

بررسی روش‌های مختلف کاهش آریبی در مطالعات مشاهده‌ای

محرری طرح: محمد رضا فرید روحانی

خلیل شفیع

فرشید جمشیدی

ماندانا طماسی نژاد

گروه پژوهشی طرح‌های فنی و روش‌های آماری

پژوهشگاه آماری

پیشگفتار

امروزه مطالعات مشاهده‌ای در حوزه‌های مختلف اقتصادی، اجتماعی، بهداشت و سلامتی، پژوهش‌های پزشکی و ... توسعه یافته است. آن‌چنان‌که در این گزارش اشاره شده است، هدف در این‌گونه مطالعات بررسی اثر علی تیمار و تبیین رابطه علی برای ظهور پدیده مورد نظر می‌باشد. بدیهی است روش استاندارد جهت بررسی وجود چنین رابطه‌ای، اجرای یک آزمایش کنترل شده است که به این روش استاندارد طلایی گفته می‌شود. اما متأسفانه در اغلب حوزه‌های مورد اشاره انجام آزمایشات کنترل شده به دلایل گوناگونی امکان‌پذیر نیست. از این‌رو اجرای مطالعات مشاهده‌ای منطبق بر مقتضیات موجود، تنها گزینه قابل دفاع در بررسی اثر علی تیمار خواهد بود. با این وجود با توجه به این ویژگی در مطالعات مشاهده‌ای، تخصیص تیمار به آزمودنی‌ها به صورت تصادفی صورت نمی‌گیرد و بررسی اثر علی با اتکاء صرف بر تفاوت‌های مشاهده شده مقادیر پاسخ در گروه‌های تیماری در معرض اریبی خواهد بود. این اریبی ناشی از تفاوت احتمالی آزمودنی‌های گروه کنترل و تیمار، در متغیرهای کمکی اندازه‌گیری شده قبل از در معرض تیمار قرار گرفتن آن‌ها می‌باشد که اصطلاحاً به آن، تفاوت ماقبل تیماری گفته می‌شود. از این‌رو برخلاف آزمایشات کنترل شده در مطالعات مشاهده‌ای با پدیده اریبی اثر علی مواجه هستیم.

گزارش حاضر با هدف بررسی روش‌های مختلف کاهش اریبی در مطالعات مشاهده‌ای تهیه و تدوین شده است. به عنوان پیش درآمدی بر بررسی روش‌های کاهش اریبی در فصل اول به معرفی مختصات یک مطالعه مشاهده‌ای و انواع آن و نیز برخی مسائل قابل تأمل در حوزه این‌گونه مطالعات پرداخته شده است. فصل دوم تا چهارم به ترتیب به معرفی سه روش زیرطبقه‌بندی، جورسازی یا نمونه‌گیری جورشده و تعدیل رگرسیونی یا تعدیل کوواریانس اختصاص یافته است. اگرچه این روش‌ها از نظر تعداد متغیرهای کمکی دارای محدودیت ذاتی نیستند اما در عمل با افزایش تعداد متغیرهای کمکی تحت مطالعه بسیار پیچیده شده و اعمال آن‌ها به عنوان گامی مهم در کاهش اریبی با مشکل مواجه می‌گردد. فصل پنجم پاسخی به چنین وضعیت‌هایی است که با تعداد زیادی متغیر کمکی روبرو هستیم. این فصل به معرفی نمره تمایل - یعنی احتمال تعلق یک آزمودنی

دارای مقادیر معینی از متغیرهای کمکی به گروه تیمار- به عنوان حلقه واسطی در اعمال روش‌های کاهش آریبی در مواردی که چندین متغیر کمکی حضور دارند، پرداخته است. به عبارت دیگر در این فصل در باره شرایط و امکان اعمال روش‌های سه گانه پیشین بر مبنای نمره تمایل برآورد شده کنکاش شده است.

اگرچه با پایان یافتن فصل پنجم عملاً معرفی روش‌های کاهش آریبی تکمیل می‌گردد، اما از آنجا که تمامی روش‌های کاهش آریبی بر اساس متغیرهای کمکی از پیش اندازه‌گیری شده اعمال می‌گردند، لذا هنوز امکان مقایسه‌پذیر نبودن آزمودنی‌ها در دیگر متغیرهای کمکی پیش‌بینی نشده وجود دارد. فصل‌های ششم و هفتم به عنوان مباحثی تکمیلی به معرفی آزمون X- تعدیل‌پذیری و تحلیل حساسیت در حوزه مطالعات مشاهده‌ای می‌پردازند.

اجرای این طرح با مدیریت جناب آقای دکتر محمدرضا فرید روحانی عضو هیئت علمی دانشگاه شهید بهشتی و همکاری جناب آقای دکتر خلیل شفیع‌ی عضو هیئت علمی دانشگاه شهید بهشتی، آقای فرشید جمشیدی کارشناس پژوهشکده‌ی آمار و سرکار خانم ماندانا طهماسبی‌نژاد کارشناس سازمان صدا و سیما انجام شده است که در این جا از یکایک این عزیزان تشکر و قدردانی می‌شود.

از سرکار خانم ناظریان که زحمت تایپ این گزارش را به عهده داشته و با حوصله فراوان آن را به انجام رساندند نیز بسیار تشکر می‌شود.

بر این اعتقادیم که این گزارش به عنوان گردآورنده روش‌های مختلف کاهش آریبی می‌تواند نخستین گام در آشنایی علاقه‌مندان به مطالعات مشاهده‌ای در امکان‌سنجی بررسی اثر علی در حوزه‌های مختلف اقتصادی- اجتماعی باشد که طبیعتاً بیش‌تر مورد توجه مراکز آمار رسمی کشور نظیر مرکز آمار ایران می‌باشد.

گروه پژوهشی طرح‌های فنی و روش‌های آماری

فهرست مطالب

عنوان	صفحه
فصل ۱- مطالعات مشاهده‌ای	۱
مقدمه	۱
۱-۱- تعریف و ویژگی‌ها	۲
۲-۱- انواع مطالعات مشاهده‌ای	۷
۱-۲-۱- مطالعات گذشته‌نگر	۷
۲-۲-۱- مطالعات آینده‌نگر	۱۰
۳-۲-۱- مطالعات مقطعی	۱۵
۴-۲-۱- مقایسه مطالعات گذشته‌نگر و آینده‌نگر	۱۷
۳-۱- ملاحظات دربارهی مطالعات مشاهده‌ای	۱۴
۱-۳-۱- گام پیوند تا علیت	۱۷
۲-۳-۱- استنباط از نمونه به جامعه	۲۶
۳-۳-۱- مواجهه و کنترل متغیرهای مزاحم	۲۸
فصل ۲- روش زیر رده‌بندی	۳۱
۱-۲- روش شناسی	۳۳
۲-۲- درصد کاهش اربیبی با یک رگرسیون خطی	۴۳
۱-۲-۲- توزیع نرمال برای متغیر کمکی	۴۶

۴۷	۲-۲-۲- توزیع غیر نرمال برای متغیر کمکی
۵۱	فصل ۳- نمونه‌گیری جورسازی
۵۳	۱-۳- روش‌شناسی یک متغیره
۵۷	۳-۱-۱- حداکثر کاهش اریبی
۶۱	۳-۱-۲- جورسازی متوسط
۶۳	۳-۱-۳- جورسازی زوجی
۶۷	۳-۲- روش‌شناسی چند متغیره
۶۸	۳-۲-۱- روش‌های جورسازی با کاهش اریبی یکسان
۶۹	۳-۲-۱-۱- جورسازی یک متغیره
۷۰	۳-۲-۱-۲- جورسازی اشتراکی
۷۲	۳-۲-۱-۳- توزیع نرمال برای بردار X
۷۵	۳-۲-۱-۴- توزیع جابه‌جاپذیر برای بردار X
۷۷	۳-۲-۱-۵- توزیع بیضی‌وار برای بردار X
۷۸	۳-۲-۲- بیش‌ترین درصد کاهش اریبی
۸۰	۳-۲-۲-۱- درصد بهینه‌ی کاهش اریبی

۲-۲-۲-۳-۳	روند تغییرات θ_{max}	۸۱
۳-۲-۲-۳	حجم نمونه‌های مینیمال	۸۴
۴	فصل ۴- روش تعدیل رگرسیونی	۸۷
۱-۴	روش شناسی	۸۹
۲-۴	تعدیل رگرسیونی و نمونه‌گیری جور شده	۹۴
۵	فصل ۵- نمره‌ی تمایل	۹۹
	مقدمه	۱۰۱
۱-۵	نمرات تعدیل و نمره‌ی تمایل	۱۰۱
۲-۵	برآورد نمره‌ی تمایل	۱۰۵
۳-۵	زیر رده‌بندی بر اساس نمره‌ی تعدیل	۱۰۶
۴-۵	جورسازی بر اساس نمره‌ی تمایل	۱۱۰
۵-۵	نمره‌ی تمایل و تعدیل رگرسیونی	۱۱۳
۶	فصل ۶- اریبی حاصل از اختلافات پیش‌تیمار	۱۱۵
۱-۶	مقدمه	۱۱۷
۲-۶	سه مثال از گروه‌های کنترل چندگانه	۱۱۸

۳-۶- تخصیص تیمار X- تعدیل پذیر	۱۲۰
۱-۳-۶- مقدمه	۱۲۰
۴-۶- (X,U)- تعدیل پذیری تخصیص تیمار	۱۲۴
۵-۶- نقش گروه کنترل دوم در مطالعات هم گروهی	۱۲۴
۱-۵-۶- گروه کنترل دوم چیست	۱۲۴
۲-۵-۶- استفاده از گروه کنترل دوم برای آزمون فرض X- تعدیل پذیری تخصیص تیمار	۱۲۵
۳-۵-۶- کنترل به وسیله تغییرات سیستماتیک	۱۲۶
۴-۵-۶- آزمون سازگار و ناریب	۱۲۷
۵-۵-۶- براکت کردن	۱۳۱
۶-۵-۶- توان آزمون X- تعدیل پذیری در صورت براکت نکردن	۱۳۳
۷-۵-۶- مقایسه پذیری جزئی	۱۳۶
۶-۶- نقش گروه «کنترل» دوم در مطالعات کنترل- مورد	۱۳۷
۱-۶-۶- مقایسه‌های مورد- نامورد	۱۳۷
۲-۶-۶- گروه‌های نامورد یا «کنترل» در مطالعات کنترل- مورد	۱۳۸
۳-۶-۶- نسبت بخت‌ها در مطالعات کنترل- مورد	۱۴۱
۴-۶-۶- نکته نهایی	۱۴۳

فصل ۷- تحلیل حساسیت.....	۱۴۵
۱-۷- مقدمه.....	۱۴۷
۲-۷- مدل سازی تحلیل حساسیت.....	۱۴۷
۱-۲-۷- مدلی بر حسب احتمالات تخصیص.....	۱۴۷
۲-۲-۷- مدلی بر حسب متغیر کمکی مشاهده نشده.....	۱۴۸
۳-۲-۷- توزیع تخصیص تیمار بعد از طبقه بندی.....	۱۵۰
۳-۷- زوج های جور شده.....	۱۵۱
۱-۳-۷- حساسیت سطوح معنی داری: حالت کلی.....	۱۵۱
۲-۳-۷- تحلیل حساسیت برای آزمون مک نمار.....	۱۵۴
۳-۳-۷- تحلیل حساسیت برای آزمون: رتبه علامت دار ویلکاکسون.....	۱۵۶
۴-۷- تحلیل حساسیت برای آمارهای نمره- علامت.....	۱۶۰
۱-۴-۷- روش کلی.....	۱۶۰
۲-۴-۷- جور سازی با کنترل های چندگانه با استفاده از آماره های نمره- علامت.....	۱۶۲
۳-۴-۷- تحلیل حساسیت با کنترل های چندگانه: مثال ویتامین C و سرطان.....	۱۶۴
۴-۴-۷- مطالعات کنترل- مورد با کنترل های چندگانه.....	۱۶۶
۵-۴-۷- مطالعات کنترل- مورد جور سازی شده: چند مثال.....	۱۶۸

نتیجه گیری ۱۷۱

منابع ۱۷۵

پیوست ها ۱۸۱

فصل اول

مطالعات مشاهده‌ای

فصل اول

مطالعات مشاهده‌ای

مقدمه:

این فصل بر آن است تا معرفی جامعی از مطالعات مشاهده‌ای و مقایسه‌ی آنها با آزمایشات کنترل شده را ارائه نماید. به علاوه به منظور شناسایی هر چه بهتر مطالعات مشاهده‌ای، انواع مختلف این گونه مطالعات معرفی گردیده و با ارائه مثال‌هایی سعی شده است علاوه بر درک هر چه بهتر مشخصات این گونه مطالعات، تا حدودی حوزه‌های مورد کاربرد مطالعات مشاهده‌ای نیز مشخص شود. در انتها نیز برخی ملاحظات عمومی در رابطه با مطالعات مشاهده‌ای و آزمایشات کنترل شده یعنی رابطه‌ی پیوند و علیت، استنباط از نمونه به جامعه و نیز مواجهه و کنترل متغیرهای مزاحم در حد ضرورت طرح گردیده است.

۱-۱ تعریف و ویژگی‌ها

ولد^۱ (۱۹۵۶) برای نخستین بار ویژگی‌های آزمایشات کنترل شده را در قالب سه معیار ارائه داد، و سپس اطلاعات مشاهده‌ای را فاقد لااقل یکی از سه ویژگی زیر معرفی نمود.

۱- تکرار آزمایش تحت شرایط یکسان انجام پذیرد.

۲- تکرارها تواماً مستقل باشند.

۳- تغییرات کنترل نشده در تکرارهای آزمایش در معرض تصادفی‌سازی به مفهوم فیشری^۲ آن باشند.

به اعتقاد ولد و پیش از آن فیشر مهم‌ترین ویژگی آزمایشات کنترل شده، گردآوری اطلاعات به گونه‌ای است که معیار تصادفی کردن در آن رعایت شده و امکان‌پذیر باشد. امری که در مطالعات مشاهده‌ای امکان‌پذیر نیست. با این وجود جامع‌ترین و روشن‌ترین بحث در زمینه‌ی ویژگی‌های مطالعات مشاهده‌ای توسط ککران^۳، برای نخستین بار در سال ۱۹۶۵ صورت گرفته است. او مطالعات مشاهده‌ای را به‌عنوان موضوعی قابل طرح در چارچوب اصول و روش‌های آماری معرفی نموده و دو ویژگی این گروه از مطالعات در جوامع انسانی را به صورت زیر بر شمرده است.

۱- هدف در این گروه از مطالعات توضیح و تبیین ارتباط علت و معلولی یا لااقل ارزیابی

ارتباط بین مجموعه‌ای از متغیرهای مشخص شده x_i و گروه دیگری از متغیرها (y_i)

به گونه‌ای است که بتوان از ارزیابی این ارتباط فرضیات مناسبی را درباره‌ی علت وقوع

پدیده مورد بررسی ارائه داد.

۲- امکان استفاده از آزمایشات کنترل شده وجود نداشته باشد. به این معنی که امکان تحمیل

روش‌ها یا تیمارهایی که کشف اثرات آن‌ها مورد توجه است، و نیز تخصیص رهیافت‌های

مختلف به صورت تصادفی به آزمودنی‌ها یا واحدهای آزمایشی وجود ندارد.

¹ Wold

² Fisher

³ Cochran

با توجه به این تعریف دو ویژگی مهم مطالعات مشاهده‌ای را می‌توان استخراج نمود. اولاً این گروه از مطالعات همانند آزمایشات کنترل شده در پی بررسی اثرات تیمارها هستند. لذا اگر در یک بررسی، تیماری حضور نداشته باشد آن بررسی فاقد ویژگی یک مطالعه مشاهده‌ای و نیز یک آزمایش کنترل شده است. بسیاری از نظرسنجی‌ها، اکثر پیش‌بینی‌ها، بخش قابل توجهی از مطالعات تشخیصی و طبقه‌بندی و گروهی از مطالعات مهم تجربی نه در زمره‌ی آزمایشات کنترل شده و نه در ردیف مطالعات مشاهده‌ای قرار می‌گیرند. ویژگی دوم مطالعات مشاهده‌ای که در حقیقت وجه ممیزه‌ی این گروه از مطالعات از آزمایشات کنترل شده است، عبارت از عدم تخصیص تصادفی تیمارها به آزمودنی‌ها است. بنابراین در مطالعات مشاهده‌ای، کنترلی از جانب آزمایش‌گر نمی‌تواند بر واحدهای دریافت‌کننده‌ی تیمارهای مختلف صورت گرفته و در نتیجه اطمینانی از مقایسه‌پذیری این واحدها وجود نخواهد داشت. هم‌چنان که می‌دانیم تیمار به مداخلاتی اطلاق می‌شود که به‌طور بالقوه هر یک از واحدهای جامعه‌ی تحت مطالعه بتوانند در معرض آن قرار گرفته یا در معرض آن قرار نگیرند (روزنباوم^۱، ۱۹۸۴). بنابراین در معرض نوعی ماده‌ی شیمیایی قرار گرفتن (مثلاً رادیواکتیو، آزبست و غیره) می‌تواند یک تیمار تلقی شود و حال آن‌که جنسیت، نژاد، سن و نظایر آن فاقد ویژگی تیمار هستند، زیرا این ویژگی‌ها حتی به صورت بالقوه تغییرپذیر نیستند. لازم به ذکر است تفاوت در پاسخ‌های مشاهده شده‌ی گروه‌های تیماری مختلف می‌تواند انعکاسی از تفاوت واقعی گروه‌های تیماری بوده، یا می‌تواند ناشی از اختلاف ساختاری و ماقبل تیماری^۲ آزمودنی‌ها باشد. اگر این تفاوت ناشی از اختلاف در متغیرهای کمکی (ماقبل تیماری) واحدهای آزمایشی باشد، در حقیقت این واحدها غیرقابل مقایسه بوده و در نتیجه برآورد اثر تیمار مستقیماً و با اتکاء صرف به متغیرهای پاسخ مشاهده شده در معرض آریبی خواهد بود. این در حالی است که در آزمایشات کنترل شده به واسطه‌ی اعمال شرایط کنترلی و تخصیص تصادفی تیمارها به آزمودنی‌ها، پدیده آریبی برآورد اثر تیمار حاصل از عدم مقایسه‌پذیری واحدهای آزمایش

¹ Rosenbaum

² Pre- treatment

به وجود نمی‌آید. بنابراین یکی از وجوه تصادفی‌سازی و تأکید فیشر بر آن در آزمایشات کنترل شده مقایسه‌پذیر ساختن واحدهای آزمایشی است. اثرات دیگر تصادفی‌سازی را در بخش ۱-۳ بررسی خواهیم نمود. مثال زیر که توسط روزن‌بام (۲۰۰۲) ارائه شده است اثر تصادفی‌سازی را بر مقایسه‌پذیر ساختن آزمودنی‌ها به خوبی نشان می‌دهد.

مثال ۱-۱- آزمایش تصادفی‌شده‌ی جراحی کرونر

مورفی^۱ آزمایش کنترل شده‌ی تصادفی‌شده‌ای را برای مقایسه جراحی بای پس^۲ عروق کرونر و معالجه دارویی این بیماری، برای بیماران دچار گرفتگی عروق کرونر انجام داد. جراحی بای پس، باز کردن مجدد عروقی است که خون را به قلب می‌رسانند و بر اثر رسوب و تجمع چربی دچار تنگی شده‌اند. در درمان دارویی، از داروهایی برای افزایش جریان خون از داخل عروق تنگ شده استفاده می‌شود. در این مطالعه ۵۹۶ بیمار از ۱۳ بیمارستان به تصادف انتخاب شده که ۲۸۶ نفر از آن‌ها تحت جراحی و ۳۱۰ بیمار نیز تحت درمان دارویی قرار گرفته‌اند. تخصیص هر یک از بیماران به یکی از دو گروه تیماری به صورت تصادفی و پس از کسب رضایت بیمار جهت شرکت در این مطالعه انجام پذیرفته است. جدول ۱-۱ خلاصه‌ای از نتایج این مطالعه را نشان می‌دهد. در این جدول گروه‌های تیماری (درمان دارویی و جراحی) بر حسب ۱۰ مشخصه‌ی مهم اندازه‌گیری شده قبل از شروع درمان که به آن‌ها متغیرهای کمکی گفته می‌شود مقایسه شده‌اند. لازم به ذکر است در این آزمایش هیچ‌گونه کنترلی بر روی این متغیرهای کمکی انجام نگرفته است.

همچنان‌که از نتایج این جدول مشاهده می‌شود، دو گروه بیماران تحت درمان دارویی و جراحی در بسیاری از متغیرهای مهم کمکی قبل از شروع درمان مشابه یکدیگر بوده و لذا مقایسه‌پذیر هستند. از این میان تنها متغیر کمکی کلسترول سرم بزرگ‌تر از ۲۴۹ است که در دو گروه تیماری در سطح ۰/۰۵ اختلاف معنی‌داری را نشان می‌دهد. بنابراین صرف وجود مکانیزم تخصیص تصادفی تیمارها به بیماران، و علی‌رغم عدم وجود کنترل بر روی ۱۰ متغیر کمکی اندازه‌گیری شده دو گروه بیماران

^۱ Morphy

^۲ By Pass

مقایسه پذیر (متعادل) شده‌اند. نتیجه آن که انتظار می‌رود گروه‌های تیماری در متغیرهای کمکی دیگری که اندازه‌گیری نشده‌اند نیز مقایسه‌پذیر و متعادل باشند. البته از آن‌جا که حتی در همین جدول متغیر کمکی کلسترول سروم بزرگ‌تر از ۲۴۹ در دو گروه بیماران اختلاف معنی‌داری دارد، عبارت دقیق‌تر چنین خواهد بود که:

اگر چه تخصیص تصادفی تیمارها می‌تواند منجر به عدم تعادل‌هایی به صورت شانسی گردد، اما استفاده‌ی مناسب از روش‌های آماری می‌تواند عدم قطعیت ناشی از این‌گونه عدم تعادل‌های احتمالی را کنترل کند (روزنبام ۲۰۰۲).

جدول ۱-۱- مقایسه بیماران کرونری در یک آزمایش تصادفی کلینیکی

متغیر کمکی	درمان دارویی (%)	جراحی (%)
در انجمن قلب نیویورک (گروه II و III)	۹۴/۲	۹۵/۴
سابقه‌ی سکته قلبی	۵۹/۳	۶۴/۰
نوار قلب (الکتروکاردیوگرام) مشخص یا احتمالی حاکی از سکته قلبی	۳۶/۱	۴۰/۵
مدت تداوم درد ناحیه‌ی سینه (بیش از ۲۵ ماه)	۵۰/۰	۵۱/۸
سابقه فشار خون	۳۰/۰	۲۷/۶
سابقه نارسایی احتقاقی قلب	۸/۴	۵/۲
سابقه بروز عروق مغزی	۳/۲	۲/۱
سابقه بیماری قند دیابت	۱۲/۹	۱۲/۲
نرخ قلب- قفسه‌ی صدری بزرگ‌تر از ۰/۴۹	۱۰/۴	۱۲/۲
میزان کلسترول سرم بزرگ‌تر از ۲۴۹ $\left(\frac{mg}{100ml}\right)$	۳۱/۶	۲۰/۶*
حجم نمونه	۳۱۰	۲۸۶

به راستی چه عواملی باعث عدم اعمال کنترل در مطالعات مشاهده‌ای می‌شود؟
برخی از عوامل مهم دخیل در عدم امکان چنین کنترل‌هایی در مطالعات مشاهده‌ای عبارت‌اند از:

* در سطح ۰/۰۵ اختلاف معنی‌داری وجود دارد.

۱- امکان دارد تیمار مضر بوده (مانند سیگار یا گاز سمی خاصی) و از نظر اخلاقی نتوان بر روی واحدهای انسانی این تیمارها را اعمال نمود.

۲- تیمار بر طبق یک مکانیزم سیاسی (که چه بسا مناسب هم باشد) کنترل می‌شود و لذا آزمایش‌گر هیچ‌گونه کنترلی بر آن ندارد، نظیر بسیاری از سیاست‌های پولی و اقتصاد کلان.

۳- تیمار فراتر از اختیارات قانونی حتی دولت قرار داشته و لذا امکان کنترل آن وجود نداشته باشد، مانند بسیاری از تصمیمات مدیریتی در یک بخش خصوصی.

۴- امکان دارد واحد آزمایش آن‌چنان به تیمار خاصی گرایش و وابسته باشد که اعمال کنترل توسط آزمایش‌گر را صلب نماید.

در تمامی موارد فوق، پژوهش‌گر کنترلی بر تخصیص تیمارها نداشته و به تبع آن نمی‌تواند از مقایسه‌پذیری واحدهای دریافت‌کننده تیمارهای مختلف اطمینان حاصل کند. از این‌رو وقوع پدیده‌ی اریبی در برآورد اثر تیمارها در مطالعات مشاهده‌ای امری اجتناب‌ناپذیر است. در مواردی که متغیرهای کمکی ماقبل تیماری اندازه‌گیری شده‌اند و اطلاعات آن‌ها در دست باشد، استفاده از روش‌های مرسوم کاهش اریبی که هدف اصلی این طرح پژوهشی است، می‌تواند بخش قابل توجهی از این اریبی را حذف کند. این روش‌ها عبارت‌اند از نمونه‌جورسازی شده، زیر طبقه‌بندی و روش‌های مدل مبنا (تعدیل رگرسیونی یا تعدیل کوواریانس) و در نهایت استفاده از نمره‌ی تمایل^۱. با این وجود امکان دارد هنوز اختلافات مشاهده شده‌ی پاسخ‌ها، ناشی از اختلاف متغیرهای دیگری باشد که اندازه‌گیری نشده و لذا با استفاده از روش‌های کاهش اریبی مزبور، اریبی برآورد اثر تیمار کاملاً رفع نشده باشد.

دو روش عمده‌ی ارزیابی اریبی ناشی از متغیرهای ماقبل تیماری مشاهده نشده عبارت‌اند از تحلیل حساسیت و آزمون X -تعدیل‌پذیری تخصیص تیمار.

¹ Propensity Score

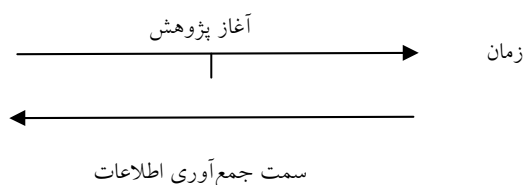
در فصل‌های آینده روش‌های کاهش آریبی برای متغیرهای ماقبل تیماری مشاهده شده و نیز روش‌های ارزیابی ناشی از متغیرهای ماقبل تیماری مشاهده نشده به تفصیل مورد بررسی قرار می‌گیرند.

۱-۲ انواع مطالعات مشاهده‌ای

مطالعات مشاهده‌ای به چند گروه اصلی تقسیم می‌شوند که این بخش به معرفی آن‌ها، ویژگی‌ها و محاسن و معایب گروه‌های اصلی آن اختصاص یافته است.

۱-۲-۱ مطالعات گذشته‌نگر^۱

مطالعاتی که در آن تیمارها (که گروه‌های مقایسه‌ای را مشخص می‌کنند) به‌عنوان معلول یا اثر و پاسخ مشاهده شده علت مفروض قلمداد می‌شوند (مک کینلی^۳ ۱۹۷۵). نقطه آغاز در این گروه از مطالعات معلول یا پیامد مورد نظر بوده و پژوهش‌گر در پی علت یا علل بروز آن می‌باشد. معمولاً این مطالعات را می‌توان با سؤال "چه روی داده است؟" توصیف نموده و سمت و سوی پژوهش رو به گذشته است (شکل ۱-۱). بنابراین به‌طور خلاصه پژوهش‌گر در مطالعات گذشته‌نگر یا مورد-شاهد از معلول در پی کشف علت است.



شکل ۱-۱

¹ Retrospective

^۲. این گروه از مطالعات تحت عنوان مطالعات "مورد-شاهد"، "مورد کنترل" و نیز (ex post facto) شناخته می‌شوند که هیچ یک از این اصطلاحات مورد اتفاق عموم نیستند (روزنباوم (۲۰۰۲)، مک کینلی (۱۹۷۵)، ککران (۱۹۶۵)).

³ Mc Kinlay

مطالعات گذشته‌نگر مخصوصاً در مواردی که معلول مشاهده شده به نسبت نادر و کمیاب باشد و همچنین در مواقعی که فاصله زمانی بین علت و معلول طولانی است، روش بسیار مناسبی به‌عنوان یک مطالعه مشاهده‌ای قلمداد می‌شود.

مثال ۱-۲- مقایسه علائم بالینی بیماران مبتلا به سندروم کارسینوئید.

پژوهش‌گران یک مطالعه "مورد-شاهد" را انجام داده‌اند که هدف آن بررسی وجود علائم و نشانه‌هایی است دال بر آن‌که بیماران مبتلا به بیماری کارسینوئید قلبی (موردها) را از بیماران غیر مبتلا (شاهدها) از نظر علائم بالینی بتوان تفکیک نمود (راس و ربرت ۱۹۸۵). به این منظور پرونده پزشکی و یافته‌های ۳۶ بیمار مبتلا به سندروم کارسینوئید مورد بررسی قرار گرفته است. از این میان ۲۱ بیمار (۵۷٪) به بیماری مزبور مبتلا بوده و ۱۵ بیمار (۴۳٪) به این بیماری مبتلا نبوده‌اند. در گروه بیماران مبتلا به کارسینوئید قلبی موارد زیر بیش‌تر مشاهده شده است: تنگی دریچه ریه، بزرگی قلب در رادیوگرافی قفسه‌ی صدری، پایین بودن ولتاژ در الکتروکاردیوگرافی. از طرف دیگر بین این دو گروه از نظر فلاشینگ، مقدار فشار خون سیستولیک یا طول مدت بیماری اختلاف معنی‌داری در سطح ۰/۰۵ مشاهده نشده است. این در حالی است که مقدار دفع ادراری HIAA - ۵ (بر حسب میلی‌گرم در ۲۴ ساعت) در میان دو گروه از بیمارانی که این متغیر برای آن‌ها اندازه‌گیری شده است (شامل ۱۶ بیمار مبتلا و ۱۲ بیمار غیرمبتلا) تفاوت معنی‌داری مشاهده می‌شود. (جدول ۱-۲)

جدول ۱-۲- خلاصه یافته‌های بالاترین مقدار دفع ادراری HIAA - ۵

بیماران غیر مبتلا به کارسینوئید قلبی	بیماران مبتلا به کارسینوئید قلبی	
۲۲۰/۶	۴۲۹	میانگین
۲۶۱/۸	۲۹۴/۷	انحراف معیار

آیا این تفاوت را به‌عنوان یک عامل علی در دو گروه بیماران می‌توان به‌عنوان یک قاعده تعمیم داده و گروه‌های تیماری مبتلا و غیر مبتلا را بر اساس مشاهده‌ی علت از یکدیگر تشخیص داد؟

گاهی یک پیشامد غیرمترقبه رخ می‌دهد، و در پی آن این سؤال طرح می‌شود که "چه عاملی باعث بروز این پدیده شده است؟". در استراژی گذشته‌نگر در ابتدا بایستی گروه‌هایی را که در متغیر پاسخ Y اختلاف دارند، مشخص نموده (که همان گروه‌های تیماری هستند)، سپس در پی آزمون این فرض بود که آیا این گروه‌ها در متغیرهای علی مورد نظر تفاوت دارند یا خیر؟ البته شناسایی و ساخت گروه‌های تیماری همواره به سادگی انجام نمی‌پذیرد. در هر صورت یکی از پرکاربردترین موارد استفاده از استراژی گذشته‌نگر مقایسه بین موفقیت‌ها و شکست‌ها در یک شغل و حرفه‌ی به‌خصوص به آن امید است که بتوان عوامل و علل آن را مشخص نموده یا لااقل پیش‌بینی‌کننده‌های مفیدی را برای موفقیت و شکست بدست آورد. به‌عنوان ساده‌ترین حالت، فرض کنید علت و پیشامد مورد نظر هر دو، متغیرهایی دو وضعیتی باشند. در این صورت برای N عضو جامعه‌ای متناهی می‌توان جدول زیر را تشکیل داد.

کل	عدم وقوع	وقوع	پیشامد
			علت
$N_{1.}$	N_{12}	N_{11}	وجود علت
$N_{2.}$	N_{22}	N_{21}	عدم وجود علت
$N_{..}$	$N_{.2}$	$N_{.1}$	کل

در روش گذشته‌نگر، در حقیقت نمونه‌ای از ستون‌های جدول فوق یعنی از گروه‌های متناظر با وقوع پیشامد و عدم وقوع آن انتخاب می‌شود. به وضوح اگر پیشامد مورد نظر نادر باشد، نمونه‌گیری از ستون‌های جدول مزبور مخصوصاً ستون وقوع، در مقایسه با استراتژی نمونه‌گیری سطوح متناظر با گروه‌های علی بسیار کم هزینه‌تر است. زیرا در نمونه‌گیری سطری برای دست‌یابی به نسبت معقولی از موارد وقوع پیشامد در هر یک از گروه‌های علی، از آن‌جا که پیشامد مورد نظر نادر فرض شده است حجم بالایی از نمونه مورد نیاز خواهد بود. به‌علاوه در روش گذشته‌نگر حتی می‌توان از ستون وقوع پیشامد نادر، با نرخ بالاتری نمونه‌گیری کرده به گونه‌ای که حتی کل