مجله تخصصی اپیدمیولوژی ایران؛ تابستان ۱۳۹۶، دوره ۱۳، شماره ۲: صفحات ۹۷-۹۰.

مقاله پژوهشی

# روش برآوردگر ذاتی برای کنترل اثر سن، همگروه و دوره زمانی و مقایسه آن با مدل کلاسیک توصیفی

رویا صاحبی[۱]، سید عباس متولیان[۲]، لیلا صاحبی[۳]، حمید شریفی[۴]

[۱] کارشناس ارشد اپیدمیولوژی، دانشگاه علوم پزشکی تبریز، تبریز، ایران

[۲] دانشیار، دکترای تخصصی اپیدمیولوژی، عضو هیئت علمی دانشگاه علوم پزشکی ایران، تهران، ایران

[۳] دکترای اپیدمیولوژی، مرکز تحقیقات مادر، جنین و نوزاد تهران، دانشگاه علوم پزشکی تهران، تهران، ایران

[۴] دانشیار، دکترای تخصصی اپیدمیولوژی، مرکز تحقیقات مراقبت اچ آی وی و بیماری های آمیزشی، مرکز همکار سازمان جهانی بهداشت، پژوهشکده آینده‌پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران

نویسنده رابط: حمید شریفی، آدرس: مرکز تحقیقات مراقبت اچ آی وی و بیماری‌های آمیزشی، مرکز همکار سازمان جهانی بهداشت، پژوهشکده آینده پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران، پست الکترونیک: hsharifi@kmu.ac.ir

تاریخ دریافت: ۹۵/۳/۲۳؛ پذیرش: ۹۵/۹/۶

**مقدمه و اهداف:** هدف از انجام این مطالعه مقایسه روش برآوردگر ذاتی با روش توصیفی سن، دوره و همگروه است که برای تشخیص روندهای سن، هم‌گروه و دوره زمانی در میزان‌های بروز، شیوع و میرایی استفاده می‌شود.

**روش کار:** برای شرح روش کار از داده‌های مربوط به دو مطالعه بررسی روند مصرف مواد غیرقانونی و الکل و روند مصرف سیگار در دانشجویان دختر دانشگاه علوم پزشکی تهران از سال ۸۵ تا ۸۸ استفاده شده است. ابتدا مدل توصیفی سن، همگروه و دوره و سپس مدل تحلیلی برآوردگر ذاتی نشان داده‌شده و مزایا و معایب آن‌ها ذکرشده و نتایج آن‌ها مقایسه شده است.

**یافته‌ها:** برای شیوع مصرف سیگار، هر دو روش برای اثر سن، همگروه و دوره زمانی یک روند افزایشی، کاهشی و افزایشی نشان دادند. روش برآوردگر ذاتی توانست اثر پیشگیری کننده عامل سنین پایین در مصرف سیگار را نیز نشان دهد. با مشاهده نمودارهای روش توصیفی همگروه در مصرف الکل برای دختران دانشجو، تقریباً در همه گروه‌های سنی، از هم‌گروه‌های قدیمی‌تر به جدیدتر یک روند کاهشی دیده شد. درحالی‌که در روش برآوردگر ذاتی، اثر همگروهی برای مصرف الکل دختران دانشجو دیده نشد.

**نتیجه‌گیری:** به نظر می‌رسد در حال حاضر روش برآوردگر ذاتی بهترین انتخاب برای حل مشکل وابستگی خطی بین سه عامل سن، همگروه تولد و دوره زمانی در این مطالعات است. ولی محققین بایستی با احتیاط و آگاهی از مشکلات و تله‌های ایجادشده توسط این نوع مطالعات از آن استفاده کنند.

**واژگان کلیدی:** اثر سن، دوره، همگروه، روش برآوردگر ذاتی

## مقدمه

مدل توصیفی سن، همگروه و دوره (سال تقویمی) مدت‌ها در مطالعات اپیدمیولوژیک در بین جمعیت‌های انسانی موردتوجه بوده است. این مدل زمانی که هر سه عامل سن، دوره و همگروه موردتوجه است، برای تحلیل همگروه‌ها به کار می‌رود و یکی از ابزارهای موردعلاقه در علوم اپیدمیولوژی و جمعیت‌شناسی، برای تشخیص روندهای سن، هم‌گروه و دوره در میزان‌های بروز یا میرایی ناشی از بیماری‌هاست (۱) که البته در علوم دیگری نظیر علوم اجتماعی، روان‌شناسی، سیاسی و اقتصاد هم کاربرد دارد (۲). اثر سن تفاوت‌هایی را در گروهی از افراد نشان می‌دهد که ناشی از افزایش سن یا همان پیر شدن است. اثر دوره ناشی از تغییرات اجتماعی یا تاریخی است که بر تمام گروه‌های سنی یک گروه یا جمعیتی به‌طور یکسان اثر می‌کند مانند رکود اقتصادی، جنگ، یا مواجهه با یک کارسینوژن (۵-۳). اثر همگروه به گروهی از افراد اشاره می‌کند که رویداد مشترکی را در سنین یکسانی تجربه می‌کنند مانند همگروه تولد (۳).

استفاده از مدل توصیفی سن، دوره و همگروه در ارزیابی اثر همگروه تولد به‌عنوان یکی از روش‌های جدول عمر با ویلیام فار در سال ۱۸۸۵ آغاز شد که نقش بسزایی در پیشرفت علوم اجتماعی داشت (۶). فراست (۱۹۳۹) نیز یکی از اولین استفاده‌کننده‌های این روش بود که با ترسیم میزان‌های اختصاصی سنی مرگ از بیماری سل بر اساس همگروه تولد و سن، میزان‌های متفاوتی از مرگ‌ومیر ناشی از سل به دست آورد (۷). به دنبال استفاده از تحلیل سن، همگروه و دوره در مطالعه کلاسیک سل توسط فراست، این مدل یکی از روش‌های موردعلاقه اپیدمیولوژیست‌ها

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

شد و پس‌ازآن بار دیگر جهت مطالعه درزمینه‌ی بروز یا میرایی انواع سرطان‌ها از آن استفاده شد. ولی علاوه بر روش توصیفی که در آن صرفاً از نمودار استفاده می‌شد از روش‌های آماری با مدل رگرسیون برای کمیت دادن به اثر این سه عامل استفاده شد (۲). مشکلی که در استفاده از این روش وجود دارد، وابستگی خطی بین سه متغیر سن، همگروه تولد و دوره است که برای مقابله با آن ماسون و همکاران (۱۹۷۳) مدلی به نام مدل خطی تعمیم‌یافته محدود به‌عنوان میانگین اثر مستقل سن، دوره و همگروه را پیشنهاد کردند (۵). کوپر و همکاران (۱۹۸۳) برآوردگر ترکیبی اصلی را برای حل مشکل معرفی کردند. فو (۲۰۰۰) نیز روش جدیدی را معرفی کرد که برآوردگر ذاتی نام داشت (۸). یانگ و همکاران (۲۰۰۸-۲۰۰۴) در چند مقاله آموزشی متد برآوردگر ذاتی را توصیف و ارزیابی نمودند و آن را با مدل خطی تعمیم‌یافته محدود مقایسه کردند و روش برآوردگر ذاتی را یک روش بدون تورش و با خصوصیات آماری مطلوب نسبت به روش مدل خطی تعمیم‌یافته، معرفی کردند (۱۰-۹، ۱). یانگ (۲۰۰۸) همچنین در مطالعه‌ای روند میرایی از بیماری‌های مزمن (۱۹۹۹-۱۹۶۰) را با روش مزبور بررسی کرد و نشان داد که روند میرایی از سال ۱۹۶۰ تا سال ۱۹۹۰ کاهش قابل‌ملاحظه‌ای یافته که این کاهش به‌طور عمده‌ای مربوط به اثر همگروه بوده است و اثر دوره با کنترل اثر سن و همگروه تولد به‌طورکلی کوچک بوده است (۱۱). مدل برآوردگر ذاتی که با روش‌های ریاضی (ماتریکس) و هندسی اثبات شده است، روشی است که در مدل‌های رگرسیون خطی مرسوم زمانی که متغیرها ارتباط خطی قوی با هم دارند، به کار می‌رود (۸). هدف از انجام این مطالعه مقایسه روش برآوردگر ذاتی و روش توصیفی قدیمی است.

### مدل سن، دوره و همگروه

اگر داده‌های فرضی از شیوع بیماری Y را در ارتباط با سن در یک مطالعه مقطعی در سال ۲۰۰۵ را بدون در نظر گرفتن سال‌های قبلی در نظر بگیریم (جدول ۱، ستون آخر). این داده‌ها نشان می‌دهند که بیماری با افزایش سن کاهش می‌یابد. درحالی‌که اگر ما داده‌های ۱۹۷۵، ۱۹۸۵ و ۱۹۹۵ را هم داشته باشیم و داده‌های هر همگروه را از بالا به پایین نگاه کنیم روند مقطعی را خواهیم دید. اما برای مثال اگر گروهی که در سال۱۹۷۵ در گروه سنی ۲۹-۲۰ سال قرار دارد را در نظر بگیریم و آنها را در طول زمان تعقیب کنیم(این گروه سنی در سال ۱۹۸۵ در گروه سنی ۳۹-۳۰ و در سال ۱۹۹۵ در گروه سنی ۴۹-۴۰ و در سال۲۰۰۵ در گروه سنی ۵۹-۵۰ سال خواهند بود)، خواهیم دید که با افزایش سن، شیوع بیماری هم افزایش می یابد. درواقع نتیجه معکوس که از بررسی داده‌ها به‌صورت مقطعی به دست می‌آید به دلیل اثر قوی همگروه بوده است. در این داده‌های فرضی دو اثر قوی وجود داشت، اثر همگروه و اثر سن، به‌طوری‌که همگروه‌های جوان‌تر میزان‌های بالاتری داشتند. اثر همگروه (داشتن میزان‌های بالاتر در همگروه‌های جوان‌تر) افزایش شیوع مرتبط با سن را می‌پوشاند و الگوی بررسی مقطعی را نشان می‌دهد.

**جدول شماره ۱**– داده‌های فرضی از مطالعات مقطعی متواتر از شیوع بیماری Y در یک جمعیت بر اساس سن و دوره در سال‌های ۱۹۷۵، ۱۹۸۵، ۱۹۹۵ و ۲۰۰۵(۵)

| گروه سنی (سال) | میانگین سن | ۱۹۷۵ | ۱۹۸۵ | ۱۹۹۵ | ۲۰۰۵ |
|---|---|---|---|---|---|
| ۱۰-۱۹ | ۱۵ | ۱۷ | ۲۸ | | |
| ۲۰-۲۹ | ۲۵ | ۱۴ | ۲۳ | ۳۵ | |
| ۳۰-۳۹ | ۳۵ | ۱۲ | ۱۹ | ۳۰ | ۴۵ |
| ۴۰-۴۹ | ۴۵ | ۱۰ | ۱۸ | ۲۶ | ۴۰ |
| ۵۰-۵۹ | ۵۵ | | ۱۵ | ۲۲ | ۳۶ |
| ۶۰-۶۹ | ۶۵ | | | ۲۰ | ۳۱ |
| ۷۰-۷۹ | ۷۵ | | | | ۲۷ |

## روش کار

در مطالعه‌ای که به‌عنوان‌مثال در این مطالعه استفاده‌شده است، تحلیل ۴ مطالعه مقطعی پشت سر هم در سال‌های ۸۵، ۸۶، ۸۷ و ۸۸ به روش تحلیل سن، همگروه و دوره برای بررسی مصرف الکل، سیگار و مواد غیرقانونی انجام‌شده است. جمعیت موردمطالعه تمامی دانشجویان رشته‌های کارشناسی پیوسته و دکترای حرفه‌ای دانشگاه علوم پزشکی تهران از سال‌های ۸۵ تا ۸۸ و به‌صورت سرشماری و ابزار گردآوری پرسشنامه بوده است. روش کار به‌طور کامل در مقاله‌های مطالعه مزبور آمده است (۱۲،۱۳).

### روش کار در مدل کلاسیک (توصیفی) سن، همگروه و دوره

در مدل کلاسیک سن، همگروه و دوره که صرفاً از جداول و نمودار و منحنی‌ها برای نشان دادن روندهای سن، دوره و همگروه استفاده می‌شود، قضاوت در روندها صرفاً از طریق مشاهده است.

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

### مرحله اول: واردکردن و تجزیه‌وتحلیل داده‌ها در روش کلاسیک سن، همگروه و دوره

در این روش برای واردکردن داده‌ها می‌توان از نرم‌افزار اکسل استفاده کرد. در صورت داشتن هر یک از متغیرهای سن یا همگروه و با استفاده از دوره می‌توان متغیر دیگر را محاسبه کرد (همگروه = دوره – سن). متغیرهای سن و همگروه برای هر دوره و میزان محاسبه‌شده برای هر همگروه و دوره، وارد نرم‌افزار می‌گردد. برای درک آسان ورود و تجزیه تحلیل داده‌ها، به ذکر یک مثال عملی می‌پردازیم. در مطالعه مربوطه (۱۲) که در زمینه اثر سن، دوره و همگروه تولد در روند مصرف سیگار در دانشجویان دانشگاه تهران طی سال‌های ۸۸-۸۵ انجام شد، ابتدا متغیرهای مستقل شامل دوره (سال‌های ۱۳۸۵ تا ۱۳۸۸)، سن (۱۸ تا ۲۳ سال)، و همگروه تولد (سال‌های ۱۳۶۲ تا ۱۳۷۰) وارد نرم‌افزار شدند. جهت محاسبه شیوع مصرف سیگار با استفاده از نرم‌افزار اکسل، تعداد افراد هر گروه سنی که حداقل یک نخ سیگار در یک سال گذشته مصرف کرده بودند و تعداد کل افراد همان گروه سنی در طی یک دوره، وارد نرم‌افزار شدند (جدول شماره ۲). این جدول ۸ ستون دارد که ۶ ستون آن وارد نرم‌افزار اکسل شد. ستون ۷ (نسبت مصرف سیگار) توسط اکسل و با استفاده از فرمول نویسی محاسبه‌شده است که اعداد مربوط در آن گرد شده‌اند. ستون هشتم جدول درصد مصرف سیگار یا شیوع مصرف سیگار را نشان می‌دهد که به‌صورت دستی و برای تنظیم جدول ۳ واردشده است.

برای داده‌های مردان هم می‌توان جدول دیگری نظیر جدول ۲ تهیه و داده‌های هر دو جنس را با هم مقایسه کرد. مقایسه شیوع مصرف سیگار، الکل و مواد در دو جنس قبلاً در دو مقاله دیگر آمده است (۱۳،۱۲). در این مقاله فقط داده‌های مربوط به دختران دانشجو نشان داده‌شده است.

### مرحله دوم: تنظیم جدول با استفاده از داده‌های واردشده به‌منظور رسم نمودار

پس از تنظیم جدول ۲، نسبت‌های شیوع به‌دست‌آمده از جدول ۲ را بر اساس سن - دوره، همگروه – دوره و یا سن – همگروه، در یک جدول دیگر و صرفاً جهت رسم نمودار مدل سن، دوره و همگروه در نرم‌افزار اکسل، وارد می‌کنیم. جدول ۳- یک نمونه از جدول تنظیمی برای شیوع سالانه مصرف سیگار در دختران دانشجو بر اساس سن - دوره را نشان می‌دهد که از داده‌های موجود از جدول شماره ۲ به‌دست‌آمده است.

**جدول شماره ۲**- داده‌های شیوع مصرف سیگار در دانشجویان دختر در دانشگاه تهران از سال ۱۳۸۵ تا ۱۳۸۸

| ردیف | دوره | سن | همگروه | تعداد افراد سیگاری | کل افراد | نسبت مصرف سیگار | درصد (شیوع) مصرف سیگار |
|---|---|---|---|---|---|---|---|
| ۱ | ۱۳۸۵ | ۱۸ | ۱۳۶۷ | ۴ | ۱۸۰ | ۰/۰۲۲ | ۲/۲ |
| ۲ | ۱۳۸۵ | ۱۹ | ۱۳۶۶ | ۱۲ | ۲۵۸ | ۰/۰۴۷ | ۴/۷ |
| ۳ | ۱۳۸۵ | ۲۰ | ۱۳۶۵ | ۱۵ | ۲۵۲ | ۰/۰۶ | ۶/۰ |
| ۴ | ۱۳۸۵ | ۲۱ | ۱۳۶۴ | ۱۷ | ۲۱۲ | ۰/۰۸ | ۸/۰ |
| ۵ | ۱۳۸۵ | ۲۲ | ۱۳۶۳ | ۱۱ | ۱۲۹ | ۰/۰۸۵ | ۸/۵ |
| ۶ | ۱۳۸۵ | ۲۳ | ۱۳۶۲ | ۷ | ۴۰ | ۰/۱۷۵ | ۱۷/۵ |
| ۷ | ۱۳۸۶ | ۱۸ | ۱۳۶۸ | ۴ | ۱۷۶ | ۰/۰۲۳ | ۲/۳ |
| ۸ | ۱۳۸۶ | ۱۹ | ۱۳۶۷ | ۱۱ | ۲۱۶ | ۰/۰۵ | ۵/۰ |
| ۹ | ۱۳۸۶ | ۲۰ | ۱۳۶۶ | ۱۷ | ۲۷۰ | ۰/۰۶۳ | ۶/۳ |
| ۱۰ | ۱۳۸۶ | ۲۱ | ۱۳۶۵ | ۲۰ | ۲۱۵ | ۰/۰۹۳ | ۹/۳ |
| ۱۱ | ۱۳۸۶ | ۲۲ | ۱۳۶۴ | ۶ | ۹۲ | ۰/۰۶۵ | ۶/۵ |
| ۱۲ | ۱۳۸۶ | ۲۳ | ۱۳۶۳ | ۴ | ۲۳ | ۰/۱۷۴ | ۱۷/۴ |
| ۱۳ | ۱۳۸۷ | ۱۸ | ۱۳۶۹ | ۱ | ۱۵۳ | ۰/۰۰۷ | ۰/۷ |
| ۱۴ | ۱۳۸۷ | ۱۹ | ۱۳۶۸ | ۵ | ۲۵۶ | ۰/۰۲ | ۲/۰ |
| ۱۵ | ۱۳۸۷ | ۲۰ | ۱۳۶۷ | ۱۴ | ۲۷۸ | ۰/۰۵ | ۵/۰ |
| ۱۶ | ۱۳۸۷ | ۲۱ | ۱۳۶۶ | ۱۹ | ۲۴۳ | ۰/۰۷۸ | ۷/۸ |
| ۱۷ | ۱۳۸۷ | ۲۲ | ۱۳۶۵ | ۵ | ۹۰ | ۰/۰۵۵ | ۵/۵ |
| ۱۸ | ۱۳۸۷ | ۲۳ | ۱۳۶۴ | ۲ | ۳۸ | ۰/۰۵۳ | ۵/۳ |
| ۱۹ | ۱۳۸۸ | ۱۸ | ۱۳۷۰ | ۴ | ۱۶۷ | ۰/۰۲۴ | ۲/۴ |
| ۲۰ | ۱۳۸۸ | ۱۹ | ۱۳۶۹ | ۳ | ۲۱۶ | ۰/۰۱۴ | ۱/۴ |
| ۲۱ | ۱۳۸۸ | ۲۰ | ۱۳۶۸ | ۱۳ | ۲۴۸ | ۰/۰۵۲ | ۵/۲ |
| ۲۲ | ۱۳۸۸ | ۲۱ | ۱۳۶۷ | ۱۱ | ۲۱۳ | ۰/۰۵۲ | ۵/۲ |
| ۲۳ | ۱۳۸۸ | ۲۲ | ۱۳۶۶ | ۳ | ۸۵ | ۰/۰۳۵ | ۳/۵ |
| ۲۴ | ۱۳۸۸ | ۲۳ | ۱۳۶۵ | ۲ | ۲۰ | ۰/۱ | ۱۰/۰ |

**جدول شماره ۳** - جدول شیوع سالانه مصرف سیگار در دختران دانشجو بر اساس سن و دوره، جهت رسم نمودار

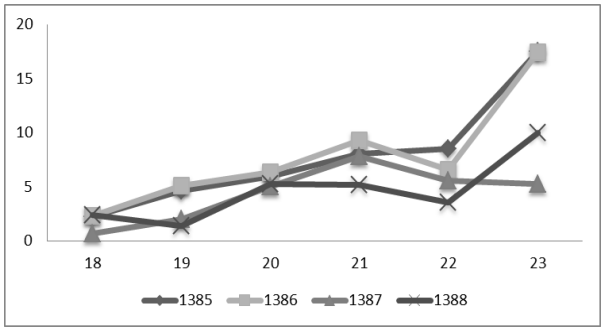
| سن | ۱۳۸۵ | ۱۳۸۶ | ۱۳۸۷ | ۱۳۸۸ |
|---|---|---|---|---|
| ۱۸ | ۲/۲ | ۲/۳ | ۰/۷ | ۲/۴ |
| ۱۹ | ۴/۷ | ۵/۰ | ۲/۰ | ۱/۴ |
| ۲۰ | ۶/۰ | ۶/۳ | ۵/۰ | ۵/۲ |
| ۲۱ | ۸/۰ | ۹/۳ | ۷/۸ | ۵/۲ |
| ۲۲ | ۸/۵ | ۶/۵ | ۵/۵ | ۳/۵ |
| ۲۳ | ۱۷/۵ | ۱۷/۴ | ۵/۳ | ۱۰/۰ |

### مرحله سوم: ترسیم نمودار

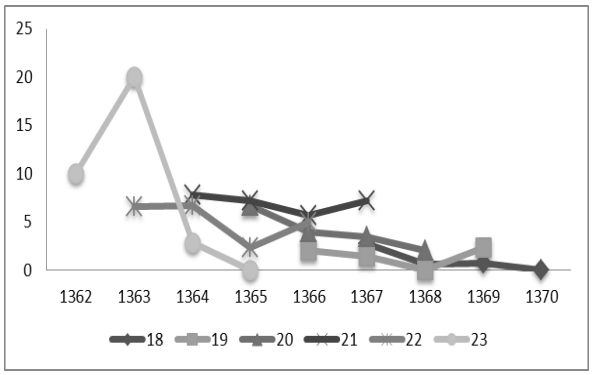
برای ترسیم نمودار توصیفی با استفاده از نرم‌افزار اکسل پس از تنظیم جدول شماره ۳ در صفحه داده اکسل، نمودار مربوطه را ترسیم می‌کنیم. شکل ۱- تصویر نمودار برای شیوع سالانه مصرف سیگار در دختران دانشجو بر اساس سن و دوره را نشان می‌دهد (۱۲). شکل ۲- تصویر نمودار برای شیوع سالانه مصرف الکل در دختران دانشجو بر اساس سن و همگروه که از جدولی مانند

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

جدول شماره ۳ به‌دست‌آمده است را نشان می‌دهد (۱۳). می‌توان جای دو عامل را در جدول تغییر داد. استفاده از این الگوها بستگی به هدف اصلی مطالعه و سلیقه محقق دارد (۵).




**شکل شماره ۱**– شیوع سالانه مصرف سیگار در زنان بر اساس سن و دوره




**شکل شماره ۲**– شیوع سالانه مصرف الکل در زنان بر اساس سن و هم‌گروه

گاهی جدا کردن اثر سن، دوره و همگروه تولد از همدیگر صرفاً با مشاهده جداول یا منحنی‌ها امکان‌پذیر نیست و با مشاهده منحنی، افزایش یا کاهش در میزان یک پیامد، ممکن است به اثر سن یا همگروه نسبت داده شود، درحالی‌که مربوط به اثر دوره است. بنابراین برای جدا کردن اثر آن‌ها از روش‌های آماری رگرسیون نیز استفاده می‌شود (۵).

مدل‌های رگرسیونی که برای کنترل اثر سن، همگروه و دوره به کار می‌روند:

۱) رگرسیون خطی که برای مطالعات اجتماعی و جمعیت‌شناسی به کار می‌رود و توسط ماسون و همکارانش در سال ۱۹۷۳ پیشنهادشده است.

۲) لگاریتم رگرسیون خطی که بیشتر در مطالعات اپیدمیولوژی به کار می‌رود به‌طوری‌که میزان‌های بیماری از طریق توزیع پوآسن محاسبه می‌شود.

۳) مدل لوجیت که در آن میرایی به‌صورت دوحالتی در نظر گرفته می‌شود (مرگ یا زنده ماندن) و به‌صورت یک متغیر دوجمله‌ای است. این مدل نیز بیشتر در تحقیقات جمعیت‌شناسی به کار می‌رود.

**محدودیت روش سن، همگروه و دوره**

چالش اصلی این روش که در اصطلاح آماری به آن مشکل شناسایی مدل[۱] گفته می‌شود، به علت وابستگی خطی کاملی است که بین سه عامل سن، همگروه و دوره در معادله رگرسیون مدل سن، دوره و همگروه وجود دارد. یعنی: همگروه = دوره – سن. این مشکل زمانی به وجود می‌آید که طول فاصله‌های زمانی در سن و دوره به‌صورت یکسان انتخاب می‌شود. مثلاً هر دو یک‌ساله یا هر دو ۵ ساله. در این مطالعات از هرکدام از مدل‌های رگرسیون خطی که برای آنالیز استفاده کنیم، به دلیل رابطه خطی کاملی که بین سه عامل وجود دارد، نمی‌توانیم برآورد درست و دقیقی از اثر سن، دوره و همگروه بدون حل این مشکل به دست آوریم.

**راه‌حل آماری برای حل مشکل**

در آمار حیاتی و اپیدمیولوژی راه‌حل‌هایی برای شناسایی این مشکلات در طول دو دهه گذشته پیشنهادشده است، هرچند هنوز هم این مشکلات تا حدودی به‌صورت حل‌نشده باقی‌مانده‌اند. روشی که فو در سال ۲۰۰۰ معرفی کرد و یانگ و همکاران در چند مقاله (۲۰۱۳- ۲۰۰۴) که در بعضی از آن‌ها فو هم جزء همکاران بوده، به شرح و توصیف آن پرداخته است، روش برآوردگر ذاتی نام دارد (۱۱-۱،۸).

## روش برآوردگر ذاتی

یانگ و همکارانش ادعا کرده‌اند که این روش به دلیل خصوصیات آماری خوبی که دارد نسبت به روش‌های قبلی ارجحیت دارد و بهتر از روش‌های قبلی می‌تواند مشکل شناسایی مدل را حل کند، این خصوصیات عبارت‌اند از: وابسته نبودن به اطلاعات قبلی و عدم نیاز برای تعیین گروه رفرنس برای ضرایب سن، دوره و همگروه؛ مناسب برای تعداد دوره‌های زمانی محدود و فاصله‌های یکسان سن و دوره (مثلاً هر دو یکسال)؛ نشان دادن اثر صفر سن، دوره و همگروه در صورت عدم وجود هر یک از این متغیرها؛ استفاده آسان و قابل‌اجرا بودن با یک نرم‌افزار آماری استاندارد مثل اس پلاس[۲] یا استاتا[۳] است (۱). البته این روش

[۱] Identification problem

[۱] S_Plus

[۲] Stata

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

مخالفانی نیز دارد، که ادعا می‌کنند این روش هیچ برتری نسبت به مدل‌های قبلی ندارد (۳،۱۴).

در این مطالعه، به علت ماهیت داده‌ها معادله ریاضی مدل، مدل رگرسیون لگاریتم خطی پوآسن[۱] هست (۱۳-۱۱).

$$\text{Log}\,(r_{ijk}) = \log\,(d_{ijk} / n_{ijk}) = \mu + \alpha_i + B_j + \gamma_k$$

Log($r_{ijk}$): نسبت (شیوع) مصرف سیگار

dij: تعداد مصرف‌کننده‌های یک نخ سیگار در یک سال گذشته که به‌صورت متغیرهای پوآسن توزیع‌شده‌اند.

nijk: تعداد افراد در معرض خطر (کل افراد مطالعه)

μ: عرض از مبدأ (میانگین تطبیق یافته شیوع مصرف)

$\alpha_i$: i امین ردیف اثر سنی برای $i = 1,\ldots, a$ (گروه سنی ۱ تا a)

Bj: j امین ستون اثر دوره برای $j=1,\ldots, p$ (دوره زمانی ۱ تا p)

$\gamma_k$: k امین قطر اثر همگروه برای $k=1,\ldots, (a+p-1)$ همگروه (۱ تا a+p-1) با $k=a-i+j$

مدل برآوردگر ذاتی با روش‌های ریاضی (ماتریکس) و هندسی اثبات‌شده است که می‌توانید برای مطالعه آن به منابع ۱ و ۸ تا ۱۱ مراجعه کنید.

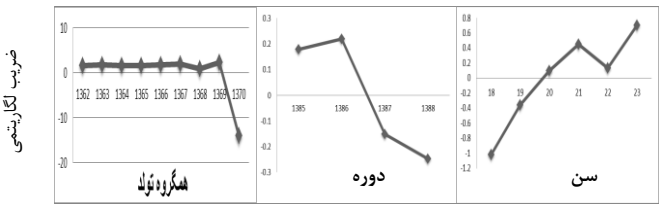
## روش کار در مدل تحلیلی برآوردگر ذاتی

برای انجام آنالیز با روش برآوردگر ذاتی، باید داده‌ها را به همان شکلی که در نرم‌افزار اکسل وارد کرده بودیم (ستون ۱ تا ۶) وارد نرم‌افزار استاتا کنیم و دستور آنالیز آن را در نرم‌افزار استاتا تایپ می‌کنیم (ضمیمه ۲).

پس ازآنالیز، با استفاده از مقدار ضرایب لگاریتمی به دست آمده از جدول برونداد داده‌ها برای سن، دوره و همگروه می‌توان نمودار داده‌ها را رسم کرد.

## یافته‌ها

همان‌طور که در شکل ۱ دیده می‌شود، شیوع مصرف سیگار تا ۲۱ سالگی افزایش و پس از یک کاهش مقطعی دوباره یک روند افزایشی را نشان می‌دهد. البته این روند برای سال ۸۸ نسبت به سال‌های ۸۵ تا ۸۷ قدری نامنظم‌تر است. در این شکل نمی‌توان روند خالص سن بدون اثر دوره را دید. شکل ۳- الف) که با استفاده از روش برآوردگر ذاتی ترسیم‌شده است، روند خالص سن برای تمام دوره‌ها را نشان می‌دهد. به‌طوری‌که همانند نمودار روش توصیفی (شکل شماره ۱) تا ۲۱ سالگی روند افزایش و پس‌ازآن پس از کاهش مقطعی برای ۲۲ ساله‌ها، دوباره روند افزایشی نشان می‌دهد. هم‌چنین برای سنین زیر ۲۰ سالگی، به‌عنوان یک عامل پیش‌گیری کننده (داشتن ضریب منفی) عمل می‌کند که احتمالاً به دلیل باقی ماندن اثر زندگی با والدین و کنترل والدین قبل از ورود به دانشگاه وداشتن زندگی مستقل یا در خوابگاه بوده است. شکل ۳ –ب) اثر خالص دوره را نشان می‌دهد، به‌طوری‌که شیوع مصرف سیگار از سال ۸۵ تا ۸۶ افزایش جزئی داشته و پس‌ازآن روند کاهشی نشان می‌دهد (۱۲). شکل ۲ – شیوع سالانه مصرف الکل در دختران دانشجو را بر اساس سن و همگروه نشان می‌دهد. شیوع مصرف الکل در دختران دانشجو تقریباً در همه گروه‌های سنی از همگروه‌های قدیمی‌تر به جوان‌تر کاهش یافته است. این در حالی است که شکل ۳- ج) نشان می‌دهد که اثر همگروه برای دختران دانشجو وجود ندارد و صفر است. همان‌طوری که در شکل دیده می‌شود، مقادیر نزدیک خط افقی هستند. کاهش ناگهانی که برای همگروه ۱۳۶۹ دیده می‌شود دارای تفاوت معنی‌داری نیست (۱۳).




الف ب ج

**شکل شماره۳- الف) برآورد ذاتی اثر سن برای شیوع مصرف سیگار در زنان ب) برآورد ذاتی اثردوره برای شیوع مصرف سیگار در زنان ج) برآورد ذاتی اثر همگروه در مصرف الکل برای زنان**

[۳] Poisson log- linear model

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

## بحث

در این مقاله بر روی روشی از مدل سن، همگروه تولد و دوره بحث شده است که ادعاشده است که خصوصیات آماری قابل‌توجهی نسبت به روش‌های قبلی دارد و همچنین نتایج روش مزبور می‌توانند جهت مقایسه تطابق روندهای دیده‌شده با الگوهای توصیفی نیز به کار روند. به عبارتی این روش نشان می‌دهد که آیا الگوهای مشاهده‌شده در روش توصیفی واقعی‌اند یا نه (۱۱). بعضی از محققین در مطالعات خود بیان کرده‌اند که اثر برآوردگر ذاتی می‌تواند برآورد بدون تورشی از اثر واقعی سن، دوره و همگروه را نشان دهد (۱۷-۱۵). یانگ (۲۰۰۴) با مقایسه روش برآوردگر ذاتی و مدل خطی تعمیم‌یافته محدود، نشان داد که روش برآوردگر ذاتی جانشین خوبی برای مدل خطی تعمیم‌یافته محدود است و نسبت به آن خصوصیات آماری بهتری دارد (۱). لانگلی (۲۰۱۱) از این روش در بررسی روند بروز شکستگی مفصل ران در داده‌های سال ۱۹۷۴ تا ۲۰۰۷ به این دلیل استفاده کرده است که معتقد بود، این روش می‌تواند اثر مجزای هریک از سه عامل سن، دوره و همگروه را نشان دهد (۱۸). یانگ این روش را بهترین انتخاب برای برخورد با مشکل وابستگی خطی بین سه عامل سن، همگروه تولد و دوره معرفی کرده است و بیان کرده که هرچند راه‌حل کاملی برای حل مشکل نیست ولی به دلیل خصوصیات آماری خوب آن، که عبارت‌اند از مناسب بودن برای جداول میزان‌ها و تعداد دوره‌های زمانی محدود و عدم نیاز برای تعیین رفرنس برای ضرایب سن، دوره و همگروه، می‌تواند به کار رود. اما محققین نباید بدون ملاحظه و مقایسه با روش‌های دیگر از آن استفاده کنند. چون هر مدل آماری بی‌شک دارای محدودیت‌هایی نیز هست و تحت شرایطی شکننده است (۹،۱). این در حالی است که لو (۲۰۱۳) ادعا کرد که روش برآوردگر ذاتی، روش بدون تورشی برای برآورد واقعی اثر سن، دوره و همگروه نیست و نه‌تنها به تعداد گروه‌های سن و دوره و همگروه بستگی دارد بلکه هیچ تفاوتی با روش مدل خطی تعمیم‌یافته محدود نداشته و نمی‌توان گفت که بهتر از آن است (۳). در مقابل یانگ (۲۰۱۳) نظرات لو را رد کرده و اظهار کرده روشی که لو با استفاده از اثبات معادلات جبری (ریاضی) در عدم شایستگی روش مذبور نشان داده است، ارزش چندانی ندارد (۸). گروتنهویس (۲۰۱۶) در مطالعه دیگری با کدگذاری ساختگی دوباره داده‌های میرایی مطالعه مستر و همکارانش (۲۰۱۴) نشان داد که برآوردگر ذاتی هم مثل روش قبلی است و می‌تواند به‌شدت دچار تورش شود (۱۴). مسترز و همکاران (۲۰۱۶) در پاسخ به گروتنهویس در مطالعه دیگری این‌طور توضیح دادند که هدف آن‌ها از بررسی روند میرایی در سیاه‌پوستان و سفیدپوستان در مطالعه قبلی برجسته کردن اندازه تفاوت‌های موجود در میزان‌های میرایی سفیدپوستان و سیاه‌پوستان نبود، بلکه آن‌ها علاقه‌مند به بررسی تغییرات گذرا در این میزان‌ها بودند و کدگذاری گروتنهویس نامناسب بوده و نشان دادند که چنین کدگذاری می‌تواند به تفسیر اشتباه منجر گردد و از مطالعه قبلی خود دفاع کردند. آن‌ها بیان کردند که محققین باید از روش برآوردگر ذاتی نیز مانند سایر روش‌های سن، دوره و همگروه با دقت و احتیاط استفاده کنند (۱۸).

لو و گروتنهویس هرچند در مطالعات خود نتیجه‌گیری کردند که روش برآوردگر ذاتی روش مناسب و بدون تورشی نیست ولی روش جایگزین یا جدیدی نیز معرفی نکردند.

یانگ (۲۰۱۳) در مقاله‌ای در پاسخ به مقاله انتقادی (لو)، چنین بیان میکند، این درست نیست که دیدگاه ما نسبت به برآوردگر ذاتی به‌عنوان یک جام مقدس یا توپ جادویی برای حل مشکل شناسایی مدل تفسیر شود. ما هیچ‌گاه چنین ادعایی نداشتیم و بر این باوریم که محققین بایستی مدل‌های سن، همگروه و دوره را بااحتیاط و آگاهی از مشکلات و تله‌های ایجادشده توسط این مطالعات استفاده کنند. ولی بااین‌حال معتقدیم که روش برآوردگر ذاتی نشان داده است که روش مفیدتری در برخورد با مشکل شناسایی و برآورد مدل کامل سن، دوره و همگروه (هر سه عامل) است (۱۹).

## نتیجه‌گیری

با توجه به مطالب گفته‌شده و با توجه به این‌که حتی کسانی که از این روش ایراد گرفته‌اند، روش جایگزین دیگری برای آن پیشنهاد نکرده‌اند؛ به نظر می‌رسد، روش برآوردگر ذاتی در حال حاضر نسبت به مدل‌های قدیمی‌تر، انتخاب بهتری برای حل مشکل وابستگی خطی بین سه متغیر سن، همگروه تولد و دوره در این نوع مطالعات است. اما محققین باید بااحتیاط و آگاهی از مشکلات و تله‌های ایجادشده توسط مطالعات سن، دوره و همگروه از آن استفاده کنند.

**ضمیمه**

۱) ابزار محاسبه:

جهت اضافه کردن ابزار برآوردگر ذاتی به کامپیوتر خود می‌توان

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

با تایپ (ssc install apc) در برنامه استاتا در زمان وصل بودن به اینترنت آن را به برنامه استاتا خود اضافه کرد. پس از نصب برنامه، برای کار کردن با آن نیازی به اینترنت نیست و می‌توان با تایپ apc_ie.hlp در بخش جستجوی help و علامت زدن در قسمت جستجوی همه موارد، جزییات را از help به دست آورد. برای استفاده از آن در برنامه اس – پلاس می‌توان یک کپی از آدرس زیر درخواست نمود (fuw@epi.msu.edu).

۲) دستور استاتا:

دستور برنامه بر اساس پیش‌فرض برنامه استاتا (Newton Raphson) و مقیاس (x2) به شرح ذیل خواهد بود.

apc_ie Y,age(a) period(t) cohort(c) family(poisson) link(log) exposure(exp_f) scale(x2(

Y/ متغیر وابسته: a/ متغیر سن: t/ متغیر دوره: c متغیر همگروه:

Family/ link: Poisson log linear model

Scale parameter (x2) = Pearson chi-squared / residual degrees of freedom

(۸،۱۳)

(برای دستور اس -پلاس به رفرنس ۲۰ مراجعه کنید).

## منابع


1. Yang Y, J. Fu W and C. Land K. A methodological comparison of age-period-cohort models: the intrinsic estimator and conventional generalized linear models. Sociological Methodology. 2004; 34: 75-110.
2. L. Kupper L, M. Janis J,Karmous A and G.Greenberg B.Statistical age-period-cohort analysis: A review and critique.J Chron Dis. 1985; 10: 811-30.
3. Luo L. Assessing Validity and Application Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort Problem. Demography. 2013; 50:1945–1967.
4. Robertson C, Gandini S and Boyle P. Age-Period-Cohort Models: A Comparative Study of Available Methodology. ClinEpidemiol. 1999; 6: 569–83.
5. Szklo M, Nieto FJ. Epidemiology beyond the Basics.2th ed. USA: Mike Brown 2007.
6. M. Keyes K and Li G. A multi-phase method for estimating cohort effects in age-period contingency table data. Ann Epidemiol J. 2010; 20: 779-85.
7. C. Kerr W, K. Greenfield T, Bond J, Ye Y and Rehm J. Age–period–cohort influences on trends in past year marijuana use in the US from the 1984, 1990, 1995 and 2000 National Alcohol Surveys.Drug and Alcohol Dependence J. 2007; 86: 132–8.
8. Yang C Y, C. Land K. The Statistical Properties of the Intrinsic Estimator for Age-Period-Cohort Analysis. 2013.Available on line at: http://yangclaireyang.web.unc.edu/files/2013/07/ResponseToLiyingLuo-2013_APConline.pdf
9. Yang Y, Schulhofer-Wohl S, J. Fu W and C. Land K. The Intrinsic Estimator for Age-Period-Cohort Analysis: What It Is and How to Use It, AJS, 2008; 6: 1697–1736.
10. Yang Y, Schulhofer-Wohl S and C. Land K. A Simulation Study of the Intrinsic Estimator for Age-Period-Cohort Analysis.2007. Available on line at: http://www.princeton.edu/~sschulho/files/YSL_apcsim.pdf.
11. Yang Y. Trends In U.S. Adult Chronic Disease Mortality, 1960–1999: Age, Period, and Cohort Variations, Demography, 2008; 2: 387–416.
12. Sahebi R, RahimiMovaghar A, Motevalian SA and Yonesian M. The Role of Age, Period, and Cohort Effects on Smoking among the Students of Tehran University of Medical Science. Social Sciences J.2014; 3: 779-89.
13. Motevalin A,Sahebi R,Rahimi- Movagar A and Yonesian M. Trend of Alcohol and Drug use in Students in Tehran University of Medical Sciences from 2006 to 2009.An Age, Period, Cohort Study. Iran Epidemiology J.2015; 11:99-108.
14. Grotenhuis M T,Pelzer B, Luo L and W. Schmidt-Catra A.The Intrinsic Estimator, Alternative Estimates, and Predictions of Mortality Trends: A Comment on Masters, Hummer, Powers, Beck, Lin, and Finch.Demography.2016. Advance online publication. DOI:10.1007/s13524-016-0476-8.
15. Fu W. Ridge estimator in singular design with applications to age-period-cohort analysis of disease rates. Communications in Statistics Theory and Method. 2000; 29: 263–78.
16. Schwadel P. Age, period, and cohort effects on religious activities and beliefs. Social Science Research. 2011; 40: 181–92.
17. Keyes K and Miech R. Age, period, and cohort effects in heavy episodic drinking in the US from 1985 to 2009. Drug and Alcohol Dependence.2013; 132: 140–48.
18. Langley J, Samaranayaka A, Davie J and Campbell A. J. Age, cohort and period effects on hip fracture incidence: Analysis and predictions from New Zealand data 1974–2007. Osteoporosis International.2011; 22: 105–111.
19. K. Masters R, A. Powers D, A. Hummer, Beck A, Lin S-F and Karl Finch B. Fitting Age-Period-Cohort Models Using the Intrinsic Estimator: Assumptions and Misapplications. Demography. Advance online publication.DOI: 10.1007/s13524-016-0481-y.
20. Yang C Y, C. Land K. Misunderstandings, Mischaracterizations, and the Problematic Choice of a Specific Instance in Which the IE Should Never Be Applied.Demography.2013; 50(6):1969–1971.
21. apc.pkg. Available at: http://fmwww.bc.edu/RePEc/bocode/a/-


[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]

Iranian Journal of Epidemiology Summer 2017; 13 (2): 90-97.
Original Article

# The Intrinsic Estimator Method for Control of Age, Period, Cohort, and Its Comparison with the Descriptive Classic Model

**Sahebi R[1], Motevalian A[2], Sahebi L[3], Sharifi H[4]**

1- MS of Epidemiology, Tabriz University of Medical Sciences, Tabriz, Iran
2- Associate Professor of Epidemiology, Iran University of Medical Sciences, Tehran, Iran
3- PHD of Epidemiology, Maternal, Fetal and Neonatal Research Center, Tehran University of Medical Sciences, Tehran, Iran
4- Associate Professor of Epidemiology، HIV/STI Surveillance Research Center, and WHO Collaborating Center for HIV Surveillance, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
**Corresponding author:** Sharifi H, hsharifi@kmu.ac.ir
(**Received** 12 June 2016; **Accepted** 26 November 2016)

**Background and Objectives**: The aim of this study was to compare the intrinsic estimator method and the age, period and cohort accounting model used for identifying age, period, and cohort trends in incidence, prevalence, and mortality rates.

**Methods:** The data of 2 studies " Age, Period, and Cohort Effects on Alcohol and Drug use Among Students of Tehran University of Medical Sciences from 2006 to 2009" and "The Role of Age, Period, and Cohort Effects on Smoking among the Students of Tehran University of Medical Science "were used. First the age, period and cohort accounting model and then The IE method are explained, their advantages and disadvantages are discussed, and their results are compared.

**Results:** Both methods showed an increasing, decreasing, and increasing trend for age, cohort, and cohort effect for the prevalence of smoking. IE could predict a preventing effect for an early age factor for smoking. In the descriptive model, a decreasing trend was seen from old cohorts to younger cohorts in nearly all age groups. However, the IE method did not show any cohort effects for alcohol consumption in female studnets.

**Conclusion**: Currently, the IE method is the best method for solving linear dependence between three variables of age, birth cohort, and period in this type of studies. However, researchers should use it with caution because it has many pitfalls.

**Keywords:** Age, Period, Cohort, The intrinsic estimator method

[ Downloaded from journals.tums.ac.ir on 2024-11-19 ]