معیارهای وضوح تعمیم‌یافته و کمینه انحراف تعمیم‌یافته برای مقایسه طرح‌های کسری نامنظم

فهیمه براتی[۱] و احمد نوراله[۲]

چکیده:

یکی از مسائل اساسی که در استفاده از طرح‌های کسری دنبال می‌شود، چگونگی انتخاب طرح بهینه از میان طرح‌های کسری دارای تعداد اجرای برابر است. معیار وضوح، عمومی‌ترین معیار برای مقایسه طرح‌های کسری منظم است. این مقاله به معرفی معیار وضوح تعمیم‌یافته، که تعمیم مفهوم وضوح به حوزه طرح‌های کسری نامنظم است، می‌پردازد. این معیار، ملاک مناسبی برای رتبه‌بندی طرح‌های کسری نامنظم است. هم‌چنین از آن جایی که وضوح تعمیم‌یافته طرح‌های کسری منظم با وضوح این‌گونه طرح‌ها برابر است از معیار وضوح تعمیم‌یافته می‌توان برای مقایسه طرح‌های کسری منظم و نامنظم با تعداد اجرای برابر نیز استفاده نمود. مزیت دیگر معیار وضوح تعمیم‌یافته این است که ایده‌ای را برای معرفی معیار کمینه انحراف تعمیم‌یافته به دست می‌دهد که معیار اخیر، تعمیم معیار کمینه انحراف است. در خلال مباحث مثال‌هایی برای تشریح کاربرد معیارهای معرفی شده، ارائه شده است.

واژه‌های کلیدی: اختلاط، الگوی طول واژه، تصویرسازی، تعامد، کمینه انحراف، وضوح.

۱ مقدمه

طرح‌های عاملی کسری با عوامل دوسطحی کاربرد وسیعی در آزمایش‌های غربال‌ساز و دیگر بررسی‌های علمی دارند. به منظور نمایش یک طرح کسری n اجرایی برای مطالعه m عامل دوسطحی، معمولاً از یک ماتریس $\mathbf{X}_{(n\times m)}$ شامل نشانه‌های ۱+ و ۱− استفاده می‌شود. در این‌گونه ماتریس‌ها هر ستون متناظر با یک عامل، هر نشانه یک سطح عاملی و هر سطر نشان‌دهنده یک ترکیب از سطوح عاملی است. در این مقاله بیشتر طرح‌های کسری متعامد مد نظر قرار می‌گیرند. لیکن معیارهای معرفی شده برای مقایسه‌ی طرح‌های کسری نامتعامد نیز کاربرد دارند. مقصود از یک طرح کسری متعامد، طرحی است که تعداد عناصر ۱+ و ۱− در هر ستون ماتریس آن یکسان هستند و هم‌چنین در هر زیرماتریس $n \times ۲$ از ماتریس مورد اشاره، چهار ترکیب (۱+,۱+)، (۱−,۱+)، (۱+,۱−) و (۱−,۱−) به تعداد دفعات برابر در قالب سطرها، ظاهر می‌گردند.

طرح‌های کسری به دو دسته منظم و نامنظم تقسیم می‌شوند. یک طرح کسری را منظم می‌نامند اگر بتوان آن را با استفاده از روابط معرف تعریف کرد. طرح‌های کسری منظم دوسطحی الگوی هم‌اثری ساده‌ای دارند که

[۱] کارشناس ارشد آمار، دانشگاه اصفهان
[۲] عضو هیئت علمی گروه آمار، دانشگاه اصفهان

طبق آن هر دو اثر عاملی یا متعامدند و یا کاملاً هم‌اثر می‌باشند. تعداد اجراها در این‌گونه طرح‌ها برابر توانی از ۲ است. در مقابل، طرح‌های کسری نامنظم دارای الگوی هم‌اثری پیچیده هستند به این معنی که در این طرح‌ها اثراتی وجود دارد که نه کاملاً هم‌اثر و نه متعامدند. تعداد اجراها در طرح‌های کسری نامنظم متعامد مضربی از ۴ است. طرح‌های پلاکت—برمن [۳] [۶] و طرح‌هایی که با استفاده از ماتریس‌های هادامارد [۴] ساخته می‌شوند (طرح‌های هادامارد) مثال‌هایی از طرح‌های کسری نامنظم هستند.

یکی از مسائل اساسی که در استفاده از طرح‌های کسری دنبال می‌شود چگونگی انتخاب طرح بهینه از میان طرح‌های دارای تعداد اجرای مساوی است. معیار بیشینه‌ی وضوح [۵] عمومی‌ترین معیار برای مقایسه‌ی طرح‌های کسری منظم است. وضوح یک طرح کسری منظم برابر طول کوتاه‌ترین واژه در رابطه معرف آن طرح، تعریف می‌شود [۱]. طرح‌های دارای وضوح کمتر، واژه‌های معرف کوتاه‌تری دارند که به تبع آن اثرات مراتب پایین‌تر در آن‌ها هم‌اثر می‌شوند. از طرفی با توجه به اصل سلسله مراتبی بودن اثرات، اثرات مراتب پایین‌تر شانس بیشتری برای معنی‌دار بودن دارند. بنابراین انتخاب طرحی که موجب هم‌اثری اثرات مراتب پایین می‌شود مطلوب نیست.

برای انتخاب طرح‌های نامنظم نیز تاکنون معیارهای متعددی معرفی شده‌اند که از آن جمله می‌توان به معیارهای وضوح تعمیم‌یافته [۶] و کمینه انحراف تعمیم‌یافته [۷] [۲]، معیار کمینه انحراف G_2 [۸] [۸] و معیار کمینه انحراف MDS [۹] [۴]، اشاره کرد. این مقاله به معرفی معیارهای بیشینه وضوح تعمیم‌یافته و کمینه انحراف تعمیم‌یافته [۲] می‌پردازد. این دو معیار به ترتیب تعمیم معیارهای بیشینه وضوح و کمینه انحراف هستند.

قسمت‌های مختلف مقاله حاضر به شرح زیر تنظیم شده است. در بخش (۲) به معرفی معیار وضوح تعمیم‌یافته برای طرح‌های کسری نامنظم می‌پردازیم و در بخش (۳) توجیه آماری استفاده از این معیار برای مقایسه طرح‌های کسری نامنظم را ارائه می‌نماییم. در بخش (۴) نیز به معرفی معیار کمینه انحراف تعمیم‌یافته خواهیم پرداخت. در خلال مباحث مطرح شده، مثال‌هایی را برای روشن شدن کاربرد معیارهای معرفی‌شده ارائه خواهیم کرد. برای حل این مثال‌ها از برنامه‌نویسی در نرم‌افزار $MATLAB$ استفاده شده است.

۲ اختصاص وضوح به طرح‌های کسری نامنظم

از نقطه نظر تصویری یک طرح منظم دارای وضوح r است اگر در تصویر آن بر روی هر $r-1$ عامل همه‌ی 2^{r-1} ترکیبات $(r-1)$—تایی از عناصر ۱+ و ۱− به تعداد دفعات مساوی در قالب سطرها ظاهر شوند. مقصود

[۳] Plackett-Burman designs
[۴] Hadamard matrices
[۵] Resolution
[۶] Generalized resolution
[۷] Generalized minimum aberration
[۸] G_2 Minimum aberration
[۹] MDS Minimum aberration

از تصویر یک طرح عاملی m—عاملی بر هر زیرمجموعه p—عضوی از عوامل آن زیرطرحی است که با حذف $m-p$ عامل دیگر طرح به‌دست می‌آید.

برای تعمیم مفهوم وضوح در طرح‌های نامنظم از منظر تصویری وضوح طرح‌های منظم استفاده می‌کنیم. به این ترتیب می‌توان گفت که یک طرح عاملی نامنظم دارای وضوح تعمیم‌یافته r است اگر در تصویر آن بر روی هر $r-1$ عامل همه 2^{r-1} ترکیبات $(r-1)$—تایی از عناصر ۱+ و ۱− به تعداد دفعات مساوی در قالب سطرها ظاهر شوند. بنابراین یک طرح کسری نامنظم با وضوح تعمیم‌یافته r یک آرایه متعامد[۱۰] دوسطحی با توان $r-1$ خواهد بود. (یک آرایه متعامد دوسطحی با توان t یک ماتریس شامل عناصر ۱+ و ۱− است که در هر زیرماتریس شامل t ستون از آن، همه 2^t ترکیب t—تایی از عناصر ۱+ و ۱− به تعداد دفعات مساوی در قالب سطرها، ظاهر می‌شوند.)

پیش از معرفی مفهوم وضوح تعمیم‌یافته، به معرفی برخی از نمادهای به‌کاررفته در این مقاله می‌پردازیم. یک طرح عاملی $n \times m$، منظم یا نامنظم، را با نماد D نمایش می‌دهیم و ماتریس این طرح را به صورت $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$ در نظر می‌گیریم که در آن $\mathbf{x}_j = (x_{1j}, x_{2j}, \ldots, x_{nj})^T$، $j = 1, \ldots, m$، نشان‌دهنده ستون متعلق به عامل j—ام است. هم‌چنین برای $1 \le k \le m$ و هر زیرمجموعه k عضوی $S = \{\mathbf{x}_{h_1}, \mathbf{x}_{h_2}, \ldots, \mathbf{x}_{h_k}\}$ از ستون‌های $\mathbf{X}$، مقدار $J_k(S)$ را که ابزاری اساسی در تعریف معیارهای وضوح تعمیم‌یافته و کمینه انحراف تعمیم‌یافته است، به شکل زیر تعریف می‌کنیم

$$J_k(S) = |j_k(S)| = \Big| \sum_{i=1}^{n} x_{ih_1} x_{ih_2} \cdots x_{ih_k} \Big|, \qquad (1)$$

که در آن $|.|$ نماد قدرمطلق است. واضح است که در طرح‌های متعامد $J_1(S) = J_2(S) = 0$ خواهد بود.

حال طرح، منظم یا نامنظم، D را در نظر بگیرید و فرض کنید r کوچک‌ترین عدد صحیحی باشد که برای آن $max_{|S|=r} J_r(S) > 0$ است، که در این‌جا $|S|$ نشان دهنده تعداد اعضای مجموعه S است. هم‌چنین $max_{|S|=r} J_r(S)$ برابر مقدار بیشینه $J_r(S)$ در بین همه مجموعه‌های شامل r ستون از $\mathbf{X}$ می‌باشد. وضوح تعمیم‌یافته D به صورت زیر تعریف می‌شود

$$GR(D) = r + [1 - \frac{max_{|S|=r} J_r(S)}{n}]. \qquad (2)$$

از آن‌جایی که برای همه مقادیر r، $0 \le J_r(S) \le n$ است، $r \le GR(D) < r+1$ می‌باشد. هم‌چنین با توجه به این که در طرح‌های متعامد $J_1(S) = J_2(S) = 0$ است، در این‌گونه طرح‌ها $GR(D) \ge 3$. با استفاده از معیار بیشینه وضوح تعمیم‌یافته، طرح‌های دارای وضوح تعمیم‌یافته بالاتر ترجیح داده می‌شوند.

در صورتی که طرح D یک طرح منظم باشد، $J_k(S)$ برای هر زیرمجموعه k عضوی $S = \{\mathbf{x}_{h_1}, \mathbf{x}_{h_2}, \ldots, \mathbf{x}_{h_k}\}$ از ستون‌های $\mathbf{X}$ برابر صفر یا n خواهد بود. رابطه $J_k(S) = 0$ نشان‌دهنده متعامد بودن ستون‌های عضو S است و $J_k(S) = n$ نمایانگر این است که عوامل متناظر با k ستون عضو S یک واژه معرف به طول k

[۱۰] Orthogonal array

تشکیل می‌دهند. بنابراین در این حالت اثر اصلی متناظر با هر یک از ستون‌های عضو S با اثر متقابل $k-1$ ستون دیگر کاملاً هم‌اثر است. به این ترتیب در طرح‌های منظم $max_{|S|=r}J_r(S) > 0$ به معنی این است که $max_{|S|=r}J_r(S) = n$ می‌باشد. در نتیجه اگر در طرح منظم D، r کوچک‌ترین عدد صحیحی باشد که برای آن $max_{|S|=r}J_r(S) > 0$ است، آنگاه طول کوتاه‌ترین واژه معرف D برابر r خواهد بود و لذا طبق تعریف وضوح D نیز برابر r می‌باشد. هم‌چنین با توجه به رابطه (۲)، وضوح تعمیم‌یافته D نیز برابر $GR(D) = r+[1-\frac{n}{n}] = r$ خواهد بود. بنابراین وضوح تعمیم‌یافته طرح‌های منظم برابر وضوح این طرح‌ها است.

در مثال زیر معیار وضوح تعمیم‌یافته برای مقایسه زیرطرح‌های ناهم‌ارز 4×20 طرح پلاکت–برمن با ۲۰ اجرا (PB_{20}) به کار رفته است. دو ماتریس $n \times m$ را هم‌ارز می‌نامند اگر بتوان هر یک را با استفاده از جایگشت سطرها، جایگشت ستون‌ها و (یا) تغییر سطوح یک یا چند عامل ماتریس دیگر به دست آورد. طرح‌های D_1 و D_2 به ترتیب با ماتریس‌های طرح $\mathbf{X}_1$ و $\mathbf{X}_2$ نیز در صورتی هم‌ارز هستند که $\mathbf{X}_1$ و $\mathbf{X}_2$ هم‌ارز باشند.

مثال ۱ ماتریس طرح PB_{20} به‌دست آمده از ترتیب چرخشی بردار سطری زیر را در نظر بگیرید

$$(+1,+1,-1,-1,+1,+1,+1,+1,-1,+1,$$
$$-1,+1,-1,-1,-1,-1,+1,+1,-1)$$

طبق یافته‌های لین و دراپر [۵] طرح PB_{20} دارای سه تصویر ناهم‌ارز 4×20 است که آن‌ها را به ترتیب با $PB_{20-4.1}$، $PB_{20-4.2}$ و $PB_{20-4.3}$ نمایش می‌دهیم.

سه طرح‌تصویر مورد اشاره را به ترتیب می‌توان با تصویر طرح PB_{20} بر روی ستون‌های {۴، ۳، ۲، ۱}، {۶، ۳، ۲، ۱} و {۱۶، ۳، ۲، ۱} به دست آورد. به دلیل این که سه طرح‌تصویر یادشده متعامد هستند مقادیر $J_1(S)$ و $J_2(S)$ در آن‌ها برابر صفر است. مقدار $max_{|S|=3}J_3(S)$ نیز برای دو طرح $PB_{20-4.1}$ و $PB_{20-4.3}$ برابر ۴ و برای طرح $PB_{20-4.2}$ برابر ۱۲ به دست می آید. بنابراین مطابق رابطه (۲) مقدار وضوح تعمیم‌یافته طرح های $PB_{20-4.1}$ و $PB_{20-4.3}$ برابر $3,8$ و مقدار وضوح تعمیم‌یافته طرح $PB_{20-4.2}$ نیز برابر $3,4$ است. بنابراین از منظر معیار بیشینه وضوح تعمیم‌یافته دو طرح $PB_{20-4.1}$ و $PB_{20-4.3}$ برتر از طرح $PB_{20-4.2}$ هستند.

۳ توجیه آماری استفاده از وضوح تعمیم‌یافته

برای ساده شدن بررسی دلیل استفاده از معیار وضوح تعمیم‌یافته بحث زیر را ارائه می‌کنیم. فرض کنید در آزمایشی قصد داریم اثرات اصلی عوامل را برآورد کنیم اما بنابه دلایلی قادر نیستیم از یک طرح دارای وضوح تعمیم‌یافته برابر ۴ یا بالاتر استفاده کنیم. هم‌چنین فرض کنید بنابر اطلاعات قبلی مطمئن هستیم که اثرات متقابل سه‌عاملی و مراتب بالاتر ناچیزند اما گمان می‌کنیم که بعضی از اثرات متقابل دوعاملی معنی‌دار هستند. سؤالی که در این جا مطرح می‌شود این است که نادیده گرفتن اثرات متقابل دوعاملی معنی‌دار در مدل برازش‌داده‌شده

بر روی داده‌ها چه تأثیری بر روی برآورد اثرات اصلی می‌گذارد؟

برای پاسخ به سؤال فوق طرح متعامد D را با $3 \leq GR(D) < 4$ در نظر بگیرید و فرض کنید که $\mathbf{X}_{(n\times m)}$ ماتریس طرح یادشده باشد. مدل شامل اثرات اصلی m عامل مشمول در طرح D به صورت زیر نوشته می‌شود

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i, \qquad i = 1, 2, \ldots, n, \quad (3)$$

که در آن y_i، مقدار مشاهده‌شده متغیر پاسخ در اجرای iـام بوده، β_0 اثر ثابت مدل است و β_jـها برای $j = 1, 2, \ldots, m$، پارامترهای متناظر با اثرات اصلی m عامل طرح هستند. با استفاده از روش کمترین مربعات خطا برآورد هر یک از پارامترهای β_j، $j = 1, 2, \ldots, m$، درمدل رابطه (۳) به شکل زیر به دست می‌آید

$$\hat{\beta}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij} y_i. \quad (4)$$

از طرفی با فرض معنی‌دار بودن بعضی از اثرات متقابل دوعاملی، مدل صحیح به صورت زیر خواهد بود

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \sum_{k<l}^{m} \beta_{kl} x_{ik} x_{il} + \varepsilon_i, \quad (5)$$

که β_{kl}، $k,l = 1, 2, \ldots, m$، در آن نشان‌دهنده پارامتر متناظر با اثر متقابل بین دو عامل k و l است.

تحت مدل صحیح (مدل رابطه (۵))، امید ریاضی $\hat{\beta}_j$ در رابطه(۴) برای $j = 1, 2, \ldots, m$، به صورت زیر به دست می‌آید

$$E(\hat{\beta}_j) = \beta_j + \frac{1}{n} \sum_{k<l} j_3(x_j, x_k, x_l) \beta_{kl},$$

که در آن $j_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l) = \sum_{i=1}^{n} x_{ij} x_{ik} x_{il}$ است. واضح است که $J_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l) = \left| j_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l) \right|$. بنابراین یک راه برای کمینه کردن اریبی ایجادشده به دلیل وجود β_{kl}ـها در برآورد β_jـها کمینه کردن مقدار $max_{j<k<l} J_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l)$ است که این امر با انتخاب طرح دارای بیشینه وضوح تعمیم‌یافته، معادل است. زیرا مطابق رابطه (۲)، $GR(D) = r + [1 - \frac{max_{|S|=r} J_r(S)}{n}]$ است و از این‌رو طرح دارای بیشینه وضوح تعمیم‌یافته در بین طرح‌های D با $3 \leq GR(D) < 4$، طرحی است که مقدار $max_{j<k<l} J_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l)$ در آن کمینه باشد. بنابراین طرح دارای بیشینه وضوح تعمیم‌یافته از طریق کمینه کردن مقدار $max_{j<k<l} J_3(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l)$، اریبی برآوردگرهای اثرات اصلی که از نادیده گرفتن اثرات متقابل دوعاملی معنی‌دار ناشی می‌شود را کمینه خواهد کرد.

۴ معیار کمینه انحراف تعمیم‌یافته

فرایز و هانتر [۳] معیار کمینه انحراف را برای تشخیص بهتر طرح‌های بهینه از میان طرح‌های کسری منظم پیشنهاد کردند. این معیار با استفاده از الگوی طول واژه طرح‌های منظم تعریف می‌شود. فرض کنید A_i نشان‌دهنده تعداد واژه‌های با طول i در رابطه معرف یک طرح کسری منظم باشد. در این صورت بردار $\mathbf{W} = (A_3, \ldots, A_m)$ الگوی طول واژه این طرح نامیده می‌شود.

تعریف ۱ دو طرح کسری منظم مانند D_1 و D_2، را به ترتیب با الگوی طول واژه $\mathbf{W}(D_1)$ و $\mathbf{W}(D_2)$ در نظر

بگیرید و فرض کنید r کوچک‌ترین عدد صحیحی باشد که برای آن $A_r(D_1) \neq A_r(D_2)$ است. در این صورت گفته می‌شود طرح D_1 دارای انحراف کمتری نسبت به طرح D_2 است اگر $A_r(D_1) < A_r(D_2)$ باشد. هم‌چنین اگر هیچ طرح 2^{f-q} دیگری با انحراف کمتر از D_1 وجود نداشته باشد، D_1 طرح دارای کمینه انحراف خواهد بود.

برای رتبه‌بندی طرح‌های نامنظم D_1 و D_2 با وضوح تعمیم‌یافته برابر نیز می‌توان از ایده‌ای مشابه با ایده به‌کاررفته در معیار کمینه انحراف، استفاده نمود. برای این منظور، دو طرح نامنظم D_1 و D_2 را با وضوح تعمیم‌یافته $r \leq GR(D_1) = GR(D_2) < r+1$ در نظر بگیرید. برابری وضوح تعمیم‌یافته دو طرح مورد اشاره متضمن این مطلب است که میزان $max_{|S|=r} J_r(S)$ در آن دو طرح برابر است. حال واضح است، در صورتی که فراوانی تعداد زیرمجموعه‌های r—عضوی S از ستون‌های ماتریس طرح D_1 که به مقدار $max_{|S|=r} J_r(S)$ دست پیدا می‌کنند کمتر از فراوانی تعداد این زیرمجموعه‌ها در طرح D_2 باشد، طرح D_1 بر طرح D_2 ترجیح داده می‌شود. اگر فراوانی تعداد زیرمجموعه‌های مورد اشاره در دو طرح D_1 و D_2 یکسان باشد، دومین مقدار بزرگ $J_r(S)$ دو طرح با یکدیگر مقایسه می‌شوند. در این صورت اگر دومین مقدار بزرگ $J_r(S)$ در طرح D_1 کوچک‌تر از دومین مقدار بزرگ $J_r(S)$ در طرح D_2 باشد، D_1 بر D_2 ترجیح داده می‌شود. اگر دومین مقدار $J_r(S)$ نیز در دو طرح مورد اشاره یکسان باشد، فراوانی تعداد زیرمجموعه‌های r—عضوی S از ستون‌های ماتریس آن دو که به دومین مقدار بزرگ $J_r(S)$ دست پیدا می‌کنند با یکدیگر مقایسه می‌شوند و این رویه تا زمانی که بتوان طرح بهینه را از بین دو طرح D_1 و D_2 تشخیص داد، ادامه می‌یابد.

ایده مطرح‌شده فوق را می‌توان با استفاده از بردار فراوانی اختلاط[11]، که در حقیقت تعمیم الگوی طول واژه طرح‌های منظم است، گسترش داد. پیش از ارائه تعریف برادار فراوانی اختلاط، مایلیم از مقادیر ممکن $J_k(S)$ اطلاع پیدا کنیم. قضیه زیر نشان می‌دهد که مقادیر $J_k(S)$ در طرح‌های متعامد الزاماً مضربی از ۴ است.

قضیه ۱ برای هر k ستون از ماتریس یک طرح عاملی متعامد شامل n اجرا و m عامل دوسطحی، مقادیر $J_k(S)$ برابر مضربی از ۴ هستند [۲].

برای مشاهده اثبات این قضیه می‌توان به دنگ و تانگ [۲] مراجعه نمود.

اکنون طرح متعامد D را با $n = 4t$ اجرا و m عامل دوسطحی در نظر بگیرید و فرض کنید f_{kj} برای $j = 1, 2, \ldots, t, t+1$ نشان‌دهنده تعداد زیرمجموعه‌های k عضوی S از ستون‌های ماتریس طرح D باشد که برای آن‌ها $j_k(S) = 4(t+1-j)$ است. از آن جایی که تعداد زیرمجموعه‌های k عضوی از ستون‌های ماتریس طرح D برابر $\binom{m}{k}$ است، $\sum_{j=1}^{t+1} f_{kj} = \binom{m}{k}$ خواهد بود و لذا کافی است تنها مقادیر f_{kj} را برای $j = 1, 2, \ldots, t$ در نظر بگیریم. هم‌چنین به دلیل این که در طرح‌های متعامد $J_1(S) = J_2(S) = 0$ می‌باشد در این طرح‌ها برای $j = 1, 2, \ldots, t$، $f_{1j} = f_{2j} = 0$ است. بنابراین در طرح‌های متعامد تنها مقادیر f_{kj} را برای $k \geq 3$

[11] Confounding frequency vector

مدنظر قرار می‌دهیم. اکنون با توجه به مطالب فوق، بردار فراوانی اختلاط طرح D را به صورت زیر تعریف می‌کنیم.

تعریف ۲ بردار فراوانی اختلاط برای طرح متعامد D با $n = ۴t$ اجرا و m عامل دوسطحی، یک بردار سطری با بعد $(m-۲)t$ است که به شکل زیر نوشته می‌شود:

$$F = [(f_{۳۱}, ..., f_{۳t}); (f_{۴۱}, ..., f_{۴t}); ...; (f_{m۱}, ..., f_{mt})].$$

بردار فراوانی اختلاط نیز همانند الگوی طول واژه طرح‌های منظم، اطلاعات مهمی در ارتباط با نحوه اختلاط اثرات ارائه می‌کند. در حقیقت بردار فراوانی اختلاط، تعمیم الگوی طول واژه طرح‌های منظم به کل طرح‌های عاملی است. شایان توجه است که بردار فراوانی اختلاط طرح‌های منظم دقیقاً همان الگوی طول واژه آن‌ها خواهد بود. زیرا همان گونه که پیش از این اشاره کردیم، مقدار $J_k(S)$ برای هر زیرمجموعه k−عضوی از ستون‌های ماتریس یک طرح منظم برابر صفر یا n است و لذا در این طرح‌ها برای $j = ۲,۳,\dots,t$، $f_{kj} = ۰$ می‌باشد و $f_{k۱}$ نشان‌دهنده تعداد واژه‌های معرف به طول k است. به این ترتیب بردار فراوانی اختلاط یک طرح منظم یک بردار $(m-۲)$−بعدی به صورت $(f_{۳۱}, f_{۴۱}, \dots, f_{m۱})$ است که عیناً همان الگوی طول واژه طرح مذکور می‌باشد.

حال ایده مطرح‌شده برای مقایسه طرح‌های دارای وضوح تعمیم‌یافته یکسان که در ابتدای این قسمت به آن اشاره شد را با استفاده از بردار فراوانی اختلاط و به عنوان معیار کمینه انحراف تعمیم‌یافته، گسترش می‌دهیم.

تعریف ۳ طرح‌های متعامد $D_۱$ و $D_۲$ را با $n = ۴t$ اجرا و m عامل دوسطحی در نظر بگیرید و فرض کنید $f_l(D_۱)$ و $f_l(D_۲)$، برای $l = ۱,۲,\dots,(m-۲)t$ به ترتیب نشان‌دهنده l−امین عنصر بردار فراوانی اختلاط طرح‌های $D_۱$ و $D_۲$ باشند. هم‌چنین فرض کنید r کوچک‌ترین عدد صحیحی باشد که برای آن $f_r(D_۱) \neq f_r(D_۲)$ است. به این ترتیب، در صورتی که $f_r(D_۱) < f_r(D_۲)$ باشد، گفته می‌شود طرح $D_۱$ انحراف تعمیم‌یافته کمتری نسبت به طرح $D_۲$ دارد. هم‌چنین اگر هیچ طرحی با انحراف تعمیم‌یافته کمتری نسبت به طرح $D_۱$ وجود نداشته باشد، $D_۱$ طرح دارای کمینه انحراف تعمیم‌یافته نامیده می‌شود.

واضح است که با توجه به یکسانی بردار فراوانی اختلاط و الگوی طول واژه در طرح‌های منظم، معیار کمینه انحراف تعمیم‌یافته در این طرح‌ها همان معیار کمینه انحراف خواهد بود.

مثال ۲ همان‌گونه که در مثال (۱) بیان شد، طرح‌های $PB_{۲۰-۴.۱}$ و $PB_{۲۰-۴.۳}$ دارای وضوح تعمیم‌یافته یکسان و برابر ۳,۸ هستند. بنابراین برای مقایسه این دو طرح می‌توانیم از معیار کمینه انحراف تعمیم‌یافته استفاده کنیم. بردارهای فراوانی اختلاط دو طرح $PB_{۲۰-۴.۱}$ و $PB_{۲۰-۴.۳}$ به شکل زیر هستند:

$$\mathbf{F}(PB_{۲۰-۴.۳}) = [(۰,۰,۰,۰,۴)_۳];(۰,۰,۱,۰,۰)_۴],$$
$$\mathbf{F}(PB_{۲۰-۴.۱}) = [(۰,۰,۰,۰,۴)_۳];(۰,۰,۰,۰,۱)_۴].$$

زیرنویس‌های ۳ و ۴ در دو بردار بالا به ترتیب نشان‌دهنده گروه‌های فراوانی $f_{۳j}$ و $f_{۴j}$، $j = ۱,۲,\dots,۵$، هستند، که گروه فراوانی k نشان‌دهنده گروهی از فراوانی‌هایی است که توسط زیرمجموعه‌های k−عضوی از ستون‌های

ماتریس طرح ایجاد می‌شود. بنابراین با استفاده از معیار کمینه انحراف تعمیم‌یافته، طرح $PB_{۲۰-۴.۱}$ برتر از طرح $PB_{۲۰-۴.۳}$ تشخیص داده می‌شود.

۵ بحث

برخلاف معیار وضوح در طرح‌های کسری منظم، معیار وضوح تعمیم‌یافته در طرح‌های کسری نامنظم قادر به آشکارساختن میزان توانایی این طرح‌ها در برآورد اثرات عاملی نیست. از نقطه نظر توانایی برآورد اثرات، اگر وضوح یک طرح منظم برابر یک عدد صحیح فرد مانند r باشد، آن‌گاه این طرح تحت فرض ناچیز بودن اثرات شامل $\frac{(r+۱)}{۲}$ عامل و اثرات مراتب بالاتر، توانایی برآورد همه اثرات شامل $\frac{(r-۱)}{۲}$ عامل و اثرات مراتب پایین‌تر از آن را دارا است. هم‌چنین در صورتی که وضوح یک طرح منظم برابر عدد صحیح زوجی مانند r باشد، طرح مورد اشاره با فرض ناچیز بودن اثرات شامل $\frac{(r+۲)}{۲}$ عامل و اثرات مراتب بالاتر از آن، قادر است همه اثرات شامل $\frac{(r-۲)}{۲}$ عامل و اثرات مراتب پایین‌تر از آن را برآورد نماید. برای مثال جدول (۱) نشان‌دهنده ترانهاده ماتریس طرح نامتعامد $D_۱$ است که توسط طالبی و اسماعیل‌زاده [۷] معرفی شده است. میزان وضوح تعمیم‌یافته $D_۱$ برابر $GR(D_۱) = ۱,۸۳۳$ به دست می‌آید. بنابراین انتظار داریم این طرح با فرض ناچیز بودن اثرات متقابل، حداکثر قادر به برآورد تعدادی از اثرات اصلی چهار عامل دخیل در طرح، باشد. لیکن با فرض ناچیز بودن اثرات متقابل سه‌عاملی و مراتب بالاتر، $D_۱$ امکان برآورد کلیه شش اثر متقابل دوعاملی بین چهار عامل را در کنار اثرات اصلی آن‌ها، فراهم می‌نماید. بنابراین به نظر می‌رسد برای تعمیم مفهوم وضوح به حوزه طرح‌های کسری نامنظم، لازم است علاوه بر جنبه تصویری این مفهوم، ارتباط آن با برآوردپذیری اثرات عاملی نیز در نظر گرفته شود.

جدول ۱. ترانهاده ماتریس طرح $D_۱$.

	اجراها											
عامل ها	۱	۲	۳	۴	۵	٦	۷	۸	۹	۱۰	۱۱	۱۲
۱	۱	۱	۱-	۱-	۱-	۱-	۱-	۱-	۱	۱	۱	۱-
۲	۱	۱-	۱	۱-	۱-	۱-	۱-	۱	۱	۱	۱-	۱-
۳	۱	۱-	۱-	۱	۱-	۱-	۱	۱	۱	۱-	۱-	۱-
۴	۱	۱-	۱-	۱-	۱	۱-	۱	۱	۱-	۱-	۱	۱-

٦ نتیجه‌گیری

در این مقاله با معرفی معیارهای وضوح تعمیم‌یافته و کمینه انحراف تعمیم‌یافته، با ارائه مثال‌هایی از طرح‌های کسری نامنظم آن‌ها را مقایسه و رتبه‌بندی کردیم. از نقطه نظر آماری، طرح دارای بیشینه وضوح تعمیم‌یافته، اریبی برآوردگرهای اثرات اصلی که از نادیده گرفتن اثرات متقابل دوعاملی معنی‌دار ناشی می‌شود را کمینه خواهد کرد. بنابراین معیار بیشینه وضوح تعمیم‌یافته معیار مناسبی برای مقایسه طرح‌های کسری نامنظم است. با ارائه مثالی نشان داده شد که ممکن است این

معیار قادر به تمایز بین دو طرح نباشد. از این رو ضرورت استفاده از معیار دقیق‌تری احساس می‌شود. ملاک کمینه انحراف تعمیم‌یافته می‌تواند این خلاء را پر نماید. در این جا نشان داد شد که ملاک اخیر در تشخیص طرح برتر، کارساز است. لازم به ذکر است، از آن جایی که وضوح تعمیم‌یافته و الگوی فراوانی اختلاط طرح‌های کسری منظم به ترتیب با وضوح و الگوی طول واژه این گونه طرح‌ها معادل هستند، معیارهای معرفی‌شده را می‌توان برای مقایسه طرح‌های کسری منظم و نامنظم با تعداد اجرای برابر نیز به کار برد.

در پایان این مقاله ارائه طرحی که ضمن برخورداری از ظرفیت بالا برای برآورد اثرات اصلی و متقابل دوعاملی، دارای میزان وضوح تعمیم‌یافته‌ای کمتر از ۲ است، روشن می‌سازد که این معیار هنوز کامل نیست. از این رو برای ساختن و اصلاح معیاری که بتواند ظرفیت برآوردی این طرح‌ها را واقعی‌تر بیان کند نیاز به تحقیق بیشتری وجود دارد.

مراجع

[1] Box, G. E. P and Hunter, J. S. (1961). The 2^{k-p} fractional factorial designs. *Thechnometrics*, 3, 311-351; 449-458.

[2] Deng, L. Y and Tang, B. (1999). Generalized and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs. *Statistica sinica*, 9, 1071-1082.

[3] Fries, A.and Hunter, W. G. (1980). Minimum aberration 2^{k-p} designs. *Thechnometrics*, 22, 601-608.

[4] Lin, C. D., Miller, A. and Sitter, R. R. (2008). Folded Over Non-orthogonal designs. *Journal of Statistical Planning and Inference*, 138, Issue 10, 3107-3124.

[5] Lin, D. K. J. and draper, N. R. (1992). projection properties of Plackett and Burman designs. *Thechnometrics*, 34, 423-428.

[6] Plackett, R. L. and Burman, J. P. (1946). The design of optimum multi-factorial experiments. *Biometrika*, 33, 305-325.

[7] Talebi, H. and Esmaeilzadeh, N. (2009). Weighted Search Probability for Classes of Equivalent Search Designs Comparison. *Communication in Statistics-Theory and Methods*, 40(4), 635-647.

[8] Tang, B. and Deng, L. Y. (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *The Annals of Statistics*, 27, 1914-1926.