویژگی‌های فیزیکی افراد مورد استفاده قرار داد. (کلانتری، ۱۳۹۱: ۲۸۲) عموماً **چارلز اسپیرمن** را پدر تحلیل عاملی و زمان پیدایش این روش را سال ۱۹۰٤ می‌دانند. در این سال، او مقالهٔ هوش عمومی، تعیین عینی و اندازه‌گیری آن را نوشت و به معرفی مدل‌های ریاضی این روش پرداخت. سه سال قبل از آن، پیرسون در سال ۱۹۰۱ روش محورهای اصلی را معرفی کرده بود. در سال‌های بعد، افرادی نظیر

تحلیل عاملی تکنیکی است برای شناسایی عواملی که به‌طور آماری اختلاف مشترک میان اندازه‌گیری‌ها را توضیح می‌دهند. به‌طور کلی، تعداد عوامل به‌صورت قابل ملاحظه‌ای کمتر از تعداد اندازه‌گیری‌هاست. در نتیجه، عوامل به‌صورت مختصر، مجموعه‌ای از اندازه‌ها را ارائه می‌کنند. از این دیدگاه، تحلیل عاملی را می‌توان به‌عنوان یک تکنیک کاهش‌دهندهٔ داده‌ها تلقی کرد. زیرا تعداد زیادی از متغیرهای اندازه‌گیری شدهٔ «همپوش» را به مجموعهٔ کوچک‌تری از عوامل کاهش می‌دهد

سیریل برت، **گودی فری تامپسون**، **ماکسول گارنت** و **کارل هولزنیگر** روش‌های تحلیل عاملی را گسترش دادند. تحلیل عاملی در سال ۱۹۳۹ به ظهور «تحلیل خوشه‌ای»[2] از سوی **رابرت تریون** منجر شد. (گال و همکاران، ۱۳۹۰: ۷۷٤) با این تلاش‌ها اصول و بنیان‌های تحلیل عاملی شکل گرفت و امروزه در سطح وسیع در شاخه‌های گوناگون علوم نظیر روان‌شناسی، اقتصاد، جامعه‌شناسی، مدیریت، کشاورزی و پزشکی مورد استفاده قرار می‌گیرد. (کلانتری، ۱۳۹۱: ۲۸۲)

تحلیل عاملی ابزاری ارزشمند در پژوهش‌های علمی است اما عوامل پدیدآمده در این تکنیک از نظر معنی و سودمندی به متغیرهایی بستگی دارند که در ماتریس هم‌بستگی وارد می‌شوند. بنابراین، پژوهشگر باید با دقت تعداد و نوع متغیرهایی را که در تحلیل عاملی وارد می‌کند، در نظر داشته باشد. اگر متغیرها هیچ یا اندکی اشتراک مفهومی با یکدیگر داشته باشند، استفاده از تحلیل عاملی مناسب نخواهد بود. (گال و همکاران، ۱۳۹۰: ۷۷۸)

مفاهیم کلیدی در تحلیل عاملی

در ادامه به چند اصطلاح مهم و عمده در تحلیل عاملی اشاره می‌شود (کلانتری، ۱۳۹۱: ۲۸۳):

- **اشتراک**[3]: میزان واریانس مشترک بین یک متغیر با سایر متغیرهای به‌کار گرفته شده در تحلیل عاملی است.
- **مقدار ویژهٔ خاص**[4]: مقدار واریانس تبیین شده به‌وسیلهٔ هر عامل است.

● **عامل**[5] **ترکیب خطی:** متغیر اصلی که حاصل خلاصهٔ متغیرهای مشاهده شده است.
● **بار عاملی**[6]: هم‌بستگی بین متغیرهای اصلی و عوامل است. مجذور بار عاملی نشان می‌دهد که چند درصد از واریانس در یک متغیر توسط آن عامل تبیین می‌شود.
● **ماتریس عاملی**[7]: جدولی که بارهای عاملی کلیهٔ متغیرها را در هر عامل نشان می‌دهد.
● **چرخش عاملی**[8]: فرایندی برای تعدیل محور عامل به منظور دستیابی به عامل‌های معنی‌دار و ساده‌تر است.
● **نمرهٔ عاملی**[9]: عامل‌های حاصل از تحلیل عاملی را می‌توان در تحلیل‌های آماری بعدی به‌کار گرفت. مقادیر حاصل از ترکیب چنین متغیرهایی نمرهٔ عاملی نامیده می‌شود.

اهداف تحلیل عاملی

هدف اصلی تحلیل عاملی تلخیص تعداد زیادی از متغیرها در تعداد محدودی از عامل‌هاست؛ به‌طوری که کمترین میزان از دست رفتن اطلاعات را داشته باشیم. (کلانتری، ۱۳۹۱: ۲۸٤) به‌طور کلی موارد استفادهٔ تحلیل عاملی را به دو دسته می‌توان تقسیم کرد:

الف) مقاصد اکتشافی و ب) مقاصد تأییدی.

چنانچه هدف پیدا کردن متغیرهای مکنون یا سازه‌های یک مجموعه متغیر اندازه‌گیری شده (نمرهٔ سؤالات یک آزمون یا ریزنمرات آزمون) باشد، از «تحلیل عاملی اکتشافی»[10] استفاده می‌شود. از لحاظ نظری متغیرهای مکنون یا سازه‌ها، علل زیربنایی متغیرهای اندازه‌گیری شده هستند. رگرسیون متغیرهای اندازه‌گیری شده روی متغیرهای مکنون وزن‌هایی فراهم می‌آورد که بارهای عاملی نامیده می‌شوند. تحلیل عامل مشترک، واریانس یا هر متغیر اندازه‌گیری شده را به دو واریانس مشترک و واریانس اختصاصی افراز می‌کند. واریانس مشترک، تغییرات مشترک متغیرهای اندازه‌گیری شده را با متغیرهای مکنون نمایان می‌کند.

در موارد اکتشافی که هدف تلخیص مجموعه‌ای از داده‌ها باشد، از «تحلیل مؤلفه‌های اصلی»[11] استفاده می‌شود. در تحلیل مؤلفه‌های اصلی، واریانس کل متغیرهای مشاهده شده تحلیل می‌شود. ماتریس هم‌بستگی متغیرهای اندازه‌گیری شده دارای قطر اصلی ۱ است، در حالی‌که در تحلیل عامل مشترک در قطر اصلی ماتریس هم‌بستگی میزان اشتراک (واریانس مشترک متغیر اندازه‌گیری شده و متغیرهای مکنون) قرار می‌گیرد. وقتی میزان اشتراک به ۱ نزدیک باشد، نتایج تمام روش‌های اکتشافی با نتایج مؤلفه‌های اصلی مشابه خواهد بود. (سرمد و همکاران، ۱۳۸۳: ۲٦۸)

انتخاب هر یک از این مدل‌ها به هدف محقق بستگی دارد. مدل تحلیل مؤلفه‌های اصلی زمانی مورد استفاده قرار می‌گیرد که هدف محقق تلخیص متغیرها و دستیابی به تعداد محدودی عامل برای اهداف پیش‌بینی باشد. در مقابل، تحلیل عاملی مشترک زمانی به‌کار می‌رود که هدف، شناسایی عامل‌ها یا ابعادی باشد که به سادگی قابل شناسایی نیستند. (کلانتری، ۱۳۹۱: ۲۸۹)

در «تحلیل عامل تأییدی»[12] که هدف پژوهشگر تأیید ساختار عاملی ویژه‌ای است، دربارهٔ تعداد عامل‌ها به‌طور آشکار فرضیه‌ای بیان می‌شود و برازش ساختار عاملی مورد نظر در فرضیه، با ساختار کوواریانس متغیرهای اندازه‌گیری شده مورد آزمون قرار می‌گیرد. (سرمد و همکاران، ۱۳۸۳: ۲٦۹) در این روش محقق انتظار دارد طرح و نقشهٔ خاصی از عوامل پنهان در ماورای متغیرها را بیازماید و متغیرها چیدمان خاصی داشته باشند. لذا در این روش محقق به آزمون فرضیات مربوط به یک ساختار عاملی خاص دست می‌زند. پژوهشگر در این شکل از تحلیل عاملی به دنبال تهیهٔ مدلی است که فرض می‌شود، داده‌های تجربی را بر پایهٔ چند پارامتر نسبتاً اندک، توصیف، تبیین یا توجیه می‌کند. این مدل مبتنی بر اطلاعات پیش‌تجربی دربارهٔ ساختار داده‌هاست که می‌تواند به یکی از اشکال زیر باشد:

● یک نظریه یا فرضیه
● یک طرح طبقه‌بندی کنندهٔ معین برای گویه‌ها یا پاره‌تست‌ها در انطباق با ویژگی‌های عینی شکل و محتوا
● شرایط معلوم تجربی
● دانش حاصل از مطالعات قبلی دربارهٔ داده‌های وسیع

چنانچه هدف پیدا کردن متغیرهای مکنون یا سازه‌های یک مجموعه متغیر اندازه‌گیری شده (نمرهٔ سؤالات یک آزمون یا ریزنمرات آزمون) باشد، از «تحلیل عاملی اکتشافی» استفاده می‌شود. از لحاظ نظری متغیرهای مکنون یا سازه‌ها، علل زیربنایی متغیرهای اندازه‌گیری شده هستند. رگرسیون متغیرهای اندازه‌گیری شده روی متغیرهای مکنون وزن‌هایی فراهم می‌آورد که بارهای عاملی نامیده می‌شوند

اینکــه کدام یــک از این دو روش بایــد در تحلیل عاملی به‌کار رود، مبتنی بر هدف تحلیل داده‌هاســت. تمایز مهم روش‌هــای تحلیل اکتشــافی و تأییدی در این اســت که روش اکتشافی باصرفه‌ترین روش تبیین واریانس مشترک زیربنایی یک ماتریس هم‌بســتگی را مشخص می‌کند. در حالی‌که روش‌های تأییدی (آزمون فرضیه) تعیین می‌کنند که داده‌ها با یک ساختار عاملی معین (که در فرضیه آمده) هماهنگ هســتند یا خیر. (Brown, 2006: 88) از ســوی دیگر، بیشتر مطالعات ممکن است تا حدودی هم اکتشافی و هم تأییدی باشند؛ زیرا شامل متغیر معلوم و تعدادی متغیر مجهول‌انــد. متغیرهای معلوم را باید با دقت زیادی انتخاب کرد تا حتی‌الامکان دربارهٔ متغیرهای نامعلوم استخراج شده، اطلاعات بیشتری فراهم آید. مطلوب است فرضیه‌ای که با تحلیل اکتشافی تدوین می‌شــود، از طریق قرار گرفتن در معرض روش‌های آماری دقیق‌تر تأیید یا رد شود.

پیش‌فرض‌های تحلیل عاملی

متغیرهایــی که در تحلیل عاملی مورد اســتفاده قرار می‌گیرند، باید دارای شرایطی باشند که عبارت‌اند از:

● **حجــم نمونه**: بــرای هــر متغیر، ماده یا شــاخص اندازه‌گیری حداقل ۵ تا ۱۰ آزمودنی وجود داشته باشد. نمونهٔ ۱۰۰ موردی قابل‌قبول اســت، ولی بیش از ۱۰۰ مورد مناســب است. (نظری و مختاری، ۱۳۸۸: ۲٤) در برخی موارد محقق مجبور اســت با نســبت ۲ به ۱ نیز بــه تحلیل عاملی بپــردازد، اما در این صورت بایســتی توجه داشت که تفسیر یافته‌ها با اطمینان کمتری انجام می‌شــود. در هر صورت تعداد نمونه‌ها نباید کمتر از ۵۰ باشد. (کلانتری، ۱۳۹۱: ۲۸۷)

● **مشــخصات داده‌های خــام**: داده‌های خام باید از نوع کمــی، دامنهٔ نمرات آن‌ها بــزرگ و دارای توزیع متقارن و تک‌نمایی[۱۳] باشــند. (زرگر، ۱۳۸٤: ٤۱۱) اما هر نوع متغیر مرتبط با مسئلهٔ تحقیق، مانند متغیرهای مجازی (با کدهای ۰ و ۱) و غیرپارامتری یا کیفی را نیز می‌توان به‌کار برد.

● **نرمال بودن**: متغیرها دارای توزیع نرمال باشند.

● **خطــی بــودن متغیرها**: خطی بودن مهم اســت، زیرا تحلیــل عاملی مبتنی بر هم‌بســتگی اســت و متغیرهای اندازه‌گیری شــده، به‌صورت خطی با عامل‌ها ارتباط دارند. اگر متغیرها تحلیل عامل شده باشند، خصوصاً اینکه گسترهٔ پاســخ‌های مورد بررســی خیلی محدود باشد (برای مثال، آزمون‌های دوگزینه‌ای درســت- نادرســت)، و توزیع‌های متغیر در «چولگی»[۱٤] متفاوت باشند، این فرض ممکن است رعایت نشــود. عدم رعایت فرض ممکن است به شناسایی عامل‌های جعلی بینجامد. اگر این فرض درست باشد، روابط بین متغیرهای مورد سنجش نیز خطی است. برای تعیین ارتباط خطی متغیرها، از «نمودار پراکندگی نقاط»[۱۵] برای دو به دو متغیرها استفاده می‌شود. (زرگر، ۱۳۸۳: ٤۱۵)

● **شــاخص رابطه**: معمولی‌ترین شــاخص رابطهٔ ضریب هم‌بستگی از نوع «پیرسون» است. بدیهی است که مفروضهٔ اصلی در محاســبهٔ این ضریب هم‌بستگی وجود یک توزیع دو متغیری نرمال است. چنانچه توزیع نمرات نرمال نباشد یا دامنهٔ تغییر نمرات محدود باشد، عامل‌ها در تحلیل عاملی کمتر مشخص می‌شوند. بنابراین هرگاه ضریب هم‌بستگی مورد اســتفاده قرار می‌گیرد، بهتر اســت از متغیرهایی که تفاوت افراد را مشخص می‌کنند، استفاده کرد. هرگاه رابطهٔ متغیرها انحنایی باشد، ضریب هم‌بستگی پیرسون نمی‌تواند براورد دقیقی از رابطهٔ دو متغیر به‌دست دهد. این حالت در مــواردی که یکی از متغیرها یــا هر دو آن‌ها به‌طور طبیعی اسمی دوارزشی باشند، مانند مواردی که برای آن‌ها ضریب هم‌بســتگی دورشــته‌ای نقطه‌ای یا فی محاســبه می‌شود، محسوس‌تر است. (سرمد و همکاران، ۱۳۷۸: ۲۷۰- ۲٦۹)

● **مستقل بودن اندازه‌گیری**: هر نوع وابستگی متغیرها بــه یکدیگر به بالا رفتن هم‌بســتگی بین آن‌ها می‌انجامد و سبب می‌شــود که این متغیرها در عامل واحدی ظاهر شوند. از جمله مواردی که این وابستگی صورت می‌گیرد، موقعی است که از نمرات زیرِ مقیاس‌ها و نمرهٔ کل مقیاس در تحلیل استفاده شود (مثلاً نمرهٔ کل بهرهٔ هوشی، نمرهٔ کلامی بهرهٔ هوشــی، نمرهٔ کلاســی بهرهٔ هوشــی و نمرهٔ عملکردی بهرهٔ هوشی تحلیل شود) یا نمرات زیر مقیاس‌ها و یا نمرهٔ کل باید در تحلیل وارد شوند. مقیاس‌هایی که در آن‌ها بعضی از سؤالات یا ماده‌های آزمون مشترک هستند نیز وابستگی ایجاد می‌کنند.

● **ترکیب ماتریس داده‌ها**: اگر محققی بخواهد ابعاد مشــترکی بین چند مقیاس اندازه‌گیری پیدا کند، باید تمام اندازه‌ها روی نمونهٔ واحدی به‌دســت آمده باشند. (همان: ۲۷۰)

● **دورافتاده‌هــا در میــان موردها**: تحلیل عاملی به موردهای دورافتاده حســاس است. این موارد نیاز است مشخص شوند یا از مجموعهٔ داده‌ها حذف شوند.

● **دورافتاده‌هــا در میــان متغیرها**: یــک متغیر با مجــذور ضریــب هم‌بســتگی چندگانهٔ پاییــن با دیگر متغیرها و هم‌بســتگی پایین با همــهٔ فاکتورهای مهم، دورافتاده محسوب می‌شود و این متغیر باید از فهرست متغیرها حذف شود. (نظری و مختاری، ۱۳۸۸: ۲٤)

پایان قسمت اول