مجله علوم تغذیه و صنایع غذایی ایران
سال دهم، شماره 3، پاییز 1394، صفحات 46-39

مقایسه مدل آمیخته متناهی چندمتغیره با روش‌های تحلیل عاملی و خوشه‌ای در شناسایی الگوهای غذایی

زهرا رزاقی[1]، یداله محرابی[2]، فرید زایری[3]، آرش رشیدی[4]، بهارک آقاپور[5]

1- دانشجوی دکتری آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران
2- نویسنده مسئول: استاد گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران، پست الکترونیکی: mehrabi@sbmu.ac.ir
3- دانشیار گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران
4- استادیار گروه تحقیقات سیاستگذاری و برنامه ریزی غذا و تغذیه، انستیتو تحقیقات تغذیه ای و صنایع غذایی کشور، دانشکده تغذیه و صنایع غذایی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران
5- کارشناس ارشد تغذیه، شعبه بین الملل، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

تاریخ دریافت: 93/3/18 تاریخ پذیرش: 93/7/16

چکیده

سابقه و هدف: امروزه متخصصین علوم تغذیه جهت بررسی تأثیر کلی رژیم غذایی بر پیامد سلامتی از دیدگاه الگوی غذایی استفاده می‌کنند. الگوهای غذایی به طور معمول با دو روش آماری تحلیل عاملی یا تحلیل خوشه‌ای تعیین می‌شود. مدل‌های آمیخته در قالب ترکیبی از چند توزیع احتمالی یک چارچوب پارامتری برای مدل‌بندی اشکال توزیعی ناشناخته فراهم می‌کنند. اخیراً از مدل‌های آمیخته به عنوان روش سوم در تعیین الگوهای غذایی استفاده شده است بدین منظور علاوه بر دو روش موجود، در این مطالعه الگوهای غذایی را با استفاده از مدل آمیخته نیز به دست آورده و با دو روش دیگر مورد مقایسه قراردادیم.

مواد و روش‌ها: در این تحقیق از داده‌های 25 گروه غذایی 400 دختر دبیرستانی شهر اهر که در مطالعه رشیدی و همکاران به شیوه نمونه‌گیری خوشه‌ای جمع‌آوری شده بود استفاده شد. ابتدا الگوهای غذایی به دو روش تحلیل عاملی و تحلیل خوشه‌ای به دست آمد و بعد مدل آمیخته نرمال چند متغیره به این داده‌ها برازش شد و 5 الگوی غذایی به دست آمده از این روش با دو روش دیگر مورد مقایسه قرار گرفت. انتخاب مدل آمیخته مناسب با استفاده از معیار اطلاع آکاییک (AIC) و معیار اطلاع بیزی (BIC) انجام شد.

یافته‌ها: در روش تحلیل عاملی 3 الگوی غذایی که براساس بار عاملی شناسایی شد و در تحلیل خوشه‌ای 2 الگوی غذایی به دست آوردیم. با مدل آمیخته متناهی نرمال چند متغیره 5 الگوی غذایی به دست آمد که نسبت مصرف این 5 الگو به‌ترتیب 6%، 12%، 34%، 28% و 20% و هر الگو با بردار میانگین بیان شد.

نتیجه‌گیری: با بررسی سه روش نتیجه می‌شود که مدل آمیخته نسبت به دو روش دیگر دارای دو مزیت است. اولاً نسبت افراد جامعه که از هر الگو استفاده می‌کنند را تعیین کند و ثانیاً میانگین مصرف گروه‌های غذایی در هر الگو مشخص است. بنابراین اطلاعات بیشتری نسبت به دو روش معمول به دست خواهد آمد.

واژگان کلیدی: مدل آمیخته، تحلیل عاملی، الگوی غذایی

• مقدمه

امروزه متخصصین علوم تغذیه به منظور بررسی تأثیر کلی رژیم غذایی بر روی پیامد سلامتی از دیدگاه الگوی غذایی استفاده می‌کنند و چنین تحلیلی را به سایر تحلیل‌های سنتی در این زمینه (مثل غذاها، مواد مغذی و گروه‌های غذایی) اضافه می‌کنند. در آنالیز الگوهای غذایی از این ارتباط بین غذاهای مختلف می‌توان استفاده کرد، چون الگوهای غذایی بر مبنای عادات غذایی طراحی می‌شوند و مداخله‌های تغذیه‌ای اگر به صورت تغییر در الگوهای غذایی طراحی و اجرا شوند آسان‌تر و کامل‌تر خواهند بود (1).

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

اکثر مطالعات تحلیل الگوهای غذایی از دو روش تحلیل عاملی یا تحلیل خوشه ای استفاده کرده‌اند (2). روش تحلیل عاملی کاربرد بسیاری داشته و می‌توان گفت اکثر مطالعات تحلیل الگوهای غذایی امروزه براین اساس انجام می‌شود، چون بعد داده‌ها را کاهش داده و اطلاعات مربوط به تغذیه جامعه مورد بررسی را می‌تواند در قالب چند الگوی محدود بیان کند. مزیت این روش نسبت به دیگر روش‌های آماری زمانی است که علاقه‌مند به توصیف ابعاد مهم تغذیه‌ای باشیم.

در تحلیل خوشه‌ای، دسته بندی براساس گروه‌های دو به دو ناسازگار طبق شباهت‌ها یا عدم شباهت‌ها در مصرف غذایی تعیین می‌شود در این روش هر فرد فقط به یکی از الگوی غذایی اختصاص می‌یابد. مشکل این روش نیز متغیرهای کمکی است و همچنین مقایسه معیارهای متفاوت خوشه‌ها با یکدیگر آسان نیست (4، 3).

مدل‌های آمیخته در قالب ترکیبی از چند توزیع احتمال یک چارچوب پارامتری برای مدل بندی توزیع‌های ناشناخته فراهم می‌کنند. در دهه گذشته حوزه کاربردهای مدل آمیخته بطور قابل توجهی افزایش یافته است. از کاربردهای مدل‌های آمیخته می‌توان به خوشه‌بندی‌ها، تحلیل کلاس‌های پنهان، تحلیل ممیزی و تحلیل بقا اشاره نمود که در این موارد علاوه بر تحلیل و استنباط، توزیع مدل را نیز می‌توان توصیف کرد. توزیع آمیخته (Mixture distribution) به صورت ترکیبی از چند توزیع با نسبت‌های متفاوت به صورت زیر تعریف می‌شود.

$$f(y|\theta)=\sum_{k=1}^{K}\pi_k\, f_k(y|\theta_k)$$

در این فرمول، $f_k(y/\theta_k)$ شامل K توزیع با پارامترهای مختلف و π_k نسبت توزیع Kام در توزیع آمیخته است. پارامترهای $\pi_1 \dots \pi_k$ نشان دهنده‌ی نسبت کلی داده‌ها با هر کدام از چگالی‌های احتمال است که مجموعشان برابر یک بوده و نسبت آمیختگی نام دارند. این مقادیر نسبت هر الگوی غذایی را در جامعه مشخص می‌کند. $f_1 \dots f_k$ تابع چگالی احتمال هر الگوی غذایی می‌باشد که با برآورد پارامترهای میانگین $\mu_1, \mu_2, \dots, \mu_k$ و ماتریس واریانس کواریانس مشخص می‌شوند (5).

در مطالعه‌ای توسط Fahey و همکاران در سال 2007 از مدل آمیخته نرمال شرطی در تحلیل الگوهای غذایی داده‌های مطالعه آینده نگر اروپایی سرطان EPIC (European Prospective Investigation into Cancer) استفاده شد. در این مطالعه برای زنان چهار الگوی غذایی و برای مردان شش الگوی غذایی با استفاده از مدل‌های آمیخته به دست آمد. معیار ارزیابی مدل‌ها در این تحقیق معیار اطلاع بیز (BIC) بود (4).

همچنین Fahey و همکاران در سال 2012 از مدل آمیخته نرمال برای یافتن الگوهای غذایی استفاده کردند. در این مطالعه الگوهای غذایی 520000 زن از 10 کشور اروپایی که در مطالعات کوهورت 23 مرکز تحقیقاتی شرکت نموده بودند تعیین گردید و برای 24 گروه غذایی مورد بررسی 8 الگوی غذایی شناسایی شد. در این مطالعه منطقه زندگی، سن و میزان انرژی دریافتی نیز به عنوان متغیرهای کمکی در نظر گرفته شدند (3). در سال 2012 Rita Gaio و همکاران از مدل‌های آمیخته متناهی در تحلیل الگوهای غذایی در نمونه‌های کوچک استفاده کردند. ایشان با استفاده از شبیه‌سازی 900 مشاهده برای 14 گروه غذایی نشان دادند که استفاده از مدل آمیخته متناهی در تحلیل الگوهای غذایی در نمونه‌های کوچک به برآورد تعداد کمتری پارامتر نیاز دارد. همچنین در این مطالعه با استفاده مدل آمیخته متناهی 6 الگوی غذایی برای جامعه مورد بررسی شناسایی شد (6).

باتوجه به قابلیت مدل‌های آمیخته در تعیین یک توزیع با استفاده از ترکیب چند توزیع مشخص با نسبت‌های متفاوت، در این تحقیق از این مدل برای تعیین الگوهای غذایی استفاده و نتیجه را با تحلیل عاملی و تحلیل خوشه‌ای مقایسه کرده‌ایم. با استفاده از مدل‌های آمیخته هر الگوی غذایی به صورت یک کلاس پنهان با توزیع مشخص شناسایی شده و تعداد افراد مصرف کننده هر الگو در جامعه نیز برآورد شده است.

• مواد و روش‌ها

در این مطالعه از داده‌های 25 گروه غذایی 400 دختر دبیرستانی شهر اهر که در مطالعه رشیدی و همکاران با استفاده از پرسشنامه FFQ و به شیوه نمونه‌گیری خوشه‌ای جمع‌آوری شده بود استفاده شد (1). ابتدا الگوهای غذایی به دو روش تحلیل عاملی و تحلیل خوشه‌ای سلسله مراتبی محاسبه شد و سپس مدل آمیخته نرمال چند متغیره به داده‌ها برازش شده و از این روش نیز الگوهای غذایی را به دست آورده و با دو روش دیگر مورد مقایسه قراردادیم. در مدل آمیخته داده‌های مربوط به مصرف گروه‌های غذایی به‌صورت توزیع چندمتغیره پیوسته در نظر گرفته شدند. یعنی برای هر فرد مورد بررسی یک بردار از مشاهدات شامل 25 متغیر تعریف شد.

تحلیل خوشه‌ای و تحلیل عاملی با استفاده از نرم افزار SPSS نسخه 16 و برازش مدل آمیخته با استفاده از flexmix package در نرم‌افزار R نسخه 3 انجام شد. انتخاب مدل

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

مناسب با استفاده از معیار اطلاع آکاییک AIC (Akaike Information Criterion) و معیار اطلاع بیز BIC (Bayesian Information Criterion) صورت گرفت(10-7).

در تحلیل عاملی چون عامل‌ها براساس ماتریس واریانس-کوواریانس تعیین شدند از مقدار خام مصرف گروه‌های غذایی برحسب گرم در روز استفاده شد. به‌این دلیل که هر گروه غذایی شامل تعداد متفاوتی از مواد غذایی است و میانگین‌های مصرف گروه‌های غذایی باهم تفاوت زیادی دارند تحلیل خوشه‌ای و مدل آمیخته تحت تأثیر این تفاوت‌ها قرار می‌گیرند. برای حل این مشکل تحلیل خوشه‌ای و برازش مدل آمیخته با استفاده از داده‌های استاندارد شده انجام شد.

• یافته‌ها

در این مطالعه میانگین± انحراف معیار وزن، قد و شاخص توده بدنی افراد مورد بررسی به ترتیب برابر با 1/10±58/68 کیلوگرم، 6/40±161/00 سانتی‌متر و 3/98±22/63 کیلوگرم بر مترمربع بود. 25/8% از کل افراد دارای اضافه وزن و 8/1 % از آنها چاق بودند و در مجموع، شیوع کلی اضافه وزن یا چاقی در این مطالعه 33/9% بود. انحراف‌معیار±میانگین سن و انرژی دریافتی روزانه به ترتیب برابر با 1/02±15/70 سال و 747±2413 کیلوکالری در روز بود.

پس از ثبت مقدار مصرف هر ماده غذایی برای هر دانش آموز، ترکیبی از مواد غذایی مشابه را به عنوان گروه غذایی تعریف شد که در جدول 1 موادغذایی مربوط به هرگروه غذایی و در جدول 2 آماره‌های توصیفی مقدار مصرف آن‌ها برحسب گرم در روز در جامعه مورد بررسی بیان شده‌اند. تحلیل داده‌ها روی این 25 گروه غذایی انجام شده است.

جدول 1. مواد غذایی مربوط به هرگروه غذایی

	گروه غذایی	اقلام غذایی
1	گوشت قرمز یا فرآوری شده	گوشت گاو و گوساله، گوشت گوسفند، گوشت چرخ کرده، سوسیس، کالباس، همبرگر
2	گوشت احشاء	دل، جگر، قلوه، زبان، مغز، کله و پاچه، سیرابی و شیردان
3	ماهی	هر نوع ماهی، کنسرو ماهی
4	طیور و ماکیان	مرغ، جوجه
5	تخم مرغ	تخم مرغ
6	لبنیات کم چرب	شیر کم چرب، شیر بدون چربی، ماست کم چرب، ماست معمولی، پنیر سفید، کشک، دوغ، سایر
7	لبنیات پرچرب	شیر پرچرب، ماست پرچرب، پنیر خامه ای، خامه و سرشیر، بستنی، سایر
8	چای و قهوه	چای، قهوه و نسکافه
9	میوه و آبمیوه	طالبی، هندوانه، خربزه، گوجه سبز، سیب، زردآلو، آلو زرد و قرمز، گیلاس، آلبالو، شلیل، هلو، گلابی، انجیر تازه، خرما، انگور، کیوی، انار، توت فرنگی، موز، خرمالو، گرمک، توت تازه، آناناس، مرکبات، خشکبار،انواع آبمیوه طبیعی و صنعتی، سایر
10	سبزیجات	انواع کلم، هویج، گوجه فرنگی و فرآورده‌های آن، اسفناج، کاهو، خیار، بادمجان، پیاز، انواع سبزی، لوبیا سبز، نخود فرنگی، کدو خورشتی، قارچ، فلفل سبز و دلمه ای، شلغم، ذرت، سیر، سایر
11	حبوبات	عدس، لپه، لوبیا، نخود، باقلا، ماش، سویا، سایر
12	سیب زمینی آب پز	سیب زمینی آب پز
13	سیب زمینی سرخ شده	سیب زمینی سرخ شده
14	غلات کامل	نان‌های تیره (بربری، سنگک، تافتون)، نان سبوس دار، جو، بلغور، سایر
15	غلات تصفیه شده	نان لواش، نان باگت، برنج، ماکارونی، سایر
16	میان وعده ها	بیسکویت، پفک، چیپس، سایر
17	مغزها	بادام، بادام زمینی، گردو، پسته، فندق، انواع تخمه، سایر
18	شیرینی جات و دسرها	شیرینی خشک، شیرینی تر، شکلات، انواع کیک و کلوچه، عسل، مربا، قند، شکر، نبات، آبنبات، حلوا شکری، سایر
19	ترشی جات	ترشی، شور، خیار شور
20	چربی‌های جامد	روغن نباتی جامد، روغن حیوانی، کره حیوانی، مارگارین
21	روغن‌های مایع	انواع روغن‌های مایع به غیر از روغن زیتون
22	زیتون	زیتون، روغن زیتون
23	نوشابه	نوشابه
24	مایونز	مایونز
25	نمک	نمک

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

جدول2. آماره‌های توصیفی مصرف هر گروه غذایی برحسب گرم در روز

گروه غذایی	میانگین	میانه	دامنه میان چارکی
گوشت قرمز یا فرآوری شده	47/60	32/98	18/07-64/87
گوشت احشاء	1/50	0/11	0/00-1/29
ماهی	25/00	15/67	5/92-31/34
طیور و ماکیان	29/01	17/10	12/82-38/47
تخم مرغ	14/82	11/11	4/65-21/66
لبنیات کم چرب	598/51	528/68	327/79-808/73
لبنیات پرچرب	89/03	11/40	2/63-39/89
چای و قهوه	731/76	600/00	480/00-960/00
میوه و آبمیوه	962/41	857/28	670/20-1211/70
سبزیجات	615/20	512/76	372/40-790/30
حبوبات	25/29	14/53	6/18-30/16
سیب زمینی آب پز	24/22	9/21	0/00-32/89
سیب زمینی سرخ شده	2/49	0/25	0/00-3/21
غلات کامل	87/75	49/61	28/63-106/21
غلات تصفیه شده	145/55	116/94	85/00-190/10
میان وعده ها	9/98	2/26	0/00-11/11
مغزها	13/41	8/48	14/01-17/36
شیرینی جات و دسرها	42/95	31/64	16/79-49/90
ترشی جات	23/57	4/60	0/27-17/31
چربی‌های جامد	6/89	0/37	0/00-7/50
روغن‌های مایع	16/03	13/43	7/07-22/27
زیتون	6/06	4/95	3/00-7/50
نوشابه	84/04	9/21	0/00-42/74
مایونز	2/92	0/00	0/00-2/42
نمک	6/87	4/95	3/75-7/50

در روش تحلیل عاملی با استفاده از نمودار Scree، 3 الگوی غذایی که در مجموع 29/6% از کل واریانس داده‌ها را در جمعیت مورد بررسی توجیه می‌کردند انتخاب شد. الگوی غذایی اول یا الگوی ناسالم با مصرف بالای مایونز، غلات تصفیه شده، سیب زمینی سرخ شده، غلات کامل، لبنیات پرچرب، گوشت احشاء و گوشت قرمز یا فرآوری شده مشخص می‌شد.

ویژگی الگوی غذایی دوم یا الگوی سالم عبارت بود از مصرف بالای سبزیجات، حبوبات، سیب زمینی آب‌پز، میوه و آبمیوه، روغن‌های مایع، زیتون، ماهی، لبنیات کم چرب و طیور و ماکیان. در نهایت، الگوی غذایی سوم یا الگوی چربی‌های جامد و نمک نیز با مصرف بالای چربی‌های جامد، نمک، میان وعده‌ها، ترشی‌جات و شیرینی‌جات و دسرها مشخص می‌شد. نام‌گذاری الگوهای اول و دوم به صورت ناسالم و سالم بر اساس مطالعات پیشین و نام‌گذاری الگوی سوم به صورت چربی‌های جامد و نمک بر پایه گروه‌های غذایی بود که بیشترین بار عاملی را در این الگو داشتند.

با خوشه‌بندی متغیرها دو خوشه به دست آوردیم که گروه‌های غذایی مربوط به هر خوشه در جدول 3 مشخص شده‌اند. خوشه‌بندی دو دسته مجزا از گروه‌های غذایی را ارائه می‌دهد و در واقع یک دسته بندی جهت کاهش بعد داده‌ها انجام شده است (12، 11).

در برازش مدل آمیخته، با توجه به معیار اطلاع BIC با 5 مؤلفه به عنوان بهترین مدل آمیخته در تعیین الگوهای غذایی انتخاب شد. نسبت مؤلفه‌های مدل و میانگین مقداراستاندارد شده هر گروه غذایی در مدل در جدول 4 مشخص شده است. بردار میانگین‌های مقدار استاندار شده مصرف گروه‌های غذایی هر الگو در نمودارهای جداگانه شکل 1 رسم شده است. در محور افقی گروه غذایی و در محور عمودی میانگین استاندارد شده مصرف هر مواد غذایی را نشان می‌دهد.

جدول3. مواد غذایی مربوط به دو الگوی حاصل از تحلیل خوشه‌ای

خوشه‌ها	گروه‌های غذایی هر خوشه
خوشه اول	ماهی، مغزها، سیب‌زمینی، زیتون، سبزیجات، روغن مایع، لبنیات کم‌چرب، میوه و آبمیوه، طیور وماکیان
خوشه دوم	گوشت قرمز یا فرآوری شده، گوشت احشاء، لبنیات پرچرب، تخم‌مرغ، چای و قهوه، شیرینی‌جات و دسرها، ترشیجات، چربی جامد، میان وعده، نمک، حبوبات، غلات کامل، سیب زمینی سرخ‌کرده، مایونز، غلات تصفیه شده، نوشابه

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

جدول 4. میانگین مقادیر استاندارد شده مصرف گروه‌های غذایی در 5 الگوی مدل آمیخته نرمال

گروه‌های غذایی	الگوی اول	الگوی دوم	الگوی سوم	الگوی چهارم	الگوی پنجم
نسبت هر الگو در جامعه	0/06	0/12	0/34	0/28	0/20
گوشت قرمز یا فرآوری شده	-0/167	0/094	0/189	0/587	-0/450
گوشت احشاء	-0/367	-0/312	-0/319	1/422	-0/136
ماهی	-0/397	-0/003	-0/269	0/576	0/178
طیور و ماکیان	0/041	-0/227	-0/317	0/875	-0/110
تخم مرغ	-0/063	-0/030	-0/463	0/612	0/162
لبنیات کم چرب	-0/058	-0/229	0/079	-0/213	0/168
لبنیات پرچرب	0/078	-0/017	-0/308	0/856	-0/214
چای و قهوه	0/116	0/460	-0/312	0/311	-0/123
میوه و آبمیوه	0/416	0/667	-0/167	0/422	0/096
سبزیجات	-0/389	0/746	-0/205	0/573	-0/192
حبوبات	-0/438	0/782	-0/326	0/712	-0/136
سیب زمینی آب پز	-0/154	0/034	-0/361	0/047	0/403
سیب زمینی سرخ شده	0/066	0/423	-0/249	0/703	-0/358
غلات کامل	-0/250	0/293	-0/088	0/599	-0/217
غلات تصفیه شده	-0/260	0/779	-0/001	0/211	-0/284
میان وعده‌ها	0/634	0/432	-0/202	0/189	-0/457
مغزها	-0/719	0/274	-0/431	0/003	-0/117
شیرینی‌جات و دسرها	-0/159	1/423	-0/284	0/153	-0/219
ترشی‌جات	-0/365	0/304	0/072	0/529	-0/272
چربی‌های جامد	0/185	-0/085	-0/137	0/894	-0/438
روغن‌های مایع	-0/105	0/344	0/036	0/133	-0/187
زیتون	-0/032	-0/407	-0/129	0/233	0/180
نوشابه	0/122	-0/066	0/035	0/481	-0/347
مایونز	-0/011	1/271	-0/233	0/140	-0/359
نمک	0/441	-0/208	-0/244	0/343	-0/128

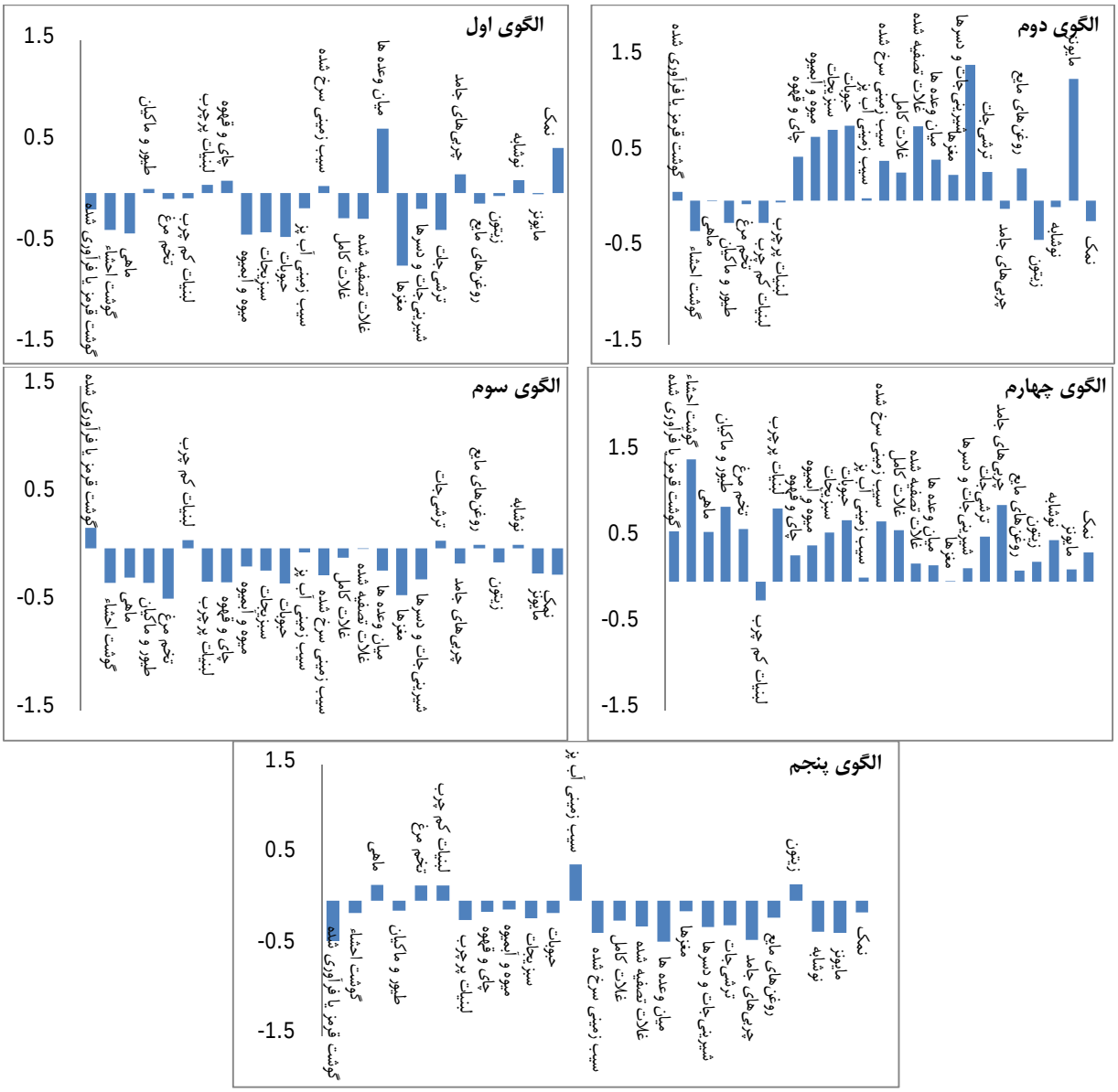
الگوی اول کمترین میزان مصرف (6%) را در کل جامعه مورد بررسی داشته است. در این الگو بیشترین میزان مصرف میان‌وعده و نمک و کمترین مصرف مغزها، حبوبات، میوه، سبزیجات و ماهی می‌باشند. در این الگو به طور کلی میزان مصرف اکثر گروه‌های غذایی نزدیک به متوسط است یعنی افراد از همه‌ی غذاها تقریباً به یک نسبت استفاده کرده‌اند.

در الگوی دوم بیشترین مصرف شیرینی‌جات و مایونز و کمترین مصرف زیتون و گوشت احشاء می‌باشند. در این الگو مصرف گروه‌های غذایی بجز گوشت احشاء، ماهی، طیور و ماکیان، تخم مرغ، لبنیات کم چرب و پرچرب، چربی‌های جامد، زیتون، نوشابه و نمک بسیار بیشتر از میانگین کل جامعه بوده است. الگوی سوم بیشترین نسبت مصرف را داشته است (34%) در الگوی سوم بیشترین مصرف گوشت قرمز یا فرآوری شده و لبنیات کم‌چرب و ترشیجات بوده و کمترین مصرف تخم مرغ و مغزها می‌باشند بطور کلی میزان مصرف اکثر گروه‌های غذایی در این الگو کمتر از میانگین مصرف جامعه است. در الگو چهارم بیشترین مصرف گوشت احشاء، طیور و ماکیان، چربی‌های جامد و لبنیات پرچرب بوده و کمترین مصرف لبنیات کم‌چرب می‌باشند در این الگو میزان مصرف همه گروه‌های غذایی بجز لبنیات کم‌چرب بیشتر از میانگین بوده است. نسبت مصرف این الگو به الگوی غذایی شماره 3 تقریباً نزدیک است در حالی‌که نوع مصرف گروه‌های غذایی برعکس می‌باشد.

در الگوی پنجم بیشترین مصرف سیب زمینی بوده و کمترین مصرف میان وعده، چربی‌های جامد و گوشت قرمز یا فرآوری شده می‌باشند. در این الگو نیز مانند الگوی 3 مصرف گروه‌های غذایی کمتر از میانگین جامعه بوده است.

در مدل آمیخته به دست آمده در الگوهای 2 و 4 میزان مصرف اکثر گروه‌های غذایی کمتر از متوسط جامعه است در حالی‌که در الگوهای 1، 3 و 5 این مقادیر بیشتر از متوسط جامعه است.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]



شکل 1. بردار میانگین‌های مقادیر استانداردشده مصرف گروه‌های غذایی پنج الگوی شناسایی شده با مدل آمیخته

• بحث

در مدل آمیخته متناهی با افزایش تعداد مؤلفه‌ها مقدار AIC و BIC کمتر می‌شود، تا جایی‌که به یک مقدار ثابتی از تغییرات برسد آن مدل را به عنوان مناسب ترین مدل انتخاب می‌کنیم، ولی در روش تحلیل عاملی از نسبت مجموع واریانس مؤلفه‌های در نظر گرفته شده به واریانس کل استفاده می‌نماییم که مقدار بیشتر از 70% مطلوب می‌باشد، ولی در داده‌های تغذیه با توجه به زیاد بودن تعداد مواد غذایی و گروه‌های غذایی تعداد مؤلفه‌های روش تحلیل عاملی نیز زیاد بوده که این در تفسیر پیچیده است بدین منظور تعداد کمتری مؤلفه در نظر گرفته شده ولی نسبت واریانس آن‌ها بسیار کمتر از 70% می‌باشد، در مدل آمیخته برای هر الگو نسبت آن در جامعه نیز به دست آمده است (9-6) که در دو روش دیگر چنین پارامتری وجود ندارد، در تحلیل عاملی الگوهای غذایی براساس بیشترین بار عاملی هر گروه غذایی با آن مؤلفه (یا الگو) به توصیف آن الگوی غذایی می‌پردازیم که در مدل آمیخته متناهی نرمال هر مؤلفه براساس مقدار میانگین ماده غذایی توصیف می‌گردد، همچنین بار عاملی در مورد میزان مصرف اطلاعاتی ندارد در حالی‌که میانگین‌های مقدار مصرف گروه‌های غذایی در هر مؤلفه در بردارنده اطلاعات کافی می‌باشد، از طرفی می‌توان آمیخته متناهی را تأییدی از تحلیل عاملی دانست که هر مؤلفه به صورت یک عامل پنهان جایگزین نمود (13)

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

در تحلیل خوشه‌ای تعداد خوشه‌ها به صورت تجربی انتخاب می‌شود و یک خوشه شامل تعدادی از گروه‌های غذایی است که این گروه‌های غذایی در خوشه دیگر وجود ندارند در حالیکه در مدل آمیخته هر مؤلفه شامل تمام گروه‌های غذایی است یعنی افرادی که در این مؤلفه هستند توزیع مصرف گروه‌های غذایی‌شان با مؤلفه‌ها دیگر متفاوت است که این تفاوت با مقایسه میانگین‌ها مشخص می شود، در عمل مدل آمیخته منطقی‌تر به نظر می‌رسد چون برای هر فرد مصرف مواد غذایی به صورت توام می‌باشد.

از طرفی تحلیل خوشه‌ای یک روش قدرتمند برای یک دسته بندی نامشخص است اما پیداکردن یک خوشه بندی نهایی نیازمند چندین دسته بندی در هر مرحله است بطوریکه اطلاعات پیشینی از داده‌ها را در نظر نمی‌گیرد (8) و همچنین در خوشه بندی در مورد نسبت مقدار مصرف هر الگو اطلاعی نداریم. از طرفی در مدل آمیخته گروه‌های غذایی در بین الگوهای به دست آمده تقسیم نمی‌شوند و این در عمل روشن است چون نمی‌توان گفت فردی از یک نوع غذای خاص مصرف نمی‌نماید. در حالیکه مؤلفه‌ها (الگوها) از هم مجزا هستند که این خاصیت اجتماعی از مزیت‌های دو روش تحلیل عاملی و تحلیل خوشه‌ای است. یکی دیگر از مزایای مدل آمیخته دسته‌بندی داده‌ها همراه با ارائه توزیع آن‌ها می‌باشد.

در مطالعه Gaio و همکاران 6 الگو برای 14 گروه غذایی به دست آمده که مقدار میانگین مصرف‌ها در شکل مشخص شده است وچون داده‌ها استاندارد نشده‌اند میزان مصرف در هر 6 الگو شبیه به هم می‌باشد. در حالی که ما با استاندارد کردن مقادیر متفاوتی از مصرف در الگوهای غذایی به دست آوردیم. غذاهایی مانند گوشت قرمز ، گوشت احشاء، لبنیات کم چرب و چربی‌های جامد در یک الگو کمترین مصرف و در الگویی دیگر بیشترین مصرف را داشت.

در مطالعه Fahey و همکاران در سال 2012 ، 8 الگوی غذایی برای 24 گروه غذایی به دست آمده است و نسبت هر مؤلفه نیز براساس منطقه محل زندگی نیز به دست آمد اما در مطالعه ما چون همه‌ی افراد در یک شهر زندگی می‌کردند یافتن این نسبت امکان‌پذیر نشد. در مدل آمیخته برآورد پارامترها نسبت به دو روش تحلیل دیگر بیشتر است و این کار با داده‌های کم مشکل ساز خواهد بود (14، 13).

با بررسی سه روش نتیجه می‌شود الگوهای به دست آمده از مدل آمیخته و تحلیل عاملی اطلاعات بیشتری نسبت به روش تحلیل خوشه‌ای ارائه می‌دهند و مدل آمیخته نیز نسبت به تحلیل عاملی کامل‌تر است. می‌توان گفت مدل آمیخته با دسته بندی افراد در مؤلفه‌ها مانند خوشه‌بندی عمل می‌کند و از طرفی مانند تحلیل عاملی هر مؤلفه به عنوان یک الگوی غذایی شامل تمام گروه‌های غذایی بوده که با میانگین‌های متفاوت مقدار مصرف هر گروه غذایی تفسیر می‌گردد.

سپاسگزاری

به‌این وسیله از شرکت کنندگان در این مطالعه و مسئولین محترم دبیرستان‌های دخترانه شهر اهر قدردانی می‌شود. این مقاله برگرفته از پایان‌نامه کارشناسی ارشد آمار زیستی است.

References

1. Aghapour B. The association of overweight and obesity with dietary patterns and some socio-economic factors among high school girl students in Ahar, Iran[dissertation]. Tehran: The International branch of Shahid Beheshti University of medical sciences, M.C. Faculty of Nutrition Sciences and Food Technology; 2013 [in Persian].
2. Esmaillzadeh A, Kimiagar M, Mehrabi Y, Azadbakht L, B Hu F, Willett WC. Dietary patterns, insulin resistance, and prevalence of the metabolic syndrome in women. Am J Clin Nutr 2007;85:910–8.
3. Fahey MT, Ferrari P, Slimani N, Vermunt JK, White IR, Hoffmann K, et al. Identifying dietary patterns using a normal mixture model: application to the EPIC study. J Epidemiol Community Health 2012; 66(1):89-94.
4. Fahey MT, Thane CW, Bramwell GD, Coward WA. Conditional gaussian mixture modelling for dietary pattern analysis. J. R. Statist Soc 2007; 170:149-66.
5. McLachlan G, PEEL D.. Finite Mixture models. 1st Edition. New York: John Wiley & Son; 2001.
6. Gaio AR, da Costa JP, Santos AC, Ramosb E, Lopesb C. A restricted mixture model for dietary pattern analysis in small samples. Statistic in medicine 2012; 31:2137-50.
7. Basford K.E, McLachlan G.J. Likelihood Estimation with Normal Mixture Models. Journal of the Royal Statistical Society 1985; 34(3):282-89.
8. Rencher AC. Methods of multivariate analysis. 2nd ed. WILEY. New York; 2002.
9. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th ed. Pearson; 2007.
10. Azar M, Sarkisian E. Food composition table of Iran. National Nutrition and Food Research Institute, Shahid Beheshti University. Tehran [in Persian].
11. Hu F.B. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002; 13:3-9.
12. Esmaillzadeh A, Kimiagar M, Mehrabi Y, Azadbakht L, Hu FB, Willett WC. Dietary Patterns and Markers of Systemic Inflammation among Iranian Women. J Nutr 2007; 137:992–998.
13. Yung YF. Finite mixture in confirmatory factor-analysis models. Sychometrika 1997; 62(3):297-330.
14. Topchy A, Jain A.K, Punch W . Mixture Model for clustering Ensemble. SIAM International conference on data mining;2008 April 24-26; Atlanta, Georgia, USA.

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]

Comparison of the results of Multivariate Finite Mixture Model with Factor and Cluster Analysis Methods in Dietary Pattern Identification

*Razzaghi Z[1], Mehrabi Y[*2], Zaeri F[3], Rashidi A[4], Aghapour B[5]*

1- PhD student of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Science, Tehran, Iran

*2-*Corresponding author: Prof, Faculty of Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran, Email: mehrabi@sbmu.ac.ir*

3- Associate Prof, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Medical Sciences. Tehran , Iran

4- Assistant Prof, Dept. of Food and Nutrition Policy and Planning Research, National Nutrition and Food Technology Research Institute, Faculty of Nutrition Sciences and Food Technology, Shahid Beheshti University of Medical Sciences, Tehran, Iran

5- M.Sc in Nutrition Science, International Branch, Shahid Beheshti University of Medical Science, Tehran, Iran

Received 8 Jun, 2014 **Accepted 8 Oct, 2014**

Background and Objectives: Today nutritionists use dietary pattern to find out the effect of food in health. Most common statistical methods to determine dietary pattern are factor analysis and cluster analysis.Mixture models are a combination of k probability distribution with different probability and provide a parametric model for unknown distributional shapes. Recently mixture model as the third method is used to determine dietary pattern. Then we compared this new method with other two methods available for this purpose.

Materials & Methods: We analyzed data from 25 food groups of 400 high school girls in Ahar (Rashidi's data), and compared the results of factor analysis, cluster analysis and multivariate normal mixture model for dietary pattern. Selection of the best mixture model was done by AIC and BIC criteria.

Results: Three, two and five dietary pattern were obtained from factor analysis, cluster analysis and normal mixture model, respectively. Prevalence of these dietary patterns in normal mixture model was 6%,12%,34%,28%, and 20%, respectively.

Conclusion: It is concluded that mixture model has two advantages over the two other methods. First, the proportion of each pattern in population is known and secondly, the average of consumption of each food group gets clear; so more information can be obtained compared to the usual methods.

Keywords: Mixture model, Factor analysis, Cluster analysis, Dietary pattern

[Downloaded from nsft.sbmu.ac.ir on 2026-11-14]