

دانشور

پزشکی

بررسی انسجام روابط در کارآزمایی‌های بالینی تصادفی نشده و مطالعات مشاهده‌ای

نویسندگان: دکتر حسین فلاح‌زاده^{۱*}، دکتر غلامرضا بابایی^۲، دکتر انوشیروان کاظم‌نژاد^۲ و دکتر محمدرضا مشکانی^۳

۱. استادیار گروه آمار و اپیدمیولوژی دانشگاه علوم پزشکی یزد و دانش‌آموخته دانشگاه تربیت مدرس

۲. دانشیار گروه آمار زیستی دانشگاه تربیت مدرس

۳. استاد گروه آمار دانشگاه شهید بهشتی

Email: hofaab@yahoo.com

* نویسنده مسئول:

چکیده

مطالعات تحلیلی مشاهده‌ای، شامل مطالعات کوهورت و مورد-شاهدی دارای مشخصه مشترکی با کارآزمایی‌های بالینی تصادفی نشده هستند و آن، عدم تخصیص تصادفی بین گروه‌های مورد مقایسه است که در حالت کلی، آن‌ها را مطالعات تصادفی نشده (Nonrandomized) می‌نامند. با توجه به این‌که تخصیص تصادفی بین گروه‌های مورد مقایسه، مبنایی برای استنباط آماری است، در تحلیل نتایج این مطالعات، آزمون‌های آماری معمول قابل استفاده نیست. در این مقاله، روش مناسب دو مرحله‌ای برای تحلیل نتایج این مطالعات ارائه شده است. در مرحله اول از یک آزمون چند متغیره ناپارامتری، برای بررسی انسجام (coherence) رابطه مورد نظر، شامل تأثیر روش درمانی یا عامل در معرض در مطالعات تصادفی نشده استفاده می‌کنیم، زیرا انسجام یک ارتباط یکی از معیارهای علی بودن روابط در مطالعات مشاهده‌ای است. در مرحله بعد به دلیل عدم تخصیص تصادفی بین گروه‌های مورد مقایسه، در مطالعات مشاهده‌ای ممکن است اریبی‌های مخفی، نتایج مطالعه را تحریف کند. بدین منظور، حساسیت نتایج به این اریبی‌ها را با استفاده از روش جدید آنالیز حساسیت بررسی می‌کنیم. نهایتاً با مثالی مناسب این دو مرحله را به صورت کاربردی ارائه می‌دهیم.

واژه‌های کلیدی: مطالعات مشاهده‌ای، آنالیز حساسیت، اریبی‌های مخفی، آزمون پوسِت (poset)

دوماهنامه علمی - پژوهشی

دانشگاه شاهد

سال سیزدهم - شماره ۶۱

اسفند ۱۳۸۴

تاریخ وصول: ۸۲/۴/۲۹

تاریخ پذیرش: ۸۳/۱۲/۲۶

۱- مقدمه

مطالعات مشاهده‌ای تحلیلی به دو دسته کوهورت و مورد-شاهدی تقسیم می‌شوند. یکی از مشکلات بزرگ این مطالعات، عدم تخصیص تصادفی بین گروه‌های مورد و شاهد است. در بعضی از کارآزمایی‌های بالینی به دلیل محدودیت‌های موجود، نمی‌توان تخصیص تصادفی بین گروه‌های مورد مقایسه انجام داد [۱].

مثلاً در جراحی به علت عدم رضایت بیماران و یا در سم‌شناسی به علت استفاده از مواد مضر، تخصیص تصادفی غیرممکن است. در حالت کلی، این دو گروه از مطالعات را به خاطر وجه اشتراک آن‌ها مطالعات تصادفی نشده می‌نامیم. وقتی نتایج مطالعات تصادفی نشده بررسی و تفسیر می‌شود، مسأله پیچیده‌ای مطرح می‌شود که آیا تفاوت معنادار ظاهر شده بین گروه‌های

و معلولی را استخراج کرد. این آریبی‌ها به دو دسته آشکار و مخفی تقسیم می‌شوند. برای کنترل آریبی‌های آشکار که ممکن است در نتیجه متغیرهای اندازه‌گیری شده در مطالعه به وجود آید، می‌توان از روش‌های تطبیق آماری، همسان‌سازی و آنالیز کوواریانس استفاده کرد. آریبی‌های مخفی در نتیجه متغیرهایی به وجود می‌آید که به آن‌ها توجهی نشده و در تحقیق اندازه‌گیری نشده‌اند. برای بررسی و کنترل آریبی‌های مخفی از روشی به نام آنالیز حساسیت استفاده می‌شود که در بخش‌های بعدی جزئیات آن را بررسی خواهیم کرد [۴].

مطالب این مقاله را به این صورت دسته‌بندی کرده‌ایم: در بخش اول جزئیات آزمون چند متغیره پوست را ارائه می‌دهیم که این آزمون تعمیمی از آزمون‌های ناپارامتری دو نمونه‌ای است، در بخش دوم آنالیز حساسیت برای بررسی آریبی‌های مخفی ارائه می‌شود. در بخش سوم مثالی را برای روشن کردن کاربرد آزمون پوست و آنالیز حساسیت ارائه خواهیم داد. بخش چهارم نیز شامل بحث و نتیجه‌گیری است.

۲- یک آزمون برای بررسی انسجام در مطالعات تصادفی نشده (معرفی آزمون پوست)

در حالت کلی این آزمون برای بررسی انسجام (Coherence) در گروه‌های مورد مقایسه به کار می‌رود، زیرا انسجام یک ارتباط که بین تیمار یا عامل در معرض بودن وجود دارد، یکی از معیارهای مهم برای علی بودن یک رابطه است. علاوه بر این، در حالت کلی این آزمون چند متغیره را به راحتی می‌توان برای مقایسه دو گروه که دارای چند متغیر پاسخ است نیز به کار برد. این آزمون در ارتباط نزدیک با آزمون من ویتنی - ویلکاکسون و آزمون گهان (Gehan) برای داده‌های سانسور شده است.

یک آزمایش بر روی N فرد را در نظر بگیریم که n نفر آن‌ها در گروه اول (گروه مورد) و $N-n$ نفر آن‌ها در گروه دوم (گروه شاهد) قرار دارند. برای این که نشان دهیم هر عضو در گروه مورد یا

مورد مقایسه مربوط به اثر واقعی روش درمانی یا عامل در معرض بوده، یا به دلیل خطاهای حاصل از عدم تخصیص تصادفی بین گروه‌ها است [۲].

در تفسیر و تحلیل نتایج مطالعات غیرتصادفی، به دلیل عدم تخصیص تصادفی بین گروه‌های مورد مقایسه، مبنای لازم برای استفاده از آزمون‌های آماری از بین می‌رود، زیرا گروه‌ها ممکن است قبل از مقایسه با یکدیگر یکسان نباشد. برای رسیدن به یک استنباط علی، معیارهایی توسط هیل در سال ۱۹۶۰ ارائه شده است که این معیارها هم اکنون نیز برای تحلیل مطالعات غیرتصادفی به کار می‌رود. یکی از این معیارها وجود انسجام و معقول بودن یک ارتباط است. انسجام به این مفهوم است که ما باید در یک مطالعه تحلیلی، فرضیه‌ای را که برای تأثیر یک عامل یا روش درمانی ارائه می‌دهیم با تئوری علمی آن موضوع تضادی نداشته و با آن هماهنگ باشد و به بیانی دیگر، در فرضیه ارائه شده یک مکانیسم علمی دقیق از تأثیر عامل یا روش درمانی مورد استفاده داشته باشیم. بنابراین انسجام ارتباط بین یک عامل در معرض یا یک روش درمانی با متغیرهای پاسخ، ملاک مهمی برای قضاوت در مورد علت و معلولی بودن رابطه مشاهده شده است. البته در اکثر کتاب‌های اپیدمیولوژی و مقالات، انسجام را به عنوان معیاری برای تفسیر علی نتایج به کار برده‌اند [۲].

در یک مطالعه مشاهده‌ای یا کارآزمایی بالینی تصادفی نشده وقتی می‌خواهیم چند متغیر پاسخ را به‌طور همزمان در دو گروه مقایسه کنیم، به‌منظور بررسی انسجام روابط از یک آزمون ناپارامتری چند متغیره به نام آزمون پوست که توسط روزنهام طراحی شده استفاده می‌کنیم [۳].

بعد از محاسبه نتایج آزمون پوست به‌منظور اعتبار بیش‌تر نتایج و توجه به این مسأله که در مطالعاتی که تخصیص تصادفی بین گروه‌ها وجود ندارد پتانسیل خاصی برای آریبی در نتایج وجود دارد و به علت وجود این‌گونه آریبی‌ها نمی‌توان به راحتی روابط علت

ندارد، یعنی اگر برای فرد i ام پاسخ y_i را مشاهده کنیم فرقی نمی‌کند که در گروه مورد یا گروه شاهد. بنابراین تحت فرضیه H_0 با تغییر نشانگر Z_i از ۰ به ۱، پاسخ y_i تغییری نمی‌کند، یعنی y_i ها مقادیر ثابتی از جامعه متناهی شامل N واحد هستند.

فرض کنیم متغیرهای پاسخ به صورت تعداد، مقدار عددی پیوسته و یا یک حادثه دوحالتی باشد. حالتی که متغیر پاسخ y_i به صورت یک متغیره باشد، واضح است. در مقایسه مشاهدات گروه مورد و شاهد وقتی از کم‌تر یا بیش‌تر صحبت می‌کنیم، منظور یک رابطه بزرگ‌تر یا کوچک‌تر معمولی است. در حالت پیچیده‌تر، وقتی متغیرهای پاسخ به صورت چندمتغیره باشند، در مقایسه مشاهدات به صورت برداری ممکن است بعضی متغیرهای پاسخ در یک بردار نسبت به بردار دیگر بزرگ‌تر و بعضی کوچک‌تر باشند و نتوانیم همه بردارهای پاسخ را مقایسه و مرتب کنیم. برای مثال بیمار ۱ ممکن است بقای بیش‌تر و کیفیت زندگی بهتری نسبت به بیمار ۲ و ۳ داشته باشد. اما بیمار ۲ ممکن است بقای بیش‌تر نسبت به بیمار ۳، ولی کیفیت زندگی پایین‌تر داشته باشد. در این حالت، پاسخ‌ها به جای این که به صورت کلی مرتب شود می‌تواند به صورت جزئی مرتب گردد [۴].

بنابراین برای تعریف فرضیه جانشین انسجام در مقابل فرضیه صفر عدم تأثیر تیمار از یک رابطه مرتب جزئی که بر روی متغیرهای پاسخ تعریف می‌شود استفاده می‌کنیم. برای سادگی، پاسخ‌های چند متغیره را به صورت برداری نشان می‌دهیم.

در تعریف رابطه مرتب جزئی باید حداقل یکی از نامساوی‌ها به صورت اکید باشد و این رابطه باید طوری تعریف شود که خواص مجموعه مرتب جزئی را دارا باشد.

بنابراین به طور رسمی فرضیه انسجام (Coherent) را به صورت رابطه (\prec_c) تعریف می‌کنیم و رابطه (\prec_c) بر روی مجموعه متغیرهای پاسخ به صورت زیر تعریف می‌شود، جایی که:

شاهد قرار دارد از یک متغیر نشانگر Z به صورت زیر استفاده می‌کنیم:

$$Z_i = \begin{cases} 1 & \text{فرد } i \text{ متعلق است به گروه مورد} \\ 0 & \text{فرد } i \text{ متعلق است به گروه شاهد} \end{cases}$$

بنابراین Z_i ها برای N عضو مورد بررسی تشکیل یک بردار را می‌دهند و داریم:

$$Z^T = (Z_1, Z_2, \dots, Z_N)^T, n = \sum_{i=1}^N Z_i$$

مجموعه Ω را که شامل کلیه ترکیب‌های ممکن قرار گرفتن n عضو در گروه مورد از کل N عضو است و هر بردار Z نشان‌دهنده یک ترکیب است در نظر می‌گیریم. این مجموعه دارای $\binom{N}{n}$ عضو است و هر ارائه یا ترکیب دارای N درایه است که n تا از آن‌ها یک و بقیه صفرند.

اگر در این آزمایش افراد یا اعضا به طور تصادفی به گروه‌ها تخصیص داده شده باشند، احتمال انتخاب هر ترکیب در Ω برابر است با:

$$P(Z=z) = \frac{1}{\binom{N}{n}}, \forall z \in \Omega \quad (1)$$

ولی در مطالعاتی که تخصیص تصادفی بین گروه‌های مورد مقایسه نباشد این رابطه برقرار نیست.

متغیرهای پاسخ را به صورت زیر تعریف می‌کنیم: فرض کنیم هر فرد K پاسخ y_1, \dots, y_k را ارائه می‌دهد که آن‌ها را در یک بردار k بعدی Y جای می‌دهیم. مجموعه مقادیر ممکن y را با Y نشان می‌دهیم. پس $y \in Y \subseteq R^k$ به صورت یک بردار K بعدی خواهد بود. حال بردار مشاهدات به صورت زیر تعریف می‌شود:

$$Y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1k} \end{bmatrix}, \dots, Y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{ik} \end{bmatrix}, \dots, Y_N = \begin{bmatrix} y_{N1} \\ y_{N2} \\ \vdots \\ y_{Nk} \end{bmatrix}$$

اکنون می‌خواهیم این فرضیه را آزمون کنیم که تیمار یا عامل در معرض بودن در گروه مورد تأثیری

تحت فرضیه H_0 ، بردار y_i ها ثابت هستند و لذا U_{ij} ها نیز ثابتند، زیرا آن‌ها تابعی از y_i ها محسوب می‌شوند. با استفاده از U_{ij} ها ماتریس $U_{N \times N}$ را تشکیل می‌دهیم و U_{ii} یعنی درایه‌های قطر ماتریس U با توجه به تعریف برابر صفر است و داریم:

$$U_{ij} = -U_{ji} \text{ و آماره آزمون عبارت است:}$$

$$T = \sum_{i=1}^N \sum_{j=1}^N Z_i(1-Z_j)U_{ij} = Z^T U(1-Z) = Z^T U1 = Z^T S$$

در این جا هدف، آزمون فرضیه صفر عدم تأثیر عامل در معرض در برابر فرضیه جانشین که عبارت است از اعضای گروه مورد پاسخ بالاتری نسبت به گروه کنترل با توجه به رابطه مرتب جزئی بر روی متغیرهای پاسخ دارند [۱].

حال اگر U_{ij} واحد گروه مورد و U_{ji} واحد گروه شاهد را در نظر بگیریم عبارت $Z_i(1-Z_j)U_{ij}$ برابر یک است اگر واحد U_{ij} پاسخ بالاتری نسبت به واحد U_{ji} با توجه به رابطه ترتیبی جزئی یعنی $y_j <_c y_i$ داشته باشد. متشابهاً عبارت فوق مقدار -1 را دارد اگر واحد در گروه کنترل پاسخ بالاتری با توجه به رابطه ترتیبی جزئی $y_i <_c y_j$ داشته باشد و در سایر موارد عبارت فوق مقدار صفر را دارا است.

داریم $S=U1$ و حاصلضرب بردارهای زیر برابر صفر است:

$$Z^T U Z = 0$$

$$1^T U 1 = 0$$

$$1^T S = 0$$

در این جا، بردار S شامل امتیازهایی که از جمع سطرهای بردار $U_{N \times N}$ به دست می‌آید و مقدار S_i عبارت از تعداد واحدهایی که مقادیر بردار پاسخ با توجه به رابطه انسجام تعریف شده کم‌تر از واحد U_{ii} است، منهای تعداد واحدهایی که دارای پاسخ‌های بزرگ‌تر از i است. می‌توانیم آماره آزمون را به صورت زیر ساده‌تر کنیم:

$$T = \sum_{i=1}^N Z_i S_i, \quad S_i = \sum_{j=1}^N U_{ij}$$

$$(2) \quad y, y^* \in Y \quad \text{برای } y <_c y^*$$

و به این معنا است که پاسخ y نسبت به پاسخ y^* سازگاری کم‌تری با فرضیه منسجم که در آن تأثیر تیمار را با توجه به متغیرهای پاسخ ارائه کردیم، دارد و اگر رابطه:

$$(3) \quad y, y^* \in Y \quad \text{برای } y <_c y^*$$

برقرار باشد، عکس این مطلب صادق است، یعنی y^* سازگاری کم‌تری با فرضیه منسجم تیمار نسبت به پاسخ y دارد. بنابراین اثر منسجم تیمار را به صورت روابط کوچک‌تر یا بزرگ‌تر و یا مساوی بین اجزای بردار پاسخ تعریف می‌کنیم.

حال با توجه به رابطه انسجام مورد نظر و تحت فرض صفر تعریف می‌کنیم:

$$U_{ij} = \begin{cases} -1 & y_i <_c y_j \\ 1 & \text{اگر } y_j <_c y_i \\ 0 & \text{Else} \end{cases}$$

U_{ij} ها طوری تعریف شده‌اند که رابطه $(<_c)$ به صورت متقارن است و سه حالت مختلفی که برای U_{ij} رخ می‌دهد، متقابلاً از یکدیگر جدا بوده، اشتراکی ندارند.

به عنوان مثال اگر بردار y_i دارای سه متغیر پاسخ باشد:

$$y_i = (y_{i1}, y_{i2}, y_{i3})$$

رابطه منسجم ممکن است به صورت زیر تعریف شود:

$$\forall i, j, \quad y_i <_c y_j \Leftrightarrow (y_{i1} < y_{j1}, y_{i2} < y_{j2}, y_{i3} < y_{j3})$$

در این حالت، وقتی روش درمانی مؤثر است که هر سه متغیر پاسخ در گروه مورد سطح بالاتری نسبت به این متغیرها در گروه شاهد داشته باشد.

در حالت کلی دو بردار پاسخ y و y^* نسبت به یکدیگر یا به صورت رابطه ۲ یا رابطه ۳ و یا مساوی هستند، البته ممکن است هیچ کدام از این روابط نیز برای آن‌ها صادق نباشد. حال با توجه به همین روابط می‌توانیم پاسخ‌ها را دسته‌بندی کنیم:

$$\text{حالت اول: } y_i <_c y_j$$

$$\text{حالت دوم: } y_j <_c y_i$$

$$\text{حالت سوم: } y_i <_c y_j, y_j <_c y_i$$

شانس متفاوتی برای قرار گرفتن در گروه‌های مورد مقایسه داشته باشند و در نتیجه ممکن است تعدادی اریبی‌های مخفی به وجود آید که این مسأله را در بخش بعدی مطرح خواهیم کرد.

۳- یک مدل برای توزیع تخصیص افراد به گروه‌های مورد مقایسه در مطالعات مشاهده‌ای (آنالیز حساسیت)

یک مشکل اساسی در مطالعات مشاهده‌ای یا کار آزمایشی‌های بالینی تصادفی نشده، وجود اریبی‌های مخفی است. این اریبی‌ها در نتیجه متغیرهای که در مطالعه به آن‌ها توجهی نشده و اندازه‌گیری نشده‌اند به وجود می‌آید. این متغیرهای اندازه‌گیری نشده را کووریت‌های مشاهده نشده می‌نامیم. در مطالعات مشاهده‌ای، چون تخصیص تصادفی بین گروه‌های مورد مقایسه وجود ندارد و لذا بعضی افراد یا واحدهای آزمایشی شانس بیش‌تری برای دریافت عامل در معرض یا روش درمانی مورد نظر دارند و این شانس برای افراد یکسان نیست بنابراین مبنای لازم برای انجام آزمون‌های آماری معمول وجود ندارد [۶]. با استفاده از مدلی که در این قسمت برای آنالیز حساسیت ارائه می‌شود اندازه انحراف نتایج از تخصیص تصادفی بین گروه‌ها را که در رابطه ۱ بیان شد مورد بررسی قرار می‌دهیم. جزئیاتی که در بخش اول برای آزمون پوسْت ارائه شد در نظر می‌گیریم.

اگر S_i مقدار رتبه‌های حاصل از متغیرهای پاسخ فرد i ام باشد، عموماً هدف ما بررسی فرضیه صفر عدم تأثیر روش درمانی یا عامل در معرض است و با قبول فرضیه صفر، S_i ها مقادیر ثابتی از یک جامعه محدود با N عضو هستند. بردار N بُعدی Z و S که شامل مقادیر Z_i ها و S_i ها است را در نظر می‌گیریم.

اگر Ω نشان‌دهنده کلیه حالات ممکن تخصیص تصادفی تیمارها برای n عضو در گروه مورد باشد. در یک آزمایش یا طرح کاملاً تصادفی شده اگر تخصیص واحدها به تیمارها را با بردار Z نشان دهیم بنابراین

یعنی آماره آزمون برابر مجموع امتیازات افراد در گروه مورد است. آماره آزمون یک آماره مجموع رتبه است و در حالت یک متغیره این آماره تشابه زیادی با آماره گهان برای داده‌های سانسور شده دارد [۱۱ و ۱۲].

تحت فرضیه صفر در یک آزمایش که تخصیص تصادفی بین گروه‌ها وجود دارد آماره آزمون $T=Z^T S$ به سادگی برابر مجموع n امتیاز ثابت که به وسیله نمونه‌گیری تصادفی ساده بدون جایگذاری از جامعه‌ای شامل N امتیاز انتخاب شده‌اند به دست می‌آید، و با توجه به این که T یک آماره مجموع رتبه (Rank sum test) است و Z_i ها تصادفی‌اند و S_i ها مقدار ثابتی دارند، به آسانی دیده می‌شود که Z_i ها دارای توزیع برنولی هستند و چون $E(Z_i)$ در توزیع برنولی برابر پارامتر آن در این جا n/N است و چون $\sum_{i=1}^N S_i$ برابر صفر است پس $E(T)=0$ خواهد بود.

با توجه به خواص توزیع برنولی و تعریف واریانس، واریانس $Z^T S$ نیز به صورت زیر است [۱].

$$\text{var}(Z^T S) = \frac{n(N-n)}{N(N-1)} \sum_{i=1}^N S_i^2$$

حال پس از محاسبه امید ریاضی و واریانس و با توجه به خواص آماره‌های رتبه‌ای خطی و نظریه این آماره‌ها و با استفاده از قضیه حد مرکزی تقریبی وجود دارد که می‌توان آن را برای توزیع تقریبی تمام آماره‌های رتبه‌ای خطی به کار برد. چون T ترکیبی خطی از Z_i ها است و Z_i ها هم، متغیرهایی تصادفی با توزیعی همانندند، لذا تعمیمی از قضیه حد مرکزی اجازه می‌دهد که نتیجه بگیریم توزیع احتمال آماره رتبه‌ای خطی استاندارد شده، تحت برخی شرایط به توزیع نرمال میل می‌کند. بنابراین در حجم‌های بالا آماره آزمون پوسْت دارای توزیع مجانبی نرمال است و در عمل می‌توانیم از توزیع مجانبی برای انجام آزمون استفاده کنیم [۱ و ۵].

حال وقتی نتایج یک مطالعه تصادفی نشده را بررسی می‌کنیم، مدل احتمال تخصیص تصادفی گروه‌ها یا رابطه ۱ صادق نیست، یعنی افراد ممکن است دارای

با توجه به این‌که متغیرهای تصادفی Z_j که به صورت دودویی هستند، یک فرم لجیت را برای ارتباط دادن آن‌ها به متغیرهای کمکی (X_j, U_j) در نظر می‌گیریم یعنی:

$$\log\left(\frac{\pi_j}{1-\pi_j}\right) = K(x_j) + \gamma U_j$$

البته در این حالت π_j ها تابعی از متغیرهای کمکی مشاهده شده و مشاهده نشده می‌باشند. جایی که $K(x)$ یک تابع نامعلوم و γ نیز یک پارامتر نامعلوم است. مدل فوق می‌گوید، تخصیص تیمارها برای افراد مختلف، مستقل است و احتمال دریافت تیمار (واقع شدن در گروه مورد)، یا قرار گرفتن در معرض عامل مورد نظر دارای یک فرم لوجیت و با جور کردن یا طبقه‌بندی بر روی X تابعی از u_i است.

اگر $\gamma > 0$ باشد، افراد دارای مقادیر بزرگ‌تر از متغیرهای کمکی مشاهده نشده u_i احتمال بیش‌تری برای گرفتن تیمار یا عامل در معرض بودن مورد نظر را نسبت به گروه شاهد دارند.

نسبت $\frac{\pi_i}{1-\pi_i}$ را بخت دریافت تیمار برای فرد i ام می‌نامیم. فرض کنیم دو واحد z و k دارای مقادیر متغیرهای کمکی مشاهده شده یکسان باشند. بنابراین داریم $X_j = X_k$ و نیز $K(x_k) = K(x_j)$ و با توجه به یکسان بودن مقادیر X یا دو واحد با هم در یک طبقه قرار دارند و یا با یکدیگر جور شده‌اند. در هر دو حالت، مقدار X در این دو واحد یکسان می‌شود. حال نسبت شانس واحد z ام به واحد k ام برای دریافت تیمار عبارت است از:

$$\frac{\pi_j (1 - \pi_k)}{\pi_k (1 - \pi_j)} = \exp \left\{ \gamma (u_j - u_k) \right\}$$

حال با توجه به خواص توزیع لجیت تابع π_j به شکل زیر تعریف می‌شود:

$$\pi_j = \frac{e^{(K(x_j) + \gamma u_j)}}{1 + e^{(K(x_j) + \gamma u_j)}}$$

بردار Z به وسیله آزمایشگر به طور تصادفی از Ω انتخاب می‌شود و رابطه ۱ را خواهیم داشت. با توجه به این‌که نتایج مطالعات مشاهده‌ای ممکن است از دو جهت، شامل متغیرهای کمکی معلوم و پنهان دچار اریبی شوند فرض کنید در یک مطالعه مشاهده‌ای N واحد تحت بررسی‌اند و هر یک از واحدها دارای یک مقدار متغیر کمکی مشاهده شده X هستند. البته X ممکن است شامل چندین متغیر کمکی باشد. فرض کنیم x_j مقدار متغیر کمکی برای j امین واحد و متغیر نشانگر Z_j ، نشان‌دهنده تخصیص واحدها به گروه‌های مورد و شاهد است.

برای رسیدن به یک مدل مناسب، تصور می‌کنیم واحد j ام با احتمال π_j در گروه مورد قرار می‌گیرد، یعنی $\pi_j = P(Z_j=1)$ و احتمال قرار گرفتن واحد j ام در گروه کنترل برابر $1-\pi_j$ است.

در این حالت، مدل را این‌گونه توصیف می‌کنیم واحدها به وسیله پرتاب یک سکه اریب به تیمارها یا گروه‌های مورد مقایسه تخصیص داده می‌شوند و حتی ممکن است برای هر واحد یک سکه اریب داشته باشیم که هر سکه دارای اندازه اریبی متفاوت و نامشخص است.

بنابراین مقادیر π_j ها مجهول است، زیرا مقادیر اریبی برای هر واحد نامشخص است. چون Z_j ها مستقل از یکدیگر و دارای توزیع برنولی هستند پس توزیع آن‌ها به صورت زیر است:

$$P(Z_1 = z_1, \dots, Z_N = z_N) = \prod_{j=1}^N \pi_j^{z_j} (1 - \pi_j)^{1-z_j} \quad (4)$$

بنابراین در توزیع فوق چون π_j ها نامعلوم است، توزیع Z_j ها نیز مجهول است و نمی‌توانیم سطح معناداری زیر را محاسبه کنیم:

$$p\text{-value} = P\{t(z, r) \geq k\} = \sum_{z \in \Omega} [t(z, r) \geq k] \cdot p(Z = z) \quad (5)$$

اگر U_j, X_j به ترتیب متغیرهای کمکی مشاهده شده و مشاهده نشده برای واحد j ام باشند در این حالت مدل توزیع تخصیص تصادفی تیمارها به صورت زیر است:

[۸و۷،۲]

$$P(z = z | n = \sum_{i=1}^N Z_i) = \frac{\exp(\gamma Z^T U)}{\sum_{b \in \Omega} \exp(\gamma b^T U)} \quad (6)$$

رابطه فوق، مدل احتمال را با توجه به متغیرهای کمکی مشاهده نشده (U_i) برای تخصیص تیمارها در مطالعات مشاهده‌ای یا کار آزمایشی‌های بالینی تصادفی نشده نشان می‌دهد.

اگر γ معلوم و u مشاهده شده باشد، سطح معناداری آماره آزمون $T(z, \tau)$ می‌تواند با استفاده از مدل احتمالی $\alpha(\tau, u)$ به دست آید. اگر $\alpha(\tau, u)$ سطح معناداری آزمون باشد، وقتی γ نامعلوم و u مشاهده نشده باشد، که در عمل نیز همین طور است، سطح معناداری به‌طور مستقیم قابل محاسبه نیست. برای رفع این مشکل، مقادیر مشخص و ثابتی را برای γ (پارامتر حساسیت) در نظر می‌گیریم و سپس کران‌هایی را بر روی سطح معناداری $\alpha(\tau, u)$ به دست می‌آوریم و حساسیت این کران‌ها را بر روی مقادیر قابل قبولی از γ نشان می‌دهیم. چون u مشاهده نشده مقیاس ثابتی برای آن نداریم. اگر بخواهیم نماد پارامتر حساسیت (γ) دارای مفهوم مناسب و قابل تعبیر باشد، باید تعدادی محدودیت بر روی u و یا u_i گذاشته شود. با توجه به این که برای γ مقدار ثابت و u را محدود می‌کنیم، می‌توانیم رابطه زیر را بنویسیم:

$$(c\gamma)u = \gamma(cu)$$

ساده‌ترین محدودیتی که می‌توانیم بر روی u بگذاریم به صورت $0 \leq u_i \leq 1$ است و در نتیجه $U \in [0, 1]^N$. این محدودیت باعث می‌شود بتوانیم تعبیر مناسبی برای γ داشته باشیم و به دو دلیل این محدودیت مناسب است. ابتدا، یک توصیف معقول و دقیق از رابطه توزیع احتمال که براساس متغیرهای کمکی مشاهده نشده و z و γ است ارائه می‌دهد. ثانیاً ممکن است در نتیجه عدم تعادل در U_i ها، برای عضوهایی که وارد مطالعه می‌شوند اریبی‌هایی رخ دهد. مثلاً دو فرد که برای X_i جور شده‌اند ولی مقدار U در آن‌ها متفاوت است، شانس دریافت تیمار (عامل در معرض بودن) با عامل e^γ نشان داده می‌شود و این

با توجه به این که Z_j ها دارای توزیع برنولی‌اند بنا به تعریفی که برای π_j ها در نظر گرفتیم می‌توانیم توزیع Z_j ها را به فرم زیر بنویسیم:

$$p(Z_1 = z_1, \dots, Z_N = z_N) = \prod_{i=1}^N [\pi(u_j)]^{z_j} [1 - \pi(u_j)]^{1 - z_j}$$

حال با توجه به این که مقادیر متغیرهای کمکی مشاهده شده معلومند، می‌توانیم با طبقه‌بندی مناسب بر روی آن‌ها یا جور کردن آن‌ها در گروه مورد و شاهد، مدل احتمالی را ساده‌تر کنیم. با محاسبه توزیع Z به شرط X ، نتیجتاً مدل برحسب U_j ها یا متغیرهای کمکی مشاهده نشده به دست می‌آید.

حال اگر برای تطبیق بر روی X از روش طبقه‌بندی استفاده کنیم، باید S طبقه مجزا بر مبنای متغیر کمی مشاهده در نظر بگیریم. فرض کنیم که هر طبقه دارای n_s واحد است و m_s نیز تعداد واحدهایی را که در طبقه S ام در گروه مورد قرار گرفته‌اند نشان می‌دهد.

اگر واحد i ام در طبقه s ام دارای تخصیص تصادفی Z_{si} و متغیر کمکی X_{si} باشد بنابراین بردار Z دارای N داریه به صورت $(Z_{11}, \dots, Z_{s, n_s})^T$ است و $m_s = \sum_{i=1}^{n_s} Z_{si}$ و بردار m تعداد واحدها در گروه مورد را در همه طبقات به صورت زیر نشان می‌دهد:

$$m = (m_1, \dots, m_s)$$

پس از طبقه‌بندی بر روی X ، طبقات به دست آمده نسبت به X همگن هستند. در این جا m_s ها یک متغیر تصادفی‌اند. حال اگر توزیع شرطی Z را نسبت به m محاسبه کنیم، شکل توزیع تابعی از γ و U_i ها به دست خواهد آمد و نتیجتاً توزیع Z به شکل زیر تبدیل می‌شود:

$$p(Z = z | m) = \prod_{s=1}^S \frac{\exp(\gamma Z_s^T U_s)}{\sum_{b_s \in \Omega_s} \exp(\gamma b_s^T U_s)}$$

Ω_s یک مجموعه شامل $\binom{n_s}{m_s}$ تا بردار Z که شامل n_s تا یک و $n_s - m_s$ تا صفر است در هر طبقه است.

در حالتی که $S=1$ باشد، بنابراین $n = \sum_{i=1}^n Z_i = m$ است و رابطه فوق به صورت زیر ساده می‌شود: [۱]

در حالتی که بخواهیم برای آزمون ناپارامتری دو نمونه‌ای از توزیع مجانبی استفاده کنیم، باید با توجه به مدل آنالیز حساسیت امید ریاضی و واریانس آماره آزمون ناپارامتری دو نمونه‌ای را محاسبه و سپس از توزیع مجانبی نرمال استفاده کنیم.

۴- یک مثال

به منظور آشنایی بیش‌تر با آزمون پوست و آنالیز حساسیت مثالی را از یک مطالعه مشاهده‌ای ارائه می‌دهیم. این مطالعه به منظور بررسی ارتباط بین هیپرلیپیدمی فرزندان زیر ۱۸ سال با بیماری عروق کرونر والدین آن‌ها در شهر یزد در سال ۱۳۷۹ انجام شده است. در این بررسی تعداد ۲۵ نفر از فرزندان والدینی که سن آن‌ها کم‌تر از ۵۵ سال و دچار سکته قلبی شده بودند، به عنوان گروه مورد و تعداد ۲۰ نفر از بچه‌هایی که والدین آن‌ها از نظر بیماری‌های عروق کرونر سالم و از نظر سنی و جنسی مشابه گروه مورد بودند به عنوان گروه شاهد در نظر گرفته شدند. متغیرهای پاسخ شامل اندازه تری گلیسرید (y_1)، کلسترول LDL (y_2) و کلسترول HDL (y_3) بود و مقادیر آن‌ها در دو گروه اندازه‌گیری شد. با توجه به این‌که در این مطالعه تخصیص تصادفی بین گروه‌های مورد و شاهد وجود ندارد بنابراین ابتدا یک فرضیه مناسب، با توجه به ارتباط بین متغیرهای پاسخ می‌سازیم. البته این فرضیه باید از نظر علمی معقول و منطقی باشد که برای تأیید آن از آزمون پوست استفاده می‌کنیم. فرضیه مقابل منسجم عبارت است از: فرزندان افراد مبتلا به انفارکتوس قلبی، همزمان سطح بالاتری از تری گلیسرید، کلسترول LDL و سطح پایین تری از کلسترول HDL نسبت به افراد گروه شاهد دارند و فرم ریاضی رابطه انسجام به صورت زیر است:

$$y_i <_c y_j \Leftrightarrow y_{i1} < y_{j1}, y_{i2} < y_{j2}, y_{i3} < y_{j3}$$

حال با استفاده از آزمون چند متغیره پوست به بررسی رابطه فوق می‌پردازیم. با توجه به جزئیات آزمون که در بخش قبل ارائه گردید مقدار آماره آزمون

عامل در بسیاری از مباحث آماری مورد استفاده و دارای تعبیر مناسبی است. وقتی $\gamma = 0$ افراد ممکن است دارای شانس دریافت تیمار یا عامل در معرض بودن با فاکتور $e^{\gamma=0} = 1$ باشند و بنابراین افراد گروه مورد و گروه شاهد شانس یکسانی دارند. در واقع همان طرح کاملاً تصادفی شده است. وقتی $e^{\gamma} = 2$ ، عدم تعادل در U_i ها ممکن است این نتیجه را بدهد که یک فرد یا عضو در گروه مورد نسبت به افراد در گروه شاهد دو برابر شانس بیش‌تر برای دریافت تیمار دارد و به همین صورت می‌توانیم تفسیر مشابهی برای $e^{\gamma} = k$ داشته باشیم، چون $\alpha(r, u)$ یک تابع پیوسته از u بر روی مجموعه بسته و کران‌دار شده $[0, 1]^N$ است. بنابراین سطح معناداری بر روی این مجموعه بسته در نقطه u^+ به سوپر موم و در نقطه u^- به اینفیمم می‌رسد. به عبارت دیگر داریم:

$$\forall u \in [0, 1]^N, \alpha(r, u^-) \leq \alpha(r, u) \leq \alpha(r, u^+)$$

حال باید u^+ و u^- را با توجه به ابر مکعب $[0, 1]^N$ که دارای 2^N حالت است مشخص کنیم، اما با توجه به قضایای ارائه شده ثابت می‌گردد که تنها $N-1$ نامزد برای u^+ و u^- وجود دارد و ما باید در یک مجموعه $N-1$ عضوی آن‌ها را مشخص کنیم.

با توجه به مدلی احتمالی که قبلاً ارائه شد کران بالا و پایین برای سطح معناداری مشخص می‌شود. البته چون محاسبات به صورت دستی عملی نیست می‌توانیم از یک نرم‌افزار کامپیوتری کمک بگیریم.

با توجه به موارد فوق $N-1$ عضو منتخب برای u^+ به صورت زیر است:

$$C = \{(1, 0, 0, \dots, 0), (1, 1, 0), \dots, (1, 1, \dots, 1, 0)\}$$

بردار متفاوت $U^+ \in C$ مقدار مختلفی را برای کران بالای سطح معناداری آزمون با توجه به مقدار مشخص و ثابت γ فراهم می‌کند که ما باید بزرگ‌ترین آن را در نظر بگیریم و با توجه به مقدار ثابتی که برای $w = e^{\gamma}$ پارامتر حساسیت در نظر می‌گیریم مقدار سطح معناداری را محاسبه می‌کنیم [۲، ۶، ۹، ۱۱].

با توجه به جدول فوق در حالتی که $W=1$ باشد با استفاده از آزمون پوست مقدار $p\text{-value}=0/013$ و همان سطح معناداری در حالت تخصیص تصادفی را خواهیم داشت. به عبارت دیگر، اگر تخصیص تصادفی بین گروه‌ها وجود داشت، تفاوت بین گروه مورد و شاهد معنادار بود. در حالتی که $e^{\gamma}=2$ یک مطالعه مشاهده‌ای را داریم که تفاوت اندکی با یک مطالعه تصادفی شده دارد و افراد ممکن است دارای شانس متفاوتی در قرار گرفتن در گروه مورد و شاهد باشند. در این حالت این شانس ۲ برابر است و یک فرد با احتمال دو برابر در گروه مورد قرار می‌گیرد و این امر به خاطر اریبی‌های مشاهده نشده که در اثر بعضی از متغیرهای نامعلوم ایجاد می‌شود و کران بالا برای سطح معناداری برابر $p\text{-value}=0/2070$ است. در نتیجه، آزمون در سطح $\alpha=0/05$ معنادار است و متغیرهای کمکی مشاهده نشده، نتیجه آزمون را تغییر نمی‌دهد، ولی مقدار $p\text{-value}$ را افزایش می‌دهد. در حالتی که $e^{\gamma}=4$ باشد، فاصله زیادی را از یک مطالعه تصادفی شده داریم و یک فرد با شانس ۴ برابر در گروه مورد نسبت به گروه شاهد قرار می‌گیرد در این حالت، کران بالا برای سطح معناداری برابر $p\text{-value}=0/43$ است و باز هم اریبی‌های پنهان نمی‌تواند نتایج را تحریف کند و ما تفاوت معناداری بین گروه مورد و شاهد خواهیم داشت. در حالتی که $e^{\gamma}=5$ باشد در این حالت پتانسیل بیش‌تری برای وجود متغیرهای کمکی مشاهده نشده خواهیم داشت و در واقع اریبی‌های پنهان می‌تواند نتایج مطالعه را تغییر دهد.

بنابراین می‌توانیم بگوییم که نتایج این مطالعه به اریبی‌های کوچک حساسیتی ندارد و این گونه تورش‌ها نمی‌تواند نتیجه مطالعه را تغییر دهد، ولی در حالتی که اریبی‌های بزرگ‌تری وجود داشته باشد، نتیجه مطالعه تغییر خواهد کرد. البته با توجه به مدلی که برای تخصیص تیمارها در مطالعات مشاهده‌ای ارائه دادیم عملاً سطح معناداری که در نتیجه W به دست می‌آید با تطبیق بر روی متغیرهای کمکی مشاهده نشده (U) حاصل می‌شود.

$T=64$ ، امید ریاضی $E=0$ و واریانس آماره آزمون $657,12$ است. با استفاده از توزیع مجانبی نرمال مقدار $p\text{-value}=0/013$ به دست می‌آید. بنابراین اگر تخصیص تصادفی بین گروه‌های مورد و شاهد وجود داشت، یک دلیل قوی برای قبول فرضیه جانشین در دست بود؛ دلیلی که با رابطه منسجم پیش‌بینی شده سازگاری داشته و این رابطه را تأیید می‌کرد و نتیجه می‌گرفتیم که بین سکنه قلبی والدین و هیپرلپیدمی فرزندان آن‌ها ارتباط وجود دارد، اما با توجه به این‌که داده‌ها از یک مطالعه مشاهده‌ای به دست آمده، بنابراین شانس افراد برای قرار گرفتن در گروه مورد و شاهد یکسان نیست و اریبی‌های مخفی می‌تواند نتیجه مطالعه را تغییر دهد. حال برای بررسی حساسیت نتایج به اریبی‌های مخفی، آنالیز حساسیت را با توجه به مدل پیشنهاد شده در بخش سوم به صورت کاربردی استفاده می‌کنیم.

قبلاً گفتیم $w=e^{\gamma}$ به عنوان پارامتر حساسیت در مدل در نظر گرفته شد و با تغییر آن و با توجه به مدل ۶ و رابطه سطح معناداری ۵ کران‌های بالا را برای سطح معناداری آزمون محاسبه و ارائه می‌کنیم. نتایج آزمون حساسیت به ازای مقادیر مختلف پارامتر حساسیت $e^{\gamma}=1,2,3,4,10$ ارائه شده است. کلیه محاسبات آماری با استفاده از زبان برنامه‌نویسی Visual fortran 98 انجام شده است.

جدول ۱: کران بالا بر روی سطح معناداری برای مقادیر مختلف e^{γ} (تحلیل حساسیت برای بررسی اریبی‌های مخفی)

$w=e^{\gamma}$	آمار استاندارد شده	کران بالا برای سطح معناداری
۱	۲/۴۶۳	$p=0/013$
۲	۲/۰۴۳	$p=0/207$
۳	۱/۸۳	$p=0/336$
۴	۱/۷۱	$p=0/43$
۵	۱/۵۸	$p=0/571$
۶	۱/۵۱	$p=0/655$
۸	۱/۳۹	$p=0/823$
۱۰	۱/۳۱	$p=0/951$

۵- بحث و نتیجه‌گیری

در قسمت اول این مقاله، آزمونی را ارائه دادیم که به بررسی انسجام روابط مورد نظر که آن‌ها را می‌توانیم در غالب یک فرضیه بررسی کنیم می‌پردازد و دیدیم که در مثال ارائه شده آزمون به خوبی توانست رابطه مورد نظر را تأیید کند. در مطالعات مشاهده‌ای یا کارآزمایی‌های بالینی تصادفی نشده، عمل تخصیص تصادفی انجام نمی‌شود و در نتیجه یک تورش بالقوه در این‌گونه مطالعات وجود دارد که این تورش‌ها یا به صورت آشکارند و یا به صورت پنهان. در مواجهه با کنترل اریبی‌های آشکار عموماً مشکل خاصی وجود ندارد ولی در مواجهه با اریبی‌های پنهان باید مسیر پیچیده‌ای را برای بیان و کنترل آن‌ها طی کنیم. برای بررسی اریبی‌های پنهان، مدلی را در نظر می‌گیریم که شامل یک متغیر تصادفی برای کوریت‌ها یا متغیرهای مشاهده نشده که اریب‌های پنهان را ایجاد می‌کنند و یک پارامتر که شانس دریافت تیمار یا عامل در معرض برای افراد یا گروه‌ها را نشان می‌دهد است و آنالیز حساسیت را با توجه به آماره آزمون و مدل مورد نظر محاسبه می‌کنیم. برای محاسبه سطح معناداری، مقدار ثابتی برای پارامتر حساسیت (ω) در نظر می‌گیریم و تفسیر نتایج با توجه به مقادیر پارامتر حساسیت و سطوح معناداری به دست آمده انجام می‌شود. اگر نتایج آزمون با توجه به مقادیر کوچکی از پارامتر حساسیت معنادار گردید، می‌توان نتیجه گرفت که نتایج مطالعه به کوریت‌های مشاهده نشده حساس نیست و در غیر این

صورت کوریت‌های مشاهده نشده می‌تواند نتایج را تغییر دهد. در مطالعه فوق تنها با در نظر گرفتن اریبی‌های مخفی خیلی بزرگ، مانند حالتی که شانس دریافت تیمار در یک گروه نسبت به گروه دیگر ۱۰ برابر باشد، نتایج مطالعه معنادار نیست و لذا نتایج نسبت به اریبی‌های کوچک و میانه حساسیت ندارد.

منابع

1. Gefeller, O., Paralle, L., A Nonparametric test for Evaluating Coherent Alternatives in Nonrandomized Studies, Proceeding of The international Conference On Nonrandomized Comparative Clinical Studies In Heidelberg, 1997.
2. Rosenbaum, Paul R., Coherence in Observational studies, Biometrics, 1994
3. Cox D., Analysis of Binary data, New York: Chapman and Hall; 1970.
4. Rosenbaum P., observational studies, New York: springer - verlag. (1995).
5. Rosenbaum P., Dropping out of high school in the united states: an observational study, Journal of Educational statistics, 1986; 11, pp.207-224.
6. Rosenbaum P., sensitivity analysis for certain permutation tests in matched observational studies, Biometrika, 1987; 14,P:13 - 26
7. Rosenbaum P. sensitivity analysis for matching with multiple controls. Biometrika. 1988, 5, pp.577 - 81.
8. Rosenbaum P., Sensitivity analysis for tow sample permutation inferens in observational studies, JASA, 1990, 85, pp.493-80.
9. Hollander M, Wolfe D., Nonparametric statistical methods, New York: Wiley: 1973.
10. Zelen, M., Randomized Consent Designs for Clinical Trial, Statistics in Medicine, 1990, 9,645-656.
11. Rosenbaum, Paul R. (2001) Signed Rank Statistics for coherent Predications, Biometric, 53.
12. Rosenbaum, P.R. (1991). Some poset statistics. *Annals of Statistics* 19, 1091-1097